1999 BROADCAST ENGINEERING CONFERENCE

PROCEEDINGS

National Association of

NAB

BROADCASTERS

J. BALLARD

1999 BROADCAST ENGINEERING CONFERENCE

PROCEEDINGS

National Association of

NAB

BROADCASTERS

These proceedings contain technical papers presented at the NAB Broadcast Engineering Conference, April 17-22, 1999.

Published by the NAB Office of Science and Technology

ISBN #0-89324-277-2

References to the papers contained in the 1999 Proceedings may be made without specific permission but attributed to the NAB 53rd Annual Broadcast Engineering Conference.

Additional copies of the 1999 Proceedings are available through NAB Services, 1771 N Street, N.W. Washington, DC 20036-2891. For price information, please write or call: 1 (800) 368-5644, or +1 (202) 429-5373.

National Association of
NAB
BROADCASTERS

Science & Technology
1771 N Street, NW • Washington, DC 20036-2891
(202) 429-5346 • Fax: (202) 775-4981

FOREWORD

The broadcast industry is in the midst of rapid, fundamental change. Nearly every aspect of radio and television station operations has been touched by digital technology; some have been completely remade. At the NAB99 Broadcast Engineering Conference (BEC), the key technologies that are driving broadcasting into the new millennium are identified and explained. This *Proceedings* includes many of the important papers given at the conference, offering detailed background information that cannot be realistically covered in a 30-minute presentation.

The presenters chosen for this year's program have targeted the important, dynamic areas of technological development for radio and television station engineers and technical managers. Clearly, digital hardware and accompanying software have displaced analog notions in nearly every area of station operation.

For radio, all-digital stations are practical, and indeed, on the air. The new tools offered by digital technologies, so well covered in the 1999 BEC and in this publication, have made practical new ways of producing audio programs. It is no secret that marketplace forces have redefined local broadcasting, and technical developments have stepped up to the challenge, making it possible to accomplish more with less. The engineers who understand and embrace this paradigm shift are finding considerable success in the new world of radio.

For television, DTV has clearly taken center stage. And while it is certainly true that NTSC pays the bills, DTV represents the future of television. The uncertainty involving the transition to DTV that was so prevalent a couple of years ago has essentially vanished with the realization that the ATSC DTV Standard works, and that consumers will want to take advantage of the new benefits and features that it offers.

The NAB/SBE Broadcast Engineering Conference Advisory Committee considered proposals for technical papers from a record number of experts. The interest in DTV and groundbreaking radio technologies, such as DAB, was enormous. The committee and the Science and Technology Department is justifiably proud of its 1999 offering.

Within this *Proceedings* you will find the future of broadcasting.

Jerry C. Whitaker
Chairman
NAB/SBE Engineering Conference Committee

Lynn D. Claudy
Senior Vice President
Science & Technology

1998-99 NAB/SBE CONFERENCE PLANNING COMMITTEE

Jerry Whitaker, Chairman
Technical Press
393 Yellowstone Drive
Morgan Hill, CA 95037

Jeff Andrew
Director/Digital TV Engineering
Gannett Broadcasting
1100 Wilson Blvd, 27th Floor
Arlington, VA 22234

Andy Butler
Public Broadcasting Service
1320 Braddock Place
Alexandria, VA 22314

Jerry Butler
Public Broadcasting Service
1320 Braddock Place
Alexandria, VA 22314

Dane Ericksen, P.E.
Hammett & Edison, Inc.
P.O. Box 280068
San Francisco, CA 94128

Robert Hess
Director, Broadcast
Operations/Engineering
WBZ-TV, WODS-FM, WBZ-AM
1170 Soldiers Field Road
Boston, MA 02134

Tom McGinley
Chief Engineer
WPGC-FM/WPGC-AM
6301 Ivy Lane
Suite 800
Greenbelt, MD 20770-1402

Robert Seidel
VP, Engineering
CBS
524 W. 57th Street
New York, NY 10019

Milford Smith
Vice President, Engineering
Greater Media, Inc.
2 Kennedy Blvd.
East Brunswick, NJ 08816

Barry Thomas
Director of Engineering
KCMG – Chancellor Media
5900 Wilshire Blvd. Suite 525
Los Angeles, CA 90036

TABLE OF CONTENTS

Broadcast Engineering Conference Opening

Sunday, April 18, 1999

9:00 am - 9:30 am

Chairperson:
Lynn Claudy
NAB, Washington, DC

***The Digits Are Coming - A Worldwide Transition**
Joseph Flaherty, CBS Inc. and Chairman, WBU Technical Committee
New York, NY
David Wood
EBU and Secretary, WBU Technical Committee
Geneva, Switzerland

*Papers not available at the time of publication

DTV: The Big Picture

Sunday, April 18, 1999

9:30 am - 12:00 pm

Chairperson: Jeff Andrew
Gannett Broadcasting, Arlington, VA

9:30 am Sharing the Vision - A Perspective on DTV Transmission from Shared Sites
Michael Thorne
NTL Group Ltd.
New York, NY

***10:00 am DTV 1998: Selected Case Studies from Stations Meeting the November 1 On-Air Date**
Kerry Cozad
Dielectric Communications
Raymond, ME

10:30 am Designing the Digital Television Transmission Network What Broadcasters Need to Consider to Protect Their Long-Term Investment
Barry Hobbs
End-To-End Systems Support, NDS Americas
Newport Beach, CA

11:00 am ATSC IS Top Down Committee Report and Findings
Shannon Skemp
Thomcast Communications, Inc.
Alexandria, VA

11:30 am Digital Television Interoperability - Issues and Progress
Eric Gsell
Advanced Television Technology Center
Alexandria, VA

---

*Papers not available at the time of publication

SHARING THE VISION
A NEW PERSPECTIVE ON DTV TRANSMISSION

Mark Aitken
COMARK Division
Thomcast Communications, Inc.
Southwick, MA

Mike Thorne
NTL
New York, NY

Abstract: *As the implementation of DTV moves forward, broadcasters find themselves struggling for capital and labor resources, and in many cases for transmission sites. As a result of this resource strain, broadcasters are asking smart questions about how best to meet the needs of their broadcast ownership, provide for the technical needs of their DTV transition, and not break the bank at the same time. Shared facility resources are making increasing sense to those with specific commercial and business needs. But is this a commercially sensible route for the broadcaster's business?*

This paper shares with its readers an understanding of the multiple technology and business issues which can lead to solutions in a shared facility environment. Some of the stronger factors for the Broadcaster to weigh are:

- *Costs to Implement and to Operate*
- *Implementation Timescales*
- *Operational Flexibility*
- *Development Potential*

For the broadcaster who is seeking to determine the extent to which a shared facility may have viability, several issues need to be addressed and analyzed. The issues discussed include:

- *Finance and capitalization cost review*
- *Project management overview for shared site compared to sole sites*
- *Transmitter building and support facility issues*
- *Interference and coverage parameters*
- *Antenna and transmission line selection criteria.*
- *Tower and location requirements and considerations*
- *Transmission equipment technology choices*
- *Viewer considerations*
- *Power versus coverage and reception performance tradeoffs*
- *Shared studio requirements for multiplexed services*

In any cooperative or shared environment, there is the need to define and agree the multiple variables, and define a decision path which provides financial and technical benefits for each of the broadcasters. Economic and technical incentives such as on-air availability, competitive cost of ownership, supplier performance control, cash flow management, and a host of others are reviewed in this paper.

We intend that the reader should become better informed to make the crucial judgements and decisions about the development route which leads to maximum shareholder value in the potentially difficult transition to Digital.

THE BACKGROUND TO EXISTING TECHNICAL OWNERSHIP

Terrestrial TV Broadcasters have a long and hard-won history of being first into the market place, and of pioneering the introduction of TV services. Markets were originally defined and subsequently defended by the original station licensees, and the ownership was often personal and proudly held. Competition over the years has been healthy, and the degree of co-operation between broadcasters has been minimal. Why should an existing broadcaster help a new competitor to get his service on the air? This rationale has driven broadcasters to construct and retain their own facilities, and this fact has colored the landscape to this day. A direct consequence is the proliferation of transmitter towers, frequently in highly visible areas, and the resulting difficulties experienced by aspiring terrestrial viewers when pointing receiving antennas for optimum reception of a number of local services. Terrestrial broadcast has a history of not sharing its infrastructure with competing broadcast services, and of keeping the day to day engineering work in-house. Strong relationships frequently exist with equipment manufacturers and consultants for the longer term engineering questions such as spares supply and new service development.

Cable Operators arrived on the broadcast scene second, and have had to carve their slice of the market. They have built their own playout facilities, implemented their own cable networks and managed the supply of cable decoders. Co-operation between cable operators has been limited prior to recent market consolidation in the cable business. Playout facilities have not been shared, cable systems have been self owned and operated, and customer service has been managed in-house. Some parts of the operation have been outsourced, however. Cables are strung on poles owned by others, and program material is taken in from terrestrial broadcasters and from national satellite operators. Trunked distribution has been purchased from telecom operators, although there is now capability for cable operators to provide telecoms services to businesses other than their own. Cable is therefore already in the business of outsourcing the supply of some of its service infrastructure, and increasingly in the business of selling its own capacity for commercial gain.

Satellite Operators are the third entrants into the TV market, and they have a new and dynamic relationship with the terrestrial and cable players. They have a similar playout requirement to the cable operator, that of producing multiple program streams for parallel distribution. The transition to digital standards has increased the number of channels to be generated, and this trend is set to continue. Satellites are proliferating, and the number of new services being distributed by this method worldwide is growing explosively. Satellite operators generally run their own playout facilities, but the distribution of the service is usually outsourced. Satellites are shared facilities, launched, owned and operated by professional satellite operators. Uplinking is, in some cases, outsourced and in others done in house. Receivers are provided through commercial collaboration between the satellite operators and the receiver/decoder manufacturers. Program material is rarely manufactured by the satellite operators, but is mostly bought-in and re-sold in some form. Satellite operators own and operate the commercially sensitive parts of their services (program stream management and selling), and outsource significant parts of the delivery chain (both input and output) to external suppliers. The price they pay for the services that they buy takes account of the supplier's capital expenditure and the staff services provided by external suppliers.

WHY CONSIDER SHARED FACILITIES?

Sharing and Outsourcing as principles

In the limit, a shared service becomes an outsourced service. If a shared facility is jointly owned by the broadcaster and one, or more, other parties, then he relies on the activities of others to ensure that his service is provided. This implies a degree of outsourcing. In the event that the shared facility is provided by a third party, the service can be said to be outsourced. Throughout the remainder of this paper we shall use the term "outsource" frequently, and it is used in this context.

Terrestrial broadcasters have historically kept the provision of their engineering services as an in-house activity. The reasons for this revolve around control of the whole process, the ability to trust others to provide services and on the lack of suitable alternatives. Newer TV players, such as cable and satellite operators, have outsourced parts of their operation with some success, as they have had both the cost constraints and the availability of suitable suppliers to drive them. They choose to concentrate their activities on the mission-critical parts of their enterprise, namely the production of program streams and the collection of revenues. Where they can earn significant revenues by using their own technical platforms to compete in a broader market place, they sometimes choose to do so. (e.g. cable telephony, Internet services)

TERRESTRIAL DTV IMPLEMENTATION ISSUES

Adding new services.

The arrival of terrestrial DTV was not predicted in the early days of analog TV. Facilities were built to accommodate the analog services, with the normal margins for expansion of existing businesses. DTV brings with it the need to simulcast for a number of years, and there is a consequential medium-term need for additional tower space for new antennas and feeders. Transmitter buildings will generally be inadequate for two sets of transmission plant, as will transmitter site power supplies and telecoms capacity. These are all major cost items for the broadcaster.

Tower overload concerns

Many towers are currently overloaded or nearing overload, following the tightening of regulations on tower capacity over the years. Some broadcasters will be able to strengthen their existing towers to add the new service, although the risks and costs in doing so are not to be taken lightly. A tower is at its most vulnerable when being modified, and the consequences of a tower collapse are potentially devastating to a business. New towers can be built, but there is a growing public lobby against the proliferation of telecoms towers of all sorts, this at a time of unprecedented growth in the radio-communications market generally. Tall tower suppliers, strengtheners and erectors are also in short supply, although their numbers will grow in time if a proven demand exists.

Station DTV profitability concerns

The profitability of DTV services will not be quick. At the predicted rates of receiver uptake it appears unlikely that advertisers will pay premium prices for DTV carriage, so broadcasters will need to take a long-term view of the new services. Cost effective implementation and operation are paramount considerations under these circumstances.

Specific Issues for analysis

Cost issues. As each broadcaster builds a new tower, antenna, buildings and program feed link, he will sustain a cost for doing so. Added to this will be the management and staff time for the zoning permissions, project design, project management and commissioning. There will be also be added costs for operating and maintaining the new services, some of which are inevitable (e.g. power) and others of which are controllable (e.g. staffing, spares, capital recovery, rentals).

A key question on cost is whether to capitalise and own new infrastructure or whether to lease service from an external agency. Different businesses will reach differing conclusions in the context of their own commercial circumstances. This is not an engineering decision, however, but a financial one.

Another key question relates to outsourcing in general as a practice, and whether the costs outweigh those of providing the service in-house. In general if an outsourced service relies on shared resources, there is scope for cost savings. This can apply to towers, antennas, buildings, playout equipment, engineering staff resources and to telecoms links. The cost of in-house resources needs to be clinically analyzed to decide whether it is really necessary

Coverage issues. The prime reason for transmitting the DTV signal terrestrially is to make it available to the end customer, the viewer. He can consume the service using a fixed receiving antenna or on a less permanent one. Nevertheless, we need to keep his situation in mind at all times. He pays the bills.

Coverage of DTV services remains an issue, and will continue to do so. Trials to date have confirmed that where a UHF DTV service is provided in conditions where an existing VHF analog service is in use, the coverage is less complete. DTV, however, is being sold as a high definition, high quality, high priced service. Receiver prices are likely to stay high for the foreseeable future. Our viewer will have high expectations of this new service, and will need to have the best available signal to fulfill them. He will now have to consider fitting an external antenna whereas he might earlier have settled for an internal one. Where will he point it? Maximum transmitter powers are also lower for DTV than for analog, further exacerbating the situation. If the available services are scattered around on separate towers the very act of pointing a high gain receiving antenna at one will discriminate against reception of the rest. The optimal solution is to group all of the services at the same transmitter site.

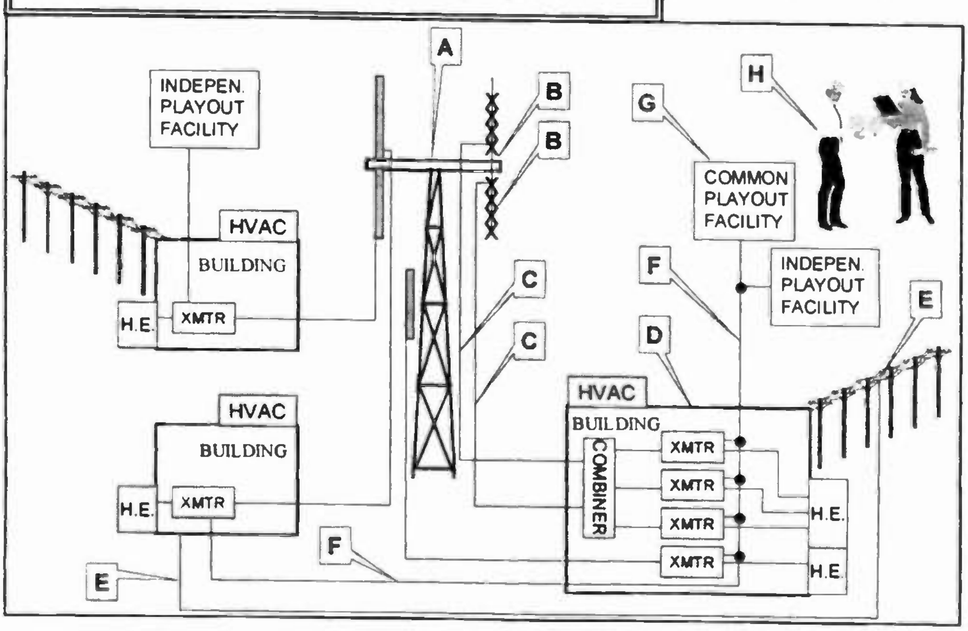
Time-scale issues. The new services need to be implemented in line with the FCC's timetable. There will also be a competitive position within each market, which will make it uncomfortable to be stranded without a solution when competitors are on air. If towers have to be replaced, and if the zoning questions become an increasingly important issue in the community with each new application, then the consequences of being last are dire. Resources also play a part in the time-scale question. If internal resources are to be used for the DTV implementation process, how sure can management be that the locally employed engineering staff have the skills and knowledge of project management in general and of DTV in particular to run the most cost effective and timely project? In many cases the staff will be of the right caliber, but in other they will not be, and outsourcing can then be an option.

WHAT CAN BE SHARED?

The extent of sharing

In practical terms it is possible to share minimally or to outsource everything. It needs to be done the right way for the specific business circumstances. There are also several commercial mechanisms for sharing and a range of service providers of varying capability and quality. We need to understand the scope, the risks and the costs of the degree of sharing which can be considered for each part of the operation. Figure 1 serves to show the vast array of possibilities.

(Figure 1) Vast Choice of Sharing Options



Shared Towers [A] are the item most likely to be shared, the issues of zoning, risk and cost being the principal drivers. The cost of building a tower for six users is approximately twice that of building one for one user. Significant savings therefore become possible with a shared tower. Shared towers imply a shared transmitter site. The terms of ownership of the site are crucial to all that follows.

Shared Antennas [B] imply a shared tower, and give rise to technically complex questions, dependent on the actual mix of frequencies and powers at the site. Two basic configurations are possible. The sharing can be achieved by having a common mechanical platform for a number of electrically separate antennas, or by having an electrically and mechanically common antenna driven by RF combining equipment on the ground. A further consideration is the mixing of shared and separate antennas on the same tower. Careful analysis is needed of the coverage patterns, frequency relationships and power handling capabilities of the antennas, but shared solutions are possible, particularly at the lower powers, and considerable savings can be made compared to building multiple separate antennas and feeder systems. Another benefit of shared antennas, often not appreciated by the broadcaster when considering the options, is that a shared panel antenna, built in two or more sections, has inherent reserve capability. A fault in one section can be isolated and repaired while the remainder of the antenna continues to operate at reduced power. This can save cost in supplying reserve antennas, feeders and sometimes towers.

Shared Feeders [C] imply shared antennas on a shared tower, and are subject to the same provisions as those outlined above. Sharing feeders can be considered at combined RF powers just below 100kW (transmitter powers, not ERP) for UHF, and 250kW for VHF, dependent on channel and frequency spacing. Remember that twin feeders would normally be used for a shared split antenna, enabling the services to continue at reduced power in the event of a feeder or antenna failure.

Shared Transmitter Buildings [D] can reduce the cost to individual users by reducing the total size of the structure. Internal walls can be made less substantial, and better use made of available land. Ventilation can be provided by a common system, housings for diesel generators, power transformers and telecom services can all be shared to save both capital and revenue costs. A downside is the increased risk of total loss in the event of fire or disaster, but common extinguisher and security systems provide an answer to such questions.

Shared Power Feeds [E] are worthy of serious consideration. Power transformers and lines are not linear cost items. Doubling the capacity does not double the cost of transformers, lines, construction or housings. A shared facility can be built for a reduced "per user" cost, usually with inherently greater reliability. This applies equally to larger shared diesel generators and to the relevant change-over switching gear. An additional revenue benefit is the cheaper purchase cost of electricity arising from bulk purchase deals with the power supplier. Increased buying power for both capital and revenue items is an important consideration.

Shared Program Feed equipment [F] is unlikely to be possible over the whole route from studio to the transmitter site, but it may be possible over part of it. Low capacity telco circuits are not maintained to the same availability as high capacity ones, and the cost per bit reduces for high capacity circuits. Traditional STL and microwave radio links can fill part of the gap in the route, but increased demands on these services are limiting availability in many cases due to spectrum congestion. Reliability can therefore rise for a lower cost if the right local solution can be found and a shared facility can be arranged. This requires specialist knowledge and the ability to work in the interests of all of the broadcasters combined, and good solutions are possible.

Shared Studio Playout Centers [G] are a new concept, and perhaps the most visionary of the possibilities discussed in this paper. They are, however, in use in some locations, and providing a highly cost-effective service. Clearly the production and programming plans of competing services need to be treated with great confidentiality, and broadcasters very sensibly will protect the manufacture of their core program product from their competitors' scrutiny. Technological advances now make it possible to assemble and play out programming on an automated basis with minimal human intervention. The equipment requires specialist knowledge, continuous monitoring and occasional personnel involvement. Program streams can be stored and played out by a professionally managed server system, with continuity inserts injected to a schedule from the minimal studio premises. While this might be too expensive for a TV broadcaster to build and operate solely for himself, a fully equipped and managed outsourced service can provide a highly cost-effective option for groups of broadcasters.

Shared Maintenance Staff. [H] When dealing with equipment it is possible to see the value of shared facilities from a relatively dispassionate standpoint. Staffing is, however, a different matter. Station managers will willingly accept that programming can be bought in, that consulting services can be hired, and that manufacturers can be relied on for spares and support. But for the more immediate issues the station manager is inclined to want his own staff, trusted, loyal and familiar, to deal with the problem. This is perfectly understandable, and beyond question in the existing environment.

There are, however, new considerations in DTV. The technical challenges are new and different, the time-scales are tight and critical, and the resources needed are in addition to the "day-job" for existing staff. It is also becoming evident that transmitter engineers are a dying breed, new graduates preferring the route to the digital regions of our business. If shared solutions really do provide potential for cost savings, whose staff is going to get the job of implementing them? Who is then accountable to whom for the outcome? This is the area where a specialist outsource can be most effective, bringing to the project the knowledge, resources and an independent view of events.

SHARING MECHANISMS AVAILABLE

A range of mechanisms

Table 1 serves to illustrate the "per user" cost savings that might result from a shared facility approach.

Tenancy on another broadcaster's facility is a tried and trusted concept. Many Public Television broadcasters rely heavily on such provision. Some commercial broadcasters do so for their main services, and a greater number do so for back-up services, sometimes on a reciprocal basis. Many broadcasters also house FM radio tenants and communications services on their towers and sites. If there is capability and space on the facility then this can be a very cost effective solution for consenting broadcasters.

Joint Ventures among local broadcasters are less common than tenancies, but some do exist. In most cases they have been triggered by the necessity to introduce new DTV services. The forming of a new company with no apparent favoritism is not an easy task amongst competing broadcasters. They have to see a compelling reason to do this, and the long-term success of such ventures has yet to be seen. Experience of the authors to date has shown that one strong individual can drive the grouping along, and that success depends on strength of character. The engineering solutions in evidence give validity to the concept of the shared site model as discussed throughout this paper. The commercial validity will emerge with time and experience, as it has for a number of commercial FM operators (and others) over the last decade.

Sharing on a commercial tower site has become a well accepted concept in the past few years, driven along by the availability of such sites and by the companies now running professionally managed site sharing businesses. Landlords range from small local operators owning a single tower, up to the large public companies owning hundreds of sites. The experience of owning and running tall towers is radically different from that required to run small

Shared Costs (Ilustrative)		
	Single User	Shared User (4)
Tower	$1,000,000	$2,000,000
Antenna	$150,000	$400,000
Combiners	N/A	$150,000
Feeder Line	$250,000	$500,000
Transmitters	$750,000	$3,000,000
Building	$150,000	$400,000
Power Services	$250,000	$500,000
Land	$500,000	$500,000
TOTAL	$3,050,000	$7,450,000
Per User Cost	$3,050,000	$1,862,500

Table 1

ones, and care is needed in entering in to sharing relationships to ensure that the supplier's technical competence is high. Services offered may be simple use of an existing tower, they may include shared building and power supplies, or they may include additional services such as maintenance. The important point is to ensure that the service provider has the competence to meet your standards, can back the story up with evidence, and has the commercial stability to be there for at least the duration of your intended contract. These relationships last a long time and are expensive to change, so they need to be right at the outset.

Sharing on a fully serviced facility is less common in the US than in other parts of the world. The essential component of the relationship is that the broadcast tenant capitalizes and owns all of the equipment dedicated to his own service. Shared antennas and feeders may be part of the site services provided, but the tenant then places all of the maintenance into the hands of the service provider on pre-defined contract terms. Performance guarantees and penalties are usually part of such contracts, and it is advisable to suspect any provider who is not prepared to accept such guarantee conditions. The benefit to the tenant is the complete outsourcing of the staffing and day to day management of the service. The downside is the continuing exposure to additional costs if the equipment fails and at the end of its life.

The fully outsourced service is one in which the broadcaster contracts with a commercial supplier to provide an agreed level of service for a fixed monthly price. There is no capital outlay and no cost variation over time. The supplier designs and builds a system at an agreed location, then runs it to provide the coverage and availability agreed. Any variable costs due to spares, accidents, replacement, obsolescence etc are absorbed by the supplier. The advantages to the broadcaster are that he does not need to fund the infrastructure, the cash flow is predictable and stable, and the service is managed by a professional supplier with the skills and resources available to meet any implementation or operational problems as they arise. The downside is the lack of direct control over the service, and the need to trust an external supplier with the delivery mechanism. The obvious requirement here is for a completely trustworthy and well-founded supplier, with proven experience and credibility.

Full outsourcing is the most common method of running transmission services in Europe, but is not at all common in the USA. The reasons for this are historical, driven originally by the centrally operated telcom providers in Europe. For example, in the UK, over 1200 shared broadcast facilities exist. However, there is now a large European customer base which has chosen to have its new transmission services provided on a fully outsourced basis by commercial service providers, such as NTL, and there are similar service providers ready to provide such services here in the USA.

THE PROS AND CONS FOR THE BROADCASTER

What questions must be addressed?

How is the broadcaster then to decide on whether to go it alone, or whether to go to the fully outsourced provision model, or whether to find a niche somewhere in between which best fits his business? The considerations are always unique to the business and to the local situation. The checklist which follows will help in isolating the important issues, although the decision must be based on the known situation.

Ownership and control issues. How important is complete ownership and control in the new environment of DTV to the business? Will well-controlled contracts with external providers do the job or is it vital to have the means of production and delivery fully in-house?

Cost considerations. What is the real cost of in-house provision, taking into account the capital outlay, the cash-flow, the exposure to debt elsewhere in the company, the full employment and running costs of staff? Can the services really be provided at a lower cost by an external provider? In making the comparison make sure that all of the costs are exposed, both internal to the business and the hidden extras in externally contracted services. What limits do suppliers set on their exposure?

Accountability for service quality. How can the quality of service be assured? What evidence can suppliers give of past success in delivering the services proposed? What contractual vehicles cover the services? What control mechanisms can the supplier show to demonstrate that the quality of service will be that needed? What happens when the service quality falls below standard? Are the penalties enforceable and meaningful? How will the service be provided under extreme conditions of weather of supplier failure?

Accountability for legal compliance. Who is responsible to the regulatory authorities for compliance on RF coverage, interference, aircraft warning lights? If the regulators hold the broadcaster accountable, how can the supplier be held to account for his part of the process? Who has accountability for public liability for towers, power equipment, environmental concerns, zoning issues? To what extent will the supplier take responsibility for these issues?

Staff and employment issues. Does my business have the available staff and resources to undertake the implementation and operation of the new service? What is the future staff situation going to be? Would specialist staff wish to transfer their employment to an external service provider? How much management cost will be spent on running operations in-house compared to the externally offered service?

Risking being different. What are the risks of being the first to do things differently? How can they be controlled or minimized? How attractive does the cost need to be to break the mould?

Quality of site support services. How seriously do the power companies and the telecom providers take my needs right now? Would they be more attentive to a bigger buyer? Would their prices be lower? How much more reliable might a centrally provided diesel generator service be? Does my business already have the skills to understand all of the site disciplines well enough to make suppliers perform?

Environmental and zoning issues. How likely is my new tower zoning application to gain approval? How many others will be attempting to gain similar approvals, forcing a major local backlash? Would zoning authorities see a shared site as a responsible solution?

Available services in the market. To what extent are external service providers available in the market place locally? Is there a tower operator prepared and ready to offer a shared tower? Who offers more than sharing only locally? Who else would wish to share a tower or service in the local market? Would we need an intermediary to make the deal comfortable for all parties?

CONCLUSIONS

The arrival of DTV in the broadcasting market has changed the nature of the business. Resource strain affects the speed, the economics and the certainty with which the changes can be made. The supplier base is ever expanding, and services of all sorts are becoming available in the market place. The services offered enable the broadcaster to approach his business in new and exciting ways, but he needs to re-think some of the basics which have served him so well in the growth of his business to its current state.

Careful selection from the range of services available can make the difference between a successful, vibrant transition into the transformed DTV environment, or a long grinding slog to establish a new service. Outsourcing, as a concept, is now readily accepted as a good way of accessing specialist skills and knowledge. Broadcasters are not unique in their need to cut costs, improve quality and streamline their businesses, and the transition to Digital forms a great entry ramp to implementing the other changes across the business by contracting services externally.

DESIGNING THE DIGITAL TELEVISION TRANSMISSION NETWORK

Barry L. Hobbs
NDS Americas Inc.
Newport Beach, California 92660

What Broadcasters Need to Consider to Protect Their Long-Term Investment

The implementation of digital technology for the broadcaster has provided an opportunity to deliver uniform signal quality across an entire network. The broadcasters have been able to realize significant operational cost reductions by utilizing less transmission bandwidth using digital techniques. The uniformity of signal quality and reduced operational costs have brought new challenges and opportunities in the broadcast industry.

This paper will cover both technical and business issues surrounding the migration of broadcasters to the digital technologies from today's analog environment. We will consider both the standard definition and ATSC digital processing and transmission formats. In addition, we will discuss a few new business opportunities that the ATSC bandwidth may allow. Hopefully, these discussions will assist those who are beginning their investigations into the digital technology migration.

This migration also requires an investment in the education of the engineers who are responsible for the studio and transmission environments. We will provide an abbreviated look at the technology and explain a few of the fundamentals that require an in depth understanding by the engineers who operate the system.

This paper is organized from the output of the studio through the transmission. The first sections will cover digitally compressing the video and audio and placing them into a multiplex with data and conditional access information. The paper discusses some of the issues surrounding MPEG video and audio. From this point the paper moves to formatting the multiplex for transmission for satellite and terrestrial transmissions. There is a short section on challenges such as MPEG transport splicing. The paper concludes with a few possible business opportunities the new digital spectrum may provide to the broadcaster.

Baseband System Operations:

Digital baseband processing systems require multiple components. Among the components are video encoding, audio encoding, data broadcast ability, auxiliary data (PSIP, closed captioning, GCR, conditional access information, subscriber management information and other user information), and multiplex management.

Digital Video and Audio Coding Technologies:

The migration for programmers and broadcasters from analog to digital compression for audio and video started in the late 1980's. The driving forces behind using digital technology have been the efficient use of bandwidth in multiple transmission mediums, a move to substantially reduce operating costs and the ability to deliver a signal of uniform quality to all authorized recipients. The issues the early adopters faced in the late 1980's with MPEG 1 are not too different than the issues that face the broadcasters today. Those issues are:

Quality of Picture:
The quality of the picture has always been of paramount concern to all adopters of digital technology. Broadcasters began by using 34 Mbs/sec compression techniques. The 34Mb system allowed cost reductions and signal security by moving to a telecommunications medium from a more expensive satellite delivery. Soon MPEG techniques began to develop rapidly. In the early 1990's cable programmers began to embrace MPEG 1 techniques at rates of greater than 8 Mbs/sec. At the same time, direct broadcast satellite operators started to embrace MPEG techniques with the goal

of moving to MPEG II. The major issue with MPEG I was that it only supported frame based progressively scanned video. MPEG II was later adopted defining methods of coding for field based interlace pictures and adaptive field/frame processing.

MPEG II defined the tools that we use today for standard definition and high definition compression techniques. It is ironic that the first use of MPEG II video by the U.S. broadcasters, outside of satellite news gathering, are actually occurring in the ATSC transmissions which began on November 1, 1998. What is ironic is the compression ratios required for the high definition 720p60 or 1080i30 formats are more stringent than those required for standard definition formats. The graph below depicts the 8-bit sampling of non-compressed versus the 8-bit sampling required for compressed pictures.

VIDEO COMPRESSION RATIOS

Format	4:2:2 8 Bit Sampling	Contribution Feeds	4:2:0 8 Bit Sampling	Distribution Feeds	% of 4:2:0 8 Bit Sampling
480I 30	166 Mbs	15Mbs = 11:1	125 Mbs	18 Mbs = 6.9:1	14%
480P 30	166 Mbs	15Mbs = 11:1	125 Mbs	18 Mbs = 6.9:1	14%
480P 60	331 Mbs	20Mbs = 17:1	248 Mbs	18Mbs = 14:1	7%
720P 60	884 Mbs	40Mbs = 22:1	663 Mbs	18Mbs = 37:1	3%
1080I 30	994 Mbs	40Mbs = 25:1	746 Mbs	18Mbs = 41:1	2%

As the graph shows, the compression requires reduction of 97% to 98% of the material in reference to the original 8 bit sampled non-compressed material. We actually transmit and reconstruct the picture with less than 2%-3% of the original picture information for a 720p60 or 1080i30 ATSC picture. At the same time a 480i30, normal standard definition picture, can utilize 5 to 7 times as much bit rate.

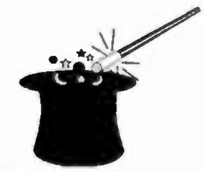
U.S. network broadcasters are just beginning to embrace MPEG II digital technology for the transmission from the network to the affiliates for their standard definition signals. This move has gained momentum as the MPEG II processes have begun to mature. The maturity has been realized in the areas of noise reduction technology in both temporal and spatial filtering, DCT algorithms for spatial redundancy, statistical redundancy (Huffman coding), expansion of motion estimation search ranges, chroma sub-sampling and statistical multiplexing. These attributes are the first keys to investing in a digital television transmission technology;

- *Make sure your choice of vendors for baseband video coding can support software updates for each of these attributes listed in the above paragraph.*

These attributes are important for both high definition and standard definition networks.

MPEG ENCODING TRICKS

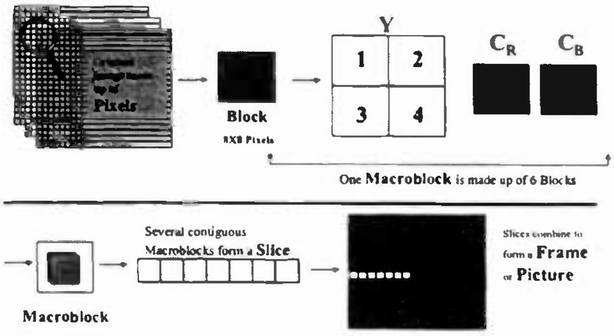
- Statistical Multiplexing
- Large Search Ranges for Motion Estimation
- Pre-Processing of the Source Video (Spatial/Temporal Noise Filters)
- Subsampling the Video at Low Bitrates

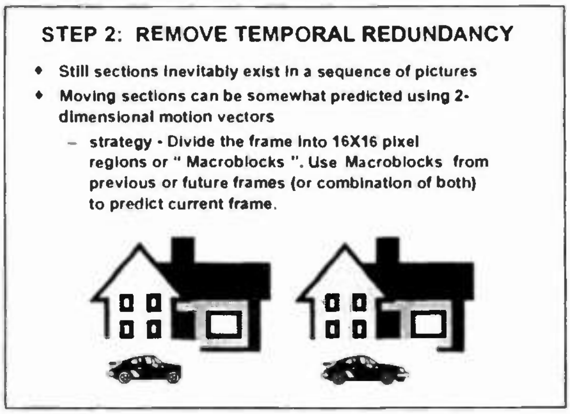


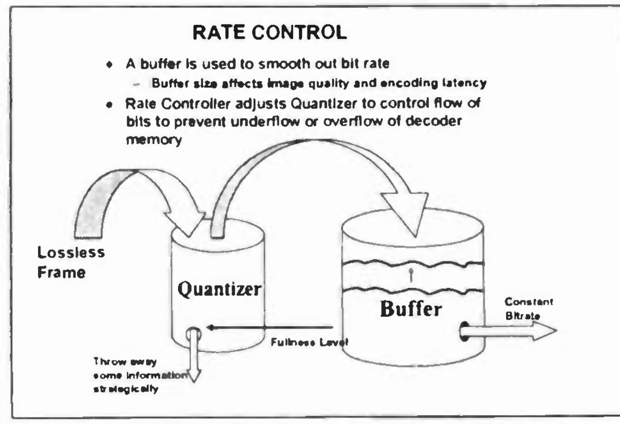
VIDEO COMPRESSION STRATEGY

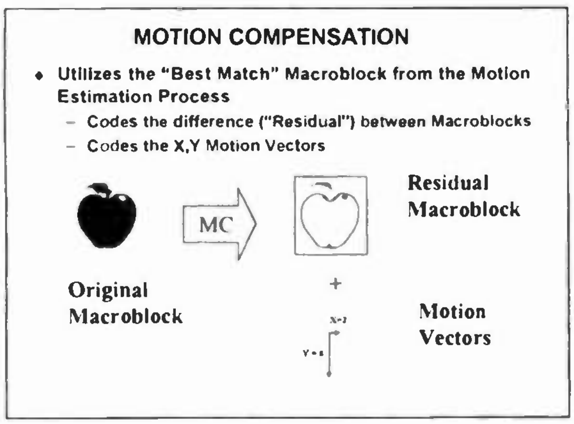
- Remove Redundancy!
  - Step 1: Color Redundancy (4:2:0, 4:2:2)
  - Step 2: Temporal Redundancy (Motion Estimation)
  - Step 3: Spatial Redundancy (DCT)
  - Step 4: Statistical Redundancy (Zig-Zag, Huffman)

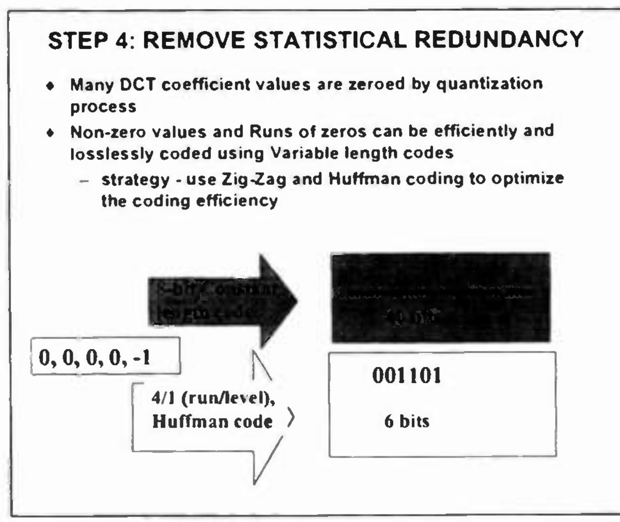
MPEG VIDEO DATA HIERARCHY

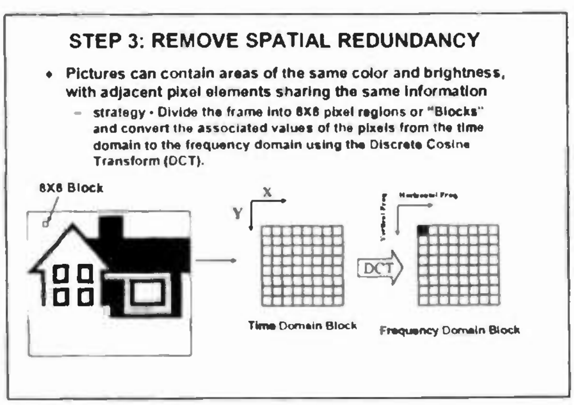


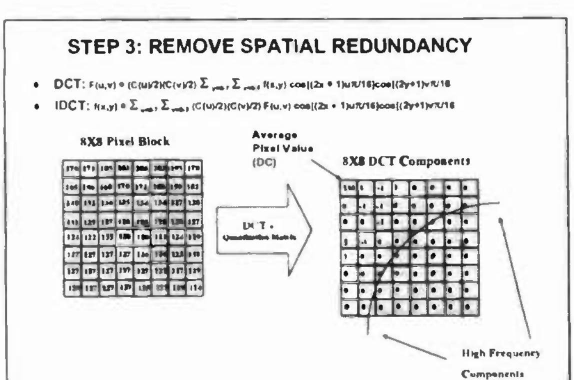












Special attributes for NTSC Standard Definition Networks:

MPEG encoders should have the ability to code in both 4:2:2 and 4:2:0 profiles. Network requirements will include the coding of Vertical Blanking Information. The coding of signals such as closed captioning information, ghost cancellation reference signals, Nielsen and other signals are accomplished through a combination of methods. The first method is to remove signals such as closed captioning and format them in user data packets for transmission. The second method is to code signals as video information. This is accomplished by turning the motion estimation off during sampling.

- Make sure the MPEG coding system you choose has the ability to code in 4:2:2 and 4:2:0 profiles as well as support and code VBI services.

Special attributes for ATSC High Definition Networks:

ATSC video encoders should be able to code multiple video formats. They should at a minimum, support 480i30, 480p60, 720p60 and 1080i30. In addition to supporting the multiple video formats, ATSC encoding system will have to support the Program Specific Information Protocol (PSIP), electronic program guide generation, closed captioning, interface to playout and scheduling systems as well support conditional access (CA) and subscriber management interfaces. Many broadcasters are seriously looking at the capability of using a portion of the capacity of the ATSC spectrum for data transmissions.

- *The ATSC system should support multiple video formats, be flexible in interfacing to third party equipment and be capable of expansion into CA and subscriber management systems.*

Audio Coding for Standard Definition and ATSC Systems:

Audio coding for broadcasters has proven to be more of a problem than video coding. The ability to code digital audio in a multiple coding environments has proven that low rate audio coding does not hold up well after 2-4 passes. This is true of both MPEG audio and Dolby AC-3 coding. Both MPEG audio and Dolby AC-3 work extremely well for the distribution applications they were designed to operate within. Broadcasters should be aware of the problems of digitally coding audio and make sure their systems will support new technologies such as Dolby E.

Data Broadcasting:

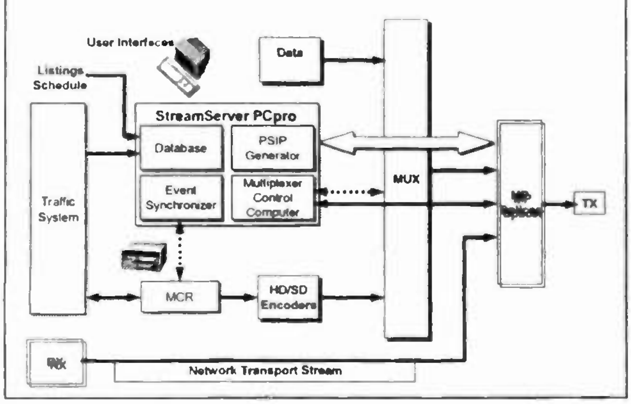
The ATSC allotted bandwidth opens doors on new opportunities. Table 3 of ATSC Document A53 permits multiple input formats for ATSC transmissions. Many broadcasters are considering the key question, "How do we make a business operate at a profit in this digital arena?"
The answers to this question loom on the horizon. The obvious answers may be in the convergence of the multimedia world where video entertainment meets computer applications. The ATSC system architecture should permit technologies such as data broadcasting, in whatever form it may take, to co-exist with technologies such as multi-casting. Today, we can easily multiplex a minimum of 1 standard definition video channel with either 720p60 or 1080i30 high definition signals. If we choose a lower vertical resolution multiple standard definition videos can be supported. Discussions for using additional channels with data are beginning to draw marketing investigations by entrepreneurs who are looking at a vast spectrum of possibilities for the ATSC spectrum.

- *Investigate your requirements for future bandwidth utilization in relationship to the system architecture. Flexibility of the architecture will prove to be an asset in the future.*

Multiplex Management:

The most often overlooked component of a digital transmission system may be the digital multiplex management. The digital multiplex manager is new to terrestrial broadcasters. Direct broadcast operators are familiar with the power of consumer interfaces. Digital broadcasting will introduce these concepts and practices to the terrestrial broadcaster.
Too often purchase decisions are based on price without regard to "system multiplex management" and capability. The digital systems for both standard definition and ATSC systems require multiplex management that is flexible and dynamic. The key to the future is the flexibility of getting ALL the information to the customer at the prescribed time in the prescribed format. The integrated multiplex management system is the gateway to building a scalable and flexible system. Broadcasters who choose to piecemeal the multiplex management issues may often find additional hidden costs in bridge software and unique interfaces. The digital multiplex management approach is new for terrestrial broadcasters, but it is inherent with the migration to digital technology and should be thoroughly investigated.

LOCAL STATION ARCHITECTURE



Digital Transmissions:

Satellite Transmissions:

Digital transmission technologies have continually evolved to higher order modulation formats. Designing the delivery system requires an in depth look at a number of parameters for network operators. In the U.S., the normal network satellite system has been built to support analog transmissions which require a very high carrier to noise ratio. The digital signals can take advantage of reduced bandwidth, higher power satellites, better low noise amplifier performance, and existing earth stations that are normally quite large.

The drawbacks to digital transmissions are few but those that exist can be fatal. Digital transmissions take advantage of numerous error protection coding techniques. When the system is capable of receiving the prescribed carrier to noise density performance requirements, there are few problems. The problems that exist come in the form of:

A. Carrier interference from a terrestrial carrier: In the analog transmission formats we used filters which could reduce or remove a portion of the spectrum where the problem existed. In a digital transmission any spectrum truncation or removal is almost always fatal to that carrier.

B. Threshold performance: We experience threshold performance effects on Ku band transmissions and networks with small, low gain, antennas. Threshold performance in the digital world are more severe than in the analog domain. The window of perfect performance to no performance is typically .5 dB to 1.0 dB wide.

C. Phase noise: Low noise converters can create reception loss anomalies which last from a few milliseconds to seconds. Typically poor phase noise parameters inhibit the performance of the error correctors to a point where normally acceptable level of carrier reception simply can not overcome the internally generated noise.

MPEG and Error Correction Overhead:

Satellite transmissions require error correction coding to overcome the inherent noise in the transmission path. One of the most misunderstood parameters is the amount of overhead (data not related to picture, audio, or data file contents) which is required. The MPEG transport packet is defined as a 188 byte packet. There are four bytes of each packet required for overhead functions. The 184/188 removes 2.1% of the payload in the actual transmitted bandwidth.

Reed Solomon is added for block error protection. The normal DVB transmission uses 188/204 Reed Solomon coding. This adds 16 bytes or approximately 7.8 % to the packets that are not dedicated to program data. At this point, after interleaving, the signal is then trellis coded with one of five DVB rates. The rates are 7/8, 5/6, 3/4, 2/3, or 1/2. This adds a minimum of 12.5% more overhead to the transmission. The graph below indicates actual payload information in relationship to total transmitted bandwidth for a QPSK transmission.

CODING AND INFORMATION RATE

MPEG Overhead = 4/188 = 2.13%
Reed Solomon Overhead = 188/204 = 7.84%
Total = 9.97%

Convolution Coding	MPEG & RS Coding	Total Coding	Total Transmitted Bit Rate	Total Payload Information Rate
7/8 = 12.5%	9.97%	22.47%	40 Mbs/sec	31.0 Mbs/sec
5/6 = 16.6%	9.97%	26.64%	40 Mbs/sec	29.3 Mbs/sec
3/4 = 25%	9.97%	34.97%	40 Mbs/sec	26.0 Mbs/sec
2/3 = 33.3%	9.97%	43.27%	40 Mbs/sec	22.7 Mbs/sec
1/2 = 50%	9.97%	59.97%	40 Mbs/sec	16.0 Mbs/sec

Satellite Transmission Parameters:

As a rule of thumb you can normally transmit a minimum total bit rate of approximately 45 Mbs/sec in a 36 MHz satellite transponder using the QPSK format. This will allow at least 2 high quality 4:2:2 broadcast channels at better than 15 Mbs/sec for video coding.
Many broadcasters are evaluating an 8 PSK transmission format. This will allow approximately 67 Mbs/sec total transmitted rate in a 36 Mhz transponder. This will allow for a minimum of 3 high quality 4:2:2 broadcast channels at better than 15 Mbs/sec for video coding. The calculations for both QPSK and 8 PSK use 7/8 rate trellis coding.

What does it cost to move to an 8 PSK type of technology?
The price to move to higher order modulation techniques is higher satellite power or larger antennas. In general it will cost you a minimum of 3 dB to move to 8 PSK for the same bit rate transmitted, or better than 4.5-5.5 dB to move to 16 QAM. In both 8 PSK and 16 QAM phase noise becomes a much more critical issue than in QPSK. The move to higher order modulation formats should be evaluated carefully with all existing network operating parameters considered.

ERROR PERFORMANCE

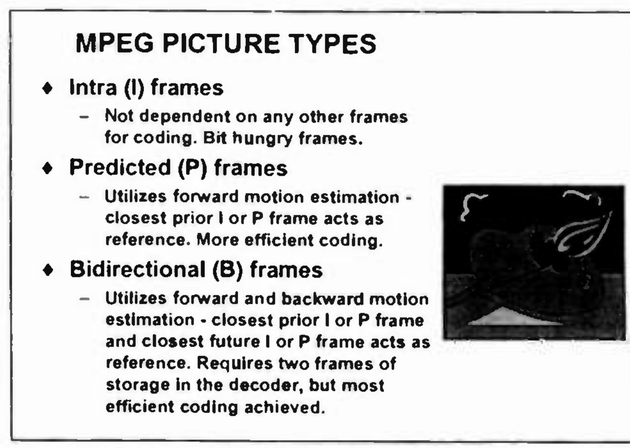
Modulation	Inner code rate	Spectral efficiency (bit/symbol)	Modem implementation margin (dB)	Required Eb/No (Note 1) for BER = 2x10(-4) before RS QEF after RS (dB)
QPSK	1/2	0.92	0.8	4.5
	2/3	1.23	0.8	5.0
	3/4	1.38	0.8	5.5
	5/6	1.53	0.8	6.0
	7/8	1.61	0.8	6.4
8PSK (optional)	2/3	1.84	1.0	6.9
	5/6	2.30	1.4	8.9
	8/9 (Note 3)	2.46	1.5	9.4
16QAM (optional)	3/4 (Note 3)	2.76	1.5	9.0
	7/8	3.22	2.1	10.7

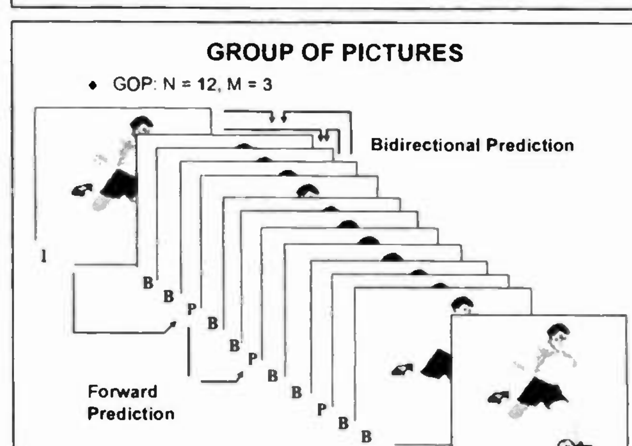
Terrestrial Transmissions for ATSC:

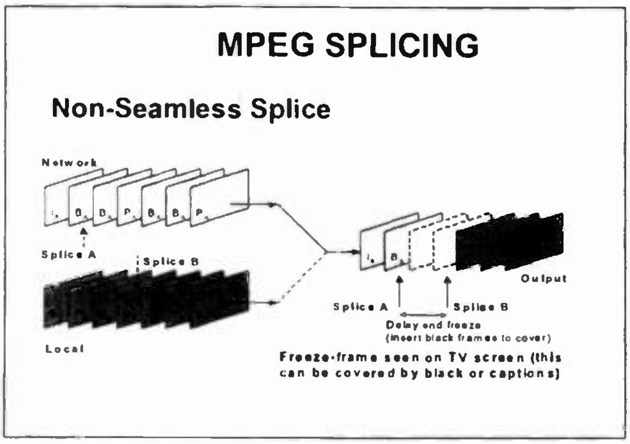
Terrestrial transmissions have been approved for 8 VSB modulation. This modulation allows a total payload of 19.39 Mbs/sec. The total transmitted bit rate is 32.28 Mbs/sec with the Reed-Solomon and trellis coding. These numbers calculate back to allow the transmission bandwidth to fall within the standard 6 MHz spectrum spacing in the U.S. The actual occupied bandwidth is 5.38 MHz. The only caution with this technology is the stability of the carrier. We have to be careful to be compliant with the SMPTE 310M specifications for the stability.
Within the 19.39 Mbs/sec we can utilize the payload for multiple applications. We will discuss some of the applications in the business section.

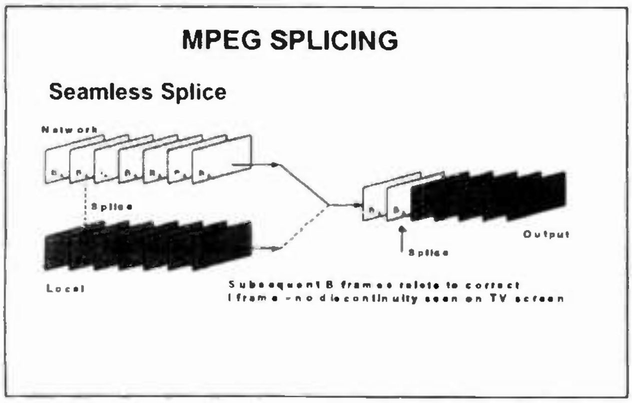
Digital Challenges:

The migration to digital baseband coding and transmission technologies have provided opportunities in many areas. We no longer have the ability to provide vertical interval switching of the video with audio following. We now deal with groups of pictures and frame types in a transmitted transport stream. We are faced with the challenge of splicing digital streams with the audio offset in the transmission path by 200 ms typically. To compound issues we must now take into consideration different resolutions, bit rates, conditional access and buffer management. These are displayed in the slides below:









Perhaps the single most significant change for the broadcast industry is a move to a new paradigm of managing a digital multiplex. The broadcast industry must take the time to understand the power of the digital multiplex and the ability it provides to reach the consumer.

Business Opportunities provided by digital transmissions:

The issues of the network distribution of programming to affiliates using digital compression and digital transmission is a straight forward engineering and economic decision. This move makes good sense for everyone. The program delivery will be of a higher quality, uniformly, across the network. The cost savings of distribution through reduced bandwidth utilization can not be argued. Again, this makes perfect business sense. The only issues here will be how to choose the vendor and system that will fit your requirements for the future.

The issues of how the affiliates will distribute digital signals has been evaluated by many parties. The cost of upgrading to allow a digital transmission is averaging between 2 and 4 million dollars per site for the tower work, transmitter, modulator and compression equipment. This price does not include studio gear. This investment requires thorough examination of the question, "Where is the payback?"

The possible paybacks:

The FCC requires a minimum of one free video channel in the multiplex. This channel must be compliant with the Table 3 requirements of the A/53 document.
Some potential additional revenue resources within this bandwidth may occur as multiple channel use for regional sports programming, movies, pay per view events, educational classes and local audio services. The bandwidth could also support data and internet type services. Some of the new services could include local news, weather, sports, local events, dining and public service applications. These applications could be served by consumer televisions, cable type set top converters, or personal computers. The addition of conditional access would open a plethora of new interactive opportunities with telephone or cable return paths. These applications could include banking services, shopping, interactive services and all types of electronic commerce. The new digital services may prove to be the true catalyst for the convergence of multi-media and broadcast video applications. Over the next couple of years, we may see new innovative approaches which will make this bandwidth invaluable.

REVIEW OF THE ATSC IS TOP DOWN COMMITTEE'S REPORT ON FINDINGS

Shannon L. Skemp
James A. Kutzner
Thomcast Communications, Inc.
Alexandria, Virginia

INTRODUCTION

With the rollout to digital television underway, television stations engineers should look toward standards setting bodies for guidance with their implementation plans. One of the most conclusive works to date is that of the Advanced Television Systems Committee's Top Down Committee. The following paper is not intended as a replacement to the ATSC Top Down Meeting Implementation Subcommittee Report on Findings. Rather, it is intended to serve as a review of the report itself. The authors of this paper encourage all interested parties to read the report in its entirety, which can be found in electronic form on the ATSC website at www.atsc.org. A hard copy of the report can also be found in the December 1998 SMPTE Journal.

GOALS OF THE TOP DOWN PROCESS

Over the last few years, several industry bodies, including the ATTC, PBS, and the Model Station, developed various implementation scenarios for DTV. To effectively serve the needs of their membership, the ATSC undertook the task of providing an all-encompassing work that would incorporate existing scenarios and draw upon new ideas within the industry, thus examining the ATSC Standard and its actual implementation from the "Top Down". The ATSC chartered its Implementation Subcommittee under the efforts of the IS-S3 Subcommittee on Station Issues to create the Top Down Committee for this purpose.

The primary goal of the Top Down Committee was to create an inventory of interface standards that could potentially exist in a typical station, regardless of the implementation scenario. This inventory would then be used as guide to pinpoint where multiple standards existed, where multiple standards conflicted, and in turn, where news standards and technology needed to be developed. Once this primary goal was achieved, the Top Down Committee was able to work towards the development of strawman implementation scenarios for each of the potential internal and external station infrastructure.

Although the work of the Top Down Committee was thorough, it is important to note that it is ongoing work which did not end with the publishing of the report. In order to insure the work continued and the ultimate goals were met, the report findings were fed back to the parent ATSC Implementation Subcommittee and other standards bodies, such as SMPTE. The ATSC Implementation Subcommittee disseminated the information and distributed the remaining tasks among its subcommittees and referred other issues to the ATSC Technology Committees for further progress.

TOP DOWN MEETING STRUCTURE

It was apparent that the focused attention of the Top Down Committee members was essential to complete the rigorous goals which had been chartered by the Implementation Subcommittee. Therefore, the Top Down Steering Committee scheduled two separate multi-day events at Sony's facilities in New York in order to attract participants with the appropriate knowledge and experience. The unparalleled collaboration of over 70 industry wide professionals with broadcast, manufacturing, and consulting backgrounds proved to be the needed formula for achieving the goals at hand.

In order to facilitate the progress of the Top Down Committee, a general drawing depicting all functional blocks and interfaces possible in a typical television station was created to serve as a type of "map" to the industry. It is important to note the drawing does not depict one preferred solution. Rather, it denotes several implementation variations and should only be used once the reader has acquired a full understanding of the Top Down Report. The Top Down Map provided an exhaustive review of a station's signal plane and allows users to view the function of the interface, as well as its hierarchy of supporting standards. Similar drawings that focused on Timing, Control, Monitoring, Data and Audio Planes were also created and are included in the report for reference.

In order to utilize the time of the participants in the most efficient manner, the work of the Top Down Committee was subdivided into several breakout sessions. The various sessions included the following topics: Multiple Video Formats, Encoding and Multiplexing, Station Inputs and Outputs, Data Services, Redistribution, Audio, Control Plane, Timing Plane, and Monitor Plane. Plenary sessions were also used in between each of the breakout sessions cycles in order to facilitate the sharing of information across the multiple groups.

GOALS OF THE TOP DOWN REPORT

The final Report on Findings was created with three main goals in mind. First, the report was intended to serve as guide to assist stations in planning new DTV implementation strategies. For the stations that have already embarked on their DTV plans, the Report can be used as a "sanity check" for their existing implementation strategies. Finally, the complete table listing of possible interfaces and associated standards can be used a reference resource for standards.

REVIEW OF THE TOP DOWN REPORT

In the following sections of this paper the key issues and findings of the various breakout groups will be highlighted and ongoing work will be noted where applicable.

Video Formats

The primary focus of the Video Formats group was on the concept of a "Plant Native Format" due to the large number of possible video formats available to the marketplace for content delivery. The philosophy behind this proposal stemmed from early indications that affiliate stations are planning to implement a single format DTV facility, while continuing operation of the legacy NTSC facility. It was noted that at a minimum, the DTV program facility must include all NTSC operations. In order to facilitate the concept of a plant native format, smart format converters that can automatically detect and switch between various video formats sources will be implemented along strategic inputs and outputs of the DTV facility.

The criteria for choosing a plant native format is dependent upon economic constraints, legacy equipment issues, and ease of conversion. The choice of emission formats is another issue that will be determined by station needs and may vary at different times of the program day.

In the Video Formats group several issues were noted for further consideration. One of the most controversial issues was that of the discrepancies between the 720 pixel/line production format and the 704 pixel/line emission format. The group noted the preferred solution would be to match the production and emission formats. This issue was carried forward following the Top Down Report and is currently under discussion. An additional recommendation of the Video Formats group was for a common serial transport standard that would encompass all production formats in the DTV environment. This continues to be an ongoing issue. The group also addressed the need for standardization of video format metadata in the uncompressed domain for use in smart format conversion. Finally, the subgroup recommended further in-depth study in the area of latency with emphasis on metadata for the time-stamping of video, audio and data.

Encoding and Multiplexing

The scope of the Encoding and Multiplexing group addressed all compression-related equipment in a television station from input to output. This included NTSC/SDTV/HDTV encoders and compressed domain multiplexers, splicers, and conditional access. The members of this group successfully completed an exhaustive inventory of signal interfaces for existing and projected practice and began work toward a preferred near-term solution. The issues noted for further consideration included format converters at boundary points, latency issues, the addition of functional blocks for synchronization, consideration of redundancy, examination of keyer and splicer details, and the exploration of the use of the AES-3 stream for carriage of ancillary/metadata. The efforts of the Encoding and Multiplexing group are continuing under the ATSC IS-S3 subcommittee on Station Issues.

Station Inputs and Outputs

The scope of the Station Inputs/Outputs group focused on the implementation of network feeds and a variety of other contribution links, to and from the station, along with communication links between the studio and transmitters for both NTSC and DTV services. The group successfully cataloged a full compliment of interfaces for both SDTV and HDTV. The SDTV interfaces covering legacy common and private carriers included satellite, microwave, copper fiber, optical laser, and off-air feeds. The HDTV interfaces covering progressive common and private carriers included digital satellite, microwave, fiber LAN, optical laser, and off-air feeds. Once the interfaces had been documented, the members of the subgroup proposed two strawmen implementation scenarios for studios that are not co-located with the transmitter. The first scenario maintained the existing analog link for transmitting the NTSC feed between the studio and transmitters, while providing a separate digital link for the HD feed. The second scenario provided for a multiplex of both the NTSC and HD feeds into a new digital radio or fiber channel.

The Station In/Out group also defined several issues for further study within the Implementation Subcommittee. These issues included further definition of interface standards. The group also felt it important to alert common carriers of the digital requirements. Specification of the contribution format and associated data rates by the Networks and other program suppliers is now underway. Finally, latency issues were once again addressed and further exploration of adding an IFB channel to the bitstream was suggested.

Data Services

The Data Services group limited its scope to range of data signals associated with the final emitted signal, which include both content and system data. More specifically, the Data group focused on four types of data including; picture user data, data carried on a separate PID for both program-related and non-program-related data, and system data. In order to facilitate examination of these four types of data, the Data group created a simplified Data Plane drawing to be used in conjunction with the Top Down Map. One of the key components of the data plane was the data bridge which serves as a conduit for the exchange of data signals between the NTSC and DTV plants. Additional functional blocks were added to facilitate data extraction and insertion between the NSTC and DTV plants. Another key component of the Data Plane is the Data Server, which handles all data broadcast signals, program and non-program related, except for the closed captioning signals.

Upon completion of the Data Services group's work, several open standards issues involving EIA-708 were identified. Work now continues in addressing the standardization of the "Closed Captioning Food Chain". It was also noted that standards development was needed for the network delivery of data to the affiliate. Finally, interfaces between the data plane and the management and control planes need to be implemented. Most of the efforts of the Data Plane group are continuing forward within various groups.

Redistribution Signals

The scope of the group on redistribution covered the issues relating to the output of the DTV signal to destinations other than the DTV transmitter. These issues included the identification of the various destinations and associated data rates of the delivered signal, which may be higher than the 19.4 Mbps. The variety of destinations likely to receive the DTV programming included: cable head-ends for broadcast feeds, independent cable feeds for news or alternate programming, satellite uplinks for distribution of syndicated programming to other stations, microwave systems which may include translator feeds, telco systems, and dedicated fiber feeds.

Issues for further consideration, highlighted by the Redistribution group, included questions related to the clarification of existing standards. More specifically, the question was raised as to whether standardization for all of the aforementioned modulation schemes existed.

Audio Plane

The scope of the Audio Plane group focussed on the audio issues from the point of view of the affiliate station for three different infrastructure scenarios. They included: 1) production and distribution of 5.1 channels of audio 2) downmixing of 5.1 channels to 2 channels for SDTV/analog services and 3) audio pass through. The Audio Plane group based their analysis for the production and distribution of 5.1 channels of audio on the assumption of an existing digital infrastructure in the plant. This infrastructure would allow for the distribution of multichannel audio via multiple AES pairs. In an alternate scenario, the digital infrastructure would allow for the distribution of rate reduced audio, multiplexed with the associated metadata into a single AES pair. In order to emphasize the importance of the metadata, which should ideally be generated during the post-production process, the Audio Plane group provided for a separate routing layer for the metadata on the Audio Plane drawing. In the same vein, the Audio group provided additional functional blocks for metadata authoring tools and bridges.

In order to facilitate the station's ability to simulcast, direct paths between the HD and SD routers must be provided. In addition, a

Dolby Surround (2 channel) downmix can be made from the HD multichannel source for the SD service. It is also important to note the concept of "Plant Native Format", previously discussed in the Video Formats section, does not apply to the audio portion.

When dealing with the pass through of the Dolby Digital (AC-3) emission rate signal, decoding of the signal for distribution should be avoided at all costs due to multi-generation coding losses. Several other issues also remain pending for further exploration by the ATSC. These issues include the possible carriage of data in the rate reduced stream, time stamping of the audio and video streams and A-V "lip-sync" dictated by the video coding for addressing latency, and better definition of metadata authoring facilities.

In order to reiterate the points already discussed, the Audio Plane group concluded their report with the following four points: 1) the distribution format for DTV programming must carry six audio channels 2) the affiliate's distribution system should carry both a six channel and a two channel soundtrack 3) a signal path must be provided for the metadata and 4) a contribution quality audio coding scheme is needed.

Control Plane

The first thing the Control Plane group did was to recognize the need for a fully integrated control plane and recognize the importance of this control plane compared to any other function within the station. In order to support the importance of this control the significance of metadata and its management in future operations was also recognized.

In order to permit broadcasters in DTV operations the same flexibility they now enjoy in analog broadcasting, "a large number of commands, parameters, and responses must be passed among devices". The control and intelligence behind that control must be far above the current basic machine control functions of starting, stopping, and queuing.

The group broke down the control tasks into Low Level Control and High Level Control functions where the low level control represented roughly the current level of control over equipment. High Level Control was then the control expected and required in future high-speed networks.

Basic functions of the Operator, System Scheduler, Facility Resource Manager, and Status Manager were identified. All of these functions exist in one form or another in today's station, many are merely manual personnel operations. In the plant of the future where additional services are delivered and operations occur at much faster speeds, and where additional staff may not be available to perform these operations even if they could, automated high level and intelligent control systems will be required.

Three areas were identified in the basic television station: the "Input" area that includes production, post-production, acquisition, etc.; the "Baseband" area that includes storage, playback, and release switching; and the "RF" area that includes everything following the release switcher. The first area has traditionally not been very automated and is not expected to require an immediate change. The second area is currently often highly automated and the level of automation is expected to increase incrementally. The third area has not traditionally been automated, but this will change dramatically and in a fairly short time period. Consequently, the group elected to focus on this third area.

Much of the work of the SMPTE-EBU Task Force was brought to the discussion of the Control Plane group. The group realized that much of the groundwork has been done and that the next steps were to apply the models developed to real situations.

Time-aware control was also noted and discussed. Three temporal levels of control were identified:

Level 1, where systems are disassociated from time (time frames of minutes or greater)

Level 2, which are time-aware devices (one video frame aware, or so)

Level 3, which are instantaneous response devices (time frames of microseconds)

Further work includes continuing to apply the SMPTE-EBU Task Force models to real applications. Also, investigations into the "Input" and "Baseband" areas need to take place.

Timing Plane

The Timing Plane group examined a number of areas with respect to the layering of DTV facilities onto existing plants.

The group felt that all broadcast facilities will operate their plants at a 59.94 Hz frame rate since they will continue to operate their NTSC plants for some time to come. The group felt that the development of a plant that used both 59.94 and 60 Hz would be impractical.

The group also discussed the conflicts that exist between time-of-day, drop frame time code, NTSC rates, and MPEG clock frequencies, and then discussed methods of bringing these into alignment. The changes would be slight but dramatic in the sense of changing basic operating conditions, however, the added complexity to automation and control functions that is created by not making these or similar changes could be substantial.

Lip sync issues were also considered crucial and methods were discussed to prevent problems from occurring. The advantages of a common reference for the industry were discussed. GPS was discussed as one potential common reference.

The problems of latency in live remote situations were also discussed. MPEG compression will bring forth substantial latency which will in turn require new solutions to the problems of remote queuing and IFB channels.

Monitoring Plane

The scope of the Monitoring Plane group was to review the new infrastructures, both compressed and uncompressed, to determine what new monitoring and measurements were needed in television plants that fed ATSC/NTSC and ATSC-only transmission or distribution facilities. Included were not only the requirements for new techniques and monitoring systems, but also confirmation of continuing old ones. Monitoring, test, and signal injection points were identified, and associated equipment to support DTV station operation and maintenance was to be identified.

The inclusion of more automated monitoring systems was also reviewed by the Monitoring Plane group. The group, taking its queue from the state of affairs in television engineering management, realized the increase in provided services did not mean an increase in staffing levels. Consequently, more intelligence will be required of the monitoring and measurement systems, and more training will be required of the station personnel.

The conclusions of the group began with confirmation that all of the existing monitoring functions and techniques in existence in today's NTSC or 125M/259M plants will remain until they are replaced with automated functions or until those plants are retired.

New test signals, test equipment and test points would need to be identified for providing the needed monitoring and testing of MPEG-based systems. Additional automated signal identification and signal tracing equipment would be needed in stations as they prepare to offer additional services.

One significant open issue is the network to station and related audio distribution format. If a higher quality format capable of concatenated codec cycles (e.g., Dolby E) becomes the norm then a new array of monitoring and test functions will be required, certainly in the compressed domain and perhaps in the uncompressed as well. In the new master control areas the monitoring and control functions will be merged as they are today, but there will also be new requirements to ensure that the compressed signals are handled and managed appropriately.

Another aspect that must follow when the service approaches reality is test and monitoring of data broadcasting systems. And, as the control plane is expanded with the control of new technology test and monitoring of those new systems must also be developed.

More test and more monitoring layered upon existing systems would simply mean more work for the same engineering staff. What is required overall is more intelligent and automated monitoring functions, not just more monitors.

CONCLUSION

The rollout to digital television can be arduous for station engineers who must implement an evolving technology while maintaining the stability of their existing facilities. In order to ease this transition it is essential for the engineers and station owners to draw upon the integral works of the ATSC, SMPTE and other standards bodies as a point of reference. The Report on Findings of the ATSC Top Down Committee is an ongoing work within the open forums of the ATSC Implementation Subcommittees. Participation in the ATSC by the ultimate users, broadcasters, is key to gaining first hand knowledge and understanding of the core issues as we move forward in the development of digital television. Readers are encouraged to participate in any way in the work of the ATSC IS.

REFERENCES

1. "ATSC Top Down Meeting Implementation Subcommittee Report on Findings", Copyright 1998. Advanced Television Systems Committee.

ACKNOWLEDGEMENTS

The authors of this paper would like to acknowledge the hard work and dedication of all of the members of the ATSC Top Down Committee. (Please see the "Recognition of Particpants" section of the ATSC Top Down Meeting Report on Findings for a complete listing of the members.) The authors would also like to give special thanks to the members of the Top Down Steering Committee and Top Down Group Chairs. They include the following:

Terry Smith, Sarnoff – Co-Chair
Shannon Skemp, CDS – Co-Chair
Art Allison, NAB – IS Chairman
Jim Kutzner, CDS
Tom Gurley , MSTV
David Horowitz, Consultant
Bruce Penney, Tektronix
Matthew Goldman, DiviCom
Bruce Miller, WHD-TV
Wayne Kube, A.H. Belo
Bernie Lechner, Consultant
Steve Lyman, Dolby
Bill Miller, ABC
S. Merrill Weiss, Merrill Weiss Group

DIGITAL TELEVISION INTEROPERABILITY – ISSUES AND PROGRESS

Eric Gsell, Charles Einolf, Walt Husak and Alexandra Loubeau
Advanced Television Technology Center
Alexandria, Virginia

Abstract

The Advanced Television Technology Center (ATTC) has provided system integration for various digital television (DTV) demonstrations around the world. This experience has shown that Interoperability is a key issue for the successful implementation of DTV. It has not been uncommon for DTV encoders and decoders to be "ATSC-compliant" and yet not be able to operate together properly. Now, with the rapid deployment of DTV broadcasting in the United States it has become crucial to ensure the compatibility of encoders and decoders and to determine the extent of functionality. ATTC has undertaken a formal program to test the interoperability between all-available DTV encoders and multiplexers with DTV receivers.

1. Introduction

As the first generation of Digital Television (DTV) encoders and commercial Integrated Receiver Decoders (IRD) reach the market place, there has become a concern regarding the interoperability of these units. The initial concern of encoder and receiver manufacturers has been accurate reproduction of video and audio material. However, more recently the issue has progressed to the inclusion of Program and System Information Protocol (PSIP)[1] material, closed captioning and ancillary data services.

While many encoder manufacturers have made efforts to ensure that their device is compatible with as many receivers as possible, the consistency and objectivity of those tests has been of a sporadic nature. The Advanced Television Technology Center (ATTC) with support from its membership has embarked on a test to verify the extent of DTV interoperability. The ATTC test verifies interoperability of all major encoder manufacturers with the IRD's of the ATTC membership.

The scope of the ATTC tests entails the research and acquisition of bit-streams from all of the major High Definition and Standard Definition encoders in as many of the 18 designated ATSC DTV standard formats. Of specific interest were encoders capable of inserting Program Association Tables (PAT), Program Map Tables PMT), PSIP tables as well as solutions regarding the carriage of Closed Captioning data. Captured Bit-Streams are then played back into the IRD under test using a simulated transmission path generated in the ATTC RF Test Bed. Each IRD is tested for its ability to properly decode the available encoded material.

This paper describes the DTV Interoperability Test Facility, the test procedures established to ensure a repeatable test process, and the initial test results. It is anticipated that DTV interoperability testing will be an on-going process as both encoder and IRD manufacturers continue to develop their products.

2. DTV Interoperability Test Design

The design of the DTV Interoperability Test involves the following seven major phases:

- Interoperability Test Facility Design
- ATSC-compliant Base-Band Material Design and Editing.
- Interoperability Test Facility Calibration
- Encoder Interfacing and Test Documentation
- Encoder Bit-Stream Acquisition
- Encoder Bit-Stream Verification
- Encoder and IRD Interoperability Testing

2.1 Interoperability Test Facility Design

Created in 1988, and continually updated, the ATTC laboratory is the premier test facility for DTV system testing in the United States. The plant was originally designed to test the proposed HDTV transmission systems solicited by the Advisory Committee on Advanced Television Service (ACATS). It was also the test facility for the Grand Alliance DTV transmission standard, which comprised the bulk of what has become the Advanced Television Systems Committee (ATSC)[2] standard. The plant is designed to generate NTSC and HDTV base-band video and audio, which was used to test the proposed Advanced Television Systems.

SDTV video and audio signals are produced using two Sony DVR-10 D-2 VTRs. Video is produced as composite analog 525i video.

HDTV base-band video and audio is produced using three Sony HDD-1000 VTRs. The VTR, which uses one-inch digital videotape, is capable of producing and recording the full bandwidth 30-Megahertz signals of standard 1125i HDTV video. It also has the capability of producing up to eight separate audio channels.

Each HDD-1000 is equipped with an ATTC specified and designed Format Converter to accommodate the use of alternate high bandwidth video signals. The original ATTC Format Converter design allowed for the storage and reproduction of 525p, 787p, 1050i and 1125i video material on the HDD-1000 videotape. Since the ATSC standard established 750p, the Format Converter was modified to accommodate this format. This change gives the HD-1000 VTR the capability of producing all of the ATSC High-Definition full bandwidth formats.

2.2 ATSC-Compliant Base-Band Material Design and Editing

The ATTC has a large library of video material used throughout the DTV development and test process. The material used for the Interoperability testing was a subset of the original ACATS test video. The material was designed to test different aspects of the encoding and decoding process. Material ranges from specially designed still, full motion video and graphic images. Film material in a variety of formats and speeds is also used.

The specific video clips used for the testing of Interoperability were chosen to best represent video styles used in television broadcasting. Clips designed to elicit undue system stress or "Bust" the encoders were omitted. Test signals were edited on the front end of the tapes to allow future reproduction of recordings. An A/V sync test was included as well.

2.3 Interoperability Test Facility Calibration

The ATTC Interoperability Test Facility is thoroughly calibrated prior to video encoding. HDTV reference signals generated by a Tektronix TSG 1001 are used to measure levels throughout the video chain. SDTV video is measured similarly using a Tektronix TSG 170A. Both video systems use Pulse & Bar to measure peak video levels. **Table 1** illustrates the required video signal levels at the various analog measurement points.

Table 1 – Required HD Video Signal Levels in the ATTC Interoperability Test Facility to Ensure Quality Signal Replication.

Peak to Peak	140 IRE	1Volt
Peak Level	100 IRE	700 mV
Black Level	0 IRE	0 mV
Sync Level	±40 IRE	± 300 mV

2.4 Encoder Interfacing and Test Documentation

Most of the encoder manufacturers use the newly implemented SMPTE 292[3] standard for Serial Digital Interface as the primary video input for their machines. The ATTC is an analog RGBS/YPrPb and parallel digital based plant. For the purpose of the Interoperability testing it was necessary for ATTC to acquire signal transcoders. The majority of the testing done at ATTC used the Panasonic HAD500 Analog to Serial Digital Transcoder. The HAD500 converts YPrPb video signals into HD SDI for 1125i video. Panasonic provided ATTC with a prototype A/D converter for 750p based video. Of interest, one encoder required the use of Cyclical Redundancy Checks in video lines to establish sync. The HDD-1000 was built before CRCs were implemented. In order to be able to lock to the encoder, ATTC used a YEM-1125 DAC, which would insert the CRC information into the video lines, unlike the HAD500.

SDTV video signals were typically input to encoders from the composite output of the DVR-10, but in several instances ATTC used a D-5 in E-E mode to transcode from composite to SDI.

Audio input differed depending on the specific design of the encoder and whether it was SD or HD capable. Most SD encoders used internal AC-3 encoders so audio would simply be input from the DVR-10 AES/EBU outputs. HD encoders required the use of an external encoder for AC-3 pass through. ATTC used a Dolby DP569 six-channel encoder to input audio to the encoders.

A thorough evaluation and documentation of the entire encoder system was completed to ensure the reproducibility of the test results. All of the encoder inputs and outputs were mapped out as well as external controls. System control software and the interface to the encoder was also documented to establish revision control. If the unit was supplied with a companion decoder values were logged for that device as well. **Table 2** illustrates the documentation parameters recorded for both encoder and decoder.

Table 2 -Codec Documentation Parameters Recorded to Ensure Repeatability of the Interoperability Tests

Encoder Parameters	Decoder Parameters
Model Number	Model Number
Serial Number	Serial Number
Software Version	Software Version
Firmware Version	Firmware Version
PSIP Generator	
Bit-Rate Control	Bit-Rate
Video Formats	Video Formats
Field/Frame Rate	Field/Frame Rate
Audio Mode	Audio Mode
Audio Rate	Audio Rate
Video PIDs	
Audio PIDs	
Data PIDs	
Chroma Sampling	
Film Mode	
Output Interface	Output Interface
Input Interface	Input Interface
MPEG Parameters	
MPEG Control	Error Concealment

2.5 Encoder Bit-Stream Acquisition

The task of scheduling available DTV encoders to be at the ATTC laboratory at the same time as the manufacturer receivers proved impractical and cost prohibitive. Consequently, bit-streams were recorded on a server prior to the actual Interoperability Testing. This process has allowed ATTC to archive the encoded test material for future tests and in-depth analysis.

Recording transport stream material entails the use of machines capable of storing vast amounts of information quickly. The average bit-rate of an ATSC Terrestrial Broadcast Signal is roughly 2.4 MB/s. ATTC utilized the Sencore Stream Station SV953 Bit-Stream analyzer for recording, analysis and playback of encoded transport material. The Sencore SV953 is also capable of interfacing in the following transport stream formats: DVB-SPI, DVB-ASI, Serial ECL and SMPTE 310[4]. **Figure 1** gives an example of the encoding set up.

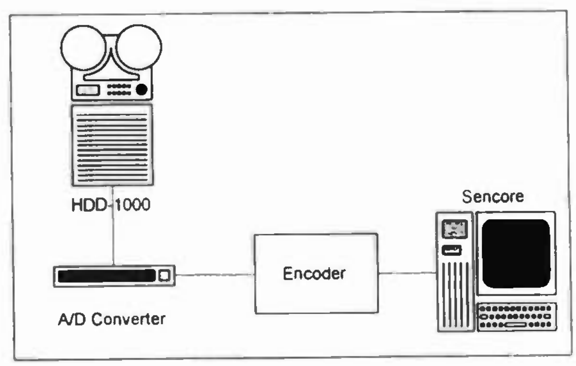


Figure 1-Encoder Bit-Stream Acquisition

Because of the extremely high bit-rate of ATSC transport material, each 14-minute transport stream fills approximately 1.9GB of hard drive space. ATTC built a 68 GB hard disk array to store the encoded transport streams, using ultra-wide SCSI disks, which have a 40 MB/s throughput.

The primary encoding was a simple transport stream of video and audio with no data services. Supplied encoders capable of multiplexing PSIP and program related data (**see Table 3**) would have a second stream created with that data multiplexed into the stream. Of all the encoders tested only one manufacturer was able to provide full PSIP data.

Table 3 -Definitions of Tables Associated with Program Information and PSIP Data

MPEG	Acronym Definition
PAT	Program Association Table
PMT	Program Map Table
MGT	Master Guide Table
RRT	Rating Region Table
VCT	Virtual Channel Table
EIT	Event Information Table
ETT	Extended Text Table
STT	System Time Table

2.6 Encoder Bit-Stream Verification

Two different systems were used to verify that the encoding was successful. First, the encoded stream is played out of the Stream Station and modulated to a broadcast frequency. Then the 8-VSB signal is received by a test IRD (see **Figure 2**). For the purposes of Interoperability Testing, ATTC used the ITS modulator. Second, the encoded material is analyzed with the Stream Station's transport stream analyzer, which looks at the packet headers of the stream. The analysis details the validity and organization of the headers for video, audio and all other tables including PSIP. When the encoding process was complete, the bit-streams would be transferred to 8mm Exabyte tape for archiving.

2.7 Encoder and IRD Interoperability Testing

The actual playback of the transport streams into the IRDs is the simplest part of the whole testing process. As with verification, the bit-stream is modulated and upconverted (see **Figure 2**). With the successful reception of a sample signal the actual Interoperability Testing begins.

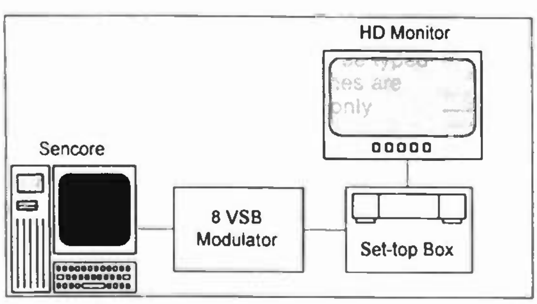


Figure 2-Interoperability Test Setup to verify successful video encoding

Interoperability was simply defined as the accurate reproduction of transport stream data. If data on the bit-stream was displayed or decoded in some manner by the IRD, the test is declared a success. Compiled test data for the IRD under test and segments of failed bit-streams are supplied to the IRD manufacturers. If a stream was unable to be decoded by any receiver or the analysis found the stream to be illegal the encoder manufacturer is notified. Complete data of the entire Interoperability Test is recorded and archived at ATTC. The following list summarizes the encoders made available to ATTC for inclusion in the Interoperability Tests.

HDTV Encoders Tested

- Mitsubishi/Tektronix 1100HD
- Scientific Atlanta Power VU HD
- NDS
- General Instruments
- Harris/Lucent Flexicoder

SDTV Encoders Tested

- Scientific Atlanta Power VU SD
- NDS
- Harris/Lucent Flexicoder
- Divicomm MV40

The following summarizes the receivers made available to ATTC for inclusion in the Interoperability Tests.

IRD's Tested

- Panasonic
- Sony
- Philips
- Mitsubishi
- Samsung
- Pioneer/Sharp

3. Conclusions of the Interoperability Tests

Preliminary results suggests that encoders and receivers have achieved a high degree of compatibility. The testing results can be categorized in several areas: Primary video and audio decoding, PAT/PMT utilization; and PSIP utilization.

The number of IRD's that successfully decoded video and audio service was almost perfect with only one IRD manufacturer thus far having a failure on one encoder's stream.

Utilization of PMT's and PAT's for basic program information by most receivers was also highly successful. This makes sense in part because PAT and PMT tables, which are actually MPEG structures, have been the first areas of the ATSC standard to be finalized. In fact, most of the early IRDs were designed only to decode these tables, since PSIP tables were not established until relatively recently.

The integration of PSIP data into IRDs varied greatly, with only one manufacturer actually utilizing the majority of the tables. Most manufacturers incorporated the use of the more critical tables such as the Virtual Channel Table and the Master Guide Table to create electronic program guides. Some groups implemented use of the Extended Text Table to provide more program information to the viewer and only one manufacturer used the System Time Table to set the IRDs master clock.

As far as closed caption systems and ancillary data are concerned, the jury is still out. As of this writing, there are currently three groups attempting to create a standard for inserting closed captioning data and at least two groups are testing systems for implementation of ancillary data services. Until those questions are finalized no manufacturer can produce those systems, but every group has built hooks to their devices for acceptance of those standards.

The first generations of DTV encoders and IRDs have made significant progress. They have been able to successfully integrate into an evolving standard and predict future changes with relative accuracy. Changes are regular, sometimes weekly especially in the case of early system designs. Every encoder manufacturer and many IRD manufacturers have built their systems to be easily updated with firmware and software upgrades.

It has become apparent to the ATTC and it's membership that Interoperability Tests will have to continue for at least the next three years as design changes and redesigns of both encoders and receivers continues. ATTC is currently working out arrangements with all of the major encoder manufacturers to update and append ATTC's library of bit-streams to ensure the operability of all IRDs.

The first year of DTV implementation has proved to be a very busy and prosperous year. In the beginning of 1998, there was only one encoder manufacturer selling HDTV encoders on the market, now there are six. Prototype receivers and laboratory models were the only decoders available, but now IRDs and home receivers are becoming available to the consumer. In early 1998, fewer than 12 stations were broadcasting DTV signals on a regular basis, now there are over 40. Broadcast DTV has come a long way since the FCC's acceptance of the ATSC standard and the DTV rollout began. The Advanced Television Technology Center continues to provide a crucial role in the development and Implementation of DTV in the United States.

References

[1] Doc A/65 ATSC Program and Specific Information Protocol for Terrestrial Broadcast and Cable 12/23/97

[2] Doc A/53 ATSC Digital Television Standard 9/16/95

[3] SMPTE 292M Television-Bit Serial Interface for High Definition Television Systems 1996

[4] SMPTE 310M Television-Synchronous Serial Interface for MPEG-2 Digital Transport Streams

DAB: The Global View

Sunday, April 18, 1999 9:30 am - 12:00 pm

Chairperson: Milford Smith
Greater Media Inc, East Brunswick, NJ

9:30 am **The Broadcasting-Satellite Service (Sound) Using 2.6 GHz Band in Japan**
Shuji Hirakawa
Toshiba Corporation
Kawasaki, Japan

*10:00 am **The Status of Digital Sound Broadcasting**
Don Messer
US Information Agency/ Int'l Broadcasting Bureau
Washington, DC

10:30 am **Draft Standard of Digital Terrestrial Audio Broadcasting in Japan**
Masayuki Takada
NHK
Tokyo, Japan

*11:00 am **DAB - The Current Situation in Germany**
Eberhard Siebert
Deutsche Telekom, AG
Freiburg im Bris, Germany

11:30 am **Audio Processing for DAB and the Internet: How Audio Quality and Intelligibility Can Be Improved In the Data Reduced Environment, Using Dynamics Control**
Frank Foti
Cutting Edge
Cleveland, OH

---

*Papers not available at the time of publication

THE BROADCASTING-SATELLITE SERVICE (SOUND) USING 2.6 GHz BAND IN JAPAN

Shuji Hirakawa (Toshiba Corporation)
Chairman of Broadcasting System Working Group for BSS (Sound)
Association of Radio Industries and Businesses, Japan

1 Introduction

This document provides the recent progress of system development to establish the Broadcasting-Satellite Service (Sound) using 2.6 GHz band in Japan. From the last September, ARIB (Association of Radio Industries and Businesses) is conducting indoor and outdoor experiments to corroborate the feasibility of its system parameters.

This system will be approved as the BSS (Sound) in Japan in the second quarter of 1999 after the corroborative experiments if they will be carried out successfully.

2 Summary of the system

Figure 1 is the system diagram of this system. This system is consisted of a broadcasting satellite, an earth station, terrestrial gap fillers, and receivers.

Broadcasting signal is transmitted from an earth station to a broadcasting satellite at first, using 14/11 GHz bands, for example. Broadcasting signal is converted from 14/11 GHz bands to S band (2.6 GHz) in the satellite. S-band signal is amplified using a satellite transponder and/or amplifier up to a desired level and this signal is broadcasted onto Japanese terrain using large transmitting antenna of the satellite. This satellite is supposed to have a geo-stationary orbit.

Main broadcasting programs of this system are sound services with CD quality in the first stage and multimedia services including data broadcasting in the next stage.

Listener / Viewer of this service can receive the broadcasting signal via the satellite using small antenna. To realise enough EIRP for mobile reception, geo-stationary satellite will be implemented a large antenna and high power transponders / amplifiers.

Major issues above Giga Hz band radio transmission are shadowing and blocking of direct satellite path. This system adopts two techniques to cope with different types of shadowing and blocking.

The first one is a bit de-interleaver in the receiver for shadowing and blocking caused by small objects. These shadowing and blocking appear as solid burst noises in the received signal up to a second, for example. A solid burst noise is distributed over several second time period using this de-interleaver to fit error-correcting capabilities of this system.

The second one is the gap filler supposed by the system design, which re-transmits a satellite signal. These gap fillers will cover the area blocked by, for example, buildings and large constructions. The gap filler has two types in this system, so-called small-area gap filler and wide-area gap filler to cover the different types of blocked areas.

Small-area gap filler only amplifies a 2.6 GHz band broadcasting signal from the satellite. Small-area gap filler cannot have a high gain amplifier inherently to avoid undesired oscillation causes by signal coupling between transmitting and receiving antennas. This gap filler covers narrow area of direct path up to 500m from the equipment.

However, wide-area gap filler covers whole area within 3km diameter from the equipment and should use high gain amplifiers. To avoid undesired oscillation, this gap filler receives a feeder signal using different frequency from 2.6 GHz, for example, 14/11 GHz bands via satellite.

In these circumstances, multi-path-fading is appeared in the area where more than two broadcasting signals are received. In this broadcasting system, we adopt CDM (Code Division Multiplex) to realise a stable reception of multi-path-fading signal. By using RAKE technique and antenna diversity in receiver, large improvement of receiver's performance is expected even in the multi-path-fading environment.

In CDM system, different broadcaster will use a different orthogonal code for spreading signal in order to broadcast its own program independently. Power Flux Density (PFD) is relatively low because CDM signal is spreaded over wide frequency band.

3 Broadcasting System

Figure 2 shows the basic block diagram of the broadcasting system and Figure 3 shows detailed block diagram of CDM part of Figure 2. In the followings, we provide the basic parameters and capabilities of channel coding and modulation scheme of this broadcasting system.

3.1 Frequency

Centre frequency is 2,642.5 MHz.

3.2 Bandwidth

Bandwidth is 25 MHz.

3.3 Polarisation

Polarisation is left-hand-circular□polarisation.

3.4 Modulation

CDM scheme is adopted for modulation both of satellite link and terrestrial gap filler link. As shown in Figure 3, one data sequence is converted from serial bit stream to I and Q data sequences at first. After that, each I and Q data are spreaded by the same unique Walsh code (#n) and pseudo random sequence These spreaded data are modulated into QPSK signal. Modulated signals, each signal is identified by Walsh code, are multiplexed each other in the same frequency band.

3.5 Chip Rate

Chip rate is 16.384 MHz and processing gain is 64.

3.6 Signature Sequence and Spreading Sequence

Walsh codes of 64-bit length and pseudo random sequences of 2048-bit length are adopted as the signature sequence and the spreading sequence respectively. This spreading sequence is obtained by truncating Maximum Length Sequence of 4095-bit length generated using 12-stage feedback shift register sequence.

3.7 Data Spreading

Signature sequences and spreading sequences are modulo-2 added to the original I and Q sequence as shown in Figure 3.

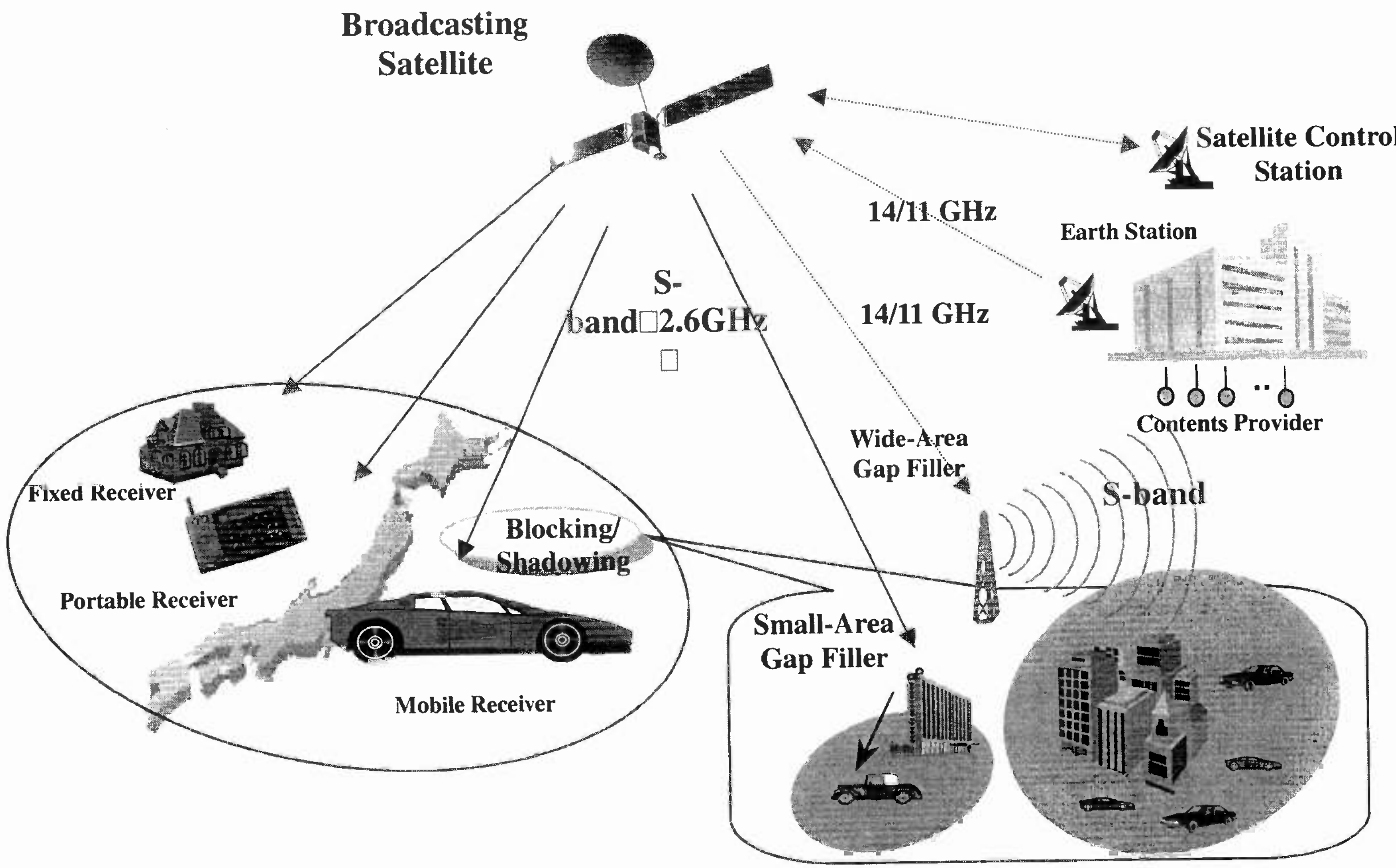


Figure 1. System Diagram

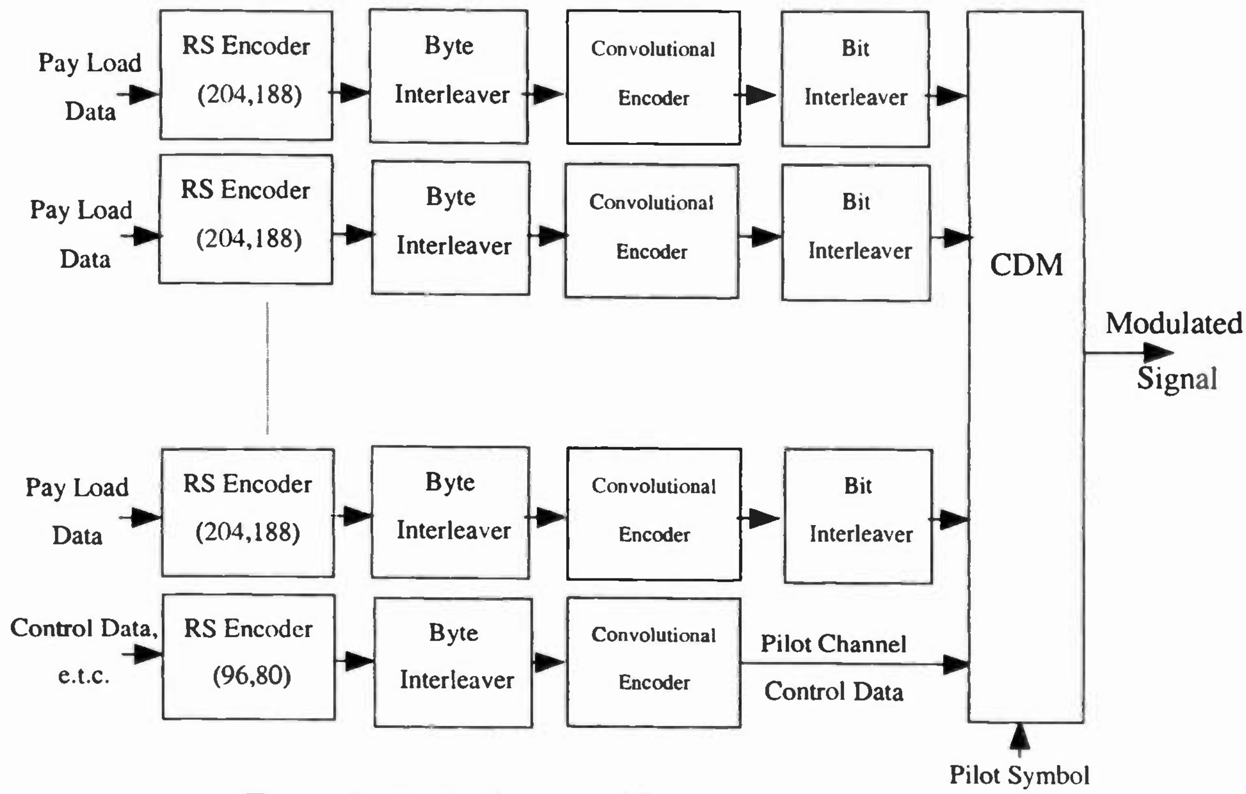


Figure 2. Block Diagram of Broadcasting System

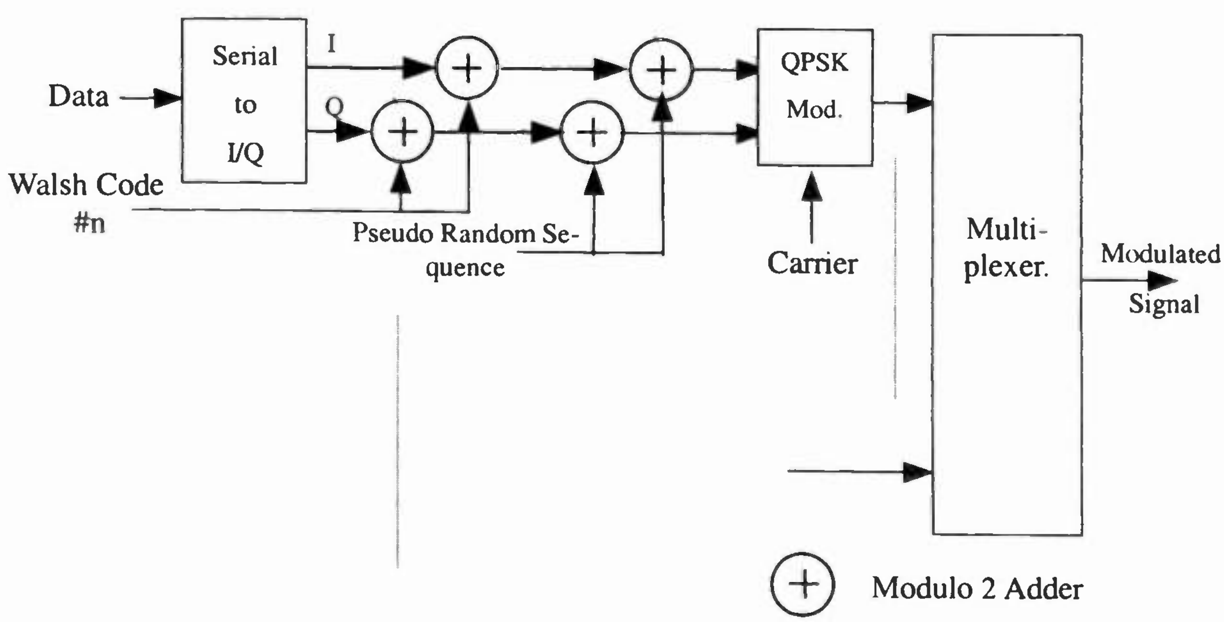


Figure 3. Detailed Block Diagram of Code Division Multiplex

3.8 Signal Filtering

The transmitted signal is filtered by square-root raised cosine filter. The roll-off factor is 0.22.

3.9 Error Correction Coding

Concatenated code with convolutional code for inner code and shortened Reed-Solomon (204,188) code for outer code is adopted for error protection scheme.

3.10 Interleaving

Byte-wise convolutional interleaving is used between outer coding and inner coding. Furthermore, bit-wise modified convolutional interleaving more than 3 seconds at most is adopted after inner coding.

3.11 Pilot Channel

Payload data are transmitted through broadcasting channels, while this system adopts pilot channel to simplify receiver's synchronisation and to transmit system control data.

Pilot channel has two functions. The first one is to transmit the unique code word for frame synchronisation and another control data to facilitate the receiver functions. The second one is shown in the following.

3.12 Pilot Symbol

Special data of this system are pilot symbols that are composed of 32-bit length continuing run of data '1'. Using these pilot symbols, receiver can analyse received signal profiles and these results assist RAKE receiver function. Pilot symbols are transmitted every 250 microseconds.

4 Satellite

In this system, geo-stationary satellite with large transmitting antenna will be used. Up-link signal is fed from an earth station in 14/11 GHz bands while service-link (down-link) to Japanese terrain in S-band. Major characteristics of the satellite are shown in the following.

(1) Up-link signal frequency:
14/11 GHz bands

(2) Service-link frequency:
2,642.5 MHz

(3) Service-link bandwidth:
25 MHz

(4) EIRP:
more than 67 dBW(Within service area, including antenna-pointing losses)

(5) Service area:
See Figure 4.

5. Receiver

Performances of typical mobile receiver of this system are shown in the following and Figure 5 depicts the block diagram of typical mobile receiver.

(1) Centre Frequency:
2,642.5 MHz

(2) Input signal bandwidth:
25 MHz

(3) G/T:
more than -21.8 dB(K^{-1})
Antenna Gain:
more than 2.5 dB
Noise Figure:
less than 1.5 dB

(4) Demodulation:
Pilot-aided coherent demodulation and RAKE receiver.

(5) Diversity:
Antenna Diversity

(6) Receiving Filter:
Square-root raised cosine filter
(Roll-off factor is 22%)

(7) Decoding of Convolutional Code:
Soft-decision Viterbi decoding

(8) Implementation Losses:
Less than 2 dB
(degradation from the theoretical value at bit error rate of 10^{-4})

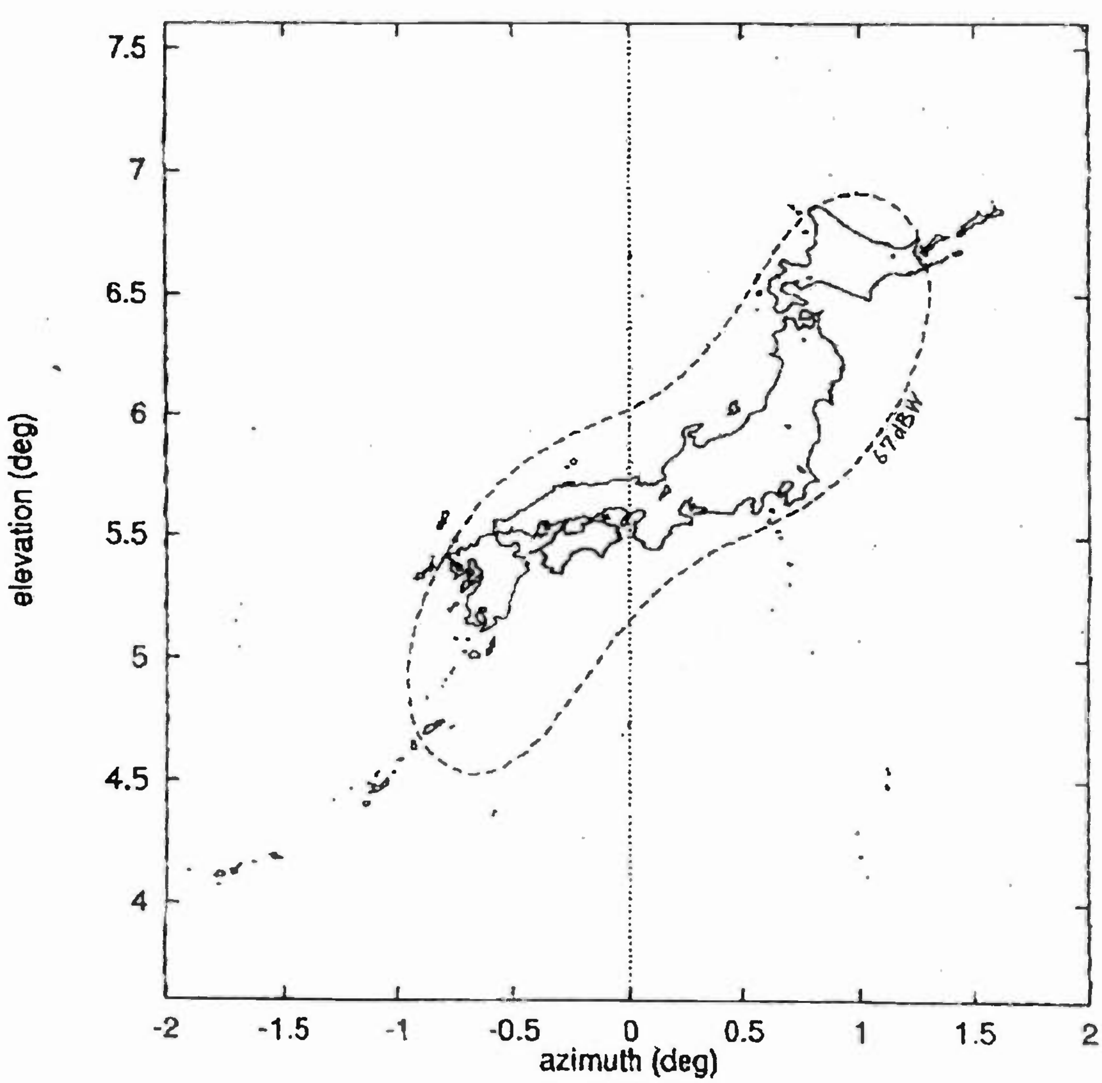


FIGURE 4 Service Area

6 Earth Station

The following shows the major characteristics of the earth station. Main function of the earth station is to transmit a broadcasting signal to the broadcasting satellite in the geo-stationary orbit.

(1) Up-link frequency band:
14/11 GHz bands

(2) Frequency bandwidth:
25 MHz

(3) EIRP:
More than 75 dBW

7 Small-Area Gap Filler

Main function of small-area gap filler is to receive the broadcasting signal directly from broadcasting satellite, to amplify, and to repeat it to the blocked area.

(1) Receiving frequency:
2,630 – 2,655 MHz

(2) Transmitting frequency:
2,630 – 2,655 MHz

(3) EIRP:
1.7 dBm

(4) Coverage area:
Direct path up to 500 m from the equipment

8 Wide-Area Gap Filler

This equipment receives 14/11 GHz bands feeder signal from the satellite, converts the signal into S-band, amplifies it up to desired level, and transmits it to the blocked area. The following is the major characteristics of the equipment.

(1) Receiving Frequency:14/11 GHz bands
Transmitting requency:2,630 – 2,655 MHz

(2) EIRP: 60.7 dBm

(3) Coverage:
Within 3-km diameter range

9 Typical Link Budget

Table 1 to 3 shows the examples of link budgets of a direct satellite link, a small-area gap filler link and a wide-area gap filler link. In these link budgets, we suppose that satellite's transmitting power is 1.2 kW and a receiver antenna has omni-directional pattern in the horizontal plane. Using these parameters, we can transmit up to 40 channels, where each channel has data rate of 236-kbit/sec, or about 9.4 Mbit/sec in total, for the direct satellite path and small-area gap filler link. Currently we are examining the total transmitting capabilities for wide-area gap filler link.

10 Conclusion

From the last September, ARIB is conducting corroborative testing, that will be followed by formal approval of this system as the BSS (Sound) system in Japan.

2.6 GHz broadcasting satellite services are expected to be major real-time and non-real-time digital media to mobile receivers. In Japan, more than 70 million automobiles are in use and this system will provide high-speed data link and downloading channel to these receivers including co-operative work with ITS system.

TABLE 1

Example of Link Budget (Satellite Link)

Centre Frequency	MHz	2642.5
EIRP	dB	67.0
The Number of Channel	ch	40
EIRP/ch	dB	51.0
Latitude of the Earth Station	deg	35.0 N
Longitude of the Earth Station	deg	140.0 E
Longitude of the Satellite	deg	154.0 E
Distance	km	37367.5
Free Space Propagation Losses	dB	192.3
Polarisation Losses	dB	0.5
Rain Attenuation	dB	0.0
Atmospheric Losses	dB	0.0
Total Propagation Losses	dB	192.8
PFD	dB(W/m^2 per 4 kHz)	-131.6
Antenna Gain	dBi	2.5
Receiver Input Signal Power	dBW	-139.4
LNA Noise Figure	dB	1.5
Antenna Noise Temperature	K	150.0
System Noise Temperature	K	269.6
System Noise Power Density	dB(W/Hz)	-204.3
G/T	dB(K^{-1})	-21.8
Receiver C/No	dBHz	64.9
Degradation due to Up-Link	dB	0.1
Adjacent Channel System Interference Allowance	dB	0.2
Total C/No	dBHz	64.6
Data Rate / Channel	kbit/s	235.9
Required Eb/No	dB	4.0
Implementation Losses	dB	2.0
Required C/No	dBHz	59.7
Link Margin	dB	4.9

TABLE 2

Example of Link Budget (Small-Area Gap Filler Link)

Centre Frequency	MHz	2642.5
EIRP	dBW	-28.3
The Number of Channel	ch	40
EIRP/ch	dBW	-44.3
Free Space Propagation Losses	dB	95.0
Distance	m	500.0
Antenna Gain	dBi	0.0
Receiver Input Signal Power	dBW	-139.3
LNA Noise Figure	dB	1.5
Antenna Noise Temperature	K	150.0
System Noise Temperature	K	269.6
System Noise Power Density	dB(W/Hz)	-204.3
G/T	dB(K^{-1})	-24.3
Receiver C/No	dBHz	65.0
Degradation due to Up-Link	dB	0.1
Adjacent Channel System Interference Allowance	dB	0.2
Total C/No	dBHz	64.7
Data Rate/ch	kbit/s	235.9
Required Eb/No	dB	4.0
Implementation Losses	dB	2.0
Required C/No	dBHz	59.7
Link Margin	dB	4.9

TABLE 3

Example of Link Budget (Wide-Area Gap Filler Link)

Centre Frequency	MHz	2642.5
EIRP	dBW	30.7
The Number of Channel	ch	40.0
EIRP/ch	dBW	14.7
Propagation Losses in urban area	dB	154.0
Transmission Antenna Height	m	30.0
Distance	km	3.0
Antenna Gain	dBi	0.0
Receiver Input Signal Power	dBw	-139.3
LNA Noise Figure	dB	1.5
Antenna Noise Temperature	K	150.0
System Noise Temperature	K	269.6
System Noise Power Density	dB(W/Hz)	-204.3
G/T	dB/K	-24.3
Receiver C/No	dBHz	65.0
Degradation due to Up-Link	dB	0.1
Adjacent Channel System Interference Allowance	dB	0.2
Total C/No	dBHz	64.7
Data Rate/ch	kbit/s	235.9
Required Eb/No	dB	3.2
Implementation Losses	dB	2.0
Required C/No	dBHz	58.9
Link Margin	dB	5.7

Draft Standard for Digital Terrestrial Sound Broadcasting in Japan

Masayuki Takada*, Makoto Sasaki*, Tamotu Ikeda**
Association of Radio Industries and Businesses (ARIB)
Tokyo, Japan

Abstract
Narrow-band ISDB-T system (NISDB-T) was determined as the draft standard for terrestrial digital sound broadcasting in Japan in November 1998. The transmission scheme of NISDB-T is the same as that of Wide-band ISDB-T for terrestrial digital television. The system can provide two kinds of bandwidth, 429kHz of one OFDM-segment and 1.3MHz of three segments. The system specification is presented along with a prototype receiver used in the preliminary experiment.

1. Introduction

Digital broadcasting is expected to begin in Japan in 2000. The draft standard for digital terrestrial sound broadcasting was approved by the Japanese Telecommunication and Technology Council (TTC) in November 1998. The system is a family of Terrestrial Integrated Services Digital Broadcasting (IDSB-T) and is temporarily called Narrow-band ISDB-T (NISDB-T). Field trials for system confirmation will begin in April 1999, and the technical report on the trial results will be submitted to the TTC this summer.

In Japan, digital terrestrial sound broadcasting will be introduced, as a new media in contrast to the introduction of digital terrestrial television broadcasting because it is will be a complete change from the current analog system. Broadcasts will continue to be made using analog AM and FM signals. Digital terrestrial sound broadcasting will use the VHF band currently used for television broadcasting. Initially, Channel 7 will be assigned for NISDB-T.

The baseline of the NISDB-T transmission scheme is the BST-OFDM of ISDB-T reported in the ITU-R recommendation for digital terrestrial television. The transmission bandwidth is one segment of 429 kHz or three segments totaling about 1.3 MHz for the extended scheme.

Experimental NISDB-T equipment has been tested in preliminary field trials. Because the transmission scheme is the same as that of Wide-band ISDB-T, various advantages should be realized from the viewpoints of receiver manufacturing and frequency usage.

2. Outline of DSB system in Japan

2.1 System features

The future broadcasting infrastructure must accommodate digital multi-SDTV, HDTV, and digital sound services, along with multiplexed services via satellite and terrestrial networks, in addition to traditional AM and FM broadcasting. Considering these requirements, the draft standard for NISDB-T should specify a system with the following features:

* NHK (Japan Broadcasting Corporation)
**Sony Corporation

(a) high-quality sound broadcasting,
(b) multimedia broadcasting,
(c) stable reception during mobile operation,
(d) base-band signal elements in common with digital BS and WISDB-T,
(e) an OFDM-segment structure for transmission that is the same as that of the WISDB-T system so that the segment-receivers can be used in common,
(f) the ability to effectively use frequency resources (single-frequency networks (SFNs), etc.), and
(g) the worldwide standards for the base-band system.

bol duration than that of the digital single-carrier scheme. Multipath interference is suppressed by inserting a guard interval in the time domain, enabling the system to operate in SFN.

A BST-OFDM channel consists of sets of frequency blocks called OFDM-segments, which have a common carrier usage structure. All of the segments have a bandwidth of 6/14 MHz to facilitate frequency usage and channel tuning. Each segment can accommodate independently several combinations of carrier-modulation schemes (DQPSK, QPSK, 16QAM, 64QAM) and coding rates of the inner code (1/2, 2/3, 3/4, 5/6, 7/8).

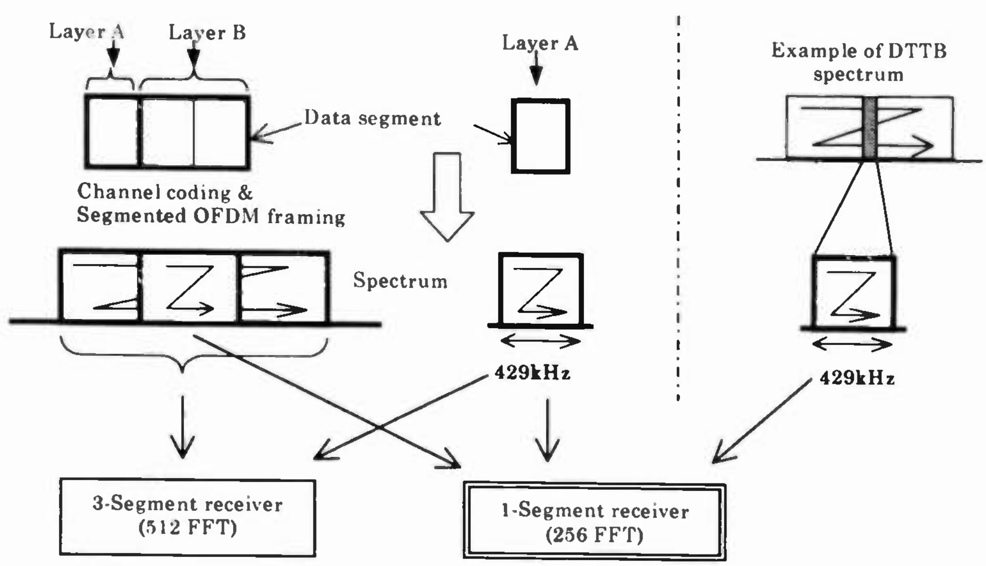


Fig.1 NISDB-T transmission and partial reception

2.2 Transmission scheme

(1) Overview

The NISDB-T transmission scheme is based on the band segmented transmission-OFDM (BST-OFDM), which is the same as that of a WISDB-T system without transmission bandwidth.

The OFDM scheme densely multiplexes many carriers while keeping them orthogonal to each other in a given transmission band. It uses a much longer sym-

The NISDB-T transmission parameters are shown in Table 1, and the concepts of NISDB-T transmission and reception are shown in Fig.1, with an example for Mode 1. As shown in the figure, two transmission bandwidths are used: 429 kHz for one-segment and 1.3 MHz for three-segment transmission.

In the case of three-segment transmission, two different transmission parameters can be set up if needed.

Table 1 Transmission Parameters of ISDB-T

Mode		Mode 1	Mode 2	Mode 3
Number of Segments		Ns = 1(One-segment transmission) Ns = 3 (Three-segment transmission)		
Bandwidth		3000/7(kHz)×N_s+250/63(kHz) 432.5···kHz (One-seg.) 1.289···MHz(Three-seg.)	3000/7(kHz)×N_s+125/63(kHz) 430.5···kHz (One-seg.) 1.287···MHz(Three-seg.)	3000/7(kHz)×N_s+125/126(kHz) 429.5···kHz (One-seg.) 1.286···MHz(Three-seg.)
DQPSK segments		n_d=1or3		
Coherent modulation segments		n_s (n_s+n_d=N_s)		
Carrier Spacing		250/63 = 3.968···kHz	125/63 = 1.984···kHz	125/126 = 0.992···kHz
Number of Carriers	Total	108×N_s + 1 = 1405	216×N_s + 1 = 2809	432×N_s + 1 = 5617
	Data	96×N_s = 1248	192×N_s = 2496	384×N_s = 4992
	SP	9×n_s	18×n_s	36×n_s
	CP	n_d + 1	n_d + 1	n_d + 1
	TMCC	n_s + 5×n_d	2×n_s + 10×n_d	4×n_s + 20×n_d
	AC1	2×N_s= 26	4×N_s= 52	8×N_s= 104
	AC2	4×n_d	9×n_d	19×n_d
Carrier Modulation		QPSK, 16QAM, 64QAM, DQPSK		
Number of Symbols per Frame		204		
Effective symbol duration		252 μs	504 μs	1.008 ms
Guard Interval (GI)		63 μs (1/4) 31.5 μs (1/8) 15.75 μs (1/16) 7.875 μs (1/32)	126 μs (1/4) 63 μs (1/8) 31.5 μs (1/16) 15.75 μs (1/32)	252 μs (1/4) 126 μs (1/8) 63 μs (1/16) 31.5 μs (1/32)
Frame Duration		64.26 ms (1/4) 57.834 ms (1/8) 54.621 ms (1/16) 53.0145 ms (1/32)	128.52 ms (1/4) 115.668 ms (1/8) 109.242 ms (1/16) 106.029 ms (1/32)	257.04 ms (1/4) 231.336 ms (1/8) 218.464 ms (1/16) 212.058 ms (1/32)
Inner Code		Convolutional code (1/2, 2/3, 3/4, 5/6, 7/8)		
Outer Code		RS (204,188)		
Inter leaving		Time and Frequency		
Information Bit Rates (TS rates)		One-seg.: 280.85kbps (DQPSK, Inner code = 1/2, GI = 1/4) — 1.7873Mbps (64QAM, Inner code = 7/8, GI = 1/32) Three-seg.: 0.842Mbps (DQPSK, Inner code = 1/2, GI = 1/4) — 5.361Mbps (64QAM, Inner code = 7/8, GI = 1/32)		

However, the frequency interleaving of the center OFDM-segment must remain within itself to maintain compatibility with one-segment receivers. This constraint applies to the center segment of Wide-band ISDB-T when that segment is dedicated to partial reception.

A functional block diagram of the NISDB-T transmission scheme is shown in Fig. 2

transmission parameters. Null TSPs are used only for signal processing on a multiplex frame and are not transmitted.

The algorithm for re-multiplexing a TS is predetermined in such a way that the receiver can easily regenerate the same TS.

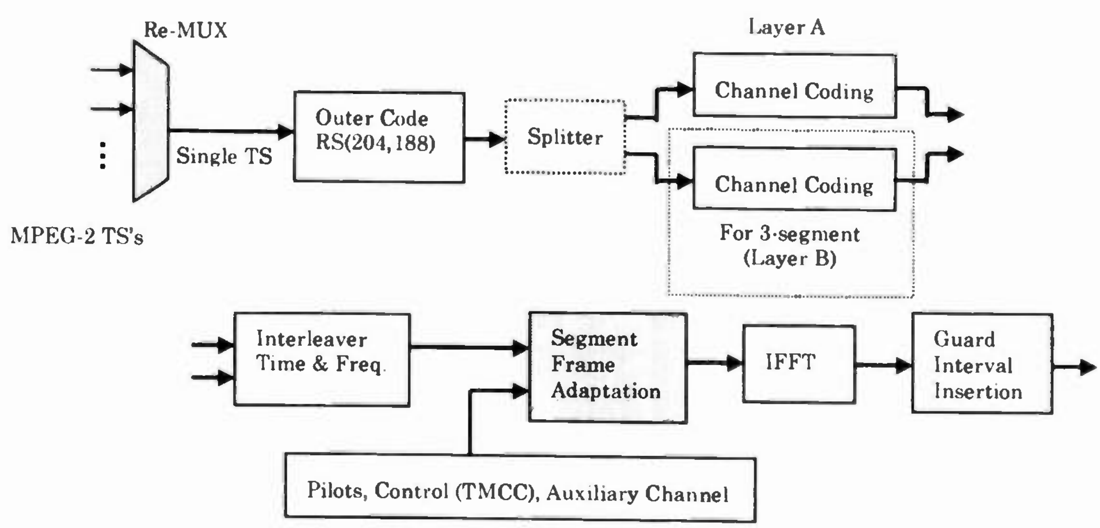


Fig.2 Functional block diagram of the transmission system

(2) Re-multiplexing

To achieve hierarchical transmission using the BST-OFDM scheme, the system defines a multiplex frame for single Transport Stream (TS) transmission within the scope of MPEG-2. In this multiplex frame, several input TSs are converted into a single stream of 204-byte TSP composed of TSP and 16-byte null data.

The duration of the multiplex frame is adjusted to that of the OFDM frame by measuring a 204-byte Transport Stream Packet (TSP) sequence at a clock rate four times faster than that of FFT sampling. Inserting null TSPs compensates for transmission-rate differences between the layers caused by layer properties, i.e.,

(3) OFDM frame

For three-segment transmission, each transport stream is divided into two layers, in which each channel coding is performed. After interleaving along the time and frequency axes, the symbol data are formed into an OFDM frame together with the control signals.

An OFDM frame consists of a sequence of 204 symbols. The number of carriers within an OFDM segment depends on the transmission mode. There are two kinds of OFDM-segment frames: differential modulation (DQPSK) and coherent modulation (QPSK, 16-QAM, 64-QAM). They are shown in Fig.3 for Mode 1. Continual pilot (CP), auxiliary channel 1

(AC1), and auxiliary channel 2 (AC2) are mainly used for frequency synchronization. Scattered pilot (SP) is only in the coherent-modulation segment and is used for channel equalization.

audio channels are provided by 16-bit PCM with 32-, 44.1-, and 48-kHz sampling rates. Low-bit-rate services, such as for speech-level audio, are now being investigated. Data-coding schemes are under standardization process in the ARIB and the TTC.

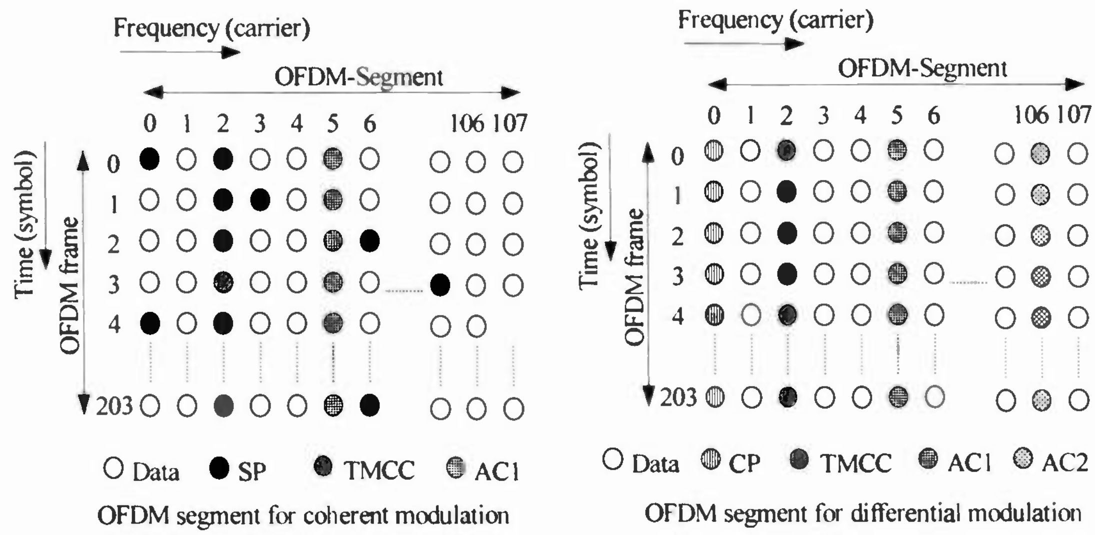


Fig. 3 Example of OFDM segment frame for Mode 1

(4) TMCC

As shown in Table 1, some carriers are used to control information. This control is called transmission and multiplexing configuration control (TMCC). It contains information on the carrier modulation scheme and the coding rate for each OFDM-segment and segment configuration. The TMCC symbols are also used for frame synchronization.

The TMCC information is encoded using shorted difference set cyclic code (184,273) and transmitted using DBPSK modulation.

3. Source-coding and mutiplexing

3.1 Source-coding scheme

The draft standard specifies MPEG-2 AAC audio (ISO/IEC13818-7) as the coding scheme because it generates high-quality audio signals. In addition to baseline service providing two audio channels with 144-kbps capability, several grades of service up to 5.1

3.2 Multiplexing scheme

The standard specifies the MPEG-2 system (ISO/IEC 13818-1) as the baseline scheme for multiplexing digital signals. A re-multiplexing scheme in which several input transports streams are combined into a single stream for hierarchical transmission is specified as part of the transmission scheme.

4. Implementation considerations

Given the very congested frequency spectrum in Japan, BST-OFDM is a promising scheme for optimizing spectrum usage as shown in Fig.4.

The NISDB-T channels of independent broadcasters can be transmitted together from the same transmitter without guard bands as long as the frequency and bit synchronization are kept the same between the channels. Most NISDB-T stations available in early stages will likely be constructed with a 1-seg. x N configuration. Because the guard band of WISDB-T corresponds to that of an OFDM segment, it can be also used for NISDB-T when the above condition is satisfied.

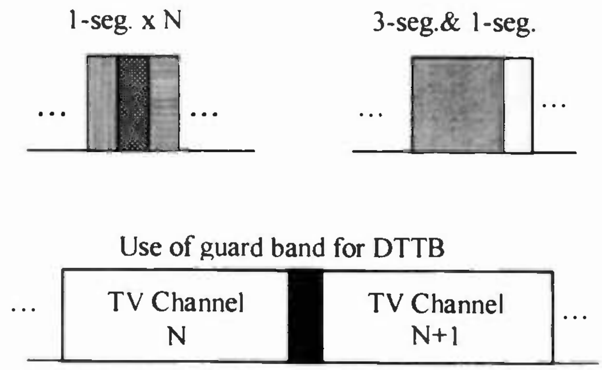


Fig.4 Transmission without guard band

5. Receiver

Narrow Band ISDB-T (NISDB-T) Basic Receiver block diagram is shown in Fig.5. Blocks inside hatched area are needed only for 3-segment receiver.

Firstly, 1 segment receiver decoding process is shown. NISDB-T signal received with Aerial Antenna

After A/D conversion of the IF signal, digital data is fed into Quadrature Demodulator followed by OFDM Demodulator using FFT with the Mode dependent number of sampling data. Referring signal both before and after FFT, synchronization recovery are executed, that is, carrier frequency, timing and FFT window recovery.

FFT output is fed to Frame Detector followed by TMCC Decoder, in which Constellation, Coding Rate and Interleave Pattern is to be extracted, and fed to Carrier Demodulator, in which coherent demodulation and equalization for each sub-carrier with scattered pilot is executed, or differential demodulation is executed.

Demodulated data is fed to Frequency De-interleaver followed by Time De-interleaver. In De-mapper, metrics to be used in Viterbi decoder are generated and fed to Bit De-interleaver followed by De-puncturing and Viterbi Decoder.

Viterbi decoded bit data sequence is converted to byte data sequence and fed to Byte De-interleaver followed by Energy Dispersal. Finally, after RS decoding, TS is

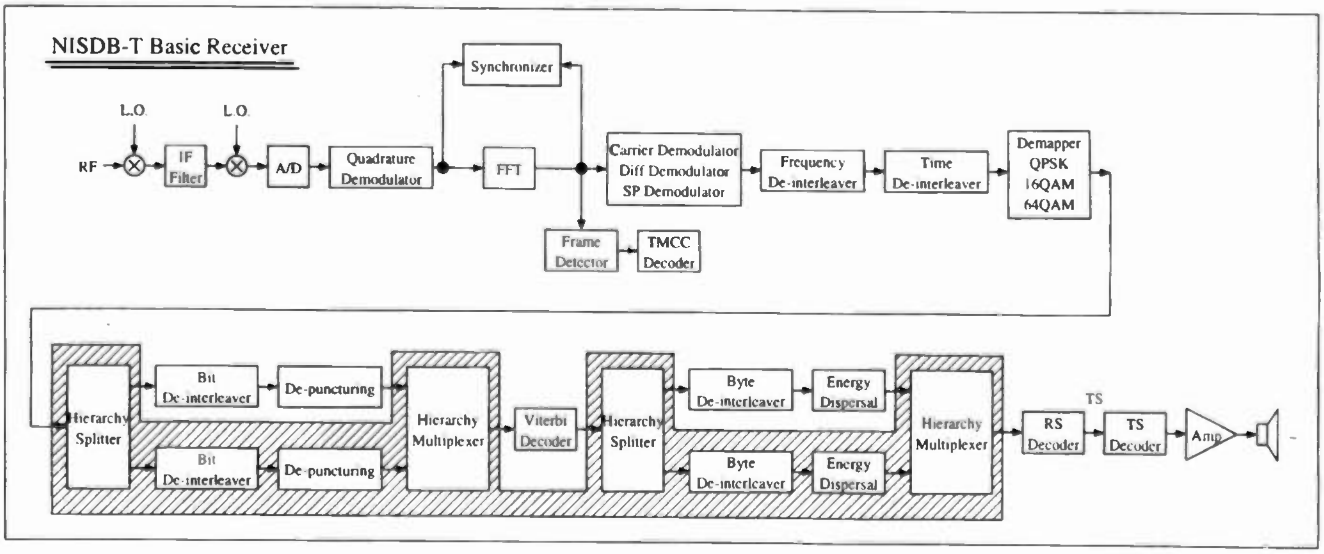


Fig.5 Receiver block diagram

to be fed to the MPEG-2 AAC decoder or to the dedicated data decoder.

Secondly, 3-segment receiver decoding process is shown. A receiver catches the center segment at first the same way a 1-segment receiver does. After decoding TMCC, IF filter of the receiver is changed to the 3-segment filter by the information for segment configuration.

In Frequency and Time De-interleavers, and De-mapper, each process is of segment independent. The output of De-mapper is fed to Hierarchy Splitter to split input data into center segment and two sided segments. For each hierarchy, bit de-interleaving and de-puncturing are executed, then two hierarchical data are multiplexed again to be decoded with only one Viterbi decoder. After Energy Dispersal, two hierarchical data are re-multiplexed again to form original TS. The prototype receiver is shown in Fig.6.

6. Conclusion

From now on, field trials are carried out for the system confirmation of NISDB-T using a transmitter set up on Tokyo tower. Accommodating as much common element as possible between other media, the system is very promising to provide multimedia broadcasting for mobile terminals. Towards the next century, it is important to construct appropriate digital broadcasting networks and to develop attractive data services in addition to hi-quality sound.

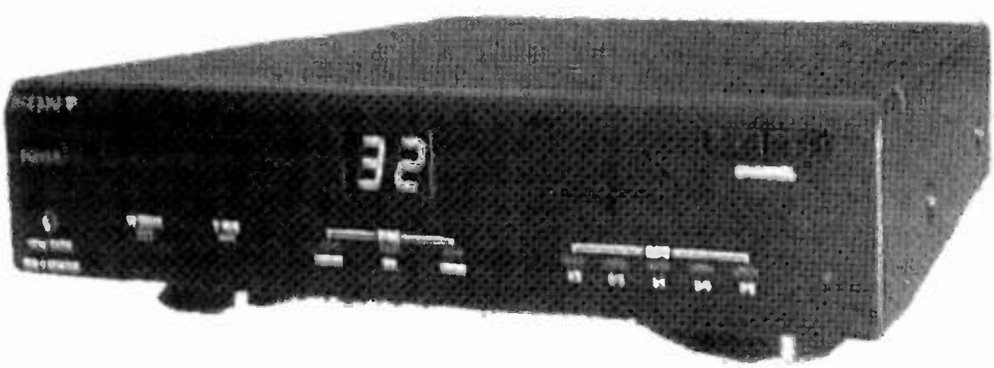


Fig.6 Prototype receiver

AUDIO PROCESSING FOR DAB AND THE INTERNET: HOW AUDIO QUALITY AND INTELLIGIBILITY CAN BE IMPROVED IN THE DATA REDUCED ENVIRONMENT, USING DYNAMICS CONTROL

Frank Foti
Cutting Edge
Cleveland, Ohio

ABSTRACT

Digital Audio Broadcasting (DAB) and the Internet are the latest broadcast media that exist today. Data bandwidth capacity dictates various degrees of quality that is possible. DAB has the potential to deliver 'near CD' quality, while the Internet provides quality ranging from 'good AM' to 'near FM like'. In each situation, some degree of audio data reduction must be applied. Use of dynamics processing can be employed as a tool to "predict" when audio conditions occur that can degrade signal quality. Utilizing a unique analysis algorithm that is modeled around the masking curve of a coding application, the processing adjusts the audio signal on a frequency dependent basis to reduce artifacts. This is especially noticeable whenever very low bitrate coding is applied.

AUDIO PROCESSING: THE "TOOL"

Is audio processing needed for the new transmission media like DAB, DTV, and the Internet? Audio purists will claim, that there is no need to create or clone the sound of FM on these new mediums. In a literal sense they are correct! But the reality is that we do, in fact, need audio processing as a *tool* so that these new mediums can be utilized with sonic efficiency and maximum intelligibility. Depending upon the method or system, one of the above mentions items will benefit from a processing *tool*.

When most people think of broadcast audio processing, they usually imagine the method used in FM and AM radio, where it's employed to create a 'dial presence.' In those instances it's a protection device for bandwidth control and to guard against overmodulation. In addition, it's a programming 'statement' that creates a sonic signature for the listener. For this presentation however, we will refer to processing in a different context. Here it will be discussed as a method that will improve the operation of the data compression required in these new mediums.

CODING IS PROCESSING TOO

When audio coding is discussed, it's usually thought of in terms of what is actually transpiring: the reduction of audio data, using an algorithm that's designed to operate within a specified medium or data bandwidth. When viewed at a bit more closely, it's also audio processing! Consider for a moment, in the following nutshell overview, what is occurring in this process:

- The audio signal is divided up using a filterbank.
- Analysis of the audio signal is applied to create a *masking curve*.
- The *masker* signal will be adjusted and moved over the spread of filterbank outputs to select the most dominant signal, and remove undesired spectrum.
- The remaining signal is then coded.

Each of these functions is an element of signal processing. In some instances, there are similarities between what transpires in a dynamics based audio processor and an audio coder. What this discussion will reveal is that dynamics based processing can operate in tandem with data reduction coding and thus create a transmission method that works together as a complete coupled system. The benefits of this coupling is better sounding audio through the coded system.

Quick Codec Review

A technique that is very popular is the use of the audio codec with transmission systems. These devices make use of "lossy" data reduction algorithms to compress the bitrate down to a size that will fit within the existing bandwidth of the system. While there are a number of specific algorithms to choose from, most have employed ISO/MPEG Layer-II, ISO/MPEG Layer-III, apt-x, and Dolby AC-2, and now AAC.

The basic operation of the "lossy" data reduction system stems from the use of a technique known as perceptual coding. Simply stated, the basic principle relies on a masking signal, or masker, that exists around a threshold curve which happens to follow that of the human auditory system. Any signal which falls below this threshold curve is basically discarded. In the digital domain, any audio data that would fall below the threshold curve is data that is then discarded by the algorithm, and thus data reduction is accomplished.

Detailed operation of the above mentioned algorithms is not needed for this discussion, as the intent will be to focus upon the effects that dynamics processing has upon data reduced audio. Suffice it to say that each system does possess many strengths and possible weaknesses for their application. It is not the intent of this discussion to compare audio coding algorithms.

Codec Transmissions, The Caveats

All audio coding methods have strengths and weaknesses. In almost all cases as bitrate is reduced, audio quality degrades due to less data bandwidth. However, it should be pointed out that some of the most recent demonstrations of AAC at lower bitrates is quite impressive!

Depending upon the transmission medium DAB, DTV, or Webcasting the audio quality of the codec will be determined by two issues: the coding algorithm and amount of data reduction. For DAB and DTV, the bitrates generally are at higher levels, usually 192kbps or greater for stereo. At these rates, the coding process affords wide audio bandwidth (20kHz) and contains a small amount of artifacts (near CD quality). But in the case of webcasting, where bitrates of 28kbps might be required, the coding artifacts and bandwidth restrictions are quite severe (AM radio-like). Audio bandwidth is sometimes reduced to 4kHz. At lower bitrates, coding artifacts differ considerably with each of the data reduction algorithms.

Given this wide range of diversity, how does processing fit in to all of this, especially given the view that transmission system processing for FM Stereo and codec STL systems do not mix that well[1]! This can be answered with two points:

• FM Stereo transmission processors employ hard limiters, or clippers, to achieve absolute peak control. It has been shown through testing and research that codecs do not perform well when a transmission processor is operated through the codec. The audio quality suffers from added distortion generated by the coding method used in the STL. Precise peak control is lost as overshoots are generated by the lossy data compression of the codec. This is the result of the preemphasis and clipper functions in the processor. Thus, in most applications of a coded STL and processor, the processor needs to be installed after the STL system to avoid the previous stated problems. In essence, the problems stem from two difficulties that occur as the harmonic content of the clipper becomes displaced by the coding algorithm and preemphasis (50µs or 75µs) hinders the masking curve of the codec from operating in an efficient manner. Coding algorithms were not designed to operate on *emphasized* audio, as this increases the audibility of coding artifacts. This is what contributes to lost peak control, and adds further Total Harmonic Distortion (THD) to the sonic quality of the signal.

• While the above comments do not bode well for processing and coding, it must be pointed out that this is true when the use of FM transmission processing is employed with a codec. As with that type of system, where processing is designed to fulfill the specific needs of the technology and augment its performance, the same thing must be done for coded systems too. Again, using the model of traditional broadcasting, we can apply the same thinking to coded transmission. As FM and AM require different processing methods, the same holds true for coded systems. Here the differences in processing will be determined by audio bandwidth, bitrate, and coding algorithms. Therefore the use for processing in this environment is dictated by the support requirements of the codec system and its attributes, just as it is in conventional analog broadcasting. What follows are some of the issues and features that a processor for coded transmission must recognize.

DIGITAL TRANSMISSION LANDSCAPE

Before specifying the attributes of a processing system for DAB, Internet, or DTV, the landscape for these mediums need to be defined and understood, as there are certain aspects that differ from the conventional analog methods. These deal with algorithm, bitrate, sampling rate, audio spectral bandwidth, digital full scale, and metadata. Processing for the digital mediums is highly dependent on how each of these issues are dealt with.

Algorithm

Definitely the most debatable item, and quite important, the choice of coding algorithm will play a significant role in the overall sonic performance of the transmission system. In most of the digital mediums, the algorithm of choice has already been made. Then the issue is to understand the usage of the specific algorithm in order to achieve best sonic performance. For this discussion, it is not necessary to delve into the aspects of each coding method. What's important is to know and understand those aspects of the algorithm for the medium used.

Bitrate

Almost as important as the choice of algorithm is the operating bitrate of the system. This will be determined by the available data bandwidth. This can range from as low as 24kbps for lower grade narrow range monophonic audio to greater than 384kbps for full range CD-like stereo or multichannel sound. As bitrate is reduced, the coding algorithm must operate more aggressively in reducing the amount of data pass-through. This will affect the available audio bandwidth and sonic quality of the audio. Generally speaking, as bitrate is reduced, both bandwidth and audio quality degrade. The amount of degradation will be different based upon the algorithm employed.

Spectrum

Each algorithm and bitrate will have a direct effect upon the available amount of audio bandwidth that the system can transport. As stated above, when bitrate is reduced, sampling rate and bandwidth follows suit. Therefor, it is crucial that any processing system be capable of managing the audio bandwidth, as this will have a direct effect on sonic performance at lower bitrates. More on this topic is covered later in this paper.

While on this topic, an item should be pointed out: preemphasis is not required in these systems. This alone will provide a major sonic improvement when compared to analog FM systems. The digital mediums all utilize a *flat* spectrum, and this negates the need for specialized high frequency control methods that operate around the emphasis networks. Later however, we will see how spectrum management through dynamics control will improve an encoder's efficiency.

Full Scale

All digital transmission systems have one important element in common. They have a specified maximum word size. In other words, they have a peak ceiling level that can not be exceeded! Overmodulation is not possible; exceeding *full scale (0dBfs)* is a nasty sounding type of distortion. Any signal processor for digital mediums must perform absolute peak control, yet eliminate the unwanted coding artifacts that would occur if a clipper was employed, as described earlier.

Metadata

Digital delivery systems usually provide some means of carrying ancillary information about the signal content, or attributes of the content. This is a function that metadata can provide. Already employed in the DTV standard, metadata has the capability to carry information about the dynamic content of the signal at the transmission point, and then apply that *knowledge* to corresponding functions in the receiver. In this manner, audio processing effects can be implemented in the system, and then each individual receiver can use this in whatever manner they choose. By example, the transmission may contain metadata information about how much dynamic level compression to employ, and then the end-user can allow that compression to be applied to the signal, or eliminated, and make use of the wider dynamic range.

The Dolby Digital system for DTV2 provides numerous functions that take advantage of metadata. These include a dialog normalization setting for loudness control, dynamic range compression, and emergency notification purposes. Some of the proposed DAB systems provide space for metadata.

From the above definitions, and functions it can be seen why a conventional FM or AM transmission processor is not applicable for the digital mediums. A simple, off-the-shelf, compressor/limiter is not the answer either, as these units only provide a generalized form of dynamics control, and are usually wideband in nature. Just as these units will not suffice in the conventional analog transmission applications, because they lack specified functions for the medium, the same analogy exists for the digital systems, too.

PROCESSING FOR DIGITAL DELIVERY

What is required of an audio processor for digital transmission are three specific items, and a possible fourth should be considered:

- Precise peak control so that 0dBfs is maintained without system distortion.
- A peak control method that does not exaggerate coding artifacts.
  - This also holds true for any dynamics control and equalization methods.
- Audio frequency response that is controlled within the bandwidth limits of the system.
- In low bitrate systems, processing should enhance intelligibility.

The following sections describe methods and means of accomplishing the above requirements. These are based upon a digital processing system that was designed specifically for digital transmission mediums. What's presented here is a global view of the implemented functions, as this provides an easier understanding of the concepts involved.

Absolute Peak Control

The simplest form of peak control is the hard limiter, or clipper, as it's quite often referred to. In the coded digital environment, the use of a clipper can cause three sonic hardships: harmonic distortion from the truncation action of the clipper, exaggerated coding artifacts of the data reduction algorithm, and clipper induced aliasing distortion, aka *digital grunge*, that results from clipper harmonics that try and exceed the Nyquist frequency of the system.[3] Therefore, a better method must be employed that controls peak levels with precision and does not generate any of the three aforementioned problems.

Look-Ahead Limiting

A *Look-Ahead, or Delay-Line* limiter is perfect for this operation. Basically, this limiter creates a gain control signal based upon the absolute peak value of the audio signal, except that while the peak level is being calculated, the audio signal is physically delayed by an amount of time, equal to the time needed to calculate the peak value. Once the control signal is ready to implement level adjustment, the audio is then sent ahead to the control element at the exact moment that the control word arrives to make the adjustment[4]. In this manner, absolute peaks can be controlled without the need to truncate the excursion, as a clipper would do. *Figure-1* is a block diagram of a look-ahead limiter..

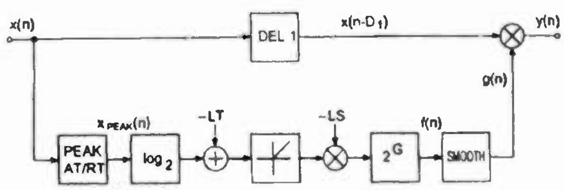


Figure-1

This results in little or no harmonic distortion generated by the limiter function. The caveat to this method is that instead of harmonic distortion, Intermodulation Distortion (IMD) can result, but that amount is dependent on the design of the look-ahead algorithm. In addition, there will be some amount of latency as the audio is delayed by a specified amount; it can be up to a few milliseconds. This could make it a bit problematic when trying to monitor oneself off-the-air, except all of the digital mediums generate some amount of latency. So the use of look-ahead limiting is not a problem here.

This is not a new concept, as this method has been utilized before in other applications. Generally, it has been employed in a wideband mode which can exaggerate IMD products. That's one of the main reasons why it was never popular in conventional broadcasting, along with the latency issue.

New research has revealed some fresh methods to implement look-ahead limiting so that IMD can be minimized, or suppressed completely. The sonic result is absolute peak control that yields a very high degree of fidelity when peak control is performed.

Prediction Analysis

In addition to look-ahead limiting, another new processing function that will aid the operation of the ensuing data encoder is a method known as *Prediction Analysis*. Its operation, much like the aforementioned limiter, will analyze signal information based upon peak level and frequency content as it relates to the coding process. The resulting analyzed information is either added or subtracted from the control signal of the final limiter based upon the prediction model that is designed around coding algorithms. The prediction model takes into consideration certain frequency and dynamics conditions that can agitate the encoder and generate codec artifacts. With Prediction Analysis, the limiter is able to allow the following encoder to operate more efficiently as it reduces coding artifacts.

Multiband Dynamics Control

When lower bitrates are used, intelligibility and overall quality is a problem. Some coding algorithms provide audio quality that has been described as sounding like a 'bad cassette' recording. Voice is muffled, music sounds thin and lifeless. Inserting a graphic equalizer is not the answer, as it will provide inconsistent adjustment on a source-to-source basis.

A multitband dynamics control section is the answer for these situations. It provides three key functions:

- It can be setup and adjusted for consistent source-to-source consistency in sound.
- The action of the frequency bands can be optimized to enhance voice intelligibility.
- The upper bands can assist the final limiter, and further improve coding efficiency.

The following block diagram, *Figure-2*, provides an overview of a processing system for digital transmissions.

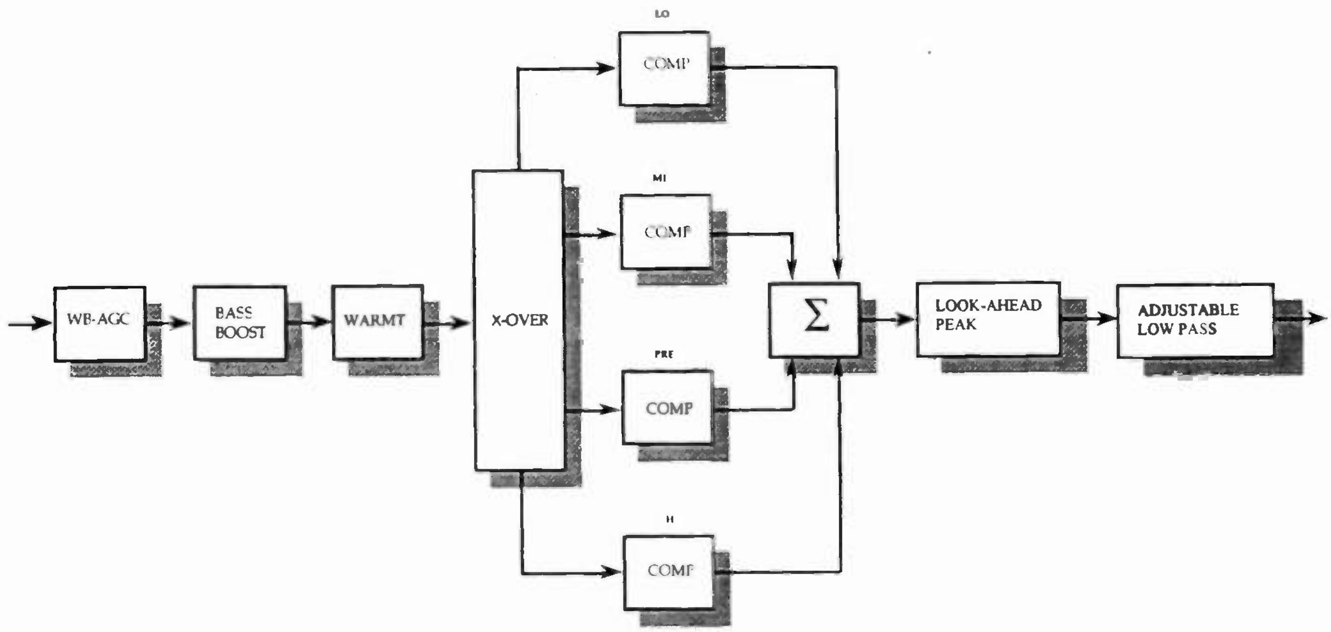


Figure-2

As with the case of conventional multiband processing, effects EQ can be inserted before the cross-over section. Here is where a gentle boost in the midrange or presence frequencies will assist in enhancing intelligibility in low bitrate systems. In addition, an adjustable cut-off low pass filter can be employed after the final limiter so that audio bandwidth control is provided. Due to the low harmonic content of the look-ahead limiter, the low pass filter will not generate any system overshoots. Using a low pass filter to remove any audio spectrum that will not be encoded further reduces the *shrillness* associated with low bitrate transmissions whenever too much high frequency content is presented to the encoder. It is desirable to know what the audio bandwidth limits of the coded system are, and then set the processor low pass filter accordingly.

It should be easily seen that each of the aforementioned items operates in a much different manner than the conventional FM/AM audio processor. Here is where the system is employed as a 'tool' instead of an effects box that's trying to create the threshold of pain on the dial! Audio processing for digital delivery will improve the overall performance of any system. Each coding algorithm has its own set of artifacts. Processing can be used to minimize, or eliminate those attributes. Additionally, in the new DTV system, it can also be used to write and implement metadata. Should that method cross over into DAB, it can be done there, as well. Here's a case where the use of signal processing is expanded from the conventional model that's employed today in analog transmission services.

LOUDNESS WARS, AND A FINAL THOUGHT...

Speaking of loudness, will DAB and Netcasters have loudness wars? Chances are some services will be concerned with competitive quality and density when compared to another. The issue will probably never end, as it just migrates onto other services. (Light humor intended!) At least loudness through overmodulation is not possible in the digital mediums! Although audio purists may scoff at the thought of a "Hot-Rockin' Flame-Throwin" digital signal, it should ease the mind that processing for DAB does not involve the extreme amounts of hard processing that's available in FM/AM broadcasting. The end result, even processed, will be superior sound, as the use of preemphasis and clipping will be eliminated.

The new digital transmission mediums create a plethora of opportunity for content providers. Therefore, the demands that these audio signals will put on the chosen coding methods will put them to the test. Hence, the employment of audio processing that suits the medium can only improve the overall end result. In this case, the analogy to the conventional processing system for FM/AM broadcasting holds true.

As DAB, DTV, and Netcasting continue growth in their respective paths, processing will find the need to reinvent itself as these new technologies break ground and flourish. The model of 'yesterday's ideals' in processing must be put to rest, as these new mediums offer a larger volume of opportunity, in both the content and technical domains. This demands that innovative research and design be performed today, as we finally bring to life these mediums for the new millennium.

REFERENCES

[1] Foti, F,: *Broadcast Signal Processing and Audio Coding: Are We Trying to Mix Oil with Water?*, 100th Convention of the Audio Engineering Society (AES), Copenhagen 1996, Preprint 4203 (I-8)

[2] Lyman, S.: *Distribution Options In The DTV Studio*, Dolby Laboratories, Digital Television 98 Conference, Chicago, December 1998

[3] Foti, F,: *Omnia.fm: An Engineering Study*, White Paper, Cutting Edge, Cleveland, September 1998

[4] Zolzer, U, : *Digital Audio Signal Processing*, John Wiley & Sons Ltd, Chichester, 1997

DAB: Migrating to a New World

Sunday, April 18, 1999

1:00 PM - 5:00 pm

Chairperson: Barry Thomas
KCMG-FM, Los Angeles, CA

1:00 pm General Performance of Amplitude Modulated PDM Transmitters for Digital Broadcasting Using the IBOC System
Wendell Lonergan
Nautel Maine, Inc.
Bangor, ME

***1:30 pm The Implementation of Digital Audio Broadcasting Using IBOC, A Complete System View**
Alan Pate
Lucent Digital Radio
Warren, NJ

***2:00 pm Interference Studies and Their Implications for AM and FM IBOC DAB**
Glynn Walden
USA Digital Radio
Columbia, MD

***2:30 pm The NRSC IBOC System Test Guidelines**
Andy Laird
Journal Broadcast Group, Inc.
Milwaukee, WI

3:00 pm Frequency Domain Reciprocal Modulation (FDRM) for Bandwidth-Efficient Data Transmission over Channels with Dynamic Multipath
Tom Williams
Holtzman, Inc.
Longmont, CO

***3:30 pm NRSC Evaluation of IBOC System Test Data**
Don Messer
US Information Agency/ Int'l Broadcasting Burr
Washington, DC

***4:00 pm Panel Discussion**
Carlos Aquirre, Grupo Radio Centro, Lomas Altas, Mexico; Rick Martinson, USA Digital Radio, Inc. Columbia, MD; Dwight Taylor, Digital Radio Express, Inc., New York, NY; Alan Pate, Lucent Digital Radio, Warren, NJ

*Papers not available at the time of publication

GENERAL PERFORMANCE of AMPLITUDE MODULATED PDM TRANSMITTERS for DIGITAL BROADCASTING using the IBOC SYSTEM

Wendell Lonergan

Nautel Maine Inc.
Bangor, Maine

ABSTRACT

This paper describes and discusses the results of tests carried out on two families of PDM (Pulse Duration Modulation) AM transmitters. A variety of tests have been done using two tone techniques to evaluate the transmitters for linearity. Testing was carried out both into a 50 ohm resistive load and into a bandpass load to simulate the effects of a limited bandwidth antenna as proposed by USA Digital Radio.

The paper discusses the various aspects of PDM transmitter design and how they contribute to its performance as a linear amplifier. The results are compared against current specification requirements for the IBOC DAB (In-Band-On-Channel Digital Audio Broadcast) system as proposed by USA Digital Radio.

1.0 INTRODUCTION

Digital radio for the US AM radio station is likely to become reality. All of the system proponents for the medium wave band in the US are proposing the addition of a digital program signal using In-Band-On-Channel (IBOC) for digital audio broadcasting. Systems provide promise of higher audio quality, more reliable reception, and additional information channels than are currently available from the existing analog AM broadcasting process. At least one of the proposed digital systems is based on Orthogonal Frequency Division Multiplexing (OFDM) technology and all will use various audio compression algorithms to encode and then deliver the digital audio to the receiver.

All of these digital broadcasting systems are likely to require similar transmitter performance. The equipment used to deliver IBOC DAB will have to meet the needs of both Consumers and Broadcasters in the terms of cost and performance and Government Regulatory Agencies to ensure a useful spectrum with minimum interference.

This paper describes the theory behind operating a PDM transmitter to generate high power signals efficiently. Just over a decade ago, such signals would have been generated using linear amplifiers. However, today's PDM transmitter provides performance which is capable of combining amplitude and phase information to meet the requirements of these digital systems while retaining high efficiency switching modes of operation for the RF amplitude stages.

Nautel first built transmitters using these actual techniques, a PDM transmitter with a Class D amplifier, eighteen years ago in 1981. This paper presents test results on Nautel's two types of PDM transmitter, the ND Series and the more recent XL Series. Test results are compared with the requirements for an AM band IBOC DAB system, as described by one of the IBOC proponents in recent technical publications (see **Reference 2-3**). Tests are also done to evaluate the impact of the restricted antenna bandwidth frequently encountered in AM installations.

Transmitter system considerations include those applicable during an initial hybrid phase when IBOC and a conventional AM transmitter signal are transmitted simultaneously and those for the final advent of an all digital world where the conventional AM signal would be phased out.

2.0 THEORETICAL BACKGROUND

A basic theoretical hypothesis is postulated as a foundation of the technique to be described in this report. This is:

A periodic waveform conveying information by means of a carrier and sideband (or sidebands) within an encompassing bandwidth of less than an octave, may be fully described as the product of an envelope (amplitude) term and a constant-amplitude, phase-modulated carrier term. Such a waveform requires only a definition of its peak envelope and voltage cross-over, as functions of time, in order to facilitate its synthesis.

The following actual waveform examples illustrate this:

2.1 Double Sideband AM with Carrier

$$E(t) = (1 + m \cos at) \sin bt$$

Where **a** and **b** are the modulation and carrier angular frequencies respectively (radians per second) and **m** is the modulation fraction.

This describes modulation with a single tone and the expression may be easily manipulated, if desired, to show the carrier and pair of sideband terms.

The envelope term in this case is just the **(1 + m cos at)** term which is always real and positive for values of m between 0 and 1.0. There is no phase modulation of the carrier term **(sin bt)** in this double sideband AM case.

2.2 Double Sideband without Carrier (i.e. suppressed carrier)

$$E(t) = (m \cos at) \sin bt$$

The envelope term in this case is

$m \cos at$

which is conveniently expressed as

$(m^2 \cos^2 at)^{1/2}$

which can be reduced to

$\frac{m}{2} (1 + \cos 2at)^{1/2}$.

This is the familiar expression for the output of a full-wave rectifier circuit and, as with the full wave rectifier, the envelope frequency has doubled.

At first glance there seems to be no phase modulation of the carrier term but the envelope term has ignored the sign (plus or minus) of the original cosine modulation and this information must be transferred as phase-reversal modulation of the carrier term to preserve identity with the original expression. Hence:

$$E(t) = \frac{m}{2} (1 + \cos 2\,at)^{1/2} \times F(t) \sin bt$$

Where $F(t) = 1$ **for** $\cos at > 0$
And $\quad = -1$ **for** $\cos at < 0$

2.3 Single Sideband with Carrier

$$E(t) = \sin bt + m \sin (a + b)\,t.$$

This is a simple spectral equation showing the addition of the carrier and sideband terms on two RF tones. It could be expanded for additional tones but two are analyzed for simplicity. This may be manipulated to a more suitable expression for the purpose of analysis as follows:

Let
$c = b + a/2$ **radians / sec**
(i.e. c = centre frequency)

Then

$$E(t) = \sin (c - \frac{a}{2})\,t + m \sin (c + \frac{a}{2})\,t$$

$$= \sin ct\,(1+m) \cos \frac{a}{2} t + \cos ct\,(m-1) \sin \frac{a}{2} t$$

$$= r \sin (ct - \emptyset)$$

Where

$$r = \text{envelope} = (1 + m^2 + 2\,m \cos at)^{1/2}$$

And

$$\emptyset = \text{phase} = \tan^{-1} (\frac{1-m}{1+m} \cdot \tan \frac{at}{2})$$

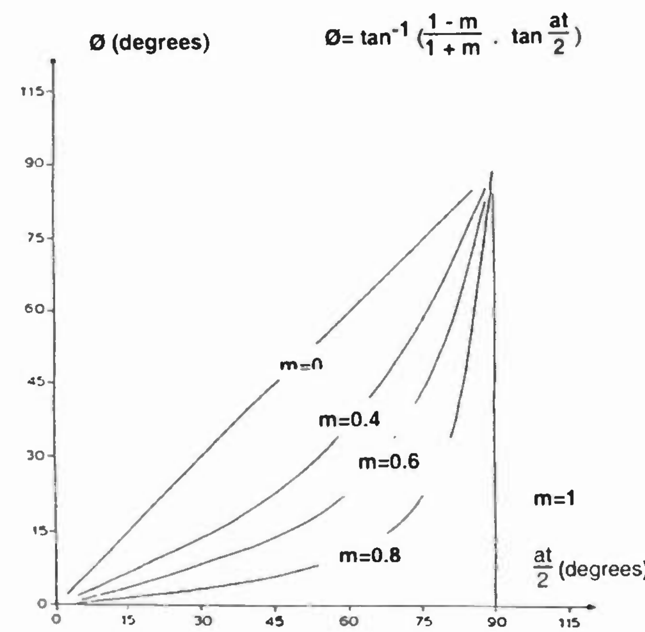


Figure 1

Plot of Phase versus (at) for Various Values of m

Figure 1 show plots of **Ø** versus **(at)** for various values of **m**. Note that when **m = 1**, simple phase reversal modulation occurs corresponding to the double-sideband, zero-carrier case.

3. AN IDEAL PDM TRANSMITTER

The basics of a Pulse Duration Modulation (PDM) transmitter are shown in **Figure 2**

The incoming audio signal is converted to a PDM signal which is used to drive a high power switching modulator. This high power switching modulator output is then filtered to remove the unwanted PDM frequencies and suppress them sufficiently to meet FCC requirement -80 dB at ±75 kHz. The output of this low pass filter is a

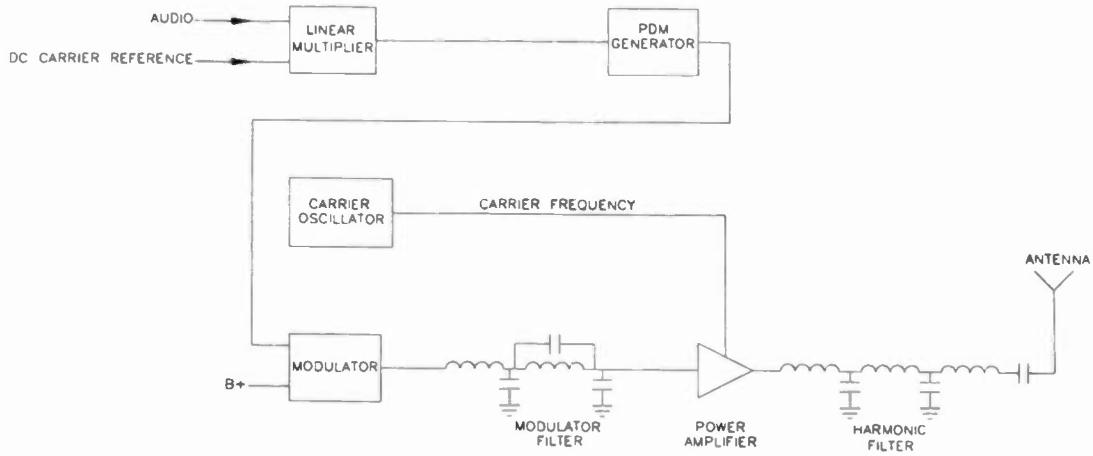


Figure 2
Pulse Duration Modulation (PDM) Transmitter
Basic Layout

combination of DC and AC which is used to modulate power amplifier stages, consisting of some form of switching amplifier, which is in turn driven with an RF drive signal at the carrier frequency. The output of the power amplifier consists of amplitude modulated and any phase modulated RF at the carrier frequency of the RF drive.

Some form of harmonic filter must be used to remove the RF harmonics, and comply with FCC requirements, prior to applying the signal, usually at a 50 ohm impedance level, into the antenna system.

It has been shown in the theoretical discussion that a signal consisting of tones could be generated as a product function of an appropriate envelope signal, **rt** and a phase-modulated carrier signal, ***(Ø)***. In practice, the envelope function would be applied as DC plus audio modulation in a PDM system and the phase function, ***(Ø)***, as the RF switching drive to the power amplifier.

The envelope function is subjected to a time delay when passing through the low pass modulator filter. Therefore, it is necessary to provide an equalizing delay in the phase channel. In a perfect system, no amplitude distortions occur and, therefore, time delays in both *amplitude* **(r)** and *phase* ***(Ø)*** channel match exactly and there is no intermodulation or noise.

4.0 PDM TRANSMITTER PERFORMANCE

Guidelines to the typical transmitter performance required by a digital system are given in **Reference 3.**

4.1 Transmitter Power

The amount of additional power required to provide the IBOC digital signal is approximately 13 dB below the carrier power in the hybrid system. Theoretically, the peak power requirement of the transmitter increases significantly with several RF tones present compared to that required in the single-tone case. It is suggested **(see Reference 3)** that in practice the peak power requirement is not much different than that required for the AM signal alone and that IBOC proponent testing has indicated that if a capability of 125% positive peak modulation index was used, then peak power corresponding to 140% modulation index is sufficient to accommodate the effect of the IBOC RF tones without introducing any significant digital error rates.

Obviously if the transmitter has insufficient peak power rating and the desire is to transmit the hybrid IBOC signal, then the alternatives are to either reduce the amount of analog amplitude modulation or reduce the carrier power or reduce both. The limiting factor in the transmitter will be its peak envelope power capability. This is particularly important in the case of solid state transmitters since the limitation on the peak envelope power capability will have been defined in the initial design of the transmitter by the number of devices and/or amplifier modules used in the transmitter.

Tests have been done on an AMPFET ND1 transmitter and an XL12. The AMPFET ND1 is an older PDM design and it is capable of 1.1 kW of carrier power with up to 125% positive peak modulation. This translates into a peak power capability of 5.5 kW. If the AMPFET ND1 was to be capable of transmitting 140% positive mod at 1.1 kW then it would need a peak power capability of 6.3 kW. Alternatively, it could be used at a carrier power of 950 W and 140% mod to give a peak power of about 5.5 kW.

The XL series of transmitter was deliberately designed with more headroom to allow for possible future needs of the customer if these new digital modulation techniques were introduced. Because of this the XL12 has a peak envelope power capability of 69 kW, 140% modulation at 11 kW would correspond to a peak power rating of 63 kW.

The all-digital IBOC system will not require transmission of a carrier. Therefore, although the peak to average power ratio of the digital waveform may be 2 or 3 dB higher than a hybrid IBOC waveform, it is starting from a lower average power point. This could result in less expensive products if transmitters are developed purely for a digital IBOC signal. Currently the research has not been done to define the necessary peak power versus digital system error. Transmitters will not need to accommodate the theoretical maximum peak that would occur corresponding to the addition of all the RF tones.

4.2 Transmitter Linearity

A very important factor in digital transmission is the transmitter noise floor level. For the purpose of this paper, the noise floor refers to noise from all sources including intermodulation distortion (IMD), thermal noise, incidental phase modulation (IPM) distortion, and electromagnetic interference (EMI). Two dominant factors affecting the transmitter noise floor in a digital broadcast transmitter system are incidental phase modulation and intermodulation distortion which are both functions of transmitter linearity.

Transmitter linearity can be measured using a two tone, single sideband, suppressed carrier signal. If the two tones are spaced 1 kHz apart, there will be spectral components above and below the frequency of the tones spaced 1 kHz apart. The spectral components that are 1 kHz from the two tones are called third order products because they are caused by the third order curvature of the amplitude transfer function.

In an ideal amplifier, the *amplitude* **(r)** and *phase* **(Ø)** signals will perfectly combine in the power output stage of an AM transmitter with the result being a perfect spectrum with only the desired signal. In practical terms, this combination has to be good enough to allow receivers to distinguish between the OFDM carriers and intermodulation products as well as meet the spectral mask defining the allowed out of band radiation of the transmitter. The ideal spectrum of this signal is shown in **Figure 3** using an arbitrary waveform generator in a closed loop using equal tones at 4 kHz and 5 kHz.

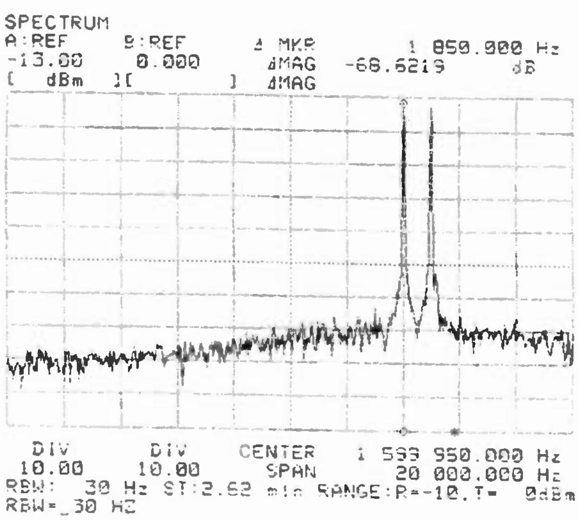


Figure 3
4 kHz and 5 kHz
Equal Tones

If the amplifier has a non-linear transfer function or amplitude to phase rotation caused by IPM, intermodulation products will be produced. The generation of intermodulation products due to phase errors can be better understood by simple vector analysis. At the point where the amplitude and phase signals are equal and 180° out of phase, the signals will cancel perfectly as shown in **Figure 4.** Any error caused by phase rotation will create error signals by imperfect cancellation of the signals. Critical alignment of the advance of the *amplitude* (**r**) signal and low transmitter incidental phase modulation are necessary to optimise cancellation.

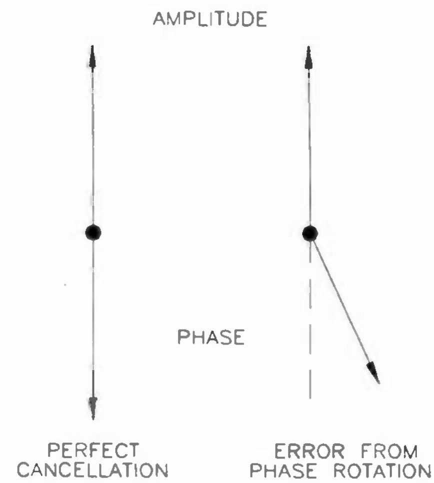


Figure 4
Vector Analysis

4.3 Transmitter Tests

To assess compatibility of PDM broadcast transmitters for use with digital transmission systems, several tests were completed both into a 50 ohm load and a band pass load to simulate the effects of a limited bandwidth antenna.

Tests were completed on AMPFET ND series transmitter and XL series transmitters. In both instances all the exciter stage audio filtering was bypassed and the audio input path was DC coupled. An arbitrary waveform generator was used to generate *amplitude* (**r**) and *phase* (**Ø**) signals to produce single sideband, suppressed carrier two tone test signals. The *phase* (**Ø**) signal was applied directly to the transmitters at the external RF input (stereo) BNC connector. The *amplitude* (**r**) signal was applied to the audio input terminals.

Adjustments were made to add the correct time advance of the *amplitude* (**r**) signal. Tests on the XL series transmitter were performed at various power levels including a carrier power of 10 kW per tone.

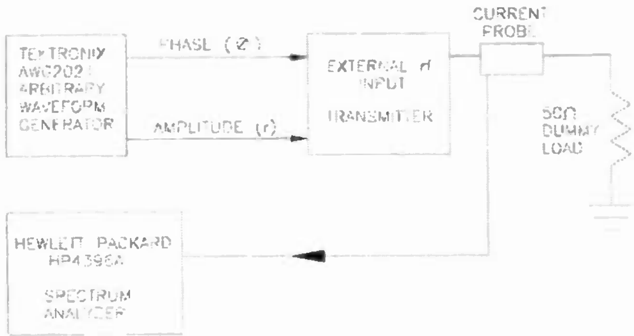


Figure 5
Test Set-up

Preliminary testing highlighted the importance of low incidental phase modulation.

Figure 6 shows the spectrum of a two tone test on an AMPFET ND1 transmitter. Equal amplitude tones at 4 kHz and 5 kHz were used. The IPM correction circuits were initially disconnected. The figure shows a high intermodulation product at -35 dB relative to the tones due to undesired phase rotation.

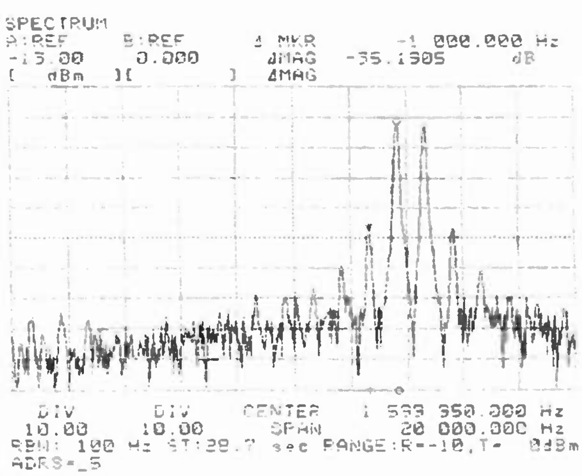


Figure 6
IQM Disconnected

The IPM correction circuitry was then adjusted to properly cancel incidental phase modulation caused by the two tones. The result was an improvement of 15 dB in the level of intermodulation product as shown in **Figure 7**. Due to the significance of low incidental phase modulation, new circuitry has been developed to maintain incidental phase modulation at about -40 dBr across a 30 kHz RF bandwidth.

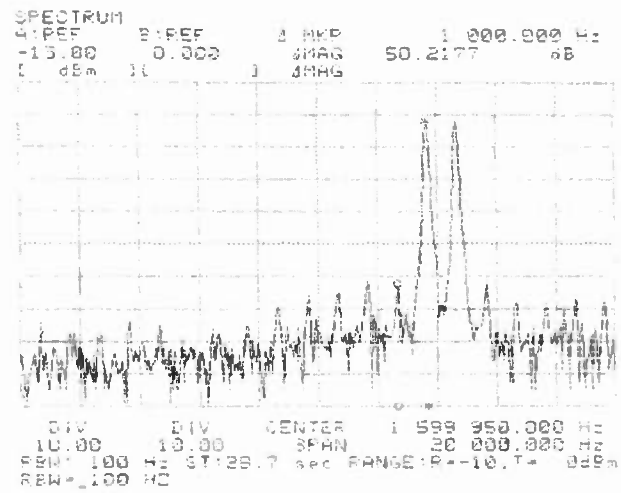


Figure 7
IQM Set-up

4.3.1 Test Results

The following test results were obtained at 1000 watts per primary tone for the AMPFET ND1 and 10,000 watts per primary tone for the XL12 transmitter. Tests were also completed at various lower power levels which yielded similar results.

TEST	AMPFET ND1	XL12
4 kHz and 5 kHz Equal tones	-50.1 dB	-59 dB
4 kHz and 5 kHz 5 kHz – 1 dB	-50.3 dB	-57.4 dB
5 kHz and 8 kHz 8 kHz – 30 dB	< -75 dB	< -75 dB
6 kHz and 10 kHz 10 kHz – 30 dB	N/A	-73 dB

Table 1
Level of Largest Intermodulation Product

The tests utilize 4 kHz and 5 kHz tones to assess transmitter linearity. A 5 kHz tone is used which, as noted, requires a bandwidth of 15 kHz to 20 kHz. In order to obtain acceptable system performance using a 32 QAM scheme, an average symbol energy to noise energy ratio in excess of 23 dB plus 5 dB margin should be achieved **(Reference 3)**. The two tone single sideband suppressed carrier tests indicate the transmitter linearity, not the actual symbol noise floor level, although directly related. A reasonable goal for intermodulation products would be –45 dB relative to the modulating tones for an all digital system based on limits of typical transmitter harmonic distortion performance. **Table 1** shows the highest intermodulation product relative to the highest modulating tone.

A more realistic test relating to hybrid IBOC systems is to apply two tones with the second tone -30 dB relative to the first tone. In a hybrid system where analogue programme is used in conjunction with digital carriers, we expect a limit of about –60 dB for the level of highest intermodulation product should be met to ensure full performance in fringe areas.

With tests completed into an ideal 50 ohm resistive load, a simple bandpass network was assembled to simulate a narrow bandwidth antenna.

Figure 8 shows the bandpass characteristics of the circuit used. The bandpass network which comprised series L and C components had a 3 dB bandwidth of ± 45 kHz. The 10 kHz sideband VSWR was just over 1.3:1. The tests were repeated using this band limited load to analyze its effects on transmitter linearity.

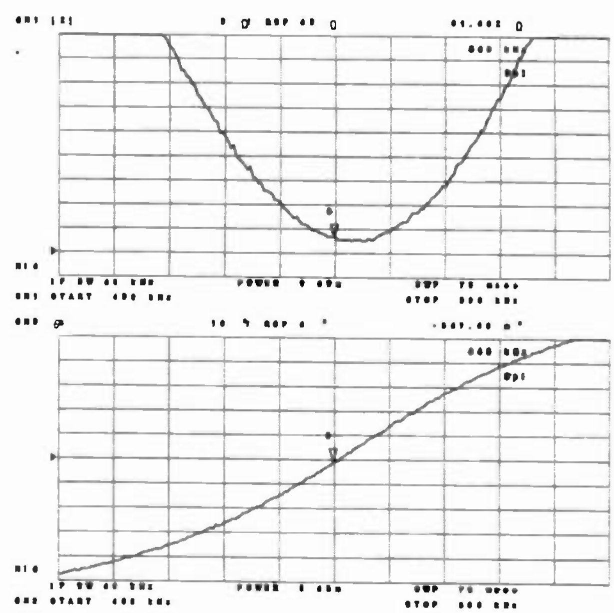


Figure 8
Bandpass Load Characteristics

Initially tests were completed with the same test settings used with a 50 ohm resistive load. Close inspection showed that the time advance for the *amplitude* (**r**) had to be adjusted to compensate for the changes caused by the bandpass load. A correction of 2 μS improved the level of unwanted signal by 10 dB. In practice this may be a routine adjustment in an IBOC exciter.

TEST	XL12
4 kHz and 5 kHz Equal Tones	-57.4 dB
4 kHz and 5 kHz 5 kHz – 1 dB	-53 dB
5 kHz and 8 kHz 8 kHz – 30 dB	-71 dB
6 kHz and 10 kHz 10 kHz –30 dB	-73 dB

Table 2
XL12 - Bandpass Load
Level of Largest Intermodulation Product

Table 2 shows level of the largest intermodulation product relative to the larger modulating tone when using a bandpass load with the XL12 transmitter operating at 10 kW.

Figure 9 is the spectrum display for the two tone test using 6 kHz and 10 kHz tones with 10 kHz tone –30 dB relative to the 6 kHz tone.

Other transmitter parameters relating to ease of facilitating IBOC transmission are the transmitters phase linearity or group delay variation, amplitude response of r channel and modulation bandwidth.

The transmitter phase linearity or group delay curves relate to the complexity required by the digital exciter to compensate for delay variation. It is implied that group delay variations less than 5μS are acceptable without correction. Both the AMPFET ND and the XL series transmitters meet this requirement. However, test results have shown a 10 dB change in intermodulation products with a 2μS change in amplitude channel advance.

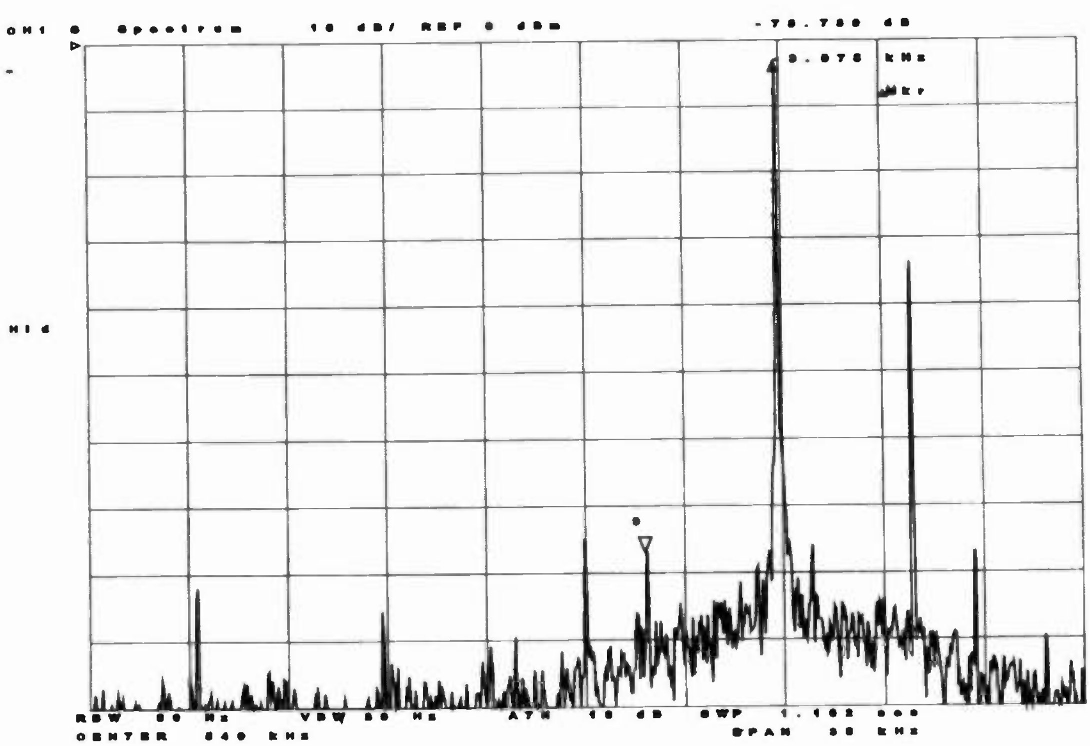


Figure 9
XL12 Transmitter
6 kHz / 10 kHz – 30 dB

This suggests IBOC exciters will require delay compensation for optimum performance especially in bandwidth limited antenna systems.

Transmitters must have sufficient modulation bandwidth to pass the IBOC signal of 15 kHz. Although achieved by both series of transmitters tested, the XL series PDM frequency is nominally twice that used in the AMPFET ND series. With a higher PDM switching frequency, the XL series transmitter is able to maintain a more linear response through the passband. The XL series transmitter requires a negligible amount of delay correction.

5.0 CONCLUSIONS

The test results provided a summary of transmitter performance specifications critical to facilitate digital broadcasting. We believe the XL series transmitter will be capable of passing the digital IBOC transmission as proposed by USA Digital Radio.

The XL series transmitter with 20% headroom above rated power of 10,000 watts plus 10% will not have amplitude restrictions due to high crest factors.

IBOC exciter manufacturers must consider delay equalization to compensate for band limited antenna installations.

ACKNOWLEDGEMENT

I would like to thank my colleagues at Nautel. Dennis Covill has provided valuable theoretical contribution to the work described in this paper. I would also like to thank Jean Légère for his patience in programming our Arbitrary Waveform Generator from my plethora of mathematical expressions.

REFERENCES

1. "Digital Audio Broadcasting For Band 6 - In-Band, On-Channel Progress", CCIR Study Groups, 13 January 1993

2. R.L. Cupo, M. Sarraf, M. Shariat, M. Zarrabizadeh, "An OFDM All Digital In-Band_On-Channel (IBOC) AM and FM Radio Solution Using the PAC Encoder", IEEE Transactions on Broadcasting, Vol 44 No.1 March 1998

3. D. Hartup, "Transmission of USADR IBOC Waveforms", NAB, Spring, 1998

ABOUT THE AUTHOR

Wendell Lonergan is Nautel's Project Leader for DAB Transmission Studies on AM transmitters. He has 20 years experience in transmitter research, design and manufacture. During this time he has had a leading role in the development of several generations of Nautel solid state AM broadcast transmitters, as well as the development of a ©C-Quam AM stereo exciter (under license to Motorola). In addition to his work on IBOC compatibility, Wendell is engaged in development of high power Medium Wave transmitters and combiner systems for super high power international projects.

Frequency Domain Reciprocal Modulation (FDRM) for Bandwidth-Efficient Data Transmission Over Channels with Dynamic Multipath

Thomas H. Williams
Holtzman Inc.
6423 Fairways Dr.
Longmont, CO 80503

Abstract

This paper discusses a new modulation technique that is designed to provide robust and bandwidth-efficient digital communications over a signal path that contains dynamic multipath distortion. The modulation technique uses two blocks of data that are sent adjacent to each other in time so the signal path applies the same echoes to each block. The two blocks contain the same information, but the second block is the reciprocal in the frequency domain of the first block. At the receiver, the two blocks are processed together to find the transmitted symbols without the linear distortion and optionally the frequency response of the channel. Receiver implementation issues and system design challenges are also discussed.

1.0 Introduction

Echoes, which are delayed versions of the original signal, are also known as multipath distortion or ghosts. Echoes, like channel tilt and group delay, are linear distortions. The transmission of digital signals through an environment that is contaminated by multipath is difficult because reflections cause inter-symbol interference (ISI). If there is a small level of ISI, the bit error rate will increase in the presence of random noise. If the level of ISI is high, error-free reception will not be possible. One solution to the echo problem is the adaptive equalizer, which was invented in the 1960's. The adaptive equalizer, which is well known by transmission engineers, uses a structure with delay elements, multipliers, and a summer. If the adaptive equalizer is programmed with the inverse of the channel's impulse response, the echoes contaminating the channel can be eliminated. In theory, the reciprocal of a single echo's impulse response is an infinitely recursive series, requiring an infinite number of stages or taps in the adaptive equalizer. However, a limited number of taps typically can provide adequate echo cancellation.

A problem arises when the echo's characteristics change, rendering the adaptive equalizer's programming wrong. Considerable attention has been devoted to the problem of rapidly reprogramming the adaptive equalizer to track a dynamic or moving echo. Solutions include training or reference signals, as well as blind equalization techniques. The presence of noise typically increases the programming time.

Echoes may change for a variety of reasons, such as swaying of a transmit tower, movement of the receiver in a multipath environment, and movement of the echo-producing objects. In some transmission environments, the received signal might be comprised of only echoes, while in other environments, the main signal might be changing characteristics (e.g. rapid fading).

In the 1960's another modulation technique was being developed that used multiple orthogonal carriers as an alternate digital signal transmission method. This technique is called orthogonal frequency division multiplexing (OFDM). The technique was not widely used until low-cost digital signal processing (DSP) integrated circuits were developed that could rapidly perform the inverse fast Fourier transform (IFFT) and fast Fourier transform (FFT) algorithms used by this transmission method. Assuming

1. the received signal is noise free,
2. a guard interval (GI) is employed,
3. the longest echo is shorter than the guard interval,
4. and the echo does not cancel the main signal at any frequency,

the echo's distortion can be eliminated from the main signal. (A guard interval is nothing more than a series of time samples cut from the end of a data block and appended to the front.) Echoes may still contaminate the received OFDM burst, but a single complex multiplication on each of the harmonically related carriers can cancel the effect of the echo. To assist in the determination of the necessary complex multiplication coefficient, some of the harmonic carriers are given a pilot tone

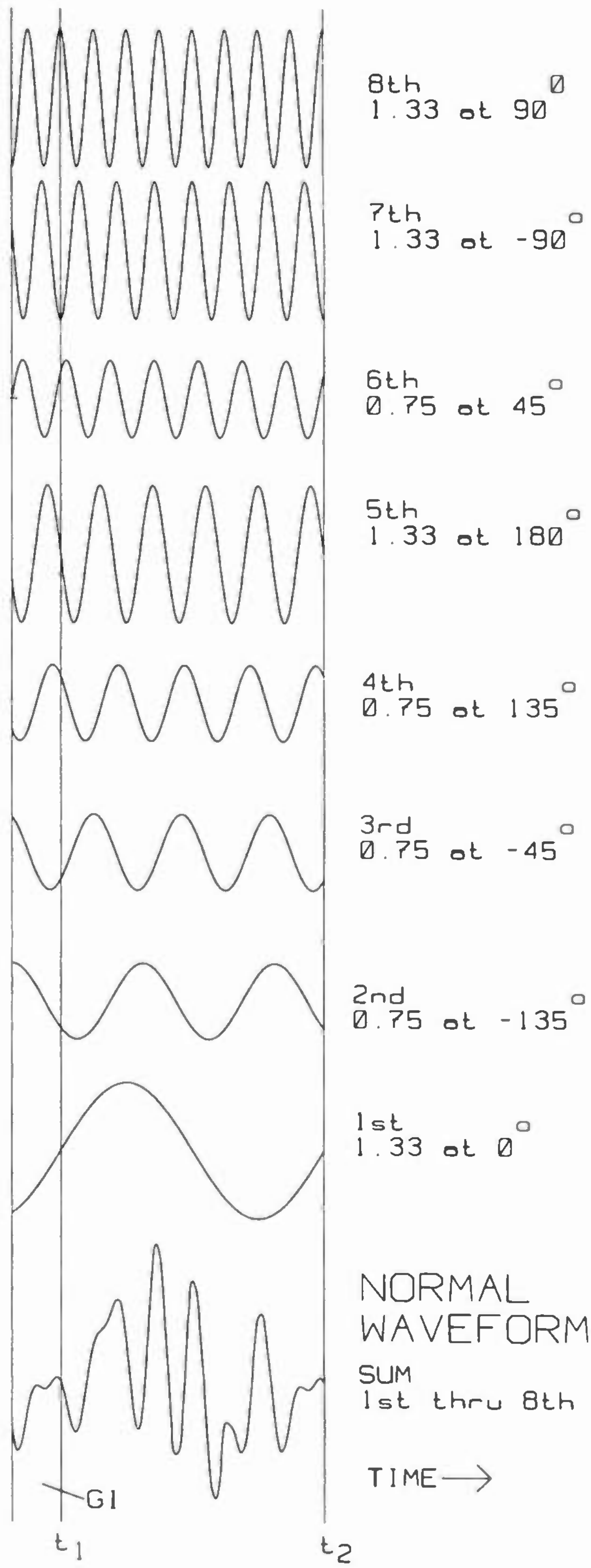


Figure 1 A Normal OFDM Waveform Comprised of Eight Harmonic Carriers (HCs)

status. A pilot signal may be viewed as a signal of known amplitude and phase. The sum of the pilot tones together may be viewed as a type of a reference signal. Another method used to counteract the effect of echoes in OFDM is the use of differential encoding.

2.0 Background on OFDM

OFDM may be viewed as a high-data rate transmission system composed of many low-data rate carriers. For example, an OFDM burst might consist of just a single sinewave oscillating for a number of complete cycles. Figure 1 is a diagram showing an OFDM burst constructed of a linear sum of eight discrete harmonically related carriers (or HCs), each with a different frequency. Each of the HCs completes an integer number of cycles between time t_1 and time t_2. By changing the magnitude and phase of each individual HC, it is possible for each HC to independently transport data. The bottom baseband waveform is the sum of the 8 HC waveforms above it. This baseband waveform may be transmitted over a baseband channel such as a telephone line, or linearly modulated onto a radio frequency (RF) carrier and broadcast into space. Reference [1] discusses modulation of OFDM.

Only eight HCs were chosen for the sake of illustration. In the European digital television terrestrial transmission system, an OFDM burst is comprised of thousands of HCs.

In Figure 1 the magnitude of each of the eight HCs is allowed to be at one of two levels and the HC's phase is allowed to be at one of 8 allowed angles. The magnitude and phase state of a single HC is frequently called a "symbol". The number of bits of information conveyed by each symbol is 2^N , where N is the total number of allowed symbol states. Figure 1 has another feature: a guard interval has been made by cutting a sample from the end of each of the waveforms and pasting it on to the beginning.

The symbols may be placed into a constellation diagram, as shown in Figure 2, by plotting the magnitude and phase data for each harmonic carrier as a single point. Thus, by modifying the phase and magnitude of each harmonic carrier, any one of 8 possible symbol states are reached, and 3 bits of information are conveyed by each symbol.

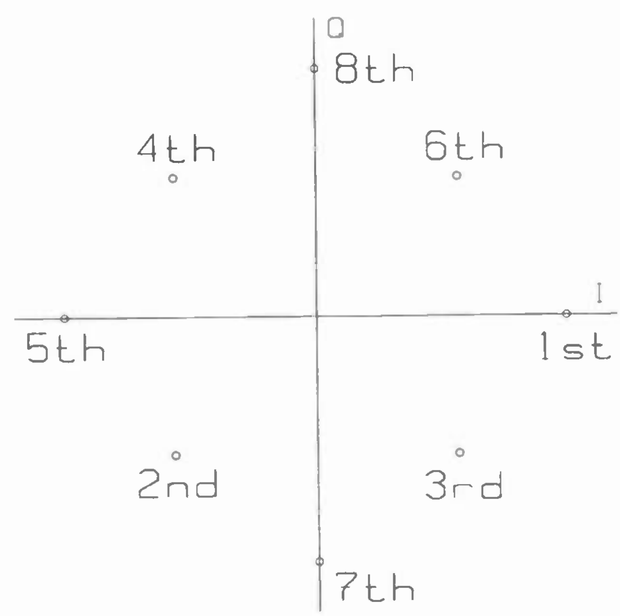


Figure 2 An Eight Point Normal Constellation Diagram for an OFDM Burst Transmission

The magnitudes and phases of the 8 harmonic carriers are pre-assigned so that each of the possible 8 symbol states in Figure 2 are hit. The harmonic number of the HC that resides in each symbol state is identified on Figure 2.

If the burst OFDM transmission is contaminated by random noise, the noise energy will cause the constellation points to be moved from their ideal positions. When many noise-contaminated points are superimposed on a constellation diagram, the composite constellation points appear to "spread out" or become larger and take on a rough outline.

To avoid having HCs interfere with each other they have to be created orthogonal to each other. A measure of orthogonality is:

$$E = \int_{t1}^{t2} e_1(t) \cdot e_t(t)dt \qquad (1)$$

where E=0 if signals $e_1(t)$ and $e_2(t)$ are orthogonal to each other over the time interval between t_1 and t_2. Sine and cosine waves are functions that are orthogonal to each other provided two conditions are met. First, the frequencies must be integer multiples of some fundamental frequency, and the time for integration must be over the period of the fundamental frequency. For example:

$$E = \int_0^{2\pi/\omega} A_4 \cos(4\omega t + \phi_4) \cdot A_5 \cos(5\omega t + \phi_5)dt = 0 \qquad (2)$$

where the fourth and fifth harmonic carriers are orthogonal, not just to each other, but all other harmonic carriers as well. A_4 and A_5 are the respective magnitudes of the carriers and ϕ_1 and ϕ_2 are the respective phase angles of the carriers. These magnitudes and phases may take on arbitrary values and the HC's will still be orthogonal.

The guard interval allows the OFDM composite burst to still maintain the orthogonal relationship between the individual carriers when there is a long echo contaminating the channel, provided that the long echo is shorter than the duration of the guard interval.

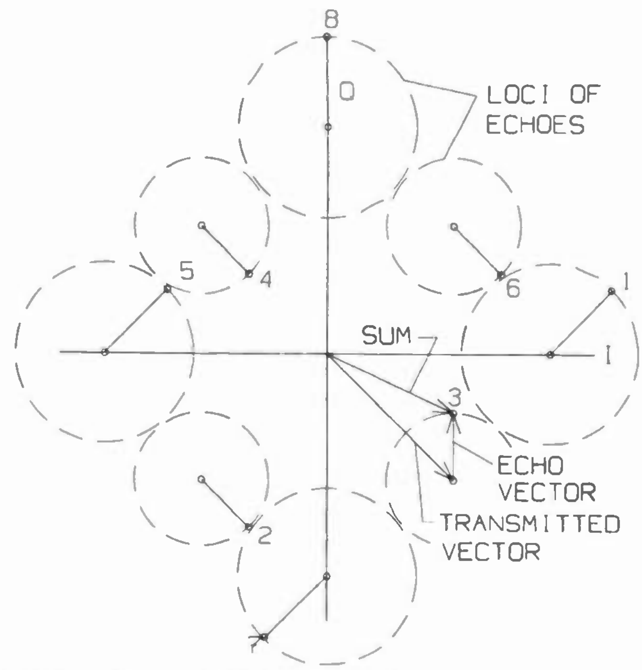


Figure 3 An Eight Point Normal Constellation Diagram for an OFDM Burst Transmission with an Echo

Figure 3 is a constellation diagram of the OFDM burst shown in Figure 2, but a single echo has been added with a magnitude of 40% of the main signal. The echo is shorter than the guard interval. Note that the constellation points are all moved by error vectors created by the echo from their nominal points to other points located on circles. The radius of each circle is 40% of the magnitude of its unimpaired vector. The third harmonic is illustrated in greater detail as an example. To cancel the echo vector, a single complex multiplication needs to be made to each of the 8 harmonic carriers in the burst transmission. If the echo is changing rapidly, the challenge for the transmission engineer is to find the correct complex coefficients needed to perform the complex multiplication.

As mentioned earlier, there are solutions available, such as pilot HCs that can be used to assist with finding the solution. However pilot carriers consume bandwidth and only reveal the exact channel characteristics at the pilot frequency.

3.0 Using Two Blocks of Data to Cancel Echoes

The problem with received data distorted by unknown echoes may be viewed as a problem with two unknowns in only one equation. The unknowns are the symbols and the characteristics of the echoes (delay and attenuation). The conventional solution is to eliminate one of the unknowns, with a training signal or with redundancy in the signal. Another new direct method, called "Frequency Domain Reciprocal Modulation" (FDRM), is discussed below.

An assumption is being made is that discrete time and frequency arithmetic is being performed on blocks of sampled data. Thus a FFT and IFFT can be used to jump between the time and frequency domains.

If a normal (N) transmitted signal $S_N(t)$ is sent through a channel with an impulse response that is H(t), the received signal $X_N(t)$ is the convolution of the transmitted signal with the impulse response:

$$X_N(t) = H(t) * S_N(t) \qquad (3)$$

If these normal (N) transmitted and received signals are viewed in the frequency domain by performing the FFT, the equation for the received signal becomes:

$$X_N(f) = H(f) \cdot S_N(f) \qquad (4)$$

where H(f) is the frequency response of the channel. Each different discrete value of (f) is the frequency of a different HC.

If one knew H(f), the undistorted signal could be found from:

$$S_N(f) = X_N(f) \cdot H(f)^{-1} \qquad (5)$$

Now assume a second transmitted signal $S_R(f)$ is made with the reciprocal (R) of the information in $S_N(f)$

$$S_R(f) = \frac{1}{S_N(f)} \qquad (6)$$

Creating a reciprocal of a harmonic carrier with a magnitude and an angle is an easy calculation. The reciprocal magnitude is the inverse of the original magnitude, and the reciprocal angle is the negative of the original angle.

If the same echo also distorts the received second reciprocal signal, the result is:

$$X_R(f) = H(f) \cdot S_R(f) = \frac{H(f)}{S_N(f)} \qquad (7)$$

If the received second reciprocal signal is inverted and multiplied by the received first normal signal the result is:

$$X_N(f) \cdot \frac{1}{X_R(f)} = H(f) \cdot S_N(f) \cdot \frac{S_N(f)}{H(f)} = S_N(f)^2 \qquad (8)$$

and if the square root is taken on the result, the undistorted normal signal can be found from:

$$S_N(f) = \sqrt{S_N(f)^2} \qquad (9)$$

Likewise, if the received normal signal is multiplied by the received reciprocal signal at the receiver location:

$$X_N(f) \cdot X_R(f) = H(f) \cdot S_N(f) \cdot \frac{H(f)}{S_N(f)} = H(f)^2 \tag{10}$$

so the frequency response of the echo contaminated channel is:

$$H(f) = \sqrt{H(f)^2} \tag{11}$$

From an implementation point of view, equation (6) represents a problem for a signal block that has zero energy at some frequency: a division by zero problem occurs. For example, a burst of an 8-VSB (8 level vestigial sideband) modulated signal will likely have frequencies at which there is very low energy. Therefore, the reciprocal will have very large frequency components at those frequencies and thus be impractical for transmission through a power amplifier with limited dynamic range. The solution is to use a signal that has non-zero energy at all frequencies where the reciprocal will be calculated. OFDM is one signal type that meets this requirement. The energy at each frequency for an OFDM burst signal may be set in the transmitter to be a non-zero value.

4.0 Square Root Transmission

Another way to implement the frequency domain reciprocal modulation (FDRM) system is to perform a square root function at the transmitter and not perform it at the receiver. That is, the transmitted normal signal becomes $\sqrt{S_N(f)}$ and the transmitted reciprocal signal becomes:

$$\sqrt{S_R(f)} = \frac{1}{\sqrt{S_N(f)}} \tag{12}$$

If a received contaminated normal square root burst is divided by a received contaminated reciprocal square root burst the result is:

$$\frac{X_N(f)}{X_R(f)} = \frac{\sqrt{S_R(f)} \cdot H(f)}{\frac{1}{\sqrt{S_N(f)}} H(f)} = S_R(f) \tag{13}$$

which is the undistorted normal signal.

5.0 FDRM for Adjacent HCs

Frequently echoes change rapidly in time but they change very little between adjacent frequencies. In this case, the FDRM may be used between two carriers that are close in frequency, but in the same data block. If this technique were applied to the example of Figure 1 the sequence [1N, 2N, 3N, 4N, 5N, 6N, 7N, 8N...1R, 2R, 3R, 4R, 5R, 6R, 7R, 8R] would be replaced by [1N, 1R, 2N, 2R, 3N, 3R, 4N, 4R.... 5N, 5R, 6N, 6R, 7N, 7R, 8N, 8R]. Where two adjacent HCs, such as 2N and 2R, would be presumed to have approximately the same echo distortion.

Thus, a two block transmission system can be made having the properties that the coefficient of each HC in the second block is the reciprocal in the frequency domain of the corresponding HC in the first block. This transmission system can have very high immunity to moving echoes provided that the same set of echoes is applied to both blocks. On the surface, it appears that half of the channel capacity is wasted since it takes two blocks of data to yield one block of information. However the

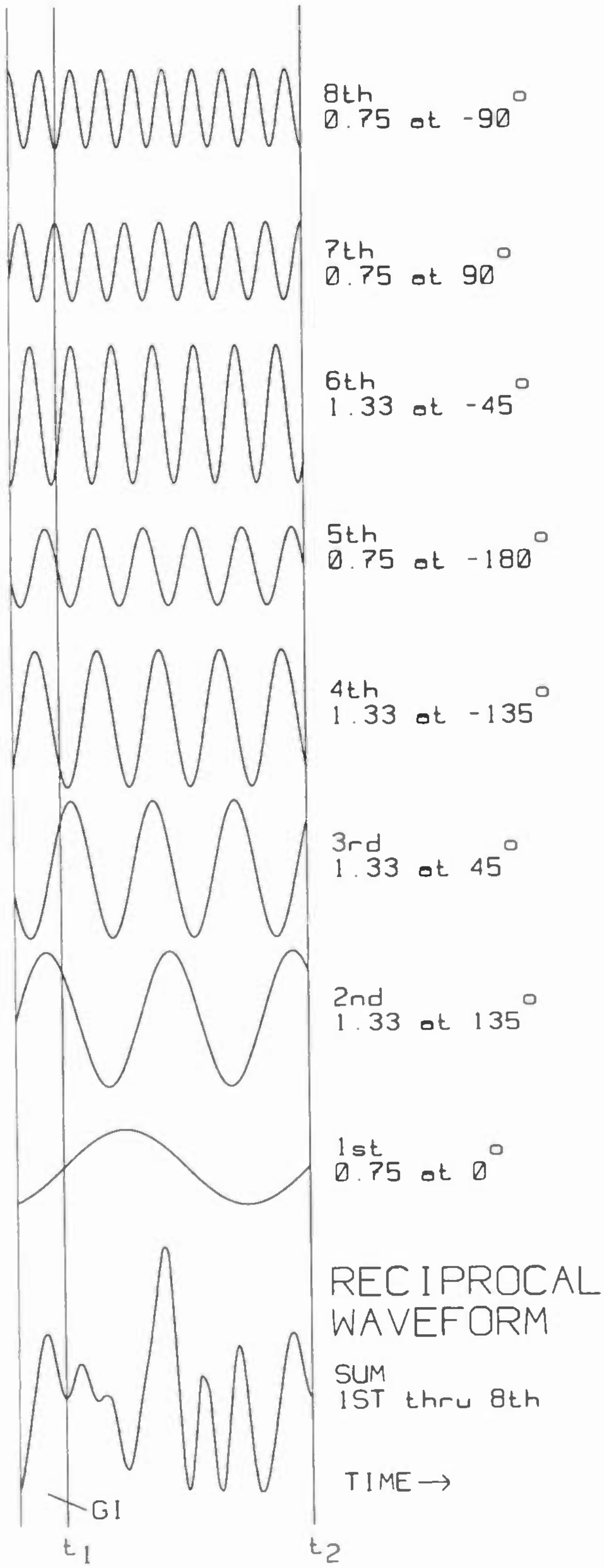


Figure 4 A Reciprocal OFDM Waveform Comprised of Eight Harmonic Carriers (HCs)

noise in each block adds on a power basis while the signal adds on a voltage basis, so there is an improvement in the signal to noise ratio of 3 dB. Thus the transmit power can be reduced by half for a comparable bit error rate performance.

Compared to other echo-tolerant modulation techniques, such as CDMA (code division multiple access) FDRM is highly bandwidth efficient. A patent is pending.

One implication of FDRM is that ghost canceling reference signals can and should transport data as well as provide a reference signal to align a conventional adaptive equalizer. Thus, the signal on line 19 of the vertical blanking interval of an NTSC television transmission, which is a static reference signal, represents a lost opportunity to provide a dynamic echo-tolerant data service.

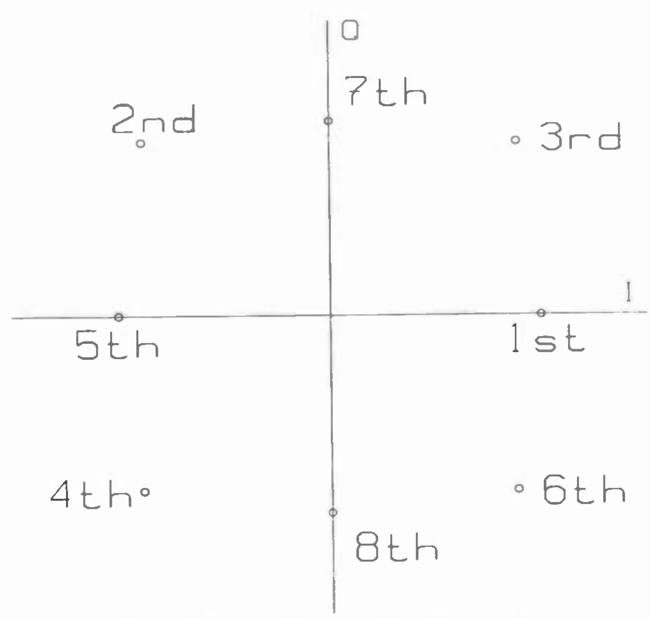


Figure 5 An Eight Point Reciprocal Constellation Diagram for an OFDM Burst Transmission

Figure 4 is a reciprocal time domain plot, showing a reciprocal waveform to the waveform of Figure 1. Note that if any number HC in Figure 1 had a large amplitude, it has a small amplitude in Figure 4, and vise-versa. Likewise, if a any number HC had a positive phase angle in Figure 1 it has the same phase angle in Figure 4, except negative.

Figure 5 is a R constellation plot associated with the reciprocal block of data. It may be compared to the N constellation plot in Figure 2.

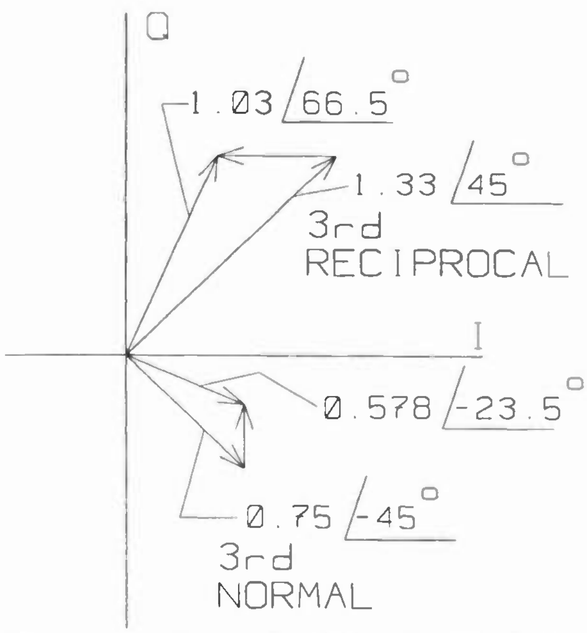


Figure 6 An Example of How the Third HC gets Deghosted Using the Third HCs from the N and R Blocks.

6.0 An Example

Figure 6 illustrates a numerical example that is useful to demonstrate the de-ghosting process. The third HC was arbitrarily chosen using the third HC from the first N waveform and the third HC from the second waveform. Referring back to Figure 3, the third HC was transmitted with a magnitude of 0.75 and an angle of -45 deg. The 3rd N HC was distorted with an echo that had a magnitude of 0.4 and a delay of 12.5% of the time between t_1 and t_2, which is 135 degrees for the third HC. This produced an error vector of 0.300 at 90 deg., as illustrated in Figure 6. The vector sum of the 3rd HC and its echo is therefore 0.578 at -23.5 deg.

The third HC in the R waveform, being the reciprocal of the third HC in the N waveform, was transmitted at 1.33 at +45 deg. The same echo (0.4 at 135 deg.) distorted it, producing an error vector of 0.533 at an angle of 180 deg. The vector sum of the received distorted R signal is therefore 1.03 at 66.5 deg.

The square of the unimpaired 3rd harmonic therefore has a magnitude of 0.562 (0.578 / 1.03) at an angle of -90 deg. (-23.47 deg. - 66.53 deg.). If the square root is taken, the correct answer for the amplitude of the transmitted 3rd HC is 0.750 at -45 deg.

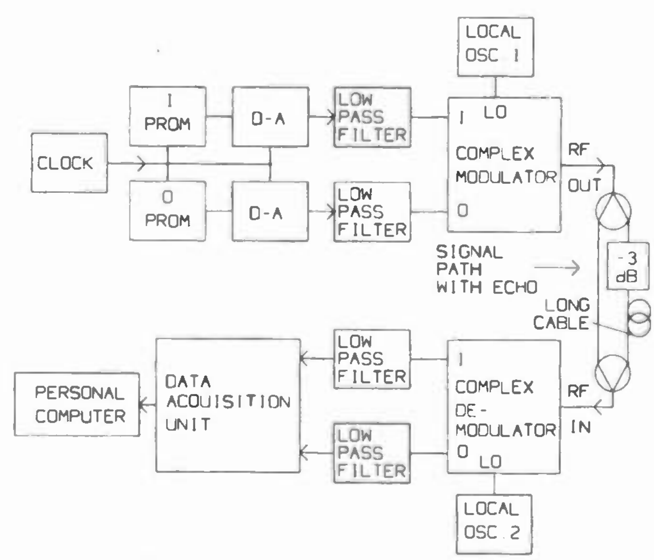


Figure 7 Demonstration Hardware Block Diagram

7.0 Demonstrating the Idea with Hardware

A demonstration system was constructed using a two-block transmission system. See Figure 7. The two blocks of random data are stored in a pair of PROMs (programmable read only memories). One PROM contained the I data and one contained the Q data. The PROMs together generate the normal block of data followed by the reciprocal block without any pause. The output of the I and Q digital-to-analog converters are low pass filtered before being applied to a complex modulator. The RF output of the complex modulator is at 70 MHz. After passing through a signal path with an optional -3 dB echo, the inverse process is used to demodulate at the receiver. A digital storage oscilloscope is used as an analog-to-digital converter, and data is passed into a PC where it is processed to produce the display shown in Figures 8 and 9.

In the signal processing the GI is discarded, the I values are used for the real numbers in the FFT, and the Q values are used for the imaginary values in the FFT.

TABLE 1 Transmission Parameters

OFDM Constellation Type	QPSK
Transmit Frequency	70.00 MHz
Occupied Bandwidth	7.808 MHz
Number of HC's	798
HC Frequency Separation	9.76 kHz.
TX and RX sample rate	10 M/sec.
N & R Burst Duration	102.4 ❍s
Guard Interval	12.8 ❍s
Total Burst Duration	230.4 ❍s
FFT / IFFT Size	1024
Low Pass Filter's Corner Freq.	4 MHz.
Number of Bits in A-D, D-A	8

Table 1 is a list of the transmission parameters used by the demonstration system.

8.0 Comments on the Display

Figure 8 is a set of plots from a signal path without an echo and Figure 9 is a set of plots from a burst with an echo. On both plots, the upper two traces are the captured I and Q time domain waveforms as they appear on the screen of the digital oscilloscope. The N waveform is followed by the R waveform without break, so the two blocks are run together. The bottom trace is the spectral energy including the upper and lower sidebands (USB and LSB) of the N burst. The set of 4 uncorrected OFDM constellations are split out by N or R blocks, as well as upper and lower sidebands. The N-USB harmonic carriers are processed with the R-USB harmonic carriers to produce the FIXED-USB constellation, and the N-LSB is processed with the R-LSB to produce the FIXED-LSB constellation.

Note that Figure 8 is supposed to be an echo-free constellation, but several impairments are evident in the spectral plot and uncorrected constellation plots. In particular, the frequency response and group delay of the low-pass filters of Figure 7 are evident. Also, the uncorrected constellation plots are rotated. Note that the FDRM modulation system corrects linear impairments from the transmission equipment. One harmonic carrier in the USB has been identified by a line connecting the constellation point to the origin. The roughness of the spectral plot is caused by IQ magnitude and phase imbalance in the complex demodulator. The LO in the receiver is free-running relative to the LO in the transmitter in Figure 8.

In Figure 9 the echo causes the ripple in the spectral plot, and some interesting patterns in the uncorrected constellations. Note that the 2 local oscillators are locked.

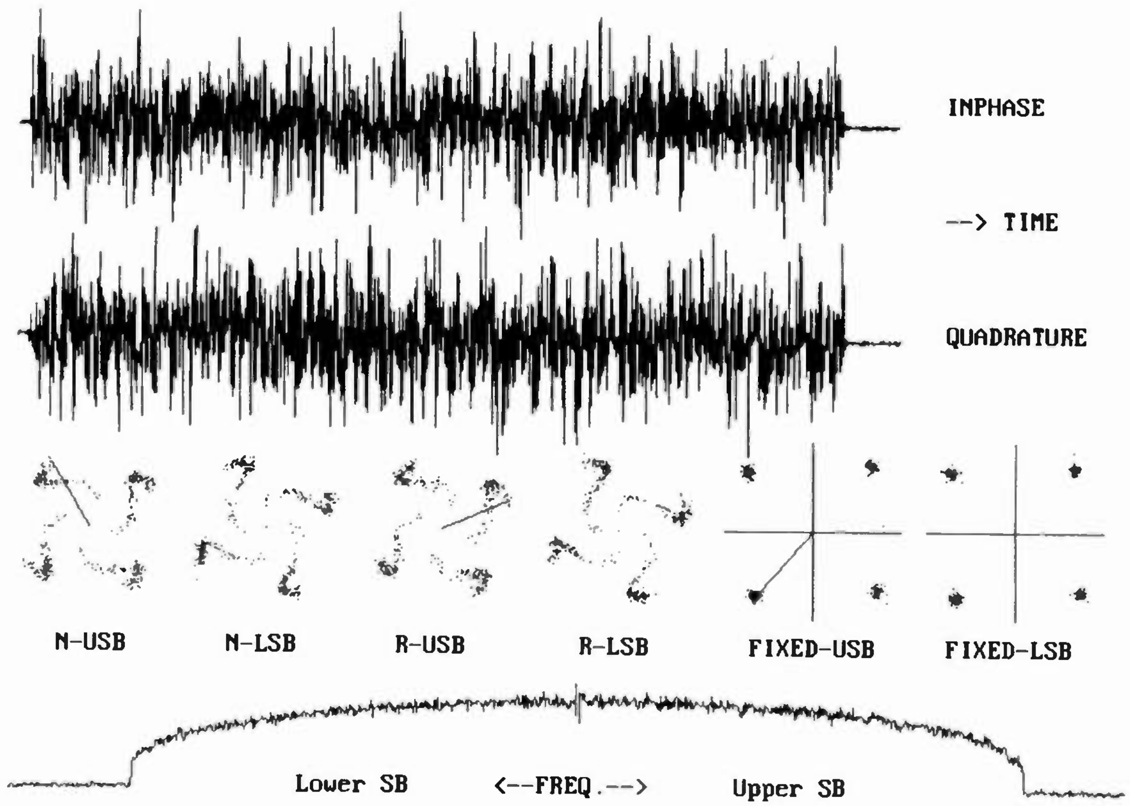


Figure 8. A Set of Plots Associated with a Signal Path without an Echo

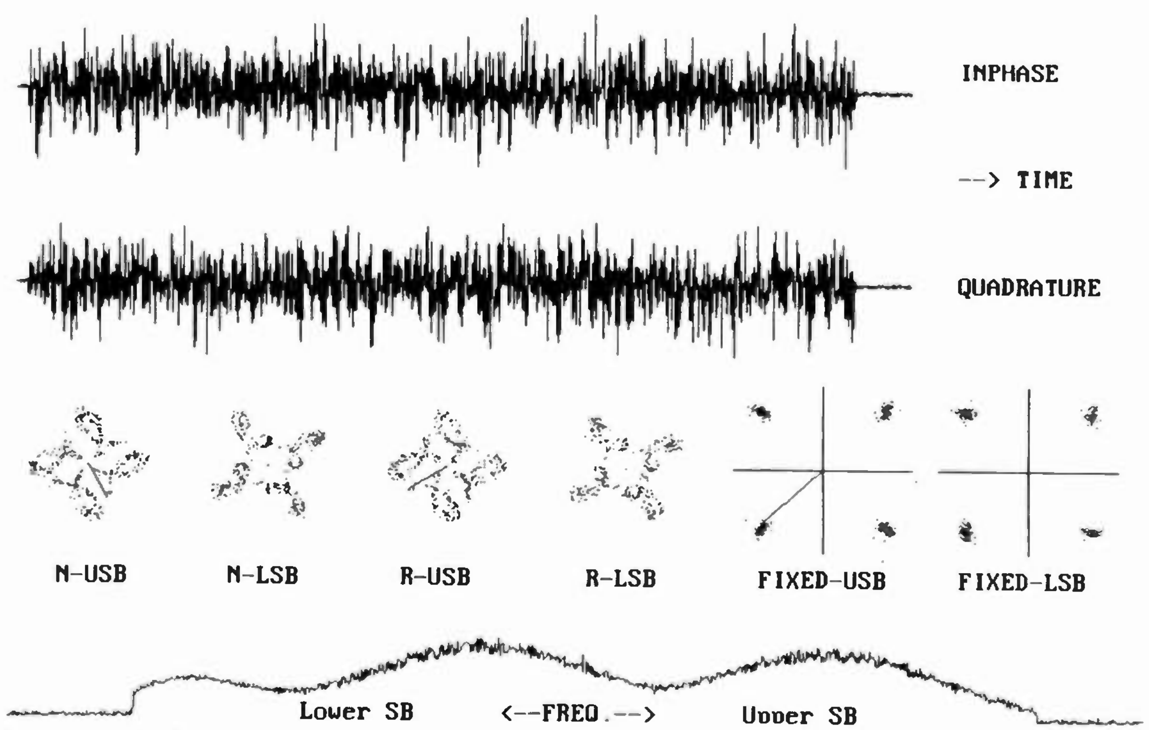


Figure 9. A Set of Plots Associated with a Signal Path with an Echo

9.0 System Design Considerations

There are a number of interesting characteristics about this novel system that should be discussed so that an engineer can evaluate the system.

9.1 Phase Locking

It is not necessary to phase lock the transmitter and receiver local oscillators. If they are not phase locked and there is a frequency difference, the corrected received constellation will still be tight, but have a rotational skew. This angular skew can be measured and used for a receiver local oscillator frequency correction. This characteristic makes this transmission system well suited for infrequent transmissions of a single or a few bursts, such as IP (internet protocol) datagrams, as well as for continuous transmissions. It is highly desirable to sample the N and R bursts with a continuous clock.

9.2 Sampling Start Error

If the start of sampling in the analog-to-digital converter has a timing error that is less than the guard interval, the recovered constellation will appear to be correct in all cases. It will actually be correct in the "square root at the transmitter" case. In the "square root at the receiver case", typically many of the symbols will be located in the wrong quadrant. This is because a HC with an unknown start time will have a 180 degree rotational uncertainty. This can be corrected by computing a "split" angle between the N and R HCs. If the "split" angle changes between adjacent frequency HCs, the rate of angular change will be the sampling delay error which can be corrected.

9.3 Randomizing

The "square root at the transmitter" constellation has a DC component because all of the constellation points are in the right half of the constellation diagram. An impulse will be created by the IFFT which may cause clipping in a transmitter. This can easily be corrected by randomization at the transmitter before performing the IFFT and de-randomization it after performing the FFT at the receiver.

9.4 Spectral Suck-Outs

It is probable that a set of multiple echoes on terrestrial channels may produce a deep frequency suck-out at some frequency. If the energy is low at some HC's frequency, the resulting symbol will be excessively corrupted by random noise. Conventional OFDM shares this problem. One solution is to use an error-correcting code, such as a linear Reed Solomon code, to provide robust error-free performance under these adverse channel conditions.

9.5 Hardware Implementation Techniques

The modulation approach that was taken in the hardware illustrated in Figure 7 is the "traditional IQ complex modulator" and is not generally advised. It was taken in this case because of the memory limitations in the digital oscilloscope. It is better to use the "modern approach" which is to generate the QAM modulation by direct synthesis of the carrier at a low frequency, such as 5 MHz, and up-convert to an IF frequency using a single mixer. At the receiver, sampling can be done with a single analog-to-digital converter. The resultant samples are multiplied by sin(•t) to recover the Q time samples and cos(•t) to recover the I time samples.

9.6 Doppler Shift
Doppler shift caused by receiver motion looks like a frequency shift, and will rotate the received corrected constellation. Reference [2] discusses the problem in a mobile environment.

9.7 Partial Use of FDRM
FDRM may be used part of the time to provide channel characterization for the case where echoes change infrequently or slowly. Another use of FDRM is to replace the static pilot tones used by conventional OFDM. The pilot tone frequencies can be made to yield channel characterization information by putting them into a reciprocal relationship with the pilot tone frequencies in the next data block. In other words, FDRM modulation can be applied only on the HC frequencies that were formerly static pilot carriers. Likewise a pair of adjacent frequency HCs, as discussed in section 5.0, may be used as information bearing pilots.

9.8 Selection of Constellations
Constellations that have points near the origin make poor choices for FDRM because the reciprocal point has a large magnitude. In OFDM, a constellation point on the origin indicates that the magnitude of that HC is zero. Circular constellations such as 8-PSK are good. 16-QAM would be a relatively poor choice because the four points nearest the origin would have a large amplitude in the reciprocal constellation. Star shaped constellations, such as illustrated in Figure 2, can also be used.

9.9 Block Length vs. Echo Rate of Change
The underlying idea of using two blocks to cancel echoes is that approximately the same echo distorts both blocks. If changing echoes are caused by vehicle motion the echoes will change quicker as the wavelength of the RF carrier shrinks for a give vehicle speed. Therefore for rapid motion at very high (e.g. microwave) frequencies the block length should be short duration. The blocks can be made with a wider bandwidth to carry the same data.

10.0 Conclusion
This paper presented a novel transmission method called Frequency Domain Reciprocal Modulation (FDRM). Its characteristics should allow data transmissions that are highly resistant to dynamic multipath and other linear distortions. After explaining the theory of operation, a numerical example was presented, followed results from test hardware. Finally a number of implementation details were discussed. The demonstration hardware shows that FDRM works. This system should find applications where the linear distortion is dynamic.

References
[1] M. Chelehmal, T. Williams, *Using Orthogonal Frequency Division Multiplexing in the Vertical Interval of an NTSC Transmission*. 1995 NAB Broadcast Engineering Proceedings, pp. 69-78
[2] L. Thibault, G. Soulodre and T. Grusec, *EIA/NRSC DAR Systems Subjective Tests Part II: Transmission Impairments*, IEEE Transactions on Broadcasting, VOL. 43, NO4, December 1997

Building the DTV Station Infrastructure

Sunday, April 18, 1999
1:00 PM - 5:30 pm

Chairperson: Jerry Whitaker
Technical Press, Morgan Hill, CA

***1:00 pm Broadcast Facility Design & Construction**
Daniel Taylor
DTA Carlson
Chicago, IL

1:30 pm High Performance Networking for Professional Video Applications
Peter Owen
Quantel Ltd.
Berkshire, United Kingdom

***2:00 pm The Reality of ATM Networks for Video Transmission**
Richard Bauarschi
Synctrix, Inc.
Glendale, CA

2:30 pm 540 Mb/s Serial Interface Pioneering New Standardization Approach at SMPTE
Dan Turow
Gennum Corporation
Burlington, Canada

***3:00 pm Real Time MPEG-2 Manipulation Over a Broadband Network**
Matthew Green
DiviCom, Inc.
Milpitas, CA

***3:30 pm Ramifications of Multi-Layered Storage for Digital Television Facilities**
Todd Roth
Leitch Incorporated
Burbank, CA

4:00 pm Asset Management Across Your Facility - A Case Study of the LucasFilm Digital Media Management Infrastructure
Timothy Campos
Silicon Graphics Inc.
Mountain View, CA

***4:30 pm Planning, Designing and Implementing Digital Terrestrial Television Networks**
Graham Warren
BT Broadcast Services
London, England

5:00 pm Old Earth Station IFL Cables and New DTV Downlink Performance
Peter Zilliox
Andrew Corporation
Richardson, TX

*Papers not available at the time of publication

High Performance Networking for Professional Video Applications

Peter Owen,
Quantel Ltd
Newbury, UK

Abstract

Networking technologies are finding their way into post production and broadcast facilities. The range of technologies is vast and the choice complex for broadcast applications where data file sizes exceed those of traditional information technology applications by factors of hundreds or thousands. Keeping the traditional values of speed is one essential attribute when incorporating networking in broadcast applications. On the other hand keeping the openness of the computer world is equally essential when interfacing between high speed broadcast dedicated discs and processors and the more general computer platforms and disc arrays.

This paper describes the attributes required for such a solution and the investigation and choices which led to the adoption of Gigabit Ethernet as the connection technology to support Clipnet™.

1.0 Introduction

Fast and efficient interchange information is the key to programme making for broadcast. We have long since mastered the vagaries of PAL and NTSC. Component analogue was a natural evolution which led the way to digital coding culminating in ITU-R 601. But now we face a new set of challenges. .
Video compression is commonplace, both 16x9 and 4x3 aspect ratios co-exist. Multiple new digital high definition formats are competing for air-time and our ever
closer working relationship with the computer industry has brought with it a plethora of different file formats, interchange protocols, physical connections together with square pixels and a demand for progressive scanning. Digitisation also brings with it the ability for interchange with allied disciplines in film, print and the world wide web. Re-purposing and therefore inter technology connectivity is no longer just an option, it is becoming increasingly essential in our never ending quest for maximum efficiency.

Key issues with regard to the interchange of data between systems and choices made by Quantel in the development of a new high performance network called Clipnet™ for interchange of audio, video and metadata over a standard computer network. are explored in this paper.

2.0 Streaming and File Transfer

Video streaming techniques rely on point-to-point connections, with guaranteed bandwidth, little or no latency and no means of communication between source device and destination device. The latest combined efforts of the EBU and SMPTE have brought us SDTI (Serial Digital Transport Interface), a new method of streaming data over a standard SDI infrastructure. SDTI brings the added benefit of faster than real-time transfer of compressed video data in its native format, thus avoiding the need for multiple codecs. Because SDTI is point to point it offers guaranteed Quality Of Service (QOS) provided the connection is maintained. The signal routing is most often set up manually.

In contrast, the computer industry has introduced us to file transfer techniques whereby data files in their native format are interchanged via standard computer networks. File transfer enables communication between multiple devices on a common network. Video file transfer is possible both faster than and slower than real-time but in either case the ability is there to deliver identical files with no degradation whatsoever provided the network is maintained. In this case routing is automatic, source and destination addresses being part of the information encapsulating the data. The Quality Of Service varies according to the type of network in use, the available bandwidth and the ability of the connected devices to sustain the transfer rate.

Whilst it is reasonable to expect both streaming and file transfer techniques to co-exist in the broadcast chain for some time, the use of file transfer is currently on the increase and nowhere more evident than in the post-production community. The greatest problem facing users and manufacturers alike is the sheer variety of file formats and network technologies. The fact remains however that the driving forces behind such development resides not in the professional video industry, but in the field of computing and networking.

3.0 Networking video, audio and metadata

The requirement exists for a high performance network capable of transporting both compressed and uncompressed video at many resolutions, together with its associated audio information and metadata at rates comparable with today's point-to-point SDI interface, or faster, but with all of the benefits of multi-point networking.

The ability to search for information can also come with the adoption of computer technologies. This attribute is key to the so called openness of networks and should operate in a familiar manner even though video file sizes are large.

The solution lies in the selection and integration of four key layers of technology to provide the highest level of performance, interoperability; network infrastructure, transfer protocol, application and file format.

3.1 Network

Crucially, the chosen network must satisfy several criteria. It must provide at least the same, or better, transfer rate provided by current 601 SDI connections. It is essential that the chosen technology be entirely open, subject to industry scrutiny and standardisation. Given that there are already a variety of available networks, many with their own particular benefits for certain applications, it is vital that the chosen technology should provide a high level of interoperability with other popular networks. Finally, it is essential that the chosen technology should enjoy a high level of popular support in the computer industry, thereby ensuring the highest level of availability and therefore enjoy a continuing downward price trend.

Five network technologies were considered for Clipnet.

- ATM (Asynchronous Transmission Mode)
- IEEE – 1394 (Firewire)
- Serial HiPPI (High Performance Parallel Interface)
- FibreChannel
- Gigabit Ethernet

Firewire is well specified but unsuitable for LAN and wider professional applications. Firewire data transfer rates are on the increase but the physical interfaces and somewhat closed architecture eliminated it from the discussion.

FibreChannel technology had already gained a good level of recognition in the industry and therefore became the first candidate for study., With experience however came doubts. FibreChannel is undoubtedly well suited as a means of interconnecting storage components such as fast disks or tape drives, but was never envisaged as a true network and this quickly became apparent when researching various hardware and software products on the market. Perhaps of greatest concern was that, despite the relatively small selection of available products, the vast majority relied on manufacturer-specific components and platforms. Several flavours of Fiber Channel exist therefore interoperability may be compromised.

Attention turned to Gigabit Ethernet. Although not a standard at the time of initial evaluation, the technology was already beginning to enjoy a high level of interest from several of the most influential players in the computing and network community. This led to standardisation in July 1998 (IEEE 802.3z) followed by widespread support and a fast growth in available products. Indeed a recent computer press survey revealed no less than twelve different Gigabit Ethernet switches. Experience thus far indicates a high level of interoperability with other popular networks notably ATM and 100Base T (Fast Ethernet) and already the market has witnessed a keen competitive pricing trend.

While the network itself is the core technology, the bulk of the integration work was found to be concentrated in the three other layers.

3.2 Transfer protocol

The choice of transfer protocol for use over an open (public) network interface is easier, the ubiquitous TCP/IP being the front runner. While TCP/IP brings with it a huge following of willing users, it also carries a large overhead – particularly when considering the high performance network solution required in this case. The major problem faced was the ability to react to the individual IP packets within the specified time - just 12 microseconds in the case of Gigabit Ethernet. In this short interval the destination device must respond to the interrupt, interrogate the packet to determine its validity for that particular device, verify the check-sum, respond if appropriate, route the data to the appropriate application before preparing to respond to the next data packet.

Alternative solutions included specialised forms of the protocol such as FTP+ and XTP as well as dedicated hardware accelerators, but all impacted on interoperability to a greater or lesser degree. This problem has to be solved ultimately by the computer industry and the widespread support enjoyed by Gigabit Ethernet provided the solution. Second generation high speed chip-sets from Alteon Inc. were found to provide all that was required (respond to the interrupt, interrogate the packet to determine its validity for that particular device, verify the check-sum, respond if appropriate, route the data to the appropriate application) - and more. Indeed the manufacturer has reported transfer rates as high as 960 Mbps, as close to the maximum theoretical limit of the network itself as anyone is likely to get!

3.3 Application

The choice of application (access method) was largely driven by our customers. Experience with industry-standard FTP for our stills network (Picturenet Plus) was good and it provided the impetus to avoid any form of proprietary access method, even at the expense of performance. The vast majority of our customers demanded standard NFS (Network File System) as their preferred method of access. Being by far the most open method, the sheer simplicity and elegance of mounting a file system on Henry™ or Editbox™ and the ability to 'drag and drop' files to it somehow outweighed any minor disadvantages of this virtually ubiquitous technology.

3.4 File formats

The choice of file formats to be supported initially has been driven from both the professional video and the computer graphics end of the market. Given the strong desire to maintain an open interface to the Quantel world, it was decided to support the Targa file format for video data in the RGB (4:4:4) domain. Targa is widely understood in computer graphics circles and is supported by most leading applications. For video data in the YUV (4:2:2) domain we naturally maintain support for the VPB format which is both entirely open and inextricably linked with the Quantel product range. The VPB format is fully compliant with ITU-R 601 and is both colour space and resolution independent.

The selection of file format(s) for the transfer of audio and metadata remains a subject for debate. Quantel wholly endorses the work of the recent EBU/SMPTE task force (Harmonized standards for exchange of programme material as bitstreams) in the area of metadata. Indeed Quantel have had considerable experience in handling metadata. Each new frame or clip generated using any Quantel machine manufactured in the last seven years carries with it a unique identifier - much like the UMID (Unique Material ID) proposed by the joint EBU taskforce. In this regard Quantel are keen to follow steps toward agreed standards for metadata and actively encourage industry dialogue on this most important aspect.

4.0 Features & Benefits

The implementation of Clipnet within the Quantel platform is designed for simplicity of operation. Network transfers are initiated directly via the Quantel man/machine interface and run as a background task, leaving the operator free to concentrate on the primary application.

The network interface itself incorporates dedicated hardware capable of formatting files to suit the nature of the target platform. For example, an uncompressed video file from Editbox transferred to a Cachebox configured for DVCPRO-50 will be automatically compressed. Likewise for files containing material of different aspect ratio or video resolution. By automatically converting files according to their source and destination, Clipnet ensures a high level of resolution independence while minimising operator intervention, and so maximising productivity.

5.0 Conclusion

Both streaming and file transfer techniques have a vital role to play in the professional video industry. SDTI offers several advantages and will certainly be widely adopted - particularly for transport of compressed video data over existing SDI infrastructure. In addition, the use of file transfer is poised to expand considerably both in broadcast and the post-production community. Quantel are committed to the provision of a high performance open network interface for video, audio and metadata and are in the process of developing Clipnet™ based on a selection of industry standard technologies including Gigabit Ethernet, TCP/IP and NFS.

References

FibreChannel Association (http://www.fibrechannel.com)

Gigabit Ethernet Alliance (http://www.giagbit-ethernet.org)

EBU/SMPTE task force report - special supplement 1998 (http://www.ebu.ch)

Quantel R&D (http://www.quantel.com)

540 Mb/s Serial Digital Interface Pioneering New Standardisation Approach at SMPTE

Dan Turow,
Product Manager,
Gennum Corporation
Burlington, Ontario, Canada

ABSTRACT

Today's broadcast and professional video market is undergoing tremendous change. Pressures to adopt ATV/HDTV, improve video quality, reduce costs and shorten time to air are driving studios to adopt new technologies and business models. In response to these, and other pressures, a 540 Mb/s serial digital interface has been adopted to address several new applications and capabilities. With these new applications becoming increasingly important, the need for a 540 Mb/s SMPTE standard was identified. This standard is currently being developed at SMPTE utilising the new "horizontal standards" structure recommended by the recent SMPTE/EBU Task Force. This paper will explore the applications and benefits of a 540 Mb/s serial digital interface and outline how several applications can take advantage of the horizontal standards currently being written at SMPTE.

INTRODUCTION

The broadcast and professional video industry is undergoing dramatic change. Pressures to adopt ATV/HDTV, improve video quality, reduce costs and shorten time to air are driving studios to adopt new technologies and business models. No one technology is optimal for all applications and all business models. Technologies are being evaluated and selected based on their fit with the particular application being considered. A 540 Mb/s extension of SMPTE 259M has arisen as one of the key technologies useful in meeting the needs of several distinct applications and user groups. A 540 Mb/s serial digital interface;

a) Enables Optimised 525p(480p) Based Systems,

b) Enables Optimised 750p(720p) @ 24 Hz Based Systems,

c) Enables Optimised 625p (576p) Based Systems,

d) Enables Effective Integration of Broadcast and Computer Video Equipment,

e) Provides a Higher Speed Data Pipe for SDTI Applications, and

f) Eliminates the Need for Complicated Dual Link (2x 270Mb/s) Interconnect Schemes.

With a 540 Mb/s serial digital interface useful in such a wide variety of applications, it became clear to SMPTE that this new interface needed to be standardised. Following on the recommendation of the recent SMPTE/EBU Task Force, a "horizontal standards" structure was adopted for this 540 Mb/s serial digital interface. With this new structure, interoperability, scalability and extensibility is ensured for 540 Mb/s enabled equipment.

MEETING THE NEEDS OF BROADCAST & PROFESSIONAL VIDEO WITH 540 Mb/s

The adoption of a 540 Mb/s serial digital interface offers a wide variety of benefits to the broadcast and professional video marketplace.

A) Optimising 525p (480p) Based Systems

Within North America, recent FCC mandates have spurred large North American broadcasters to adopt ATV in some form or other. Some broadcasters have selected a 1080i based system. Other broadcasters have chosen a 720p based system and/or a 480p based system. Adoption of the 1080i or 720p formats, and the required 1.485 Gb/s infrastructure, will require significant capital expenditures in the future. Although these expenditures are undesirable, most large broadcast networks will be able to support the required investment.

As the need to address ATV migrates down to second and third tier broadcasters, the ability to choose between a wide range of formats and pieces of equipment will become increasingly important. Knowing that many of these second and third tier stations are struggling to remain profitable, it is clear that cost will become a very large portion of the equation. The option of working with a 480-line progressive scan video format has arisen as a very attractive way to optimise this trade-off. In addition to defining a system that enables low cost ATV video equipment, the 480p format provides a system that helps avoid some of the drawbacks associated with interlaced video signals. The 480p format provides better quality images in fast moving pictures and acts as a good "base format" from which to derive other interlaced formats with either higher or lower vertical resolution.

Mapping the 480p signal format onto a 540 Mb/s serial digital interface is a very simple process. With this mapping, 480p systems can take advantage of the increased performance and lower cost that a 540 Mb/s serial digital interface provides. Although the 480p signal format could also be mapped into a 1.485 Gb/s serial data stream, the costs associated with this make it unpalatable and sub-optimal for several reasons:

- Compatibility with Installed Infrastructure:

  In facilities today, coaxial cable is the main method of interconnecting equipment. It is typical to see run lengths in excess of 200m. With a 540 Mb/s serial interface, such cable length performance can easily be achieved. SMPTE 292M interfaces do not have such cable length capabilities and as a result would require re-wiring and/or re-designing existing facilities. The costs associated with this are enormous and should not be underestimated.

- Added Expense For HDTV Spigots

  Today, there is a large premium on an HDTV interface when compared to an SDTV interface. These costs are primarily driven by the fact that completely new design techniques are utilised when dealing with 1.485 Gb/s interface signals as defined in SMPTE 292M. In some cases, additional costs are introduced because designers may be forced to migrate to high cost, high power Fabry-Perot lasers used in HDTV fibre optic interfaces.

- Added Latency and Memory Requirements

  When investigating the mapping of 480p signals onto a SMPTE 292M HDTV serial interface, potential issues regarding latency arise. Although analysis is continuing, it appears that substantial amounts of memory (framestores?) may be required at each and every interface where the 480p mapping takes place. As a result, the cost associated with using a SMPTE 292M physical layer for 480p transport may be much higher than expected, making it even more inappropriate for cost sensitive 480p applications.

Today, a 540 Mb/s serial interface is the **same cost** as a 270 Mb/s serial interface. Since 540 Mb/s capabilities are embedded within chipsets that support all SMPTE 259M data rates (143 Mb/s, 177 Mb/s, 270 Mb/s, 360 Mb/s), the cost of a 540 Mb/s serial interface will follow a more attractive price curve.

Notwithstanding the above, it is clear that several facilities will pursue the installation of 1.485 Gb/s infrastructure for their 1080i or 720p based systems. In these cases, it does not make sense to invest in a completely separate infrastructure for the small amount of 480p material that may be used in the facility. In these circumstances, it is logical to re-use the 1.485 Gb/s infrastructure to carry 480p signals. Defining a mapping of 480p onto SMPTE 292M is clearly advantageous in this situation.

B) Optimising 750 (720p) @ 24 Hz Based Systems

In addition to enabling 480p based systems, a 540 Mb/s serial interface enables another low cost ATV format with even higher vertical resolution. Specifically, a 540 Mb/s capable serial digital interface can support low cost 720p @ 24 Hz based systems. This format may be interesting to facilities that have a heavy focus on film. To understand the importance of this format, one must be aware of the fact that television receivers perform 3:2 pull-down on 720p @ 24 Hz signals that they receive. This means that 720p @ 24 Hz signals would be presented at 60 Hz even though the distribution format was running at 24 Hz. Recognising and taking advantage of this means that 24 Hz signals could be broadcast with relatively low levels of compression. Lower levels of compression means higher quality video. At the same time, 720p @ 24 Hz can be mapped onto a lower cost, high performance 540 Mb/s serial interface. Once again, the benefits of operating with a 540 Mb/s serial interface cannot be ignored. It is a low cost, high performance and complementary alternative to SMPTE 292M that enables broadcasters and equipment manufacturers the flexibility to select the optimal technology for their particular application.

C) Optimising 625p (576p) Based Systems

Market dynamics within Europe are significantly different than those in North America. Right now, there are no plans within Europe to migrate to a SMPTE 292M-based HDTV system. However, there have been discussions within the EBU about potential applications for a 625-line progressive scan video format. The 625p (576p) video format has >80% of the vertical resolution of some HDTV formats widely being adopted within North America. As such, the 625p format provides a high quality, progressive scan video format for ATV in Europe. This format also provides an excellent base video format which would enable the distribution of high quality ATV European content into the North American HDTV broadcast environment. A quick analysis of the 625p format reveals that it is a simple process to map 625p data onto a 540 Mb/s serial digital interface. This mapping enables a low cost physical layer transport mechanism for the 625p signals that is compatible with the installed base of coax, has superior cable length performance and is easy to integrate into existing facilities.

D) Enabling Effective Integration of Broadcast and Computer Video Equipment

In addition to being excellent broadcast video formats, the 525p (480p) and 625p (576p) formats ensure that compatibility with computer based video equipment is maintained. Computers are inherently progressive scan. As a result, compatibility issues often arise when computer based video equipment is interfaced with interlaced video signals. With a 540 Mb/s serial interface, the 625p and 525p video formats can be simply and easily transported into and out of computer video equipment. Although these signals may be carried over a SMPTE 292M serial digital interface, the dynamics of the NLE market demand that interfaces be cost effective and easy to implement. A 540 Mb/s serial interface provides this ideal physical transport mechanism. As previously mentioned, a 540 Mb/s serial interface is the same cost as a "mainstream" 270 Mb/s serial interface since such capabilities are embedded with SMPTE 259M compatible chipsets. In addition to this, a 540 Mb/s serial interface avoids the introduction of unnecessary cost, latency and data buffering that may be required with mapping 480p signals onto SMPTE 292M serial interfaces. Finally, a lower speed and more robust 540 Mb/s serial interface is more conducive to integration and operation in the high density, high noise environments typically generated in PCI based computer video boards.

E) Providing a Higher Speed Data Pipe for SDTI

Although the current SMPTE 305M SDTI standard currently operates only for 270 Mb/s and 360 Mb/s systems, there is clearly a trend towards migrating SDTI to higher speed applications. Reducing file transfer times increases productivity, optimises equipment utilisation and reduces time to air. A 540 Mb/s serial interface, in conjunction with a natural extension of SMPTE 305M, significantly increases the speed at which compressed or uncompressed file transfers can be accomplished. With a 540 Mb/s serial interface, over 400 Mb/s of SDTI payload can be transferred. This capability proves extremely interesting for transporting very lightly compressed ("Mezzanine Level") HDTV signals. For DV based compression, a 540 Mb/s serial digital interface enables 8x transfers of 25 Mb/s, 4x transfers of 50 Mb/s or 2x transfers of 100 DV data.

Data Rate	SDTI Payload	Application
270 Mb/s	200 Mb/s	4x/2x @ 25/50 Mb/s
360 Mb/s	270 Mb/s	6x/3x @ 25/50 Mb/s
540 Mb/s	400 Mb/s	8x/4x @ 25/50 Mb/s

F) Eliminating the Need for Complicated Dual Link (2 x 270Mb/s) Interconnect Schemes

Less related to ATV is the application of 540 Mb/s to bandwidth intensive SDTV applications. These include, amongst others, the transmission and reception of 4:4:4:4 signals typically used in post-production and high-end compositing/effects applications. To date, transmission and reception of these signals has required dual link systems (2 x 270 Mb/s) for basic connectivity. Such systems comply with SMPTE RP175. With a 540 Mb/s serial interface, connection of this equipment can be done much more effectively utilising SMPTE RP174. SMPTE RP174 defines a single link variation of SMTPE RP175. With this single link interconnection scheme running at 540 Mb/s, unnecessary cabling cost and complexity can be avoided. Additional cost savings can be realised by eliminating the need to "autophase" two incoming serial data streams and by eliminating the need to have two chipsets on a board each and every time a signal is transmitted or received.

In addition to enabling low cost Single Link 4:4:4:4 applications, a 540 Mb/s serial interface enables a cost effective way of doing 2x real time transfers of standard 270 Mb/s signals. In an environment where time to market is key, the increased productivity offered by faster than real time data transfers over a lower cost and higher speed 540 Mb/s serial interface is difficult to pass-up. A simple twist on this application would also enable the transmission and reception of two independent 270 Mb/s video over one wire. Such capabilities can prove extremely useful in a variety of applications including inter-studio or inter-facility links where cabling issues can be significant hurdles to overcome.

ADOPTING A NEW "HORIZONTAL STANDARD" STRUCTURE AT SMPTE

The recent SMPTE/EBU Taskforce has identified the need to transition from what may be called "vertical standards" to "horizontal standards". A "vertical standard" describes within a single document all of the features necessary to make a particular application operable. Vertical standards make it possible to use one source to know everything necessary to interconnect a system. The main drawback of using vertical standards relates to the fact that vertical standards do not permit the flexibility to mix and match solutions to meet the needs of specific applications. Scalability and extensibility are severely hampered when utilising vertical standards[1].

Horizontal standards are inherently designed to stack on top of each other so that a suite of horizontal standards are used to satisfy the needs of a particular application. Horizontal standards may be mixed and matched to optimise any one particular application. System scalability and extensibility is enhanced in systems that use horizontal standards[1].

When deciding on how to standardise the 540 Mb/s serial digital interface, it became clear that the large number of applications for 540 Mb/s made the use of vertical standards completely inappropriate. Only by using horizontal standards could the number of existing and future 540 Mb/s applications be accommodated.

After considering the trade-offs of various approaches it was decided that the introduction of "mapping documents" was appropriate. These mapping documents would essentially map specific application information onto another SMPTE standard that strictly defines the 540 Mb/s physical layer interface. This approach is analogous to the OSI model used in the data communications.

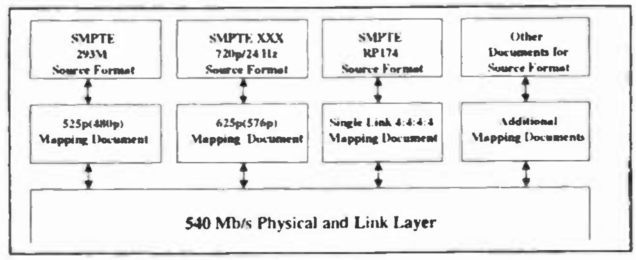


Figure 1 - New Physical Layer and Mapping Document structure.

In this structure, the Source Format document specifies key parameters like the frame rate, the sampling structure, the colorimetry and other items necessary to define the base video format.

The mapping documents then define how to generate an appropriate 54 MHz multiplex that can be serialised in accordance with the physical and link layer document. These mapping documents reference key source format parameters like the frame rate, the number of lines per frame, the number of samples per line and legal word values when defining how to properly create an appropriate 54 MHz data multiplex.

Finally, the physical and link layer parameters are specified with the lowest layer document. Parameters like signal data rates, signal amplitudes, rise times, fall times and other key parameters like return loss are specified in the document. In addition to this, specific requirements for synchronisation characters (TRS sequences) are included in the physical and link layer document. Without including these synchronisation words in the physical and link layer document, a 540 Mb/s SDI receiver would be unable to properly synchronise and word align a parallel representation of the incoming serial data.

With this structure, it is easy to see that any particular source format may be mapped to a number of different physical and link layers. This provides the flexibility to select the optimal physical layer transport on an application by application basis. Historically, vertical standards have forced end users to select only the physical transport mechanism specified in the original vertical standard. Having the ability to select a lower cost, higher performance 540 Mb/s serial digital interface for the 480p video format is a perfect example of the benefits derived from this new structure.

In addition to this, new applications for the physical and link layer can quickly and easily be added by simply defining a new mapping document. For example, the available bandwidth in an SDTI system can easily be expanded by simply defining a mapping document on the 540 Mb/s physical and link layer document.

As the number of applications for 540 Mb/s grow, additional mapping documents will be defined. These new capabilities will be easily integrated into 540 Mb/s capable equipment and enable the interoperability, scalability and extensibility long sought by both end users and equipment manufacturers.

CONCLUSIONS

In response to several key market needs, a 540 Mb/s serial digital interface has been adopted as a piece of technology that is key to the future of broadcast and professional video. A number of applications will benefit from the capabilities of a cost effective and high performance 540 Mb/s serial digital interface. These include 525p(480p), 625p(576p), 750p(720p) based systems as well Single Link 4:4:4:4 and high bandwidth SDTI applications. The "horizontal standard" approach taken in standardising the 540 Mb/s serial interface will enable interoperability, scalability and extensibility in equipment designed to support these new capabilities and standards.

REFERENCES:

1. Weiss, Merrill, Turning Television Standards on Their Heads, DV Business, November 1998.

ASSET MANAGEMENT ACROSS YOUR FACILITY
A CASE STUDY OF THE LUCASFILM DIGITAL MEDIA MANAGEMENT INFRASTRUCTURE

Timothy Campos and Beth Rogozinski
Silicon Graphics
Mountain View, California

Contributors:

Paul Stevens, Technical Lead for StudioCentral, Silicon Graphics
Ken Fischer, Technical Lead for StudioCentral, Silicon Graphics
Tony Behen, Director of Operations, Information Technologies, Lucasfilm Ltd.
Charlie Ledogar, Lead Software Engineer, Lucasfilm Ltd.

ABSTRACT

This paper documents the continuing implementation of an asset management system at Lucasfilm. The Lucasfilm environment is an integration of heterogeneous systems and applications moving towards a common media management framework. The solutions put in place include media servers from Silicon Graphics, database servers from Oracle, workstations, and RAID arrays tied together underneath a media management operating environment, StudioCentral. In examination of this environment, we will investigate the imperatives that spurred this company to move to a digital asset management system (DAMS) and the issues that arose while implementing the digital asset management environment – execution, technology, training and support, as well as the solutions.

Overview

The need for broadcast and media production facilities to incorporate asset management systems has increased rapidly in the past several years. By tracking versions of media through the creation process, accessing archived media, and moving media through the production process in a more efficient way, asset management systems cut costs, increase productivity, and forge new potential revenue streams via media asset resale. To be effective, an asset management environment must work with all existing computer and visualization technology platforms. It must be able to seamlessly manage all media files and formats. It must be flexible enough to use desktop systems, servers, RAID storage systems, tape robot archives, and all other currently employed and to-be-implemented storage systems. It must be able to access, file, and archive into and out of legacy file systems. And, it must work on the fastest available networking protocols.

Needs Analysis of DAMS at Lucasfilm

As Lucasfilm began to embark on their Star Wars prequel series of films, they determined that they needed to devise a more efficient mechanism for tracking their media assets. They desired a system that could store, catalog, sort, and retrieve all of the digitally created assets as well as analog film assets and props. This unified digital system would allow various internal departments at Lucasfilm to access, manipulate, and distribute their assets more readily. With this system, time and money would be saved in the reuse of digital assets on related media productions. In addition, the media management system would allow them to streamline the approval process for licensing and allow their catalog of assets to be browsed directly by clients via a CD-ROM based "catalog," or potentially a secure connection into their asset repository.

The digital asset management system needed to be flexible, scalable, customizable and robust. The specific needs of each of multiple identified groups of primary users needed to be met while the DAMS remained flexible enough to scale to future needs of Lucasfilm. Additionally, the system needed to integrate into existing multi-platform servers and workstations and be accessible via a cross-platform interface.

At the initial phases of researching such a system, a number of digital asset management solutions existed from various software and hardware vendors. However, none matched closely enough to their requirements. The option of building a DAMS from scratch was considered, but was rejected because this approach is not the primary business focus of Lucasfilm. A proposal was made in the fall of 1997 to build upon and customize the StudioCentral asset management environment from Silicon Graphics. It was determined that the StudioCentral API provided much of the total solution required and the remainder could be created in house.

An important aspect of the StudioCentral environment was its openness and flexibility. No commercially available media management product provided an out-of-box solution that could be tailored to the workflow and data management needs of Lucasfilm. A customized StudioCentral environment could be matched to the asset management needs of Lucasfilm and their internal groups. This approach minimizes direct capital costs and development expenses. The project to implement this customized digital asset management system was given the code name "Chess."

Implementation in Phased Approach

The Chess project was originally conceived to be a large system that would be able to manage all the digital assets from the new Star Wars prequel series. The original goal was for Chess to manage all existing assets, input assets directly from the post-production and art departments' workflow, and access assets across all the Lucas Companies (which include: LucasArts Entertainment, Lucasfilm, Lucas Licensing, Skywalker Sound, Industrial Light and Magic, and Lucas Learning). Lucasfilm's IT department was to lead the effort.

A common pitfall when implementing a DAMS is to attempt to completely overhaul an existing infrastructure and integrate all processes with the DAMS. The main reasons that overhaul approaches typically don't work can be broken down into three key issues: budget and schedule constraints, user buy-in and workflow issues, and technology ramp-up issues. The IT team was aware of these concerns, and understood the multi-platform, multi-application, and diverse user group environment that the system would need to accommodate. After careful evaluation of the goals, expectations, technology, and human implementation demands such a project would entail, they decided to abridge their system to one that could provide return on investment in a realistic and responsive timeframe, but that laid a flexible back-end foundation for extensions to the system.

System Users

The users (both present and future) of the Chess DAMS at Lucasfilm represent several diverse groups, each with different requirements: the Art Department, the Production Department, the Library, the Archives, Licensing, Marketing, the Internet Group, and Finance.

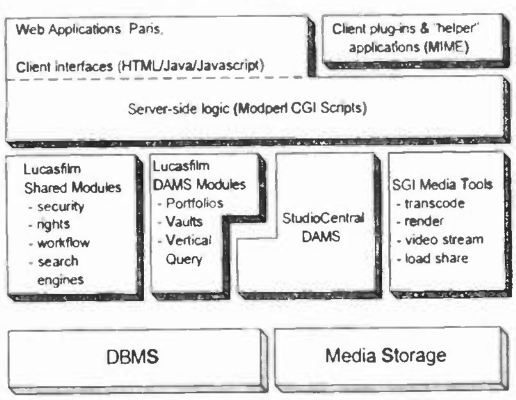
The general requirements for the DAMS represent the overlapping needs of these principal user groups. These include: seamless transition of media through various workflows, performance, simplified cataloging, customizable metadata fields, powerful search and replace features, managing collections of assets, multiple file format support, multiple resolution support, transcoding of assets, revision/version control, storage management, rights management, support for wide area networks (WANs), and off-line browsing/catalog capabilities.

In addition to these general requirements, many of the user groups had system requirements that were specific to their area of expertise. This was especially true of the metadata that users wanted to associate with the media stored in the DAMS. The metadata requirements of each department were different, and the IT department strove to encourage buy-in by providing a system in which each department could define its own set of metadata. Another major challenge was to understand and formalize the paper- and telephone-based media workflows that exist within and between each of the departments.

The Lucasfilm IT Infrastructure

To meet the needs of such a diverse user group, the IT team required a development environment that could rapidly evolve to meet new requirements from each of the many teams that would use the DAMS. Furthermore, Lucasfilm is a heterogeneous environment that includes Silicon Graphics, Macintosh, Windows 95, Windows NT, Sun, and LINUX client platforms. This led the IT team to implement their client applications using an intranet architecture that standardized on Netscape/JavaScript clients and an Apache server running on an Origin 2000 using modPerl CGI components. The components include modules for workflow, database information retrieval, and user access control. The StudioCentral DAMS environment is integrated into this system via StudioCentral's Perl interface. The layered architecture that Lucasfilm devised for all their applications enabled rapid application development (RAD), minimized deployment cost, and provided a cross-platform client solution with minimal development cost.

Lucasfilm Chess Architecture



Distributed Architecture

Lucasfilm's goal for their DAMS is an infrastructure that provides access to media regardless of its geographic location or the type of system it is stored in. One of the advantages that Lucasfilm gained using StudioCentral is the distributed nature of the product. A StudioCentral Repository is made up of a database system and one or more content storage systems. Distributed content is managed using StudioCentral content servers that abstract the interfaces to the underlying storage systems. The database is used to persistently store metadata information that describes the media, including its location as well as user-defined metadata. The StudioCentral software provides the integration between the database system and the content servers and maintains the integrity between the two. It also provides a set of services that facilitates management of the media under StudioCentral control.

When building Chess, Lucasfilm chose to configure StudioCentral to use their existing DBMS, allowing them to fully leverage their investment in that infrastructure. Rather than building interfaces to all of the various systems that would house content for Lucasfilm, the abstracted content management architecture allowed Lucasfilm to focus on interfacing StudioCentral into Chess, and later integrate with other products such as video streaming systems. This implementation also allows Lucasfilm to consider implementations where distributed content is managed over a wide area network while providing a centralized point of access for locating media.

Datamodel Design and Implementation

Chess leveraged the metadata capabilities provided by StudioCentral, which provide a framework for defining dynamically extensible metadata attributes using entities known as "datamodels." A datamodel defines a set of metadata attributes: names, types (integer, string, vector, structure, and so forth), and optionally their allowed and default values. Subject to certain constraints, the datamodels can be extended, removed, or modified after they have been loaded.

StudioCentral internally maps the datamodel attributes into database tables in the underlying DBMS. If not specified by the datamodel designer, StudioCentral provides default mappings of datamodel types into type-specific tables. However, it is also possible to perform a "direct mapping" whereby the attributes are mapped into tables created independently by the system designer. When the metadata for an asset is read or written via the StudioCentral tools and APIs, the datamodel tables are located, and all the metadata associated with that asset is retrieved or stored without having to understand how it was mapped in the underlying database.

Using StudioCentral's metadata framework in Chess did not initially require customization other than the definition of the datamodels. However, Lucasfilm's integration of StudioCentral was not trouble-free. The fact that StudioCentral retrieves all the metadata from various database tables whenever the metadata is accessed was soon recognized as a performance bottleneck. In some instances, only a few attributes need to be read or written to implement a given user interface, but StudioCentral provided only an all-or-nothing API for reading or storing attributes. Fortunately, the ability to define a direct mapping of these attributes to specific database tables that were defined by Lucasfilm provided a solution. Once their application became aware of how these tables would be directly mapped, it could access the tables directly when performance was a concern.

Media File Typing, Transcoding, and Versioning

Lucasfilm required that their DAMS be capable of dealing with at least 26 various file formats, including formats of still images, video and film files, audio files, and text and other documents. The Chess system had to be able to transcode files from one format to another so that they might be used across the facility without users having to launch software transcoding applications and manually change the file of each asset prior to use. In addition, the system had to be able to produce and associate various resolutions of media. Lastly, the system had to be able to support the tracking of changes to an asset's media and metadata so that an audit trail of modifications could be maintained.

The Chess system, using the StudioCentral digital asset management environment, supports all types of digital assets. The StudioCentral framework provides fundamental support for integrating format conversion or transcoding software, creation of assets with multiple resolutions, as well as the capability to logically group together all formats and resolutions of the same asset. The development environment also lends itself well to integrating format conversion software through its plug-in architecture that allows the Chess DAMS to manage any file type. As versioning is a native feature of StudioCentral, tracking of multiple versions of the same asset in Chess was relatively simple. The Chess DAMS also supports automated creation of thumbnails and low-resolution proxy images via the IRIX Digital Media Libraries.

Storage Management

Though the cost of hard drive storage continues to drop, on-line storage capacity continues to be a major issue for digital media systems. The IT department is evaluating the use of near-line or off-line tape archives managed by a hierarchical storage manager (HSM). The HSM will interface with the Chess DAMS via StudioCentral's ability to manage media simultaneously in a variety of different types of storage locations.

Another future need for Chess is the ability to manage media stored in separate distributed systems. StudioCentral's ability to associate multiple distributed content stores will provide them with the flexibility they need to grow their DAMS to fit their distributed environment. This will allow Lucasfilm to share media with remote subsidiaries while maintaining centralized control.

User Rights Management

Lucasfilm's IT organization supports many different applications in addition to their DAMS applications, including a request tracking application, accounting systems, and licensee management. Lucasfilm desired centralized user rights management across all these applications, and determined that the permissions structure that is built into StudioCentral was not appropriate. Therefore, the IT team implemented a custom security and rights management module at the application layer.

Flexible Dynamic Configuration in Distributed Environment

During the development process, Lucasfilm made good use of StudioCentral's distributed architecture, in which the content server, database, repository server, and client application are separate objects that communicate via CORBA services. By changing configuration files, Chess clients could quickly switch to different media repositories. This allowed Lucasfilm to separate their production system from development and demonstration systems in a logical fashion, and simplified changes to the system environment such as porting, mirroring, and upgrades. More importantly, it provides scalability for future demands on the system, particularly WAN-based media management.

PARIS: THE FIRST APPLICATION OF CHESS

The IT department chose to build their first Chess application for the licensing, library, and marketing departments. This application, named Paris, is the first of several planned applications that will build on the Chess environment. The Paris application was designed to replace the laborious task of managing the approval process of releasing still images of sets, props, movie frames, and so forth, to external parties. Prior to Paris, this was achieved using a standalone mini-database and paper forms. Images were stored on CD-ROM systems, and could not be retrieved over a network, making accurate audit trails impossible. The IT department looked to resolve these issues using the Chess DAMS environment.

Developing Paris with End-User Feedback

Lucasfilm, as with any facility that would be installing a new system infrastructure, had a central requirement that the system not have a negative impact on current workflow practices or the performance of users and groups during transition and full implementation phases. This is a difficult requirement, because any new system or tool put in place requires training and ramp-up time. But, by allowing the user groups to have influence in the design of the system, they had an understanding of and an enthusiasm for what the system would provide. In this way, the IT group was able to make the user groups more receptive to the transition process. Furthermore, the RAD nature of Chess allowed the IT group to incorporate additional feedback from the user groups during the test process quickly, which made them more accepting of new tools and have a smaller learning curve in mastering the new tools.

Automated Cataloging and Customizable Metadata

Paris is built to automatically capture simple asset information such as file format and file type. In addition, Paris provides a cataloging interface that allows the user to add additional metadata in a structured fashion. While developing this cataloging interface, the IT department was faced with a challenge. StudioCentral's ability to dynamically modify the datamodel definitions throughout the prototyping and development process provided great benefit. However, the open-ended nature of the metadata infrastructure soon became a daunting disadvantage; there were just too many requirements coming from end users. Lucasfilm found they had to limit themselves initially to not get bogged down in endless iterations on possible datamodels. In a valuable lesson for future developments, the IT team chose to use a single unstructured datamodel and implemented a process of soliciting but limiting end user input. As Chess becomes more central to the Lucasfilm environment, and as users become more familiar with the system, a structured approach to metadata will be more easily implemented.

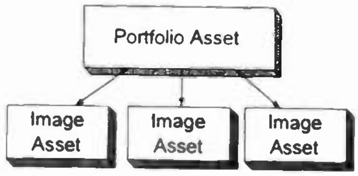
Managing Content Storage and Access

Paris centralizes storage for Lucasfilm's stills using StudioCentral's managed filesystems. The Chess architecture allows users to browse this content on their desktop over the network using a web client. Most of the images are stored in PhotoCD format which allows the users to select the resolution that they'd like to see, from postage size thumbnails to full resolution 6144x4096 images. The cost savings achieved by this are substantial as users no longer need to physically go to a Librarian to request access to the PhotoCD files in order to fulfill requests from external parties.

Workflow in Paris

Equal in benefit to centralized content management in Paris will be integrated workflow. Users deal with groups of images through a folder-like metaphor called a portfolio. An example of a typical workflow event is the licensing approval process. A requestor would search and place one or more images within the portfolio and submit it for approval. Users within the Marketing department would automatically "see" portfolios waiting for their review and could approve or reject individual images as well as add comments to justify their actions. After review, portfolios would automatically be routed back to the submitter..

Portfolios are implemented as assets within the StudioCentral environment. An asset is the basic building block for media objects within StudioCentral. Portfolios are assets that contain only references to other assets instead of directly containing the content. The state of a portfolio is managed by metadata stored in a StudioCentral "portfolio" datamodel. As a user reviews a portfolio, Paris determines the proper values for the attributes of the datamodel and sets them in StudioCentral when the review is complete.



When StudioCentral assets are created, they can be instructed to track their revision history. Since the portfolios are created as versioned objects within StudioCentral, the revision history acts as an audit trail that can be reviewed to see what elements of a portfolio were approved and rejected, by whom, when and why. A downside of the StudioCentral versioning implementation, however, is that even minor changes to portfolios create new versions within the system. To eliminate clutter in the version tree for portfolios, Lucasfilm can use StudioCentral's pruning feature to remove older versions of portfolios that do not contain important change history.

Present and Future of Chess

Chess's architectural approach to digital asset management has been well received throughout the organization. While Paris was designed to manage still images, the flexibility of the Chess architecture and StudioCentral have allowed the Lucasfilm IT department to easily extend the Paris interface to demonstrate management of MPEG and MJPEG video content. This includes support for transcoding video files, managing them via vaults and portfolios, generating thumbnails, and splicing together simple video sequences, and viewing content by streaming it out of the StudioCentral content store via one of two video-streaming applications.

By proving asset management solutions through Paris, Lucasfilm's IT department is confident on implementing and deploying future production DAMS environments. This success has raised awareness with other groups within the company that a DAMS system can be used to streamline workflow and provide valuable cost savings. Lucasfilm has already started to design a new application for the Art and Production departments that will also use the Chess infrastructure.

CONCLUSION

Just as relational databases revolutionized information management in the 80s and early 90s, digital asset management systems provide opportunities for substantial cost savings, and enable new revenue streams for companies whose business is media. While there are a number of excellent DAMS available on the market today, and these will continue to improve to meet the varying needs of broadcast and media production facilities, many who are beginning to implement asset management plans are looking to have a system that they can mold and manipulate themselves. The Chess DAMS in place at Lucasfilm is an excellent example of asset management that was designed to meet established and evolving needs. For a relatively small amount of work, StudioCentral allowed the IT team at Lucasfilm to develop an open solution that meets their specific needs. The Chess system in use at Lucasfilm is not yet in its final iteration, and this system will continue to grow and change. Chess has been a productive addition to the work environment and is expected to dramatically increase asset value, increase user productivity, and decrease time to market of future productions and products.

In the increasingly competitive world of media production and distribution, managing assets, saving time and money, and delivering products more quickly to market will demarcate successful organizations. The key element to achieve all these ends will be a flourishing, flexible, maintainable and scalable digital media asset management system.

ACKNOWLEDGMENTS

The authors of this paper would like to express their sincere gratitude to Tony Behen, Charlie Ledogar and the IT staff of Lucasfilm and Paul Stevens, Ken Fischer, and Ardene Kannely of Silicon Graphics for their assistance and contributions to this document.

OLD EARTH STATION IFL CABLES AND NEW DTV DOWNLINK PERFORMANCE

Peter C. Zilliox
Andrew Corporation
Richardson, Texas USA 75082

INTRODUCTION

Existing earth stations, used to receiving network TV program material via satellite, were designed to receive analog FM modulated RF signals. This paper presents an analysis of the suitability of the interfacility link (IFL) cabling used in these installations between the antenna and the indoor equipment for use in distributing the new broadcast TV network digitally modulated RF signals. These IFL subsystems usually transported broadband L-Band analog FM TV signals from the Low Noise Block Converter (LNB) to either active or passive power splitter assemblies for distribution to multiple receivers. Therefore, the old designs, the implementation techniques employed at the time, and the cable aging processes associated with these IFL cable runs may result in an unacceptable level of performance degradation when passing digitally modulated signals. This paper addresses the impact of the potential signal quality (E_b/N_o) degradation that these cables may have, and specifically, it examines the effects of the IFL cable's termination mismatches. Poorly terminated coaxial transmission lines may produce echo distortion of significant enough level to distort the digital signal and cause an increase in the downlink's bit error rate (BER).

This paper presents an detailed analysis of the transmission channel model. The model suggests certain performance figures of merit needed for acceptable performance. A recommended procedure is offered to characterize the performance of an existing L-Band downlink chain. Recommended L-Band IFL designs for various requirements and/or distances are suggested. Finally, a summary of the findings of this paper are presented.

TRANSMISSION LINE MISMATCHES

Transmission line mismatches within the L-Band coaxial cable distribution network produce echo distortion that distorts the downlink signal. This type of distortion degrades the reception performance of both analog FM and phase-modulated digital transmissions. However, certain forms of digital modulation may be more severely degraded, depending on the data rate and modulation type.

The magnitude of the amplitude and group delay distortion produced in the IFL by the echo mechanism is primarily a function of the return loss (VSWR) at each interface point, the loss of the cable, and the cable's electrical length. For example, the VSWR presented at each cable connector and every active or passive component's input and output port determines the system's overall performance. In most simple terms, reflections, due to imperfect VSWR at the interfaces, generate a series of reflected waveforms that echo through the transmission line. Numerous such discontinuities exist for the signal as it makes its way from the LNB to the Integrated Receiver Decoder (IRD). Signal reflections bounce between discontinuities and produce interfering signals (or echos) that are delayed from the main signal. In an actual downlink chain, there are numerous discontinuities and multiple echo paths. The resulting composite echo signal is a complex waveform consisting of a summation of all of the individual echo pairs. The interaction of this composite echo waveform with the desired waveform produces a degradation to the detection and decoding processes of the network signal. The echo composite signal produces group delay (with the resulting phase nonlinearity) and amplitude distortion in the earth station's downlink chain.

GROUP DELAY DUE TO ECHO DISTORTION

The group delay produced by an echo between a single pair of discontinuities on an ideal transmission line can be characterized by the following idealized approximation (assuming a perfect transmission line with a constant VSWR versus frequency at each discontinuity):

$$\tau_g(\omega) = 2\alpha_p^2\tau_p\rho_1\rho_2\cos\left(2\tau_p\omega\right) \qquad \textbf{(Eqn 1)}$$

Where: $\tau_g(\omega)$ is the group delay function of frequency
- α_p is the one-way transmission coefficient (loss factor)
- τ_p is the one-way propagation time between discontinuities
- ρ_1 is the reflection coefficient of the first discontinuity
- ρ_2 is the reflection coefficient of the second discontinuity

$$\tau_p = \left|\frac{L}{CV_f}\right| \qquad \textbf{(Eqn 2)}$$

Where: L is the physical length of the transmission line
- C is the speed of light
- V_f is the velocity factor for the transmission line

The non-idealized (real world) group delay function is more complex than depicted in Equation 1, but this model will serve to help us calculate some expected values for group delay distortion for given typical component VSWRs for various downlink cable configurations.

Equation 1 is based on an intuitive deduction of how the signal echo levels (due to reflection coefficients, ρ_1 and ρ_2) combine with the desired signal at various signal frequencies as a function of the total round trip propagation time between a pair of transmission line discontinuities. Field and laboratory group delay measurement data has good correlation with group delay values predicted by Equation 1.

PHASE NONLINEARITY DUE TO GROUP DELAY

By definition, the group delay function is the derivative (rate of change of the phase versus frequency) of the phase transfer function with respect to the radian frequency:

$$\tau_g(\omega) = \frac{\partial\phi(\omega)}{\partial\omega} \qquad \textbf{(Eqn 3)}$$

Or conversely, the phase transfer function is the integral of the group delay function as follows:

$$\phi(\omega) = \int_\omega \tau_g(\omega)\partial\omega \qquad \textbf{(Eqn 4)}$$

The phase nonlinearity (variation from linear phase) is the important component in determining how the digital signal will be degraded due the echo distortion. The deviation in phase from a linear response (over the channel bandwidth) is the phase nonlinearity. Digital satellite modem manufacturers sometimes specify the maximum Root Mean Square (RMS) value for the transmission channel's phase nonlinearity when guaranteeing operational bit error rate (BER) performance for their modem through the channel. Alternately, modem manufacturers specify the peak-to-peak (P-P) variation in the channel's group delay performance when guaranteeing operational BER performance. Either or both can be specified, but channel group delay limits are specified more often.

The phase nonlinearity can be expressed by substituting an expression for group delay variation into Equation 4, as follows:

$$\phi(\omega) = \int_{\omega} \frac{-\partial[\Delta\phi(\omega)]}{\partial\omega}\partial\omega = \int_{\omega}^{\omega+\Delta\omega} \Delta\tau_g(\omega)\,\partial$$

(Eqn 5)

In Equation 5, we see that the magnitude of the phase nonlinearity is a function of the rate of change of the group delay over the channel's bandwidth. Therefore, a channel that exhibits a slowly varying group delay function versus frequency (such as a parabolic or linear component across the channel) produces the largest phase nonlinearity, as seen when the group delay function is integrated over the channel bandwidth. A group delay function that rapidly varies throughout the channel bandwidth may integrate to produce less of an effect on the phase nonlinearity and therefore contributes to less signal degradation (some laboratory work confirms this; however, this work and analysis is incomplete at this time).

TYPICAL EARTH STATION EXAMPLE

Consider an IFL cable installation found in a typical TV network receive earth station. This cable would likely consist of a 150-foot run of 75 ohm 1/2" diameter low loss coaxial cable running between the LNB output and the input of a L-Band splitter/divider with an 8-foot jumper cable on each end. The typical jumper cable is likely to be a 75 ohm 1/4" diameter cable. The VSWR interface characteristics for the LNB and splitter are both typically 2.5:1. The group delay expected, due only to this simple cable run, can be estimated using the equations presented above.

The echo magnitude on the IFL is proportional to the product of the reflection coefficients (ρ) associated with the VSWR at each end of the transmission line. Referring to Equation 1, the magnitude of the group delay on the long cable, in this example, is as follows:

$|\text{GD Level}|_{\text{worst possible case}}$ =
$2\tau_p\, \alpha_p\, \alpha_p \rho_1 \rho_2$ =
27.48 ηsec or, 54.97 ηsec P-P

Where:

ρ_1 (reflection coefficient) =
0.43, due to a typical LNC output VSWR = 2.5:1

ρ_2 (reflection coefficient) =
0.43, due to a typical input VSWR of the 4-way signal divider = 2.5:1

α_p =
0.66 (3.65 dB), the 1-way propagation transmission line loss coefficient for LDF4-75A 1/2" HELIAX® cable run for 150 ft between the LNC and the 4-way divider

τ_p =
173.18 ηsec, the 1-way propagation delay time for a LDF4-75A 1/2" HELIAX cable run for 150 ft

The actual expected group delay performance is between 1/3 to 1/2 of that calculated using the above method. In actual practice, the reflection coefficients are not all at their worst-case values at every frequency. Therefore, the group delay will be well below 54 ηsec P-P calculated. The expected group delay in this actual system will be typically 18 ηsec P-P.

In this idealized model (i.e., a perfect transmission line with only two discontinuities having constant VSWR versus frequency), the group delay variation function can be mathematically expressed in the following manner:

$$\Delta\tau_g(\omega) = A\cos(2\tau_p\omega) \qquad \textbf{(Eqn 6)}$$

Therefore, from Equation 5, the phase nonlinearity is:

$$\Delta\phi(\omega) = \int_{\omega} \Delta\tau_g(\omega)\, d\omega$$

$$\Delta\phi(\omega) = \int_{\omega} A \cos(2\tau_p \omega) d\omega \qquad \textbf{(Eqn 7)}$$
$$= \frac{A}{2\tau_p} \sin(2\tau_p \omega)$$

Substituting values into Equation 7, the magnitude for the phase nonlinearity is:

$$|\Delta\phi(\omega)| = \frac{A}{2\tau_p}$$

$$= \frac{9.07x10^{-9}}{2\,(173.2x10^{-9})}$$

$$= 0.0262 \ \mathit{Peak\ Radians}$$

The RMS value of the phase nonlinearity is:

$$\Delta\phi(\omega)_{RMS} = \frac{1}{\sqrt{2}} (0.0262)$$

$$= 0.0185 \ \mathit{Radians}$$

Understanding how the calculated RMS phase nonlinearity value, the amplitude distortion, or the P-P level of the group delay distortion may contribute to intersymbol interference, reducing the Euclidean distance in the detection process, or other negative factors affecting the downlink's BER performance, is beyond the scope of this paper. But some guidelines can be established for amplitude and group delay distortion limits based on a set limit for a tolerable BER performance degradation.

DISTORTION SPECIFICATION LIMITS

In order to limit the signal impairment in the earth station downlink chain, the following channel amplitude versus frequency performance is recommended from LNB input to demodulator input:

Gain vs Frequency Response
$\leq$ 0.5 dB P-P over the channel bandwidth

In a "soon to be published" article about the degradation of digital satellite signals due to group delay, the researchers (Steve Back, Globecomm Systems, Inc., and Mark Weigel, EFData) report the results of data collected on modems for various levels of channel group delay. They report that with Nyquist filtering in accordance with DVB, a modem performance degradation of 0.5 dB or less is realized when the total SR x GD (Symbol Rate times Group Delay Product) for the entire link— including uplink, transponder and downlink— does not exceed the values presented in the Table below over the symbol rate bandwidth:

End-to-End Group Delay Channel Recommendations

Form of Digital Modulation	(Symbol Rate)*(Group Delay)
QPSK	< 0.62
8PSK	< 0.58
16QAM	<0.41

It is necessary to keep track how much group delay is apportioned to the uplink, the satellite, and the downlink earth station. When the amount of group delay is excessive, then corrective action is necessary. An amplitude/group delay equalizer can predistort the uplink signal to compensate for linear and parabolic amplitude and group delay components encountered in the uplink earth station and the satellite input and output filter/diplexers. However, amplitude and group delay ripple components can only be treated by better terminating the transmission lines in the system at the VSWR sources and loads.

If we consider designing the downlink system to operate 8PSK with a transmission symbol rate of 30 Msym/sec (a full 36 MHz transponder signal), then the end-to-end group delay is recommended to be:

Total Group Delay Response
$\leq$ $(0.58)*(30 \times 10^6)^{-1}$ ηsec P-P, over the channel bandwidth
$\leq$ 19.3 ηsec P-P, over the channel bandwidth

Because the uplink and the satellite are likely to produce similar parabolic group delay, it is not reasonable to root sum square the allocations for distortion among the three contributors. Therefore, it may be prudent to suggest an equal sharing of allocation. Thus, the earth station downlink group delay might be best specified within 7 ηsec P-P over the transponder bandwidth for this example.

RECOMMENDED L-BAND IFL DESIGNS

L-Band hardware, designed specifically for digital TV distribution, is universally designed using a 75 ohm characteristic impedance. LNB output ports, coaxial cables, line driver amplifiers, signal splitters, bias tees, polarization/routing switches, and integrated receiver decoders (IRDs) all operate at a 75 ohm impedance and thus require a source and/or termination with a 75 ohm impedance. Without proper source and termination impedance matching, the system generates higher levels of echo distortion, as described earlier. Therefore, it is important that the L-Band IFL cable run be presented a good match at each end of the link. This does not mean that the IFL must operate with a characteristic 75 ohm impedance. On the contrary, it may consist of any type of transmission line, as long as the IFL presents a matched impedance at its interfaces. As we will see, it is desirable to use 75 ohm transmission lines in some applications and 50 ohm lines in other applications with an appropriate impedance transformation to 75 ohms at the IFL interfaces.

IFL Runs 200 Feet or Less

When the required L-Band IFL cable distance is 200 feet or less, the installation of a 1/4" small diameter, low VSWR, 75 ohm cable is recommended. The 1/4" diameter cable exhibits a relatively high attenuation factor (loss/foot); therefore, it presents a significant attenuation to the echo signal's round trip path. The high attenuation reduces the echo signal level to a point where it has little detrimental effect on the desired digital signal[1]. Additionally, to ensure a low VSWR at the interfaces, the 1/4" 75 ohm cables should be connectorized with quality 75 ohm type F male connectors. Type N 75 ohm connectors should be avoided because their small, fragile pins fall prey to being broken by the larger pins used on the common 50 ohm Type N connectors found on test cables and test equipment! If it is necessary to interface to BNC, use quality 75 ohm impedance BNC connectors. Figure 1 illustrates in block form the IFL configuration and the graph presents the expected IFL performance versus frequency. Note the G/T degradation and gain slope illustrated in Figure 1 are within the criteria that was set as part of the IFL design requirement.

IFL Runs Between 200 Feet and 350 Feet

If the IFL cable length must be extended beyond 200 feet, a 1/4" IFL link becomes excessively lossy. When this happens, the system G/T is

---

[1] Estimated Group Delay Ripple to be less than 6 ηsec P-P

degraded[2], the downlink's "gain slope[3]" becomes excessive, and the absolute signal level is lowered. Therefore, a lower loss (larger diameter) cable is required for runs exceeding 200 feet. Andrew Corporation recommends a low VSWR, low loss 50 ohm coaxial transmission line (such as HELIAX 1/2" diameter LDF4-50A cable) for IFL cable runs between 200 feet and 350 feet in length. An Andrew Transformer/Pad (ATP) should be placed at each end of the cable run. This device is a simple, passive, minimum loss (5.7 dB) matching pad. It provides two very important functions. First, it transforms the 50 ohm transmission line impedance to the 75 ohm source and/or load impedance. Second, it introduces an additional 11.4 dB return loss to reflections resulting from very poor source or load VSWR interfaces. Figure 2 shows the schematic representation of the ATP. Note that the ATP has an internal dc shunt choke that allows LNB supply voltage (if present) to pass through the ATP without being subjected to the minimum loss pad resistors. The ATP's Type N male 50 ohm connector terminates directly to the IFL cable's mating female Type N (Figure 3). This mating connector configuration provides the necessary low VSWR interface at each end of the long 50 ohm IFL cable. However, the ATP's Type F precision 75 ohm female connector is terminated with the aid of a short 1/4" 75 ohm jumper cable to the downlink chain hardware (LNB or IF splitter, etc). Figure 3 graphically illustrates the expected G/T performance of this length IFL design and presents a block diagram of the recommended configuration.

IFL Runs Between 350 Feet and 700 Feet

Andrew recommends an IFL configuration as presented in Figure 4 for IFL runs exceeding 350 feet but less than 700 feet. The recommended configuration consists of 7/8" cable, 50/75 ohm transitions, and 1/4" jumper cables with precision type F connectors.

Summary of Findings

Existing IFL installations should be examined and their amplitude and group delay characteristics should be characterized. Cable impedances should be matched with their source and termination impedances as closely as possible in order minimize the amplitude and phase distortions that impair the digital reception process. Generally, if the group delay can be limited to under 10 ηsec P-P, and the amplitude variation can be limited to 0.5 dB P-P over a 36 MHz channel, the resulting impairment to digital signal reception should be within acceptable boundaries.

If it is determined that new IFL cable installations are necessary, the recommended IFL cable designs are as follows:

1. When IFL runs are under 200 feet, use 1/4", 75 ohm cable (Andrew HELIAX FSJ1-75A cable with Type F 75 ohm connectors).
2. For IFL cable lengths exceeding 200 feet and but less than 350 feet, convert to 50 ohm 1/2" transmission line (Andrew HELIAX LDF4-50 with L4NF Type N female connectors).
3. For IFL runs up to 700 feet, use 7/8" Andrew HELIAX LDF5-50.
4. Any IFL run exceeding 600–700 feet, use a low cost L-Band fiber optic system.

HELIAX is a registered trademark of Andrew Corporation.

[2] System G/T degradation less than 0.3 dB for 60 dB gain LNB/LNAs

[3] Gain Slope should be no more than 0.25dB/36 MHz

Figure 1 Recommendation for IFL Cable runs under 200 feet

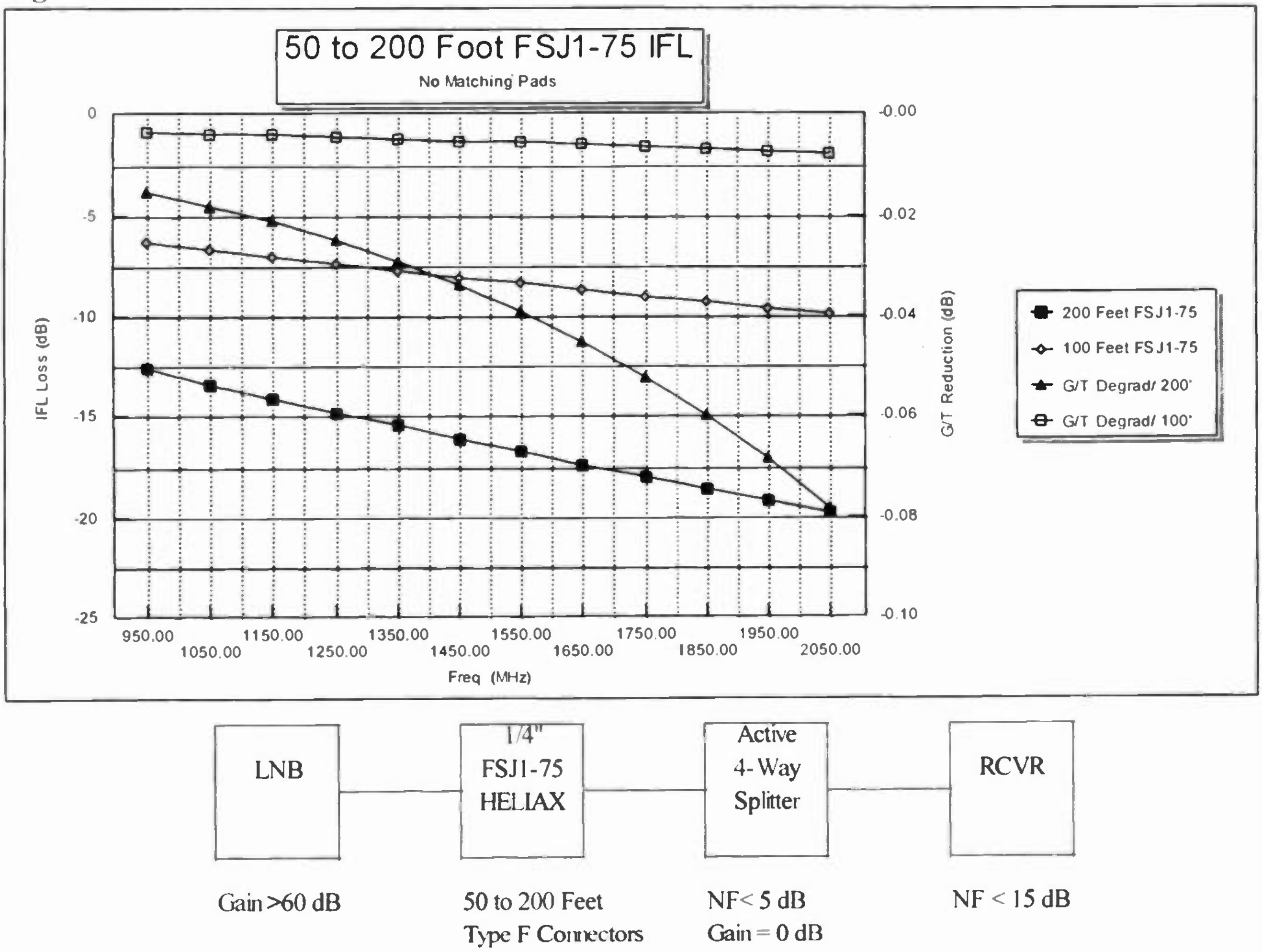


Figure 2
Andrew ATP Transformer/Pad Schematic Diagram

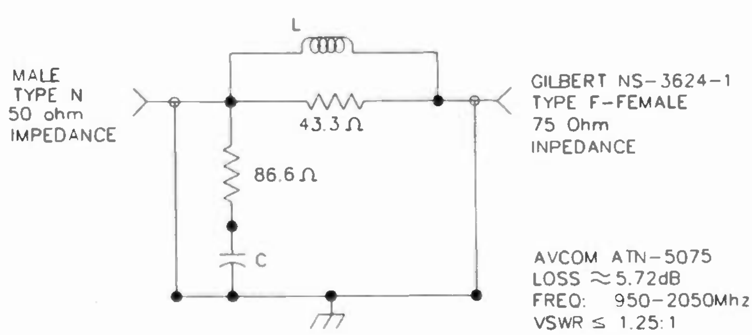
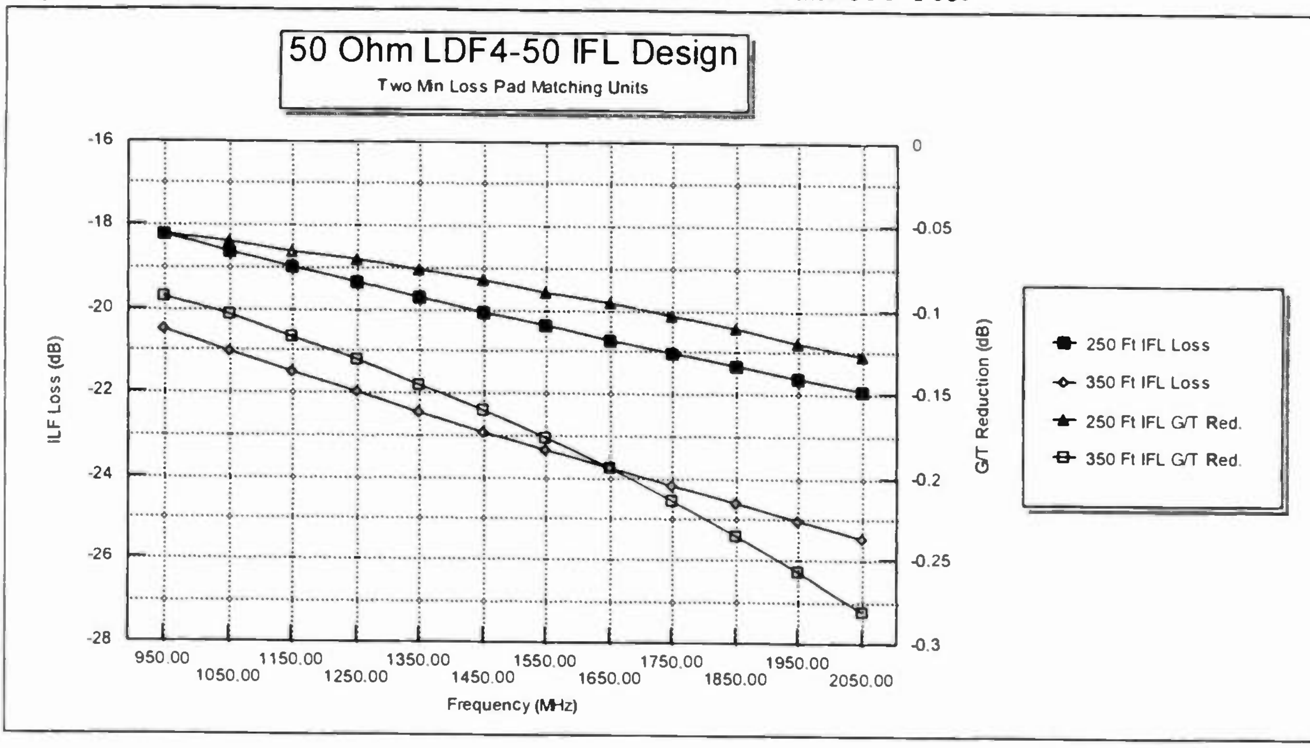


Figure 3 Recommendation for IFL Cable runs between 200 and 350 Feet



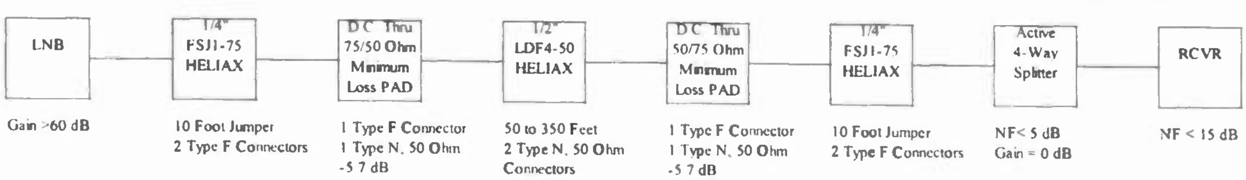
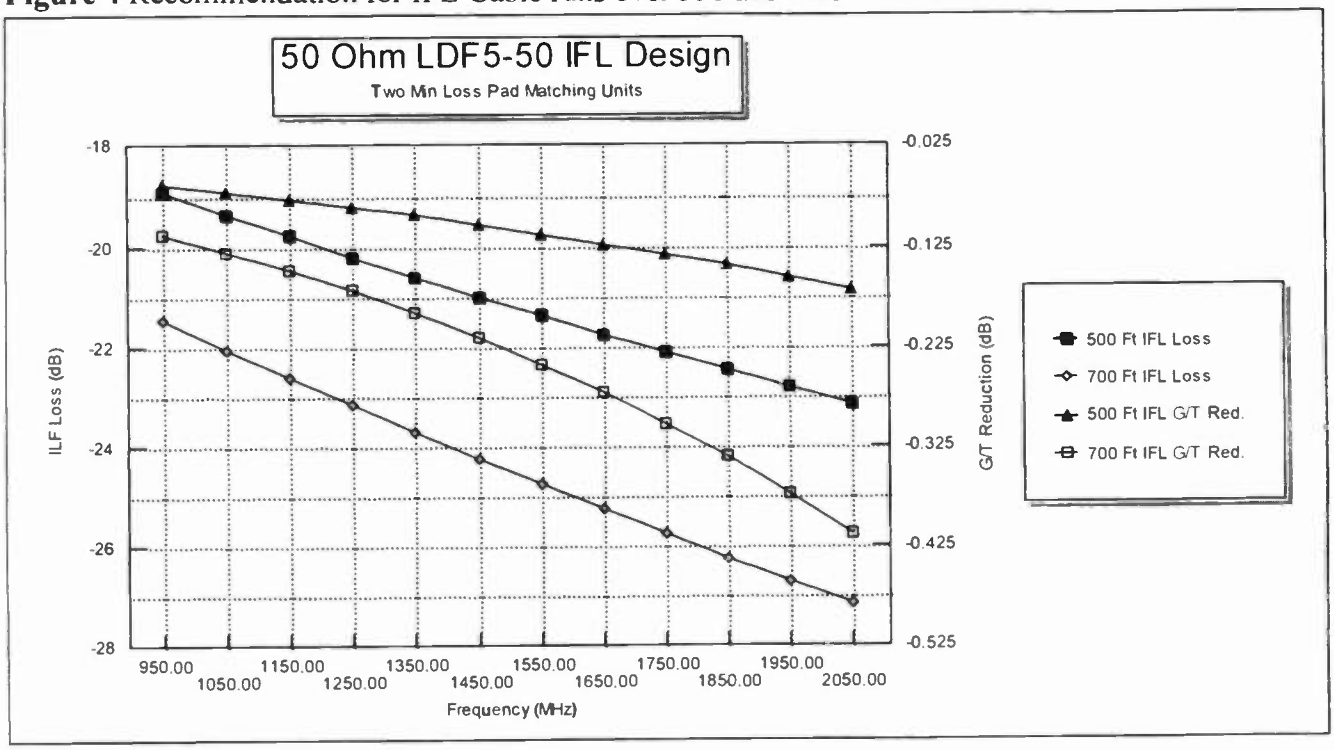
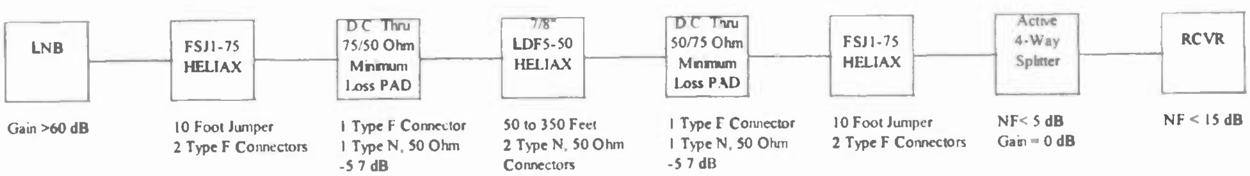


Figure 4 Recommendation for IFL Cable runs over 350 and below 700 feet





Data Broadcasting: Where Is the Money?

Monday, April 19, 1999

10:30 am - 12:00 pm

Chairperson: Robert Hess
WBZ -TV, Boston, MA

10:30 am Understanding ATSC Datacasting – A Driver for Digital TV
Giri Venkat
Philips Semiconductors, Inc.
Sunnyvale, CA

***11:00 am Digital TV Delivers New Opportunities for Broadcasters, Advertisers and Consumers**
Abe Peled
NDS Ltd.
West Drayton, United Kingdom

***11:30 am Broadcasters as Datacasters**
Alan Rosenberg
StarBurst Software
Concord, MA

*Papers not available at the time of publication

UNDERSTANDING ATSC DATACASTING-A DRIVER FOR DIGITAL TELEVISION

By Giri Venkat
Philips Semiconductors, Digital Television Product Group
Sunnyvale, CA

Abstract
In this paper, I will show how the ATSC (Advanced Television Systems Committee) digital television standard capability to support data content (datacasting) may be a more important factor in its adoption by consumers than the digital standard's stunning high definition video and multi-channel audio. This paper will focus primarily on Enhanced Television applications, datacasting content that is synchronized with video program content.

Introduction
Because of the excitement about wide-screen, high definition video and multi-channel audio, the significance of the new ATSC broadcast standard's ability to simultaneously transmit huge amounts of digital data has been overlooked by many broadcasters. One reason is that broadcasters often assumed that viewers would have to attach a PC or similar device to their ATSC receiver in order to use the data content. When in fact, at least six leading consumer electronics manufacturers are planning intelligent ATSC DTV receivers capable of processing datacasting content.

Another reason is that many broadcasters are uncertain how to use this new capability. One important application is for advertising. A combination of new technologies including new digital recording and time shifting technologies may make datacasting vital to support advertising revenue. Recently announced digital recording devices (Replay TV and TiVo are two examples) allow consumers to easily record programs for later viewing, or to simultaneously record and replay a program shifting the viewing time up to 45 minutes. The result is that viewers will easily be able to skip conventional television advertising. Fortunately a new generation of intelligent receivers, combined with innovative uses of datacasting and seamless integration with the Web or Internet-like services offer the potential for new advertising and revenue generating services.

Types of Datacasting
The industry has defined two major categories of datacasting:

1. Enhanced Television – data content related to and synchronized with the video program content. For example, a viewer watching a home improvement program might be able to push a button on the remote to find more information about the product being used or where to buy it. A viewer watching a cooking program might be able to obtain a recipe with the push of a button, even order the ingredients from a local market. A news broadcast could include transcripts and related news information that a viewer could display, save, or print on demand.

2. Data Broadcast – data services not related to program content. An example would be current traffic conditions, stock market activity, or even subscription services that utilize ATSC conditional access capabilities. Data broadcast bandwidth could also be sold to third parties for a wide range of applications not related to broadcast television.

The distinction between Enhanced Television datacasting and Data Broadcast could be a significant issue pertaining to how the FCC applies "must carry" rules to local cable operators. The FCC could consider Enhance Television datacasting part of the program content and force the cable operators to include it in their signal, but not require the cable operator to carry data that is not directly linked to the program content.

Enhanced Television Examples
First, let's review a range of applications that are being considered and the ATSC receiver standards that will support them. These examples

demonstrate how ATSC datacasting could be used to enhance program content based on the receiver and transmission standards that the ATSC is currently developing, and the capabilities of intelligent ATSC receivers (both complete television sets, and set-top converter boxes) that leading consumer electronics companies are developing.

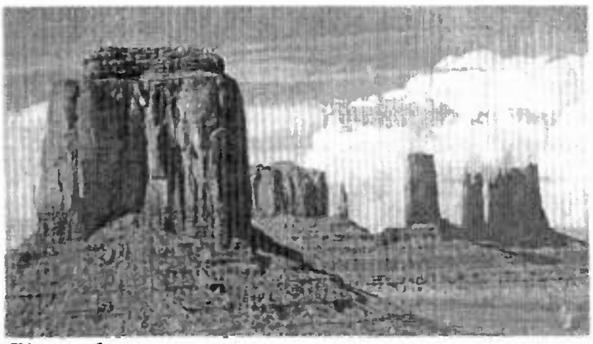


Figure 1.
Icon Indicating Data Content

Our viewer is watching a travel program. When there is datacasting content related to the program, an icon could appear on the screen in an unobtrusive manor (Figure 1). If interested in seeing more information, the viewer could push a button on the remote.

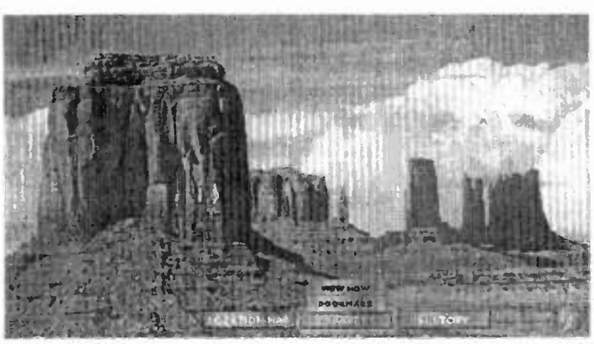


Figure 2.

The next level of information, also unobtrusive, could identify the types of content available (Figure 2).

If the topics look interesting, the viewer could select the category —let's say "travel", then choose whether to view the information immediately or book mark it and retrieve it later. Bookmarking would instruct the receiver to store the data in local memory for later use.

If our viewer decided to look at the datacasting content immediately, the moving video image could shrink, revealing the additional menus and levels of information (Figure 3). Again our viewer could bookmark, view on screen, or even print this information.

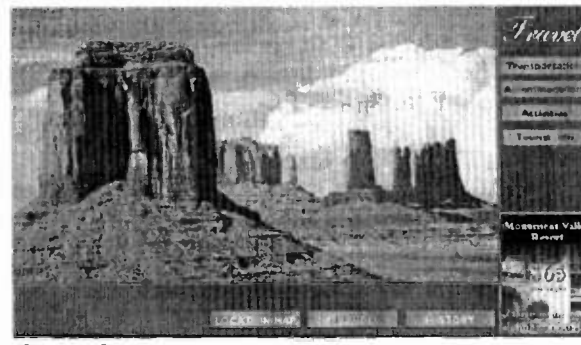


Figure 3.

Another feature being developed by many receiver manufacturers is a built-in reverse communications channel (such as a modem) and an Internet browser. The ATSC Digital TV Application Software Environment (DASE) standard, the specification for digital receivers to handle datacasting, will likely support links to related Web sites allowing receivers with Web browsers to automatically connect to a Web site to support interactive applications and ordering goods and services.

In this example, if there's a reverse communications channel, our viewer could request additional information, generate an automatic e-mail or transparently link to an Internet site to make a reservation at the hotel advertised on screen.

Keeping the Viewer's Attention

Producers, broadcasters, and most importantly advertisers have voiced concerns about distracting a viewer with data content during the program, and the possibility of moving a viewer's attention away from the television program to the Internet. Advertisers are particularly concerned about loosing the viewer's attention. Several of the major networks, the Corporation for Public Broadcasting, advertising agencies, and consumer products manufacturers are conducting focus groups and user tests to determine effective techniques for using datacasting content to enhance program content and advertising. Among the techniques being developed:

Intermission – Specific times in the program when the viewer is given the opportunity to chose from a menu of possible options. Menu items might include music or news clips related to the program content. The viewer might even be able to buy CDs of the music and other products.

Product Placements – At any time during a program, a viewer might be able to obtain more information about anything in the program. Perhaps the viewer wants to know more about the BMW that the hero is driving in a drama or the jewelry the heroine is wearing. Or, the viewer may be interested in the furniture, the art, or location. At any time a viewer could request additional information about anything included in the program. The value to advertisers of product placements would increase significantly if a viewer could obtain more information about the product and where to purchase it.

Figure 4.

Multilevel Advertising – With datacasting, the audio/video content could be designed to appeal to the appropriate market segment. In this example, bicycling enthusiasts. But the single video stream could support multiple advertising messages. In this example (Figure 4.) the advertisement could support biking events, helmets, clothing, and bicycles as well as where-to-buy information.

There's enough data-carrying capacity in the ATSC standard to download information to support multiple menus. Of course the receiver would have to be equipped with enough memory or storage to save the various options. The proposed DASE standard will likely have a provision to carousel data, that is broadcast the same data multiple times in a short period of time, making it available to support a viewer's selection within seconds even if the receiver doesn't have sufficient memory to store all menu options. There's also enough broadcast data capacity to send pictures and multimedia content – video and audio – and use new techniques such as immersive photography and JAVA applets.

Datacasting in the News

Figure 5.

Datacasting can also enhance news programs in many ways. For this example (Figure 5.), the video content is in a 4:3 aspect ratio allowing the news program to be simultaneously broadcast for NTSC and ATSC viewers. But the ATSC viewers would see additional menus providing immediate access to information such as transcripts, still images, and related stories. A digital viewer could also access local information such as weather, traffic, and more.

DASE –Digital television Application Software Environment

The ATSC DASE specialist group defining the software application environment for digital receivers is moving toward a specification that addresses the following issues:

1. **Open Architecture** – Receiver manufacturers want independence from any particular vendor of hardware or software subsystems, freeing them from the PC industry model where a small number of companies dictate product specifications.
2. **JAVA** – The standard will likely support JAVA.
3. **Wide Range of Datacasting Services** – Receiver manufacturers want to be able to offer products that support different levels of datacasting and browsing features. The core functions can be supported with a memory footprint of 500 kb (or less), making it possible to design low-cost receivers that can still process and display some datacasting content. A basic receiver with limited memory might be able to process and display unidirectional content, while a more sophisticated receiver might include more memory and storage plus support bi-

directional communications with a complete Internet browser.

4. **Use of Existing WEB Authoring Tools** – The standard will be based on HTML (HyperText Mark-up Language), the programming language used for the Internet. Consequently ATSC datacasting services will be able to make use of the pool of experienced authors already creating content for the Web, as well as support reuse of existing Web content.
5. **Use of Existing Web Transaction and Order Fulfillment Services** – The proposed standard will allow ATSC datacasting services to use existing services developed to support the Web.
6. **Web Links** – The datacasting content will be able to automatically link to a Web site if the receiver is equipped with an Internet browser. In addition, the standard will allow the broadcaster or advertiser to limit a viewer to predefined sites from that connection, minimizing the chance of losing the viewer to the Internet. The DASE standard would require to viewer to return to the program content.
7. **Synchronize Data to Program Content** – The datacasting standard will provide techniques to synchronize data content to specific segments of a video stream and provide precise layout control for the data content to co-exist with the video image.
8. **Extensible** – The standard will support new media types by using content decoders that are extensible with downloadable software code.

Summary

ATSC datacasting may be a more important factor in driving consumers to acquire digital receivers than high definition video and multi-channel audio. Effective use of datacasting could have far reaching effects on advertising and commercial broadcasters' business models. A new generation of intelligent ATSC receivers with built in Internet browsers and reverse communications channels will integrate Internet services with broadcast television. Broadcasters and producers have an exciting new dimension to add to television.

Receiving DTV Broadcasts

Monday, April 19, 1999
10:30 am - 12:00 pm

Chairperson: Jerry Butler
PBS/Public Broadcasting Service, Alexandria, VA

10:30 am The Evolution of Front Ends for Digital TV
Simon Wegerif
Philips Semiconductors, Inc.
Sunnyvale, CA

***11:00 am Open Cable: Establishing a Road Map for Tomorrow's Network Platform**
William Wall
Scientific-Atlanta Inc
Norcross, GA

11:30 am Television Home Server for Digital HDTV Broadcasting
Tatsuya Kurioka
NHK Science & Technical Research Labs
Tokyo, Japan

---

*Papers not available at the time of publication

THE EVOLUTION OF FRONT ENDS FOR DIGITAL TV

Simon Wegerif
Philips Semiconductors
Sunnyvale, CA, USA

ABSTRACT

This paper describes the evolution of hybrid analog digital front ends for receiving ATSC (VSB) and existing NTSC signals from the perspectives of component count (increasing integration & cost reduction), flexibility (decoding platform independence) and performance.

INTRODUCTION

After a decade of standard setting, digital TV has moved into the implementation phase, and progress, both at the transmission and receiver side has already been rapid. For the front end of the digital TV receiver, it is both digital and mixed signal IC technology which has allowed substantial increases in performance, flexibility and a reduction in parts count – all of which are necessary if DTV is to hit mass market consumer price points and levels of reliability in the near future. Some of the major drivers & opportunities for these trends will be explored in the following sections.

COMPONENT COUNT

A typical first generation VSB front end employs separate IC based modules for the tuner, IF, digitization (A/D) and demodulator / decoder. Much of this functionality is replicated in order to acquire existing NTSC cable or off air signals.

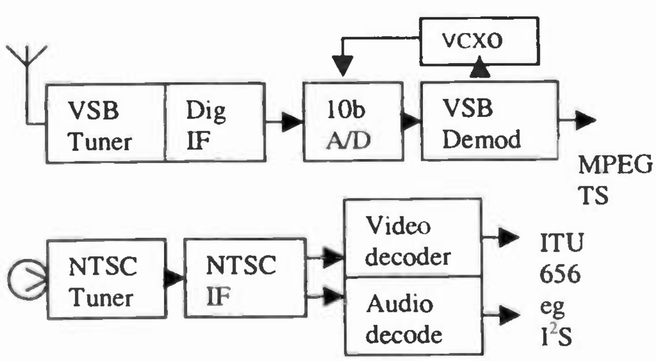


Fig. 1. Typical first generation DTV front end

An analysis of the requirements of different receiver types identifies some common needs, which can be implemented using a very few ICs. The core architecture of NTSC and ATSC decoding can be extended in a modular fashion to encompass other modulation schemes in a cost-effective way.

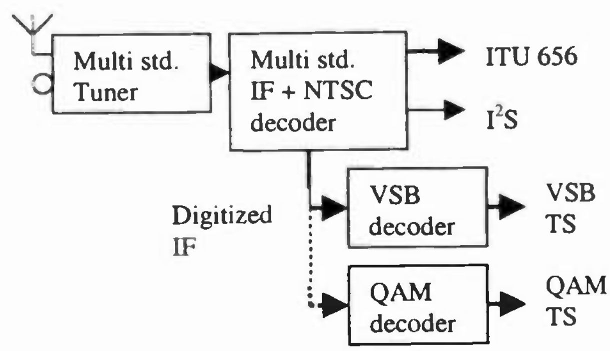


Fig. 2. Second generation hybrid DTV front end

Here, a multi standard tuner, capable of tuning to digital VSB, cable (QAM) and NTSC signals throughout the 50-860 MHz band is connected to a new mixed signal IC[i]. This IC combines the functions of NTSC IF, picture and sound demodulation with VSB / QAM down conversion. The NTSC signals are digitized and output as ITU-R 656 and Inter IC Sound ready for processing in a digital receiver architecture. Analog component (YUV, S-Video) inputs are also provided, allowing the IC to act as a source selection and format conversion center for the many legacy inputs required on a digital hybrid receiver.

When VSB or QAM signals are being received, the IC outputs a digitized representation of the IF, for glueless interfacing with VSB & QAM demodulators. The VSB demodulator can also act as a Transport Stream switcher, to eliminate glue logic where more than one digital demodulator is required to interface with a source decoder having only a single input. Other component reduction measures include the provision of a sample rate converter at the input of the VSB demodulator, removing the need for an external VCXO regulating the synchronous sampling of the IF input.

PERFORMANCE

The first generation design has been optimized with respect to certain performance parameters. It has been tested extensively in the Lab (including the facilities at the Advanced Television Technology Center in Arlington, VA), and also in the field at several different locations, using both indoor & outdoor antennas. Feedback from these tests is incorporated into the second generation, without adversely affecting the achievable price point.

First generation receivers have typically been designed to perform similarly to the Grand Alliance reference system (known affectionately as the 'Blue rack'). During 1998 a number of new transmitters have come on line, allowing testing in reception environments that were not included in the original Grand Alliance program. Some of the most notable impairment mechanisms resulting are dealt with in the following sections.

Short delay (dynamic) multipath

This is perhaps the most severe problem that VSB has to cope with. It is caused principally by two phenomena:

1. Urban canyons with high rise walls. Modern steel & (metallized) glass buildings can reflect incident radio waves in the VHF/UHF range with very high efficiency. The overall effect when trying to receive signals in an urban environment is an 'urban clutter' effect, characterized by many close together high energy, short duration echoes.

2. The use of indoor antennas. Indoor reception conditions represent an additional challenge, as often there may be no line of sight reception, the signal instead having to travel via doors & windows which do not absorb or reflect the signal. Additional reflections can be caused by people / other moving objects in the vicinity. Indoor antennas are typically not very directional, so most reflected / delayed energy has to be rejected using intelligent techniques in the VSB decoder.

Some parameters in the receiver have to be optimized to cope with dynamic multipath:

1. The Decision Feedback Equalizer (DFE) has to be updated rapidly to cope with echoes whose strength and delay is continually changing. The training sequence built in to VSB transmissions repeats every 24ms – a rate which is too slow to compensate for indoor reception changes e.g. people moving around the living room, or vehicles moving outside. An extra blind update path has been added which can update the equalizer at VSB symbol rate continuously, in addition to the trained path. The use of an equalizer with the fewest number of taps needed also helps improve the inherent tradeoff between equalizer introduced white noise and adaptation time. In addition to the measures deployed in the equalizer, a four stage AGC circuit allows a fast response to changes in signal level, whilst preserving headroom & resolution at each stage in the chain from antenna to transport stream output.

2. Synchronizer enhancement in the demodulator front end. No matter how good the equalizer and its update algorithms, the VSB decoder cannot operate without reliable front end lock. A complex equalizer placed before the main adaptive equalizer decreases the apparent Inter Symbol Interference (ISI) by compensating for certain types of e.g. group delay related channel distortions. This circuit aids in the recovery of the ATSC synchronization symbols, yielding a robust lock in the presence of interference.

Long static ghosts

Evidence has appeared that certain man made & natural geographic features can cause long static echoes at particular reception sites. Examples of this occur at a particular site in Washington DC where a discrete echo of 37us has been measured and several sites in Chicago where strong 26us echoes have been found, in this case caused by signal reflections from the City's three tall skyscrapers.

In order to cope with field test findings, and with possibly even more difficult conditions in the future, a DFE architecture has been implemented which can cope with discrete echoes of up to 80us duration. The architecture is chosen so as not to compromise unduly the desirable adaptation time, equalizer introduced white noise and chip area / power dissipation characteristics of a shorter equalizer.

NTSC adjacent & co-channel interference

It has been known for some time that, due to the FCC channel allocations for DTV having to occupy the same spectrum as that for analog during the transition period, some problems are likely with strong NTSC signals causing interference with their (typically 12dB) less powerful DTV counterparts on the same, or adjacent channels.

The Grand Alliance decoder used a single tap comb filter generating regularly spaced notches and placed before the equalizer in the VSB receiver to remove energy around the NTSC picture, chroma sub carrier and sound carrier frequencies. The problem with this relatively simple implementation is that it degrades the basic S/N performance of the receiver by approximately 3dB when switched in.

Much improved performance can be obtained by using a multiple tap, adaptive FIR filter placed between the output of the equalizer and the input of the trellis decoder. The NTSC co-channel interference filter operates in a trained mode, using the field sync segment synchronization signal as a training sequence. Unlike the Grand Alliance scheme, this filter is left permanently in circuit and adaptively tracks & removes interference with little or no loss of AWGN performance.

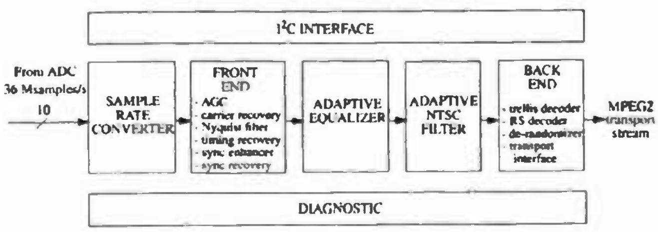


Fig. 3. Second generation VSB demodulator[ii]

PLATFORM

It is hoped and expected that users of TVs, STBs and PCs will all benefit from DTV transmissions, however these devices have distinctly different architectures. Interfacing, together with other hardware and software requirements, focussing on the differences between the three architecture types, will be explored in the following section.

TV

The basic multi function front end block in Fig. 2 feeds a TS de-multiplexer / source decoder block which can be either dedicated hardware or, for improved flexibility and future proofing can be a software based media processor, such as the Philips TriMedia. In either of these situations, the VSB decoder can be used as a routing / switching device for the 656 and an additional MPEG-2 transport stream.

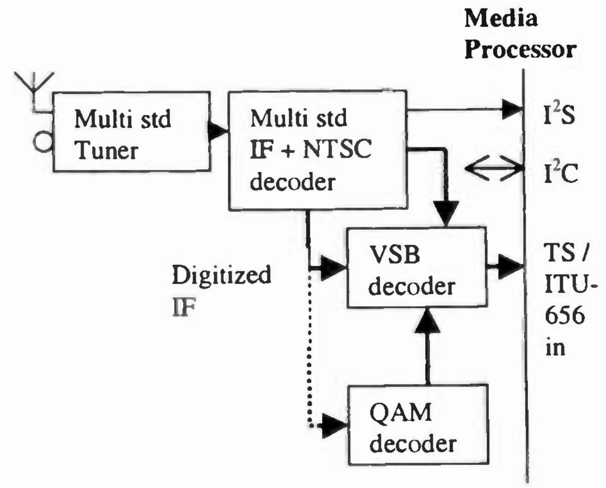


Fig. 4. Simplified TV front end architecture

PC

As CPUs and graphics controllers become more powerful it is fast becoming a realistic proposition to add DTV tuner capabilities to the Multimedia PC. This can be achieved using the same basic front end block as for the TV, but this time interfacing to a PCI bridge designed especially for video & data handling. As can be seen, the PCI bridge needs to handle the MPEG Transport Stream and ITU-656 inputs, and the provision of Inter IC Sound & Control ports greatly simplifies interfacing.

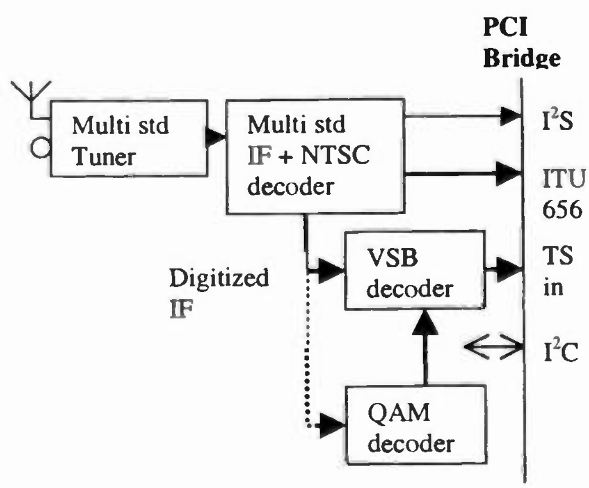


Fig. 5. Simplified PC DTV front end architecture

Fig. 6. shows how the ATSC Transport Stream and digitized NTSC signals are handled once they have been carried over the PCI bus to the host. The MPEG-2 HL assist block may or may not be required, depending on the capabilities of the CPU and graphics device, and the display resolution required.

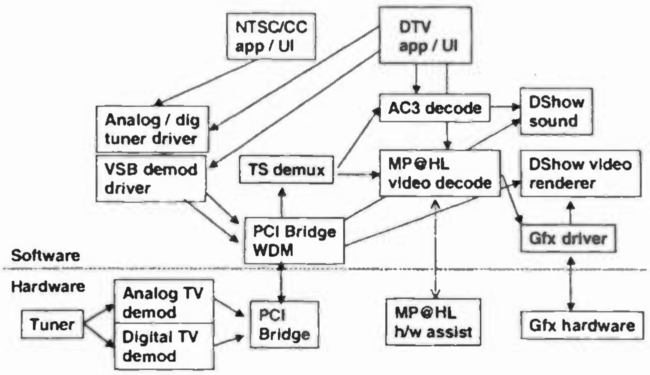


Fig. 6. PC DTV system architecture

Modular home appliance

The advent of the 1394 'Fire wire' bus and the Device Bay physical / electrical format allow the possibility of creating front end modules for access to different distribution media and service providers.

A Device Bay module for digital satellite TV, for instance, would contain the required tuner, QPSK demodulator and a service provider specific CA slot in a box sealed to make it 'tamper evident'. To prevent content piracy over the host link, the signal could be encrypted using a standard copy protection mechanism, before being sent according to the 1394 Trade Association's AV/C protocol.

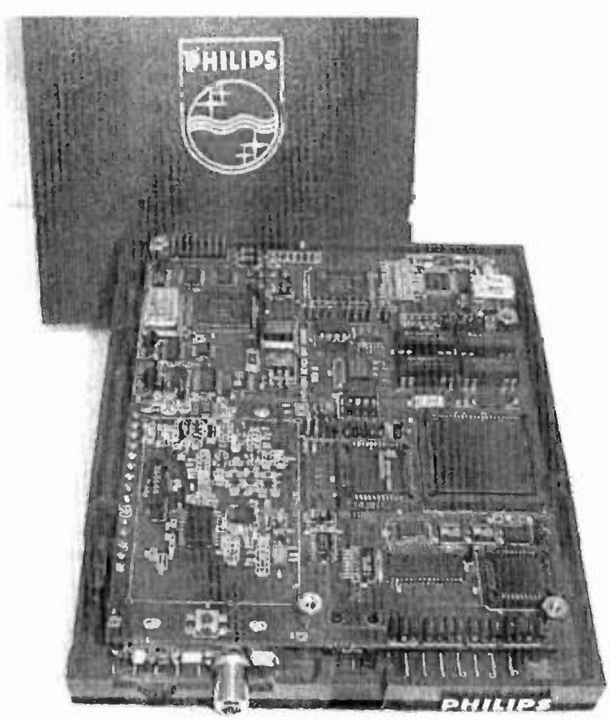


Fig. 7. Prototype ATSC Device Bay receiver

It is foreseen that 1394 could become a standard interconnection mechanism for tuner peripherals on both advanced CE appliances and home multimedia PCs. A Device Bay tuner unit for ATSC terrestrial reception has been designed & built, as shown in the photograph in Fig. 7.

The provision of reference design information, such as schematics, parts list and detailed PCB layouts enables platform providers to concentrate on aspects of the system such as user interface and applications which allow them to differentiate themselves from their competitors.

NEXT STEPS

Finally, some thoughts on where front end architectures for DTV can go next.

Front end modules are a logical extension of the reference design concept, where all of the tuner and channel decode functions are contained in a ready to use module, most likely a shielded can to start with. Where the specific channel format requires sophisticated software for control, it may make sense for this to be contained in an embedded microcontroller within the front end module also, simplifying the software interface with the rest of the system.

The trend towards higher levels of integration will also continue. Even in the limit though, it is likely that two silicon die will be required, since RF amplification, mixing and A/D conversion require IC process characteristics substantially different from deep sub-micron CMOS needed for ever increasing amounts of DSP. Architectures, partitioning and novel signal processing techniques, both in the analog and digital domains will be the future sources of competitive advantage for IC vendors.

ACKNOWLEDGEMENTS

The author would like to acknowledge the help & support of his colleagues in Philips Semiconductors Nijmegen & Sunnyvale sites, and those in Philips Research Briarcliff.

---

[i] Available commercially as the TDA8980
[ii] Available commercially as the TDA8961

Television Home Server for Digital HDTV Broadcasting

Tatsuya Kurioka, Hiroki Minami, Toshiyuki Fujisawa,
Haruo Okuda and Junji Numazawa
NHK (Japan Broadcasting Corporation)
Tokyo, Japan

ABSTRACT

In the age of digital broadcasting, demand will increase for a television Home Server that automatically records a viewer's favorite programs for subsequent viewing at any time. We propose a hierarchical storage management system (HSMS) as the architecture for developing a high-speed Home Server with very large capacity. Using this HSMS, we have developed a prototype Home Server. This Home Server, using video hard disk technology that we have developed, can play back a recorded HDTV program while recording another digital HDTV program on a single hard disk. The development of a hierarchical storage management method will also enable it to record over 4 hours of HDTV programs and play them back without having to wait. The hardware of this Home Server will be sufficiently compact to be built into a wall-mounted television set such as a plasma display panel.

1. Introduction

Digital broadcasting has significantly widened the scope of program selection, hence there is an even greater chance of viewers missing programs they want to watch. To reduce such missed opportunities, home storage devices will become increasingly important. Future storage devices for the home should provide not only large recording capacity, but also many other useful functions not provided by conventional home VCRs. The most important function is an "anytime function" that will allow viewers to watch favorite programs at any time, rather than the time at which such programs are broadcast. Such home storage systems with the "anytime function" we call Home Servers [1].

Integrated services digital broadcasting (ISDB[2]) offers a variety of broadcasting services including high-quality digital HDTV programs, an electronic program guide (EPG), program-related information, and multimedia news, as shown in Fig. 1. A television set installing a Home Server can provide up-to-date news and weather forecasts at any time of the day, enabling viewers to interactively enjoy broadcasting programs 24 hours a day, free from the set-time broadcasting hours. We conducted viewer questionnaire surveys on two occasions, in May 1996 and March 1997, concerning installing Home Servers in television sets. The results showed that 99 percent of those surveyed supported the idea. About 80 percent said that 7 hours would be adequate as the recording capacity of a Home Server. As for the question when they would like to have Home Servers installed in their television sets, 61 percent said within 3 years, and 93 percent within 4 years. These answers reflect the high demand for the early appearance of Home Servers.

This paper first describes the required specifications and functions of Home Servers for digital TV broadcasting, and then proposes a new HSMS consisting of several storage devices, such as semiconductor memory, hard disk, optical disk and

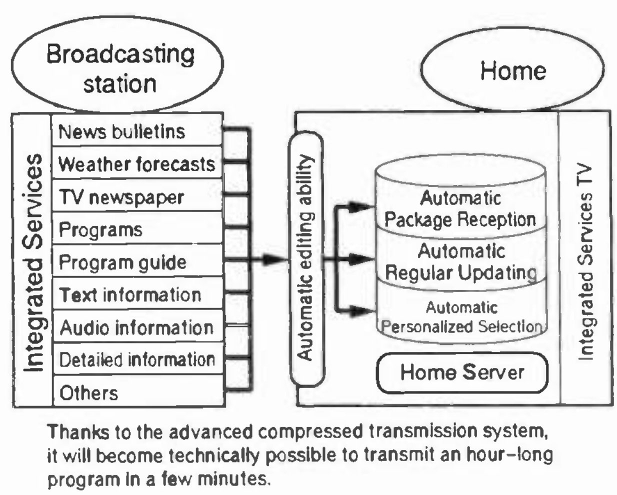


Fig. 1 Storage and Home-delivery of Broadcasting Services

tape, as the architecture for developing a high-speed, large-capacity and compact Home Server. Furthermore, it describes the element technologies, such as the video hard disk, hierarchical storage management method, storage agent, recording data format, new user interface, etc., and outlines the prototype Home Server and results of performance tests.

2. Required Specifications and Functions

Home Servers must be able to provide not only large recording capacity, but also many other convenient functions not available on conventional VCRs, as shown in Table 1. These functions include simultaneous recording and playback of several programs, high-speed retrieval, and interactive playback. Further, it should be compact and available at low price.

Home Servers must also provide two contrasting functions at the same time: the processing of news and other programs that are constantly updated, and that of movies and other programs kept in a video library. The storage devices of Home Servers must have a high-speed random access capability for the former function (primary memory), while possessing a large recording capacity for the latter (secondary memory), as shown in Table 2. To satisfy these requirements, we propose a new HSMS (hierarchical storage management system) consisting of several storage devices, such as semiconductor memory, hard disk, optical disk and tape, as the architecture of a Home Server [3].

This architecture allows step-wise expansion of memory configurations in the Home Server depending on the types of broadcasting services desired by viewers. Table 3 shows the grades of a Home Server ranked by available services and memory configurations.

3. Prototype Home Server

Digital HDTV programs (MPEG-2 MP@HL) are to be provided at the data transfer rate of about 23Mbps in BS digital broadcasting in Japan. To offer "anytime services", the Home Server must be able to play back a stored program while recording another one on the air. It should also be possible to play back a program while it is being recorded. This requires a Home Server to transfer data at 50Mbps or faster. Also, the recording capacity should be at

Table 1 Specifications of Home Server

	Home Server (future)	Home VCR (conventional)
Recording signals	Digital Broadcasting (MPEG2-TS, data), CATV services, Internet	Analog Broadcasting (NTSC)
Operation mode	Simultaneous recording and playing back	Recording or playing back (not both together)
Multi-program recording/ playback	○	× (1 program only)
Delayed viewing	○	×
Digest viewing	○ (nonlinear playback within a program)	×
Zapping viewing	○ (nonlinear playback between programs)	×
Program search	Intelligent, fast	Manual, slow

Table 2 Memory Construction of Home Server

	Primary memory (temporary memory)	Secondary memory (archive memory)
Functions	• Temporary storage of programs • Multi-task processing • Playing back while recording	• Program storage as a private library • Recording of many programs • Long-term program storage
Requirements	• High-speed random access • Recording/playback at high bit rate	• Large memory capacity • Inexpensive • Compact • Easy to handle
Examples	RAM, HDD	Optical Disk, Tape

Table 3 Grades of Home Server

Home Server	Memory	Service
Grade 1	Semiconductor only	Information-based programs, still pictures
Grade 2	Semiconductor + hard disk	Temporary moving pictures in addition to Grade 1
Grade 3	Semiconductor + hard disk + tape or optical disk	Private library for large-capacity moving pictures in addition to Grade 2

least 40Gbyte in order to store HDTV programs for up to 4 hours.

Fig. 2 compares the recording capacity and data transfer rate of storage devices that are expected around the year 2000. DVD as a removable storage device will be attractive, but currently does not have sufficient data recording rate for real-time processing of HDTV programs. It will take some time before we will be able to transfer data at a bit rate high enough for simultaneous recording and playing back of several programs. In this paper, the following storage devices were studied as an example of the Home Server [4], as shown in Fig. 3.

- Primary memory : semiconductor memory and hard disk
- Secondary memory : tape

The overview of the prototype Home Server that we have developed is shown in Fig. 4. Although based on a personal computer, the storage device controller for the hard disk and the tape drive has been newly developed. The hardware is compact enough to be built into a PDP (plasma display panel) and other wall-mounted television sets. Fig. 5 shows the block structure of the Home Server. The input/output signal of the Home Server is MPEG-2 Transport Stream (TS). The hard disk and the tape drive are connected by a single bus of Ultra-Wide SCSI for recording and playback control. Fig. 6 shows the software structure of the Home Server. To realize a HSMS that utilizes the respective strengths of each storage device, we developed storage device drivers for the hard disk and the tape drive each and also a storage agent for controlling several storage devices at the same time.

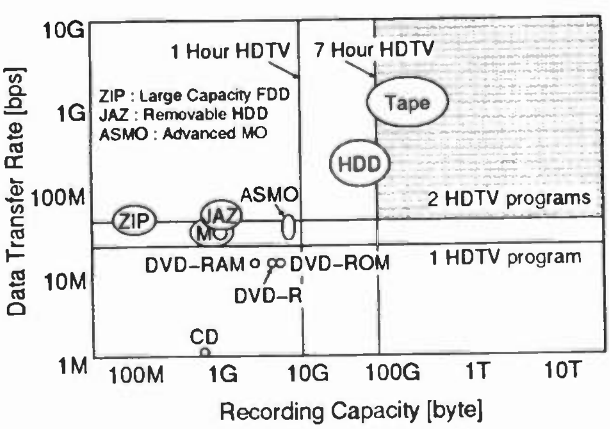


Fig. 2 Comparison of Storage Devices

The semiconductor memory (RAM) generates basic data blocks for recording on the hard disk or tape and also reproduces MPEG TS streams when playing back from the hard disk or tape. The hard

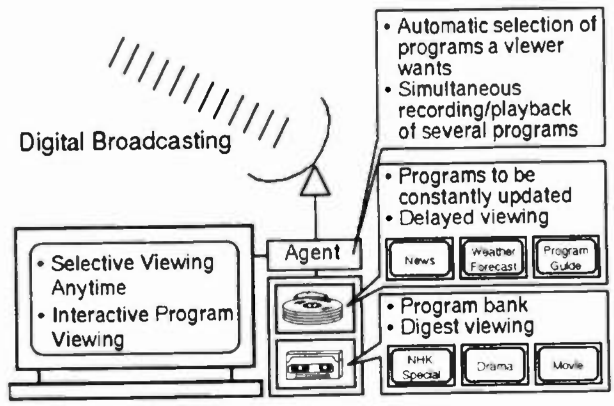


Fig. 3 An Example of the Home Server

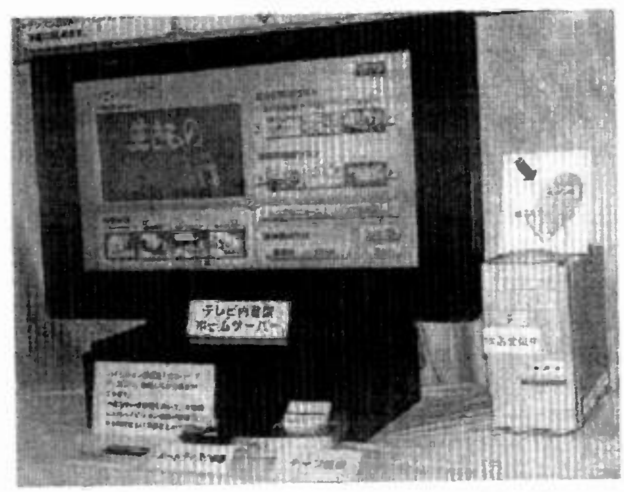


Fig. 4 Overview of the Prototype Home

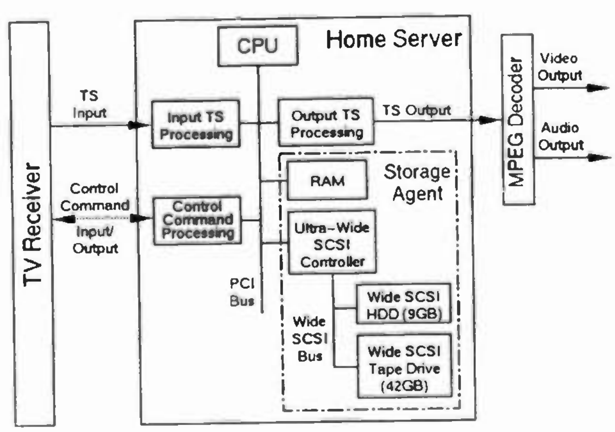


Fig. 5 Block Structure of the Home Server

disk is used for storing and playing back update-type programs (news, etc.) and supports multi-channel recording onto the tape. The tape stores programs that the viewer wants to keep in the private library. In addition to this conventional function, the tape in the Home Server lets the viewer retrieve a program much faster and offers such interactive services as zapping playback and digest playback.

4. New Element Technologies

The main element technologies newly developed for the Home Server are as follows.

4.1 Input/Output TS Processing

We developed a technology to identify event IDs

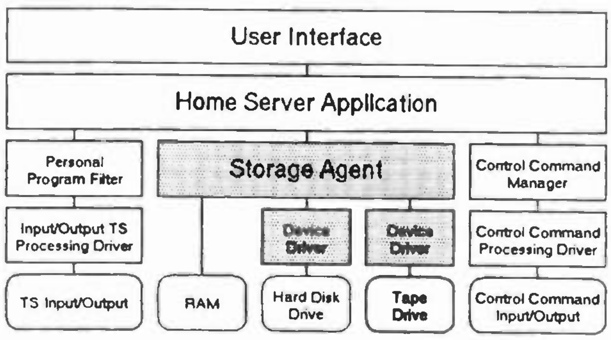


Fig. 6 Software Structure of the Home Server

and other program information multiplexed over the MPEG-2 TS. This allows the viewer to select and record the desired programs. Fig. 7 outlines the operation of the input/output TS processing board that we have developed. The Home Server input signal is the 28.305Mbps TS. The input TS processing leaves only those data packets related to programs to be recorded, subjecting the remainder packets to null-data packets of partial TS. During this processing, all of the RS error-correction codes attached to the data packets are removed. When a recording stream is produced, the null-data packets are also removed, and 4-byte inter-packet time information is attached to each data packet for the number of removed null-data packets can be counted later. The output TS processing inserts the null-data packets based on the inter-packet time information, and the 28.305Mbps partial TS, the same as the input TS processing, is reconstructed to be used as the Home Server output signal.

As the result of this processing, the recording stream for digital HDTV programs is handled at about 23.31Mbps, thus reducing the recording/playback data rate by 18% as well as allowing long-hour recording to make use of the recording capacity of the storage devices. As this input TS processing can also identify program IDs over the MPEG-2 TS, the prototype Home Server can accommodate

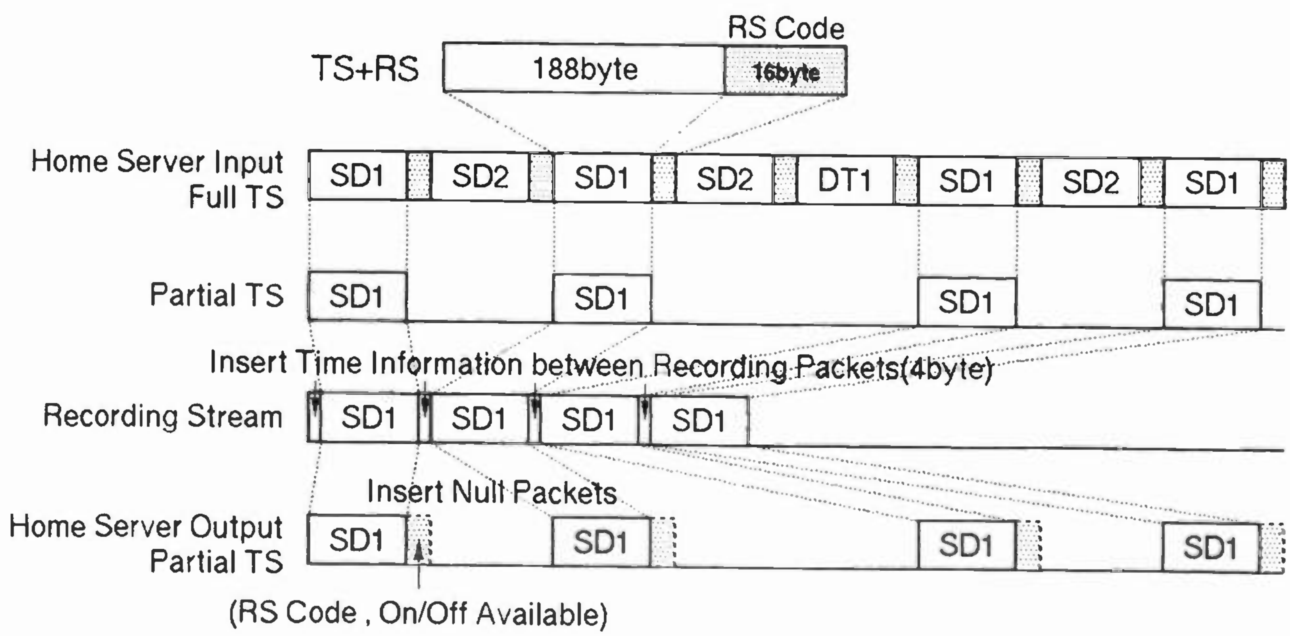


Fig. 7 Operation of the input/output TS processing

broadcasting services combining HDTV and SDTV (standard-definition television) programs. The hardware, including both input and output TS processing, is sufficiently compact and is the same size as a standard board of the PCI bus.

4.2 Video Hard Disk

A hard disk plays an important role as the primary memory in the HSMS. This is an essential storage device for quickly accessing television programs stored on tape or optical disk used as the secondary memory in the HSMS. However, a hard disk is basically designed to record computer data, and it has some problems to effectively record video data as a primary memory of the Home Server. Although a hard disk can transfer video data at the maximum burst rate in an instant, its continuous data transfer rate is quite slow so that zero-failure processing of a certain amount of data within a fixed time is required. We have developed a video hard disk technology [5] capable of recording and playing back the data with the same block size as the basic video data size, while guaranteeing continuous data transfer at high speed. With this technology, we can significantly reduce the seek frequency during simultaneous recording/playback of hard disk data, thus fully exploiting the advantages of hard disks in video applications.

In the prototype Home Server, we measured the read/write performance of the hard disk. The results are shown in Fig. 8. This measurement can be used to evaluate the multi-channel access capability of the hard disk, providing a basic design guideline for the Home Server. In order to realize simultaneous recording and playback on a hard disk, a bandwidth for 2 channels (46.62Mbps for this system) is needed as the data transfer rate of the hard disk. Fig. 8 indicates that the data can be processed without problem up to 50 seeks per second. If the same data are processed using an OS filing system in the ordinary computer environment, at least 100 seeks per second would be required for the hard disk, thus causing data processing to fail. Our video hard disk technology enables the size of the recording data block of the hard disk to be set to 1Mbyte with the original video filing system. As a result, only 9 seeks per second would be required for the hard disk, and the hard disk uses just 56 percent of its maximum capability in this data processing. The

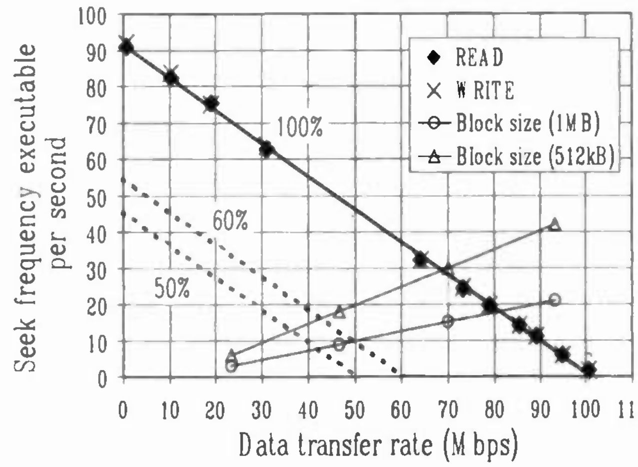


Fig. 8 Multi-channel Access Characteristics of the Hard Disk

prototype Home Server, using a video hard disk technology, can simultaneously record and play back digital HDTV programs on only a single hard disk, which contributes to the compactness of the Home Server.

4.3 Hierarchical Storage Management Method

The prototype Home Server uses three different storage devices: a semiconductor memory, a hard disk and a tape. This is not just a simple combination of three types of storage devices; rather, the features of each of these devices are finely adjusted to record and play back television programs, which makes this system unique. The Home Server can dynamically access several storage devices at the same time, which conventional computers cannot do. We thus achieved quick program retrieval, digest playback, simultaneous recording of several programs, and synchronized interactive playback from the tape, as explained below.

4.3.1 Quick Program Retrieval from Tape

A conventional VCR usually takes several minutes before the beginning of a program recorded toward the end of the tape can be accessed. The Home Server stores the beginning of each program on the hard disk, instantly picking up any server program from the tape. It can seamlessly switch from the hard disk to the tape for playback after searching for the program on the tape is completed. Viewers can thus watch a program recorded on the tape without waiting for searching of the program.

4.3.2 Interactive Playback from Tape

A tape is a storage device that cannot be accessed randomly at high speed because it is mechanically driven. Its high-speed data transferring, however, enables a program to be recorded and played back faster than real time. If a search for the next playback scene can be completed by utilizing this time difference, we can realize seamless, nonlinear playback, as shown in Fig. 9. This method improves the search capability of the tape by utilizing the high-speed data transfer rate, thus achieving interactive playback from the tape.

In the prototype Home Server, we measured the nonlinear access (seamless playback) characteristics from the tape. The result is shown in Fig. 10. The curve shows that, with the playback time set at 30 seconds per scene, it is able to play back scenes up to 20 minutes ahead in seamless non-linear playback.

4.4 Storage Agent

In this Home Server, tasks are divided among three storage devices to record and play back a television program. It is important, therefore, to control these storage devices in an intelligent manner, automatically managing stored programs and information. A storage agent performs this function. It intelligently selects and controls recording/playback devices and their operations.

The prototype Home Server, for instance, has six operation modes as shown in Table 4, which are selectively controlled by the storage agent. In this way, viewers perceive that what they are watching comes from a single storage device so that they are unaware of the tape, disk and other storage devices. They simply select a program and specify the form of its program presentation. With the storage agent switching these six operating modes while managing their operation time, the Home Server can simultaneous record and play back several programs in real time.

Dramatic improvements in many different storage devices are expected over the next few years. The storage agent technology will play a greater role as it does not limit the types of storage devices used in Home Servers.

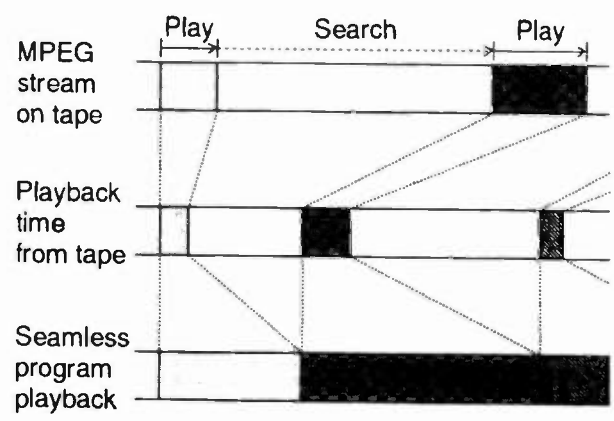


Fig. 9 Nonlinear Playback from Tape

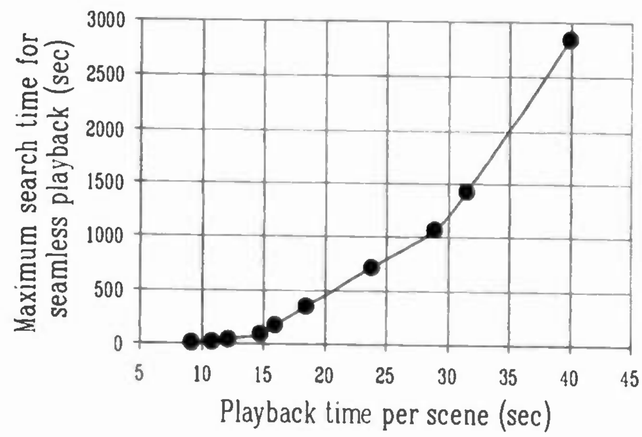


Fig. 10 Nonlinear Access Characteristics from the Tape

Table 4 Operation Modes

Modes	Recording	Playback
Mode 1	Program Reception, Recording on HDD	No Playback
Mode 2	Program Reception, Recording on HDD	Playback from HDD
Mode 3	Program Reception, Recording on HDD	Playback from Tape
Mode 4	Program Reception, Recording on HDD & Tape	No Playback
Mode 5	Program Reception, Recording on HDD & Tape	Playback from HDD
Mode 6	Program Reception, Recording on HDD & Tape	Playback from Tape

4.5 User Interface

A user-friendly interface is important for viewers to make use of the functions of Home Servers. Fig. 11 shows the user interface that we have developed. The viewer can move to a desired interface screen simply by selecting the icon in the menu. Thanks to the storage agent technology mentioned earlier, the viewer can operate the Home Server without being aware of the existence of storage devices such as tape and disk. The Home Server provides the "anytime services" as well as new viewing control functions not provided by conventional storage devices.

The "anytime services" allow viewers to directly choose the most up-to-date news and weather information at any time of day. Among the new viewing control functions, viewers can choose one from "Mark / Play from Mark," "Beginning," "Digest," and "Zapping." The first two allow truly time-free program viewing so that the viewer can temporarily stop watching a program on the air or watch a program from the beginning of the program on the air. "Digest" allows the viewer to watch only the highlight scenes of a program. "Zapping" allows the viewer to freely choose more than one program while enjoying television.

5. Summary

There is strong demand for early implementation of the television Home Server, which is a future video storage system for interactive TV. We proposed a new HSMS consisting of several storage devices, such as semiconductor memory, hard disk, optical disk and tape, as the architecture for developing a high-speed, large-capacity and compact Home Server. This architecture enables viewers to upgrade the memory configurations in the Home Server depending on the grade of the necessary services, from using a simple semiconductor memory to a hard disk and a tape.

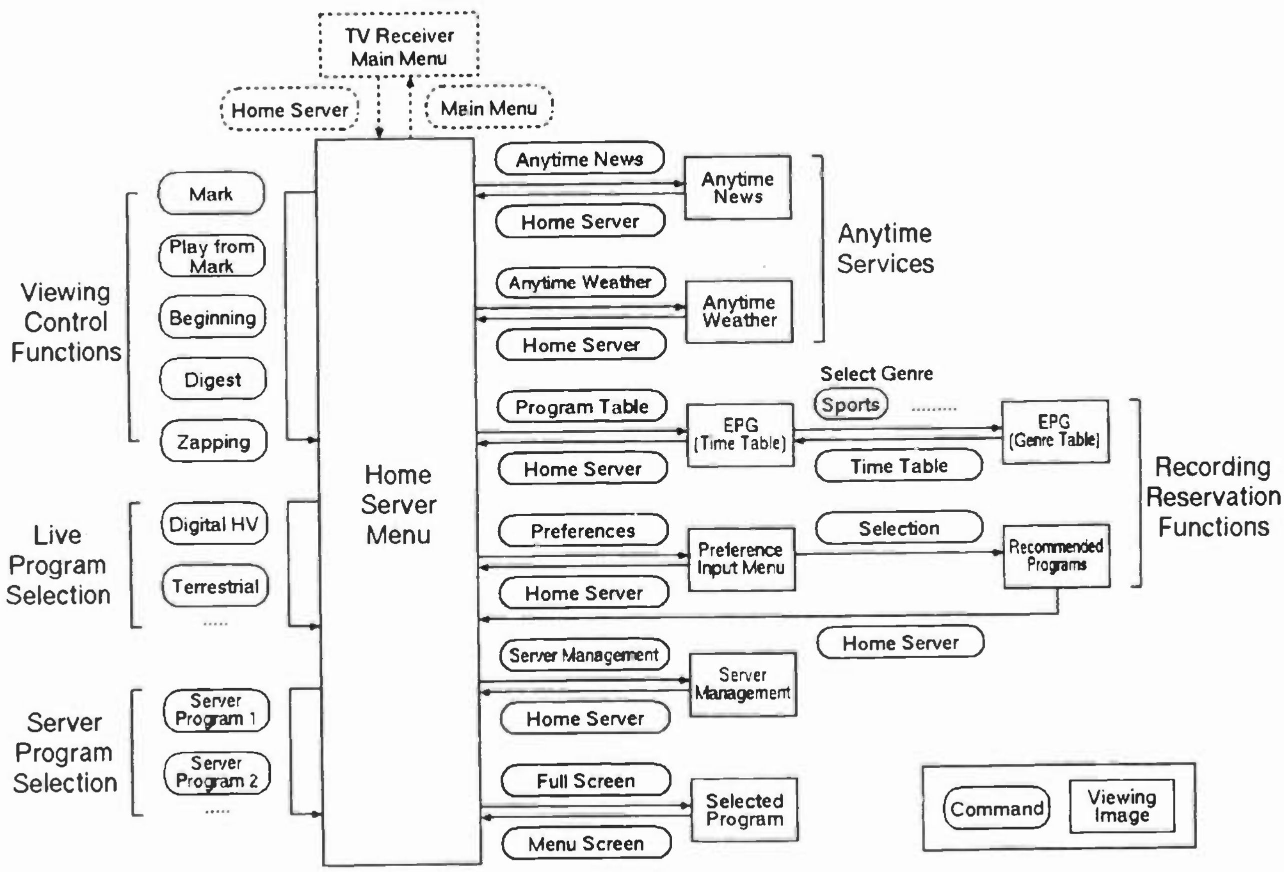


Fig. 11 User Interface of the Home Server

We have developed a prototype Home Server, compact enough to be built into a wall-mounted television set, with an "anytime function" for digital HDTV broadcasting. Our video hard disk technology enables the Home Server to play back a stored HDTV program while recording another digital HDTV program on only a single hard disk. The development of the new HSMS also enables the Home Server to record over 4 hours of HDTV programs and play them back without having to wait, even though the device uses a tape. In this HSMS, the storage agent automatically controls the recording and playback of programs and related information. Viewers can thus enjoy programs without being conscious of the storage device on which the programs are recorded, with new viewing control functions not provided by conventional home VCRs.

References

[1] T. Kurioka, "Digital Broadcasting and Home Server," ITE Annual Conference, S1-4, pp.438-441, July 1998

[2] T. Isobe, A. Ohya, K. Usui, M. Ueno, A. Yanagimachi and T. Nishizawa, "Interactivity in Broadcasting and Its Application to ISDB Services," IEEE Transactions on Broadcasting, Vol.42, No.3, pp.179-186, September 1996

[3] T. Kurioka, H. Minami, H. Okuda, J. Numazawa, A. Yanagimachi and H. Ohshima, "A Hierarchical Storage Management System Toward the Realization of ISDB Anytime Services," ITE Technical Report Vol.21, No.31, pp.33-38, May 1997

[4] T. Kurioka, H. Minami, H. Okuda, J. Numazawa and A. Yanagimachi, "Television Home Server for Integrated Services," IEEE Transactions on Consumer Electronics, Vol.44, No.4, pp.1195-1200, November 1998

[5] T. Kurioka, H. Minami, H. Okuda and H. Ohshima, "A New Video Hard Disk System for Multimedia Program Production," IEEE Transactions on Broadcasting, Vol.42, No.3, pp.222-228, September 1996

Radio: The Computer Connection

Monday, April 19, 1999

10:30 am - 12:00 pm

Chairperson: Barry Thomas
KCMG-FM, Los Angeles, CA

10:30 am SAN Enabled Standard Network File System
Doug Anderson
Transoft Networks
Santa Barbara, CA

***11:00 am Averting Obsolescence: How an Open-Platform Architecture Can Foster Compatibility and Greater Development**
Mike Palmer
Associated Press Broadcast
Washington, DC

***11:30 am From On Air to On Line: The Internet Provides New Revenues for Radio Stations**
Steve Carley
On Radio
Scotts Valley, CA

*Papers not available at the time of publication

SAN Enabled Standard Network File System

Douglas B. Anderson
Transoft Networks, Inc.
Santa Barbara, CA

ABSTRACT

Data sharing solutions for digital imaging markets have come of age. Third generation SAN solutions are available today, and they are suitable for a wide array of data-intensive applications with diverse workflows. The SAN Enabled Standard Network File System is a prime example of this state-of-the-art technology and a harbinger of things to come.

1.0 INTRODUCTION

There are numerous and compelling reasons for sharing data in the broadcast and post-production industries. When sharing is done efficiently multiple editors and artists can complete pieces faster, easily incorporate existing material, utilize expensive equipment better, and move work between different work centers with ease.

Digital non-linear editing for broadcast and post-production not only provides faster turnaround and higher quality productions, but also places tremendous demands on the workstations, interconnect, and storage devices. Large files and real-time streaming required for on-air systems also preclude all but the highest performance networking solutions.

Conventional networks based on common Local Area Network (LAN) technology like Ethernet may seem enticing, but today they cannot deliver the performance the broadcast and post-production industry demands. To overcome these limitations, the industry has gravitated toward Storage Area Networks (SAN) for data sharing.

Fibre Channel technology has quickly become the underlying technology of choice for data sharing SAN solutions; but more is required than just this core technology. In order to obtain the true benefits of shared storage networks a complete solution is needed which includes software for file sharing and storage management.

2.0 SHARING APPROACHES

Two distinct approaches have evolved to create data sharing environments: conventional corporate networking, and the SAN. The fundamental difference between the approaches is where the network resides relative to file system. In conventional corporate networks the shared file systems reside on servers that are separated from applications by the LAN or Wide Area Network (WAN). With SANs the applications and file system reside on the same machine (e.g. the user's workstation) and the storage devices are directly attached via a high-speed network.

2.1 Conventional Networks

Conventional corporate networks share data though the use of file servers interconnected to user workstations (clients) via LANs or WANs. Network security and file system protocols like NFS and SMB/CIFS provide the transactional semantics for data sharing. While this approach provides the connectivity and ensures file system coherency, performance is limited.

The vast majority of corporate network infrastructure connecting to user workstations is only capable of supporting 10/100BaseT Ethernet or similar speed communications technologies. With a maximum bit rate of 100 Mb/Sec, only two users streaming data concurrently from a server at 3 MB/sec can be supported. When real-time performance counts or more users are required, it's just is not enough!

Clearly, a higher speed connection must be installed for demanding digital video applications. The question is what kind of connection. Proponents of Gigabit Ethernet suggest that it is the solution for high-speed networking. Though the bit rate of Gigabit Ethernet rivals the 1,062 Mb/sec rate of Fibre Channel, there are fundamental issues that hinder Ethernet performance scaling.

These factors include the 1,500 byte Ethernet frame size and the transport control protocols like TCP that must be layered on top of connectionless network protocols like IP to provide guaranteed data delivery. The resulting "network protocol stack" and small frame size imposes a significant burden on the client and server central processing units (CPU). This overhead includes packet sequencing, error detection, and recovery mechanisms. Combine the Ethernet limitations with server system bottlenecks, and the result is throughput rates that are approximately half of what can be achieved using a Fibre Channel based SAN.

2.2 Storage Area Network (SAN)

Storage attached to a dedicated high-speed network, like Fibre Channel (FC), is the basis for a SAN. When combined with management software, it becomes a network where many computer systems can share the storage.

SANs go straight to the heart of the performance issues that plague conventional networks. In SAN implementations the user workstations are connected directly to storage devices via a high speed interconnect like Fibre Channel, completely eliminating the file server. Performance is dramatically improved because the SAN eliminates server data path bottlenecks and reduces network protocol overhead. This reduction in CPU overhead is attributed to the relatively large 64,000 byte FC frame size and the thinner network protocol stack resulting from the guaranteed network data delivery mechanisms handled by the FC host bus adapters.

3.0 SOFTWARE SOLUTION

As computer operating systems evolved along with conventional networking technology, file systems were developed to meet the demands of the client/server paradigm. The result was network operating systems (NOS), like Novell NetWare, Windows NT, UNIX, etc. In this environment, file system integrity is only guaranteed if one computer is directly accessing and managing a disk volume. If more than one computer running a standard NOS directly accesses the same volume, corruption results; each computer assumes it owns the volume and uses the data blocks as needed. Since none of the computers are aware that another may also have direct access to a disk volume, the computers inadvertently overwrite each other's information, commonly known as the cache coherency problem.

3.1 Software Options

Though the SAN hardware provides the physical paths for computers to directly access disks, it does not enable multiple computers running a standard NOS to share disk volumes, because the hardware does not solve the cache coherency problem. In order to share the data on a disk volume, additional software must be provided to manage access. This software may take the form of a distributed lock manager, replacement file system, or standard file systems with extensions.

First/second generation SAN data sharing solutions predominately use the distributed lock manager approach to achieve volume level sharing. More recent solutions either provide a replacement file system or add extensions to a standard file system, to achieve a finer level sharing granularity and a more dynamic sharing environment. This evolution represents another step towards the goal of delivering a data sharing solution with the ease of use associated with conventional networks without any of the LAN/WAN performance limitations.

3.2 Decision Factors

Each software design approach has positive features and limitations or drawbacks, so the key is selecting one that maximizes reliability/availability, performance, and ease of use. Developers must also consider what approach will bring the technology to market in a timely manner without compromising the primary objectives.

For example, before Transoft Networks selected the software architecture for their 3rd generation data sharing solution they considered many factors. The distributed lock manager approach was not chosen, because experience suggested that it would be more challenging to extend that approach than to select another method that already had provisions for finer granularity access and dynamic control. The replacement file system approach was not selected, because any potential benefits were heavily out-weighed by the onerous task of trying to create a new non-standard file system (historically it has taken years to stabilize and refine a new file system). Instead, Transoft Networks chose to implement their 3rd generation by extending a standard network file system.

4.0 SAN FILE SYSTEM

Transoft Networks 3rd generation data-sharing solution adds extensions to Microsoft's® New Technology File System (NTFS) that maintain file system coherency for the SAN attached storage. By choosing to add extensions to NTFS, the basic file system mechanisms for file locking and access privileges are preserved, Windows NT network security is maintained, standard protocols like SMB/CIFS are available to facilitate heterogeneous platform support, and storage network adapter drivers can be used without modifications. The result is a "SAN Enabled Standard Network File System".

To understand how the SAN Enabled Standard Network File System works, it is necessary to understand the terminology and architecture for what is known technically as an asymmetric clustered file system with distributed I/O.

4.1 Terminology

What is a Cluster?

A Cluster is a type of parallel or distributed system that consists of a collection of interconnected nodes (computers) used as a single, unified computing resource.

What is a Clustered File System?

A clustered file system distributes command (metadata) and data transfer operations to the nodes within the cluster, allowing these operations to occur in parallel on multiple nodes.

What is an asymmetric file system?

An asymmetric file system is one in which critical file system on-disk structures, security, and range locking are centralized in a coherent unit that may be separated from the nodes performing data transfer operations.

What is an asymmetric clustered file system?

Asymmetric clustered file system acts as a single, unified computing resource to present a file system among nodes connected to the clustered file system. The file system input/output (I/O) operations are split into metadata and data transfer portions, enabling each node accessing the file system to perform direct I/O data transfers with network attached storage, thereby distributing the I/O and acting as a cluster.

4.2 Metadata portion of an I/O

The metadata portion of an I/O concerns itself with what is known as file system metadata operations. File system metadata is used to maintain the on disk structure of the file system. All operations on the file system, such as range locking, directory services, and managing file attributes, need to be coordinated so that the integrity of the file system is maintained. The SAN node responsible for handling the metadata portions of I/O for a file system is known as the Metadata Server for that file system.

4.3 Data Transfer portion of an I/O

The data transfer portion of an I/O concerns itself with the movement of data between SAN nodes and storage devices. When a logical file I/O request enters the shared file system it is split into metadata and data transfer portions. The I/O request is forwarded to the Metadata Server if the desired data is not already in the node's disk cache. The Metadata Server converts logical file I/O operation into a physical block list and returns it to the originating node. The physical block list is used to manage the direct data transfers to/from storage, without further Metadata Server interaction. (See figure 1.)

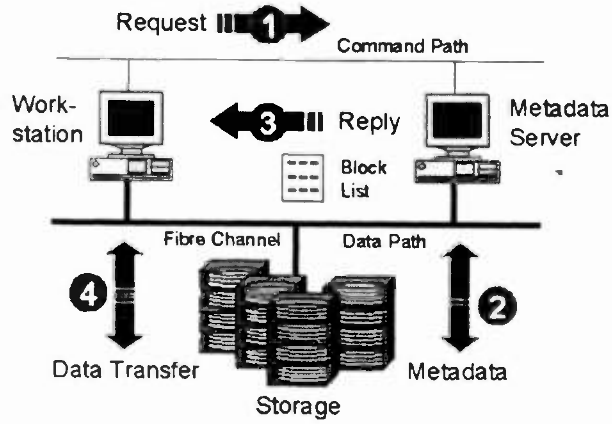


Figure 1: Direct I/O Operation

5.0 FS FEATURES & BENEFITS

Storage Area Networks with clustered (shared) file systems create an ideal environment for data-intensive applications such as non-linear video editing, audio, animation, graphics, medical imaging, and geo-sciences. By combining the high-throughput capabilities of interconnect technology like Fibre Channel with the ability to share storage across multiple heterogeneous machines, SAN Enabled Standard Network File Systems create the possibility for entirely new applications and work flow models.

5.1 Features

Though there are many different approaches to data sharing, a robust 3rd generation solution should have a feature set like the one provided by Transoft Networks' SAN Enabled Standard Network File System.

- Globally shared file systems
- Multiple readers/writers per file system
- Dynamic file and byte-range locking
- Direct data I/O from storage to DS nodes

- Heterogeneous system support
- Cross platform data sharing
- Standard file privileges and security
- On-line storage & file system management
- Journaling, integrity checks, and mirrored metadata

Not *all* solutions are created equal, though it may not be apparent when comparing a basic feature list to the one presented above. Judicious use of existing technology, adherence to industry standards, and careful attention to network adapter, protocol, and media independence provides additional benefits.

These principles were applied to the SAN Enabled Standard Network File System to ensure that the solution would be compatible with existing network infrastructure, to insulate it from changes in the base file system, and so an exceptionally robust solution would be available to address today's data sharing needs.

5.2 Benefits

The jury is in! The benefits of SAN data sharing solutions have been proven in the digital imaging markets. Only two years ago data sharing appeared to be a novel solution promoted by a handful of entrepreneurs. Today SAN technology has advanced and blossomed into a mainstream component of the industry. The fundamental reasons that production managers and editors have been eager to adopt this technology are:

- Increased productivity
- Higher data availability
- Cost saving from better storage utilization
- Ability to stream data at higher rates
- Seamless integration into workflow

In particular, SANs are helping news-rooms meet critical deadlines by enabling editing to occur in parallel with digitization, and work to be passed rapidly from one suite of tools to another in process. Sharing reduces production time by providing the workgroup with instant access to libraries of archival footage and new material from locally digitized clips or from satellite feeds. Once a job is completed, it is ready for immediate access, because it is not necessary to copy the master to tape and carry the tape over to the on-air system.

6.0 Y2K - A VERY GOOD YEAR

Despite all the gloom and doom forecasted for the year 2000, it should be a very good year for the SAN industry. The benefits accrued from continuing industry wide emphasis on Fibre Channel interoperability testing and anticipated deployment of next generation SAN software solutions is something to really be looking forward to.

The technology underlying the SAN Enabled Standard Network File system is providing a spring board that Transoft Networks is using to develop a data sharing solution with more even more flexibility and higher availability than any solution available today. This year 2000 deliverable is called FibreNet™ H/A.

FibreNet™ H/A will provide continuous distributed clustered file system services despite availability changes of system servers. FibreNet™ H/A network policies will allow relationships between nodes and file system resources to be defined for diverse sets of circumstances, including transparent client I/O recovery and load balancing. Policies will be enforced transparently, so scaling will continue gracefully as more nodes and storage are added to the SAN, offering an increase in availability as well as throughput. Hence, the SAN Enabled Standard File System is just the beginning!

DTV Production

Monday, April 19, 1999

1:00 PM - 5:30 pm

Chairperson: Robert Seidel
CBS Inc., New York, NY

1:00 pm MPEG Splicing
Mike Knowles
NDS Limited
Hampshire, United Kingdom

1:30 pm Emerging Technologies for Content-Based Access of Digital Video
Michael Smith
ISLIP Media, Inc.
Pittsburgh, PA

2:00 pm Workflow Implications of an Integrated Digital Newsroom
Charlie Bernstein
Leitch Technology Corporation
Burbank, CA

***2:30 pm A Progressive Route to Film Mastering for DTV**
David Bancroft
Philips Digital Video Systems
Reading, Berkshire, United Kingdom

3:00 pm Reframing *A Bugs Life*
Craig Good
Pixar
Richmond, CA

3:30 pm European Commission ACTS - MIRAGE: A Success Story in Virtual Reality
Chas Girdwood
Independent Television Commission
Winchester, United Kingdom

4:00 pm Remote NLE
Robin Rowe
SAIC
San Diego, CA

4:30 pm Design and Implementation of an Integrated News Production System
Steve Owen
Quantel, Ltd.
Newbury, Berkshire, United Kingdom

5:00 pm Tape Libraries in Digital Video Storage Area Networks

Steve Georgis
Exabyte Corporation
Boulder, CO

---

*Papers not available at the time of publication

MPEG SPLICING

Mike Knowles, Nigel Dallard, Alois Bock

NDS Broadcast,
Southampton
U.K.

ABSTRACT

Traditional analog network and affiliate broadcasters re-broadcast contributed content (e.g. network programming), and switch in locally sourced news, advertisements or other content. These sources are edited and switched into a program stream according to a schedule. The switching is easily achieved as the program material is all in the same format and the picture frames, making up the material, are easily identified and labeled.

In the digital world the production requirements are the same. However, the content will now be delivered to the affiliate in a pre-coded state. The broadcaster will need to pass through the High Definition content through the transmission network without decode/recode or editing, and insert the local material including news, advertisements and other content into the digital feed. This is an extremely complex process, which requires the use of MPEG splicing. It will inevitably lead to compromises in program production, but is a very cost-effective alternative to the network affiliates budgeting large capital expenditures to equip their transmission studios with expensive High Definition equipment.

This paper will examine the challenge of MPEG splicing, discussing the principles behind MPEG splicing and the developments being made to provide simple, cost-effective systems to allow the broadcasters to address the needs of digital broadcasting.

INTRODUCTION

Bitstream splicing is the process by which MPEG compressed programs are switched, concatenated edited, etc. in the compressed domain, i.e. without the need for decoding and re-encoding. It involves joining each of the component elementary streams of one program with the corresponding streams of another. Due to the complexity of the encoding and decoding algorithms, attention has so far concentrated mainly on splicing issues of the compressed video signal itself[1,2,3] with little consideration given to other constituents of MPEG transport streams.

Therefore, the techniques for splicing MPEG video elementary streams are well understood and need not be considered in detail in this paper, although a brief overview will be given. Instead, this paper concentrates on the implications of transport stream splicing for broadcasters and aims to clarify some of issues concerning the application of MPEG transport stream splicers in a real broadcast environment.

Applications

There are many potential applications for MPEG transport stream splicing, including but not limited to

- local insertion of pre-compressed sources (e.g. advertisements) from video servers into a live feed;
- live-to-air local program insertion (e.g. news) into a national network feed;
- insertion of pre-compressed network feeds into local broadcasts;
- switching of other (non-network) content into the program.

Advantages of transport stream splicing

Transport stream splicing, compared with decoding and re-encoding solutions, has a number of benefits for broadcasters in cases where frame accurate switching is not required, such as

- low cost;
- compact solution;
- possibility to concatenate sequences with different resolutions;
- low processing latency, typically less than 1 ms;
- no loss of picture quality;
- interoperability through the proposed SMPTE-312 standard;
- seamless or near-seamless transitions.

MPEG VIDEO SPLICING

Probably the most important aspect of MPEG video splicing is the management of the decoder buffer fill level during and after the splicing operation. In an MPEG bitstream, there is a constant delay from the input of the encoder to the output of the decoder, but the relative fill-level of the encoder and decoder buffers varies greatly over time, depending on the type of video frame being transmitted, or on the complexity of the scene, among other factors. Thus, it may be required to splice from a bitstream with a full decoder buffer to one with the occupancy close to empty. This is almost certain to cause a decoder buffer overflow after the splice if no measures are taken during the splice to reduce the decoder buffer occupancy.

Video elementary stream constraints

To make continuous decoding possible, the first picture in the bitstream after the splice has to be intra coded. This is fairly easy to arrange if you are splicing to a program coming off a server, or to a co-sited local encoder. If, however, the two bitstreams are coming from remote sites, or there is no control over them for other reasons, the splicer has to manage the transition between the out point of one stream and the in point of the other. This can be done in a number of ways as shown in the next section.

If B-pictures are used, it is necessary to ensure that the first Group of Pictures (GOP) after the splice is "closed", that is, there are no B-pictures dependent on I or P pictures from the preceding GOP. This would result in bad block distortion because the previous I or P picture in the receiver came from a different bitstream.

In 60Hz environments, where 3:2 pull-down is used, it is necessary to ensure that the field-parity sequence is maintained over the splice-point. SMPTE-312 specifies "out-points" after a bottom-field and "in-points" before a top-field.

Video splicing types

MPEG defines two types of splicing.

- **Seamless splicing**, where the decoder video buffers are carefully equalised before the splice takes place and
- **non-seamless splicing**, where the decoder buffer is allowed to underflow and decoding restarts on the new bitstream.

As is implied by the names, seamless splicing results in a much cleaner join of the two video sequences – in fact the join is invisible. Non-seamless splicing will normally result in a short freeze-frame at the decoder whilst the buffer re-fills and decoding restarts. Alternatively, it is possible to inject a black I frame during the splice transition if a freeze-frame of the video signal is considered undesirable.

In addition to the splice types defined by MPEG it is possible to perform what has become known as **near-seamless splicing** whereby the freeze-frame or black period is reduced to a minimum - usually just a few video frames. In this method the decoder buffers are not allowed to empty nor are they required to be exactly equal at the splice point, but are deliberately managed during the transition across the splice. This method combines the advantages of non-seamless splicing in so far as the two buffer occupancies are not equalised (and hence the video coding algorithm does not have to be constrained) and seamless splicing because the transition is often fast enough not to be noticed.

Therefore, with only a small amount of processing power in the splicer it is possible to avoid sacrificing encoding efficiency for bitstream splicability. This can be achieved with the following measures.

- Decoder buffer management during the splicing operation;
- GOP closure in the splicer rather than in the encoder.

Video bit rate constraints

When splicing between programs compressed to different data rates it is necessary to make sure that there is enough capacity in the output multiplex to cope with the higher of the two bitrates. Since multiplexed transport streams are usually running to full capacity this means that spliced-in streams have to be of the same or lower bitrate.

The much bigger problem is statistically multiplexed bitstreams. More and more broadcasters are looking to utilise such systems to improve the efficiency of the encoding. In such systems, the bit-rate of the video component of any one program is continuously varying over a wide range. To splice into such a bitstream, you either need to ensure that the program you are splicing to is coded at the lowest bit-rate utilised by the program you are replacing, or the program in question is taken out of the statistical multiplexing algorithm for the duration of the insert. The use of multiple logical groups of statistically multiplexed programs, with real-time allocation of programs to those groups, enables the broadcaster to combine the advantages of statistical multiplexing with the flexibility of bitstream splicing.

CLOCK REFERENCES AND TIME STAMPS

The chances of any two independently encoded programs to be spliced together having the same clock reference (PCR) values are extremely remote. Therefore the splicer has to either

- change all PCR, PTS and DTS values of the 2nd bitstream to match those of the first, such that the decoder does not see any discontinuity or
- signal a system timebase discontinuity at the splice point and leave the decoder to recover according to the MPEG standard.

The second of these two options has a number of disadvantages:

- The MPEG-2 standard does not allow the presence of PTS or DTS fields referring to the "old" PCR after the first occurrence of the "new" PCR. This can potentially cause problems with audio and other non-video streams due to the transmission delay between video and audio (and other) frames with the same PTS.
- Timebase discontinuities in the transmitted bitstream may enable "commercial killers" to be built.
- Commercial decoders/IRDs on the market today will not have been tested against a spliced bitstream, and may misbehave when presented with a timebase discontinuity (despite the MPEG-2 Conformance specification requiring that they cope).

It is therefore better for the splicer to alter the PCR, PTS and DTS fields in the 2nd stream to make it adhere to the same timebase as the 1st.

PIDs

An MPEG-2 decoder determines the relevant PIDs for each of the elementary streams of a program from the Program Specific Information (PSI). This information is only transmitted every few hundred milliseconds and the response time to changes is specific to decoder implementation. Therefore, the splicer must re-map the PIDs of the second bitstream to match those of the first to ensure there is no delay in starting to decode the second bitstream.

Furthermore, consideration has to be given to the possibility that the programs being spliced have different numbers of elementary streams. For example, a network feed may have a second-language audio component, whilst the advertisement to be inserted only has one. For a combination of technical and commercial reasons, including the reaction time of decoders to changes in PSI, and to defeat "commercial killers", it is necessary to define the number of elementary streams being transmitted as fixed and either drop or duplicate streams as appropriate.

A couple of examples may help to explain. First, consider a local affiliate station receiving a network feed which always has English and Spanish language audio components. If the local affiliate station produces programmes and advertisements which only have English language audio, then it may be best for the splicer to drop the Spanish audio stream, and modify the PSI accordingly when the network feed is spliced in. Now consider a Canadian local station which usually produces bilingual English/French programming, but the incoming network feed only has English audio. In this case, it may be best for the splicer to duplicate the English audio on the French audio PID whenever the network feed is spliced in, to ensure that the French speaking listener is never left without audio for a few seconds whilst the decoder detects and responds to the PSI change indicating the loss of French audio.

OTHER SERVICES

ATSC Closed Captions

ATSC systems carry Closed Captions in the user_data of the video elementary stream, according to EIA-708. They will therefore automatically be spliced along with the video. If there are different numbers of closed caption streams, e.g. for different languages, then similar problems are encountered to those described for audio above, and either the PSIP has to be adjusted at the time of splice, or the closed caption stream inside the video bitstream has to be processed by the splicer. A more pragmatic solution when the closed caption stream contains more languages than required is to simply not describe the "extra" languages in the PSIP.

DVB Subtitles

DVB systems carry subtitles as a separate packetised elementary stream, according to ETS 300 743. This means that the splicer can easily drop or duplicate subtitle streams, however, a single PID stream can carry multiple language subtitle streams, and thus similar problems and solutions exist to those described for ATSC above.

DVB subtitles streams need some simple processing at the splice point to ensure that the subtitle decoder is cleanly "reset" at the splice point and does not continue to use regions and objects defined prior to the splice point.

DVB Teletext

World System Teletext is conveyed in DVB bitstreams according to ETS 300 472. This specification defines that Teletext is conveyed in PES packets that are exact multiples of Transport Packets. There is no retained "state" between PES packets, so the actual splicing is very straightforward. The issues are similar to that for audio - what to do if the old and new transport streams have different numbers of Teletext streams defined.

DSM-CC Carousels

DSM-CC Data and/or Object Carousels are now being designed in to systems to convey programme-related data for Interactive TV applications. There are two splicing-related issues that need to be considered when implementing such systems.

- The splicer needs to be DSM-CC aware and splice the carousel in such a way that the decoder does not receive partial or corrupted data.
- The splicer needs to take in to account the PCR-related time-base information associated with/carried in the carousel to tell the decoder when to render the data.

Because DSM-CC carousels are conveyed in Sections, not PES packets, conventional PTS/DTS fields cannot be used.

DSM-CC defines the concept of Normal Play Time (NPT), which is used to describe all "stream events" and uses NPT Reference Descriptors to lock NPT to the PCR. If the splicer re-stamps PCRs and PTS/DTSs, then it also needs to restamp NPT Reference Descriptors if this method of synchronisation is used.

The draft ATSC data broadcasting specification carries PTS/DTS for the data objects in a header in the carousel. In this case, if the splicer re-stamps PCRs and PTS/DTSs, then it needs to also be capable of re-stamping those in the DSM-CC carousel.

ENCRYPTED BITSTREAMS

Bitstreams which are scrambled at the transport layer have the entire PES packet including PES header scrambled. It is thus not possible to read or change any of the content. This means that only non-seamless splicing, with timebase discontinuities, is possible with scrambled bitstreams. For either seamless or near-seamless splicing, the bitstream needs to be in the clear, i.e. if conditional access is required, scrambling takes place after splicing.

REAL-TIME CONTROL

Unlike uncompressed signals, video, audio and data signals which are to be presented simultaneously occur in the compressed bitstream at very different times. For example, compressed video is transmitted well ahead of the associated compressed audio due to the longer decoding delays.

Thus, manual control of splicing is becoming extremely difficult unless the operator can see the program prior to compression. Even then the actual splicing point has to be delayed in line with the encoder delay which is variable. If the operator only has access to the decoded program, there will be a variable delay between the time when the 'splice' button was pressed and the point where the splice actually occurs, depending on decoder buffer occupancy at the splice point. For these reasons work is concentrating on automated splicing according to predetermined schedules.

Figure 1 shows a block diagram of a possible configuration of a local station. In this diagram the transport stream splicer switches between live-to-air local programs 1 and 2, pre-compressed files from the server and MPEG-2 network feeds. The splicer is controlled by a local control computer with scheduling information.

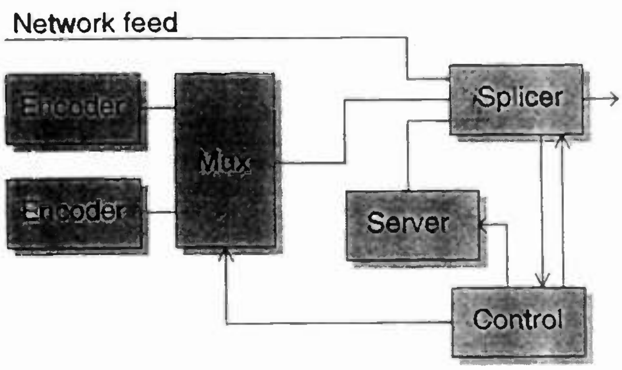


Figure 1 **Block diagram of local station with transport stream splicer and scheduling control.**

The proposed SMPTE splicing standard, SMPTE-312 defines a means for a network to download local opt-outs schedules using splice_info_sections that are carried as part of the program and defined as such in the PMT. This data is then extracted from the network feed by the splicer and fed to the local control computer. The splicer then modifies the PMT to remove any reference to the splice_info PID, and also deletes the PID from the outgoing transport stream to defeat "commercial killers".

The local station staff then define which of the scheduled spice-points in the network feed they want to use, and what source the splicer should splice to when the splice-point occurs. This data is then uploaded to the splicer.

On the same PID as the splice schedule, the network can send "pre-roll" and "execute" commands that inform the splicer when the various splices in the network feed are approaching. This is a digital version of audio cue-tones used by many (cable) networks in today's analogue world.

CONCLUSIONS

Issues concerning MPEG real-time splicing have been discussed. It has been demonstrated that there is more to MPEG splicing than solving video buffer management problems. In particular, the requirement to splice streams other than video and audio puts more demands on the splicer, and broadcasters should ensure that the splicer they purchase has the required functionality. Also the question of real-time splicer control will need careful attention of system designers and new ways to operate (local) broadcast stations. Nevertheless, transport stream splicing, properly integrated into the rest of the digital broadcast chain, can provide compact and cost-effective solutions for today's broadcasters.

REFERENCES

1 O'Grady, W., Balakrishnan, M., Radha, H., Real-time switching of MPEG-2 bitstreams, Proceedings of 1997 International Broadcasting Convention, pp. 166 to 170.

2 McGrath, E., Digital insertion of advertising into a digital stream (DID), Proceedings of 1997 International Broadcasting Convention, pp. 258 to 261.

3 Birch, C.H., MPEG splicing and bandwidth management, Proceedings of 1997 International Broadcasting Convention, pp. 541 to 545.

Emerging Technology for Content Based Access of Digital Video

Michael A. Smith Robert Mozenter
ISLIP Media, Inc.
Pittsburgh, Pennsylvania

ABSTRACT

Content-based access of digital video is possible through recent advances in image and audio understanding technology. Automatic indexing and retrieval of digital video allows users to process vast collections of data in a short time and with less manual effort. In the broadcast industry, the ability to search and retrieve content and information will be required to reuse and resell content that would otherwise remain locked in storage. Broadcasters will leverage new gains in content-based access to better affect their bottom line.

1.0 INTRODUCTION

Image understanding technology provides robust systems for identifying important video content, such as, scene changes, camera motion, video captions and the presence of human faces. Advances in audio understanding technology enable users to verbally annotate video segments and provide an accurate and efficient means for converting audio into textural transcripts. There are many practical applications as well as limitations for this technology in the broadcast industry. Content-based retrieval techniques will be utilized to establish large and extremely valuable archives of information that can be used for information retrieval, data mining and content reuse. We will describe the current state of the art for automated extraction of image and audio content from video, as well as current research in this area.

2.0 MOTIVATION

Unlike text or images, video must be encoded, cataloged and indexed so that lengthy analog tapes can be broken down into a small "clips" searchable by the following:

- Image (drag a picture of a tree into a window and the software will find video containing similar trees),
- Text (enter a name, subject or phrase – similar to the search method on search engines such as Yahoo or Lycos)
- Fielded search (pick choices from a list, such as: subject, description, lighting, angle, format, time, date, reporter, or location).

Video Indexing - The ubiquity of the Internet and corporate Intranets have offered us multiple pathways to disseminate information in addition to the network broadcast and cable television channels. The main difference between the two is the notion of broadcast video (fixed time 30 to 60 minutes) and on-demand video (specific information regardless of time and location). In order to serve the on-demand users, we need to quickly find and deliver the most relevant portion of the video and not a 30 or 60 minute program. Therefore, it is necessary to consider effective ways and means to index and store video so that it can be retrieved instantly on-demand and shared.

The effort and frustration involved in locating video information is enormous today. The onslaught of new tools and technologies to create and share even more information makes the magnitude of this data management problem even more difficult. The information architects of the future have to also focus on providing tools and methodologies to efficiently navigate the video information. Proper indexing is the core ingredient of this effort.

Feature Film Production - One of the largest expenses in the feature production industry is the cost of processing film. The ratio of footage shot to footage used in final production is typically between 10:1 and 40:1. With such a wide shooting ratio, many financial and labor resources could be saved if production houses could print only the footage that will actually be used in the feature.

But because they lack an effective system for logging, managing, sorting, and proofing footage before it is printed, most production houses spend resources processing all of the footage shot in a day—even bad takes.

In addition to managing daily footage, production houses are also challenged with the problem of storing footage from past shoots. This problem drains resources as well. For example, production houses often spend money and man-hours reshooting scenes, sometimes in expensive, remote locations, simply because previously shot footage with similar content cannot be located. Without a system to archive and index past footage and make it accessible, there is no return on work invested in past projects.

Currently, production houses lack an inexpensive and efficient mechanism for sorting and tracking daily footage as well as an effective means of searching archives to repurpose previously processed footage. Without such a system, production houses will continue to incur costs needlessly.

Stock Footage Delivery - Today, when customers need stock footage, they typically contact a house that offers the type of footage they are looking for. Stock footage researchers then look through their collections and send a preview tape (also called window dub) to the customer via mail. A day (or several days later) the customer receives and reviews the tape. If they like any of the content, they order a high-resolution version (again via mail). If not, the process repeats itself. In a deadline driven business, this process is tedious and cumbersome.

Cataloging, search, and retrieval of stock footage via Intranets, Extranets or the Internet saves significant time and increases productivity. Users can search thousands of hours of video content and instantly receive video clips. For example, users can type in keywords or select visual criteria, such as image, color or camera angle to return the exact content they are looking for. Depending upon the format or resolution chosen, digital or analog versions of the purchased clips are instantly viewed, downloaded or sent to the user.

Broadcast News - The current amount of stored video footage in the news industry amounts to over an astounding 10 million hours. Simply managing the physical inventory of that video is an enormous task. And, the amount of video continues to grow, as stations broadcast up to 20 hours daily and receive incoming video feeds which can amount to over 300 hours per day. While all broadcasters manage these video assets in some way, the requirement for on-time, up-to-the-minute news has put considerable strain on today's outdated management systems. To remain competitive and offer viewers the most in-depth coverage possible, news organizations must have immediate access to both live, incoming video as well as archived video from previous stories.

With the advent of networked computers, faster PCs and greater storage capacity, news broadcasters now have the capability to easily and inexpensively convert archived and incoming analog video to a digital format searchable over standard PCs or the World Wide Web.

Internet Video Delivery - Today, finding and delivering video content over the Internet is problematic, with no way to search or preview video content without downloading entire files. With segmented video, you can preview video clips, easily navigating through search results and downloading only the video you want.

3.0 IMAGE TECHNOLOGY

Certain image-based features may be extracted without rigorous analysis of the actual content from the video. These features include such analytical features as scene changes, motion flow and known production structure in such formats as news video.

In addition to analytical methods, many image processing systems approximate the actual content of an image or video. For many users, the query of interest is text based, and therefore, the content is essential. The desired result has less to do with analytical features such as color, or texture, and more with the actual objects within the image or video.

3.1 SCENE SEGMENTATION

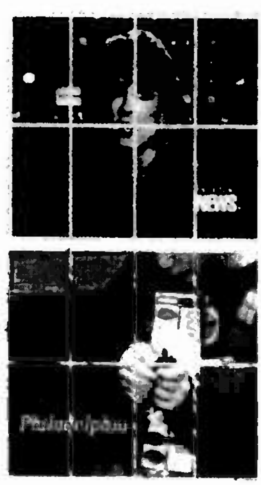
Scene changes are used to separate visual content in video. The most fundamental scene change is the video cut. For most cuts, the static difference between image frames is so distinct that accurate detection is not difficult. Cuts between similar scenes, however, may be missed when using only static properties. There are a variety of more complex scene changes used in video production, but the basic premise is a change in visual content. The video cut, as well as other scene change procedures are discussed below.

Dissolves and Fades – Dynamic imaging effects are often used to change from one scene to another. A common effect in all types of video is the Fade. A Fade occurs when a scene changes over time from its original color scheme to a black background. This procedure is commonly used as a transition from one topic to another. Another dynamic effect is the Dissolve. Similar to the Fade, this effect occurs when a scene changes over time and morphs into a separate scene. This transition is less intrusive and is used when subtle change is needed.

Wipes and Blends - These effects are most often used in news video. The actual format of each may change from one show to the next. A wipe usually consists of the last frame of a scene being folded like a page in a book. A blend may be shown as pieces of two separate scenes combining in some artistic manner. Like the fade and dissolve, wipes and blends are usually used for transition to a separate topic, therefore detection is extremely important.

Several research groups have developed working techniques for detecting scene changes [Arman94], [Zhang93], [Hampapur95]. Several survey publications for image and video access methods will be available in the near future.

A histogram difference is less sensitive to subtle motion, and is an effective measure for detecting scene cuts and gradual transitions [Arman94], [Zhang93]. By detecting significant changes in the weighted color histogram of each successive frame, video sequences can be separated into scenes. In the difference, *D(t)*, peaks are detected and an empirically set threshold is used to select scene breaks. This technique is simple, and yet robust enough to maintain high levels of accuracy for our purpose. Using this technique, many researchers have achieved high accuracy in scene



$$D(t) = \sum_{v=0}^{N} | H_t(v) - H_{t+1}(v) |$$

$H_t(v)$: Histogram of Color in Image(t)

Figure 1, Histogram Scene Segmentation

segmentation. Figure 1 is an example of histogram-based segmentation, with the image divided into sub-regions. Sub-region analysis provides segmentation that is less sensitive to motion.

Edge Based Segmentation

An alternative form of scene segmentation involves the use of traditional edge detection characteristics. Edges in images are useful information about the changes in background and object distribution between scenes. An effective algorithm for detecting cuts and gradual transitions was developed using edge detection technology [Zabih95].

Motion Based Segmentation

An analysis of the global motion of a video sequence may also be used to detect changes in scenery. When the error in optical flow is about some threshold, this is usually attributed to its inability to track a majority of the motion vectors from one frame to the next. Such errors can be used to signify scene changes. A motion-controlled temporal filter may also be used to detect dissolves and fades.

3.2 MOTION ANALYSIS

An analysis of optical flow in video sequences can serve as an additional feature for retrieval. There are many algorithms with different levels of emphasis, such as, processing speed, and pixel accuracy. Most of these algorithms require some computation, and more often, researchers are exploring methods to extract optical flow from video compressed with some form of motion compensation. Section 4 describes the benefits of using compressed video for optical flow and other image features. Optical flow fields may be interpreted in many ways to estimate the imagery in video. Two such interpretations are the action of camera and object motion.

Camera Motion

One important aspect of video characterization is based on interpreting camera motion [Akutsu94b]. Many scenes have beautiful visual effects, but offer little in the description of a particular segment. Static scenes, such as interviews and still poses, contain essentially identical video frames. Since the skim must congregate large amounts of visual information into a short sequence, we avoid video with excess camera motion or visual redundancy.

Knowing the precise location of camera motion will provide another tool for video segmentation. Fast camera shots often appear as scene breaks when only the intensity distribution is measured. The quality of the camera motion estimate is used to detect video with extreme motion.

Object Motion

An important kind of video characterization is defined not just by motion of the camera, but also by motion or action of the objects being viewed. The global distribution of motion vectors distinguishes between object and camera motion. Object motion typically exhibits flow fields in specific regions of an image, while camera motion is characterized by flow throughout the entire image.

3.3 OBJECT RECOGNITION

Identifying significant objects that appear in the video frames is one of the key components for video characterization. Several working systems have generated reasonable results for the detection of a particular object, such as human faces, text, or automobile. These limited domain systems have much greater accuracy than do broad domain systems that attempt to identify any object in the image.

Human Faces

The "talking head" image is common in interviews and news clips, and illustrates a clear example of video production focusing on an individual of interest. A human interacting within an environment is also a common theme in video. The detection of humans in video is possible using a number of algorithms. Most techniques are dependent on the face size, and rely heavily on lighting conditions, limited occlusion, and limited facial rotation. Current research has shown promise in the ability to detect faces of different orientation, scale, position [Rowley95]. An example of the results from the Neural Network face detection system by Rowley is shown in Figure 2. Companies such as Visionics have produced products that offer face matching and detection functionality.

Captions

Text and graphics are used in a variety of ways to convey content to the viewer. They are most commonly used in news broadcast, where information must be absorbed in a short time. Examples of text and graphics in video are discussed below. An example of an image with captioned text and graphics is shown in Figure 2.

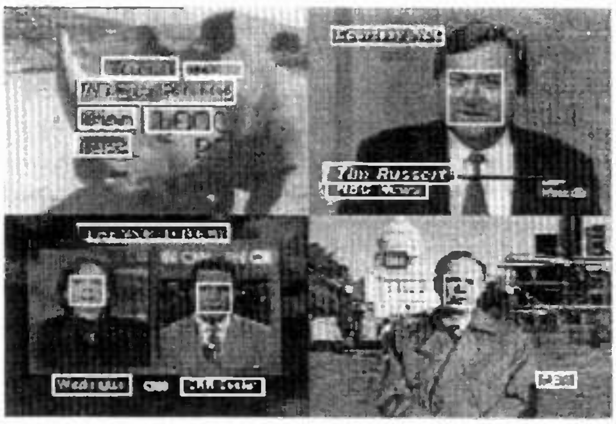


Figure 2, Video Captions and Face Detection

Video Captions - Text in video provides significant information as to the content of a scene. For example, statistical numbers and titles are not usually spoken but are included in captions for viewer inspection. Moreover, this information does not always appear in closed captions.

In news video, captions of the broadcasting company are often shown at low opacity in a corner without obstructing the actual video. A ticker-tape is widely used in news broadcast to display information such as the weather, sports scores, or the stock market. In some news broadcast, graphics such as weather forecast are displayed in a ticker-tape format with the news logo in the lower right corner at full opacity. Captions that appear in the lower third portion of a frame are almost always used to describe a location, person of interest, title, or event in news video. In Figure 2, the subject's names and affiliations are listed.

Captions are used less frequently in video domains other than broadcast news. In sports, a score or some information about an ensuing play is often shown in a corner or border at low opacity. Captions are sometimes used in documentaries to describe a location, person of interest, title, or event. Almost all commercials use some form of captions to describe a product or institution, because their time is limited to only 30 to 50 seconds.

Most captions use high contrast text such as the black and white chyron commonly found in news video. Consistent detection of the same text region over a period of time is probable since text regions remain at an exact position for many video frames. This may also correct for the false detection of text regions that move or fade in and out when captions are placed in a scene.

For some fonts a generic optical character recognition (OCR) package may accurately recognize video captions. For most OCR systems, the input is an individual character. This presents a problem in digital video since most of the characters show some degradation during recording, digitization and compression. For a simple font, we can search for blank spaces between characters and assume a fixed width for each letter [Sato98].

Graphics

A graphic is usually a recognizable symbol, which may contain text. Graphic illustrations or symbolic logos are used to represent many institutions, locations, and organizations. They are used extensively in news video, where it is important to describe the subject matter as efficiently as possible. A logo representing the subject is often placed in a corner next to an anchorperson during dialogue. Detection of graphics is a useful method for finding changes in semantic content. In this sense, its appearance may serve as a scene break. Recognition of corner regions for graphics detection may be possible through an extension of the scene change technology. Histogram analysis of isolated image regions instead of the entire image can provide a simple method for detecting corner graphics.

Articulated Objects
A particular object is usually the emphasis of a query in image and video retrieval. Recognition of articulated objects poses a great challenge, and represents a significant step in content based feature extraction. Many working systems have demonstrated accurate recognition of animal objects, segmented objects, and rigid objects such as planes or automobiles. Research at the University of California at Berkeley, University of California Santa Barbara, and the University of Kentucky has shown promising results in the recognition of complex objects.

The recognition of a single object is only one potential use of image based recognition systems. Discrimination of synthetic and natural backgrounds, or an animated or mechanical motion would yield a significant improvement in content based feature extraction.

3.4 IMAGE STATISTICS
A variety of image statistics may be used without extensive processing of the image content. An analysis of the shape and contours of an image can provide useful information to the content. This is particularly the case in video with rigid object models, such as military footage.

Analysis of image texture is useful in the discrimination of low interest video from video containing complex features. A low interest image may also contain uniform texture, as well as uniform color or low contrast. Perceptual features for individual video frames were computed using common textual features, such as coarseness, contrast, directionality, and regularity.

4.0 AUDIO TECHNOLOGY
In addition to image features, certain audio features may be extracted from video to assist in the retrieval task. Loud sounds, silence, and single frequency sound markers may be detected analytically without actual knowledge of the audio content. Loud sounds imply a heightened state of emotion in video, and are easily detected by measuring a number of audio attributes, such as signal amplitude or power. Silent video may signify an area of less importance, and can also be detected with straight-forward analytical estimates. A video producer will often use single frequency sound markers, typically a 1000 Hz. tone, to mark a particular point in the beginning of a video. This tone may be detected to determine the exact point in which a video will start.

Audio segmentation is often needed to distinguish spoken words from music, noise and silence. Further analysis through speech recognition is necessary to

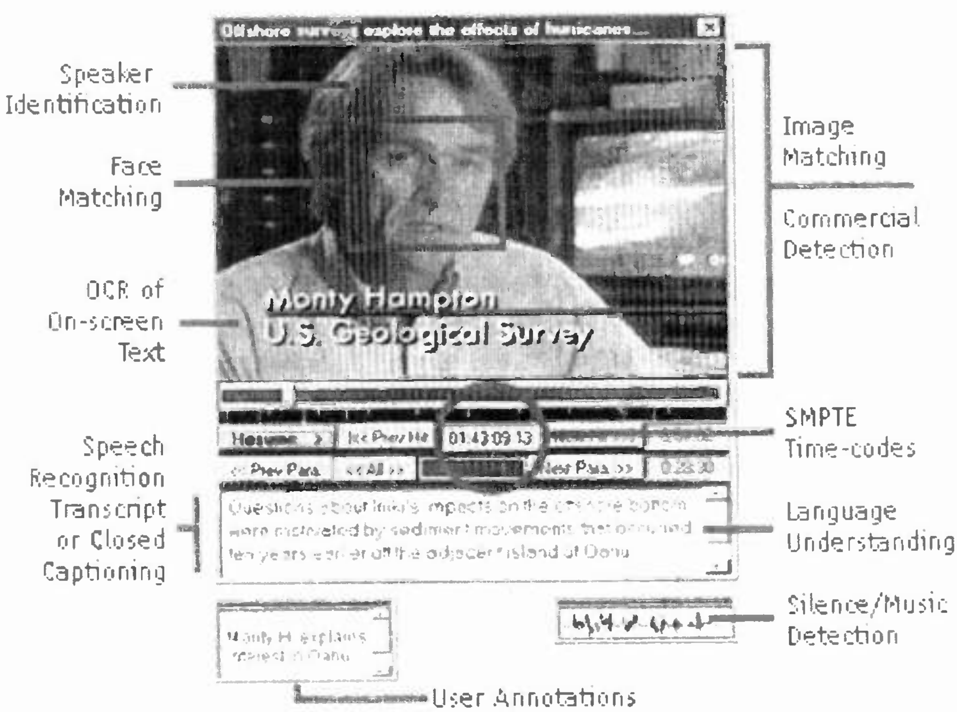


Figure 3, Video Meta-data

align and translate these words into text. For many video editors, audio selection is made on a frame by frame basis, so it is important to achieve the highest possible accuracy. At a sampling rate of 8Khz, one frame corresponds to 267 samples of audio. Techniques in language understanding may then be used for selecting the most significant words and phrases.

4.1 SPEECH RECOGNITION

An important element in video indexing creation is the audio track. Audio is an enormous source for describing video content. Words specific to the actual content, or "Keywords" can be extracted using a number of language processing techniques [Mauldin91], [Salton83]. Keywords may be used to reduce indexing and provide abstraction for video sequences [Wactlar96], [Smith98]. There are many possibilities for language processing in video, but the audio track must first exist as an ASCII document or speech recognition is necessary.

In order to use the audio track, we must isolate each individual word. To transcribe the content of the video material, we recognize spoken words using a speech recognition system. Speaker independent (any user) and speaker dependent (single user) recognition systems have made great strides as of late and offer promise for application in video indexing [Frowein91], [Hauptmann97]. Speech recognition works best when closed-captioned data is available. Captions usually occur in broadcast material, such as sitcoms, sports, and news. Documentaries and movies may not necessarily contain captions. Closed-captions have become more common in video material throughout the United States since 1985 and most televisions provide standard caption display.

4.2 LANGUAGE ANALYSIS

Natural language processing is applied to understand and expand the user's query and to associate it with correct but inexact matches from the library's content to go beyond limited keyword matching in the search [TREC93], [Salton83], [Mauldin91]. Natural language processing is applied to both query processing and during library creation – spoken and typed free-form query processing, ranked retrieval and summarization for use in title generation and video abstract. Many retrieval engines implement a probabilistic matching to return a rank-ordered result list. Varying relative thresholds enables either precision or recall to be adjusted by the user. It also enables multiple types of similarity matching and tolerates errors in speech recognition of the spoken query and generated transcripts.

5.0 VIDEO META-DATA

Sections 3 and 4 describe methods for creating image and audio meta-data through automated techniques in image, language and speech processing. An example of the types of meta-data that may be extracted with this technology is shown in figure 3. In this section we discuss alternative forms of meta-data that do not require extensive processing. In many cases, the data described in this section may be extracted from previously derived manual data, such as geo-spatial information, and video production structure and standards.

Global Position Systems (GPS)

The incorporation of explicit or implied geo-references is a useful form of meta-data for video indexing and retrieval. In its simplest form this means the inclusion of the corresponding GPS location data as an additional dimension and modality of information search. This "location data" corresponding to each video segment is represented as a single, set or range of values. The user can include a named location or location coordinates in the query and search for events at that location or within some "distance" of that location. The distance and location may also be expressed as a region, and refer synonymously, or hierarchically, to political or geographically defined boundaries that determine a region (e.g. queries about Tuzla also optionally refer to the Balkan region or Yugoslavia (historically)).

The geo-spatial location information can be captured through alternate means:

- GPS data captured concurrently with the audio and video content and embedded within.
- Human annotations added to the video information as named locations or coordinates.

Production Standards

Video production manuals provide insight into the procedures used during video editing and creation

[Bordwell93], [Pryluck82]. One of the most common elements in video production is the ability to convey climax or suspense. Producers use a variety of different effects; ranging from camera positioning, lighting, and special effects to convey this mood to an audience. Detection of these procedures is beyond the realm of present image and language understanding technology. However, the previous sections describe technology that represents a subset of basic editing techniques.

Video Structure from Content

Structural information as to the content of a video is a useful tool for indexing video. For example, the type of video being used (documentaries, news footage, movies and sports) and its duration may offer suggestions to assist in object recognition.

There are also many visual effects introduced during video editing and creation that may provide information for video content. The scenes prior to the introduction of a person usually describe their accomplishments and often precede scenes with large views of the person's face. A person's name is generally spoken and then followed by supportive material. Afterwards, the person's actual face is shown.

6.0 Content Based Matching

Content matching as opposed to feature matching attempts to correlate actual objects with a given query. In news footage, the anchorperson will generally appear in the same pose and background at different times. The exact locations of the anchorperson can then be used to delineate story breaks. In documentaries, a person of expertise will appear at various points throughout the story when topical changes take place.

Several Working systems have demonstrated the potential of content based matching for identifying specific objects and stories. Three of the more interesting systems are discussed below.

- Name-It, is a system for matching a human face to a name in news video [Satoh97]. It approximates the likelihood of a particular face belonging to a name in close proximity within the transcript. Integrated language and image understanding technology make the automation of this system possible.

- Spot-it, is a topological system that attempts to identify known characteristics in news video for indexing and classification [Nakamura97]. It has reasonable success in identifying common video themes such as interviews, group discussions, and conference room meetings.

- Pictorial Transcripts, a working system at AT&T Research Laboratories has shown promising results in video summarization when closed-captions are used with statistical visual attributes [Shahraray95]. CNN video is digitized and displayed in an HTML environment with text for audio and a static image for every paragraph.

Queries: Image or Text

For most image and video retrieval systems, the query is an image. When the comparison is based on analytical features, the results can often be ambiguous. Content based features provide a more accurate match to the given query, but the results are based on image processing technology which is currently capable of recognizing a specific type of object.

Text queries eliminate ambiguity in the query, and work only with content based features. There is still a dependence on content based feature extraction, but there is limited uncertainty in the query. This type of the query may also be used to match the title of the image or the transcript of the video.

Video Access and Browsing

With the size of the video collections growing to thousands of hours, technology is needed to effectively browse segments in a short time without losing the content of the video. Simplistic browsing techniques, such as increased playback speed and skipping video frames at fixed intervals, reduce video viewing time. However, increased video rates eliminate the majority of the audio information and distort much of the image information; and displaying video sections at fixed intervals merely gives a random estimate of the overall content. An ideal browser would display only the video pertaining to a segment's content, suppressing irrelevant data.

A multimedia abstraction ideally preserves and communicates the essential content of a video segment via a compact representation. Examples of multimedia abstractions include short text titles and single thumbnail images. Another commonly used abstraction presents an ordered set of representative, "thumbnail" images simultaneously on a computer screen [Arman94], [Mills92], [Taniguchi95], [Rorvig93], [Zhang94], [Zhang95a]. Image statistics, such as histogram analysis and texture, camera structure and scene changes are the dominant factors in these systems. While these abstractions have proven useful in various contexts, their static nature ignores video's temporal dimension.

Recently, researchers have proposed browsing representations based on information within the video [Tonomura94], [Zhang95b]. These systems rely on the motion in a scene, placement of scenes breaks, but not on integrated image and language understanding.

In addition, these abstractions often concentrate exclusively on the image content and neglect the audio information carried in a video segment. There are a number of efforts that combine language and image understanding as of late. The application of technology integration is different for these systems, however, they all demonstrate the advantages of using multiple modalities in video characterization and summarization. Examples of these systems are discussed below:

- Video Skimming - The video skim was the first system to integrate technology in image, language, and audio understanding for browsing and summarization [Smith97]. It identifies significant image and audio regions in the video and produces a compact representation without apparent loss in content. The results of this work were tested with several user studies.

- Browsing through Clustering - This system was designed to cluster image regions for browsing digital video [Yeung95]. It uses many of the image statistics mentioned earlier, but it attempts to process scene transitions rather than just process individual frames.

- High Rate Keyframe Browsing - The Digital Library Research Group at the University of Maryland, College Park, MD, has conducted a user study to test optimal frame rates for keyframe based browsing. They use many of the same image analysis techniques mentioned earlier to extract keyframes, and they quantify their research through studies of a video slide show interface at various frame rates.

- Video Abstracts - The Movie Content Analysis (MoCA) group in Mannheim, Germany has created a system for movie abstraction based on the occurrence of image statistics and audio frequency analysis to detect dialogue scenes [Pfeiffer96].

Video Library Case Studies

There are many researchers working in the area of image matching. A few systems with unique characteristics are listed below.

- UC Berkeley - Object extraction and recognition system.
- UC Santa Barbara -Image matching system based on region segmentation.
- Carnegie Mellon University - The Informedia digital video library project Wactlar96] has established a large, on-line digital video library by developing intelligent, automatic mechanisms to populate the library and allow for full-content and knowledge-based search and retrieval via desktop computer over local, metropolitan, and wide-area networks. Their approach utilizes several techniques for content-based searching and video sequence retrieval. Content is conveyed in both the narrative (speech and language) and the image. The collaborative interaction of image, speech and natural language understanding technology allows for successful population, segmentation, indexing, and search of diverse

video collections with satisfactory recall and precision.

- IBM - One of the first well-known image matching system, Query by Image Content, or QBIC. This provides fast indexing through condensed hierarchical tree structure.
- VIRAGE - Image and video retrieval company from research at the University California, San Diego.
- ISLIP Media - Multimedia Data indexing and content retrieval through mass storage systems and Internet distribution. Primary technology licensed from Carnegie Mellon University Informedia Project. An example of the interface is shown in Figure 4.

7.0 COMPRESSED DOMAIN FEATURES

In typical applications of multimedia databases, the video is compressed. To process this video, a straightforward approach is to decompress all the data, and utilize the same features as mentioned in previous sections. This has certain important disadvantages. First, the decompression, and subsequent recompression require extra computation. Second, the process of decompression and re-compression, often referred to as "recoding," results in lost of image quality. Finally, since the size of decompression data is much larger than the compressed form, most operations become much heavier when applied to the decompressed data.

The solution to these problems is to extract features directly from the compressed data [Chang95]. We

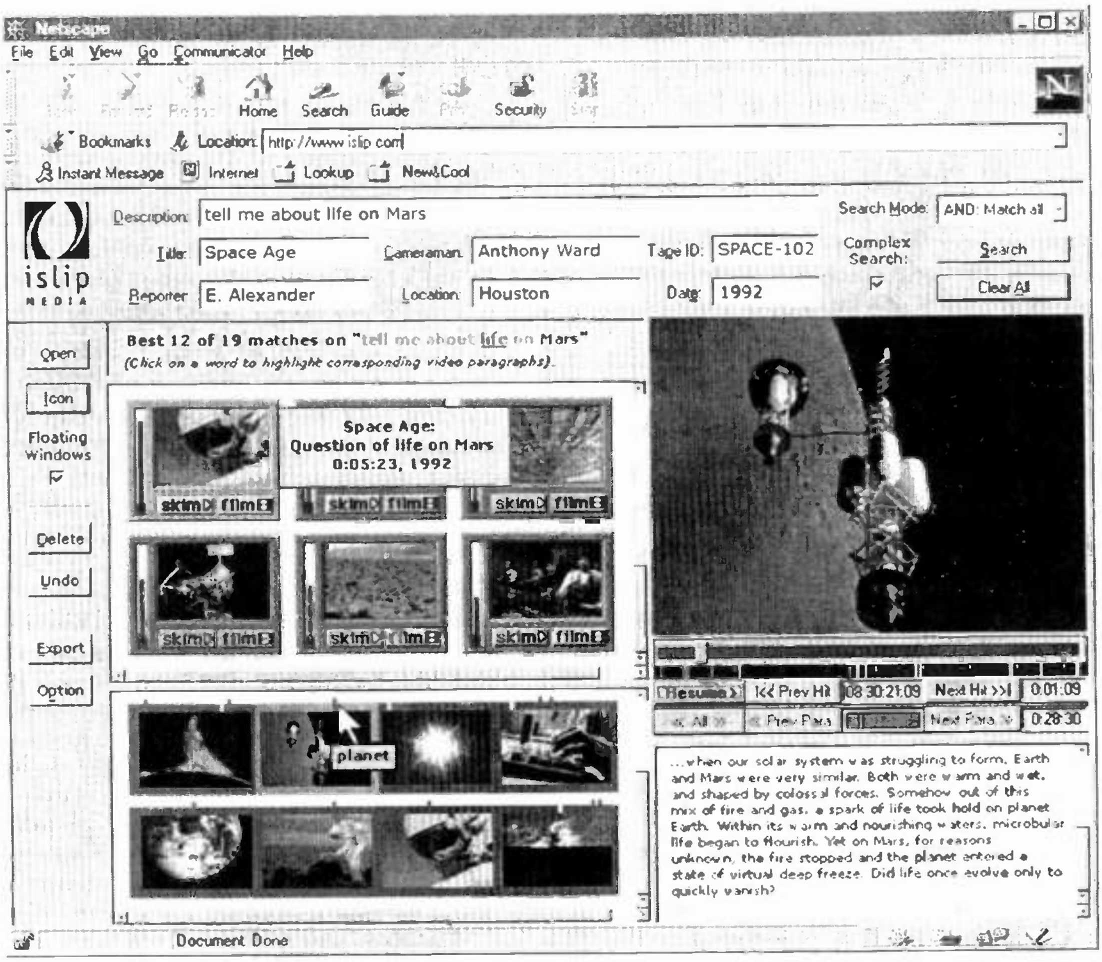


Figure 4, Video Retrieval User Interface

call these the compressed-domain features. Commonly used compressed-domain features include:

- The motion vectors that are available in all video data compressed using standards such H.261/H.263 and MPEG-1/2 are very useful. They can be used to detect scene changes and other special effects such as dissolve, fade in and fade out. They can be used to detect moving objects and track their positions. They can also be used to derive camera motion such as zoom and pan [Akutsu94a], [Tse9].
- DCT (Discrete Cosine Transform) provides a decomposition of the original image in the frequency domain. Therefore, DCT coefficients form a natural representation of texture in the original image. In addition to texture analysis, DCT coefficients can also be used to match images and to detect scene changes.

The compressed-domain approach does not solve all problems. Each compression technique poses additional constrains, e.g., non-linear process, rigid data structure syntax, resolution reduction. The compressed-domain approach provides significant advantages but also brings new challenges.

8.0 VIDEO STANDARDS

MPEG-7, is an ongoing effort by the Moving Picture Experts Group towards the standardization of meta-data for multimedia indexing and retrieval. MPEG-7 is an activity that is triggered by the growth of digital audiovisual information. The group strives to define a "Multimedia Content Description Interface" to standardize the description of various types of multimedia content, including still pictures, graphics, 3D models, audio, speech, video, and composition information. It may also deal with special cases such as facial expressions and personal characteristics.

The goal of MPEG-7 is to enable efficient search and retrieval of multimedia content. Once finalized, it will transform the text-based search and retrieval (e.g., keywords) as is done by most of the multimedia databases today, into a content-based approach, e.g., using color, motion, or shape information. MPEG-7 can also be thought of as a solution to describing multimedia content. If one looks at PDF (Portable Document Format) as a standard language to describe text and graphic documents, then MPEG-7 will be a standard description for all types of multimedia data, including audio, images, and video. There is also a video meta-data standard being developed by the Society of Motion Picture and Television Engineers (SMPTE), which addresses many of the same concerns as MPEG-7.

9.0 ACKNOWLEDGEMENTS

We wish to thank Tsuhan Chen of the Department of Electrical and Computer Engineering at Carnegie Mellon University for his contribution to the Compressed Domain and MPEG7 sections of this paper.

10 CONCLUSION

Digital library systems have traditionally used statistical analysis of textual descriptors to access relevant data. Video provides a medium with image, language and audio content, in which advances in content-based analysis may be used. With an increase in feature-based analysis and extraction, image and audio understanding systems are becoming usable and efficient in retrieving perceptual content. Powerful feature-based indexing and retrieval tools can be developed for image/video archives, complementing the traditional text-based techniques. In the future, these technologies will allow for more advanced systems such as searchable television and content-based video-on-demand.

11 REFERENCES

[Akutsu94a] A. Akutsu, *et al*, "Video Indexing Using Motion Vectors," *SPIE Visual Communications and Image Processing* 1992, Vol. 1818, pp. 1522-1530.

[Akutsu94b] Akutsu, A. and Tonomura, Y. "Video Tomography: An efficient method for Camerawork Extraction and Motion Analysis," *Proc. of ACM Multimedia '94*, Oct., 1994, San Francisco, CA, pp. 349-356.

[Arman94] F. Arman, R. Depommier, A. Hsu, and M. Y. Chiu. "Content-based browsing of video sequences," *Proc. of ACM Multimedia '94* , October, 1994, San Francisco, CA pp. 97-103.

[Bordwell93] Bordwell,D., Thompson, K., "Film Art: An Introduction," 4th Ed., McGraw Hill, Englewood Cliffs, NJ, 1993.

[Chang95] S.-F. Chang, "Exploring Functionalities in the Image/Video Compressed Domain," ACM Computing Surveys, Vol. 27, No. 4, Dec. 1995.

[Frowein91] Frowein, et. al, "Improved speech recognition through videotelephony: experiments with the hard of hearing," IEEE Journal on Selected Area in Communication, vol. 9, no, 4, May 1991.

[Hampapur95] Hampapur, A., Jain, R., and Weymouth, T. Production Model Based Digital Video Segmentation. Multimedia Tools and Applications, 1 (March 1995), 9-46.

[Hauptmann97] Hauptmann, A.G., and Witbrock, M.J. Informedia: News-on-Demand Multimedia Information Acquisition and Retrieval. In Intelligent Multimedia Information Retrieval, M. Maybury, Ed. 1997, AAAI Press/MIT Press: Menlo Park, CA.

[Mauldin91] Mauldin, M. "Information Retrieval by Text Skimming," PhD Thesis, Carnegie Mellon University. August 1989. Revised edition published as "Conceptual Information Retrieval: A Case Study in Adaptive Partial Parsing, Kluwer Press, September 1991.

[Mills92] Mills, M., Cohen, J., and Wong, Y.Y. A Magnifier Tool for Video Data. In Proceedings of the ACM CHI'92 Conference on Human Factors in Computing Systems. (Monterey, CA, May 1992), 93-98.

[Nakamura97] Y. Nakamura. ..."Spot-It, Association of Objects and Text" *Proceedings of ACM Multimedia 1997*, October, 1997, Seattle, WA.

[Pfeiffer96] Pfeiffer, S. Lienhart. R., Fischer, S., Effelsberg, W., "Abstracting Digital Movies Automatically," *Journal of Visual Communication and Image Representation*, Vol. 7, No. 4, pp. 345-353, December 1996.

[Pryluck82] Pryluck, C., Teddlie, C., Sands, R., "Meaning in Film/Video: Order, Time and Ambiguity," *Journal of Broadcasting* 26, pp. 685-695, 1982.

[Rorvig93] Rorvig, M.E., "A Method for Automatically Abstracting Visual Documentaries," *Journal for the American Society of Information Science* 44, 1, 1993.

[Rowley95] Rowley, H., Baluja, S. and Kanade, T. Human Face Detection in Visual Scenes. Carnegie Mellon University, School of Computer Science Technical Report CMU-CS-95-158 (Pittsburgh, PA, 1995).

[Sato98] Sato, T., Kanade, T., Hughes, E., Smith, M., IEEE Workshop on Content-Based Access of Image and Video Databases (CAIVD'98), Bombay, India, January, 1998.

[Satoh97] "NAME-IT: Association of Face and Name in Video." Sato, S., Kanade, T., In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Juan, Puerto Rico, 17-19 June, 1997.

[Salton83] Salton, G., and McGill, M.J. "Introduction to Modern Information Retrieval," McGraw-Hill, New York, McGraw-Hill Computer Science Series, 1983.

[Shahraray95] Shahraray, B., Gibbon, D., "Authoring of Hypermedia Documents of Video Programs," *Proceedings of the Third ACM Conference on Multimedia*, pp. 401-409, San Francisco, CA, November, 1995.

[Smith97] Smith, M., Kanade, T. Video Skimming and Characterization through the Combination of Image and Language Understanding Techniques. In Computer Vision and Pattern Recognition. (San Juan, PR, 1997).

[Taniguchi95] Taniguchi, Y., *et al*. An Intuitive and Efficient Access Interface to Real-Time Incoming Video Based on Automatic Indexing. In Proceedings of the ACM Multimedia Conference. (San Francisco, CA, November 1995), 25-33.

[Tonomura94] Tonomura, Y., Akutsu, A., Tonomura, Y., Hamada, H., "Structured Video Compting," *IEEE Multimedia Magazine* 1, pp. 34-43, 1994.

[Tse91] Y.T. Tse, and R. L. Baker, "Global Zoom/Pan Estimation and Compensation For Video Compression" *Proceedings of ICASSP* 1991, pp.2725-2728.

[TREC93] "TREC 93," *Proceedings of the Second Text Retrieval Conference*, D. Harmon, editor, sponsored by ARPA/SISTO, August 1993.

[Wactlar96] Wactlar, H.D., Kanade, T., Smith, M.A., and Stevens, S.M. "Intelligent Access to Digital Video: Informedia Project," *IEEE Computer*, 29, 5 (May 1996), 46-52.

[Yeung95] Yeung, M., Yeo, B., Wolf, W., and Liu, B. Video Browsing Using Clustering and Scene Transitions on Compressed Sequences. *Proceedings IS&T/SPIE Multimedia Computing and Networking*, February, 1995.

[Zhang93] Zhang, H., Kankanhalli, A., and Smoliar, S. "Automatic partitioning of full-motion video," *Multimedia Systems* 1993 1, pp. 10-28.

[Zabih95] Zabih, R., Miller, J., Mai, K., "A Feature-Based Algorithm for Detecting and Classifying Scene Breaks," *Proceedings of the ACM International Conference on Multimedia,* San Francisco, CA, November, 1995.

[Zhang95a] Zhang, H.J., Smoliar, S., and Wu, J.H., "Content-Based Video Browsing Tools," *Multimedia Computing and Networks,* . Multimedia Systems, 2, 6 (1995), 256-266.

[Zhang95b] Zhang, H.J., Tan, S., Smoliar, S., and Yihong, G."Video Parsing, Retrieval and Browsing: An Integrated and Content-Based Solution," *Proceedings of the ACM International Conference on Multimedia,* San Francisco, CA, November, 1995.

Bibliographies

Michael A. Smith received his Ph.D. in Electrical and Computer Engineering at Carnegie Mellon University in 1998. His dissertation topic was the Integration of Image, Audio and Language Understanding for Video Characterization and Variable Rate Skimming. He received his B.S. in Electrical Engineering from North Carolina A&T State University in 1991, and his M.S. in Electrical Engineering from Stanford University in 1992. He is a member of IEEE as well as Eta Kappa Nu, Tau Beta Pi, and Pi Mu Epsilon. He has published papers in the areas of pattern recognition, biomedical imaging, video characterization, and interactive computer systems. His current interests are image classification and recognition, and content based video understanding. He was a member of the Informedia Digital Video Library Project, and serves as a research consultant with ISLIP Media.

Robert Mozenter is the vice president of engineering of ISLIP Media, with over 12 years of management and development experience in the high technology industry. Prior to joining ISLIP, Mr. Mozenter managed a variety of engineering projects at Intel Corporation in the areas of video conferencing and networked multimedia. In addition, he has several years of experience as a commercial software developer, delivering products and technology in the fields of digital audio and video, streaming multimedia, and personal conferencing. Mr. Mozenter holds an undergraduate degree in engineering from Rutgers University and a master's degree in computer engineering from Villanova University.

Workflow Implications of an Integrated Digital Newsroom

Charlie Bernstein
Leitch Technology Corporation
Burbank, CA 91505

Abstract
Digital Newsrooms solutions have been proposed over the past four or five years in an attempt to streamline operations in the hectic, chaotic world of the newsroom. With new technology and tighter integration, the reality of a real Digital Newsroom solution has become possible. Industry experts have predicted that these solutions are still two to four years away. New developments in processor speed, integrated control, and integrated applications are making these applications a reality today.

OVERVIEW

The operation of news organizations today have not changed for many years. To date, most changes that have occurred involve new tape formats (digital tape) or new more powerful applications performing the same tasks (3D graphics). These changes have enhanced the content of broadcasts, but to date have not changed the way news is produced.

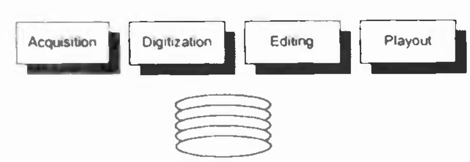
The news environment is hectic and chaotic, created by the time sensitive nature of the business. The first station to report a breaking story in a market is reporting news, everyone else is reporting history. Raw footage comes in from all over the world. Local and national stories are covered by news crews. New stories need to be edited and ready to air at specific times every day and to add to the excitement, you cannot predict when the next big story is going to hit. Late breaking stories are edited at the last minute with people running down halls with tapes to get the story to air. This is repeated two to three times per night and then when your done, you get to do it all over again -- tomorrow.

Deadlines are tight and the pressure to get stories to air is great. Things in the newsroom are frantic all day as the news director decides what stories to run, and how to fit in late breaking news. The technical director needs to make sure that all the stories chosen for the rundown are available and ready for playout. Run-downs need to be created, tape machines need to be loaded and monitored, and the news automation system needs to be programmed to run the story, the graphics overlays, and the character-generator.

THE DIGITAL NEWSROOM

Digital Newsrooms solutions have been proposed over the past four or five years in an attempt to streamline operations. With the competition for news content and delivery heating up every year, the need for better acquisition, editing, and playout tools are stronger than ever. Today, news operations in many parts of the world still rely on manual tape operations. These facilities use tape to tape editing suites, store these tapes on shelves and may require live manual loading of tape during their newscasts. Many operations have migrated to semi-automatic operation for playout by integrating a cart machine and news automation into their operation

Digital Newsroom systems are based on centralized disk based storage systems that provides simultaneous access to multiple users. Access to storage can be for acquisition of raw material, digitization of archive footage, or multiple simultaneous access of stored material for editing or playout. Simultaneous access to material is made possible by the support of multiple I/O channel each with equal access to the storage system.
for rundown playout.

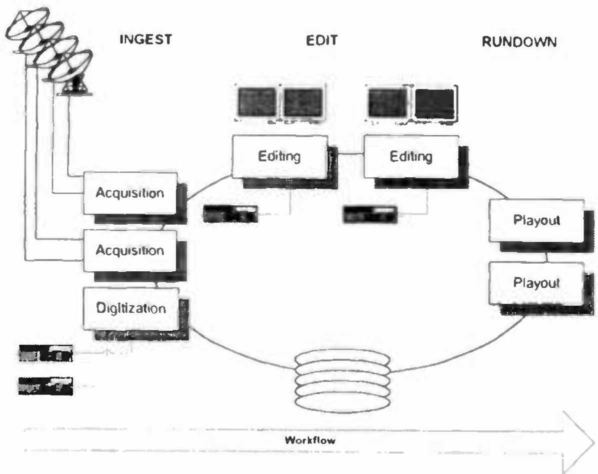


Basic Functions of a Digital Newsroom System with centralized storage

SYSTEM ARCHITECTURE

The Digital Newsroom is comprised of one or more of the following functions or "sub-systems":

Acquisition, Digitization, Editing, and Rundown Playout.



Digital Newsroom Video Server System

A new system architecture has been proposed that supports all of these functions. Integration of one or more non-linear editors with centralized storage along with integrated channels for acquisition and playout is possible. The basic system is built around the architecture of a multi-channel video server. The basic concept of the video server is multiple simultaneous access to shared storage.

Ingest & Acquisition

Using a Digital Newsroom System, acquisition of video is now streamlined. Acquisition of video can be recorded directly into central storage from one or more ingest channels on the server.

Loaded material enters an I/O module and is encoded using a chosen encoding format. Using a modular approach to I/O allows for systems to be designed to support various compression schemes. Digital compression formats include Motion JPEG, MPEG 4:2:2, MPEG 4:2:0, DVCPro 25, and DVCPro 50 can be supported. Intermixing of multiple compression formats is available with limitations. New chipsets today can support both MPEG and DV formats in a single chip allowing the creation of a system that supports four formats simultaneously in a single system.

In addition to real-time transfer of video to the server, new standards for moving information in the facility have been ratified by SMPTE that will allow for compressed video to be moved from acquisition to the server as data. This new standard, SMPTE 259M, otherwise known as SDTI, will allow for files to be moved from cameras and VTRs at speeds up to four times real time, streamlining acquisition further.

Once material is loaded, it is available to all other channels in the server. With optimizations in system design, advances in disk based storage and increased processing power, newly acquired material can be accessed by other I/Os in the system in as fast as three seconds, depending on the size of the system and compression quality used.

Digitization

Additional source material can be loaded into the system via machine control at the ingest point. Edit stations can also control VTRs for last minute digitization of video into the system directly. Machine control of the VTRs are handled by the editor. Other machine control options can be used in the digitization process from automation systems to other machine control devices to load material from tape.

Editing – Linear & Non-Linear

The heart of the Digital Newsroom system is the editor. Much of the source material that is shot during the day must be editing into a finished story for air. Today many news organizations use VTRs connected to an edit controller and an audio mixer to cut their stories. Most stories are cuts only due to time constraints and limitations of the equipment being used.

With the introduction of the Digital Newsroom System, editing can now be supported by a video server and centralized storage. There are a number of configurations that can be used to connect an editor to the server. A linear editor can access a channel of the video server as it would a VTR. The editor has access to the source material and can stream clips to edit and then record them to tape or back to another channel in the server.

A non-linear editor can be connected to the system. Raw footage would be loaded into the non-linear editor from the centralized storage, and once loaded, a story can be cut on a timeline that represents the order and length of clips used in the story. Each clip can be trimmed and the story can be made to "Fit" or "Fill" a specific length of time. Once the story is completed, it is transferred back to the centralized storage system for playout to air.

Editor is in the Server

Connecting an editor to the server is an effective solution for digital newsrooms, but new levels of integration now allow even more efficient solutions to be built. This approach saves time during the editing process, but still requires the editor to load all material and spool the finished story back to the server. By connecting an editor to the server, news organizations are still required to transfer raw footage from the server to the non-linear editor prior to editing. This can take some time especially if there are a number of clips to be used. Once the story is complete, the final story needs to be transferred back to the server for playout. This can be accomplished in real-time as video. This also can be transferred faster than real time using high speed network connections or SDTI.

A new architecture has been proposed to put the non-linear editor into the I/O of the video server. By doing so, the editor actually has instant access to all material. Since the server keeps track of the files in storage, as soon as new source is recorded, is available to the edit station. Access to material is easy because each channel of the server has access to all files. In addition, transition effects, graphic overlays, and character generation technology can be placed in the server to provide an even higher level of integration to the system.

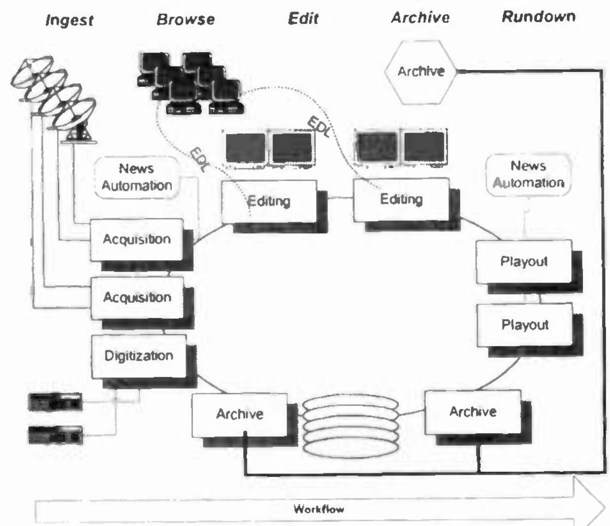
New stories, created as timelines, reference clips and source material on the server. When the story is complete, the timeline file is saved directly to shared storage. Once saved it is available for playout by all other I/Os in the system. Playout I/Os can be configured with the same functionality as the non-linear editor with effects, graphics and text so stories are played out real-time with no rendering required.

Rundown Playout
Playout in some stations involves an operator manually loading tapes into a VTR during live broadcasts. Others use cart machines with stories loaded into bins and scheduled appropriately. As late breaking stories unfold, news organizations still have someone running down the hall with tapes to make late deadlines.

A server based digital newsrooms system solves this problem. Once the story is edited and saved to disk, it is available to play to air on one or more output channels connected to the system. The output channels can be controlled manually via a computer driven interface, by GPIs, or via news automation.

The computer driven interface consists of a sequencing program that has a list of the stories to play. The order of the stories can be changed on the fly prior to air and new stories added to the list as needed. Playout can be controlled by mouse clicks, keyboard entries or via GPI triggers set up for each event.

ADDITIONAL NEWSROOM FUNCTIONS
In addition to the functionality offered in a digital newsroom system, other synergistic systems can be integrated into the newsroom to provide additional functionality. These include: Low-resolution Shadow Browsers, Data Tape Archive Libraries, Asset Management Systems, and News Automation.



Digital Newsroom System with integrated Control, Browsing, and Archiving

Automation – Automated Machine Control
During the past decade computer based control systems have been developed that can control multiple devices in a facility for broadcast and news. New news automation systems have evolved to control multiple devices in the newsroom from VTRs, character generators, teleprompters, and now video servers. Automation systems can be configured to control acquisition of material into a video server as well as for rundown playout. Solutions for news automation is available from multiple vendors.

Low Resolution Browsing
New systems have been created that work along side acquisition and storage systems that shadows the in coming footage at low resolution. Using computer and internet technology, a computer base server can now encode the same video stored in central storage in formats such as MPEG-

1 or Quicktime at bit rates under one megabit per second. These shadow browsers can be accessed by computers on journalist's and editor's desktops for viewing of new stories or checking a story for air without using conventional broadcast video equipment.

These browse systems log each clip with the appropriate data and store material on a smaller computer based server. Client terminals can now be located throughout the facility (and even remotely) with access to material over a local area network. Viewing is done using specialized software on the workstation or via standard internet style browsers with support for streaming video. As the journalist views the footage, edit decisions can be made. These edit points can be saved in an edit decision list and then sent to a full non-linear editor for conformance on the broadcast system.

Content Archiving and Asset Management

Although the cost of hard disk based storage has dropped, it is not practical to keep all source material and archived source material on a disk based system. There are limits to the storage capacity of large disk arrays. Many disk arrays today can be expanded to hundreds of hours of storage. Even with this expandability, archive material over time will certainly take much more than this space. Practical solutions for archiving this information is via tape.

As server systems become more sophisticated, new data tape solutions have emerged that can be integrated into the digital newsroom system. Using computer based control systems, selection of archive material, and control of moving files to and from a data tape library is streamlined. Importation of lists of files to save or direct control by an automation system provides methodical control. All error detection and correction is handled by the system as is flow control and bandwidth allocation. Multiple tape drives can be connected to a single server.

Material stored on the data tape library system can be stored "Near-Line" or "Off-Line". Tapes loaded into the bays of the system can be accessed by a robotic arm that loads the tape into the tape drive are considered to be near-line. They are available to the system without operator intervention and can be loaded into the server for playback in seconds or a few minutes. File transfer speed will vary based on the length of the file, the speed of the tape drive, the speed of the connection to the tape drive, and the bandwidth available in the server system.

Tapes can also be considered "Off-line". Data tapes can be taken out of the archive and stored. The system is aware that the tape exists, and marks it off-line. If a user on the system needs footage that is off-line, the system can tell the user that the tape needs to be loaded and also indicate which tape or tapes is required.

BANDWIDTH

In order to support scalability required for multiple high quality input feeds, multiple non-linear editors, and multiple output channels, centralized storage must be based on a high bandwidth storage subsystem. By building a system based on high bandwidth, all the components of the system can have access to media simultaneously.

Using technology such as fibre channel arbitrated loop bandwidth systems can be built with bandwidth up to 1 gigabit per second. When this is divided up into channels of compressed video, systems can be built with multiple input channels, multiple non-linear editors, multiple output channels, and multiple archive connections.

WORKFLOW

There are many benefits to a integrated digital newsroom system. All of these benefits assist in streamlining the daily workflow of the newsroom. Workflow is streamlined by having an integrated system to handle ingest, editing, and play to air.

Acquisition is streamlined by having all feeds recorded directly into centralized storage. Access to all feeds is available in multiple locations in the facility without having to keep track of which tape it is on, where the tape is or concern over dubbing copies of the tape once received. Using an integrated browser, source material can now be viewed by editors and journalist without leaving their desks or offices. Time is saved at every step.

Editing is now streamlined. No time is taken to make copies of source. By putting the editor into the video server, access to material is literally instantaneous. Utilizing Non-linear editing techniques the system can cut together clips with a clip of a mouse. Timelines can be played out real-time in cuts only mode or with two dimensional effects, text and graphics.

Rundown Playout is streamlined. Once a timeline is saved in the system it is available to play to air on an output channel of the server. No more running down the hall with tapes, no waiting for the finished story to rendered or copied to tape.

SUMMARY

Digital newsrooms systems will change the way that news organizations work. Using a video server based system with built in non-linear editors and integrated third party applications this system reduces operational complexities and provides new functionality not available in the past. From ingest -- to edit -- to playout, the system increases productivity, simplifies workflow, and saves time.

REFRAMING *A BUG'S LIFE*

Craig Good
Pixar
Richmond, CA

ABSTRACT

Wide screen film formats were largely a reaction to television and are thus inherently difficult to transfer. The best solution is a full image, or *letterbox* transfer, although this suffers from spatial resolution problems. Most consumers still expect full screen transfers. A few 2.35:1 films are shot with spherical lenses and masked, allowing the matte to be pulled out to get to, or close to, the 1.33:1 aspect ratio target. This can work well, but easily leads to compromises in composition. For anamorphic films, the only option is cropping or so-called *pan and scan* using a telecine chain. The drawbacks are numerous, including the pathological case of disembodied noses speaking to each other from opposite sides of the screen. As a result, filming a wide screen film with the thought in mind of a subsequent video transfer can lead to compromises in staging and composition.

Pixar's second feature-length film, *A Bug's Life* was produced as an anamorphic 2.35:1 picture. This aspect ratio was chosen to match the large-scale, epic nature of the story. Production was able to proceed with no television-driven compromises because of a new option provided by the computer graphics technique used. *Reframing* is the name given to the process of remaking the film in a new aspect ratio. Using 2D image processing tools, many shots were "panned and scanned", albeit with a level of control impossible using telecine. Roughly half of the film was reframed at the new aspect ratio and recomputed. In some cases characters and props were moved to facilitate the narrower frame. In others the camera was simply recomposed and/or reoperated. The result is a feature film transferred directly to D1 video, both letterboxed and reframed, without leaving the digital domain. Both versions benefit from perfect registration, and the 1.33:1 version has been redesigned to work in the narrower, smaller, and lower resolution world of television. Outside of animation, and computer animation in particular, there is no practical way to achieve this result. That is why *A Bug's Life* is the first film to be reframed.

INTRODUCTION

"Story is King" at Pixar as director John Lasseter says. Every choice is designed to support and enhance the storytelling of the film. During preproduction on *A Bug's Life* the choice of aspect ratio got serious attention. Even though *Bug's* would be set in the miniature world of insects, the cast and story were large scale. A big picture of a small world was a tantalizing prospect. The Layout Department, which is in effect the camera crew, was an early proponent of a wide screen aspect ratio. The Layout department and the Art Director studied a number of wide screen films in order to better understand what the format had to offer. They were impressed with the sense of scale and the compositional possibilities of the format. The Art Director's wide-aspect composition studies convinced the Director that it was the right way to go.

The decision to go with a wide, 2.35:1 aspect ratio was not taken lightly. It would affect every aspect of production, but especially the framing, lighting and rendering. It would require more compute power and more disk space. Not only was it new technical territory for Pixar but it presented a significant problem for the eventual video release of the movie.

Wide screen film formats have existed since the very early days of film but didn't become popular for many years. By the 1930s the film industry had settled on a very pleasing aspect ratio, still known as the *Academy Aspect Ratio*, of 1.37:1. Nearly every major Hollywood film up through the end of the 1940s was produced in this format, including classics such as *Casablanca* and *Gone With The Wind*. . When standards were being set in the early days of television the screen was naturally made about the same shape as the movies.

As the popularity of television grew, Hollywood feared that movies would be replaced by the new medium. The studios reacted by giving audiences something they couldn't get at home on TV: wide screen. A number of wide formats were trumpeted through the 1950s and 60s including *VistaVision, Cinerama, Todd-AO* and *Cinemascope*. The latter is an anamorphic 2.35:1 format which prints the image on the film "squeezed" into a normal 1.33:1 frame. *(Fig. 1)* The image is later "unsqueezed" by a special lens on the projector. Today nearly all wide screen films are made this way and (even though *Cinemascope* was a trademarked brand name like *Kleenex*) are often referred to as *Cinemascope* or *'scope* films.

Fig. 1 Anamorphic, or "squeezed" image as recorded on film.

Now Hollywood recognizes home video as a valuable friend and not an enemy, but is left with the thorny problem of transferring wide-screen films to the narrower video frame. In recent years home theatre enthusiasts, especially the owners of laserdisc players, have recognized that the best method is the so-called "letterbox" transfer where the entire wide screen frame is shrunk to fit the video screen. It's not a perfect solution because a large area at the top and bottom of the screen is left unused, losing the potential benefit of its resolution and because most consumers still expect a full-screen transfer. In addition, throwing away part of the monitor's resolution becomes a significant problem for those watching low-resolution VHS tapes.

The standard solution is the so-called "pan and scan" method. Using a telecine transfer machine to go from film to video, large parts of the image are selectively thrown away in order to fill the video screen with picture. Sometimes it is simply cropped right down the middle, and sometimes the image is "scanned" from one side to another. The drawbacks to this technique are numerous. The original composition is lost forever. Extra movement is often added to the camera while trying to "scan" from one part of the screen to another. Sometimes a single shot of two characters becomes a series of one-shots, ruining the editorial flow of the film. And then there's the pathological case of disembodied noses speaking to each other from opposite sides of the screen.

One can, of course, shoot a wide screen film knowing that it will go to video one day. James Cameron (*Terminator 2*) likes to shoot on Super-35, a flat (*ie:* not anamorphic) format which is masked in the theatre for wide screen. The result is that he has extra movie at the top and bottom of the screen to use in his sophisticated style of "pan and scan". The only other alternative is to try to compose and stage for both aspect ratios at once. This can obviously lead to compromises in both the theatrical version and the home video version.

With the phenomenal, record-setting performance of *Toy Story* on home video, Pixar understands the importance of the video release. As filmmaking purists we wanted to make a no-compromise wide screen film. As storytellers we also knew that millions of people would be watching the movie over and over again on video. We didn't want to compromise their experience either.

DIRECT DIGITAL TO VIDEO

Very early in the discussion of reframing *Bug's* Pixar realized that the film to video transfer step could be avoided entirely. Both the 1.33:1 and letterboxed images could be delivered directly on D1 digital videotape. This decision alone accounts for much of the stunning quality of all of the video versions. Instead of the generation loss of going from digital image to film and then back to digital video, there is nearly no generation loss. The images are perfectly stable, something that is hard to achieve in even a pin-registered film transfer. In fact, even the 16x9 anamorphic video for DVD was done this way. So *A Bug's Life* is the first feature film that could use the label "DDD" on a home video release.

DIVIDE AND CONQUER

A Bug's Life has 1680 shots. Each one had to be examined and processed for the video aspect ratio. Layout and the Editorial Department divided it into four classes. There were two 2D categories, which mimic the traditional "pan and scan" of a video transfer. This image editing task was performed directly on the film resolution frames, making it very fast compared to rendering. There were two 3D categories, which involved re-rendering the shots, albeit at video resolution rather than for film. The four categories appear here in their order of difficulty.

- **Crop**. Using Pixar's 2D imaging tool *comet* a 1.33:1 crop window is sampled out of the original film-resolution shot.

- **Scan**. This is a moving crop, where the crop window animates across the film-resolution frame.

- **Frame Height**. The original shot is copied to a new one and the aspect ratio of the camera is changed to reveal what was previously outside the top and bottom of the film frame. This is like selectively opening the masking on a spherical (non anamorphic) film. The name comes from this effect of making the frame "taller". Sometimes the camera is re-operated as well. The *xyz* position of the camera never changes, but it may be zoomed, panned or pitched to get the needed composition. This ensures that the perspective never differs from the original shot.

- **Restage**. The original shot is copied as for a Frame Height, but characters and/or props may be moved to fit in the new framing.

GETTING STARTED

The film was divided into the same sequences used to produce the movie. A sequence is a series of related shots which naturally belong together. Before any work could begin on a sequence the Layout team reviewed each shot on videotape, assigning it to one of the four classes and giving any additional notes which might help later in production. Several factors came into play. Not only were we re-composing the film for a narrower aspect ratio, but remaking the film for a smaller screen. The large size of a movie screen and the high resolution of film allow characters, faces and details to be proportionally much tinier in the frame than on the smaller, lower-resolution video screen. The video audience could miss many important story points if they couldn't see the acting at critical moments.

Another problem is that the standard for video frames includes the notion of an *Action Safe* area *(Fig. 2)*. Most television sets and monitors, with the exception of a few professional models, trim the video frame much the same way a magazine may "bleed" an image in order to fill the page with picture. So there is an area around the edge of a video frame which may or may not appear. A television that meets the standard will at least show the *Action Safe* area. In practice, very few show anything outside that line.

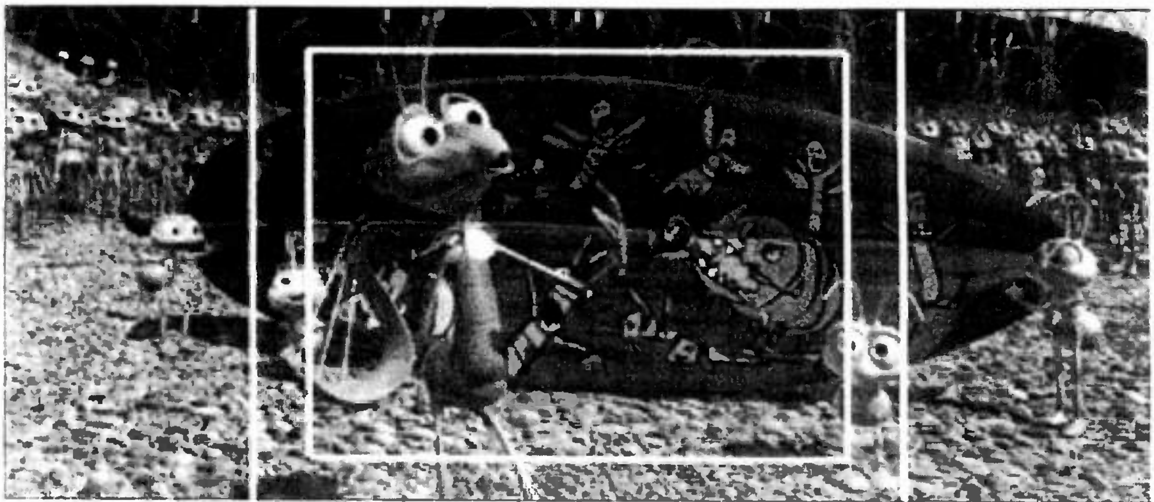


Fig. 2 Action Safe Area compared with wide screen format. Inner line is Action Safe. Outer lines are the limits of the video crop window.

THE TEAM

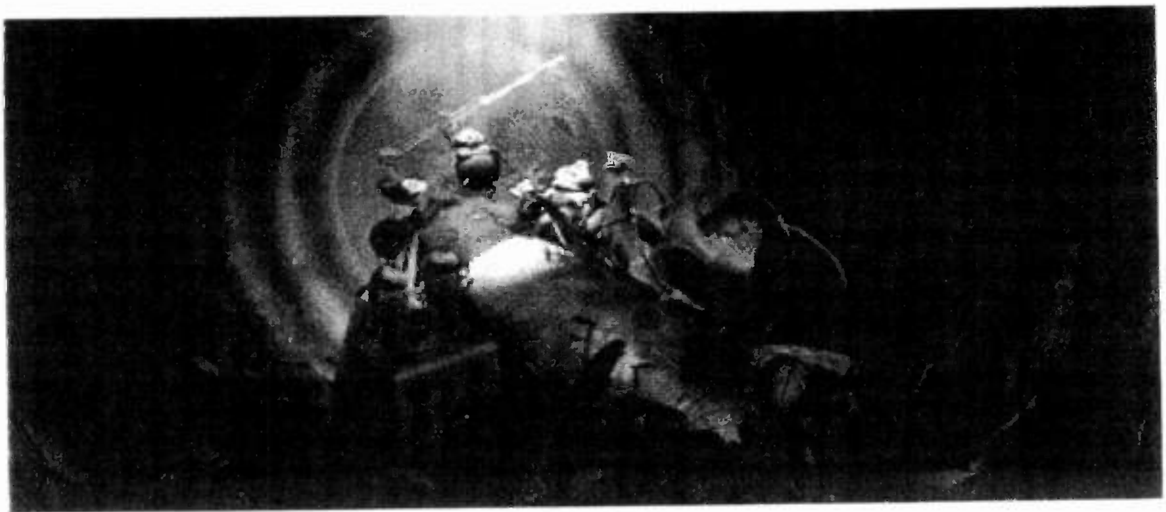
All of these reframing notes were entered into Pixar's production database so that the work of the entire team on all 1680 shots could be tracked. The first on the scene was typically the Layout Department. They did the reframing on each shot that determined what would appear in the video frame. In the case of rerenders, it was common to reveal a problem in the animation that was outside of the film frame. For example, a character's hand hanging below the old frame line might now appear "dead", or antennae which used to be cut off at the top of the frame suddenly looked like a stiff pair of rabbit ears on an old television. A dedicated fixer, assisted for a few weeks by members of the *Bug's* animation staff, attended to the animation fixes. A team of Technical Directors and Render Wranglers rendered, painted and repaired rendered frames as needed.

THE CATEGORIES

Once each shot in a sequence was assigned to a reframing category the Layout Artists could start working. For the two 2D classes, Crop and Scan, the original film frames had to be restored to Pixar's disk farm. For the 3D groups new shots were copied and work could begin right away.

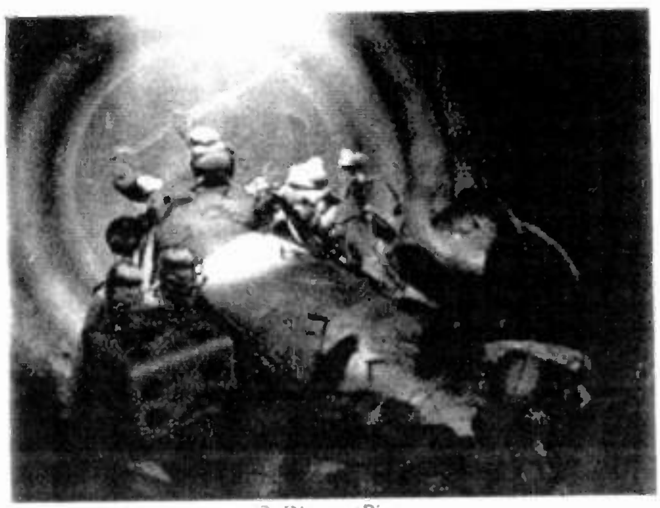
CROP

Pixar's proprietary animation software, *Marionette*, uses "Articulated Variables", or *avars* to control the animation. In some commercial systems each avar would be considered a channel. Every translation, rotation or other animation control uses its own avar. Pixar's camera package uses avars to control all aspects of the camera, such as position, orientation and field of view (FOV). A special set of avars which doesn't affect the final render allows the animation system to control how comet crops the image. *Marionette* allows a reduced resolution copy of the final film frames to form the background image of the camera tool*(Fig. 3a)*. A representation of the *crop window* over the image tells the artist which part of the original frame will be cropped later by *comet* to form the video frame *(Fig. 3b)*. The layout artist sets the value of the East/West scan avar, which positions the crop window.



© Disney/Pixar

Fig. 3a Film Frame before cropping

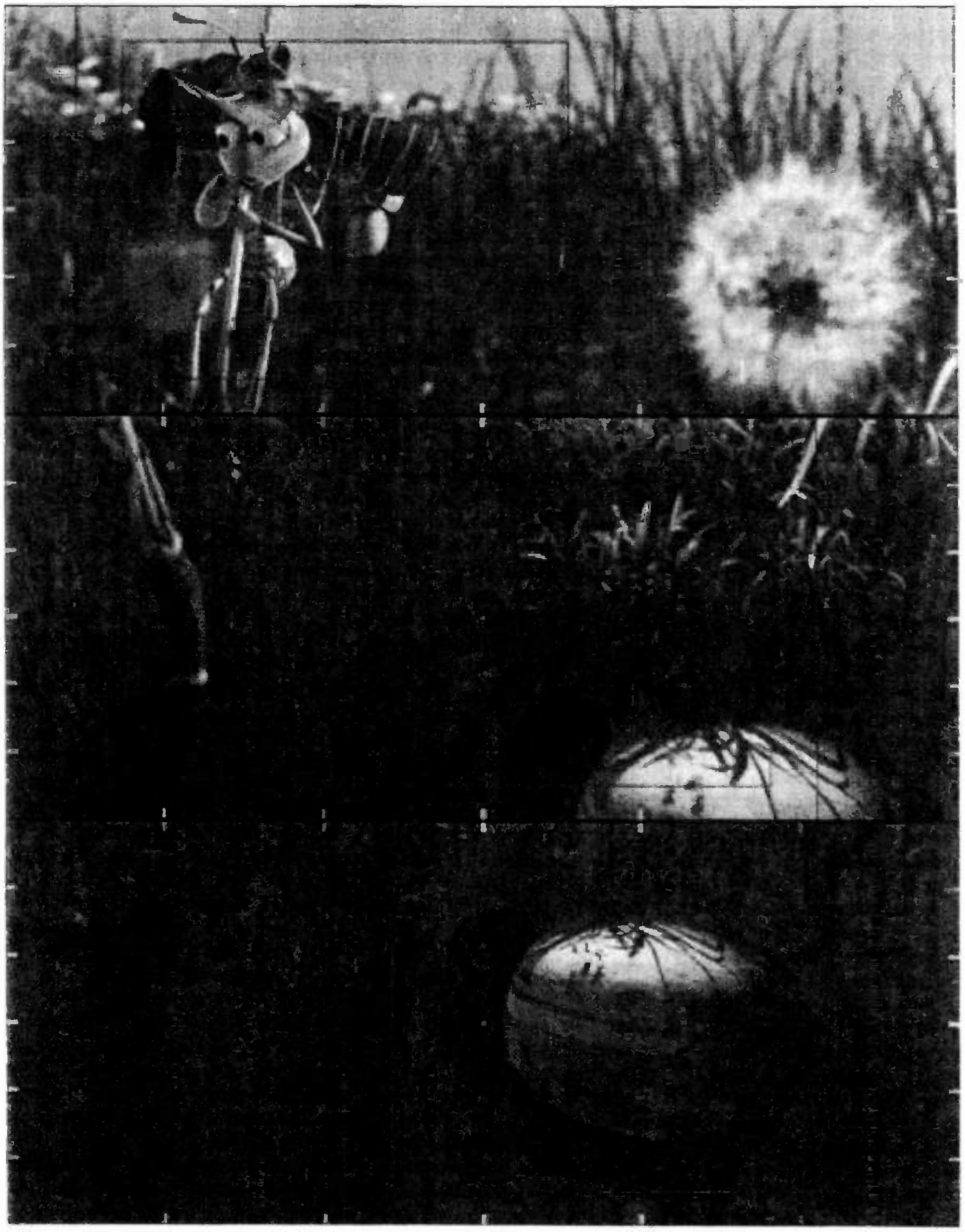


© Disney/Pixar

Fig. 3b Video Frame after cropping

SCAN

A scan is simply an animated crop. The crop window can be moved left or right during the shot. Traditional telecine machines have this capability. But Pixar had two distinct advantages. One was that the crop window is animated using normal avars, which means that the shape of the curve can be precisely controlled. On a telecine machine the only options are linear or a fixed ease at the start and end of the move. The second advantage is related to the first. Not only was the scanning being done by the same crew who operated the camera in the original film, but the curves used to operate the camera are available for matching.



© Disney/Pixar

Fig. 4a,b,c Start. ,middle and end of scan

The effect of matching the spline curves is that the camera now looks down and to the right as naturally as if the pan were part of the original move. It was a simple matter to match the bezier *handles* of the scan, shown in the following illustration as the top spline, to the original pitch and pan, shown here as the bottom splines. The method used for shaping curves in Pixar's system will look familiar to those who have used graphics software which uses click-and-drag handles to change the shape of the spline. *(Fig. 5)*

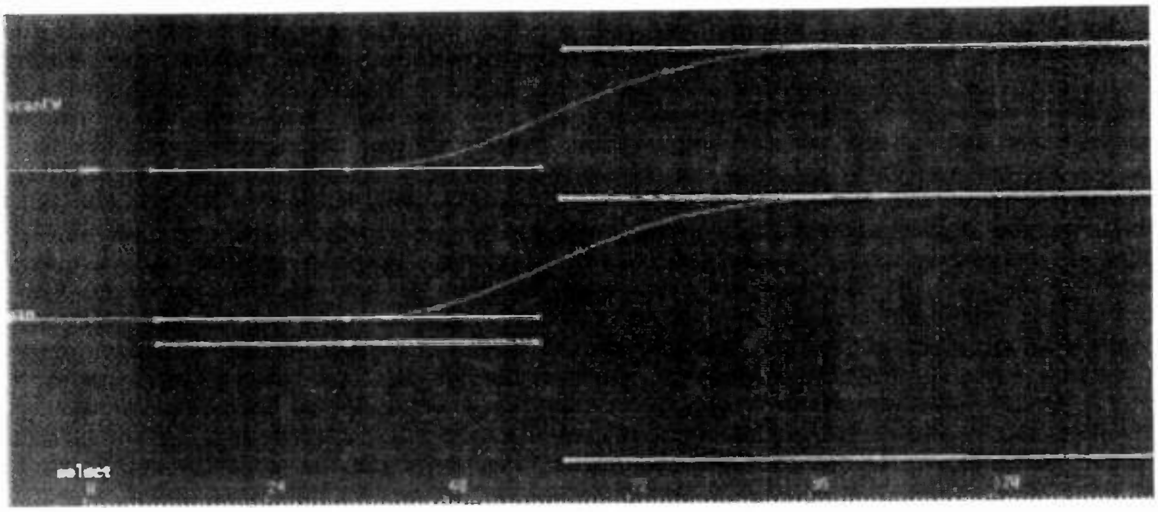


Fig. 5 Scan movement matched to original camera move

The scan becomes a seamless part of the original camera move, avoiding a clunky telecine artifact. That way the audience isn't distracted even subliminally by a camera move that isn't as smooth and cohesive as those in the original film. This technique was used in a number of shots.

FRAME HEIGHT

Even though the results are better than with a telecine, both crops and scans are imitations of a normal "pan and scan" transfer in that they simply crop out a region of the original film frame. They sufficed for about 45% of the shots in the film. The real reason for the reframing project, though, was in the 55% of the shots that needed to be recomputed. Frame Height seems like a strange name until, perhaps, one considers Figure 6:

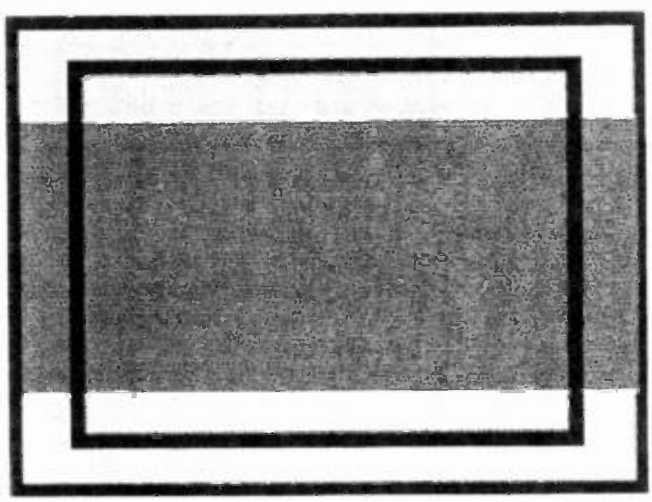


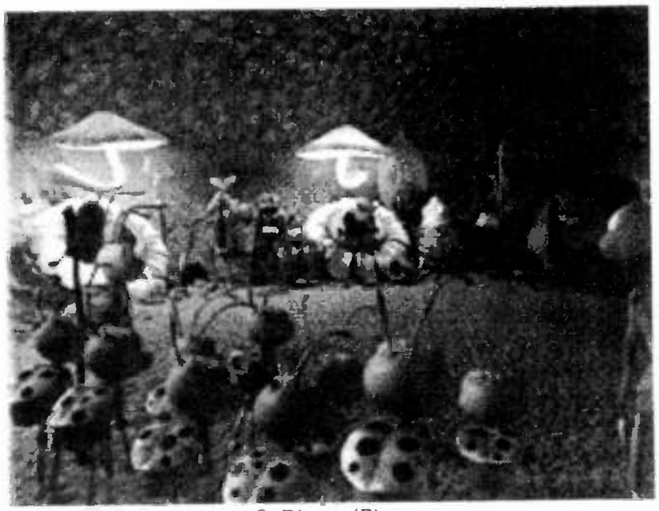
Fig. 6 Letterboxed area compared to "Frame Height" intermediate framing.

The dark grey region represents the shape of the original wide screen frame. The outer black box represents the shape of a video frame. This is exactly what the relationship would be in a *letterbox* transfer: The full movie frame is centered vertically in the video frame. The light grey area would be the unused portion of the TV screen.

We very often wanted to see just a little bit beyond the top and bottom edges of the original frame in order to get a good composition for video. The inner black box in the illustration represents an example of this "frame height" adjustment. In some cases, the frame was heightened until the width was equal to the original frame. This was called a "Full-width Frame Height". Clearly, we spent more time making the film *look* good than coming up with good names for how we did it.

If shaded boxes aren't clear, here's an example taken directly from the film. The original frame included a panoramic staging of both the Blueberry Troop and the circus bugs. Any attempt to crop this image would throw away over half of the cast. It was an obvious candidate for a full-width Frame Height. The video frame adds new wall above and Blueberry Troop below.

Fig. 7a Original wide screen frame



© Disney/Pixar

Fig. 7b Reframed as a "Full-width Frame Height"

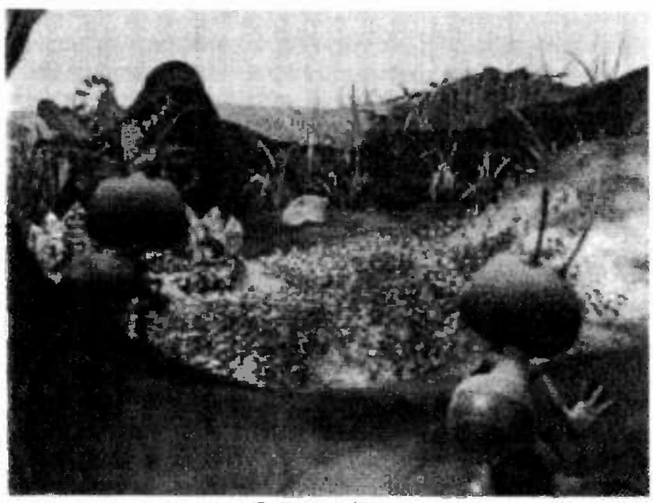
It wasn't always simply a matter of opening the frame. Sometimes the camera move itself had to change during the shot. Fortunately, computer graphics allows the camera operation to be changed after the acting is done. Several shots take advantage of this ability to re-operate the camera. A good example, though impossible to illustrate in this paper, is the shot where Flik is fretting to the circus bugs that his children's children will be taunted with "Look! There goes the spawn of Flik, The Loser!". This was a complex tracking and panning move in the film, and proved impossible to scan. Instead, we re-timed the animation of the camera to properly compose for Flik and the circus bugs as he hits each mark.

RESTAGE

In a few cases we didn't want to open all the way up to a full-width. The characters would have ended up too small in the frame, or distracting elements might have been revealed above or below the original frame lines. In those cases we moved or reanimated some or all of the characters in order to get them into the new frame. In the establishing shot of the banquet sequence, the two girl ants were so far apart on the screen that it was impossible to get both without going all the way wide and ending up with a strange shot of lots of sky on top and lots of leaf on the bottom *(Fig. 8a)*. So the right-hand girl was moved left and her animation was fixed.
Fig. 8b).

© Disney/Pixar

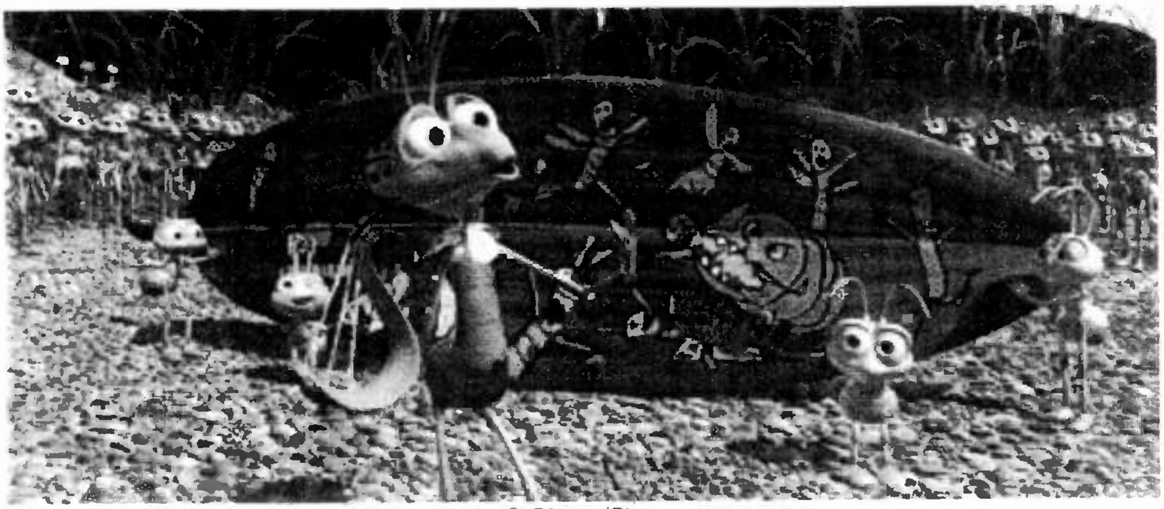
Fig. 8a Original film frame



© Disney/Pixar

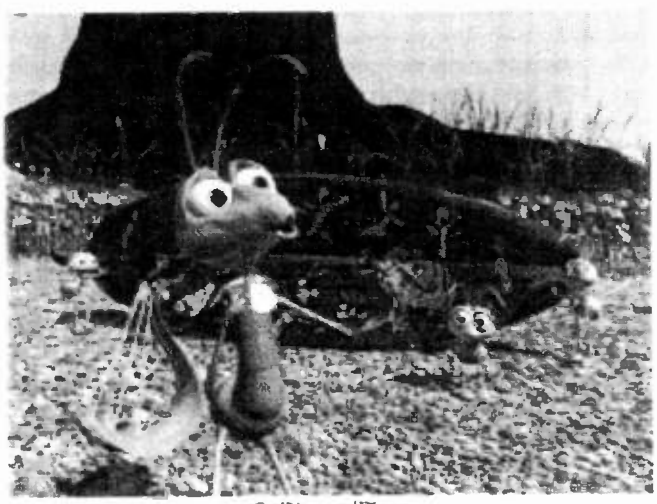
Fig. 8b Restaged

Another good restaging example is from the same sequence(*Fig. 9a*). In this shot, we restaged Mr. Soil closer to the camera for better composition and to hide the fact that his legs weren't completely animated (*Fig. 9a*) . Without being able to open up the frame, we never would have seen all of the mural, nor all of the "actors" at the end of the shot.



© Disney/Pixar

Fig. 9a Original wide screen frame



© Disney/Pixar

Fig. 9b Reframed and Restaged

A different sort of restaging involved tweaking the animation. The animator gave Francis a big wind-up when he delivered his pie to Heimlich *(Fig. 10a)*. To make the animation fit on video, we modified the animation so that Francis didn't anticipate quite so far screen left *(Fig. 10b)*.

Fig. 10a Francis uses the whole frame...

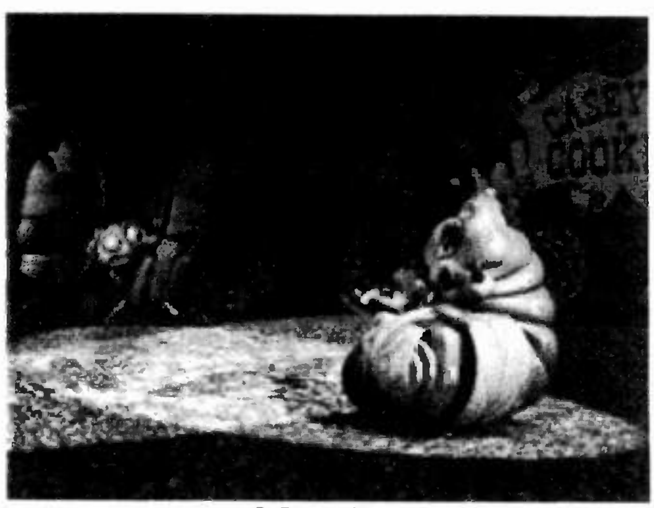
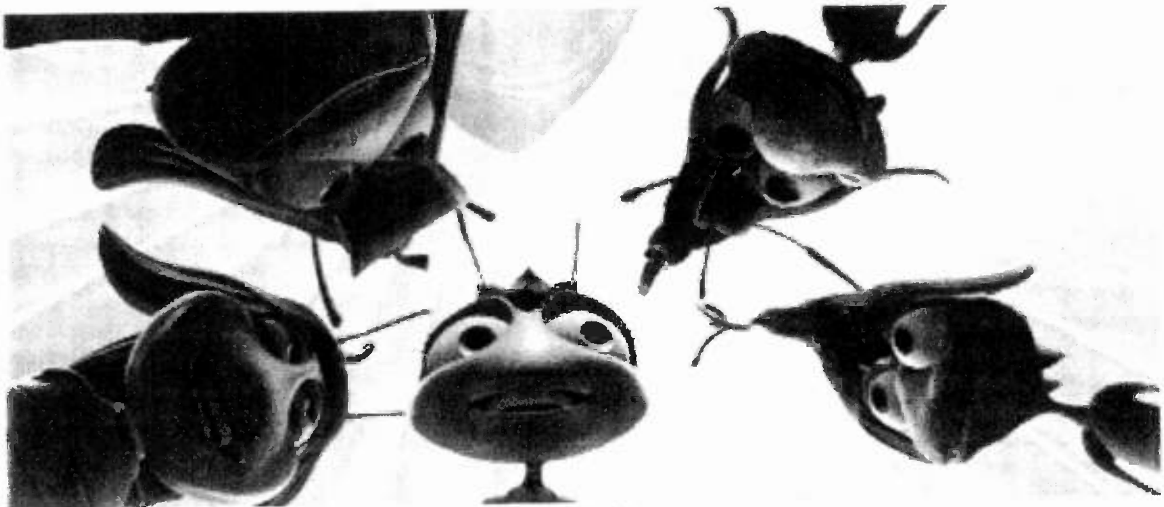


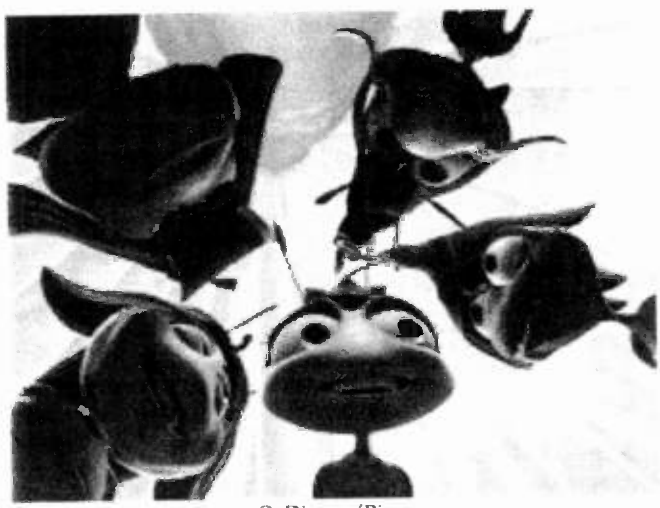
Fig. 10b ...so he gets restaged.

Perhaps the most obvious, and obviously necessary, example of restaging involved the "royal huddle". After restaging the characters, only slight animation tweaks were needed to get them looking at each other again.



© Disney/Pixar

Fig. 11a Wide Screen



© Disney/Pixar

Fig. 11b Restaged

PLEASANT SURPRISES

There are those who insist that a wide screen aspect ratio is *better* than 1.33:1. There are also those who insist that apples are better than oranges. Some of the most beautiful cinematography ever done has been shot in the Academy aspect ratio. There's nothing inherently wrong with it. In fact, we found a few pleasant surprises in the form of shots which looked better reframed at 1.33:1 than they did in wide screen. *(Figs. 12 - 13)*

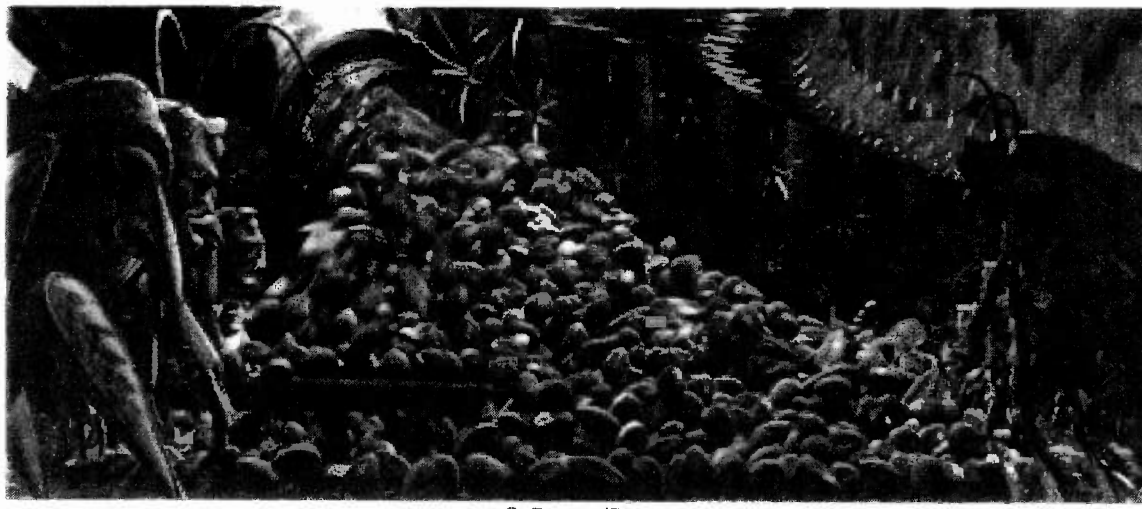


Fig 12a This grain pour was very good in wide screen...

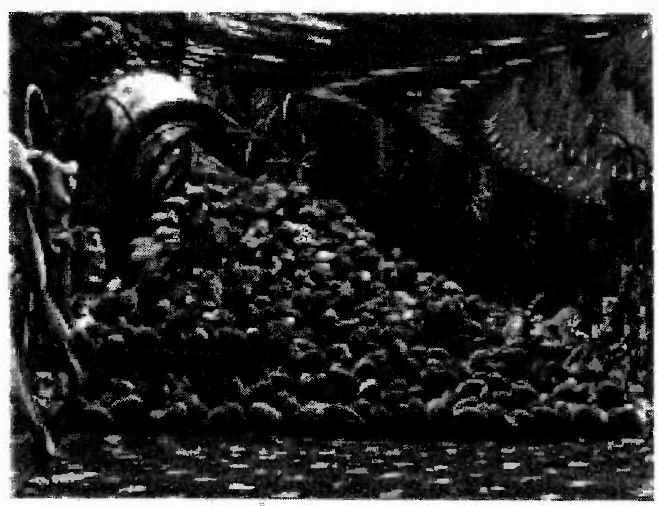
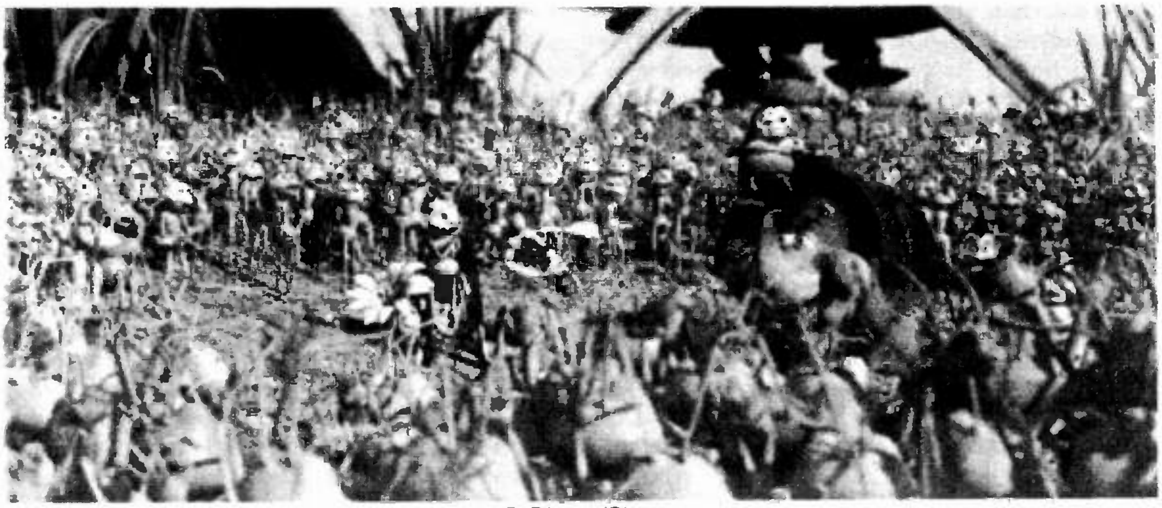
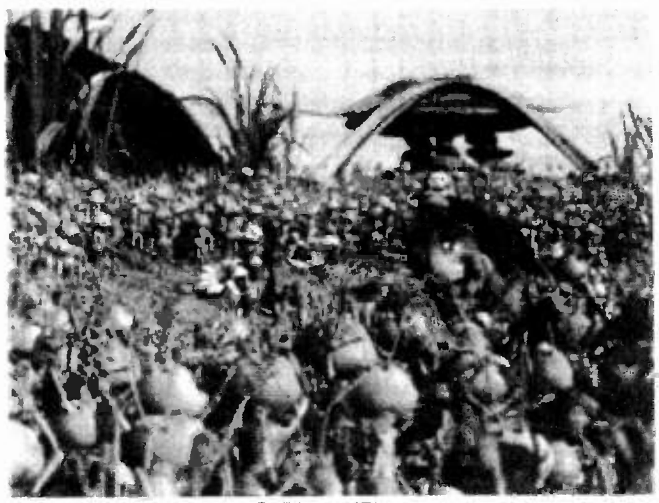


Fig 12b ...but works very well reframed for video.



© Disney/Pixar

Fig. 13a Good in wide screen...



© Disney/Pixar

Fig. 13b ...better on video?

DELIVERING THE GOODS

Once every frame had been processed, all that remained was delivering the final product. For delivery the film was divided into the same reels used by Editorial during production of the original film. Each reel was laid up frame by frame onto Sierra Design Labs Quickframe Digital Disk Recorders. Going reel by reel meant that no more than about ten minutes of the film had to be on the DDRs at a time. *Comet* was used to resize the renders and convert to 4:2:2 digital video at both PAL and NTSC resolutions. One Quickframe was configured for PAL and used for all of the PAL transfer work. The D-1 deck is reconfigurable between NTSC and PAL.

In all there were six versions of the film laid off: 1.33:1, letterbox and a 16x9 anamorphic letterbox in each of PAL and NTSC. Since the anamorphic versions were digitally resized straight from our final renders, owners of DVD players and anamorphic televisions will get the full resolution benefit possible in that format.

As each reel of movie was approved, it was rolled onto D-1 tape using a Sony DVR-2100. There it was quality checked and set aside for later color correction and online assembly. The tape-to-tape color correction was done at Western Images in San Francisco. As we suspected, only minor tweaks were needed. It couldn't have hurt that we were not correcting for film, processing and telecine transfer. Final post was performed at Todd-AO in Hollywood. The DVD master was encoded from the D-1 tapes.

IT'S A WRAP

Reframing *A Bug's Life* breaks new ground in delivering a single motion picture to two vastly different markets: The big screen of the theatre on 35 mm film and the small screen at home on VHS tape. It allowed the film to be made with no compromises imposed by a future film-to-tape transfer. And it allows millions to watch it at home without the usual compromises of choppy editing or talking elbows. We at Pixar love to tell stories. We make movies for "all those wonderful people out there in the dark", whether they be in theatre seats or on their favorite couch.

E. C. ACTS – MIRAGE: A SUCCESS STORY IN VIRTUAL REALITY

Chas Girdwood
Independent Television Commission
Winchester, United Kingdom

ABSTRACT

The EC ACTS project, MIRAGE, started its three year term in October 1995 and finished in September 1998. During this time it has developed affordable and practical virtual production systems and techniques, and successfully demonstrated 3D programme production and display. Papers by Consortium Partners were given at the International Broadcasting Convention in Amsterdam which effectively gave progress reports along the way and this paper now sums up the project. The result is the development of several systems which are now complete and have been fully integrated for trials and demonstrations. Commercial exploitation has commenced with systems being installed within European television companies in the real world of making programme content and in Hollywood for motion picture production.

INTRODUCTION

Advanced Communications Technology and Services (ACTS) is one of the Programmes in the European Community activities in the field of research and technological development and demonstration. It focuses on the effort to accelerate the deployment of advanced communications infrastructure and services, which when tried and in place, create the need for more originated content.

Over the last decade there have been significant changes in the world of television broadcasting with more channels and an accompanying increase in the use of television for multimedia, entertainment, telepresence and the transfer of information. All this has meant the production of more programmes and specialised content with the resulting demand for more studio time, faster turn-around between programmes, more facilities and greater economies. MIRAGE (Manipulation of Images in Real-time for the creation of Artificially Generated Environments) was set up to address this.

MIRAGE AND VIRTUAL REALITY

Virtual reality production can satisfy some of these demands but only if this is cheap, efficient, can be operated without the need for highly specialised staff who, naturally, are costly to hire and is used for the right programmes. This was not the case about four years ago when some television programmes, or inserts, were being made in the first generation virtual studios. The production capability then was limited, the equipment was expensive, consisting of super-computers not really built for the purpose, production could not be done by staff with traditional studio disciplines because computer scientists and programmers were required, and use was generally on the lavish spectacular productions which naturally attract large budgets and where any savings brought about by virtual reality are negligible.

The MIRAGE Consortium partners, mainly small production and media companies, are involved in the making of successful television programmes which use virtual reality for creating fantasy worlds and putting the live action, realistically into these. They know the industry and the business of television and they recognise the limitations. They recognised what was required to make virtual studio production, effective, straightforward, cost effective and possible for traditional studio staff.

THE MIRAGE PROJECT

The Objectives

The MIRAGE project was born with objectives to meet these needs:

* the development of virtual production tools and techniques at affordable costs.
* the demonstration of hardware and software systems for use by traditional programme makers.
* the integration of production techniques, the definition of working practices and the formulation of preliminary standards.

The Consortium

Independent Television Commission (UK)
Televirtual Limited (UK)
A&C Limited (UK)
Comtec Studios Gmbh (Germany)
de pinxi sa/nv (Belgium)
TYVE Idea é Imagen SL (Spain)
AEA Technology Limited (UK)

Most of these are small specialist companies and the ITC, the managing partner, considers this is the ideal size for optimum working in such a project.

The Areas of Work

The work was carried out in a number of sections concerned respectively with systems for studio production, programme creation and post production, motion capture and character animation, and 3-D (stereoscopy). The trials of these systems at each stage of project development usually included the making of some form of television programme which resulted in not only the technology being developed but also the techniques and working practices as well. This all helped to progress development to the point where system performance was optimised and the exact requirements of studio working were met. All the stated objectives have been met at the end of the project.

STUDIO PRODUCTION

Three virtual studio systems have been built to provide different levels of complexity in a wide price range. Two of these are PC based and use pre-rendered graphic or pre-shot video backgrounds and the third is a real-time system based on a high-end graphics work station.

The use of pre-shot video or pre-rendered (at field rate) computer generated images as backgrounds guarantees the very highest, photo-realistic, picture quality usually surpassing that achieved by a super computer struggling to render in real-time. The graphic backgrounds can be built in any standard 3D package and in operation can be changed in a matter of seconds making it easy to change between scenes in one programme and between different programmes being made in the same studio. Foreground and background pictures are combined in a proprietary chromakeyer.

First Level Pre-Rendered VSS

MIRAGE VSS, the basic virtual studio system, provides a complete and comprehensive facility with up to four studio cameras to place the action into some other environment not existing in front of the cameras. It is fully robotic and controlled by a single operator to select up to 20 different shots with computer-generated graphic or pre-shot video backgrounds (virtual sets). Control is from a user friendly touch screen containing thumbnail pictures of the actual camera outputs for the different shots, or a custom designed switch panel. Camera positions, pan and tilt, and lens positions; iris, focus and zoom, are set up out of vision and cut to air at the director's request. This is the most common operation in the majority of television programmes and occurs with much greater frequency than in-vision pans, tilts and zooms. Various effects like the inclusion of video walls in the scene, windows for sports results etc., flying sponsor's logos and captions can also be included in the picture.

An early development of the system was installed in a vehicle to create a self-contained Mobile Virtual Studio (MVS) for trials and demonstration purposes. It was equipped with one Pentium PC (120 MHz and 32MB RAM) having a hard disc of 850 MB and an external disc store of 4.3 GB which could store up to 14 minutes of broadcast quality (Beta-SP) video material. Backgrounds were recorded and replayed using a Perception component video recorder which has been replaced in the latest version with Matrox Digisuite for an all digital operation. The MVS was used during 1996/97 for the making of three live programmes one of which was a two and a half hour prelude to the Rael Madrid v Barcelona football match. Development was continuous with major trials and demonstrations being set up at the Montreux International Television Symposium in 1997 and IBC in Amsterdam in 1998. Complete systems have now been set up in a number of small TV companies in Spain giving new economies and exciting creative possibilities like the marble-hall sports set shown in Figure 1.

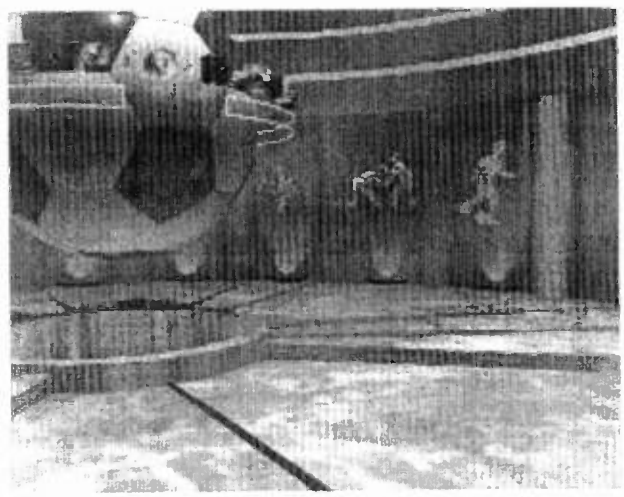


Figure 1

Second Level Pre-Rendered VSS

The first level system is taken to the next level with full camera movement and the corresponding changes to background in vision. This is achieved using the POWERpod 2000 precision camera motion head developed in the project. This allows simultaneous pan, tilt and roll with full focus, zoom and iris control using backgrounds which are again either computer generated or real video. The POWERpod 2000 is described in detail in the paper [1] given at IBC in 1998.

The system is again PC based with recording being done to hard disk and developed software supports 100 different camera positions containing the movement and lens positional data with up to 28 user defined moves between them. Protection is built in so that moves are always legal, i.e. the starting position of one move is the same as the finishing position of the previous one.

Camera moves are worked out as usual by the director before the programme. These are rehearsed and entered into the motion head then repeated to capture the corresponding background in the case of pre-shot video. During the studio production the same camera moves are done on the presenter or other foreground scene with the backgrounds being played out in synchronism from the hard disk. When computer generated backgrounds are used the pre-defined shots, animations and moves are rendered off-line and played back during studio production in the same way.

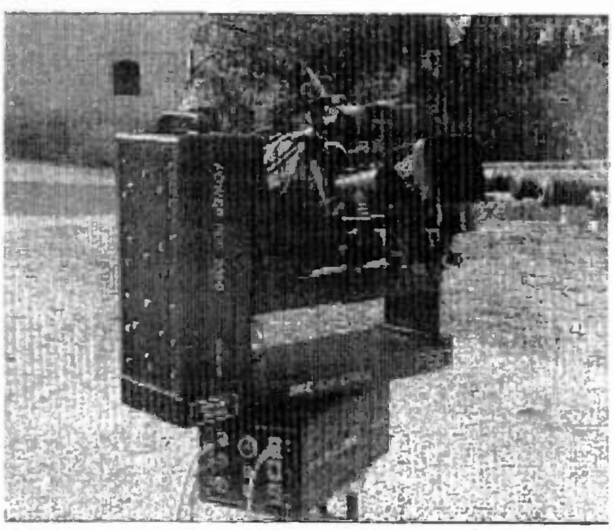


Figure 2

Figure 2 shows the motion head configured for just pan and tilt operation during location shooting of backgrounds for a virtual studio production in Germany.

Third Level Real-time VSS

The third level system is based on a high-end SGI graphics work station and runs in real-time to give high quality graphics without the need for frame by frame rendering. It controls cameras, lighting and character animations during production and has been used to produce what is believed to be the world's first 3-D virtual studio which was used to make an entertainment feature for a theme park and demonstrated extensively at IBC. It is based on the MIRAGE Virtual Edit Suite which will now be described.

VIRTUAL EDIT SUITE

The Virtual Edit Suite is a system developed as a highly advanced event manager operating in real-time to create, control, adjust and edit virtual environments and virtual reality programmes. It uses advanced facilities for the control of cameras, virtual lighting, animations and events and puts these at the finger tips of production staff through a carefully conceived user interface with control modes familiar to post-production staff. It can be used as a programme creation and production tool, in post-production and with its 3-D capability as a tool for creating futuristic experiences for theme parks.

The system contains all the features required and although these can be found in competitive systems MIRAGE VES is the only one to bring them all together and present them in an easy to operate way. The host computer can be either an SGI work station or a PC running Windows NT. Mechanical camera tracking is employed for synchronised background play-out during recording, realistic lighting can be controlled during recording, and there is full control of virtual actors moving in the virtual set. As well as the ability to create virtual worlds the VES can modify their properties during a live show and modify them in post production if required. A Generalised Edit List is employed giving the capability to replay and add virtual and real features in post production. During production, automated directing is available giving the facility for automated camera (virtual and real) control.

VIRTUAL CHARACTER ANIMATION

MIRAGE has addressed all issues concerned with character animation and developed advanced and realistic virtual characters and presenters which can work in live television productions. The work included recognition and the solving of problems that exist when actors or a studio audience have to work live with a character animation that is not actually on the set with them. Speech recognition and advanced motion capture systems have been developed for the control of face and body movements running in real-time under MS Windows on a high-end PC.

Investigations [2] were carried out into the efficiency and accuracy of various motion capture systems and the sampling of data, its correction and subsequent use in the control of animations. Motion sampling / recording software for the Windows platform was developed. Used with the Ascension Motionstar magnetic body suit, the software drives scaleable rectangular shapes which represent the "safe area" or "field of influence" for a body part. The accuracy of the positioning of these rectangles has been the subject of much experimentation to give seem-less performance.

An advanced performance animation system, working on a PC, has been unveiled. This is a synthetic acting system which is capable of injecting real human talent into computerised characters. It has been designed to change the face of 3D animated film and TV drama and is a powerful real-time capture and live performance system which gives completely realistic speech and expression to live computerised comperes, virtual TV hosts and animated drama characters.

Although PC based the system capitalises on the recent rapid advance in 3D acceleration offered by Pentium IIs. The resultant low cost system is nonetheless high performance, typically running characters composed of 5,000 fully textured polygons at a mean frame rate of 40+ fps, more than adequate for live television broadcast purposes. Currently this allows head and upper torso performance but further, imminent advances in the Windows / PC platform are expected to advance the system for full-body real-time capability.

3D (STEREOSCOPIC) TELEVISION

3D video production was introduced into MIRAGE because virtual studios started with 3D computer graphics sets although these were never seen in 3D (stereoscopic) TV and real video should be given the same capability to enable seem-less interchange. Technology exists for stereoscopic programme transmission and display in the home. Computer games generate 3D graphic images and there are future home entertainment possibilities that would benefit from 3D display.

Early work included the making of a 3D television programme covering all types of programme genre from presentation to drama and sport. This used a studio camera, shown in Figure 3, built by one of the consortium partners in a previous project and a small lightweight unit which was the prototype of a small telepresence / broadcast camera to be developed by MIRAGE.

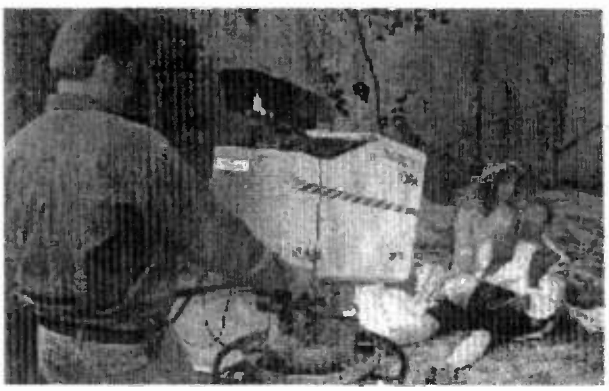


Figure 3

The 3D programme, *Eye to Eye* was demonstrated to a wide audience including key decision makers using every available display technology and showed that the inclusion of depth information is an important factor in any future service offering a real or immersive experience to the viewers. The making of the programme is described [3] in detail elsewhere.

The development of the small camera started with a thorough evaluation of the requirements for telepresence which concluded that the main use would be little more than static video conferencing. It was therefore decided that a camera should be built close to broadcast specifications for use in the acquisition of entertainment / theme park 3D video sequences though maintaining the originally planned small size. This became possible because of the rapid decrease in cost of small camera heads giving increased performance over the cameras used in the *Eye to Eye* prototype. The main work was to fully specify the camera, evaluate heads that could be incorporated, design and then build the camera. This was done and sequences of 3D programme material were used for subjective evaluation and have been supplied to other ACTS projects for their work.

As part of the subjective viewing trials the camera was used to make a 25-minute programme, *Afternoon on the River* . This was made as part of the camera evaluation process, for use in immersive TV experiments and for showing at IBC '98. The programme consists of essentially one continuous shot presenting the viewer with an un-interrupted journey along backwaters on the Norfolk Broads in East Anglia. It was made to show some of the possibilities offered by telepresence for first-person experiences and travel and is accompanied with forward stereo sound and one backward effects sound channel that can be processed into Surround Sound for additional studies in the future.

Figure 4

VIRTUAL ARENA GAMES SHOW

This is a television programme which pulls together all the technologies and techniques developed by MIRAGE for competitive games play between teams from different European countries in a 'cyberspace' arena. The concept has been demonstrated in a trial programme and negotiations are in progress for the transmission of a programme series.

EXPLOITATION OF RESULTS

All developments have been well received with enthusiasm in the intended market-place. This has led to nine sales of the MIRAGE Virtual Studio System in Spain and one in the UK, over twenty confirmed orders for the POWERpod 2000 motion camera head, use of the Virtual Edit Suite for making a theme park entertainment extravaganza and an agreement for the hardware production of the system, interest by decision makers and television producers in the 3-D (stereoscopic) TV work, the use of the character animation systems in daily TV production and the marketing of a proprietary system. The consortium believes that this is a firm endorsement of the success of MIRAGE.

CONCLUSIONS

Always bearing in mind the requirements of television production, MIRAGE has successfully demonstrated several new tools for virtual television studios at affordable prices. All the project objectives have been met and the value given is increased economic and artistic benefit.

ACKNOWLEDGEMENT

The author wishes to thank all members of the MIRAGE consortium for their help in publishing this paper and the ITC for permission to give it.

BIBLIOGRAPHY

1. Calvo, R., Chiwy, P., Girdwood, C., Wells, M., 1998. ACTS MIRAGE: A success story in virtual reality and 3D TV technology and production. Proceedings of 1998 International Broadcasting Convention.

2. Tutt, M. and Wells, M., 1998. Practice and application of motion capture. Proceedings of Computer Vision Colloquium, IEE.

3. Dumbreck, A., Girdwood, C. and Wells, M., 1997. The production and display of a 3D television programme. Proceedings of HDTV '97.

Remote NLE

Robin S. Rowe
SAIC
San Diego, CA

ABSTRACT

Remote non-linear video editing is a natural progression to the next generation in video editing technology. A paradigm shift, editors are no longer chained to the edit suite. RNLE operates upon a medium or low bitrate mirror copy of broadcast-quality video clips. With this approach a low resolution rough cut is created on a PC in a simple cut-and-paste interface. Sometime later, the resulting edit decision list (EDL) is transferred back to the edit suite for the finish cut and committed at high resolution suitable for broadcast or as film or HDTV.

RNLE has important implications to the efficiency of TV newsrooms and other production staffs. Journalists can edit at their desktops freeing up time in the edit suites. Mobile RNLE detaches the RNLE from the network by replicating the low bitrate video clips from the network server onto a laptop. A director can conveniently create a director's cut on his or her laptop while on an airplane en route to a shooting location. The lightweight EDL can be easily emailed back to an editor based in LA or NYC who has the same footage in high res. This enables the director to work from a rough cut rather than uncut dailies. The director can verify at the shoot that they have the necessary shot coverage as well as judging the pacing of the show.

The Leitch BrowseCutter is an RNLE implemented in Java via a web browser interface. At NAB an untrained volunteer from the audience will operate the BrowseCutter software (running on a laptop) and demonstrate the ease of quickly creating an MPEG-1 rough cut from raw footage.

1.0 Client/Server Video Editing

Everyone knows that broadcast-quality digital video files are too large to handle over a corporate LAN running 10 or 100 mbs Ethernet. That's why television stations use Fibre Channel running at 1gbs to connect digital broadcast servers, such as the Leitch VR-300 server and NewsFlash editing system. Video systems with a client/server architecture present a tremendous advantage in the creation of television news programming and play-to-air because a single high-quality digital video file residing on the server can be shared among multiple users. In a true client/server design there is no need to wait while a video file copies from an editor to a play-to-air server. It is the same file.

Unfortunately, the expense of high bandwidth pipes and video mainframes discourages the widespread deployment of client/server video editing. It is simply too costly for most TV stations to provide an editing station to every journalist. The result is a traffic jam in the edit suites as too many users want access to too few machines. What would be ideal is if every journalist could have his or her own inexpensive client-based video editing station running on the PC already on their desk, and have the video files available from a server using the existing corporate LAN. An adaptation of an advanced technology developed for commercial database servers actually makes this possible.

2.0 Replication Servers

For many years there has been a technology in database servers called the replication server. A replication server keeps a copy of the data on a server so that if that server would fail or lose connectivity to remote users it would still be possible to continue business as usual by connecting to the alternate database server. The design of such a system presents a challenge because when the connection between servers is later restored they must automatically resynchronize their data. A typical application would be for a company that has a server on each coast handling sales orders. The servers must track each other so that more items aren't sold than exist in inventory even if the two servers temporarily lose contact with each other.

What makes replication servers relevant to video mainframes is that except when playing to air most video users don't have the requirement that the replica be a perfect copy. On a broadcast server video is typically encoded at bitrates from 24 mbs to 48 mbs, while on a web server the same video content is encoded in MPEG-1 at 1.1 mbs (or less). This MPEG-1 video has a picture quality roughly comparable to VHS SLP-mode. That isn't good enough for broadcast, but is quite adequate for shot logging and rough cuts.

Combine the facts that MPEG-1 can be effectively transmitted over conventional corporate LAN's and will play back in Microsoft ActiveMovie which is included with Windows and things become very interesting. All that is required is to keep the broadcast server and the low bitrate replication server in sync. The web server recorders do this automatically. Any content changes made to the broadcast server will be mirrored in the databases and MPEG-1 files on the web server. The users access the high bitrate version for play-to-air or the low bitrate version for distribution to the desktop.

3.0 Java Applets

The hottest trend in desktop interface design is to use Java applets. Java has the advantage of running on almost any platform. Using Microsoft Internet Explorer (which is included with Windows) or Netscape enables the user to access the replication server without installing any additional software on the client machine. Java applets are programs that reside on the server but are executed on the client in the browser. For network systems administrators this greatly simplifies their world because they never need to install new versions of software to the users' desktops. Upgrades are made at the server and everyone is inherently running the same version of the software.

Java has other advantages. Unlike most languages, with Java the same software will run on multiple platforms. Windows machines, Macintoshes, and Unix-based workstations all run the same Java applets. The network administrator installs just one version of the Java applet to support all these platforms. All platforms run the same program and users avoid the headache so prevalent in the Mac/Windows world of one platform supporting a newer version than the other.

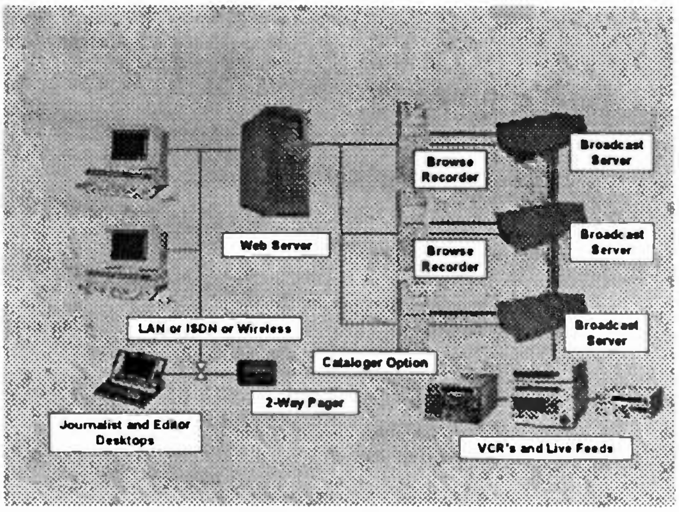
4.0 File Sizes

Even though MPEG-1 files are relatively small as video files go, these are still huge files to routinely move across a corporate LAN. An MPEG-1 file is typically compressed at about 10 MB/min. A three minute spot is therefore 30 MB in size. In studying how users manipulate video files for editing, much of their time is spent scanning for the appropriate footage and marking in and out points. Using a filmstrip representation of the video enables these tasks to be done without transmitting the MPEG-1 files.

In effect yet another replication database is kept of the video, but this time as JPEG thumbnails sampled at one frame every two seconds. Each JPEG still image is about 2k in size. A three minute spot is therefore roughly 200k in size as a filmstrip, and 30 MB in size as an MPEG-1, and 550 MB in size as the original high bitrate digital video on the broadcast server.

Keeping three copies of a video obviously takes more storage, but the relative sizes of the copies at 5.5% and 0.04% respectively makes that insignificant. In fact, holding the copies after the originals have gone to tape archive is practical.

5.0 Architecture



This architecture diagram provides an example of how these components may be laid out. Each browse recorder is connected to a broadcast video server. This browse recorder can march in lock step with the broadcast server recording the same clips as they are being ingested. Or, if the web server and broadcast server lose synchronization with each other the browse recorder can take machine control of the broadcast server to play out the missing segments or edited videos back into the web server.

The browse recorders place the MPEG-1 data and the JPEG stills onto the web server. A database is also created there to provide the data the Java applets need to display at the client desktops.

VANA is a video cataloger compatible with web-based architectures (see architecture figure). A video cataloger is conceptually the reverse of a browse editor in that it takes completed television content and disassembles it back into individual stories. It can take a network news broadcast and dice it into each individual story so that it can be keyword searched. This enables news analysts to query broadcast television content as conveniently as they would a web site. For example, the user can instantly retrieve a list of today's broadcast news stories that contain the word "president."

6.0 HDTV Implications

This replication strategy has important implications for HDTV. Replication databases can mirror video encoded at any resolution or even film. A browse editor can easily support letterboxing at 16:9 and Panavision aspect ratios. And, NTSC and PAL are both converted into MPEG-1 enabling the same browse system to support many video standards interchangeably. No matter what the source format, the result is the same international-standard MPEG-1 video files that can play back on any PC.

With HDTV moving the video around the TV station is a huge cost. Once again it requires Fibre Channel connectivity or something like it to have adequate bandwidth to move such massive files. The expense of routing HDTV video everywhere in the manner in which stations have become accustomed to with analog television has discouraged some TV stations from embracing HDTV in the near term. However, with browse editors the amount of high bitrate routing needed is decreased.

The MPEG-1 replication copies can travel across the existing corporate LAN while keeping the HDTV video mainly in the central rack room. As edits are applied to the MPEG-1 clips the EDL (Edit Decision List) is propagated back to the original high resolution video. The EDL is simply a list of the names of the video clips being assembled with their in/out points and related data. In other words, it's just a text file.

Using an EDL to synchronize edits across video replication servers means that users may detach from the network and carry the browse editor with them on their laptops. In conventional replication databases this same approach is used so that traveling sales reps can take orders without being continually connected to the server. Each day the sales rep dials up the server and reserves the stock he expects to sell that day keeping those items in a subset replicated database on his laptop. After taking orders during the day on the laptop its replication database is relinked temporarily with the main server and the two systems resynchronize. Likewise, an editor can copy to his laptop those files he expects to edit. Later the extremely lightweight EDL can be emailed back from the detached laptop to the server (even using a two-way pager) and applied against the original footage there.

7.0 Remote Clients

Handling video content efficiently makes it practical to have the browse editor user connected by ISDN or cable-modem to a distant browse server. A TV station with a news bureau in another country could access their browse database remotely to extract just the video clips they need. Even though the resultant high bitrate video edit would be pretty big, it may be cheaper to send electronically by ftp than the alternative of putting a tape in a bag and having a courier deliver it.

By trimming the video duration down to the exact footage needed, it becomes practical to browse content over conventional communications infrastructures. For instance, a TV station in Manhattan could connect to its archive in New Jersey and cherry pick just the right clips to bring back. Even if a courier (sneaker-net) approach is more economic for transporting video in a particular situation, having a lightweight remote browse capability into the archive is a tremendous improvement for selecting what footage to have fetched back.

8.0 Web Sites

TV stations are embracing web sites as a means of promoting their brand and as a secondary source of income. The act of converting content into a browser-friendly format can help repurpose video content for the Internet.

If placed on an external web server the JPEG filmstrips can be accessed easily over conventional modems. However, MPEG-1 (at CD-ROM quality) is too heavy for such low bitrate connections. An alternative is to reduce the picture quality (and size) by tweaking the encoder settings or to use a secondary encoder that converts the video into a format tailored for low bitrate Internet carriage (such as Microsoft NetShow or RealVideo). Combining the browse filmstrips with just the audio is another alternative to neck down the content to fit it on today's Internet.

As this new web-based video editing approach becomes widely available it will change the economics of television production. New uses and applications will become possible. For example, it will be simple to burn promotional reels and other video content as MPEG-1 on CD-ROM (CDR) or MPEG-2 on DVD rather than tape. Browsing and editing video with the convenience of web-based systems will help make video presentations as simple to prepare and as ubiquitous as PowerPoint presentations are today.

*BrowseCutter, VR-300, and NewsFlash are trademarks of Leitch Technology Corporation (www.leitch.com). VANA is a trademark of Science Applications International Corporation (ike.saic.com). NetShow, Internet Explorer, PowerPoint, ActiveMovie, and Windows are trademarks of Microsoft Corporation (www.microsoft.com). Netscape is a trademark of Netscape Communications Corporation (www.netscape.com). RealVideo is a trademark of RealNetworks, Inc. (www.real.com). UNIX is a registered trademark of The Open Group (http://www.opengroup.org). MPEG-1 is a standard of the Moving Picture Experts Group (www.mpeg.org). JPEG is a standard of the Joint Photographic Experts Group (www.jpeg.org).

Robin Rowe is a Chief Technologist at SAIC, a $3B international R&D company based in San Diego. He is the designer of BrowseCutter and VANA. Robin's broadcast experience includes Technical Director of Broadcast News at WICD-TV, a mid-market NBC affiliate in central Illinois. As a consultant he worked on the installation of the robotic studios at Chicago NBC station WMAQ-TV. He can be reached at Robin.S.Rowe@cpmx.saic.com.

Design and Implementation of an Integrated News Production System

Guy Eaton, Steve Owen, Ian Stewart
Quantel Limited
UK

ABSTRACT

The production and transmission of television news is a complex and demanding task. The professionals involved have a tradition of making the best use of new technology both to get stories to air faster and to get better stories to air. Recent advances in newsroom systems such as video servers, non-linear editing and automation all bring benefits to news, however the greatest advance is when all the new developments can be integrated together giving journalists unprecedented freedom to tell their stories.

Introduction

News production and transmission is a complex task involving the co-ordination of many people and technical facilities in the preparation of a bulletin. Story writing, video editing, graphics preparation, programme structuring all proceed in parallel to come together in the studio gallery as the bulletin airs. Different craft skills combine to support the journalists in their job to tell and explain the news. Different technologies are also combined to get the programme to air and recently these production technologies have taken large steps forward.

Newsroom systems need no longer be restricted to text based information, the world of multi-media PCs brings the potential of graphics and sound to the journalist desktop. Searching and browsing can expand from stories and wires to include raw footage and packaged stories.

Non-linear editing potentially frees on-line from the restrictions of tape allowing multiple versions to produced almost as easily as producing a single tape version. Editors can simply change the length of a package to fit the time available and can build stories in modules before all the shots have even arrived.

Video servers promise new transmission flexibility with the ability to add and drop stories at a moments notice and to do it with far less manual intervention than at present. Incoming lines can be recorded to disk and made available more quickly in more places than with multiple tape dubbing.

Automation systems offer the ability to manage more of the process freeing up valuable resources to front line roles.

The need for integration

Impressive though the above technical advances are, as stand-alone systems they can all promise more than they actually deliver in the real world of news. Multi-media desktops are frustrating unless the journalist can actually use the pictures available at their desktop on-air. Even more frustrating if they are available for use on air but not yet, 'Please wait...' is not a message that goes down well in the newsroom. Non-linear editing has undoubted advantages in producing the second and subsequent versions of a story but loading times mean tape can win on the critical first version. Servers offer real transmission flexibility but media must be instantly available every time or their benefits can too easily be lost in increased administration- moving people from putting tapes into VTRs to searching networks and copying files! Automation systems can suffer similar problems introducing media managers into the team when they should be simplifying the process, clearing the technology and allowing stories to get to air more easily.

The difficulties arise because to take full advantage of the potential of new technology it cannot be stand-alone, it has to be integrated together, to form a team, to mirror the team that makes the news. Without integration the people, the journalists, editors, and directors have to divert their effort into managing more technology. With integrated technology the actual technical infrastructure becomes transparent, it assists and speeds the workflow rather than becoming an additional process to manage. But what is meant by integrated technology? It is obviously not about one box but rather about ensuring connectivity and understanding between a large number of systems. And as with any system it is important to define the specification upfront.

Real world news production system goals

The main components of the news production process have already been discussed. They are a multimedia journalist desktop with the ability to browse and search video and audio as easily as text. An extension to the journalist workstation is journalist editing allowing shot selection and editing either for transmission or as a rough cut. The newsroom must talk to the automation system enabling journalist decisions to be reflected in the on-air control system. It follows that journalist edits should be available for transmission, if required, without manual intervention. Incoming lines need to be recorded and automatically made available on the journalist desktop. On-line editing, be it for sophisticated packages, fixes, or the refinement of journalist edits must be available and work reflected onto the journalist desktops. Complete status monitoring of work in progress is needed so that the news editor and director remain in control. Real time control and predictable performance are vital. It is clear that integrating news production is a far from trivial task. Several issues and their ramifications on the system architecture are critical to final performance.

Access to material In systems where more than a handful of journalists need access to video it is impractical to provide them with direct access to broadcast video, it simply costs too much. A 'shadow' server storing lower resolution desktop video allows practical and economic distribution of media to 50 or more desktops. A separate broadcast video server stores the broadcast quality video. Recordings can be made in parallel into both servers. The broadcast server must also be able to support on-line editing suites. Any limitations on access to material, for example copying it between servers, or dubbing to make an edit will degrade performance and introduce extra administration in an attempt to manage the limitations. For reliable and predictable performance instant access to any material that is currently on-line is essential. In practice this means a broadcast server with sufficient storage, ports and random access flexibility must form the heart of the system. Scalability beyond the capabilities of single servers will be dealt with later in the paper.

Journalist editing As previously discussed lower quality 'browse' video is available at the journalist desktop. This video must be of sufficient quality to make accurate editing and shot selection possible. A 1 Mbit/s stream of MPEG-1 video gives sufficient quality and is within the capability of affordable networks, for example 100 Mb/s Fast Ethernet can support more than 50 users. Technology is available that enables frame accurate and time code aware editing of MPEG-1 streams and this is used on the browse server. Journalists can select and edit shots, previewing the edit on their desktop. However it is a requirement that these edits can get to air quickly. This is achieved by passing the description of the journalist edit (a form of edit decision list, EDL) to the broadcast server when the journalist saves the desktop edit. The broadcast server then instantly 'conforms' the journalist edit and makes it available to the automation system for playout. Clearly for this to be possible the browse and broadcast server must remain in sync and more fundamentally the broadcast server must have instant access to all the broadcast media. If broadcast media is distributed across a number of servers then delays as material is moved between them are inevitable. Even if material exists on a single server the conform must happen without any need to copy material else delays are inevitable and that last minute edit won't get to air.

On-line editing Complex edits and repairs will continue to require on-line editing facilities even where journalist editing is widely adopted. The on-line editors need access to the same video available at the desktop including any edits made by journalists. If editors are to have flexibility to refine journalist edits then all the edits must remain live with tails or handles providing trim flexibility. The essential capability of the broadcast server to conform without copying allows 'live' edits to reside on the server. Instant access to these edits can only be provided if the on-line editing takes place direct on the server, else once again delays and the consequent need to manage delays have been introduced into the system. Once an on-line edit is saved the automation system needs to know so that the piece can be made available for playout.

Workflow The News Editor is in control of the process, deciding running orders and allocating staff. The newrsoom system provides the facilities that make this possible. The automation system provides a vital link between the newsroom and the technical facilities being used. As the news editor updates the running order then the automation system must follow amending playlists and flagging back status information. Close communication between the newsroom and automation system gives news editors an accurate and up to date picture of how the bulletin is developing.

Integrated news production system specifications

From the above it is obvious that there are a number of critical interactions between systems that are essential to any real world performance specification. To specify these interactions the concept of 'flight time' is helpful. Flight times define the delay between some user action and the intended outcome, the analogy arises from the data which must 'fly' via various routes from source to destination. The Integrated News and Sports Production system defines the following flight times:

Journalist edit to air:	less than 3 seconds
On-line edit to air:	less than 5 seconds
Newsroom running order change:	less than 2 seconds

Flight times are dependent upon system configuration, quoted are for a medium size system with up to 50 journalist workstations. Following the journalist edit through to air gives some insight into the tight integration of different systems and the essential capabilities of the broadcast server. The journalist presses 'save' on their desktop as their browse quality edit is finished. The browse server flags to its automation device controller that an edit has been saved and uploads the EDL. The EDL is passed via the server automation controller to the Clipbox broadcast server. The instantly Clipbox 'conforms' the edit in broadcast quality video and adds the new clip into its library. The automation system notices the new clip available in the broadcast library and that it is required in the current on-air run down. It changes the status of the clip in the run down from 'unavailable' to 'on-line' and the journalist edit is ready to air - all in less than 3 seconds.

Integrated news production system overview

The integrated news production system has been developed in close co-operation with several leading suppliers, most notably Associated Press (AP) and OmniBus. (See Figure 1.)

The journalists' newsroom system is Electronic News Production System (ENPS) from AP. ENPS was designed by journalists - for journalists. The database and search engines are uncompromising. Instead of forcing the journalist to know where to search for information, built-in information management means that the user need know only what they want. The one step 'Briefing' facility finds and presents it on the journalist's workstation - everything from scripts and news wires to browse-quality audio and video. It is clear from the system architecture that other newsroom systems can be used in place of ENPS.

Incoming raw footage, as well as completed stories, are simultaneously recorded to both a Quantel Clipbox and a low-resolution browse server. Each can have up to 100 hours of storage. The browse server is based on SpectreView from Telemedia Systems. The browse client, with built-in video editor has been developed by OmniBus Systems as another in the suite of automation applications for their OmniBus network. Low-resolution material is distributed over a conventional computer network for

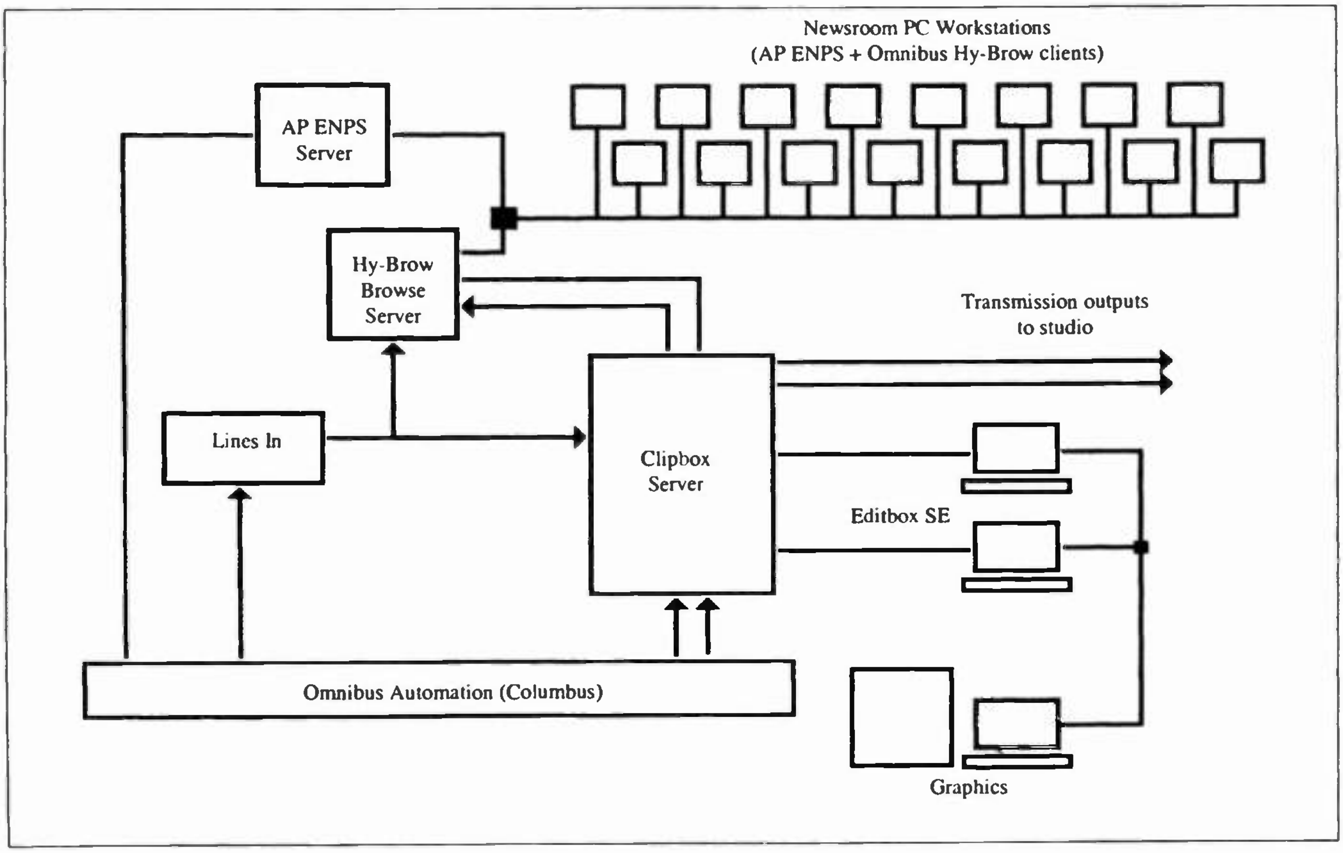


Figure 1

Intergrated News Production
System Overview

journalists to browse and edit on the desk-top. The client application is easy to use, and supports frame accurate editing. It is designed to integrate directly into the ENPS workstations, providing a single point of focus for text, video and audio in the newsroom.

Clipbox lies at the heart of the system. Its unique simultaneous true random access ensures all the bandwidth for all the users all the time. The ability to support direct server editing is the vital key which enables edits performed on the desk-top to be instantly available for play-out as complete stories - in full broadcast quality. Multiple Editbox SE on-line editors can be accommodated, all with direct access to all the material in the system. Every frame of every story – including those edited on the journalists' workstations - are available for finishing, if required. Open Picturenet allows fast access to graphic elements.

OmniBus provides automation and control. It is a distributed system based on a network connecting all controlled devices with an integrated suite of dedicated applications. Like the conductor in an orchestra, the controller initiates commands via the network, but the event itself, like that of a talented musician, is managed locally by one of any number of intelligent device controllers, or 'engines' attached or embedded in each controlled device.

Recording, browsing, routing, presentation and archiving are all provided for within an integrated suite of dedicated applications.

The benefits of integration

The architecture described above, developed by treating the news production process as a whole rather than focusing on individual elements, realises much of the potential of the new technologies outlined earlier. The multi-media desktop delivers real control to the journalist enabling them even to edit their own underlays. Where journalist editing is not desired then shot selection at the desktop speeds on-line editing.

When all incoming feeds are available on the server non-linear editing has real advantages over tape, even for that first critical edit. In fact edits can begin even before recording has finished. 'Chunking' is a parameter that describes the smallest duration of material that can be handled. When one chunk has been stored and the recording has moved on to the next chunk the first chunk becomes available for replay/editing. With Clipbox the chunk is one video frame but even the browse server can manage chunks of a few seconds duration.

The benefits of parallel simultaneous access are many. In traditional news production there is a hierarchy with TV usually taking precedence over other services such as radio and increasingly internet news. The availability of all media to all the services at the same time is a major journalistic step forward.

The integration of direct server editing allows more time for repairs and fixes, getting material to air that perhaps otherwise could not be used. Stories can confidently be worked upon until seconds before transmission.

The ability of the system to allow work in progress to be browsed gives news directors and chief editors an accurate helicopter view of the bulletin as it comes together. Stories and edits can be checked earlier whilst there is still time for change.

The combined effect of these changes is profound and can be applied by different organisations to meet their differing priorities. Stories can get to air faster, better stories can be made, efficiency can be increased, journalists can be given more control and new services can be introduced. However none of the above are a given and news providers need to consider carefully their priorities and how real world issues can affect any system implementation.

Real World Issues

One of the biggest issues to be faced when introducing an integrated news production system is that of staff training. Whilst there are obvious new skills to learn for the various groups involved, for example teaching journalists how to use their new desktop editing, there is a need for many of the groups to understand how areas outside their immediate expertise have changed. Recent experience shows that journalists need to understand on-line non-linear editing if its potential is to be realised. During a recent bulletin a live interview was being recorded so that it could be included in a closing headlines package. Although the on-line editor could start to edit the piece almost immediately, thanks to 'chunking' the item producer could not understand the concept of starting to edit before the piece had finished being recorded. At an instinctive level there was concern about 'stopping the tape' even though there was no tape involved. Changes this fundamental in the way items can be produced take time to be understood.

The role of on-line editors also needs thought. Desktop journalist editing means items can go straight to air without any on-line intervention. There is no doubt that many journalists can learn editing skills but consideration needs to be given to the journalist role of telling a story. Editing that helps tell a story is useful to the journalist, more complex editing, for example to correct video, may not fit so easily into the journalist role. The availability of powerful on-line suites is sometimes questioned within the news environment however a recent example shows their worth. Footage from a significant story was declared 'unfit for broadcast' by the regulatory authority because of the number and intensity of flashes from photographers. Many networks only took stills to air however the sophisticated on-line editing capabilities of Editbox SE was able to remove the flashes and allow the moving footage to be broadcast. Even within environments where journalist editing will become the mainstay of the operation there is a need for on-line editing facilities.

The mission critical nature of news places great demands on system reliability. It is most important that wherever possible control loops are closed. Controlling devices that issue commands need to check that the controlled device has actioned the command in the way intended. Media management quickly becomes impossible without the system doing most of the checking. The system must provide safeguards against accidental deletion of material. If something can go wrong it will. The classic example is what happens to an edit, which as has already been discussed contains only pointers to the original footage, if the original footage is deleted. Managing this at high level quickly becomes unwieldy, it is necessary for the servers themselves to understand and handle this process taking the load of the automation system. In Clipbox the internal media management database tracks each current usage of every frame in the system. Only if there are no users can material be deleted.

Scalability is another key real world issue. Most broadcasters only certainty is change. Systems must be able to grow and adapt to changing requirements. The integrated architecture relies on the broadcast and browse servers making all material available simultaneously across the system. Current technology limits are around 100 hours of DVCPRO video to 14 simultaneous users from Clipbox and about 50 users per browse server. Paralleling browse servers is already giving 200+ users and this can be further extended with due consideration to network topology. Depending on the application paralleling broadcast servers is also possible splitting the servers between the different services to be provided. The large agile server is an enabling technology for the integrated news production system, it does not prevent larger multiple server systems being built. Indeed new networking technology based upon Gigabit Ethernet minimises inter-server bottlenecks [1]. Again by managing the process at low level it reduces the load on the automation system.

A solid foundation for the future

The integrated news production system described is based upon an architecture that meets the requirements of news production. By concentrating on certain fundamentals essential for news production the integrated system has proved that the potential offered by new technology can be realised. Future developments can build upon this foundation without needing to redesign it. Newsroom developments, such as Media Object Server Protocol (MOS) [*2], will further integrate the newsroom system with the automation system blurring the distinction as far as the users are concerned between them. Native compression schemes and fast loading from tape will further reduce the time taken to get a story to air [*3]. The architecture is extensible into HDTV without modification and basic HD systems could be built today. The system represents the inevitable way forward for large system design, many manufacturers each doing what they are good at with open and tight connectivity between them. Above all the integrated news production system is empowering the journalists, using technology to hide the technology.

References

(1) Resolution Independence - A Creative Bottleneck? - J P Smith, Quantel UK - IBC 1998 Conference Publication

(2) MOS - ww.enps.com/mosproto

(3) Getting the most in news production from compressed acquisition formats - K Brindley, Quantel UK - IBC 1998 Conference Publication.

Acknowledgements

The authors would like to thank their colleagues for their contribution to this work. They would also like to thank Quantel for permission to publish this paper.

Tape Libraries in Digital Video Storage Area Networks

Steven P. Georgis
Exabyte Corporation
Boulder, CO

ABSTRACT

The migration to digital video for broadcast and post-production facilities creates a requirement for low-cost storage of long- and short-format content which can be rapidly accessed, managed and shared. With DTV requiring over 10GB of digital storage per hour of content, digital tape libraries are the only cost-effective solution for long-term storage. New generations of tape libraries combined with high-performance drives, Fibre Channel networking, and management software form the backbone of a storage area network to provide a compelling solution to the digital video storage problem.

1.0 INTRODUCTION

The advent of DTV creates stunning requirements for storage of digital video content. An SDTV system with a bit rate of 25Mb/sec requires 11GB of storage per hour of video. Hard disk drives and RAID are the preferred storage medium for buffering and short-to-intermediate term storage, but are prohibitively expensive for long-term and archival storage of large amounts of content. Consequently, digital tape libraries are generally recognized as the best cost-effective solution for long-term, high-capacity storage.

When deploying tape libraries as a central repository for video content, additional requirements are created:

- The need to have shared access to the content
- The need to physically locate storage where it can be managed
- The need for centralized management of the storage systems

2.0 TAPE LIBRARIES TODAY

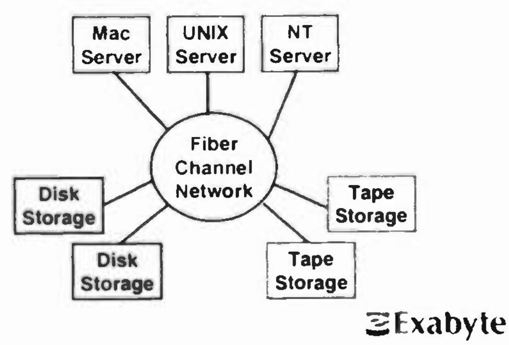
Today's state-of-the-art digital tape libraries offer all of the features required for video storage:

- scalable capacity (up to 10's of Terabytes)
- high data rates (parallel operation of drives)
- high-bandwidth network interfaces
- remote management over the network
- high reliability and data availability
- low cost per GB of storage

3.0 THE STORAGE AREA NETWORK

The Storage Area Network (SAN) is a high-speed network dedicated to shared storage devices. (Figure 1) The SAN is based on a standardized Fibre Channel (FC) backbone with a bandwidth of 1Gbit/sec. Using FC, networking components are added to permit routing and

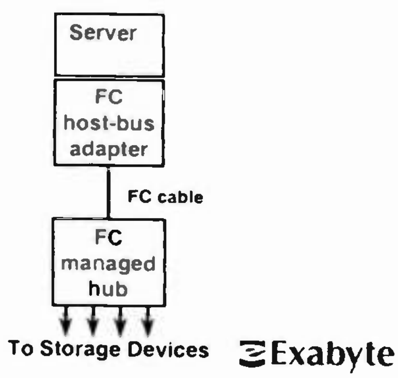
Figure 1
Storage Area Network (SAN)



shared access to the storage devices. Figure 2 shows a minimum FC network. Fibre Channel has been adopted as the interface of choice for the SAN for several reasons:

- high bandwidth (1 Gbit/sec today)
- long cabling distance (up to 20km)
- supports large number of devices (up to 126)
- protocol independence
- can transport standard SCSI command-set protocol for device and application compatibility

Figure 2
Minimum FC Network



Servers and storage subsystems plug into the SAN. Servers connect via a FC host-bus adapter (HBA). The FC protocol is hardware and operating system independent, permitting a mix of system types to connect to a single SAN (see Fig. 1).

Storage subsystems can connect to the SAN in one of two ways: native FC connection, or via a FC-to-SCSI bridge. Today, most RAID vendors offer native FC. However all tape libraries are still based on SCSI interfaces, requiring a FC-to-SCSI bridge. In the future, tape libraries will also offer native FC interfaces.

SAN Software

In addition to the hardware connections to the SAN are three levels of software requirements:
- SAN device and network management
- Storage and media management
- Content-sharing applications

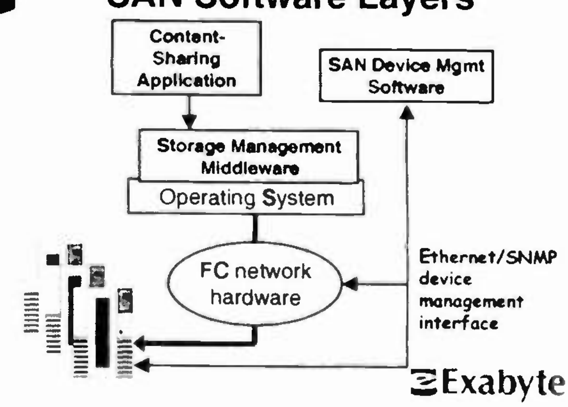
SAN device and network management is a software utility which enables configuration, monitoring and control of the SAN hardware from a single console, preferably using a Web-based interface.

Storage and media management is a software utility which allows the SAN manager to manage the stored video content and removable media from a single console. This utility also permits shared access to the storage by compatible applications. This software can be either built into the operating system, be a "middleware" layer between the operating system and the applications, or be embedded inside an application.

Content-sharing application is software which interfaces to the storage management utility to access the video content (i.e.: non-linear editing, SFX, broadcast server).

A true SAN implementation will typically require all three levels of software in the system.

Figure 3
SAN Software Layers



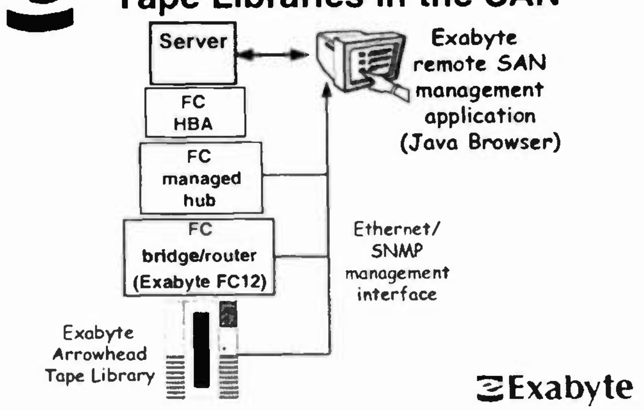
Benefits of the SAN architecture

- Puts the bandwidth-intensive storage onto a dedicated network
- Centralizes storage resources
- Scalable for capacity and performance
- Sharable storage resources
- Centrally manageable
- Configuration and connection flexibility
- Supports a heterogeneous systems environment (i.e.: UNIX, NT, Mac)

4.0 TAPE LIBRARIES IN A SAN

Figure 4 shows an example of a tape library in a SAN configuration. Besides the basic FC networking hardware, a FC-to-SCSI bridge/router is required to attach to today's SCSI libraries. A SCSI library normally includes several independent SCSI interfaces: one for each tape drive and one for the library robot. The commands and data for all of the SCSI devices in the libraries can be transported simultaneously over a single FC interface. The FC-to-SCSI bridge/router performs three critical functions in attaching the tape library to FC. The first is to convert the physical FC serial interface to the parallel SCSI interface. Second, the SCSI command protocol is de-serialized into the parallel byte-wide format. Finally, since both FC and SCSI support multiple device addresses (channels), the router section directs the SCSI commands and data to/from the target device at the appropriate address.

Figure 4
Tape Libraries in the SAN



5.0 DEVICE MANAGEMENT SOFTWARE

Figure 4 also shows the remote SAN device and network management console. This example is based on the capabilities of the Exabyte Arrowhead family of 8mm, Mammoth and DLT tape libraries, the FC12 FC-to-SCSI bridge/router, a third-party managed hub and third-party FC HBA, all combined with the Exabyte Web-based remote management software console. Web-based remote management enables the SAN to be managed from anywhere on the network. This can be attached to the server, within an existing Ethernet LAN, or across the internet/intranet. Device management from within a single application is critical to ensure the compatibility and interoperability of the various elements of the FC network.

In order to enable remote device management which does not interfere with the operation of the SAN, each of the devices in the SAN should include an "out-of-band" management port based on the standard Ethernet SNMP interface. Out-of-band management is required to prevent conflicts between the device management software and the other software applications involved in the data flow to the libraries.

6.0 STORAGE MANAGEMENT OF TAPE LIBRARIES

The storage management software layer is primarily responsible for three primary functions:

- Control of the library robots
- Keeping track of the location of each file or piece of content within the library
- Controlling access to the content, including sharing of content

Some storage management software will provide features such as tracking media usage and tape life.

The most commonly used storage management middleware for video server applications is Avalon Archive Manager™ (AAM) from Avalon Consulting (Denver, CO). AAM is currently available for leading UNIX platforms such as SGI and Sun.

With the introduction of Microsoft Windows 2000™ (NT 5.0), a middleware layer called RSM (Removable Storage Manager) will be built-in to the operating system. This will bring tape library management to the mainstream NT platform for the first time.

7.0 VIDEO CONTENT-SHARING APPLICATIONS

Applications for sharing video content in a tape library fall into two basic categories: video servers and media asset managers.

Video server applications are used to move the video data from a shared tape library to a hard disk. From disk, the data can then be directly broadcast or transmitted to another application, such as an editing station. A typical video server in this environment is a Tektronix Profile™.

Media asset managers are multimedia database applications which store the digital video content (assets) and provides a user interface with which to access that content. Media asset management is typically used to maintain an archive in a production or post-production environment. A common media asset manager is Bulldog Media Management™.

8.0 PUTTING IT ALL TOGETHER

The first steps in designing a storage area network to access shared tape libraries is the selection of the libraries themselves. The tape libraries should house high-performance drives with high-capacity media suitable for digital video content.

The tape libraries should also offer:

- scalability in both the number of cartridges and number of drives in order to meet growing needs over time
- redundancy of power supplies, cooling fans and other critical components to ensure system uptime
- hot-swap capability for the drives and power supplies to eliminate the need to power-down the system for service
- Ethernet/SNMP "out-of-band" management interface

The next step is to provide Fibre Channel connectivity with a FC-to-SCSI bridge/router. It is important to certify that the FC bridge is compatible with the library (since all are not). Depending on the number of drives in the library and the data rate of the drives, you may require multiple FC bridges to handle to bandwidth.

If more than one FC bridge is required or if you also plan to attach a FC RAID system, you will need a managed FC hub to connect the multiple FC devices.

Because interoperability problems abound today with FC devices, it is often advantageous to look to a single manufacturer for a complete certified library solution that is compatible with your storage management and content-sharing application software.

Radio Transmission Systems - Digital and Analog

Monday, April 19, 1999

1:00 PM - 5:00 pm

Chairperson: Tom McGinley
WPGC-FM, Greenbelt, MD

1:00 pm **Protecting Broadcast Facilities From Lightning**
Cris Alexander
Crawford Broadcasting
Dallas, TX

1:30 pm **Frequency Agile Medium Wave Solid State Transmitter**
David Solt
Omnitronix Inc.
North Wales, PA

2:00 pm **A Medium-wave Urban Antenna**
Grant Bingeman
Continental Electronics Corporation
Dallas, TX

2:30 pm **Digital Peak Modulation Control: An Alias Free Limiting/Filtering Method Utilizing 48 kHz Sampling and No Overshoots**
Frank Foti
Cutting Edge
Cleveland, OH

3:00 pm **Three Egyptian MW Broadcast Crossed Field Antennas**
Brian Stewart
Glasgow Caledonian University
Glasgow, Scotland, United Kingdom
F.M. Kabbarry
Egypt Radio and TV Union
Cairo, Egypt

3:30 pm **Using Synchronized Transmitters for Extended Coverage in FM Broadcast**
Bill Gould
Harris Communications Intraplex Transmission Solutions
Littleton, MA

4:00 pm **A Trial of AM/Digital Multiplexing Transmission**
Atsushi Shinoda
Kenwood Corporation
Yokohama City, Japan

Protecting Broadcast Facilities From Lightning

W.C. Alexander
Director of Engineering
Crawford Broadcasting Company

Abstract

Lightning strikes cause more damage to broadcast site equipment than any other natural phenomenon. With radio towers taller than surrounding terrain, lightning tends to be attracted to these sites. This paper will describe some of the elementary physics of lightning, describe the characteristics of a "typical" lightning strike and detail specific steps that broadcast engineers can take at the tower, tuning units, transmitter building and studio to shunt lightning currents away from equipment. It will deal with treating power lines, transmission lines, control cables and telephone lines in addition to specific pieces of equipment, such as solid-state transmitters.

1.0 Introduction

Since the first antenna tower was erected in the early days of radio, lightning has been a hazard which radio engineers have had to deal with. Man has known throughout history that lightning is an arbitrary, random and unpredictable phenomenon. Despite our ever-improving technology, lightning remains beyond man's ability to control.

The insurance industry tells us that in the US alone, lightning causes more than 26,000 fires with damage to property in excess of $2 billion. One damage claim results from every 57 lightning strikes. Worldwide, there are typically 2,000 thunderstorms in progress at any one time in which 100 lightning strikes occur each second.

While we may not be able to accurately predict or control lightning, there are steps that can be taken that will mitigate the amount of damage that a lightning strike causes.

2.0 Strike Characteristics

The type of lightning that does damage to broadcast installations is the discharge of energy from an electrically charged cloud to the ground. Cloud-to-cloud discharges seldom cause damage on the ground.

When lightning strikes the earth or an object on the earth, a more-or-less usual sequence of events occurs. First, downward "leaders" from a highly-charged thunderstorm cloud pulse toward the earth, seeking out electrical ground targets. Objects on the ground, such as buildings, trees, power lines and radio towers, emit differing amounts of electrical activity during this event. Streamers are launched upward from some of these objects. Some of the downward-going leaders connect with some of the upward-reaching streamers. It is at this point that the circuit is completed and current flows. The arc is then visible, and the superheated air displaced by it creates the thunderclap.

There are ways to dissipate static charge from a tower structure that reduce the attractiveness of the structure for a strike. These treatments are worthwhile in many cases and may be cost-effective for some installations. Most broadcast facilities do not have static dissipator arrays, however, and we must deal with the problem from another direction. If we cannot prevent a strike from occurring, the goal becomes to give its energy a place to go that bypasses sensitive equipment.

A "typical" lightning strike has a peak amplitude of 20,000 amps and lasts 40 microseconds to half amplitude. Some lightning pulses can reach 400,000 amps and reach temperatures of 50,000 degrees Fahrenheit. The rise time of a typical strike is about 5 microseconds to peak amplitude.

The current path in a lightning strike is from the cloud to what we call "ground". A perfect ground connection, however, does not really exist, and any real ground connection will have both a finite DC resistance and AC impedance of from several ohms to several hundred ohms. Applying Ohm's Law, you can see that a large potential can be developed from a ground connection to "real" ground. Several million peak volts or more can easily be developed in such a situation.

In a typical broadcast transmitter or tower site, there is a ground at the tower base and a number of other ground points. The current from a lightning strike will see several parallel paths to ground. For example, the ground rod(s) at the tower base will be one path, the outer jacket of the transmission lines through the equipment cabinets to the transmitter building ground will be another, and the AC safety ground wiring to the distribution panel ground on the tower light wiring still another path. If you can imagine an equivalent circuit of these several resistive paths in parallel, you will be able to grasp the idea that even with a solid ground at the base of the tower, large and damaging potentials can be developed across the other paths. In addition, the fast-rise-time currents that will flow in all these paths will produce large magnetic fields that will induce significant unwanted currents in nearby conductors such as AC wiring, control cables, audio lines and the like. These are sometimes the most damaging by-products of a lightning strike.

3.0 Grounding Systems

3.1 Ground Rods

The most important principle of lightning protection, as most of already know, is to provide the best, lowest DC resistance and lowest AC impedance local ground connection possible as close to the tower base as possible. This is usually best achieved by using an array of at least four ground rods driven around the tower base pier and tied together with a large copper conductor. The rods should be separated by at least twice their length, and ideally they should penetrate below the deepest frost level into the water table. Exothermic (welded, not clamped) 1/0 or larger bare copper wire to the rods, making a ring connecting all the rods and then connecting each rod with a separate length of wire to the tower base. The tower connection should also be welded. A wire connected to the tower by way of a lug or using a bolt and washer will have a lot higher resistance than a welded joint.

In some areas where the soil is particularly dry and non-conductive (such as a mountaintop with no water table and little top soil), there are chemical ground rods available to lower the impedance of the ground connection. These rods contain a chemical paste that over the life of the rod seeps into the soil into which the rod is driven through weep holes in the rod. Once the chemical paste has been exhausted, the rod must be replaced. The service time of the various chemical rods is listed in their specifications.

3.2 AM Ground Considerations

Those with AM towers should not be fooled into thinking that the ground screen and radial system provides a good lightning ground. In some areas with very conductive soil, this may be true, but in many locations it is not. A set of rods should be installed at the tower base and connected with 1/0 cable or larger to the ground side of all the arc gaps. Unless it is a certainty that the local soil is conductive enough to make the screen and radial system an adequate lightning ground, use an array of rods.

AM tower bases should also have their antenna tuning unit (ATU) chassis connected to the tower base ground rod array. Even with the best ground rod array, some portion of the current is going to flow in the parallel path presented by the tower feed tubing to the ATU. Once it hits the ATU chassis, it needs a low-impedance path to ground to prevent it from flowing through ATU components and into the transmission line. Most modern ATUs have a horn or ball arc gap right at the point where the tower feed tubing leaves the chassis. The ground side of this gap needs to be tied into the ground rod array. If your ATU does not have an arc gap at this point, you can purchase one inexpensively from Kintronic Laboratories, Phasetek and other manufacturers.

4.0 Transmission Lines

In most cases, the tower at a broadcast transmitter site is located some distance from the transmitter building. Whether this is twenty feet or several hundred, the transmission line outer conductor needs to be firmly connected to the ground rod array at the point where the line leaves the tower in FM and grounded-base AM towers. Transmission line manufacturers offer grounding kits for their various lines that provide a secure, weatherproof ground connection to the outer conductor.

A component of the current from a lightning strike that hits the top of a tower with one or more transmission lines will flow down the tower structure and a portion will flow down the parallel path presented by the transmission line outer conductors. If the transmission line outer conductors are not properly bonded to the tower structure at the top and bottom (and at the manufacturer's recommended interval along the length of the lines for long runs), large potentials can develop between the lines and the tower structure. When the potential exceeds the breakdown voltage of the outer jackets, it will arc through. Such an arc can be sufficiently hot to actually create a pinhole in the outer conductors, making a way for pressure to leak out and water to get in.

This parallel current in the transmission line outer conductor needs a place to jump off to ground before it travels into the transmitter building and into your equipment. This is why the outer conductor of every line leaving the tower needs to be bonded to the ground rod array in addition to being bonded to the tower structure.

For long horizontal transmission line runs, it is a good idea to provide one additional grounding point for the outer conductor just outside the transmitter building. This ground should be the central point of the ground array for the transmitter building. Shorter lines, where the tower is within ten or fifteen feet of the transmitter building, do not need the additional ground.

5.0 Tower RF Feeds

A piece of copper tubing is usually employed to carry the RF current from the AM antenna tuning unit to the tower itself. As already noted, this tubing presents yet another parallel current path for lightning currents.

While a path to ground for this current by way of an arc gap at the output of the ATU can and should be provided, what is in essence a "pi" low-pass filter can be created with the tower feed and the tower and ATU arc gaps, making this parallel current path very unattractive for lightning currents. This is done very simply by winding one or two turns into the tubing on a 12 or so inch diameter to form a series inductor. This series inductance will present a high impedance to the fast rise time lightning current while the arc gaps at either end present a very low impedance to ground. Adding a couple of turns to the feed tubing will often require a slight retuning of the output leg of the ATU network, but the payoff can be greatly reduced potential for ATU damage due to lightning strikes.

6.0 Arc Gaps

Arc gaps consist of two conductors spaced a certain distance apart with an air space between then. They give lightning current a path to ground on an insulated tower. When the potential between the two conductors exceeds the breakdown voltage of the dielectric (air), ionization occurs and a very low-impedance path between the conductors develops.

Commonly seen in AM installations are ball gaps, which are common at tower bases, and horn gaps, which are more commonly seen in phasing/coupling equipment and transmitters. Many times, transmitters feature gas-discharge gaps with a specified voltage rating across the RF output terminals.

Proper spacing of air gaps makes the difference between a gap providing the proper level of protection and having little effect in terms of lightning protection. At sea level, the breakdown potential of air is about 5 peak kV per 0.1 inch, or 1 peak kV per .020 inches. As altitude increases, the breakdown voltage decreases. A good rule of thumb is to reduce the breakdown voltage by 20% for every 5,000 feet AMSL.

The peak modulated RF voltage across the base of an AM tower can be calculated by the following formula:

$$V_{PEAK} = 3.182 \times Z_A \times I_A$$

Where: Z_A = antenna impedance in ohms
I_A = antenna current in RMS amps

Once the peak modulated RF voltage is known, multiply the voltage in kV by 0.020 to determine the proper ball gap spacing. Horn gap spacing can be calculated using the same method, but the sharper points on a horn gap may require slightly wider spacing.

On a practical level, the optimum spacing for an arc gap is that which is just wider than the point which produces arcs during normal full-power modulated operation. Remember that the wider the spacing, the greater the potential which must develop across the gap before the air ionizes and the gap conducts. This translates to higher voltages applied to ATU components, isocouplers and the like, which beyond a certain point will cause serious damage and result in down time.

7.0 Station Equipment Protection

7.1 Station Reference Ground

The heart of any effective lightning protection scheme is a central ground system. Such a system is shown in Figure 1. Some call this a "star" grounding scheme because of the way all the ground conductors return to a central point or reference ground. If the transmitter building is located very near the tower, this ground can be the same as that for the tower itself. In most cases, however, there will be some distance between the tower and transmitter building, and in those instances, another array of ground rods should be provided.

All conductors operating at ground potential that enter or leave the transmitter building, including transmission lines, control cable shields, and conduits, should be bonded to this ground array.

A single conductor from the ground rod array should be brought into the transmitter building via the shortest and straightest route possible. The point where it enters the building becomes the center of the "star", or the point to which everything in the building is grounded. This is called the "station reference ground". All grounds in the building, including the safety ground of the electrical system (service entrance ground) and the ground conductors from all the equipment and outlets, then connect to this point. Figure 1 is a diagram of a properly-designed station grounding scheme.

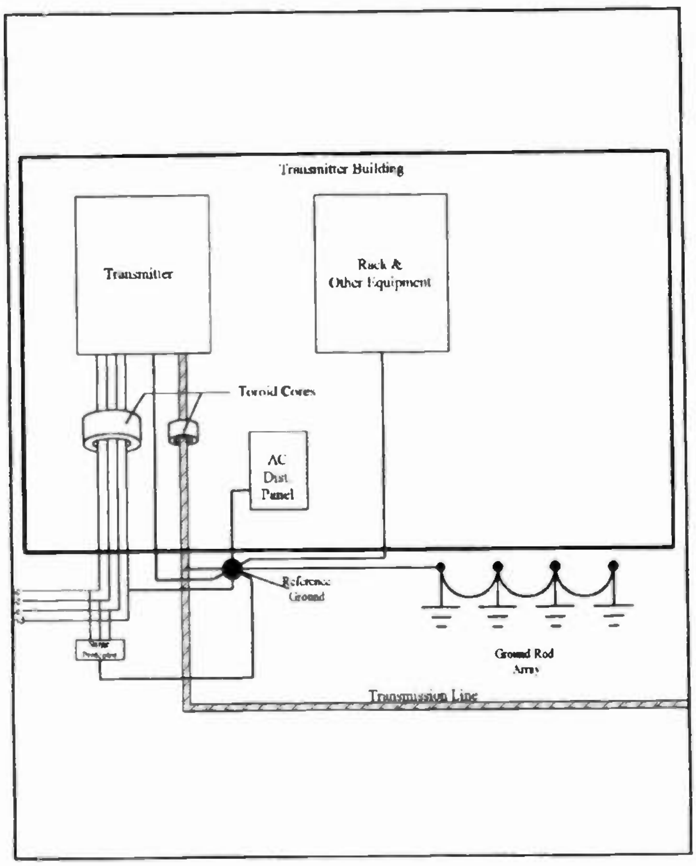


Figure 1

Beware of having a separate ground rod connected to the electrical service entrance. Such separate rods are standard practice, but having a separate rod connected can spell trouble as a huge potential can develop between the station reference ground and a separate rod outside the system.

If you have a ground strap or terminus of an AM radial ground system coming into the transmitter building from the tower(s), be sure to connect it to the station reference ground. If there is no such strap coming from the tower(s) you do not specifically need one, but an advantageous location for a transmitter building is often at the end of the ground system at the end of the transverse ground strap. If this is the case, that strap may have been extended to connect inside the building.

When connecting transmitters, racks and other equipment to the reference ground, it is important to do so in such a way that lightning currents will not flow through the equipment cabinets enroute to ground. On a transmitter, for example, make the ground connection as close to the RF output connection as possible. In that way, residual lightning currents coming in on the transmission line outer conductor can flow through the short copper path to the ground conductor and not through the metal of the cabinet. Such currents create a strong magnetic field that will induce currents into nearby unshielded conductors. By keeping surge currents out of the cabinet steel, this keeps them out of the transmitter's wiring harness as well.

7.2 Transformer Configuration

If the site has three-phase power, when it comes to lightning protection, a "wye" secondary on the utility power feed is preferable. This type of connection has several advantages, the most important of which are that every leg is referenced to ground (balanced with respect to ground) and the lower voltage (208) is easier to clamp in surge conditions.

Unless you specify a 208-volt wye, the utility company will probably provide a delta. Worse, they will probably save themselves a transformer and provide an *open* delta, which is terrible from a lightning protection standpoint.

Most all broadcast transmitters will operate just fine on anything over 200 volts, so switching to 208 volts will pose no problems. A change of taps should be all that is required. Beware, however, of the increase in current. Service conductors and disconnects sized for 240 volt operation may be too small for use at 208 volts. If undersized conductors are not replaced with those of appropriate size, a fire hazard will exist.

7.3 Surge Suppression

A good surge suppressor is the only way to minimize lightning transients on the incoming utility power. These devices range from inexpensive "kamikaze" devices that work one time and have to be replaced, to very expensive series/shunt devices. Somewhere in between is an economical device that will adequately protect the equipment at most every broadcast facility without breaking the bank.

The metal-oxide varistor (MOV) is at the heart of most shunt-type surge suppressors. These devices conduct when the potential across them exceed a threshold voltage. The devices must be rated to carry most of the anticipated lightning current. This may seem like an impossible specification, but the device only has to carry the current for a very short period of time.

Modern surge suppressors are available with fused MOVs in many voltage ratings that will hold up well under typical lightning surge conditions, clamping the AC line to ground during the surge and thus protecting equipment downstream. The fuses are designed to act slowly, holding their state for the short duration of the surge but blowing if the MOV becomes shorted as a result of excess current. The affected MOVs and fuses can then be replaced and the effectiveness of the surge suppressor restored.

Be sure to install the surge suppressor downstream of the main fused disconnect at the site. The ground connection from the surge suppressor must connect to the station reference ground. All the conductors to the surge suppressor must be relatively large, as the instantaneous currents that they will be called upon to carry can be substantial.

When it comes to surge suppression, the best policy is to buy all you can afford. If the budget can sustain a $10,000 series-shunt type, this will provide a high degree of protection. If the little "kamikaze" cans are all that the budget will stand, buy and install them. Any working surge suppression is better than none. In practical terms, the insurance deductibles and premium increases you will save may well pay for one of the more expensive units in just a few years.

7.4 Cable Protection

The final step in creating an effective lightning protection scheme is to build a low-pass filter into all your power, control and monitor cables. This is easily done by placing a toroid core over the conductors. This effectively forms an RF choke that presents a very high impedance to fast rise-time lightning energy. Such cores are available from most mail-order electronic parts houses, and they come in a variety of sizes.

One such core should be placed over each of the cables entering a transmitter cabinet or rack. Run all the AC power wires through a single toroid. Pass the remote control cable through a core, and do the same with any small coaxial feeds (RF drive, mod monitor sample, etc.). Finally, for transmission lines up to and including 1⅝", install one or more cores on the cable just above the connector.

Larger, rigid transmission lines should be installed so that they form a "trombone" section, making at least three 90-degree turns before connecting to the transmitter. The 90-degree bends also present a high impedance to lightning energy.

8.0 Conclusion

Although the focus of this paper has been on transmitter site lightning protection, the same principles can be applied at studio and other locations as well. There is no substitute for good surge suppression on the incoming studio AC power feed, and provide good grounds and install toroids on transmission lines coming in from the STL tower. The studio is a place where multiple grounds can easily exist, especially if the building has been expanded over the years, so close attention should be paid to this.

There is no way to completely

lightning-proof a site, but by taking some effective steps, the probability of sustaining significant damage and down time can be greatly reduced.. Money spent on shunting lightning away from a site and equipment will pay dividends many times over during the life of the site.

A 5 KW Frequency Agile MW Solid State Transmitter

David L. Solt, P.E., President
Omnitronix, Inc.
North Wales, Pennsylvania

ABSTRACT

Whether in the commercial broadcast industry or the non-commercial environment, the AM broadcast medium offers a service that is universally accepted worldwide. To help fulfil the need for information dissemination in portable and harsh environments Omnitronix has developed a light weight 5 KW frequency agile transmitter. This unique product can be used worldwide for refugee management, disaster relief, or propaganda. This paper will discuss the technology used in the frequency agile
5 KW transmitter.

Frequency Agile Applications

Frequency Agile transmitters are used by the military for applications involving propaganda and jamming. Frequently, the frequency agile system is used for both applications. Often, these applications require that frequency changes occur in seconds. Omnitronix was approached by the US military to design such a system based on the unique Omnitronix 700 watt broadband amplifiers. The project, SOMS-B, involved putting a frequency agile AM transmitting station into a shelter mounted on the back of a HUMVEE vehicle. Kintronic Labs developed a lightweight portable antenna to use with this system. The system application was to be fielded with the troops in environments that required immediate broadcast capability.

Several foreign governments also approached Omnitronix about a similar product. Their needs required light-weight and rugged solid state transmitters. Several systems have been installed in the Middle East where their frequency agility has allowed a rapid response, thereby protecting their citizens from clear reception of un-wanted broadcast signals.

Emergency broadcast environments have also benefitted from the frequency agility capability whereby, a single transmitter can serve as a back-up to many transmitters with minimal hassle.

Medium Wave Transmitter Topology

The self-contained linear power supply conditions all voltages required by the transmitter. The power supply contains isolation transformers, Metal Oxide Varistors (MOV's), power contactors, fuses, circuit breakers, rectifiers, filter capacitors, and the control circuitry. Front panel LED's indicate proper operating voltages and MOV status. The transmitter utilizes a step-start sequence to initiate charging the high voltage filter capacitors.

The power mains are protected from transient surges with 60 mm MOV's. Each MOV's is bypassed with a high voltage capacitor for additional transient suppression. The MOV's only absorb a certain amount of energy and then they fail short circuited. When this occurs, circuit breakers take the MOV's out of the circuit and a front panel LED illuminates. The isolation transformers are manufactured to rigorous standards to protect the transmitter from power mains high voltage transients. The isolation transformers appear to be large inductors when a power main high voltage transient encounters them. The inductance slows down the rising edge and transient energy of the impulse. Any transient energy that is passed through the isolation transformer is efficiently absorbed by zener diode transient suppressors, filter capacitors, and voltage regulators. These protection

Power Supply

circuits are in the power supply and on each printed circuit board.

The Power Supply's isolation transformers contain various taps for differing line voltages. The low voltage transformer taps are chosen so that the +18 volt DC positive supply is between +18 and +18.5 volts DC under full power transmitter operating conditions. The 300 volt DC high voltage transformer taps are selected for a 300 to 330 volt DC output under full power transmitter operating conditions. The -18 volt DC negative power supply is not critical and can range from -18 to -23 volts DC. The 24 volt DC is used exclusively in the power supply for operation of the control relays and contactors.

If desired, the customer can order the transmitter with an optional switching power supply. While this option reduces the robustness of the power supply, it does afford a significant advantage in weight reduction. This option includes a ±18 volt DC switching supply and a 300 volt DC high voltage switching supply.

Carrier Exciter

The carrier exciter is a standard Motorola synthesizer chip incorporating 9/10 KHz channel spacing. The synthesizer chip needs to be programmed with a 32-bit digital word whenever the transmitter is powered on. This can be accomplished via a microprocessor or hardwired digital logic, which is the approach in Omnitronix transmitters. It is possible to have different channel spacing as an option. The synthesizer utilizes an analog phase detector for lower noise.

The reference for the synthesizer can be the internal high stability 5 MHz reference or an external reference. There is a provision for the external reference to be locked to a standard such as the GPS satellite network. This would allow an array of transmitters to be phase locked to each other with possibilities for steered beams and synchronous broadcasting.

The MOSFET's inherent pulse distortion requires that the duty cycle of the carrier drive signal be maintained precisely. Pulse distortion is the difference in MOSFET turn-on and turn-off times. Because of the absolute nanoseconds difference in pulse distortion, and the variation in period with frequency, the duty cycle changes as the transmitter operates in different parts of the band. Typically, duty cycles at the high end of the band are in the 25% range while duty cycles at the low end of the band are in the 40% range.

The duty cycle is controlled digitally with the use of a programmable digital delay generator. Its inputs come directly from the digital thumbwheel switches on the front panel and are gated into the digital delay generator with a simple clock circuit.

Control of the duty cycle is critical in a bridge amplifier, the type used in Omnitronix transmitters, because same leg current values would exceed safe operating limits for the MOSFET's. This can be understood intuitively if one pictures the bridge as and "H". There is a MOSFET switch in each element of the "H" except for the horizontal element which is a transformer. If one of the upper MOSFET's is switched on when a lower MOSFET on the same side is switched on, there is a short circuit across the supply to ground. Hence, great care is exercised in finding the optimum switching time where efficiency is maximized and crossover operation is minimized.

Broadband Power Amplifier

The 700 watt power amplifiers utilize an "H" bridge topology with no tuned circuits. The four MOSFET amplifiers are highly efficient. They are broadband from 0.5 MHz to 1.7 MHz and are in a compact physical size that allows easy replacement while the transmitter is on-the-air.

Each amplifier module contains a PDM modulator using one MOSFET followed by a PDM filter. The PDM filter converts the digital PDM 0-300 volt signal to an high level audio signal. This high level audio signal is the high voltage feed to the RF "H" bridge amplifier.

The RF amplifier needs to see an inductive impedance at the third harmonic in order to maintain good efficiency.

Great care must be exercised to maintain an inductive impedance around 50 ohms.

The transmitter utilizes two stages of hybrid combining to achieve high efficiency across the band. The first stage of combining is a 4-way combiner that combines the outputs of four 700 watt amplifier modules. The hybrid combiner is a form of parallel combining that when used with balancing resistors allows an operator to remove power amplifier modules while the transmitter is on-the-air. The combiner requirements are defined as less than 0.1 dB insertion loss and greater than 25 dB of return loss on each port. Additionally, an isolation of greater than 25 dB was required between the input ports.

To prevent the ferrites in the hybrid transformers from operating in their non-linear region during modulation peaks, the flux density is only allowed to reach several percent of the ferrite flux saturation level. The transformers are wound with coaxial cable which affords low loss and consistent performance.

The 6 KW Combiner is similar to the 3KW 4-port hybrid combiner in every way except, with higher KVA rated transformers and coaxial cable. Care must be exercised in the design of these high power combiners to implement high frequency manufacturing techniques.

Harmonic Filter

The harmonic filter utilized is totally dependent on the application and a variety of options is available. Some applications are solely interested in light weight and,

Combiners

where this is the utmost criterion, no harmonic filtering is installed. Because of the digital nature of the bridge amplifier, the RF output of the amplifier module is a square wave. The Fourier series of a square wave indicates that there will only be odd order harmonics present. Other customers require the transmitter to comply with all applicable FCC specifications. In these cases, the broadcast band is divided into 7 sub-bands which are manually or automatically switched in line with the output of the 6 KW combiner. Most customers want something in-between and may have two sub-band filters with modest attenuation of the harmonics.

In order to provide the absolute minimum in weight, a simple low pass harmonic filter can be provided. While this does not attenuate the third harmonic from the low end of the band, there will be modest attenuation of the harmonics when the transmitter is operated at the higher frequencies.

Mechanical Stability

While the transmitter was not designed for hostile vibration environments, it is fairly rugged, having endured qualification in a shelter mounted on the back of a HUMVEE vehicle. In this application, the transmitter is mounted on shock absorbers typical of military installations.

All hardware is stainless steel for corrosion resistance and, the wire insulation is Teflon material for long life. The chassis are irridited aluminum. All mechanical construction is conservative and self locking pem nuts or lockwashers and Locktite are utilized throughout.

CONCLUSION

This paper has presented Omnitronix' new 5 KW Frequency Agile Solid State Transmitter and described its operation and technology. Various aspects of its potential application were presented. Omnitronix has made the 5 KW transmitter available commercially. Higher power frequency agile transmitters are under consideration. Further information is available at info@omnitronix-inc.com.

A Medium-wave Urban Antenna

Grant Bingeman
Continental Electronics
Dallas, Texas

Many medium-wave transmitting sites designed to cover urban areas are located in nearby rural areas, which can result in poor signal levels within the concrete and steel canyons of a city, especially at the high-frequency end of the band. The effective ground conductivity within a city may appear artificially low because of the shielding effects of tall structures. And of course the higher frequency ground-wave signals are attenuated much more quickly than those at the low end of the standard broadcast band anyway, regardless of ground conductivity.

This article presents a practical urban antenna design for use in the standard broadcast band, especially above 1605 kHz. This vertically polarized antenna does not require a ground system or conventional tower, yet produces ground-wave field intensities equivalent or superior to normal quarter-wave monopoles operating over extensive ground systems. Locating the antenna inside the urban target area helps to overcome the effectively low ground conductivity within a city.

A high-power AM transmitting antenna operating close to office buildings raises several questions, however. The high field intensity near the radiator may elicit complaints regarding RFI to telephone systems and other equipment, overloaded receiver front-ends, strange inter-modulation products, induced high RF voltages on floating metal objects, arcing between poorly bonded structural steel building members and subsequent RF noise, or possibly hazardous electromagnetic field exposure levels (or the misinformed perception of same). The existing buildings may tend to shield their inhabitants, but at the same time severely distort the radiated pattern. If this pattern disturbance is not predictable, or cannot be de-tuned there may be little point to erecting the antenna. Certainly the cost of installation is affected by the number of man-hours spent de-tuning nearby structures, and the cost of operation is affected when there are a lot of RFI complaints that have to be addressed. These and other possible difficulties need to be assessed on a case-by-case basis before significant permit application and construction costs are incurred. Clearly site selection is important for a number of reasons.

The urban antenna proposed in this article is otherwise omni-directional, and configured as an inverted monopole with a horizontal two-wire counterpoise on top. As such it can be hung from an existing building, or between existing buildings. EZNEC2 and NEC4 were used to model the antenna and its supporting structures. In light of this antenna design's probable close proximity to occupied structures, a study of the near electric and magnetic fields at 10 kW is included within this article. Power levels within the extended standard broadcast band are 10 kW daytime, and one kilowatt at night.

Benefits of this inverted monopole antenna include its low parts cost and a minimal real-estate requirement. Its impedance bandwidth is reasonable even with the electrically short examples included in this text. The inverted monopole is excited at the top of the vertical wire per Figure 1. A coaxial feeder line can be run up the center from ground, or along one of the horizontal wires from an adjacent building. An isolation coil in the feeder line is required at the far end of the antenna to keep currents off of the outer conductor of the co-ax beyond the radiating portions of the antenna, and to minimize effects on the input impedance and radiation pattern. Within the antenna structure, the outer conductor can be used as the radiating element if it is strong enough to withstand the mechanical tension. Normally however the co-ax would use a steel antenna cable for support, and the RF current would be shared between the co-axial outer conductor and the support wire.

The suspended wires throughout this article are assumed to be aluminum-clad one-half inch diameter steel. Copper clad steel would also be suitable. Earth conductivity is assumed to be 5 mS/m and its dielectric constant 13. The one-kilowatt groundwave field intensity at a mile from a standard base-insulated quarter-wave tower operating over a radial ground system would be about 130 mV/m rms over this lossy ground at 1700 kHz. This 130 mV/m value is a good reference against which to compare the inverted monopole performance. A quarter wavelength at 1700 kHz is about 145 feet in free space, and somewhat shorter on the antenna structure.

Moment method analysis using ten foot long current segments was used in all cases. The programs exploited for this article were Roy Lewallen's EZNEC 2.0, which uses NEC2, and Nittany Scientific's GNEC, which uses NEC4.1. Unless you are modeling buried wires, EZNEC is preferable because it allows input power specification and calculates RMS voltages, currents and fields, whereas GNEC presents its results in peak values based on a specific input voltage rather than on an input power. Thus GNEC requires more work to obtain useful results, although some of the embedded graphics functions are attractive.

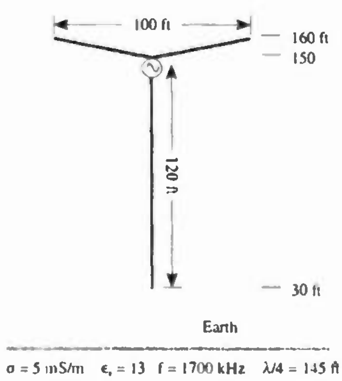


Figure 1

kHz	Fig. 1 Impedance	Resonated	VSWR	mV/m @ mi
1690	21.8 – j494.5	21.8 – j9.4	1.53	
1700	22.1 - j488.0	22.1 + j0	1.00	137
1710	22.3 - j481.6	22.3 + j9.3	1.51	

As you can see, the impedance bandwidth of the inverted monopole is a bit narrow compared to that of a base-insulated tower, because of the relatively narrow and electrically short (less than a quarter wavelength) wires. The vertical wire is only about 74 degrees long, and the two counterpoise wires are each 31 degrees long. We can reduce the reactance by increasing the length of the counterpoise wires, but first let's try adding a second vertical wire per Figure 2 to make the vertical element look wider and longer. The sideband VSWR drops as expected, even though the input resistance falls about ten percent.

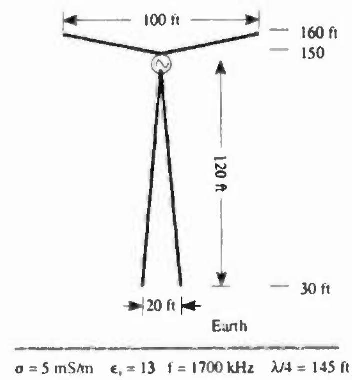


Figure 2

kHz	Fig. 2 Impedance	Resonated	VSWR	mV/m @ mi
1690	19.2 – j370.8	19.2 – j6.8	1.42	
1700	19.4 – j366.2	19.4 + j0	1.00	153
1710	19.6 – j361.6	19.6 + j6.8	1.41	

If we add some inverted top-loading in the form of a 20 foot horizontal wire between the Figure 2 vertical wire ends 30 feet above ground, we gain a substantial improvement in bandwidth, and we recover the resistance lost in the previous configuration change. Our 74 degree tall antenna is starting to look a bit longer now.

Figure 2 with 20 feet of *bottom-loading*:

kHz	Fig. 2 Impedance	Resonated	VSWR	mV/m @ mi
1690	22.4 – j349.8	22.4 – j6.6	1.34	
1700	22.7 – j345.2	22.7 + j0	1.00	152
1710	23.0 – j340.6	23.0 + j6.6	1.34	

Before we try even more *bottom-loading*, let's take a look at increasing the counterpoise length. In the interest of saving real-estate, we can fold the counterpoise wires back on themselves and still achieve the desired effects: reduced input reactance and improved bandwidth.

In general, increasing the counterpoise length serves to adjust the input reactance independently of the input resistance. However, changing the vertical element length(s) tends to affect both resistance and reactance. Theoretically one could adjust this antenna for a resonant 50 ohm condition by changing the various wire lengths, but in practice there may be more interaction than expected. In other words, the process may take longer than anticipated. I would not rule out the possibility of tuning this antenna without a lumped-parameter impedance matching network, or perhaps with only a single series component, but there is a trade-off with the amount of time required to obtain an impedance match.

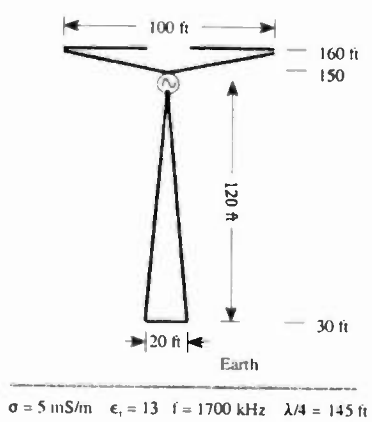


Figure 3

kHz	Fig. 3 Impedance	Resonated	VSWR	mV/m @ mi
1690	23.7 – j212.6	23.7 – j5.3	1.25	
1700	24.0 – j208.5	24.0 + j0	1.00	152
1710	24.3 – j204.4	24.3 + j5.3	1.25	

Now if we increase the amount of bottom-loading from 20 feet to 40 feet, we see another improvement in bandwidth, but a drop in field intensity, because we are now coupling more power into the lossy earth from the horizontal wire 30 feet above ground. However the field intensity is still quite a bit better than what would be obtained from a ground-mounted quarter-wave monopole operating over an extensive ground system.

Figure 3 with 40 feet of bottom-loading:

kHz	Fig. 3 Impedance	Resonated	VSWR	mV/m @ mi
1690	28.6 – j193.0	28.6 – j4.9	1.20	
1700	28.9 – j188.8	28.9 + j0	1.00	146
1710	29.3 – j184.6	29.3 + j4.9	1.20	

If we increase the length of the counterpoise wires to 60 feet each, and close the gap across the top per Figure 4 (this saves us two insulators), some additional improvement is noted. I like to see sideband VSWR less than 1.2, because then the question of impedance symmetry at the transmitter PA is no longer a significant issue. That is, audio distortion and changes in depth of modulation apparent at the output of an AM receiver's envelope detector because of unsymmetrical sideband components are minimal when the +/- 10 kHz VSWR is less than 1.20. Of course this is a moot question when a synchronous detector is used, such as is the case with a few AM stereo receivers. Unfortunately most automobile AM stereo receivers convert their synchronous detectors to envelope detectors!

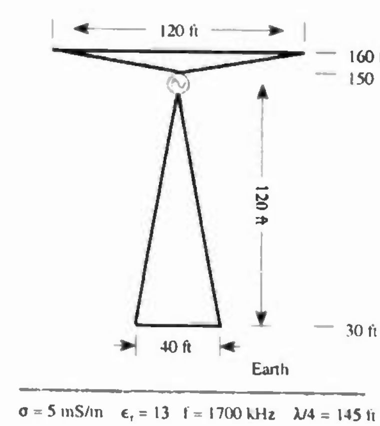


Figure 4

Figure 4 over 5 mS/m earth:

kHz	Fig. 4 Impedance	Resonated	VSWR	mV/m @ mi
1690	29.4 – j108.5	29.4 – j4.7	1.17	
1700	29.8 – j104.4	29.8 + j0	1.00	146
1710	30.2 – j100.4	30.2 + j4.6	1.17	

If we made this structure a bit larger, we could obtain an input impedance close to 50 ohms resonant, and theoretically would not need a matching network. However while physically modeling this antenna at 14.2 MHz I found that the isolation between the feeder co-ax and the antenna is a bit sensitive, so an external tuning network capable of both resistance and reactance adjustment is recommended. Independent phase rotation for impedance symmetry at the transmitter PA is not necessary since the bandwidth is quite good at this upper end of the standard broadcast band.

Another reason for having an external impedance matching network is to provide a means of dealing with the impedance changes caused by weather-related variations in ground conductivity. But the magnitude of the effect is small in above-ground antennas that have a reasonably low Q such as that of Figure 4. Assume that the local earth has a ground conductivity of 5 mS/m when dry, but 30 mS/m when saturated with water. Without the stabilizing effect of an extensive copper ground system and well-designed gravel bed drainage, the previous table is only slightly transformed. So unless the antenna is electrically much shorter, smaller, and higher Q than our example, then weather variations should not pose much of a problem.

Figure 4 over 30 mS/m earth:

kHz	Fig. 3 Impedance	Resonated	VSWR	mV/m @ mi
1690	29.2 – j107.1	29.2 – j4.7	1.17	
1700	29.6 – j103.0	29.6 + j0	1.00	206
1710	30.0 – j98.9	30.0 + j4.7	1.17	

Additional input impedance changes occur when we add support cables and insulators, which tend to make the antenna look a bit longer. In the case of our present antenna, which is operating a bit shy of its first series resonance, the support cables will tend to bring the input impedance closer to resonance. So we don't want to select antenna dimensions until we have taken these loading effects into account, which is addressed later in this article. In any case, a series inductor at the center feed point of the antenna would be useful to reduce the VSWR and stress on the co-axial feeder cable, if the counterpoise cannot be made large enough to reduce the capacitive reactance of the electrically short examples in this text. At 1700 kHz, a 100 ohm, 9 uH inductor would only require five or six turns of half-inch diameter tubing on a ten inch diameter form, so would not appreciably increase the wind-loading and weight of the antenna.

Referring to Figure 4, one can see that this antenna is ideally suited for hanging between buildings. The main suspension cable would require an insulator at the corners of the folded counterpoise wires. Another pair of insulators would be required at the corners of the bottom-loading wire, and a fifth insulator would separate the counterpoise from the vertical wires where the feed is located. An insulator may be desirable at the attachments to the support buildings, and of course in any de-tuning wires that need to be added to adjacent structures to minimize parasitic re-radiation and subsequent pattern distortion. Two anchors would be required on the ground to tension the vertical and bottom-loading wires.

The near electric and magnetic fields beneath the antenna are low near the grounded anchors, and reach a maximum near the insulators. But a person standing on the ground below the antenna would experience the highest electric field near the center of the horizontal wire. For ten kilowatts input to the antenna of Figure 4, the maximum electric field six feet above ground (24 feet below the bottom-loading wire) is about 160 rms volts/meter, which is mostly in the z or vertical direction. The 10 kilowatt maximum magnetic field six feet above ground is less than 0.1 rms amps/meter, which is mostly in the x and y directions, parallel to the ground. The maximum electric and magnetic fields do not occur in the same location, and the power density is not likely to exceed a few milliwatts per square centimeter. Both fields are smaller closer to the ground, and greater closer to the antenna. These results assume there is no system of buried ground wires or concrete pavement with embedded steel, but that the metal anchors do extend a few feet into the earth. Thus it appears that pedestrian traffic beneath this particular antenna would not be subjected to electro-magnetic fields in excess of safe limits, which at this frequency are 614V/m, 1.63 A/m or 100 mW/cm^2 averaged over six minutes. However the results I obtained are constrained by my assumptions, and before pedestrian traffic is allowed near the antenna, a rigorous set of field intensity measurements would certainly have to be made with calibrated test instruments under various weather conditions (wet, dry, frozen, etc).

What would happen if an automobile were to park beneath the antenna? Assuming the car is insulated from ground by its tires, what level of RF voltage could be expected to exist between the car's chassis and ground? Would the antenna input impedance and radiation pattern also be affected by the presence of the car? If we model the car as a box 12 feet long, 6 feet wide, and 4 feet tall hovering one foot above ground directly below the center of the antenna, the input impedance and far-field pattern do not change. The near-fields are perturbed in the immediate vicinity of the car, but overall remain the same. Since there is no conductor in the earth against which to measure the car's RF potential, a passenger stepping from the car to the earth is not likely to set his shoes on fire. However he may feel something on his fingertips if he is touching a conducting portion of the car when he steps onto the earth, depending on the conductivity of his shoes (leather, rubber, bare-feet) and the moisture content of the earth. Another concern would be the possibility of damage to the front-end of a car radio. Therefore it may be necessary to prevent automobile access to the area immediately beneath the antenna.

If the Figure 4 antenna were moved up another 70 feet above ground, the far field intensity would increase to 170 mV/m at a mile, the input resistance would drop significantly and the reactance would become slightly more negative. The 2.3 dB increase in gain compared to a quarter-wave monopole is naturally paid for with a decrease in bandwidth. There is also a high-angle radiation lobe evident in the elevation pattern (Figure 5). So careful consideration must be given to the overall height of the structure above its ground image, since an above-ground antenna behaves more-or-less as a vertically stacked array of two antennas. If the vertical space is available, consider increasing the length of the vertical wires to bring up the resistance and reduce the reactance, but not until you have modeled the support cables and insulators, plus the adjacent buildings and their de-tuning arrangements.

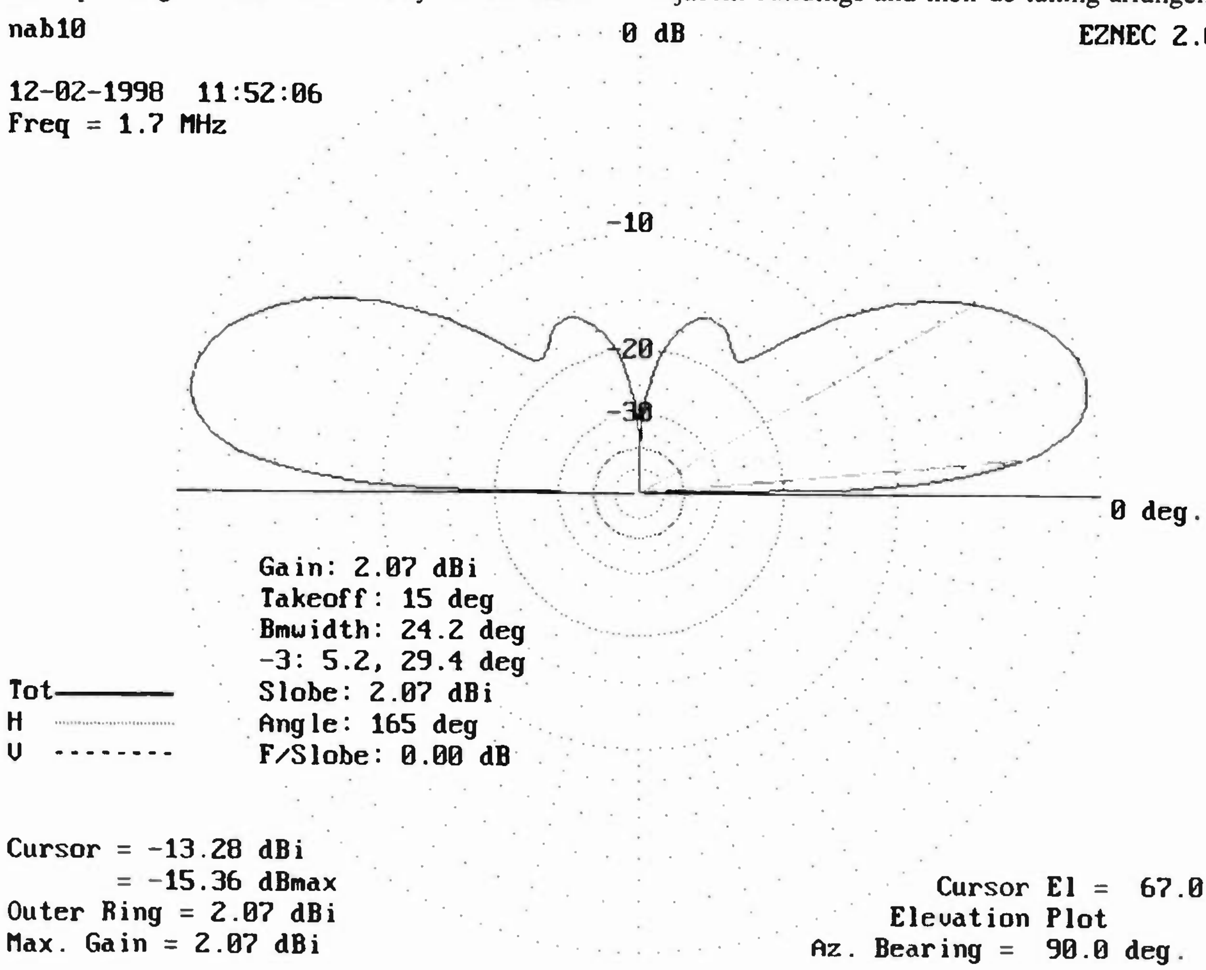


Figure 5

Often the additional loading afforded by the support cables will make the antenna look electrically longer anyway, so you may end up with a 50 ohm feed impedance without having to increase the antenna dimensions. The current in the support cable insulators may be small, but it can significantly affect the input impedance to the antenna, since it changes the current distribution on the principal radiating elements, and allows some current in the structural elements.

Figure 4 raised 100 feet above ground:

kHz	Fig. 4a Impedance	Resonated	VSWR	mV/m @ mi
1690	19.4 – j116.6	19.4 – j4.8	1.28	
1700	19.6 – j112.5	19.6 + j0	1.00	170
1710	19.9 – j108.4	19.9 + j4.8	1.27	

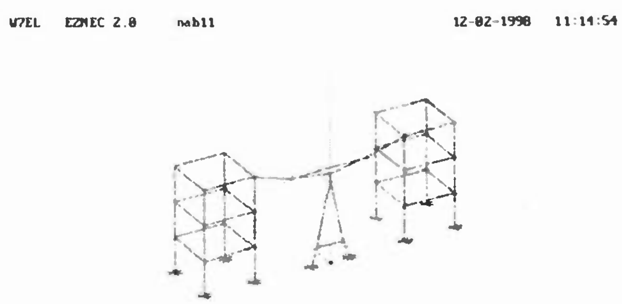


Figure 6

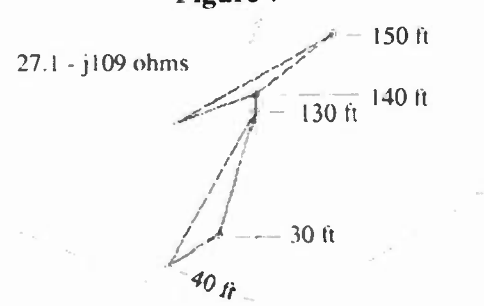
Returning to the antenna of Figure 4 (30 feet above ground), let's add the support wires and insulators per Figure 6 to see how the voltage stresses look across the insulators, and what kinds of loading effects occur. Note the simplified building structures nearby which are supporting the antenna. The resulting azimuthal pattern of Figure 8 tells us that re-radiation or scattering from nearby buildings could be a major problem from an urban antenna site. The input impedance also changes very significantly from 30 - j104 ohms (Figure 4) to about 40 - j47 ohms. If we were to add some outrigger de-tuning wires to the buildings, the impedance effect would be reduced, but not entirely since some of the impedance change is caused by incidental loading via the insulators and support cables. The insulators and the support cables are part of the antenna, and should not be ignored, since they have RF currents in them that can contribute substantially to the radiated fields. Also keep in mind that the moisture content of concrete buildings may change according to the weather, and this can affect antenna performance as well.

If the insulators (square symbols in Figure 6) look like -j10k ohms, then for one kilowatt input to the antenna, the peak insulator voltages are as follows:

location	carrier	1 kW, 125% AM	10 kW, 125%AM
feed	440 volts	990 volts	3130 volts
building	410	920	2910
upper	1260	2840	8970
bottom	1950	4390	13900

Note that these voltages increase by 2.25 times the carrier condition when modulated 125 percent. The insulators closest to ground have the highest voltage stress in this particular antenna at 1700 kHz. Thus you would need bottom insulators with a wet working rating of at least 6 kV peak to be safe at one kilowatt, or 20 kV peak at 10 kW. These figures include some safety factor for the inevitable accumulation of dirt on the insulators. If your site is more than a mile high, you should also add some safety factor to allow for the lower breakdown voltage of the thinner air at higher altitudes.

Figure 7



This inverted monopole can also be hung at a slant from a single building, if the counterpoise is hung below the roof soffit. The radiation pattern does include a horizontally polarized component, but in the case of Figure 7, this is 23 dB below the v-pol component, which is still omni-directional in the azimuth plane. Figure 7 is the same as Figure 4, except the top of the antenna is 150 feet off the ground (ten feet less than Figure 4), and the bottom-loading cable is displaced 40 feet from plumb, but still 30 feet off ground.

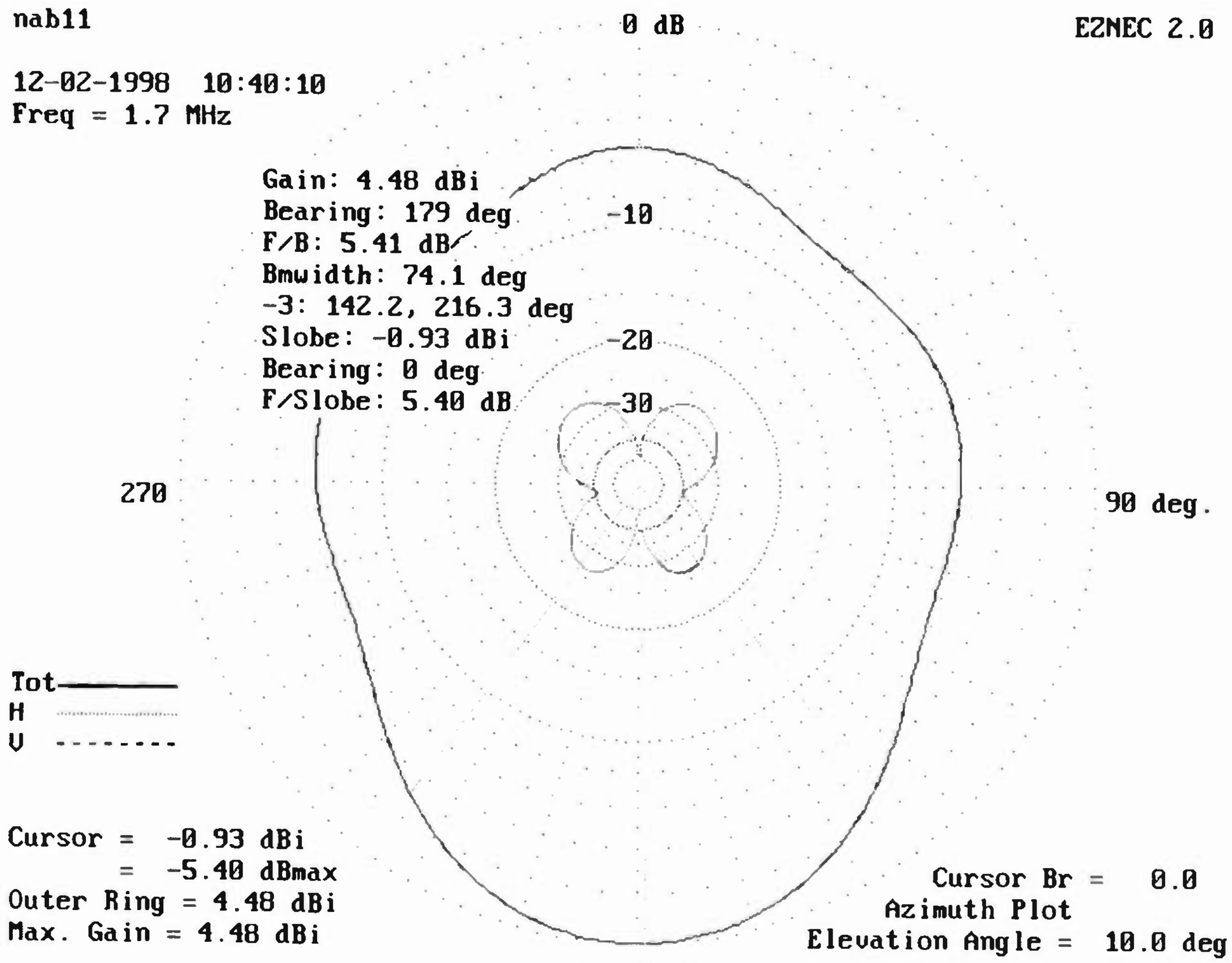


Figure 8

In conclusion it appears that the inverted monopole is a viable radiator for limited-space applications if existing support structures are available from which to hang the antenna. However in some situations 10 kilowatt operation may create RFI problems for existing tenants, so operation as a low-power repeater may be more practical, such as the drop-wire installed by Hatfield and Dawson in the Seattle Kingdome in 1978. As in all antenna installations, each site must be carefully considered in order to arrive at a design which produces the optimal performance for the money over the operating life of the station. If there are new or unique aspects to the design, then careful planning and analysis are critical to a timely and cost-effective outcome. Risks must be weighed, and the experiences of previous inner-city transmitter plants should be researched. For example, I heard of a case where an emergency horizontal dipole antenna and one-kilowatt alternate transmitter were used at the studio location of a station when its main transmitter plant was temporarily out of service. There were RFI problems, and this antenna was eventually removed. But perhaps a broadcast site in an industrial portion of the same city could have co-existed peacefully with the indigenous daytime population.

DIGITAL PEAK MODULATION CONTROL: AN ALIAS FREE LIMITING/FILTERING METHOD UTILIZING 48KHZ SAMPLING AND NO OVERSHOOTS

Frank Foti
Cutting Edge
Cleveland, Ohio

ABSTRACT

Peak modulation control is easily accomplished in the analog domain. The digital counterpart requires more consideration and sophistication, as sampling rate, dynamics generated aliasing distortion, and the transmission medium all play an important part. There are questions and concerns about the utilization of 32kHz sampling as the sole means of interconnection in FM transmission. Real world experience shows that whenever peak limiters sampled at 48kHz are integrated into a 32kHz sampled transmission path, problems such as overshoots arise. This presentation offers mathematical reasons why this occurs. Discovery is presented about a digital peak limiter that eliminates overshoots, enabling a 48kHz sampled limiter to exist in a 32kHz sampled environment.

NOT EXACTLY...

The radio broadcast transmission path, in the all-digital domain, has been marketed as a *plug and play* system: Just purchase the processor, STL, and exciter of your choice, connect the gear together via AES/EBU, and away you go. A wonderfully clean, loud, and precise peak controlled signal, all rolled up into one! Well, as one rental car commercial tells us..."Not Exactly!" While some of us might be fooled by the mystic wonders of what a digital system brings to the table, the rest of us have found, or are finding out, that the digital system has some *issues* of its own. It's how we choose to deal with these issues that determine whether or not the digital transmission path is providing added benefit, or is merely a weak clone of the analog predecessor.

Within a digital processor and transmission system, numerous factors can affect the absolute peak control. Some of these are related to processor design, and others are a product of overall system performance. Sample slipping, sample rate converters, low pass filters, and emphasis can all contribute to generating overshoots in a transmission system. Later in this presentation, you will see how the implementation of these functions and their placement within the system will have an important effect on peak control.

It is imperative that precise peak control be attained or loudness will be lost due to overshoots. Most countries institute critical modulation limits and restrictions, deeming that any overshoots that occur must be compensated for by reducing the overall modulation by an equal magnitude. Therefore, it is mandatory that any overshoot be minimized–usually to 2% or less. This ensures maximum modulation density, which yields increased perceived loudness. The following sections discuss each of the aforementioned items and how they create overshoots.

SAMPLING RATE AND ALIASING DISTORTION

Recently, there has been significant discussion and debate within the broadcast industry about sampling rate for transmission purposes. At issue is the choice of either 32kHz or 48kHz sampling. Both will work, yet each bring different benefits and caveats to the table.

The choice of 32kHz sampling was employed in older digital transmission systems. At the time, both DSP hardware and service bandwidth for signal transportation were at a premium, as is, the FM Stereo broadcast system only uses 15kHz audio bandwidth. So, theoretically, the 16kHz Nyquist frequency of the system will fit in an efficient manner and maximize spectrum space in the digital system. While this looks very nice on paper, real-world performance has indicated otherwise.

Research has shown that the 32kHz based dynamics processors generate a high level of aliasing distortion due to weaknesses in their final peak control methods. Even when over-sampled, these systems still generate aliasing products that are audible in program material. A recent AES paper[1] detailed research and testing on this matter, and strongly recommended that dynamics processing should employ at as high a sampling rate as possible to reduce dynamics generated aliasing artifacts. In addition, the severity of slope for the 15kHz low pass filter will affect the sonic performance. This will be covered in a later section.

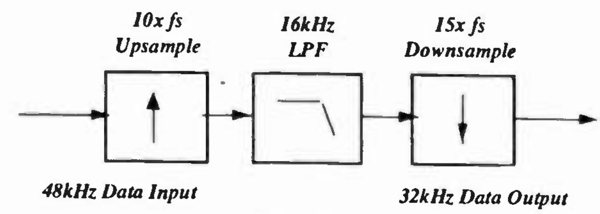
Through virtual up-sampling methods, innovative final limiter algorithm design, and sampling at 48kHz, the dynamics generated aliasing problem is reduced to insignificance, thus rendering it inaudible. Yet a concern of a 48kHz sampled system is this: How well will it interface to a lower sampled transmission system? Converting the sampling rate is not the issue, that's easily done using a device known as a *Sample Rate Converter*. The issue for discussion is what happens to the peak control integrity when the conversion process is applied.

Sample Rate Converters (SRC)

This device transforms one system sampling rate to another, which is necessary when interfacing digital equipment that uses different sampling rates.

The conversion is accomplished by scaling up, or interpolating, the incoming signal by a factor that allows the *desired* rate to be divisible by the system's internal rate. The signal is then low pass filtered at the Nyquist frequency of the desired sampling rate. This filter is required to smooth out the added samples, which would otherwise create aliasing products. Finally, the signal is scaled down, or decimated, by the factor needed to achieve the desired rate. When converting from 48kHz to 32kHz sampling rates, for example, a 10x multiplication rate will up-sample the incoming signal to 480kHz, which can then be divided by 15 to achieve 32kHz. *Figure-1* shows a block diagram of a SRC.

While this sounds quite simple—and basically it is—there are a few issues to consider. Of primary interest is the interpolation filter, typically an FIR filter. It must provide a stop-band rejection of 96dB at 16kHz in order to suppress aliasing distortion. A steep slope in the transition area will be required.



Sample Rate Converter
Figure-1

Since nearly all audio processors apply some form of overshoot control in conjunction with the output filtering section, the overshoot component can be determined by the Gibbs Phenomenon[2]. Should the slope of the up-sampled interpolation filter be greater than the slope of the final filter in the audio processor, then output overshoots may result in the sample rate conversion process. But since these overshoots are generated after the audio processor, removing them requires another limiting device—thus a need for an added limiter further downstream in the system.

Of interest is the direction of rate conversion. When converting from a lower rate to a higher rate, the chances of overshoot are small, because the frequency of the up-sampled filter is set to a higher frequency than the Nyquist frequency of the incoming signal. Overshoots are a significant problem only when transforming a higher rate to a lower rate, as described above.

Our testing has shown that the use of sample rate converters in the digital audio path between an audio processor and exciter will cause overshoots whenever down-converting from 48kHz to 32kHz sampling. In our test lab, we have also confirmed that any 32kHz sampled system, for transmission processed audio, must have tight low pass filtering at the Nyquist frequency of 16kHz. Any non-linear products that exceed 16kHz will cause overshoots in succeeding SRCs or additional low pass filter stages. This constrains any processing system, regardless of sampling rate, to a tight 16kHz bandwidth. Because of this restriction, it renders other benefits of a higher sampling rate useless. Why should this penalty be paid? It has been proven time and again that 48kHz sampling is a superior rate for digital audio, even when a lower audio bandwidth, such as 15kHz for FM Stereo broadcasting, is used.

It is possible to implement a 15kHz low pass filter with a tight 16kHz stop-band in a 48kHz sampled system. There is no problem with doing that. There is however, a subjective and sonic choice for using a filter with a broader slope, as it sounds better. Consider the landscape in the analog processor/transmission system. The need for the 15kHz low pass filter is to protect the 19kHz pilot frequency. Thus, analog low pass filters all were designed to create their stop-band somewhere around 18kHz. Creating an analog low pass filter with a stop-band at 16kHz is theoretically possible, but quite difficult in the real-world, as component tolerances and group delay issues are a problem. Using that example, it shows that low pass filtering with a stop-band beyond 16kHz is not detrimental to the FM Stereo system used in broadcasting, as it has been done for many years and with no problems. It's only important for usage in a 32kHz sampled transmission system.

Another point to consider: All existing analog processing systems, when digitized and connected into a 32kHz sampled transmission system, will overshoot! This is due to the same reasons as stated above, where energy beyond 16kHz will ring in the low pass filters of the transmission system or SRC, if a conversion is being applied. Considering that analog processing equipment will not become extinct overnight, this problem of 32kHz sampled transmission system overshoots exists even when a hybrid of analog processing is coupled to a digital transmission path.

Therefore, those who argue about the strict use of 32kHz sampling for processing and transmission purposes have given no regard to older technologies that must continue to exist in today's environment. In essence, the 32kHz sampled system is not backward compatible, whereas a 48kHz sampled system is, as far as overshoot control is concerned.

The following persistence display is of a digital oscilloscope that measured the output of the mpx test point in a Harris Digit[R] FM Exciter. The test used an Optimod[R] 8100 connected to the exciter through a Symetrix[R] 20bit A/D converter. This is the exact same configuration that a radio station choosing to use an analog FM processor and digital exciter would have. This test configuration was setup as follows:

- Program audio connected to 8100 processor.
- Left/Right Output "Test Jacks" of 8100 connected to A/D converter inputs.
- A/D converter AES/EBU output connected to digital input of exciter.
- Test Point J-1 of Digit mpx board connected to scope.

Following is the result of that test:

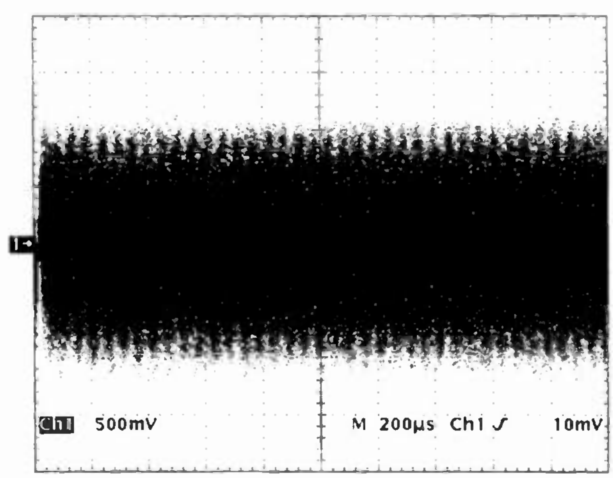


Figure-2

As *Figure-2* shows, overshoots occur! Thus proving the point that subsequent filters in SRC's and the exciter are the culprits as the bandwidth control in the 8100 does not have the required stopband suppression at 16kHz! The above illustrated problem exists today with every radio station that is choosing to use an analog processor that operates in a 32kHz sampled transmission path.

Remember that a system using 48kHz sampling does not require tight filtering at 16kHz; it must provide tight filtering at 19kHz to protect the pilot frequency and the remainder of the composite spectrum. Thus, there will be some non-linear products beyond 16kHz which could overshoot when down-converted to 32kHz sampling. Note that 32kHz sampling is not a standard for the FM transmission system, nor was it intended to be. A recent AES Journal[3] recommendation for sampling rate instructs that 32kHz *may* be used for broadcasting, but it does not suggest it as a standard. Digital FM exciters should be able to accept a 48kHz sampled signal and modulate it without generating any overshoot. The broadcaster should not be penalized for desiring to use 48kHz sampled systems in the transmission path of their radio station!

Since this sampling rate issue is primarily based in the broadcast environment, here's another perspective to add into the mix: Digital Audio Broadcasting (DAB) specifies a 20kHz audio bandwidth. Since many existing broadcast facilities will, no doubt, employ this technology when it becomes available, they will need to provide transmission systems capable of 20kHz bandwidth. Each of the proponents who argue in favor of 32kHz sampling will be out in the cold with regard to DAB. So here is another reason to embrace 48kHz sampling throughout the broadcast facility and transmission path.

DIGITAL TRANSMISSION SYSTEMS AND THEIR EFFECT ON SOUND QUALITY AND PEAK CONTROL

FM Exciters

These are the latest entry to the digital audio transmission path. Capable of exceptional modulation performance, they offers two forms of signal input: Analog composite (MPX) for the non-digital transmission site and AES/EBU.

The composite input connects to the modulator by way of a high speed A/D converter, and requires a faster sampling rate than normally used for the discrete channels. Since the modulation spectrum for FM can range up to 99kHz, the exciter must use a sampling rate of at least 200kHz, for a Nyquist at 100kHz, which covers the baseband spectrum.

The AES/EBU input accepts the signal in the discrete left/right format. Thus, the exciter must perform the stereo generator function. Here is where the story gets interesting.

Consider the AES/EBU input signal to the exciter. It might be at a different sampling rate than that of the exciter. If so, a sample rate converter is employed to make the proper transition. This can pose problems, as the digital filter within the sample rate converter can generate overshoots, adversely affecting the tightly peak-controlled audio data being converted.

The audio, having already been emphasized, peak-controlled and band-limited by the audio processor, needs only matrixing and MPX encoding for stereo modulation to occur. But what's present in most digital exciters is a sample rate converter, another low pass filter, and in some cases, the addition of, yet again, pre-emphasis. There is a final limiter included in some digital exciters to help remedy some of these overshoot problems, albeit with adverse sonic consequences.

In essence, the signal that only needed to be matrixed and MPX encoded now has additional conditioning applied to it which can degrade sonic performance and modulation efficiency. To learn why, let's review low pass filters and emphasis networks.

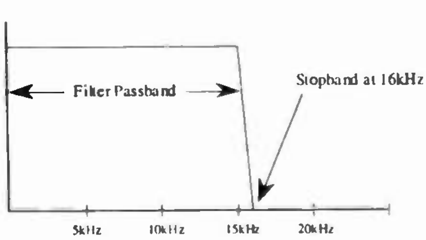
Low Pass Filters: The Sonic Effects

Not all low pass filters sound the same, even when they are designed to the same cutoff frequency and are of the same type. Differences in their transition range will affect how they sound.

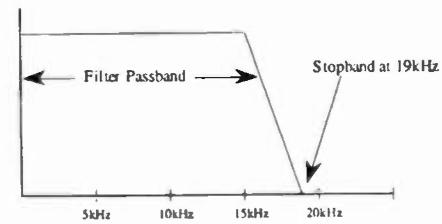
Let's take a look at a few of the restrictions in using tight low pass filtering to provide stop-band rejection at 16kHz in a 32kHz sampled system. The stop-band must provide 96dB of rejection at 16kHz in order to be effective, or aliasing distortion will result. To achieve a filter of this magnitude along with phase linear group delay, a FIR filter is used. With the design specification of 96dB stop-band rejection and 0.1dB passband variance, an equiripple style of filter is suited for the job. Unfortunately, this filter will require 119 taps to create the tight slope that provides 96dB of stop-band rejection. A filter of this length will create *1.8ms* of throughput delay.

By contrast, a 15kHz low pass filter designed to provide 96dB stop-band rejection at 19kHz to protect the pilot, and operating at 48kHz sampling, requires only 47 taps. This generates a throughput delay of only *0.47ms*, almost 4 times less than the above-mentioned filter. When we consider that time delay in digital transmission systems is a cumulative function, every millisecond counts, as it can add to the comb-filter effect that disk jockeys perceive when monitoring themselves off-the-air.

Another aspect to consider is the sonic differences of filters with different slopes. Psychoacoustic tests have proven that the transition slope of a filter will affect the timbre of the audio. As the filter slope is made tighter, "ringing," which degrades the clarity of the audio, is increased. Therefore, a low pass filter which utilizes a gentler slope is sonically superior. *Figures 3-4* depict the differences in the slopes of 2 different 15kHz low pass filters.



Example of Tight Slope
Figure-3



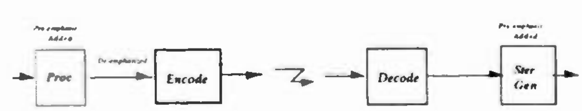
Example of Broader Slope
Figure-4

The design of transmission equipment is both simplified and results in superior performance with a 48kHz sampling rate. Down-conversions don't affect tightly-controlled audio, and filtering requirements are eased.

Pre-emphasis/De-emphasis Considerations

Most exciters let you add pre-emphasis. Optimally, however, the addition of pre-emphasis is best left to the audio processor, as it employs specialized high frequency control sections that provide both the boost and control of the high frequency energy. In this manner, high levels of modulation are easily obtained, since the processor is designed to balance the tradeoffs between pre-emphasis and high frequency limiting.

In situations where a codec-based STL system and audio processor are inserted before the stereo generator, the codec must pass "flat" (non pre-emphasized) audio. This requires adding de-emphasis to the output of the processor; pre-emphasis is then re-applied in the exciter's stereo generator. *Figure-5* illustrates this:



Codec-based STL System
Figure-5

A flat signal is required by the codec because of its reliance on masking principles. Any significant change or imbalance of the frequency spectrum can cause the codec to expose artifacts that would normally be masked.

Whenever multiple stages of frequency contouring are applied, the phase response of all stages must match, or overshoots will result. To eliminate the added overshoot, another limiter must be employed as a "band aid" in the exciter. Even though emphasis networks are derived from a first order filter process, it is possible to create networks that may not match up in phase with each other.

The following are the formulas for determining emphasis response that correlate to first order analog RC networks for pre-emphasis and de-emphasis. Specific frequency gain, along with phase response, can be calculated. Any emphasis networks implemented in DSP should follow these calculations:

To Calculate Pre-emphasis/De-emphasis:

$$\text{Ratio} = \sqrt{(2\pi f_r T)^2 + 1} \qquad \textit{(equation - 1)}$$

$$\text{dB} = 20\log \text{Ratio} \qquad \textit{(equation - 2)}$$

Where: Ratio = Emphasis gain at a given frequency
f_r = Audio Frequency
T = Time in milliseconds (50μs or 75μs)

Equation-1 is used to calculate the gain at a specific frequency along the emphasis curve. *Equation-2* converts the gain ratio to dB. Taking the reciprocal of the ratio in *equation-1* provides the calculation for de-emphasis. These values represent the exact response that is obtained in an emphasis network that is implemented with a single pole RC filter in the analog domain. The phase relationship for both pre-emphasis and de-emphasis are represented in *equation-3*.

$$\varnothing = \text{atan } 2\pi f_r T \qquad \textit{(equation - 3)}$$

Where: Ø = Degrees of phase shift at a given frequency
f_r = Audio Frequency
T = Time in milliseconds (50μs or 75μs)

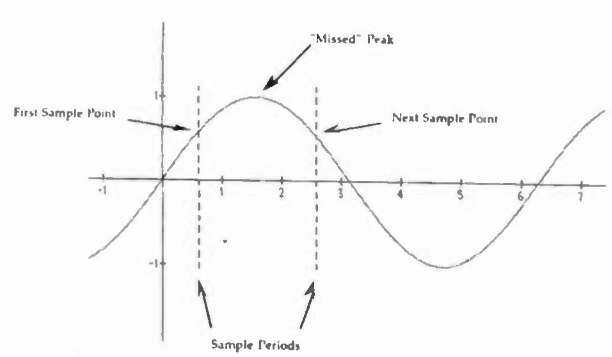
Unless each processor, STL, and exciter manufacturer follow these same equations when designing emphasis networks, the resultant phase mismatches will cause overshoots.

Based upon the previous discussion, you can see why it's best to install the audio processing system as close to the exciter as possible and to use the processor's pre-emphasis. By doing so, internal limiting in the exciter becomes unnecessary and allows the processing system to provide all of the required peak control.

Sample Slipping

Within a digital system, the resolution of the audio data is determined by the number of samples for a given frequency. Lower frequencies will be sampled more often than higher frequencies. According to Nyquist theory, there will be at least two samples at the highest frequency. This does not leave much resolution when trying to determine the exact peak level at the upper portion of the spectrum, since the two sample points can occur over a 360 degree range. If this happens within the hard limiter algorithm of an audio processor, overshoots will result!

When hard limiting is performed, the precise level of the upper frequencies in the spectrum can be missed, as some of their peaks will occur between sample points. Should these peaks exceed the threshold of the clipper, what the final output level will be after the clipping function is performed becomes uncertain. This is technically known as *unquantized intra-sampled peaks*, or *sample slipping*. *Figure-6* shows a worst-case example of this:



Worst-case example of unquantized intra-sample peak
Figure-6

Notice how the *missed peak* reaches its crest factor exactly between the two sample periods. At each sample point, the value that is registered as data is significantly less than the peak value. If this missed peak is at a level that would cross the clipper threshold, nothing would happen, as the clipper is not aware of it. The problem is most severe when the signal in question approaches the Nyquist frequency. We can calculate the acquired level, and hence the error between the acquired level and the peak level, by using the following equations:

$$\textbf{Phase } \emptyset = 360 * (f_a/f_s) \qquad \textit{(equation - 4)}$$

Where: Ø = Degrees between Upper Audio Frequency & Sampling Rate

f_a = Upper Audio Frequency
f_s = Sampling Rate

$$\textbf{Acquired Level (A)} = \cos(\emptyset/2) \qquad \textit{(equation -5)}$$

Let's have a look at some examples. With 32kHz sampling and a test frequency of 15kHz (the upper bandwidth limit in FM broadcasting), the acquired level equals 0.098, or 10%.

$$168.75° = 360 * (15\text{kHz}f_a/32\text{kHz}f_s)$$

$$0.098A = \text{Cos}(168.75°/2)$$

In other words, there is less than 10% level acquisition, or 90% detection *error* in a 15kHz peak, sampled half-way between two samples with 32kHz sampling. On the other hand, with 48kHz sampling, there is 55% level acquisition, or 45% error. A 128kHz system generates 7% error. In a virtual 192kHz sampling method, there is 97% level acquisition, which generates only 3% error. *Table-1* summarizes the effect of sampling rate on the efficacy of peak acquisition:

Sampling Rate	Acquired Level A	% Acquisition	% Error
32	0.098	10	90
48	0.55	55	45
128	0.93	93	7
192	0.97	97	3

Summary of Peak Acquisition Error as a Function of Sampling Rate
Table-1

This is why using a higher sampled system will reduce this problem to insignificance. Thus, when an audio processing system is being evaluated for any tendency for peaks to "slip between the samples," you only need to determine what sampling rate is used. Furthermore, a peak limiter will control peaks regardless of the sampling rate, and nothing will "slip between the samples," as long as a clever limiter algorithm is utilized.

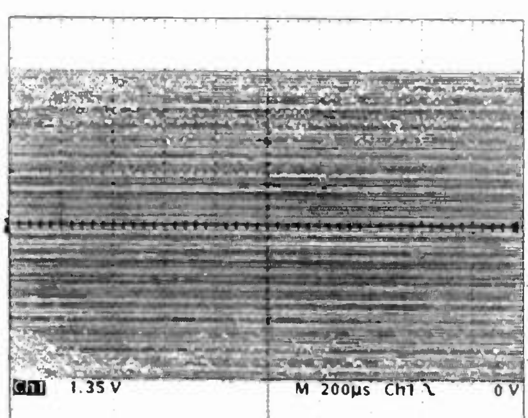
LOOKING AT PEAK CONTROL PERFORMANCE

Before having a look at peak control performance through a digital path, let's first verify the normal peak control operation at the output of a processor under test. A digital processor that employs 48kHz base sampling rate will be used. This processor employs a 192 kHz virtual up-sampled hard limiter. Intra-sample peak problems are virtually non-existent, being limited to about a worst-case 3% error. Following is a test that describes a look at the discrete Left/Right outputs of the system as viewed by a digital storage oscilloscope, to verify peak control.

Left/Right Channel Overshoot Test Methodology

Using program material, the audio processor was set to process *aggressively*. The song "The Real Thing" by Lisa Stansfield was used, because it contains substantial low frequencies and clean high frequencies, thus providing a good challenge for the control of overshoots. The analog output was connected to a Tektronix TDS-744A digital storage oscilloscope. The 'scope was set to the *infinite persistence* mode, which will "hold" the monitored waveform on the screen. Each waveform was stored for at least one minute. The Tek 'scope can store its display as a bitmap file; these files were used for this document.

Over time, the persistence will "fill in" the block with traces of audio waveforms, and the "flat" lines along the top and bottom of the *filled in* section represent clipper performance. Any little "dots" that exceed the reference level of 1.35 volts are overshoots. *Figure-7* shows the performance of the system.



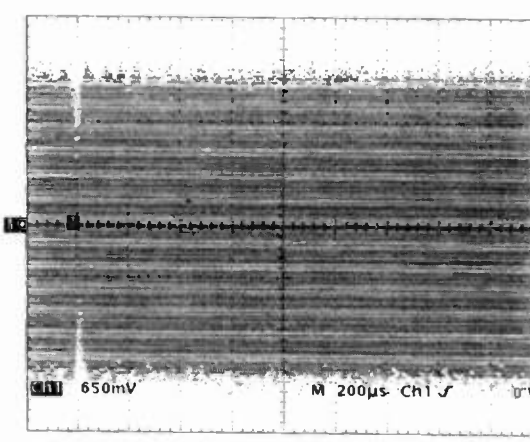
Persistence Display of Processor's Left Channel Analog Output
One Minute Time Period
Figure-7

Notice that there are few little "blips" above the 1.35 volt reference level. These are of insignificant level and of very short duration: approximately 200µs. In real life, they wouldn't be detected by any modulation monitor!

A SOLUTION TO AES/EBU TRANSMISSION OVERSHOOTS: *PREDICTION ANALYSIS CLIPPING*

In each of the above discussions, it is shown how and why overshoots can develop using the AES/EBU connection between processor and exciter. It does not matter if they are co-located or separated by an STL system. As discussed earlier, this is especially true whenever a down-conversion is required between 48kHz and 32kHz sampling, where overshoot components can reach 20%. Using the final limiter in the exciter as a remedy has its own disadvantage—degraded audio quality. What is needed is a final limiter that can *analyze* and *predict* what will happen to the signal downstream, and correct for that—a *Prediction Analysis Clipper.*

Early performance of the 48kHz processor connected to a digital exciter via AES/EBU exhibited the overshoot phenomenon described above, compromising ultimate loudness by up to 2dB. The following oscilloscope image, *Figure-8*, was taken from a test point within the digital exciter after the discrete left/right input has been stereo encoded, and it shows the overshoot components.



Persistence Display Showing Overshoots
Figure-8

In this display, there are "spikes" representing overshoots 15 to 20 percent beyond the reference peak level of ±650 mv. Compare this figure with that of the earlier figure, which showed the tightly-controlled output at the output of the processor. Clearly, there is a loss of peak control as the signal makes its way to the output of the MPX generator in the exciter, this can be attributed to all the problems detailed in the above discussions. What can be done?

Prediction Analysis

In trying to devise a solution to what seems to be an unsolvable problem, let's consider what is known about the problem:

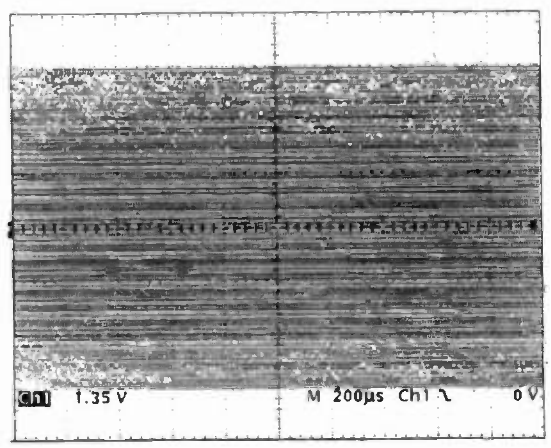
- Overshoots occur whenever down-conversion of 48kHz to 32kHz sampling is performed.
- The tight transition slope of the 16kHz filter in the sample rate converter is a significant contributor to the problem.
- The problem occurs only with signal components above 5kHz.
- It is not desirable to reduce the slope of the low pass filter in the audio processor, as it degrades sound quality.
- Adding more clippers and filters only increases distortion.

Might it be possible to pre-compensate for predicted occurrences of overshoots by the use of supplementary control signals applied to the upper audio spectrum—some type of dynamic, self-adjusting coefficient that could anticipate an overshoot situation, and then correct for it in advance? The answer, amazingly, can be found within the system's main clipper algorithm—the same one employed to eliminate aliasing distortion[4]...aka: *digital grunge!*

Since it's known what mechanisms contribute to overshoots, the severity of the overshoots can be calculated. Then, this information can be combined with the effects of a network that simulates the sharp slope of the 16kHz filter in a sample rate converter. This analysis provides the actual overshoot components that could occur later in the system. By dynamically applying both results to the non-aliasing clipper algorithm, the predicted overshoots can be eliminated!

Note that when analyzing the effects of the 16kHz low pass filter used in the SRC, it is not desirable to actually bandlimit the audio for the tighter requirements of the SRC filter. The broader low pass filter in the processor's design is maintained, which provides two benefits: it does not add further time delay to the system, and it preserves sound quality.

The use of the *Prediction Analysis Clipper* method reduces overshoots in the sample rate converted signal path from a worst case of 20% to considerably less. Testing was done using very aggressive processing settings, under normal processing operation, overshoots were controlled to within 3% or less. As *Figure-9* shows, overshoots in the AES/EBU sample rate converted path are insignificant.



Persistence Display Showing Performance of Prediction Analysis Clipper with Sample Rate Converter
Figure-9

The *Prediction Analysis Clipper* eliminates overshoot problems associated with the use of lower sampling rates in the transmission path. Now, the processor can be utilized with 32kHz digital uncompressed STL systems and 32kHz exciters, and tight peak control will be achieved. Systems can "mix and match" sampling rates with little or no problem incurred regarding overshoot.

It is still recommend that, when using a coded STL link, the processor be located at the transmitter site, as it is proven that codecs will *undo* the tight peak control of any processing system. For further discussion on this topic, please refer to the technical paper "Broadcast

Signal Processing and Audio Coding: Are We Trying to Mix Oil with Water?" This can be found at our web site: *www.nogrunge.com*.

While this new clipping method solves the overshoot problems associated with sample rate conversion, it may not be able to compensate for additional variables that may exist in a broadcast chain. Furthermore, it does not remedy the sonic degradation associated with the added amount of up/down conversions and increased time delay associated with an AES/EBU connection.

CONCLUSIONS

The sampling rate of the audio processor and transmission system have a direct effect on both system peak control performance and subjective sound quality. It has been discussed and shown through research and on-air evaluation that usage of higher sampling rates improves the overall performance in each of these areas. Yet, we live in a world where older technologies, that employed sample rates at 32kHz are in use. What has been shown here is an example where the use of a higher sampled system for processing can co-exist in a lower sampled environment, and without modulation overshoots. Unfortunately, the same process cannot be applied to remaining analog processing systems that must make use of a 32kHz sampled digitized transmission system. There, the lack of backward compatibility is impossible to overcome.

With DAB already on-air in some countries, and hopefully here soon in the USA, it makes all the more sense to realize that we will soon live in a world where 48kHz sampling is at least the minimum.

REFERENCES

[1] Mapes-Riordan, D.: *A Worst-Case Analysis for Analog-Quality (Alias-Free) Digital Dynamics Processing*, 105th Convention of the Audio Engineering Society (AES), San Francisco 1998, Preprint 4766

[2] Steiglitz, K.: *A Digital Signal Processing Primer*, Addison-Wesley Publishing, 1995

[3] AES5-1998: *AES Recommended Practice For Professional Digital Audio – Preferred Sampling Frequencies For Applications Employing Pulse-Code Modulation*, Journal of the Audio Engineering Society, Volume 46, Number 10, October 1998

[4] Foti, F.: *Digital Audio Broadcast Processing: Finally The New Frontier!*, National Association of Broadcasters Convention, Las Vegas, April 1997

FOUR EGYPTIAN MW BROADCAST CROSSED-FIELD-ANTENNAS

F M Kabbary[a], M Khattab[a], B G Stewart[b], M C Hately[c] and A Fayoumi[a]

[a]**Egyptian Radio and TV Union**
Cairo
Egypt

[b]**Dept of Engineering**
Glasgow Caledonian University
Glasgow, Scotland, UK

[c]**Hately Antenna Technology**
Aberdeen
Scotland, UK

ABSTRACT

Crossed-Field-Antennas (CFAs) are novel, small, broadband, high power antennas commonly less than 2 to 3% of λ in height. Currently there are a number of MW broadcast CFAs in service in Egypt. Information relating to four of these broadcast antennas is presented. The paper details: the basic CFA design principles which result in their novel size-wavelength independent nature; near field measurements showing the existence of minimal induction field; vertical plane radiation field patterns; evidence of strong ground-wave and diminished sky-wave radiation; input impedance and bandwidth evaluations of the four CFAs showing their broadband frequency characteristics; and finally, advantages and benefits of CFAs over conventional MW and/or LW antennas.

1.0 INTRODUCTION

Crossed-Field-Antennas (CFAs) originated around 1988 at the Robert Gordon University in Aberdeen, Scotland. [1,2,3] These antennas derived from a research project, the main aim of which was to develop a technique to synthesise directly radiated Poynting vectors from separate *E* and *H* field sources. Over the past few years CFAs have been built and put into service for MW broadcasts by the Egyptian Radio and Television Union (ERTU).[4] To enable an appreciation of the novelty, design and benefits associated with CFAs as compared to standard antennas it is helpful first to review some important features of conventional antenna theory.

Broadcast and antenna engineers will appreciate that effective medium and long wave transmissions are possible with tower antennas that are $\lambda/4$ to $\lambda/2$ in size. For the MW and LW bands this often results in antennas of significant height. For example at 1600kHz, a $\lambda/4$ antenna tower is about 46m (150ft); at 600kHz, a $\lambda/4$ tower is about 125m (406ft). Not only are such towers expensive to manufacture, install and maintain, but they also introduce a significant hazard in relation to electromagnetic safety due to the substantial resonant voltages and currents flowing on the antenna structures.

A further issue relating to conventional antenna theory concerns radiated power. It is well known that radiated power from a dipole or tower antenna has low efficiency. The radiated power for these antennas occurs in the "far field" (generally thought of as the region extending beyond a distance λ from the antenna). In the far field, the *E* and *H* fields are in time-phase, and the ratio *E/H*, often called the wave impedance Z_w, matches space impedance $Z_{space} = 377\Omega$. In this region the Poynting vector $\mathbf{S} = \mathbf{E} \times \mathbf{H}$ produces real power radiation. Two key points also arise in this respect. Firstly, the strong *E* and *H* fields in the "near field" are 90^0 out of time-phase close to the antenna resulting in reactive or non-radiated power in the the near vicinity of the structure. Secondly, the *E* and *H* fields in the far field, which produce the radiated power, are much weaker than the reactive field components located in the near field. These details explain why conventional antennas possess large inductive fields and are not efficient radiators.

What then is the CFA? To put it simply, the CFA is an antenna which achieves the following features:

- it synthesises *E* and *H* fields to be in time-phase in the "near field";
- it designs Z_w to match space impedance Z_{space}.

In other words, a CFA is fundamentally an antenna which is designed to move the radiated power production from the conventional far field region to the near field, thus saving land and minimising the reactive power or inefficiency problems associated with standard antenna designs.

The content of this paper is as follows. Section 2 introduces the essential design concepts of Ground Plane (GP) CFAs outlining the basic techniques underpinning Poynting vector synthesis. Section 3 discusses the improvement to the basic design of GP CFAs for ground-wave broadcasting purposes through the addition of extended cones, and details four MW broadcast CFAs currently in daily service in Egypt. "Near field" measurements on a broadcast CFA are presented showing the non-inductive capabilities of these antennas and thus indicating their high radiation efficiency. In addition, vertical plane radiation patterns for two CFAs are presented showing the relationship between ground-wave and sky-wave radiation. Section 4 presents wide-band input inpedance measurements of all four antennas and discusses the extended zone broadcast capabilities of CFAs. The final section, 5, presents a general summary of the advantages of CFAs over conventional MW and LW antenna towers.

2.0 BASIC CFA DESIGN PRINCIPLES

The fundamental principle underpinning CFA design is that electric and magnetic fields are produced from separate field stimuli, or field electrodes, and crossed-stressed in–phase within a small volume, called the *interaction zone*, close to the CFA structure.

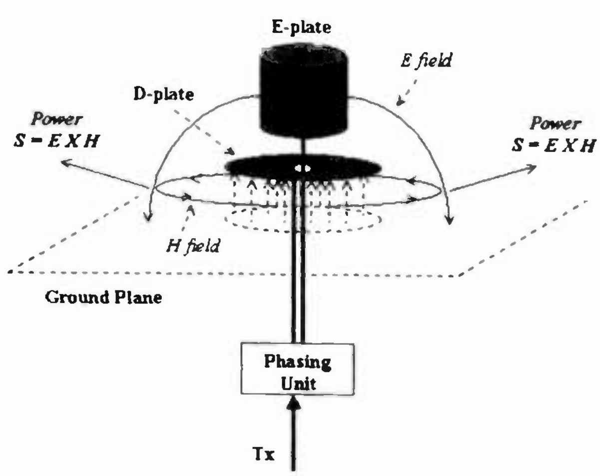


Fig. 1 The basic operation of a GP CFA

Fig. 1 shows the general concept of a GP CFA. Power from a transmitter is fed into a phasing unit from which two voltage feeds are taken to the respective electrodes. One feed is taken to what is called the E-plate, a hollow metal cylinder which produces curved *E* field lines to the GP. The other feed connection is taken to the D-plate, a circular metal disk, which in conjunction with the GP forms a parallel-plate capacitor. The time varying electric field lines between the *D*-plate and the GP produce *H* field lines around the capacitor as shown in Fig. 1. This induced *H* field now links with the *E* field from the cylinder to produce significant power radiation when the following conditions are met:

- both *E* and *H* are in time-synchronism; and
- the field strengths are such that Z_w matches Z_{space}.

The fact that a time-varying electric field creates a magnetic field is a well known phenomenon. The 4th Maxwell equation, viz.

$$\nabla \times \mathbf{H} = \mathbf{J} + \mathbf{D'}$$

indicates that a magnetic field is created from either a charge current *J* (Ampere's Law $\nabla \times \mathbf{H} = \mathbf{J}$) or a displacement current *D*' (Maxwell's Law $\nabla \times \mathbf{H} = \mathbf{D'}$) or from both *J* and *D'* together (note that $D = \varepsilon_0 E$, and ' represents time derivative). To help appreciate the magnetic field production nature of a time-varying *D* field creating an *H* field, Maxwell's 4th equation (omitting the charge current component) may be expressed in the *reversed* Maxwell Law form: [2]

$$\mathbf{D'} \xrightarrow{\quad} \nabla \times \mathbf{H}$$

i.e. *D*' creates an *H* field such that the curl of the *H* field is equal to *D*'. The function of the D-plate is now self-evident. In addition, from the Maxwell Law, when a sinusoidal voltage is applied to the D-plate, the created *H* field close to the plate is 90^0 phase advanced from *D* field. To achieve radiated power the D-plate voltage must therefore be 90^0 phase advanced from the E-plate voltage for time synchronism of the fields and for outward $\mathbf{S} = \mathbf{E} \times \mathbf{H}$ to occur.

With the above information the role of the phasing unit now becomes clear. Firstly, it provides the 90^0 phase difference between the voltage feeds on the **E** and D-plates to provide *E* and *H* in time-phase within the interaction zone, and secondly, it controls the voltage levels on the plates in order that Z_w is able to match Z_{space}. When these condition are met then effective Poynting vector synthesis, i.e. of $\mathbf{S} = \mathbf{E} \times \mathbf{H}$, is achieved and radiated power flows from the interaction zone outward into free space.

It is important to emphasise that as a consequence of this design methodology, i.e. Poynting vector synthesis,

CFAs are *not* resonant antennas like conventional λ/4 or λ/2 antennas.
Two significant features of CFAs therefore arise from these design concepts.

- **Wavelength independent antenna sizes**

Firstly, the synthesis of E and H does not depend critically on CFA size thus ***CFAs can be made extremely small in comparison with the desired radiated wavelength***. As will be seen below it is not uncommon for CFA heights to be less than 2 or 3% of λ. In other words size of the CFA is not wavelength dependant as conventional antenna theory stipulates. The only stipulation on size arises as a consequence of the power requirements of the CFA as necessitated by power engineering criteria.

- **Minimal inductive field**

Secondly, when the time-phase and space impedance conditions are satisfied, there is ***minimal inductive field around the CFA***. The reason is obvious - the field energy in the interaction zone has been designed directly to provide radiated and not reactive power. This can be contrasted with the significant inductive fields from standard conventional antenna structures.

3.0 GROUND-WAVE ENHANCEMENT, FIELD STRENGTH MEASUREMENTS AND VERTICAL PLANE RADIATION FIELD PATTERNS

3.1 Four Broadcast CFAs with Improved Ground-Wave Radiation

The standard GP CFA can be modified with the addition of extended conic sections to the E-plate (see Fig. 2)[4]. These extensions have the effect of confining the curved E field lines in the interaction zone to low angles, such that Poynting vectors produced from the interaction zone are now limited to lower radiation angles. The intended outcome of this arrangement is to produce a significant increase in ground-wave radiation accompanied by a highly desirable decrease in sky-wave radiation.

Four main broadcast CFAs with extended conic sections are now in daily operation in Egypt. Table 1 details basic broadcast information for these antennas, including their location, power level, frequency, CFA height and also the CFA height as a % of the radiated wavelength. Photographs of the two Tanta CFAs and the Barnis CFA showing the extended conic sections are given in Figs. 3 and 4. It may be noted from Fig. 3 that the two Tanta CFAs have been positioned on the rooftop of the same building separated by about 6m (19.5ft) (see later).

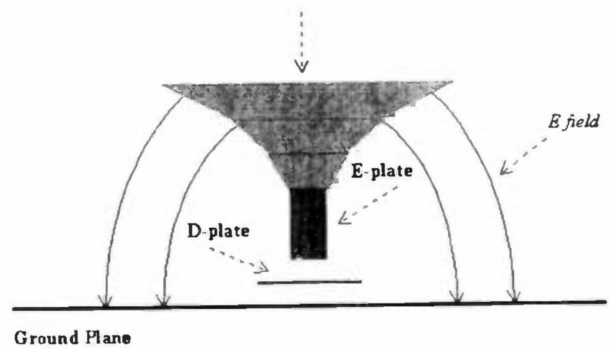


Fig. 2 The addition of conic sections to the E-plate

CFA	*f(kHz)*	*λ*	*Height*	*% of λ*
Tanta 30kW	1161	258.2m (840ft)	8.2m (26.7ft)	3.5%
Tanta 100kW	774	387.6m (1260ft)	9.0m (29.3ft)	2.3%
Barnis 100kW	603	497.5m (1617ft)	9.0m (29.3ft)	1.8%
Halaieb 7.5kW	882	340.1m (1105ft)	6.0m (19.5ft)	1.8%

Table 1 Details of 4 Egyptian Broadcast GP CFAs

Fig. 3 The 100kW and 30kW Tanta CFAs situated on the same rooftop, separated by 6m (19.5ft)

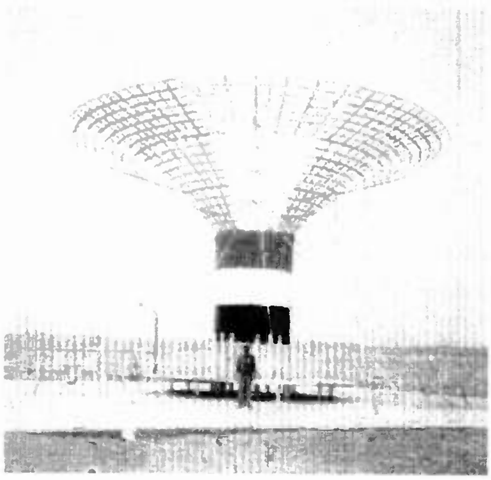


Fig. 4 The 100kW Barnis CFA

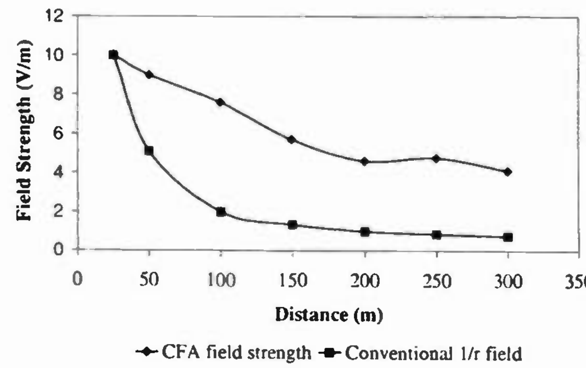


Fig. 5 Near field CFA measurements on the 30kW Tanta CFA

3.2 Near Field Measurements

To investigate near field characteristics of broadcast CFAs, field strength measurements (at reduced power) were taken at near ground level on the 30kW Tanta CFA. These measurements were obtained with a Potomak field strength meter over distances from 25m to 300m. The results are shown in Fig. 5. For comparison, the effective $1/r^2$ field strength values expected from inductive fields is also plotted on the same figure. The CFA shows approximate 1/r proportionality in the near field – there is no sign of the inverse square law proportionality within the first λ/π as associated with the inductive field of a classical dipole antenna. The CFA therefore exhibits very little inductive field in its close proximity.

The significance of this result has resulted in the ERTU recently constructing the 100kW Tanta CFA and positioning it approximately 6m (19.5ft) from the 30kW CFA on the rooftop of the same building as pictured in Fig. 3. There is no evidence of inductive coupling between these antennas, and both operate independently and efficiently without interference.

Measured voltages on the E and D-plates of CFAs also show that voltage levels are about $1/6^{th}$ of those on conventional broadcast antennas carrying the same input power. This feature is again indicative of the non-resonant like behaviour of CFAs. These reduced voltage levels also provide a safer environment near CFA structures.

3.3 Vertical Plane Radiation Field Patterns

Measurements of the vertical plane radiation field patterns of the 30kW Tanta and the 100kW Barnis CFAs have also been taken. Fig. 6 shows the relative vertical plane radiation field pattern of the Tanta CFA. Measurements were taken at a distance of about 610m (1980ft) (using a nearby tall TV tower) utilising an RF meter. Fig. 7 displays the relative vertical plane pattern of the 100kW Barnis CFA, measured at a distance of about 70m (228ft) to a height of around 37m (120ft) using a kite floating a battery powered RF meter. Unfortunately vertical elevation angles of less than about 30^0 were not measured at Barnis as a consequence of the limited height restrictions on the kite. However, the plot shows expected *interpolated* values (dotted line) consistent with what might be expected in relation to the nature of the Tanta CFA pattern.

Fig. 6 shows that a significant proportion of the radiated power goes into ground-wave radiation. For example, the field strength at an elevation angle of about 20^0 is approximately 0.32 that of the ground-wave strength, indicating that the radiated power at this angle is close to 10% (i.e. 0.32^2) of the ground-wave power. At higher elevations, the radiated power is seen to be less than 10%. The Tanta CFAs broadcast to residential populations across a region of 100km – 250km over land based soil, which produces little attenuation of the ground-wave. These service areas are therefore constantly provided with strong signal strength broadcasts.

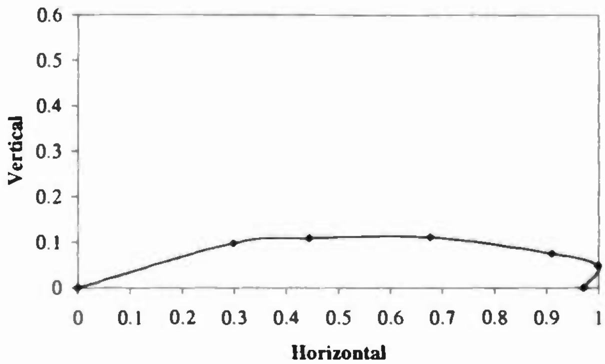


Fig. 6 Tanta 30kW relative vertical plane radiation field pattern measured in the vertical direction at a distance of 610m

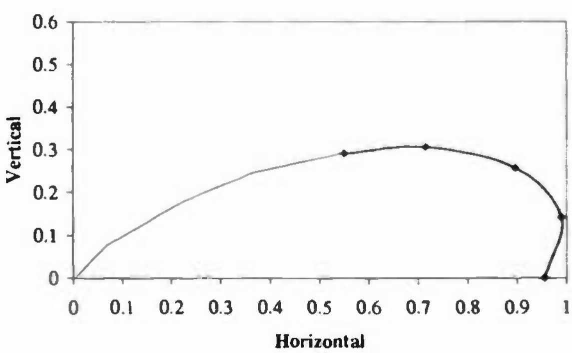


Fig. 7 Barnis 100kW relative vertical plane radiation field pattern measured in the vertical direction at a distance of 70m

For Barnis in Fig. 7 at elevation angles greater than 30^0, the sky-wave radiated power is less than about 40% of the ground-wave power. There is clearly a difference in the radiation patterns between the Tanta and Barnis CFAs, arising for a number of reasons. For example, different heights and angular conic sections on the antennas plus different sizes and separations of the D-plate. These contribute to a variation of the interaction zone field geometries and thus variations in the radiation field patterns. It should also be commented that the Barnis CFA is situated in a region of dry desert, which introduces attenuation on the ground wave thus resulting in what may be expected as a different radiation characteristic pattern than Tanta in the extreme far field.

4.0 BANDWIDTH MEASUREMENTS and EXTENDED SERVICE ZONES

4.1 Frequency Bandwidths

The bandwidth of an antenna is usually presented in terms of input impedance and/or SWR measurements. A fascinating feature of CFAs is that the input impedance to the antenna can always be adjusted to match any desired input impedance at the required broadcast frequency. Using an HP Network Analyser attached to the input of the phasing unit, Smith Charts were obtained for the four broadcast CFAs and these are presented in Figs. 8-11. Table 2 details the bandwidth frequencies and % frequency bandwidths (i.e. bandwidth/broadcast frequency) assuming an SWR of 2:1 side-band down points. It can be seen that all CFAs show remarkable broadband characteristics. In all cases the bandwidths accommodate easily the AM audio spectrum and beyond. If the SWR were to be extended to a conservative 3:1 then it will be obvious that the bandwidths will increase beyond that which can be determined from the Smith Charts presented here. For all CFAs, % bandwidths based on this premise will extend well beyond 10%. In addition, due to the large bandwidth requirements of digital transmissions, these results indicate that it should be possible to transmit higher data rate digital signals at MW using CFAs.

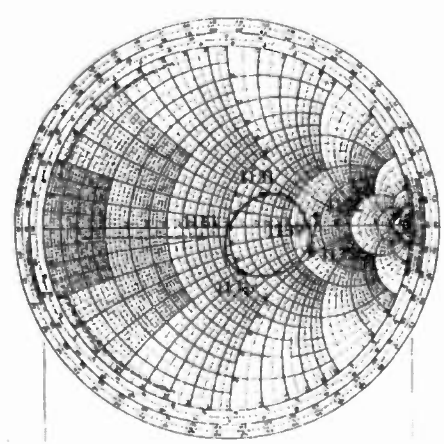


Fig. 8 Smith Chart for the 30kW Tanta CFA

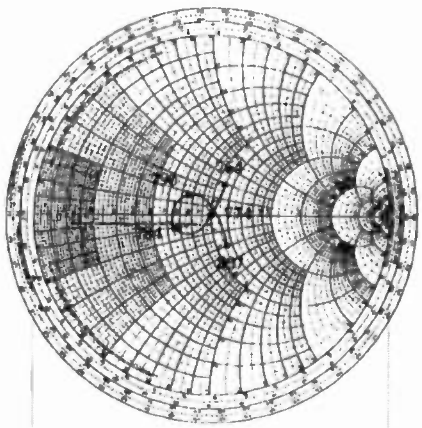


Fig. 9 Smith Chart for the 100kW Tanta CFA

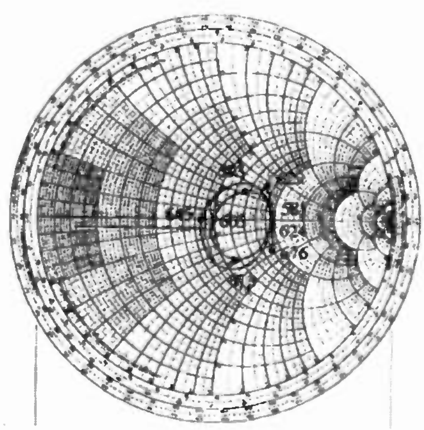


Fig. 10 Smith Chart for the 100kW Barnis CFA

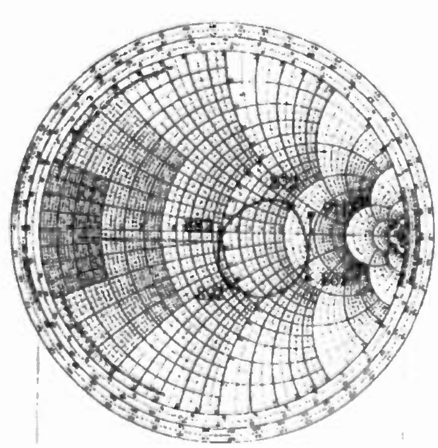


Fig. 11 Smith Chart for the 7.5kW Halaieb CFA

CFA	*2:1 SWR freqs (kHz)*	*Bandwidth (kHz)*	*% Bandwidth*
Tanta 30kW (1161kHz)	1148 1175	27	2.3%
Tanta 100kW (864kHz)	759 814	55	7.1%
Barnis 100kW (603kHz)	579 627	48	8.0%
Halaieb 7.5kW (882kHz)	875 894	19	2.2%

Table 2 SWR 2:1 CFA Bandwidth evaluations

4.2 Extended Service Zones

In terms of useful service zones all CFAs produce strong signal strengths. A comparison between the field strength of the 30kW Tanta CFA and a nearby 30kW $\lambda/4$ antenna tower has previously been reported detailing that the CFA consistently out-performed the $\lambda/4$ antenna by 3-10dB per μV. [4] In this respect it has often been reported that CFAs have been audible a considerable distance from their intended broadcast regions. The BBC (British Broadcasting Corporation) recently performed reception checks on the 30kW Tanta CFA from Nicosia in Cyprus, situated approximately 480km (approximately 280 miles) across both desert *and* mediteranian water from Tanta. Their results are summarised briefly in Table 3.

Time (GMT)	*Reception*	*Signal Strength*
0700	Good	Variable Fair-Strong
1200	Good	Variable Fair-Strong
2000	Poor/Fair	Fair

Table 3 BBC reception reports on 30kW Tanta CFA from Cyprus (4th September 1998)

An interesting feature is that the signal strengths were reported fair-to-strong during both morning and daytime. These simple checks further evidence the significant radiated ground-wave radiation and diminished sky-wave radiation possible with broadcast CFAs, and thus show the extended broadcasting capabilities of these antennas. The fair signal strength report in the evening and the associated poor/fair reception report is due to interference from other broadcast stations operating on 1161kHz e.g. Moscow (at a power of 1MW) and Sofi (power 600kW). It is important to note that these stations have no influence during the daytime.

Recent reports have also indicated it is possible to receive reasonable audible signal levels from the 7.5kW Halaieb CFA during daytime in Khartoum, a distance of approximately 600km (375miles) south from Halaieb over desert and land.

5.0 CONCLUSIONS AND ADVANTAGES OF CFAs

CFAs are small, compact, high power radiation antennas. The construction of these antennas is radically different from conventional antenna techniques due to the fact that by their very nature they are designed to synthesise radiated power in a small interaction zone surrounding the antenna structure. CFAs appear to have minimal induction field, as measured and also evidenced by the fact that two CFAs located 6m (19.5ft) apart on the same rooftop do not interfere. They also possess superior bandwidths in relation to conventional MW antennas, and show vertical plane radiation patterns which exhibit strong ground-wave and reduced sky-wave characteristics.

Taking all the above features into account a number of distinct operational benefits and advantages of CFAs may be summarised as follows:

- Increased broadcast service areas with useful signal strength
- Reduced transmitter power and capital costs thus long term reduced electricity costs – additional benefit includes longer life for transmitter components
- CFAs require no planning structure licence due to small height
- Reduced hazards for aircraft
- CFAs can be mounted unobtrusively on rooftops
- Different CFA antennas can operate in close proximity with no interference due to minimal coupling, i.e. CFAs are EMC friendly
- No tower construction
- Saving on tower maintenance such as lighting, upkeep, guys, insulators etc.
- Reduced insurance costs
- No large real-estate required
- Night-time broadcasts possible due to reduced sky-wave characteristics
- Improved safety due to lower voltage levels of CFAs
- High quality of received audio signal due to broadband characteristics
- Possible use of CFAs for higher data rate broadband digital transmissions at MW

As a consequence of the success of CFAs, the ERTU now plan to replace all conventional MW and LW broadcast antennas with CFA systems over the years ahead.

REFERENCES

1 Hately M C and Kabbary F M
US Patent No. 5155495, Radio Antennas

2 Kabbary F M, Hately M C and Stewart B G
"Maxwell's Equations and the Crossed-Field-Antenna", Electronics and Wireless World, Vol 95, pp216-218, March (1989)

3 Hately M C, Kabbary F M and Stewart B G
"CFA: Working Assumptions", Electronics and Wireless World, Vol 96, pp. 1094-1099, December (1992)

4 Kabbary F M, Khattab M and Hately M C
"Extremely Small High Power MW Broadcasting Antennas", IEE International Broadcasting Conference (IBC), Amsterdam, 10-12th September (1997)

ACKNOWLEDGEMENTS

We would like to thank Miss Heba Said Mohammed for her help on the field and antenna measurements and also the BBC for their assistance in securing the reception checks at Nicosia in Cyprus.

USING SYNCHRONIZED TRANSMITTERS FOR EXTENDED COVERAGE IN FM BROADCAST

Bill Gould and Jai Eu
Harris Communications
Intraplex Transmission Solutions
Littleton, MA

INTRODUCTION

This paper provides an overview of "simulcasting" — the technique of extending radio broadcast coverage areas by using multiple transmitters operating on the same frequency. The paper discusses how simulcasting works and new solutions that address specific technical obstacles that have historically limited use of this technology.

This paper is primarily directed at chief engineers, directors of engineering, professional engineers and consulting engineers who desire a technical introduction to simulcasting and transmitter synchronization. Non-technical readers who would like a brief, high-level overview may wish to read the executive summary that follows, and then review the more detailed sections as their interests dictate.

EXECUTIVE SUMMARY

What is simulcasting?

Simulcasting is the use of multiple, overlapping transmitters, operating on the same frequency, in a market. Through simulcasting, radio broadcasters can ring a metropolitan area ("metro") with low-powered suburban stations and cover the entire market. Simulcast transmitters placed in a line can cover the length of a major highway, population corridor, or rating service survey area. On-frequency boosters can fill in important coverage gaps.

Why simulcast?

Simulcasting can provide dramatically increased coverage. In fact, using simulcasting, smaller stations can be combined to create expanded coverage maps that are larger than a metro powerhouse. Operating costs of smaller transmitting facilities are a fraction of those associated with a single major market station transmitter. Station acquisition costs are far less, and the resale value of the combined system can be many times the price of its individual component stations.

Why transmit on the same frequency?

Simultaneous transmission of programming on multiple frequencies has yielded marginal results. Promoting multiple dial positions is both expensive and confusing to listeners who often don't remember to retune their radios when traveling between coverage areas. It's also important to note that rating services use dial positions to score survey responses. Realistically, the only way to keep people listening is to allow them to move transparently between coverage areas on the same frequency.

Why SynchroCast™?

Broadcasting from two nearby transmitters on the same frequency can lead to serious reception problems in the overlap areas. SynchroCast technology, originally developed for use in two-way radio systems, makes this type of broadcasting possible by integrating state-of-the-art technologies, GPS satellite receivers, and precision digital delay management.

SynchroCast technology can be applied to multiplexed digital studio-transmitter links (STLs), across leased T1/E1circuits, microwave radio links, or fiber optic links.

Carrier frequency and program audio timing at all transmitters are locked to the GPS timing standard, reducing or eliminating unwanted artifacts at the listener's receiver.

APPLICATIONS OF SYNCHROCAST

Applications of SynchroCast include expanding station coverage, filling difficult shadow areas, and conserving frequency usage. Some application possibilities and benefits are outlined in the following sections.

Ring a market with low-powered stations

Major metropolitan stations typically utilize a powerful transmitter, driving an antenna located on a tall in-town building or in an antenna farm. Operating high-powered transmitters and leasing tower space both represent significant expenditures for station operators. Simulcast systems permit joint operation of two or more lower-power suburban stations, with far lower operating costs, to achieve comparable coverage of that same large metropolitan area, with extended reach into the suburbs as well.

The purchase price of a suburban radio station suitable for creating a simulcast system can be a fraction of that of a metro facility. Often times, these stations struggle for profitability when attempting to compete with signals from the "big guns." Simulcasting provides the opportunity to leverage these smaller facilities and enhance their competitive position significantly. Capital outlay for the individual facilities is relatively low and the future resale value of the simulcast system can be many times the price of the individual stations that make it up.

Ownership of all the stations on the same frequency also creates the opportunity for power increases at one or all transmitters or use of directional antennas to further improve coverage. Since any potential for interference would be between co-owned stations, waivers can be simple to obtain.

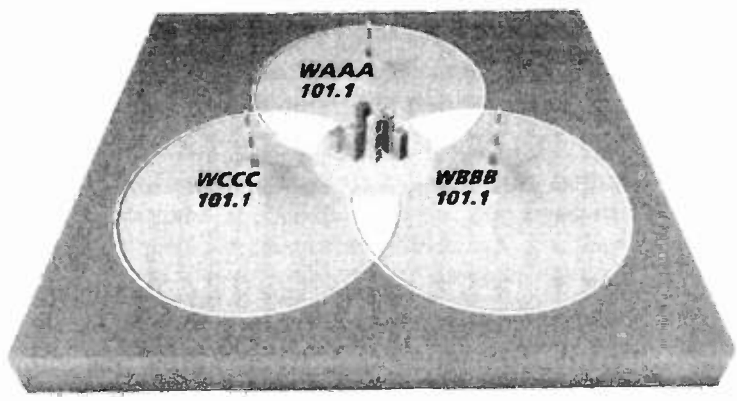


Fig.1: Ringing a Market with Low Powered Stations

Intercity travel corridors

Extended travel corridors cannot be covered by a single transmitter. Simulcasting allows coordinated transmitting stations to be placed along such a corridor so that travelers can enjoy uninterrupted programming without the need to retune their receivers.

Booster stations

Terrain obstacles, such as mountains or steep valleys, can create shadow areas with limited or no coverage within a broadcaster's service area. Traditionally, this problem has been addressed by the addition of either booster stations on the same frequency or translator stations on a different frequency. Translator stations, however, are subject to the availability of usable frequencies, and in many cases, booster stations can create undesirable interference with the main transmitter. In these cases, simulcast systems can be used to eliminate the undesirable effects of booster stations while gaining operators the advantage of transmitting on a single frequency.

European applications

In Europe, the driving force for simulcasting is conservation of frequencies. Most European broadcasters already enjoy regional or national coverage, achieved by programming a network of transmitters on different frequencies. The number of available frequencies is often severely limited, especially in metropolitan areas, restricting the number of stations that can be licensed.

Another problem that arises is the phenomenon of "shadows," or areas where coverage is blocked by hills or mountains. While a repeater or fill-in transmitter can cover the shadowed area, this requires exclusive use of yet another frequency from a highly limited pool. Simulcasting allows broadcasters to cover extended areas, and to fill shadowed areas within a transmitter's primary footprint, without consuming additional frequency. This can dramatically reduce the total number of frequencies required by the overall network and open the possibility for additional services to a given geographic area.

SYNCHROCAST AND THE STL PLUS SYSTEM

Harris' simulcasting technology, SynchroCast, operates in conjunction with the Intraplex STL PLUS digital studio transmitter link, which can be used on leased T1/E1 circuits, microwave radio links, or fiber optic links. The Intraplex STL PLUS multiplexing system provides bidirectional transmission paths for program audio STL/TSL, data for remote control and LAN interconnect, and voice channels for off-premise extensions and intercoms, all on a single digital circuit.

The Intraplex STL PLUS supplies the transmission capability for the STL and SynchroCast. Program audio is delivered to the transmitter by a 15 kHz stereo linear uncompressed digital system. Inputs and outputs can be either analog or AES/EBU. SynchroCast timing signals accompany the program audio in the outbound direction.

An optional TSL package provides a 15 kHz stereo linear uncompressed return audio path, well-suited for off-air audio from the modulation monitor. Since one or more of the transmitters in the simulcast system will be out of over-the-air reception range, this provides a convenient means of monitoring the actual air sound of the transmitters. Two bidirectional voice-grade audio paths are also provided. One accommodates transmitter remote control and return telemetry. The second may be used for backhauling audio from an EAS receiver or RPU gear at the remote transmitter site.

The system requirements for a simulcast application are as follows:

- One Intraplex STL PLUS multiplexer for each studio to transmitter link
- One GPS receiver for the studio and one for each transmitter in the system
- One SynchroCast simulcasting package, which includes all GPS modules, timing transmission modules, and digital delay modules needed for the studio and two transmitter sites

THE EVOLUTION OF SIMULCASTING

A historical perspective

Attempts to increase coverage by broadcasting the same programming on multiple transmitters using different frequencies has been only marginally successful. The same is true for stations using translators to fill in coverage gaps. While continuous coverage can be demonstrated, one important factor has remained out of the control of the broadcaster: the listener.

Communicating a laundry list of dial positions is confusing to listeners. In this situation, the station's ratings are dependent on the listeners' ability and motivation to remember where to tune their radio dial based on their location.

The issue for in-car listeners, a high percentage of today's audience, is even more complex. When the listeners drive out of the coverage footprint of one transmitter and forget to retune their radios, they are lost from a rating and competitive standpoint.

Effects of same-frequency simulcasting

Simply broadcasting the same signal from two nearby transmitters on the same frequency produces a cascade of effects. Listeners located closer to any one of the transmitters where the signal is significantly stronger (usually 6 dB or greater) will hear only the closer transmitter due to the "capture effect" of the FM receiver. Listeners in moderate signal areas will hear one or the other transmitter and may transition between them. Listeners located in "equal-signal" areas or fringe areas can experience serious reception problems where transmitter footprints overlap.

For simulcasting to work effectively, the broadcast signal from each transmitter must arrive at the receiver at a precisely controlled time. A signal leaving the studio will be subject to two delay factors: uncontrolled studio transmitter link path delay and the "flight time" in the air from the transmitter to the receiver, over two different paths. The arrival time of each signal at the receiver can differ significantly, causing distortion, echoes and other artifacts. In addition, each transmitter's local oscillator frequency will be slightly different, causing phase errors between carrier frequencies in the overlap area.

GPS satellite technology

The advent of Global Positioning System (GPS) satellites has, for the first time, created a truly effective method for synchronizing transmitters. Harris' simulcasting technology was originally developed and refined for use in land mobile radio systems carrying mission-critical traffic such as police, fire, and emergency communications. Extension of this type of application for FM radio broadcasting is made possible by the integration of two state-of-the-art technologies: timing from GPS satellite receivers and precision, digitally induced delay.

A GPS receiver delivers a precise timing reference to the studio and to each transmitter site in the simulcast system. At the studio, a timing signal is sent along with the program audio over the studio's transmitter link, to each transmitter site. At the transmitter, SynchroCast technology compares the timing reference received from the studio to the local timing signals, to determine the actual path delay.

Once this delay is established, digital delay modules calculate and introduce a precisely controllable delay, causing exact alignment of the transmitted audio signals. In addition, the GPS receivers at all transmitter sites provide the same 10 MHz reference signal to each transmitter exciter, locking all of their carrier frequencies to the same satellite-delivered timing reference.

Effects of multipule transmitters, each broadcasting a locally generated 19 kHz stereo pilot in a simulcast system, must be considered. Proper decoding of the stereo L-R signal depends on the accurate reception of the stereo pilot. Just as the transmittter carrier frequencies in a simulcast system must be locked to a GPS-delivered reference, the stereo pilot should be locked to a master reference as well.

The system operates automatically once the initial installation and alignment is complete. It continually monitors the timing of each link, keeping the total delay to each transmitter constant, even if the actual path delay changes. This can occur, for example, if a T1 circuit gets rerouted to an alternate path due to network interruptions.

CRITERIA FOR GOOD RECEPTION

When a receiver is in range of more than one transmitter, the criteria for good reception include *relative signal strength* and *total transmission delay*.

Relative signal strength describes the relationship of two or more transmitted signals, based on the location of the receiver. Take the case of two overlapping transmitters, for example. Within the capture area of the transmitters, the signal level of one transmitter is stronger than that of the other.

Total transmission delay is the elapsed time interval calculated from when the signal leaves the studio to when it reaches the receiver. This delay can differ from one transmitter to another, based on the signal path of the specific studio-transmitter link.

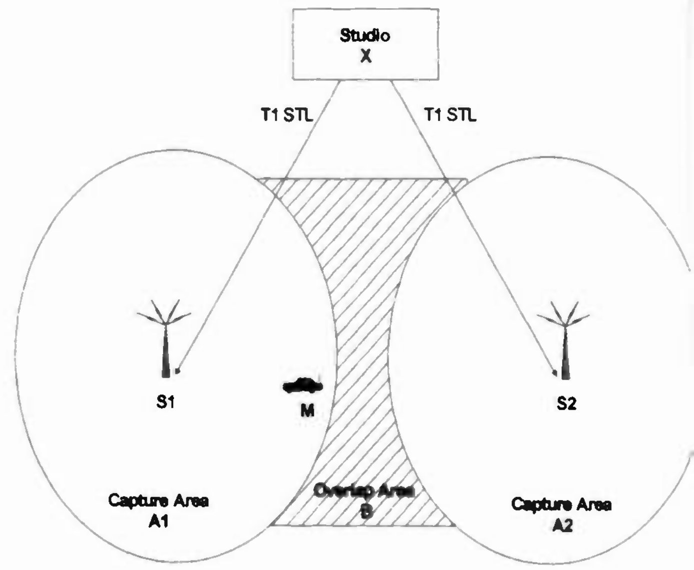


Fig. 2: Impairment Compared to Relative Delay and Protection Ratio

Figure 2 illustrates the relationship between delay and relative signal strength. The figure shows a broadcast system in which the same audio program is simultaneously transported from Studio X over T1 or E1 STLs to two transmitter sites (S1 and S2).

In this example, both sites have equal transmission power. The total transmission delay between the studio and each transmitter is different, based on audio processing time and path delay.

When the FM receiver "M" is located in capture area A1, the receiver will lock in the program transmission from site S1. This is because the signal from S1 is much stronger in capture area A1 than the signal from S2. In this case, the signal from S2 can be considered an interfering signal. When the receiver is located in capture area A2, the reverse occurs.

When the receiver is located in the overlap area "B," however, it receives signals of almost equal strength from both transmitter sites. These signals interfere with each other.

Figure 3 depicts the contours of relative signal strength from both sites. In the overlap area, the relative power levels differ by less than 6 dB.

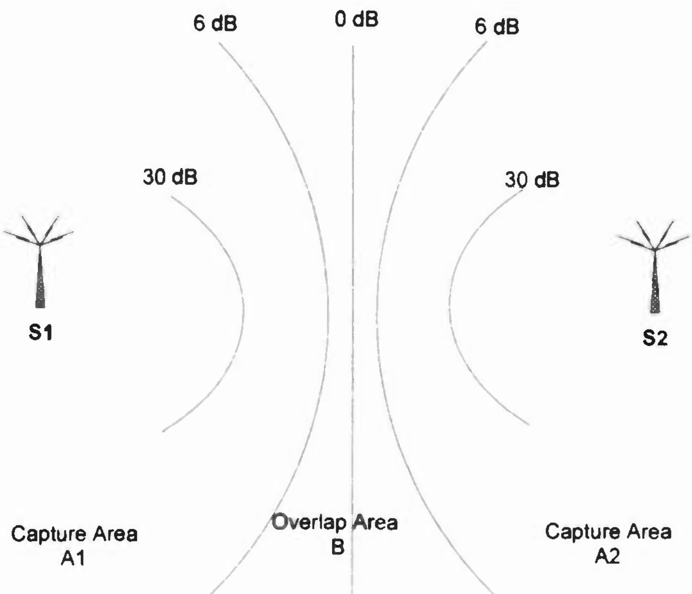


Fig. 3 Overlap and Capture Areas

Because the studio distributes the same FM audio program over STLs using T1 or E1 circuits in the public-switched telephone network, different time delays will occur between the studio and the receiver in the overlap area, based on their location. Factors affecting the total transmission delay time can include audio processing delay at the studio and/or transmission sites, T1 or E1 network path delay, and air path delay.

In the overlap area between two adjacent transmission sites, equalization of the time delay and phase alignment of the audio base band are required for good reception.

Now, consider the relationship between the air path delay from the two transmitter sites. Regardless of the signal power strength of the transmitter, the contours of the air path propagation delay are determined by the distance between the receiver and the transmitter.

As shown in Figure 4, if each transmission site transmits the same signal at exactly the same time, there is a line of equal delay that lies exactly halfway between them, perpendicular to a line connecting them. A receiver located anywhere on this line will receive exactly the same signal at exactly the same time from both transmitters. This is because the speed-of-light delay from each transmitter is exactly the same for all points located on this line.

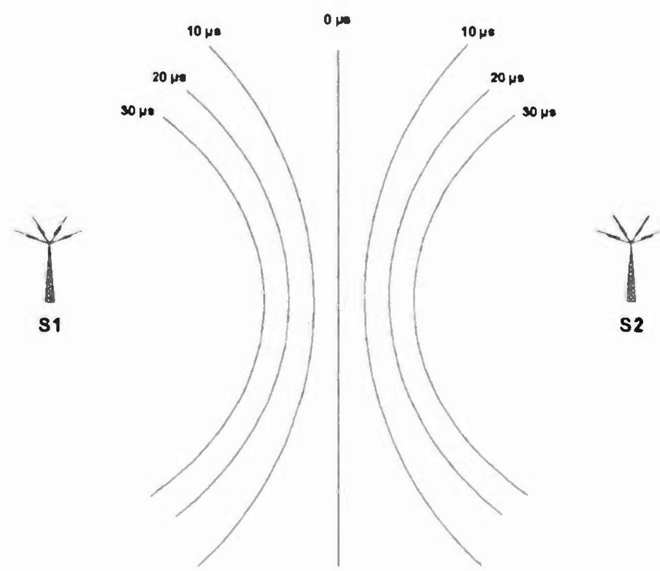


Fig. 4 Propagation Delay Difference in Microseconds

Other lines can also be defined along which a receiver will receive the signal from one transmitter at a constant specified interval before the other. These lines are in the shape of mathematical hyperbolas, with one transmitter or the other at the focal point.

For example, a time-delay difference of 5.364 μs corresponds to a corridor about 1.0 mile (1.6 km) wide, crossing directly between the two transmitters, because the FM receiver gets the signal from the closer transmitter 2.682 μs (0.5 miles) earlier and 2.682 μs later from the farthest transmitter, for a total difference of 5.364 μs.

As indicated in Figure 4, the width of these corridors, at their narrowest, is independent of the actual distance between the transmitters; only the shape of the curves to either side changes.

A similar set of curves can be drawn that represent the relative signal strength that a receiver obtains from each transmitter.

Mathematically, these curves are sections of ellipses, not hyperbolas (because of the inverse-square law). As a result, their basic shapes resemble, but do not correspond perfectly to the delay curves. Furthermore, signal-strength curves are distorted by the antenna patterns, multipath, and terrain conditions. If the transmitter ERP levels are not equally matched, the center of the equal-signal corridor is offset toward the weaker transmitter.

In cases of unequal transmitter power balance, where the point of equal field strength is not located at the equal distance point, signal delay at one of the transmitters must be intentionally and precisely altered. This alters the position of the delay curves relative to the signal level curves, eliminating problem areas or allowing them to be shifted to unpopulated areas such as mountaintops or over bodies of water.

Distribution Networks

At the speed of light, a signal traverses about 981 feet (300 meters) in one microsecond (μs). If the time that a signal leaves a transmitter varies by ± 4 μs, the location of the equal-delay curves will shift by about a third of a mile (0.6 km) to either side.

This delay-curve change is equal to half of the transmitter signal propagation delay change. To maintain this degree of control, the delay of every element in the signal path, from the studio to each antenna, must be controlled to this level of precision. This becomes problematic when sending the signal through a public or private distribution network whose characteristics do not allow this degree of control.

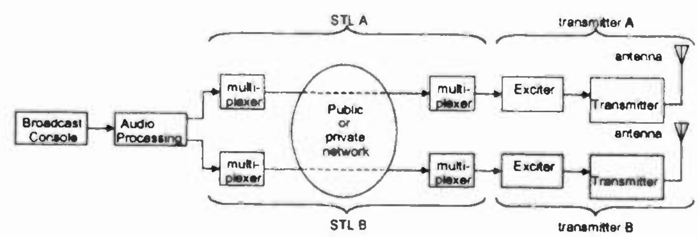


Fig. 5: Signal Path for Two Synchronized Transmitters

Use of digital T1 (1.544 Mbps) or E1 (2.048 Mbps) circuits in the public-switched telephone network have become an increasingly popular way to distribute high-quality audio signals between studios and transmitters. These circuits tend to have fairly stable delay characteristics, typically in the 3-8 ms range. Public networks, however, are subject to rerouting. This means that the phone company or service provider can shift the data to a different physical network path if a hardware fault or excessive congestion occurs.

Rerouting can cause a sudden and dramatic change in the overall circuit delay, and can happen as often as several times each day, without warning.

STLs using private networks or microwave links can also be subject to variable delays on the order of tens of microseconds, as a result of data buffering in modems or other equipment. Long microwave links can have unequal amounts of delay shift due to path differences. Clearly, some sort of mechanism for compensating for these unpredictable delay variations is required for successful transmitter synchronization.

How SynchroCast solves these problems

The Harris Intraplex simulcasting system solves these problems by automatically adjusting for any differences and variability in the STL path delay, and by providing a GPS-locked frequency reference to the transmitting equipment to eliminate carrier frequency drift.

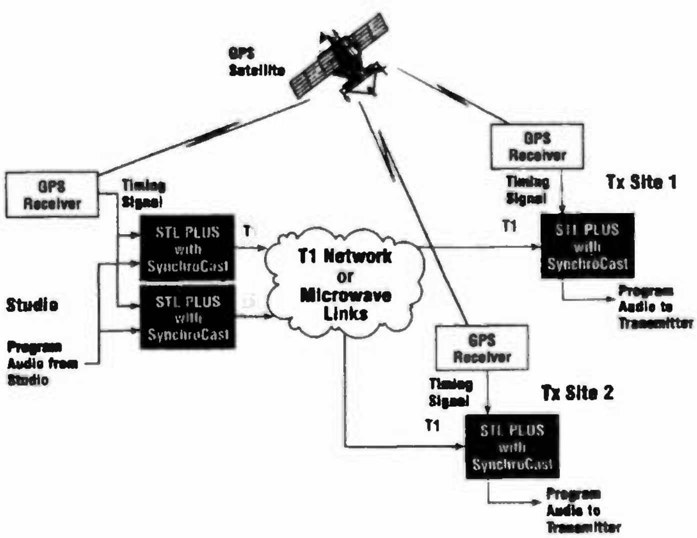


Fig. 6: Block Diagram of System Configuration

The system, shown in Figure 6, addresses these problems with a three-pronged approach.

First, the program is distributed as a discrete-channel digital audio signal, which facilitates amplitude and frequency response matching. Second, GPS receivers at all sites in the network provide extremely precise frequency and time references—to one part in 10^{13} in frequency and to less than 1 μs in time. The GPS frequency reference is used to calibrate the transmitter frequency directly, while, the time reference is used to adjust a variable time delay mechanism that automatically compensates for delay changes in the distribution network.

The timing comparator receives a local timing reference from a GPS receiver located at the transmitter site. Its other input is the master timing reference that arrives over the STL, with exactly the same network delay as the audio signal it accompanies.

The comparator measures the time offset between these two inputs and sends commands to a precision digital delay line to create the overall delay required to compensate for any variations in the network delay.

Overall, the system can control the path delay from point A at the studio to point B at each transmitter to within ± 2 μs. The individual path delays can be offset in steps of 0.1 μs to optimize the performance of the system in the overlap regions. It is also possible to configure the system to absorb path delay variations of up to 84 ms.

It is important to note that the audio and other multiplexed signals are not interrupted or perturbed in any way, even when delay adjustments are made. This is called "hitless" operation, and a patent covering the mechanism that accomplishes this has been applied for.

The system also delivers the signals to the inputs of the exciters with the desired degree of precision. It is important that the signal chain from that point, to the antenna at each transmitter, maintain the same fixed delay, or at least delays that track each other. This is most easily accomplished by using identical processing and amplifying equipment at each transmitter site.

PLANNING A SIMULCASTING SYSTEM

The first step in planning a simulcast system is identifying an appropriate opportunity. For example, a metro market with two or more nearby stations on the same frequency (or one that could be changed to the same frequency) would represent a target for simulcasting. A travel corridor, ratings service survey area, SMSA, or other area with geographic or demographic commonalities would also be possibilities. Station coverage areas can be mapped to determine if their locations would meet simulcasting requirements and business objectives. Marketers may also wish to consider the location of desired population areas, primary highways and commuter routes, and commercial areas where advertisers are located. This process requires both technical knowledge and creativity.

The second step in a feasibility study would involve investigation of existing transmitter facilities and possible future improvements that could enhance these systems. For example, a power increase at one or more of the existing transmitters may be possible if one owner controls all stations on the frequency.

Directional antennas can be used to concentrate signals in desired areas while protecting them from exceeding strength limitations in other directions. A frequency search of the fundamental, as well as the first and second adjacent frequencies, can help identify future expansion possibilities as well as any "short spacing" that might exist.

Distributing the Signal

Once the actual transmitter sites have been determined, the next step is to design the signal distribution network. As noted in the previous section, the key requirements for synchronization of multiple transmitters are ensuring that the carrier frequencies are locked together, and aligning the modulation in both amplitude and phase.

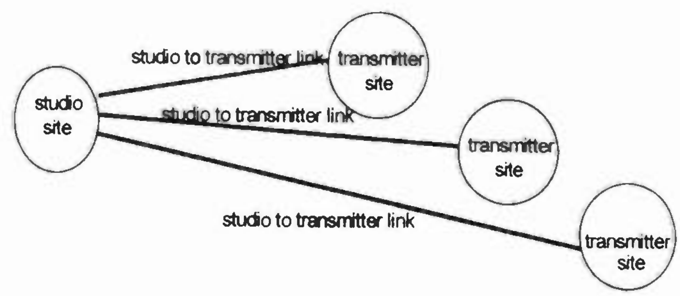


Fig. 7: Block Diagram of Distribution Network

Figure 7 shows an overall distribution network schematic. One link is required from the studio to each transmitter site. In addition to the 15 kHz stereo linear uncompressed channels for transmission of program audio to the transmitter, SynchroCast timing information is sent in the outbound direction. The optional TSL support provides a 15 kHz inbound stereo channel for an air monitor and a bi-directional path for remote transmitter control. An additional voice grade circuit is available for carrying Emergency Alerting System (EAS) audio, remote pick-up, etc. (See Figure 8)

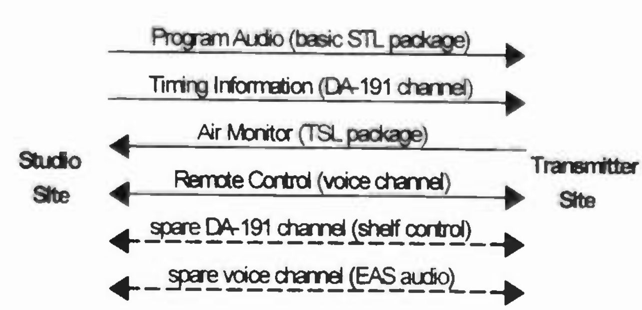


Fig. 8: Traffic Details of Each Studio-Transmitter Link

Each link requires the equipment previously listed. The studio site requires one multiplexer shelf for each link, but a single GPS receiver can be shared for timing reference. Each transmitter site requires one multiplexer shelf and a GPS receiver to generate its local timing and frequency references. Each additional transmitter site added to the system requires a multiplexer shelf at each end of the link and a GPS receiver at the transmitter site.

Audio Processing

The simulcast concept requires exact timing of the signals from all transmitters. Management of transmitted audio is an important factor for seamless reception in a simulcast system. While individual station operators have personal preferences when it comes to audio processing, it is important that the audio transmission be as consistent as possible, including density of processing and modulation levels, across all transmitter sites.

One way to achieve this is to locate the processor at the studio and split its output to the STLs. (Some amount of final peak limiting may be employed at the transmitter, but care should be taken to make this identical at each transmitter site.) Modulation levels at all transmitters should be maintained as close as possible to the others.

IMPLEMENTING A SIMULCAST SYSTEM

Once the basic functionality has been installed, the primary parameter that needs to be set is the total nominal system delay between the studio and the transmitters. Initially, each transmitter site should be set to the same delay. This value should be 2 to 5 ms larger than the longest link delay expected to any one of the transmitter sites, taking rerouting into account.

A typical installation will begin with a nominal value of 10 ms. The SynchroCast system will be able to compensate for any actual link delay to any of the transmitter sites up to this value. If the actual delay to a transmitter site exceeds this value, that site will be out of synchronization with the rest of the system until the delay is reduced again or the system delay value is adjusted.

On the other hand, picking a system delay value that is much larger than necessary can create excessive delay in the off-the-air monitors.

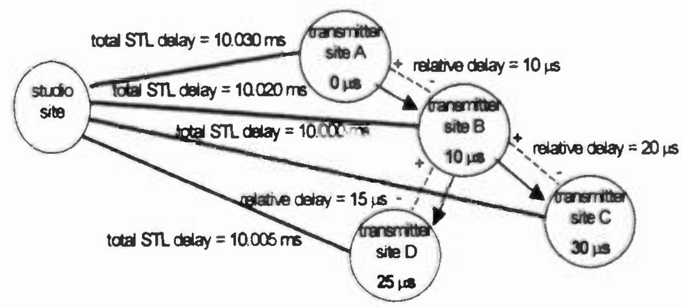


Fig. 9: Adjusting individual site delays

If the planning phase indicates the need for delay offsets between pairs of transmitters, the delays at some of the sites will need to be increased beyond the nominal system value (never reduced).

It can be useful to create a "dependency" graph by drawing lines between pairs of transmitters that overlap, then putting an arrowhead on each line, pointing to the transmitter that requires the smaller delay. (See the example in Figure 8.)

The final graph should indicate one or more transmitters that have only "outbound" arrows and no "incoming" arrows. If this is not the case, some individual compromises in the design of the system will need to be made, based on the specific circumstances.

The methodology for creating the graph is as follows:

1. Write the value "0" in these transmitters, as shown on transmitter A.
2. Follow the arrows outward; for each new transmitter you reach, add the value of the delay on the arrow to the value written on the previous transmitter, and write the total on the new transmitter. In our example, we write "10 μs" on transmitter B, and then "30 μs" on transmitter C and "25 μs" on transmitter D. If two or more arrows enter one transmitter, take the largest value for that transmitter.
3. When all of the transmitters have been marked, at least one of them will have the largest delay number written on it, and there will be no arrows exiting it.
4. Leave this transmitter set at the nominal system delay. In the example, it's transmitter C, which is set to 10.000 ms. Follow the arrows backward and forward to visit all of the other transmitter sites.
5. Each time you follow an arrow backward, increase the delay at that site relative to the delay of the site you just left, and each time you follow an arrow forward, decrease the delay. In the example, we follow the arrow backward from transmitter C and set transmitter B to 10.020 ms, then follow the arrow forward to transmitter D and set its delay to 10.020 – 0.015 = 10.005 ms. Finally, we go back to transmitter B and follow the arrow backward to transmitter A and set its delay to 10.030 ms.

Establishing Delay Alignment

Once the overlap areas have been identified, it is necessary to decide where performance should be optimized in each area. The relative timing between pairs of transmitters can then adjusted to place the line of equal delay in the desired location.

If the delays of two transmitters are identical, the line will be exactly halfway between them. Increasing the delay of one transmitter by 1 μs will shift the line toward that transmitter by about 500 feet (150 meters). Also, the line will no longer be straight; it will start to curve around that transmitter in a hyperbolic shape.

This set of adjustments is relatively straightforward for simple combinations of transmitters. However, if specific transmitters overlap with several others, it may be impossible to fully optimize the relative delays because of conflicting requirements, and a compromise will have to be made.

Power levels and antenna patterns are dictated by station license and are difficult to change. Increasing or decreasing the power level of one transmitter relative to another can be used to shift a problematic overlap region away from, or toward, that transmitter. Adjustments of one or both antenna patterns can achieve the same result.

The final performance of the system can be verified by once again driving through the overlap areas, this time with the same tone being broadcast from all transmitters. Listening for the "purity" of the tone, especially in the "problem" areas located previously, will give a good worst-case indication of the system performance on normal program material.

Verifying Overlap Regions

While the location of the overlap regions can be roughly predicted from the ERP and antenna patterns of the individual transmitters, it is a good idea to verify the results in the field.

One simple way to do this involves transmitting two different audio tones on the two different transmitters being tested. Note that a subcarrier (SCA channel) can be used for this testing to avoid interrupting the main channel's programming.

If you drive around the expected overlap region with a receiver, it will be easy to tell from moment to moment which transmitter is being received, and how much the receiver is jumping between them. Coloring a map of the region with two different colors based on which tone is heard will produce a convenient visual summary of the test results.

Conclusion

Simulcasting programming on multiple overlapping transmitters on the same frequency in a market can provide radio broadcasters with a significant competitive advantage in increased coverage, higher ratings and lower operating costs. The effectiveness of same-frequency overlapping transmitters depends on accurate synchronization of the carrier frequencies and broadcast audio. Only when this timing is precisely controlled can the listener enjoy seamless reception with a minimum of artifacts.

The Harris Intraplex SynchroCast system unites GPS exact timing references and precision digital delay with the Intraplex STL PLUS digital studio-transmitter link to produce a fully integrated simulcast synchronized transmitter STL/TSL solution.

59 Porter Road, Littleton, MA 01460-1410

Phone: (978)486-9000 Fax: (978)486-0660
URL: http://www.intraplex.com

A Trial of AM/Digital Multiplexing Transmission

Atsushi Shinoda, Kenichi Shiraisi, Shuichi Okamoto, Shoichi Suzuki, Ryuichi Okazaki, Toshiyuki Takegahara, Masaoki Takai
KENWOOD CORPORATION
JAPAN

1. Outline

In Japan, traffic reports, weather forecasts, and other text broadcasting is being performed with the so-called DARC (DAta Radio Channel) system, which is an FM multiplex digital data broadcasting system. The FM band is wide band enough to broadcast multiplex digital data in addition to audio signals. In the case of AM broadcasting, however, the entire AM channel band is occupied by the audio signal, so the signal cannot be easily multiplexed. Though sound quality gets worse, audio signal band is made narrow and the band which digital data transmit is produced.

We have devised a method of multiplexing digital data and audio signal together in the AM channel band, in keeping compatibility with existing AM receivers without losing audio properties. And we conducted computer simulations, trial manufacture of equipment, and indoor tests, all of which are described in this report.

2. Introduction

We developed a system of transmitting multiplexed digital data within the AM channel band. This system can multiplex a digitally modulated wave on an amplitude-modulated wave, and transmit and receive digital data and analog audio signals simultaneously. And the system has already been tested indoors.

In the test facility, the phase of a QPSK signal (about 4kbps of digital data) digitally modulated wave is inverted to the upper and lower side-bands in the AM band and superimposed on the monaural amplitude-modulated wave in symmetry with the AM carrier.

The AM audio signal is demodulated using a synchronous detection system. At that time the digitally modulated wave is canceled and only the audio signal is fetched.

When the digital data are demodulated, the AM audio signal is canceled and only the digital data are fetched by inverting and subtracting all the AM and digitally modulated waves.

3. Modulation System

Fig. 1 is a block diagram of the modulation part of the test system. In the test system, QPSK was used for the digital modulation system.
Expressing the output from modulator as $v_{AM}(t)$,

$$v_{AM}(t) = \{1 + kv_m(t)\}\cos\omega_c t \qquad \textbf{Eq.1}$$

Expressing the output from orthogonal modulator 1 as $v_{DH}(t)$,

$$v_{DH}(t) = I_n \cos(\omega_c + \omega_1)t + Q_n \sin(\omega_c + \omega_1)t \qquad \textbf{Eq.2}$$

Expressing the output from orthogonal modulator 2 as $v_{DL}(t)$,

$$v_{DL}(t) = -I_n \cos(\omega_c - \omega_1)t + Q_n \sin(\omega_c - \omega_1)t \qquad \textbf{Eq.3}$$

And then, the output from adder 1 as $v(t)$ is following.

$$\begin{aligned} v(t) &= v_{AM}(t) + v_D(t) \\ &= \{1 + kv_m(t)\}\cos\omega_c t \\ &\quad + I_n \cos(\omega_c + \omega_1)t + Q_n \sin(\omega_c + \omega_1)t \\ &\quad - I_n \cos(\omega_c - \omega_1)t + Q_n \sin(\omega_c - \omega_1)t \end{aligned} \qquad \textbf{Eq.4}$$

This modulation process is shown Fig.2.

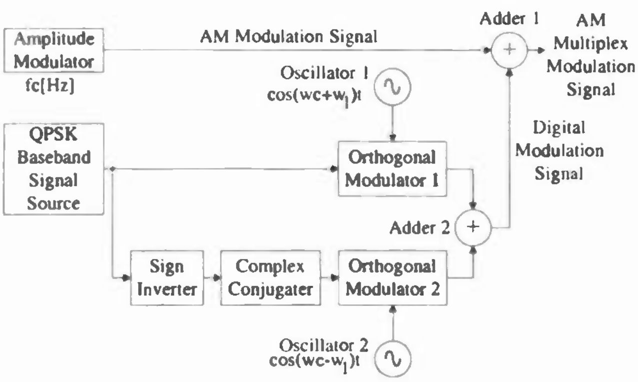


Fig. 1 The block diagram of modulator

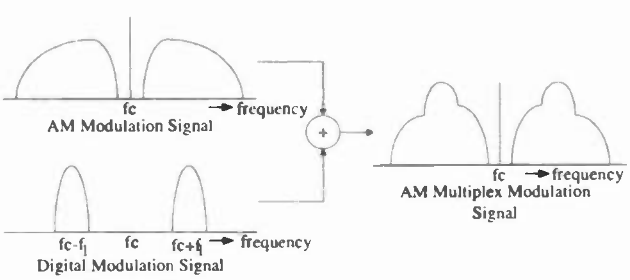


Fig. 2 Image of modulated wave

4. Demodulation System

Fig. 3 is a block diagram of the demodulation part of the test system. To demodulate the AM audio signal, the modulated wave should be synchronously detected. When the AM multiplex modulation signal $v(t)$ is multiplied by the carrier wave $\cos\omega_c t$,

$$\begin{aligned} &v(t)\times\cos\omega_c t \\ &= \{1 + kv_m(t)\} + I_n \cos\omega_1 t + Q_n \sin\omega_1 t \\ &\quad - I_n \cos\omega_1 t - Q_n \sin\omega_1 t \\ &= 1 + kv_m(t) \end{aligned} \qquad \textbf{Eq.5}$$

is obtained, indicating that the audio signal can be easily fetched. Here, for the purpose of simplification, the after-LPF is taken into consideration and the higher frequency items are omitted beforehand.

To demodulate the digital data, the AM audio signal must first be canceled. An example of a circuit that cancels the AM audio signal is shown in Fig. 4.

First, the AM multiplex modulation signal $v(t)$ frequency is converted with frequency oscillators 1 and 2 by carrier oscillator 1 ($\cos\frac{3}{2}\omega_c t$) and carrier oscillator 2 ($\cos\frac{1}{2}\omega_c t$). Expressing their respective outputs as $v_{UPPER}(t)$, $v_{LOWER}(t)$.

$$v_{UPPER}(t) = \frac{1}{2}\{1 + kv_m(t)\}\cos\frac{\omega_c}{2}t + \frac{1}{2}\left\{\begin{aligned} &-I_n \cos\left(\frac{\omega_c}{2} + \omega_1\right)t - Q_n \sin\left(\frac{\omega_c}{2} + \omega_1\right)t \\ &+ I_n \cos\left(\frac{\omega_c}{2} - \omega_1\right)t - Q_n \sin\left(\frac{\omega_c}{2} - \omega_1\right)t \end{aligned}\right\} \qquad \textbf{Eq.6}$$

$$v_{LOWER}(t) = \frac{1}{2}\{1 + kv_m(t)\}\cos\frac{\omega_c}{2}t + \frac{1}{2}\left\{\begin{aligned} &I_n \cos\left(\frac{\omega_c}{2} + \omega_1\right)t + Q_n \sin\left(\frac{\omega_c}{2} + \omega_1\right)t \\ &- I_n \cos\left(\frac{\omega_c}{2} - \omega_1\right)t + Q_n \sin\left(\frac{\omega_c}{2} - \omega_1\right)t \end{aligned}\right\} \qquad \textbf{Eq.7}$$

is obtained. For the purpose of simplification here, also, the after-LPF is taken into consideration and the higher frequency items are omitted beforehand.

This is then subtracted in a subtractor. Expressing the results as $v_D(t)$,

$$
\begin{aligned}
v_D(t) &= v_{LOWER}(t) - v_{UPPER}(t) \\
&= I_n \cos\left(\frac{\omega_c}{2} + \omega_1\right)t + Qn \sin\left(\frac{\omega_c}{2} + \omega_1\right)t \\
&\quad - I_n \cos\left(\frac{\omega_c}{2} - \omega_1\right)t + Qn \sin\left(\frac{\omega_c}{2} - \omega_1\right)t
\end{aligned}
\qquad \textbf{Eq.8}
$$

is obtained. This equation clearly shows that the AM audio signal has been canceled. Then, the digital modulation signal can be demodulated in the normal manner. The image of the AM signal canceling process is shown in Fig. 5.

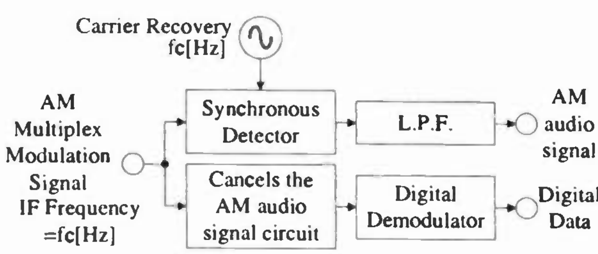


Fig. 3 A block diagram of the Demodulator

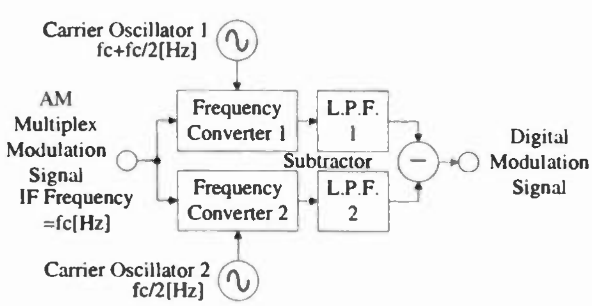


Fig. 4 An example of circuit that cancels the AM signal

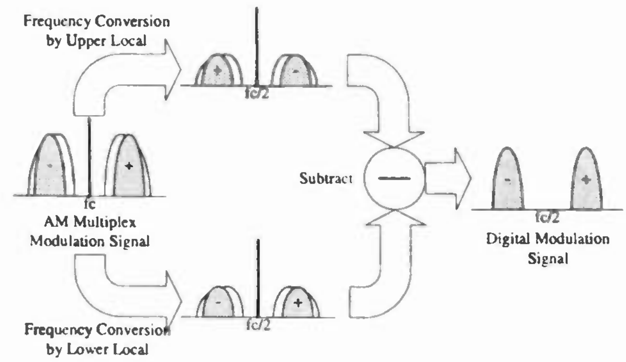


Fig. 5 The image of the AM signal canceling process

5. Prototype System

Figs. 6 and 7 show our test system.

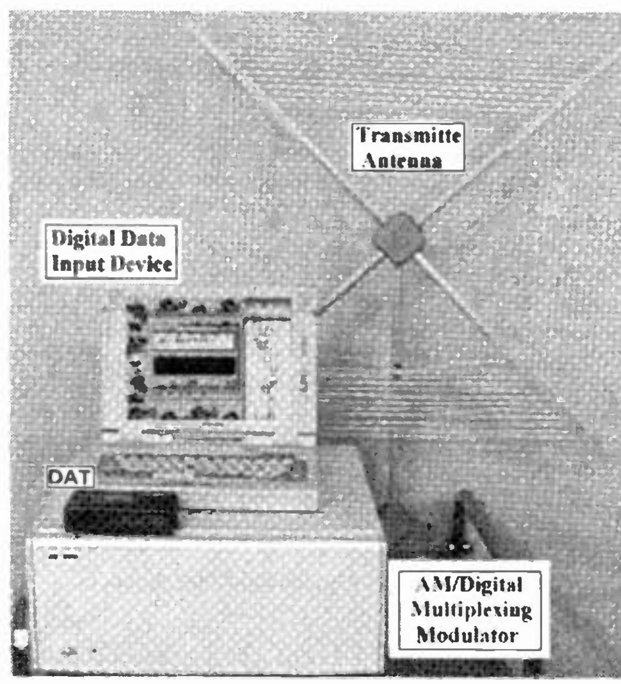


Fig. 6 The prototype of modulator

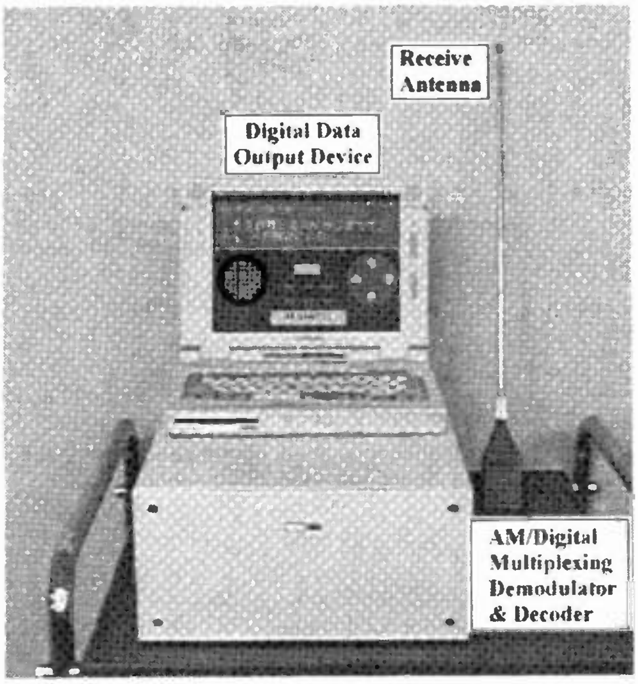


Fig. 7 The prototype of demodulator

The modulator is made up of an AM/Digital Multiplexing Modulator, DAT, a Digital Data Input Device, and a Transmit Antenna. The audio signal reproduced with the DAT and the digital data generated by the Digital Data Input Device are multiplexed by the AM/Digital Multiplexing Modulator, and the radio frequencies are modulated. The carrier frequency uses 855kHz, which is not used in Tokyo and its environs. The modulated wave is transmitted from the Transmit Antenna.

The demodulator consists of an AM/Digital Multiplexing Demodulator and Decoder, a Digital Data Output Device, and a Receive Antenna. The AM multiplex modulation signal received by the Receive Antenna is demodulated by the AM/Digital Multiplexing Demodulator and Decoder, and the digital data are displayed by the Digital Data Output Device. A commercially available synchronous-detection type AM radio and diode-detection type AM radio are used to receive the audio signal.

In the Fig. 5 image diagram, the audio signal and digital data levels are drawn practically the same. If that were the case, however, the digital data would be heard as noise on the diode-detection type AM radio; hence, the digitally modulated wave level is lowered 20dB below the AM carrier wave frequency. For test purposes, moreover, QPSK (transmission rate = 4kbps) is used for the current digital modulation system.

The test results indicate that, in practical terms, the digital data do not have most interfere on the audio signal. Though it is natural, no noise hear on the synchronous-detection type radio, and it does not become a problem so much on a diode-detection type radio.

Digital data BER is nearly zero. But errors were produced at times due to the RF analog part performance, when loud sound is included in the audio signal. When the system is put to practical use, however, this problem can be solved by adding an error correcting code.

6. Outlook

The current test facility uses the system described in Par. 4 as an AM signal canceling method. There are other AM signal canceling methods that can be considered.

For example:

- A digitally modulated wave is synchronously detected and canceled, and then it is modulated again and subtracted from the original modulation wave.
- The data modulation signal contained in the AM multiplex modulation signal is synchronously detected directly, and a code is configured to cancel the audio signal.
- The portion, on a time axis, of the AM multiplex modulation signal that is not affected by the audio signal is sampled, and only the digitally modulated signal is fetched.

We executed computer simulations and obtained a large variety of experimental data.

The current tests were conducted to corroborate our theory; consequently, QPSK is used for digital modulation because the circuit scale is made small. Nevertheless, once an OFDM signal with 200 carriers in the AM channel band is used and digital multi-level modulation is adopted, musical broadcasts with high sound quality similar to the digital audio broadcast IBOC method being studied in United States can be a reality.

DTV Propagation Issues and Field Studies

Tuesday, April 20, 1999

9:00 am - 12:00 pm

Chairperson: Robert Hess
WBZ -TV, Boston, MA

9:00 am Understanding SNR in the Digital World
Brett Jenkins
COMARK Division of Thomcast Communications Inc.
Southwick, MA

9:30 am On the Validity of the Longley-Rice (50,90/10) Propagation Model for HDTV Coverage and Interference Analysis
Oded Bendov
Dielectric Communications
Cherry Hill, NJ

10:00 am Results of TV Receive Antenna Field Testing For Correlation of Digital and Analog Television Reception Using the CEMA Antenna Selector Map Program
Ray Conover
U.S. Satellite Broadcasting Company
St. Paul, MN

***10:30 am Digital Television Propagation - An Update of Current Field Measurements Projects**
Dennis Wallace
Wallace & Associates
Washington, DC

11:00 am Comparing Panel and Slot Antennas in Shared Digital and Analog Service
Anthony Magris
Radio Frequency Systems
North Haven, CT

11:30 am DTV Taboo Channel Interference into NTSC at High Power Levels
Stanley Salamon
Advanced Television Technology Center
Alexandria, VA

*Papers not available at the time of publication

Understanding SNR in a Digital World

Brett Jenkins
COMARK Division
Thomcast Communications, Inc.
Southwick, MA

Abstract

With the transition to DTV, some of the old challenges in maintaining signal purity through a high power transmitter become more complex as the signal can be much less tolerant to distortions. New technologies are essential to minimize and compensate for distortions in DTV transmitters. This paper will explain why correction of these distortions impact power requirements, coverage and cost of operation. The paper will highlight methods being used in DTV transmitters to insure that the highest possible quality signal is transmitted at all times. Real world examples from high power DTV transmitter installations will be presented and explained along with detailed discussion of the multiple transmitter system elements that contribute to SNR reduction and consequently, loss of signal quality.

Causes of Distortion in a DTV Signal

Many things can contribute to the degradation of quality of a DTV signal. We can lump distortions into three broad categories. The first is linear distortion. Linear distortions are generally manifested as amplitude or group delay variations over the frequency band. A system is considered to be linear when you can apply the principle of superposition.[1] That is, given an input with a particular system response and a second input with a second system response, the system response to the sum of the inputs will be the sum of the system responses. This also must be true when the inputs are multiplied by a constant. That is, multiplying the original input to the system by a number yields a response which is the original system response multiplied by the same number. Generally, a system ends up being linear if the system does not change when the magnitude of the input changes, but simply passes the scaling factor along to the system output. In a transmitter plant, the most common components which introduce linear distortion are system output filters and tube cavities.

The second type of distortion is non-linear. Non-linear distortion occurs when a signal undergoes any non-linear operation. To visulize a non-linear operation in a high power transmitter we usually look at the Power_{in} versus $\text{Power}_{\text{out}}$ curve. If a transmitter were perfectly linear, then the gain would have to remain constant for any input power. But we know that all "real" amplifiers cannot be linear over an infinite range. If this curve were a perfect straight line, then the amplifier would be linear. Mathematically, the amplifier would then be a simple multiplication, which is a linear operation and adheres to the rules of superposition. The major contributor of non-linearity in a transmitter will be the final output stage amplifier.

The third type of distortion is noise. Noise can be added into the signal because of a variety of effects, e.g. local oscillator phase noise, thermal noise, etc. In general, a well designed modulator and transmitter system will not suffer any penalty from these effects, so this paper will focus on the linear and non-linear forms of distortion.

The Effects of Distortion on the DTV Signal

Non-linear Distortion

Non-linear distortions cause intermodulation products. These products are distortion components that appear at sum or difference frequencies of whole number multiples of the frequencies in the original signal. Depending on the frequencies involved in making the intermodulation product, the distortion can show up at various places either in the frequency band of interest, or outside of it. In this paper, we are concerned with the effect of the products on the 8VSB signal, so we will interested in looking at only the in-band intermodulation products.

Linear Distortion

In order for any digital communication system to function well, we need to find a way to send bits, or pulses, in such a way that the first pulse doesn't interfere with the second and so on. It was Nyquist who defined a class of filters to allow transmission of pulses without this type of interference, called inter-symbol interference, or ISI.[2] One type of filter which is practically realizable and satisfies Nyquist's symmetry requirements for ISI free transmission is a raised-cosine filter. This is the filter used in the 8VSB system. With this type of filter, the system response to each "pulse" of digital data has a zero-crossing at each symbol instant. This means that no previous or future pulses of data will interfere with the current pulse. It is easy to see that changing the system transfer function by changing the group delay or amplitude response of the channel will distort the impulse response and cause the responses to be non-zero at symbol instances other than the pulses occurrence in time. It should be obvious why this is called "inter-symbol interference."

This problem becomes even more complicated when you consider that the impulse response of an 8VSB symbol lasts for about 70 symbol times. That means that any particular signal could possibly have interference with the 35 consecutive symbols that occurred previously and the 35 consecutive symbols that will occur later!

Receiver Adaptive Equalization

The good news about linear distortion is that all receivers in DTV need to have some kind of correction for linear distortion. This correction is necessary to compensate for the kinds of distortions which are typical of almost any terrestrial broadcast channel. Without any type of linear equalizer in a receiver, it is unlikely that the signal, no matter how clean it was out of the transmitter, could be received reliably over the entire desired coverage area.

But the bad news is that these equalizers will not have infinite correction capability. So any linear distortion created by the transmitter will tend to "eat up" the receiver's headroom in its ability to compensate for channel distortion. The ultimate effect of this is to potentially reduce the coverage area. Not only that, but as the receiver equalizer compensates for linear distortions, it necessarily adds noise to the signal. This happens because the equalizer works by adding and subtracting delayed and scaled copies of the main signal back into itself. As these copies are added and subtracted, the noise power of the copies are always added back into the signal. The result is that more noise power ends up in the final signal degrading the C/N ratio.

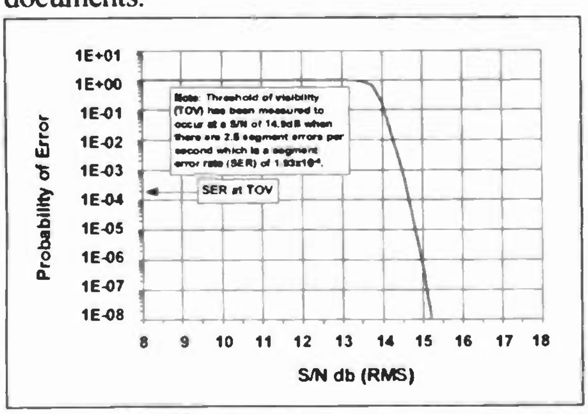
Loss of Signal Quality

The effect of all types of distortion and noise can be measured using a number called Error Vector Magnitude (EVM). The error vector magnitude is a ratio of the average power in several symbols to the average noise power in the same sample set. This number is frequently expressed as a ratio in dB and called Signal-to-Noise ratio (SNR).

Any type of distortion in-band, whether linear, non-linear or random will increase the probability that a received signal will be misinterpreted and assigned an incorrect value in the data stream.

The 27dB Specification

To understand the real coverage effects of signal distortions, it is necessary to define a receiver's threshold given an ideal signal (that is no transmitter distortion) in the face of random noise. Of course, this has been done for the ATSC system and the results have been published in applicable ATSC standards documents.[3]

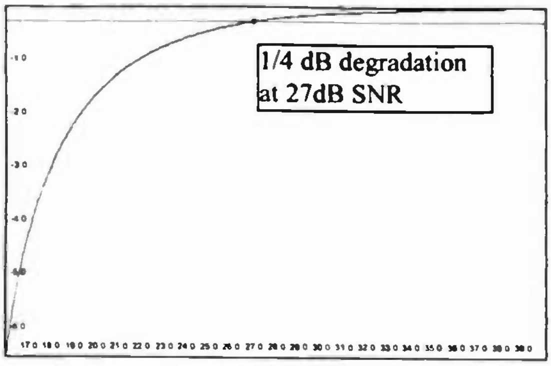


• Figure 2 - The TOV for the ATSC system (taken from ATSC document A/54)

As shown in Figure 2, the TOV for 8VSB is about 14.9dB. One must find out what the new threshold is for a given SNR at the same segment error rate in order to understand how in-band distortion affects the receiver threshold. Or alternatively, given the distortions are present to start with, one must find how much noise the receiver can tolerate before it can no longer achieve the SER required for acceptable reception.

It turns out that one interference specification has already been published and is becoming an accepted industry practice.[4] Given a transmitted signal with a 27dB SNR, the receiver threshold will degrade about 0.25dB. This means that the new TOV becomes about 15.15 dB. Clearly, this does not represent a significant difference in the receiver threshold. Consider that a 5dB improvement in the transmitted SNR only degrades the receiver threshold by about 0.1dB (a threshold of 15.0dB). So picking up 5dB in SNR only buys you 0.15dB in margin. Conversely, having a transmitted SNR of worse than 27dB will begin to degrade reception margin very rapidly. Where a 5dB change for the better from 27dB to 32dB only picked up .15dB in margin, a 2dB change for the worse from 27dB to 25dB results in an additional 0.2 dB degradation. In other words, the total reception margin changes by about 0.45 dB.

Figure 3 shows a plot of the effect of transmitted SNR on receiver threshold. This calculation assumes that the interference causing the transmitter SNR will be statistically independent of distortions which are added through the broadcast channel. An SNR degradation caused by only linear distortion will not affect the received threshold in the same way since linear distortion is not a random process and does not add to noise the same way other distortions do, e.g. intermodulation products.



• Figure 3 - The degradation in the received C/N in dB based on the transmitted SNR in dB

It is clear that this issue of signal quality will directly relate to coverage area. While it may not be necessary to improve a transmitter's output signal to much better than 27dB, it is critical to make sure that the performance never degrades beyond this level. This is one reason why manual analog correction is not advised for DTV service, since the level of performance is guaranteed only after immediate set up. Once parameters begin to change, the reception margin may degrade very rapidly.

Transmitter Correction Schemes

DAP

One way to attack the problems encountered in digital transmission is with digital technologies. Using a technique called Digital Adaptive Precorrection (DAP), both linear and non-linear transmitter distortions can be compensated for automatically without operator intervention. This assures that the transmitted signal is always of a high quality.

The basic but essential function of the precorrector is to generate a perfect complementary amplitude/amplitude and phase/amplitude curve in such a way that the response of the precorrector and the response of the HPA cascaded together results in a perfectly linearized amplifier.

There are two parts of correction that must be done. The linear equalizer is responsible for pre-distorting the forward signal in such a way as to cancel out linear distortions that occur further on in the transmitter system. It is essentially a complex digital filter whose coefficients are determined by the signal processing unit. The signal processing unit calculates new coefficients by comparing the input signal from the forward path with the reference signal fed back from the transmitter system output.

In much the same way, the non-linear corrector is made up of a large look-up table. The values of the look-up table are updated from the signal processing unit. This allows both the phase and gain characteristic of the forward signal to be manipulated. The signal processing unit calculates the HPA's complementary curve. That is, it first finds the transfer curve of the HPA by comparing the input signal and the feed back signal. It then finds the inverse curve and feeds the look-up table with new values. The result of pre-distorted signal passing through the HPA is a "perfect" signal at the output, with all distortions canceling.

This idea has been used in analog transmitting equipment for many years now. The difference between those analog correctors and the DAP is the methodology used to "tune" the correction. Because of the digital processing involved, the operator involvement is eliminated since the correction is now computed. Also, the computations are very precise, allowing for the overall result to be much more accurate.

If the principle seems quite straightforward at a first glance, its success relies on a deep knowledge of the precorrection principles, a sound global transmitter system approach and a mastery of leading edge digital technology and simulation tools.

Conclusion

It should now be clear that transmission of digital signals can be very difficult compared with analog transmissions. Because the digital signal can be sensitive to linear distortions and because its extremely high peak to average power ratios make it harder to amplify linearly, a transmitter must be designed with a different focus. Since quality of the digital signal will directly translate to coverage and not just to a small degradation in the picture, it is incumbent on transmitter manufacturers to allow broadcasters to deliver the highest possible quality signal with the best power and efficiency. The Comark digital exciter was designed to take these issues of digital transmission into account when integrated with an entire transmission facility, not just the final amplifier. In the end, it will be the application and mastery of new digital technologies which will allow DTV transmission to be implemented in the most efficient fashion.

References

[1] McGillem, Clare D. & Cooper, George R., *Continuous and Discrete Signal and System Analysis*, Holt, Rinehart and Winston, New York, NY, 1984.

[2] Proakis, John G, and Salehi, Masoud, *Communications Systems Engineering*, Prentice Hall, Englewood Cliffs, NJ, 1994.

[3] "Guide to the Use of the ATSC Digital Television Standard", ATSC Document A54, Advanced Television Systems Committee, Washington, DC, October, 4, 1995.

[4] "Transmission Measurement and Compliance for Digital Television", ATSC Document A64.

On the Validity of the Longley-Rice (50,90/10) Propagation Model For HDTV Coverage and Interference Analysis

O. Bendov
Dielectric Communications

Introduction

In 1948 the FCC issued a "freeze" order, suspending new and pending applications for construction of television-broadcasting facilities pending further study of channel allocation and the method of service and interference prediction. The "freeze" was lifted in 1952 as part of the then FCC's 6th Report & Order, which included the NTSC table of channel allocation and propagation curves. Thereafter, the UHF propagation curves were deemed inaccurate. It took 271 engineers three years, from 1956 to 1959, to complete the experiments leading to the establishment of the propagation curves now in use for NTSC stations. That task was carried out by the Television Allocation Study Organization (TASO), set up by the FCC.

Fifty years later, in 1998, the FCC issued another 6th Report & Order (an ironic coincidence!) that included the channel allocation table for digital television. Unfortunately, the Longley-Rice (LR) propagation model, as used by the FCC for interference and coverage was not validated prior to the issuance of the allocation table. To-date, there is no known concerted effort to validate the LR model. The validity of the LR model can be questioned on several grounds. Among the questions are -- will history repeat itself and will a new TASO have to be convened?

Chances are, history will repeat itself. Unacceptable interference by HDTV stations to cochannel and adjacent channel NTSC stations have been reported. Mandatory standards for minimal receiver performance do not exist. The threshold level by multiple interferers with a single victim remains unknown. Indoor reception is uncertain. And, the few field tests, limited in scope and with widely varying and unexplained results, have not answered some critical questions. One critical question, fundamental to channel allocation, interference and service prediction, is the concern of this paper.

Examples

A. Channel 53 in Charlotte, NC

Field testing of the Grand Alliance's HDTV transmission subsystem was conducted during 1994 in Charlotte, NC. The terrain was variable. The radiation center was approximately 1940' above mean sea level on a tower 1337' above ground. The Effective Radiated Power (ERP) was set at 31.6 kW.

The expected coverage calculated using the LR propagation model with the parameters recommended by the FCC is shown in Figure 1. The circled locations are the locations were HDTV reception had failed. The radio horizon map is shown in Figure 2. This map provides a more realistic prediction of coverage than the LR model. To match the coverage shown in Figure 1 with that of Figure 2, the percentages of locations and time availability would have to be raised from 50,90 to 99,99 and the confidence margin would have to be raised from 0 to 10 dB.

B. WHD, Channel 30, in Washington D.C.[1]

The model station, WHD, operates with an Effective Radiated Power (ERP) of 440 kW (56.43 dBw) from an antenna 405 feet above ground. The measurements analyzed here were performed within the arc over which the ERP is relatively constant. The coverage predicted by the LR model is shown in Figure 3. The circles in Figure 3 are the locations were HDTV service had either failed or is deemed unreliable. Failed or unreliable service locations are defined here as those sites with carrier-noise margin of ≤ 1 dB. The percentage failure was 32% (84 of 263 outdoor sites).

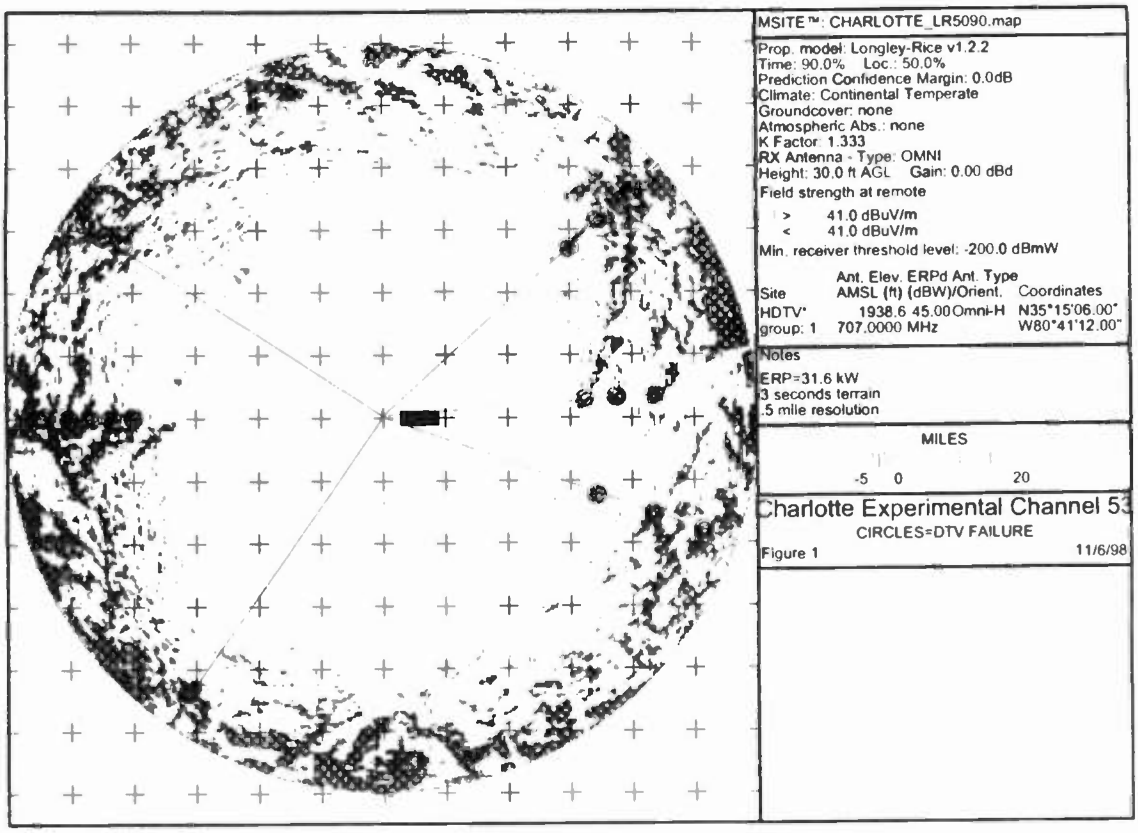
Figure 4 shows the failed sites plotted on the radio horizon/shadow map. While the radio horizon map may be an improved indicator of coverage over the LR(50,90), it is useless for interference analysis.

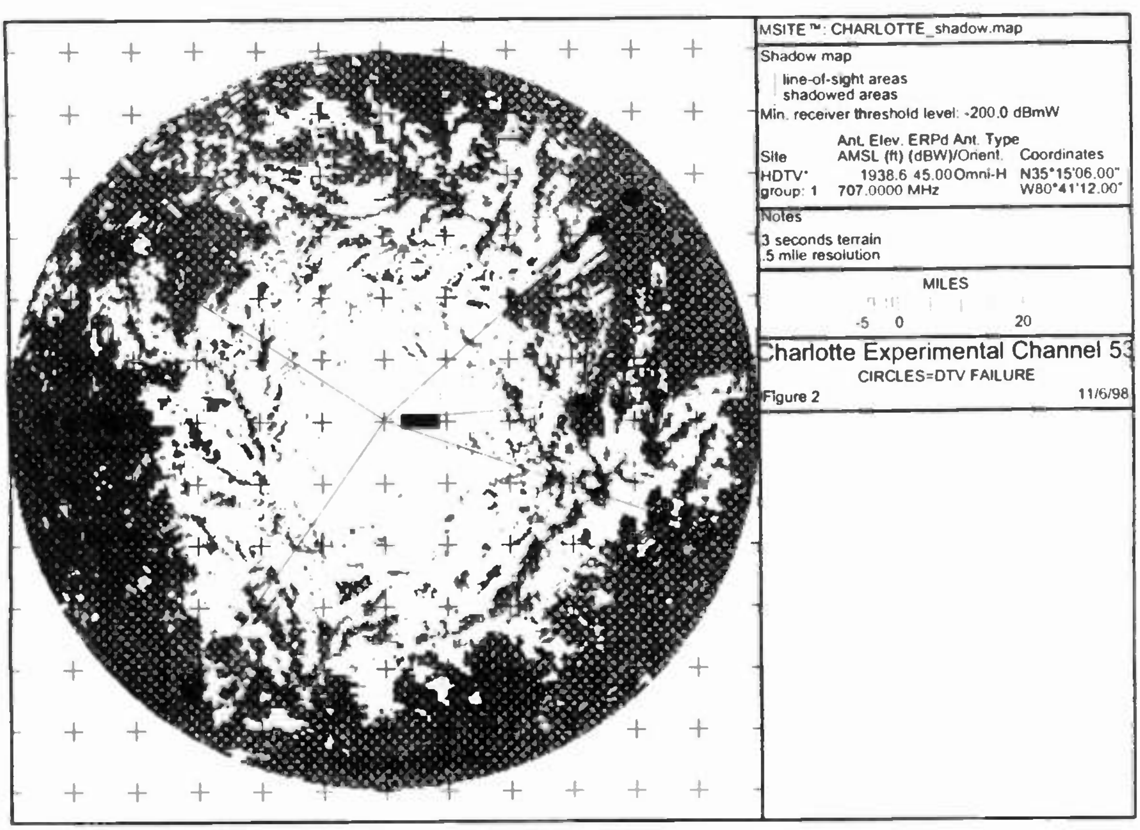
In all of the failed sites, the measured "Field Strength[2]" of the HDTV signal was well below that predicted by

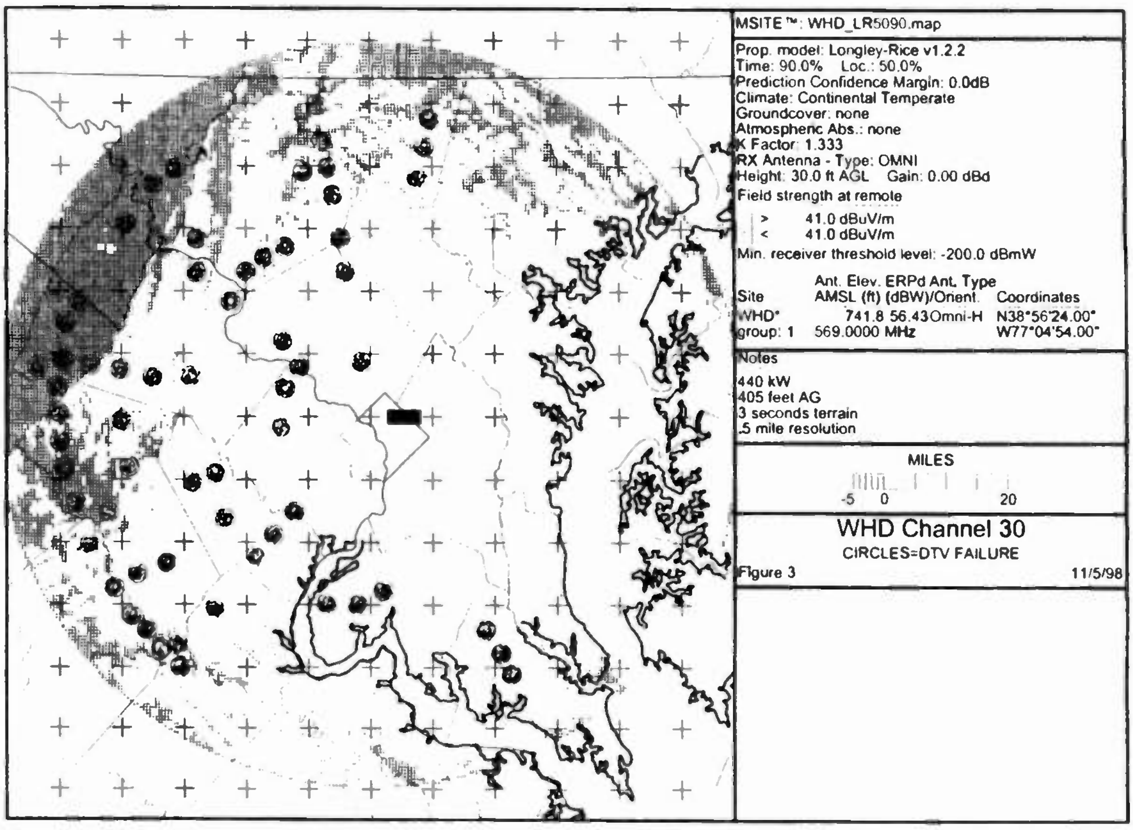
---

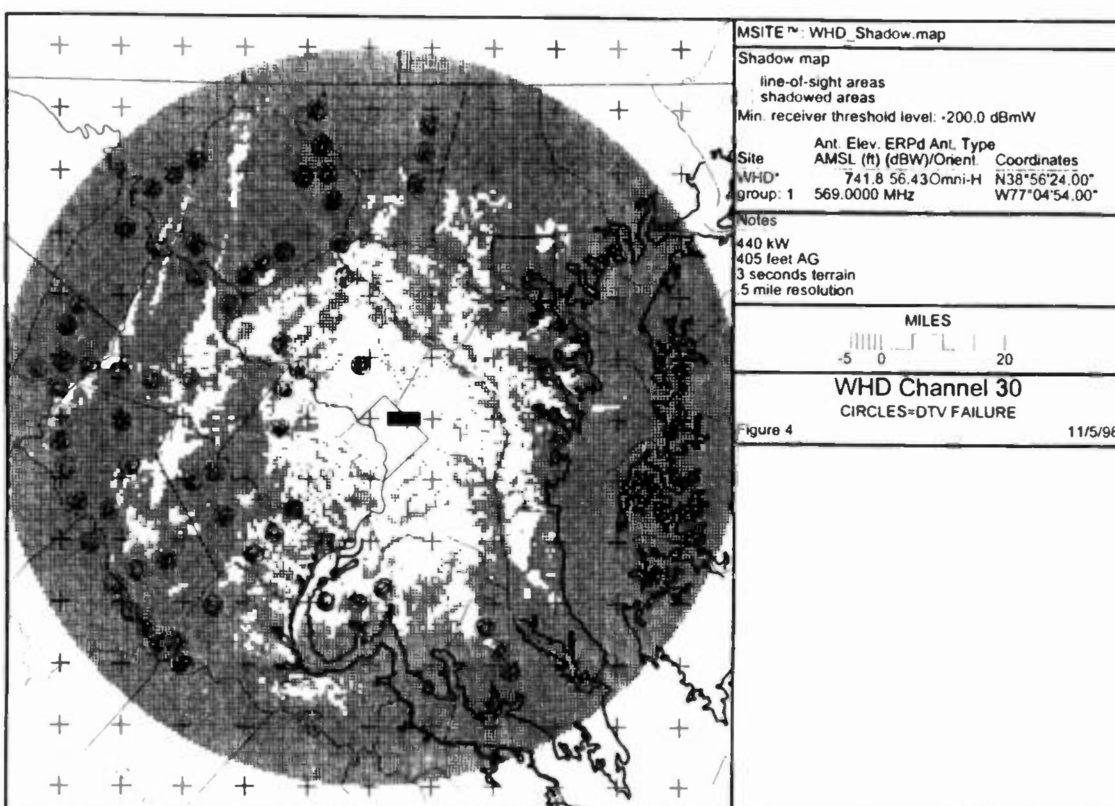
[1] Test data is published in the interim report (September 1998) of the Model HDTV Station Field Test Program. I am grateful to V. Tawil of AMST Inc. for providing the data on a disk.

[2] The field strength of HDTV cannot be measured directly. It is calculated from the total power in 6 MHz and does not include the effective loss of power due to multipath. For a detailed discussion of this point, see "Predicting HDTV Coverage," Broadcast Engineering, March 1966.


MSITE™: CHARLOTTE_LR5090.map
Prop. model: Longley-Rice v1.2.2
Time: 90.0% Loc.: 50.0%
Prediction Confidence Margin: 0.0dB
Climate: Continental Temperate
Groundcover: none
Atmospheric Abs.: none
K Factor: 1.333
RX Antenna - Type: OMNI
Height: 30.0 ft AGL Gain: 0.00 dBd
Field strength at remote
> 41.0 dBuV/m
< 41.0 dBuV/m
Min. receiver threshold level: -200.0 dBmW
Site Ant. Elev. AMSL (ft) ERPd (dBW) Ant. Type /Orient. Coordinates
HDTV* 1938.6 45.00 Omni-H N35°15'06.00"
group: 1 707.0000 MHz W80°41'12.00"
Notes
ERP=31.6 kW
3 seconds terrain
.5 mile resolution
MILES
-5 0 20
Charlotte Experimental Channel 53
CIRCLES=DTV FAILURE
Figure 1
11/6/98


MSITE™: CHARLOTTE_shadow.map
Shadow map
line-of-sight areas
shadowed areas
Min. receiver threshold level: -200.0 dBmW
Site Ant. Elev. AMSL (ft) ERPd (dBW) Ant. Type /Orient. Coordinates
HDTV* 1938.6 45.00 Omni-H N35°15'06.00"
group: 1 707.0000 MHz W80°41'12.00"
Notes
3 seconds terrain
.5 mile resolution
MILES
-5 0 20
Charlotte Experimental Channel 53
CIRCLES=DTV FAILURE
Figure 2
11/6/98

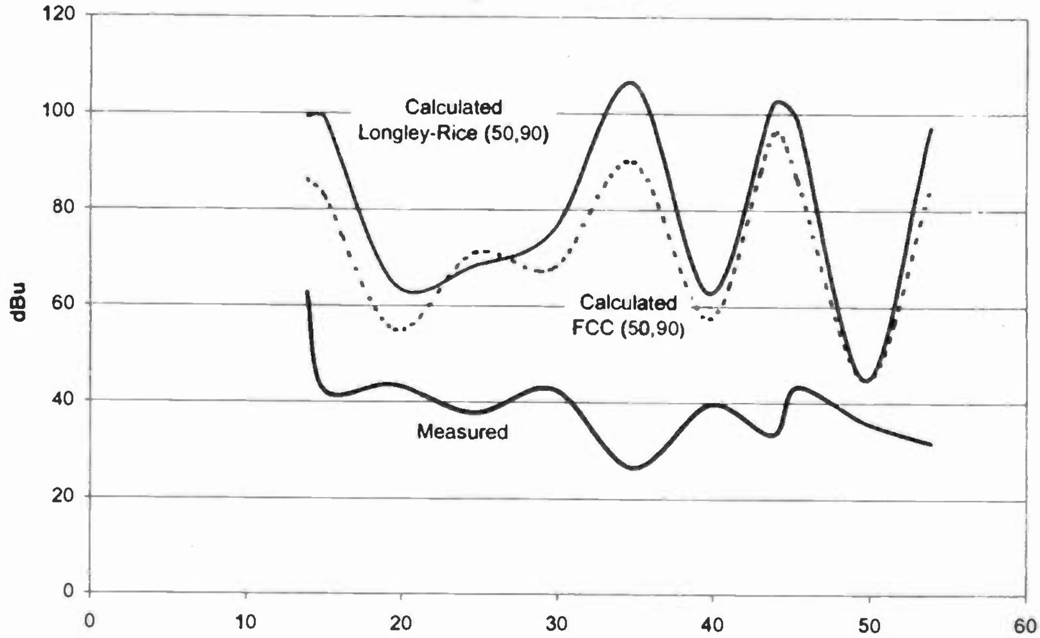

MSITE ™: WHD_LR5090.map
Prop. model: Longley-Rice v1.2.2
Time: 90.0% Loc.: 50.0%
Prediction Confidence Margin: 0.0dB
Climate: Continental Temperate
Groundcover: none
Atmospheric Abs.: none
K Factor: 1.333
RX Antenna - Type: OMNI
Height: 30.0 ft AGL Gain: 0.00 dBd
Field strength at remote
> 41.0 dBuV/m
< 41.0 dBuV/m
Min. receiver threshold level: -200.0 dBmW
Site Ant. Elev. AMSL (ft) ERPd (dBW)/Orient. Ant. Type Coordinates
WHD* 741.8 56.43 Omni-H N38°56'24.00"
group: 1 569.0000 MHz W77°04'54.00"
Notes
440 kW
405 feet AG
3 seconds terrain
.5 mile resolution
MILES
-5 0 20
WHD Channel 30
CIRCLES=DTV FAILURE
Figure 3
11/5/98

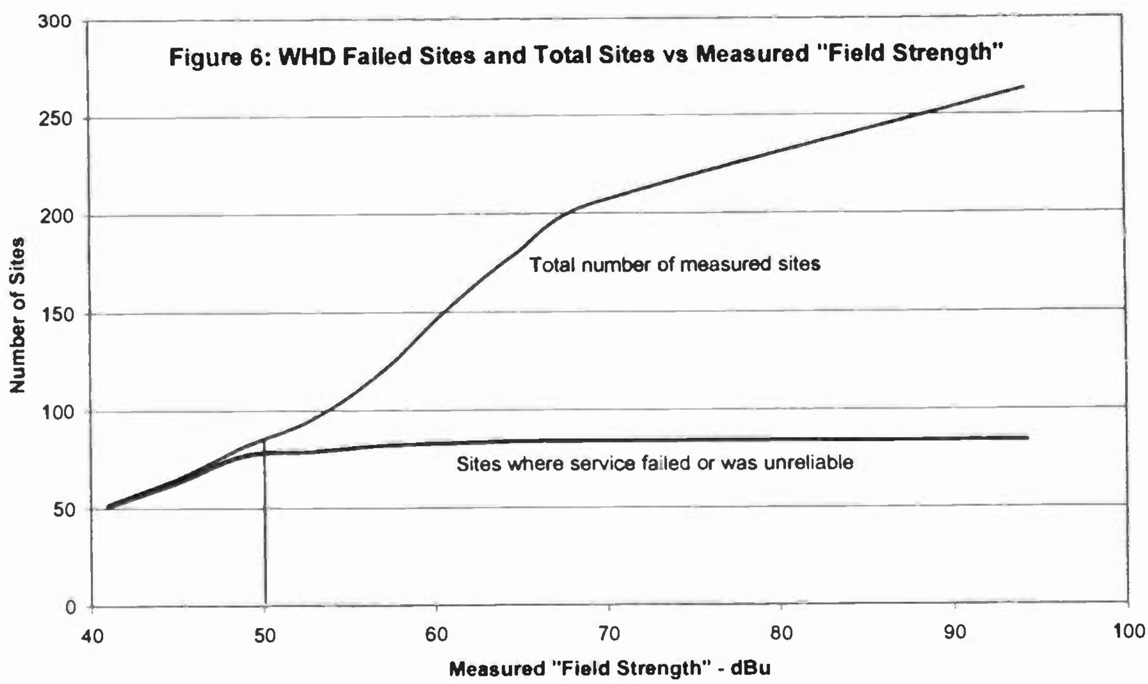

MSITE ™: WHD_Shadow.map
Shadow map
line-of-sight areas
shadowed areas
Min. receiver threshold level: -200.0 dBmW
Site Ant. Elev. AMSL (ft) ERPd (dBW)/Orient. Ant. Type Coordinates
WHD* 741.8 56.43 Omni-H N38°56'24.00"
group: 1 569.0000 MHz W77°04'54.00"
Notes
440 kW
405 feet AG
3 seconds terrain
.5 mile resolution
MILES
-5 0 20
WHD Channel 30
CIRCLES=DTV FAILURE
Figure 4
11/5/98

either LR(50,90) or by the F(50,90) curves as shown in Figure 5. Clearly, a "Field Strength" of 41 dBu cannot serve as reliable predictor of coverage.

Even with higher "Field Strengths," service has failed in many locations. As shown in Figure 6, service has essentially failed at 94% of the locations in which the "Field Strength" did not exceed 49 dBu. The difference between the number of sites visited and the number of failed sites begins to spread for "Field Strengths" > 49 dBu.

Figure 5: WHD "Field Strength" at Sites Where HDTV Service has Failed or is Unreliable
Average of All Directions





Evaluating the coverage of just two stations is hardly definitive, but the results suggest that the LR(50,90) propagation model may not be generally applicable for reliable HDTV coverage prediction unless its input parameters are properly adjusted. Similarly, the usage of LR(50,10) for interference analysis may need reevaluation. The shortcomings of the LR model are:

- Unrealistic statistical margins.
- Multipath, the most serious cause of HDTV service failure, is ignored.
- The LR algorithm applies to a single carrier, not to wide band (6 MHz) signals.

Realistic Propagation Modeling

In producing the HDTV channel allocation table, the FCC has relied, simultaneously, on two different propagation models. One model was the empirically derived curves for NTSC. The second was the LR. From a technical standpoint, there would be no reason for mixed use of different models, with widely divergent prediction of service, if one of the two were appropriate for the purpose.

A mathematical propagation model requires that the user enter several critical parameters before calculation can proceed. As is always the case with computer modeling, the axiom "garbage in – garbage out" applies. What are the critical parameters and what values did the FCC assign these parameters?

Height of Receive Antenna

The median gain/loss of signal due to a change in the height of a UHF receive antenna from a reference height of 30 feet is[3]:

H is the height in feet, A=4, 6, and 8 are respectively, for Rural, Suburban and Urban areas.

$$G \approx \frac{A}{6} 20Log\left(H/_{30}\right) \quad dB$$

The FCC set the antenna height above ground at 30'. That elevation may have been appropriate during the 1950's. Nowadays, the average height of outdoor antennas would be lower. For example, if the height of the receive antenna were 15' above ground, a received UHF signal would be –6.0 dB below that expected from an antenna 30' above ground. It would take quadrupling of the transmitter power to make up this loss.

Receiver Noise Figure

For UHF-NTSC receivers, the FCC mandates a minimum noise figure of 14 dB. A minimum noise figure specification for HDTV receivers has not been mandated. The FCC has used a noise figure of 7 dB as a planning factor for UHF-HDTV but that noise figure is not binding on the manufacturers of HDTV sets.

Even if the manufacturers specify 7 dB as a noise figure, that noise figure cannot be used for service prediction without further modification. The reason for that is that the factory noise figure is measured with the receiver ("load") matched to the noise generator ("source"). A household antenna is rarely matched to the receiver. A Standing Wave Ratio (SWR) of 5:1 across the UHF band is not unusual and a ratio of 2:1 is common. The effective noise figure increases for a mismatched antenna/receiver. For a 2:1 mismatch, the effective noise figure increases by 3 dB over the factory's noise figure. It would take more than doubling of the transmitter power to make up this loss.

The factory specification for noise figure of production run UHF-HDTV receivers will probably be closer to 10 dB. That is 3 dB higher than that specified in the planning factors. The effective noise figure, accounting for the SWR in the downlead cable, would then be closer to 13 dB than to 7 dB. All these shortfall could be mitigated by a "smart" receive antenna[4].

The Effective Earth Radius Multiplier Factor

The variations in propagation conditions require that the US be divided into three zones. The 1959 TASO report recommended that the multiplier factors for the three zones be:

Eastern Seaboard	Ka=1.75
Gulf Coast	Ka=1.85
Remainder of US	Ka=1.33

Therefore, the radio horizon, which is the line-of-sight for UHF waves, is not uniform around the US. Based on effective earth radius recommended by TASO, the radio horizon for various antenna heights around the US are:

	Gulf Coast	Eastern Seaboard	Remainder of US
Effective Earth Radius:	7326 m	6930 m	5280 m
	Radio Horizon (miles)		
Antenna @ 1000 feet	53	51	43
Antenna @ 1500 feet	65	63	55
Antenna @ 2000 feet	75	72	63
Antenna @ 2500 feet	83	81	71

[3] ITU Recommendation P.370-7, 1995.

[4] First proposed by the author at the PS/WP3 meetings and later elaborated in two papers presented by the author at the 1994 and 1997 NAB conventions.

The FCC's LR model assumes an effective earth radius of 5280 miles everywhere.

Statistical Margins

The received power level of over-the-air transmission depends on factors that are neither constant nor exactly accountable for. Local terrain and weather variations are examples. Therefore, statistics and probability are used to supplement the calculation of the received signal level. With a statistical description of measured data at hand, a safety margin is assigned to the unpredictable degrading factors. The margin, expressed in dB, is a power-loss penalty subtracted from the ERP to ensure that the statistics of HDTV reception would meet at least the preset levels.

The LR model allows for an assessment of most but not all of the required margins. Included in LR are margins for location and time variations and confidence level. Not included in LR, but a necessity in digital RF links, is a margin for multipath.

a. Reliability margin

This margin, in dB, is subtracted from the median signal level to insure a higher probability than having the desired signal at just the best 50% of the location at least 50% of the time. What the correct percentages are for digital television where the picture abruptly disappears rather than fades, no one knows. The FCC has set the same reliability statistics for HDTV as those in use for NTSC. That is, HDTV is considered to be reliably available if it can be viewed at least 90% of the time at the best 50% of the locations. At the other 50% of locations, any percentage of time availability would be acceptable. A loss of picture 10% of the time within the best 50% of the locations would be acceptable as well.

b. Confidence margin

The LR model requires as input of the confidence level that the designed reliability (time and location) will be met by a percentage of broadcasters in any one market. The FCC set the confidence level at 50%. That means that at least 50% of the stations in any one market will have HDTV signals available at least 90% of the time at the best 50% of the locations. The signal availability statistics of other 50% of the stations in the same market may be lower.

It has been pointed out[5] that the result of using the LR model with a 50% confidence level is not comparable to that of using the FCC's (50,50) curves, and that confidence levels of at least 90% must be used in the UHF band.

c. Receiver equalizer/multipath margin

Experience to date has shown that multipath propagation is a major detriment to HDTV reception, even when received by outdoor antennas. The effect of multipath propagation appears as in-band amplitude and phase distortion. If the multipath does not exceed a certain time delay and/or certain magnitude, the equalizer at the receiver will provide the correction – at a price. The price, in dB, amounts to a power penalty that effectively increases the threshold level of the carrier – noise ratio.

The use of a multipath ("dispersion") margin in wideband digital radio links is well known and thoroughly documented[6]. The addition of this safety margin is important because the LR model is presently designed for single-frequency signals only.

Interference Prediction

The assumptions in the FCC's Planning Factors are that the outdoor receive antenna is 30 feet above ground and that the antenna's backlobe (-14 dB @ UHF) is oriented toward the interfering station. But what about cable head-ends with receive antennas 300 feet above ground? What if the home user cannot rotate the outdoor antenna in such a way that the backlobe is toward the interfering station? What if rotating the antenna to minimize the interference also causes a loss of the desired channel? These questions may have to be faced in the months to come as the levels of cochannel and adjacent channel interference into NTSC channels becomes clearer[7]. Here too, a validation of LR(50,10) may be timely.

Excessive interference could also result from higher then planned ERP in some directions. There are two potential causes for excessive ERP:

- Filing based on an incorrect antenna gain.

Until November 1998, licensing applications for HDTV used form FCC 302 which did not require that the antenna's RMS and peak gains both be listed. The peak gain can exceed the RMS gain by several dB.

[5] Louis A. Williams, Jr., private communication sent to the Association of Federal Communications Consulting Engineers (AFCCE) DTV Committee dated August 2, 1997.

[6] For example, "A Simplified Method for Prediction of Multipath Fading Outage of Digital Radio" by Y. Serizawa and S. Takeshita, IEEE transactions on Communications, Vol. COM-31, August 1983. Also, EDX Engineering MSITE™ manual.

[7] As of November 1, 1998, unexpected cochannel interference has been reported in WI and adjacent channel interference has been reported in PA.

Figure 7: RMS AND PEAK ERP FIT TO THE FCC's ASSIGNED HDTV DIRECTIONAL ANTENNA PATTERNS

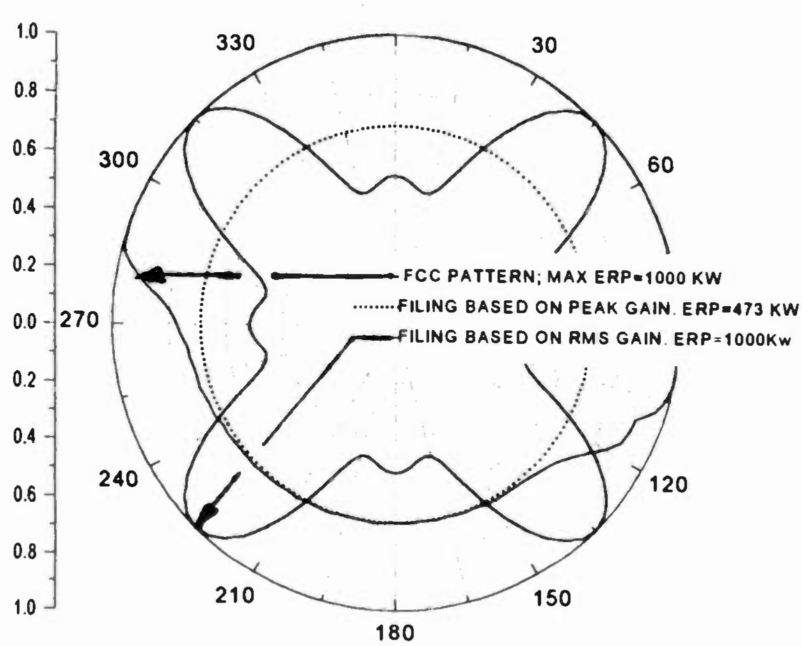
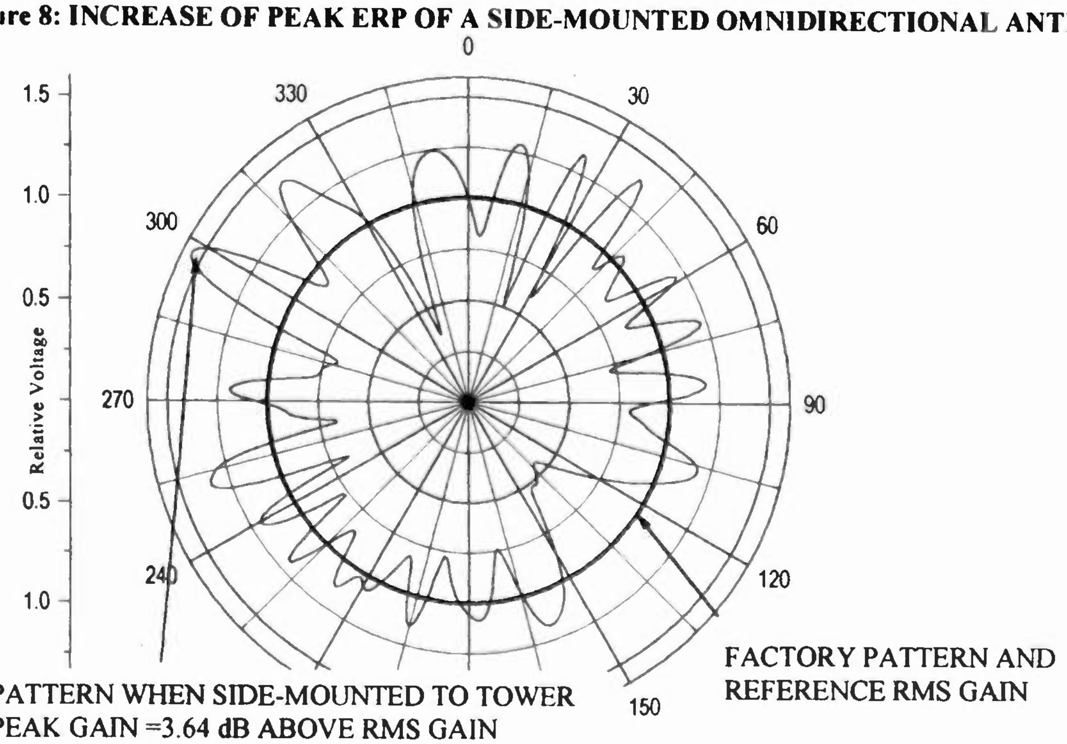


Figure 8: INCREASE OF PEAK ERP OF A SIDE-MOUNTED OMNIDIRECTIONAL ANTENNA



Since the Commission's rules are not specific on this matter and RMS gain has been used for omnidirectional NTSC stations, it could be assumed that either gain is acceptable. The relation between the two definitions of gain is shown in Figure 7.

The peak gain of the omnidirectional antenna pattern is 3.25 dB above the RMS circle. Therefore, if filed on the basis of RMS gain, the antenna whose pattern is shown in Figure 7 could be used as part of a checklist application and licensed for ERP of 473 kW (RMS) which would not exceed the allowable ERP anywhere. However, the peak ERP of that antenna would reach 1,000 kW, exceeding the allowable ERP in two directions.

In November 1998, a new form, FCC 302-DTV, was issued. The new form requires that both peak and the RMS gains of the antenna be specified. That requirement closes only half of the loophole. For example, antennas may require a larger diameter support pole in order to meet certain structural specifications. Now, should the RMS and peak gains of the antenna assuming the supporting pole does not exist be submitted or should the as-installed patterns be submitted?

- As-installed antenna pattern.

Antennas that are side-mounted on the tower shaft and antennas that are placed next to one another on one tower will have their patterns modified by the proximity to other conducting objects. An example of the modified pattern of an omnidirectional antenna, side-mounted next to a tower with an 8-feet face, is shown in Figure 8. Clearly, in some directions the ERP has increased by as much as 3.64 dB.

The potential problems that could arise from the situations described are easily correctable. First, license applications should include the antenna patterns and gain specifications with the support poles (if used). Second, for antennas installed within 50 feet of a metallic obstruction, factory patterns as well the as-installed patterns would be required as part of the license application.

Conclusion

Analysis of the available field test results coupled with key theoretical considerations shows that a modification of the LR model will be required before it could be effectively used for HDTV coverage and interference prediction[8].

This paper has also demonstrated that a "Field Strength" 41 dBu is inadequate and inappropriate as a measure of HDTV service contour and as a measure of service within that contour.

The limits to reliable HDTV service will not be known until several thousand sets are in place in each of the major markets. In the meantime, the radio line-of-sight could serve more reliably as a predictor of HDTV coverage than LR(50,90).

Several steps should be taken now to help smooth the transition period ahead:

First, the LR model should be modified for wide-band signals, its statistical margins adjusted and multipath margins for urban and suburban areas added. The multipath margins would be related to the performance of the receiver's equalizer.

Second, future field tests should be expanded for the purpose of determining the correct parameters used by the LR model for coverage and interference predictions.

Third, the FCC should establish minimum performance standard for HDTV sets. The standard should cover noise-figure, selectivity and equalization.

Fourth, the FCC, in cooperation with industry and academic institutions, should establish TASO II as an advisory group to help resolve key transmission and allocation problems related to digital television broadcasting.

[8] There may also be software code implementation error. Hammett & Edison, a consulting engineering firm, has asserted to the FCC that, on average, 18% of the population reside in cells arbitrarily assumed to get HDTV service even though a proper LR(50,90) calculation could not be performed in those cells.

RESULTS OF TV RECEIVE ANTENNA TESTING FOR CORRELATION OF DIGITAL AND ANALOG TELEVISION RECEPTION USING THE CEMA ANTENNA SELECTOR MAP PROGRAM

Ray Conover
US Satellite Broadcasting Co.
St. Paul, MN

ABSTRACT

Field testing of digital television signals from WBTV-DT Charlotte, NC was undertaken to determine whether the receive digital signal quality comported with the assumption that the CEMA Antenna Selector Map Program is also a useful tool for consumer reception of digital television. The receive picture quality from NTSC stations in the Charlotte market was compared to WBTV-DT's digital television reception. Correlation of these results included examination of the CEMA Antenna Map Program's demarcations of appropriate receive antenna performance categories and their application to digital TV reception.

1.0 INTRODUCTION

The the antenna mapping program described herein is a project of the Consumer Electronics Manufacturers Association. (CEMA) The mapping project originated at US Satellite Broadcasting Co. and has now been licensed to CEMA for industry wide use. The project is administered by two groups within CEMA, the Antenna Subdivision of the Accessories Group, and by the R-5 (Antennas) Standards Committee.

The goals of the CEMA antenna mapping project are to promote the use of TV antennas and to promote the growth of the TV antenna manufacturing industry without favoring any particular brand or manufacturer. Many antenna manufacturers have donated antennas, preamplifiers, and other reception accessories to USSB and then to the CEMA R-5 (Antennas) Committee for purposes of executing the program. Since this project is an industry wide effort specific antenna model numbers are not referenced in this report in keeping with the goal of not favoring any particular product.

The CEMA Antenna Selector Map Program is designed to provide instant antenna selection expertise to the retailer wishing to sell antennas to consumers. The hope is to increase the rate of appropriate antenna selection to a 70% to 80% success range.

The Antenna Selector Map Program combines the Longley-Rice propagation model with unique algorithms developed for the mapping program. A variety of databases are then used to produce a required antenna prediction for a set of TV stations listed on each map. The maps are calculated by using the technical characteristics of the listed TV stations at each pixel on the map. The signal strength component of the prediction is that of the weakest signal of the listed TV stations. Other algorithms take into account the performance of the television transmission on paths that are obstructed or in the interference pattern caused by large structures.

The map program has two components, the color coded Antenna Selector Map, and the companion Antenna Selector Guide. The consumer finds their location and corresponding color on the relevant map, and then answers a few qualifying questions about their immediate neighborhood in the Antenna Selector Guide. If the consumer location is in a low area, or an area with nearby tall structures the Antenna Selector Guide will recommend a different more appropriate color coded antenna. The consumer then selects any antenna bearing the recommended color coded CEMA antenna logo.

Antenna Selector Map technical details

The maps predict the minimum acceptable antenna required to produce a CCIR grade 3.5 or better picture from an antenna mounted at 30' and distributed to four TV receivers. The color contours on the map correspond to antenna categories defined in the CEMA engineering bulletin CEB/6A that can be summarized as follows:

Small multidirectional antenna - suitable for use in the yellow contour when the immediate area around the receive location is not a low area compared to neighborhood terrain, and the area around the receive location is free of tall structures within about half a mile. Technically it is a low gain antenna that may have nulls, but no useful directivity. Small disk shaped antennas are typical of this category.

Medium size multidirectional antenna - suitable for use in the green contour when the immediate area around the receive location is not a low area compared to neighborhood terrain, and the area around the receive location is free of tall structures within about half a mile. Technically it is a low gain antenna that may have nulls, but no useful directivity, but does have more low band VHF gain. Dipole and wing shaped antennas are typical of this category.

Large size multidirectional antenna - suitable for use in the light green contour when fitted with a preamplifier and mounted at roof level. This antenna can be used when the immediate area around the receive location is not a low area compared to neighborhood terrain, and the area around the receive location is free of tall structures within about half a mile. Technically it is an antenna that behaves like a dipole and has close to 0dBd gain on all channels. Larger dipole shapes and small Yagis are typical of this category.

Medium size directional antenna - Suitable for use in the red map contour, the antenna is large enough to have a good front to back ratio on all channels making it a useful ghost killing antenna. The red map area represents problem receive areas caused by obstructed paths or multipath. It is the default antenna when antennas otherwise suitable for the Yellow, Green, or Light Green map contours cannot be used due to low terrain or structure induced multipath in the immediate area of the receive antenna.

Medium size directional antenna with preamp - Suitable for use in the blue map contour the antenna is large enough to have a good front to back ratio on all channels making it a useful ghost killing antenna. The required preamp compensates for line and splitting loss, making the antenna useful at lower field strength levels.

Large size directional antenna with preamp - Suitable for use in the violet or pink map contours the antenna is large enough to have a good front to back ratio on all channels making it a useful ghost killing antenna. The required preamp compensates for line and splitting loss. The antenna's higher gain receives the maximum possible signal for use at great distance or other low signal areas. Technically it meets or exceeds the FCC's digital reference antenna. When used in the violet map area all of the listed stations will be above the FCC's no service signal strength. When used in the pink map area at least one of the listed stations will be above the FCC's no service signal strength.

Consistent categorization of TV receive antennas required an accompanying antenna test procedure which is contained in the EIA-774 standard created by the R-5 committee.

Picture Quality Measurement

Throughout this paper the CCIR 5 Point Grading Scale for picture quality has been employed where 5.0 is a perfect studio quality picture and 0 is no picture. While this scale is subjective, all the ratings in the CEMA mapping program, including the Charlotte tests, have been rated by the same individual. In the CEMA antenna mapping program tests the best rating given to an off air NTSC signal is 4.1 due to the limitations of AM vestigial sideband transmission systems. For purposes of this report on transmission quality where the principle impairment being evaluated is multi path, and the resultant ghosting, the scale is applied to ghosting only. Impairments due to noise or other factors will be noted separately. Successful reception of WBTV's digital picture is noted by either a "Y" for yes or "N" for no.

Digital compatibility

The CEMA R-5 committee employed extra effort to assure that the CEMA Antenna Selector Map program would be compatible with digital television reception since its introduction at this time would, whether desired or not, be linked with the introduction of digital television. The parties involved wanted to make sure that the mapping program was indeed compatible with digital television and the committee needed to answer the question whether or not the map would select antennas suitable for digital TV, hence the tests in Charlotte.

2.0 CHARLOTTE TV MARKET

Charlotte, NC was an ideal location for conducting such a test since there is a wide variety of terrain that represents all common receive antenna locations except large mountains. In addition the television transmitters in the Charlotte metropolitan area are located in two distinct groups separated by substantial distance. This is one of the cities where the group separation was so large that it was necessary to make two separate maps of the Charlotte area. (Two or more maps are used when groups of transmitters are separated by distances that would not allow a single map to make sense for use in a consumer antenna selection process.) Thus the Charlotte area allows us to functionally test the performance of the map as if testing two separate cities since every test location has a group of stations with a nice mix of channels in two different directions and at two different distances.

The group of transmitters referred to as the Northwest group includes WBTV-DT and is the basis for one Charlotte TV Antenna Selector Map. The Group of stations transmitting from Northeast of Charlotte are referred to as the Northeast group.

3.0 TEST FACILITIES

WBTV had outfitted one of it's older ENG vans with test equipment for purposes of conducting tests of their digital transmissions. The van had a telescoping pneumatic mast with a Pan / Tilt head. The ENG microwave antenna had been replaced by a professionally cut UHF Yagi antenna that feeds approximately 60 feet of RG/6 coax. The normal WBTV test antenna has a preamp that can be switched in and out from the control position in the truck, but for these tests neither the antenna or preamp was used. Instead when a preamp appeared to be called for, or required, a typical consumer preamp along with the consumer antenna was installed.

The coax from the antenna on the telescoping mast enters the truck's system via a patch bay and is routed into a 4-way splitter that is routed again via the patch bay to an ATSC signal analyzer, an HP 8558 spectrum analyzer, an ATSC receiver, and a conventional NTSC tuner. The ATSC digital receiver supplied an RGB signal to a small HDTV monitor.

The CEMA antenna selector map assumes coax line loss and loss from a 4-way split. The losses in the WBTV van between the antenna and the individual devices were similar to the line loss and splitting loss assumed in the CEMA Antenna Mapping Project.

Measurement locations and distance from the WBTV-DT transmitter was determined using a GPS receiver mounted on the WBTV test van.

3.1 Test equipment limitations

Several limitations were discovered in the test facilities that must be considered when reviewing the test results. It appears that the splitter employed in the test truck was probably rated to 500 MHz or so and performed quite well on the VHF band through the lower portion of the UHF band. Though accomplishing its task of being a good device for WBTV's digital signal on Channel 23, it was found to have excessively high splitting loss on higher UHF channels. Typical loss from the patch bay through the splitter, through the patch bay again, to the four receiving devices was 8 dB through channel 23. The insertion loss increased to 19 dB by the time Channel 55 was reached. This did not present a major problem since most tests involving the higher UHF channels involved evaluating them for ghosting in relatively high signal strength areas, and this could be done while ignoring the added noise. Also it was also noted that there was some interaction between the displayed signal strengths on the upper UHF channels and the tuning of the NTSC receiver. This suggests that the tuner input impedance was not constant across the band, changing its impedance with tuning.

It was originally intended to connect a portable computer to the digital TV receiver for purposes of collecting data from the receiver related to receive signal strength and equalizer tap weights. Unfortunately, after working through two locations it was realized that the computer connection was damaging our digital performance by allowing digital crosstalk into the RF sections of the digital receiver. The computer was disconnected to allow the receiver to perform its best. Thus all digital reception at locations 1 and 2 was impaired, and it became clear that the receiver could perform much better than those first two sites indicated.

4.0 FIELD TESTS

Throughout the testing measures were taken to make sure that tests are conducted in a manner similar to the way the consumer would use antenna products. In the following field tests the antenna that would have been recommended by the Antenna Selector Map / Antenna Selector Guide combination is noted by displaying the map color in bold type. The test locations are shown in Figure 1 below.

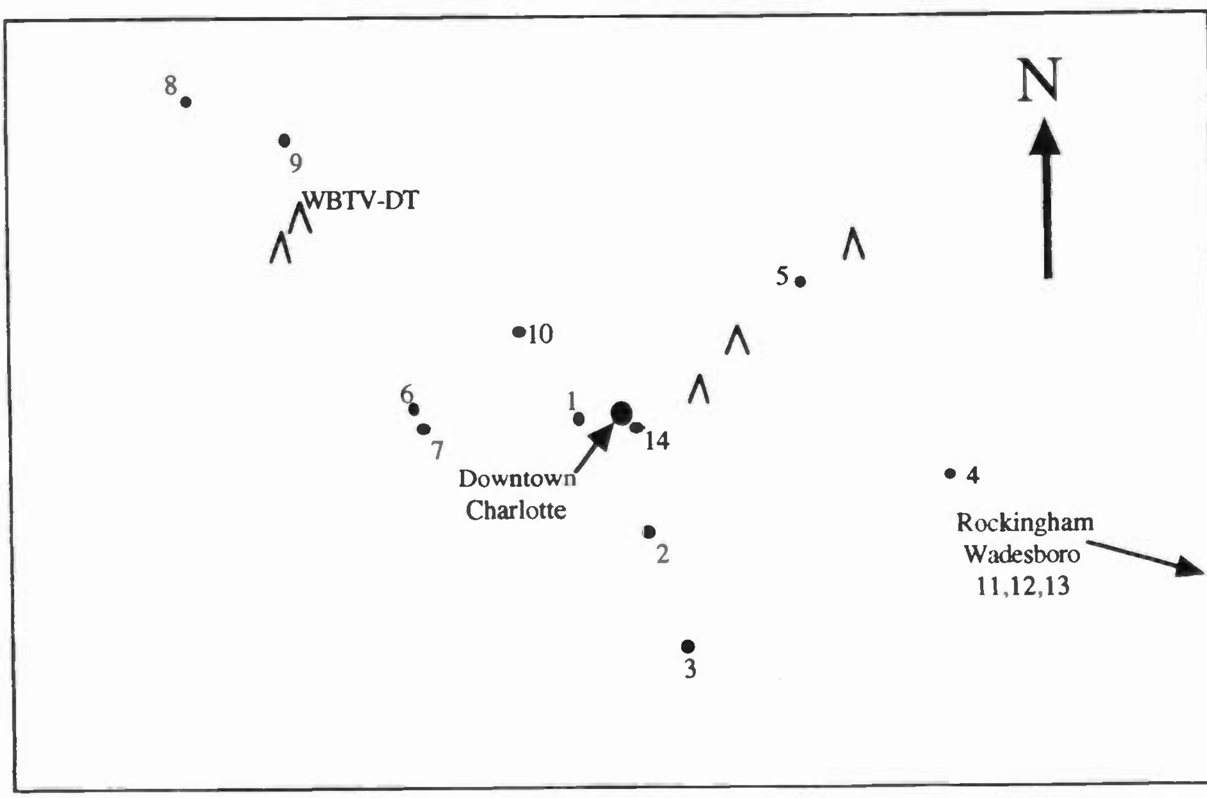


Figure 1

The field testing with the ENG truck and supply van full of antennas started on the roof of WBTV's parking garage which is roughly at ground level. The parking garage is built down the side of a hill, hence the roof being at ground level. This location was an obstructed path towards WBTV's transmitter and the other stations located with it. It was, however, a fairly clean path towards the Northeast group of transmitters with the signal path passing just north of downtown Charlotte. This location allowed the test of a wide range of antennas while keeping supply lines short. It allowed the number of antennas transported into the field to be reduced since antennas with like performance from different manufacturers could be reduced to one antenna of each type. The following tables summarize the performance from the WBTV parking ramp, location number 1.

4.1 Location #1

WBTV Parking Ramp

Antenna #1: Unamplified Omnidirectional disk

NW Stations - Red

Channel	3	23	36	46	55
12' Quality	1.7	N	3.3	1.8	.8
30' Quality	1.7	N	3.3	2.5	.7

NE Stations - **Yellow (Near Red)**

Channel	9	18	42	58
12' Quality	3.8	3.7	3.8	3.9
30' Quality	4.1	3.9	4.0	3.8

Antenna #2: Medium directional

NW Stations - **Red**

Channel	3	23	36	46	55
12' Quality	3.5	Y	2.9	3.5	3.6
30' Quality	3.9	Y	4.1	4.1	4.1

NE Stations - Yellow (Near Red)

Channel	9	18	42	58
12' Quality	4.1	4.1	4.0	4.0
30' Quality	4.1	4.1	4.0	4.0

Antenna #3: Corner reflector with bow tie feed (UHF Only)

NW Stations - Red

Channel	3	23	36	46	55
12' Quality	-	Y	3.2	3.6	3.3
30' Quality	-	Y	4.0	4.0	4.0

NE Stations - Yellow (Near Red)

Channel	9	18	42	58
12' Quality	-	4.1	4.1	4.1
30' Quality	-	4.1	4.1	4.0

Antenna #4: 8 element UHF Yagi

NW Stations - Red

Channel	3	23	36	46	55
12' Quality	-	N	2.5	3.4	2.9
30' Quality	-	Y	3.9	3.9	3.7

NE Stations - Yellow (Near Red)

Channel	9	18	42	58
12' Quality	-	4.1	4.1	4.1
30' Quality	-	4.1	4.1	4.1

Antenna #5: 10 element UHF yagi with corner reflector

NW Stations - Red

Channel	3	23	36	46	55
12' Quality	-	N	3.2	3.9	3.9
30' Quality	-	Y	4.0	4.0	4.0

NE Stations - Yellow (Near Red)

Channel	9	18	42	58
12' Quality	-	4.0	4.1	4.1
30' Quality	-	4.0	4.1	4.1

Antenna #6: UHF Yagi with 2 VHF elements

NW Stations - Red

Channel	3	23	36	46	55
12' Quality	2.8	N	2.2	3.2	3.2
30' Quality	2.9	Y*	3.3	3.4	3.4

* Near threshold

NE Stations - Yellow (Near Red)

Channel	9	18	42	58
12' Quality	4.0	3.9	4.0	4.1
30' Quality	4.0	4.0	4.0	4.1

At the first location on WBTV's parking garage a pattern was becoming evident that had shown itself in the original testing used to creating the map. Early in the original test program it became clear that an antenna's front to back ratio is the single most important characteristic in determining the antenna's ability to reject ghosts for analog and to provide a suitable signal for digital. As can be seen from the above tables, antennas with suitably large reflectors provided the best signal quality performance. Location #1 was well over 100 feet from large structures in the direction of the Northwest transmitters. A 30' antenna height would not come close to clearing the structures or hills in that direction, but a little bit of added height still helped in improving picture performance even when the path was obstructed.

4.2 Location 2

Location 2 was a residential area featuring large homes on the border of a forested area that was toward the low elevation end of terrain in the immediate neighborhood. When using the CEMA map program, it is important to take into account this kind of detail that is below the resolution of the map. (roughly 2 block pixels)

Antenna #1: Amplified Medium Multidirectional

NW Stations - Lt. Green

Channel	3	23	36	46	55
30' Quality	3.9	N*	4.0	4.0	3.9

NE Stations - Yellow

Channel	9	18	42	58
30' Quality	4.0	3.3	3.4	3.4

Antenna #2: Amplified Medium Multidirectional (Larger version, different Mfg.)

NW Stations - Lt. Green

Channel	3	23	36	46	55
30' Quality	3.4	Y	3.2	-	3.2

NE Stations - Yellow

Channel	9	18	42	58
30' Quality	3.9	2.8	-	2.8

Antenna #3: Medium directional antenna

NW Stations - **Lt. Green**

Channel	3	23	36	46	55
30' Quality	4.0	N*	4.1	4.0	4.0

* Receiver problems identified

NE Stations - **Yellow**

Channel	9	18	42	58
30' Quality	4.0	3.8	3.5	3.4

*It is at this location that it became clear that the digital receiver was not performing as well as it should have. Good quality analog reception was obtained from the medium size directional antenna and yet there was no digital picture despite good looking spectrum and signal level. It was at this site that the digital receive problem was resolved by disconnecting the computer interface. While time was not available to retest for digital reception at this location it was clear that all of the antennas tested should have produced a digital picture had the receiver been working correctly.

This site was one that the consumer may well have made a bad choice using the Antenna Selector Guide. This site was low ground in the area, but that was a subtle distinction based in the gently rolling forested terrain.

4.3 Location 3

This location, 33.5 miles from the transmitter, was in a residential area consisting of gently rolling terrain near the lower elevation limits of the rolling terrain. A large brick church was about a block behind the test location which would have put the church in a position of roughly 240° for the northwest stations and 180° for the northeast stations referenced to the signal direction. The church produced a very consistent and strong ghost lagging the main image on channels coming from the northeast group of stations.

Antenna #1: Medium directional antenna

NW Stations - Green

Channel	3	23	36	46	55
30' Quality	3.9	Y	4.0	-	4.0

NE Stations - Yellow

Channel	9	18	42	58
30' Quality	3.4	2.4	-	3.2

Antenna #2: UHF Yagi with 2 VHF elements

NW Stations - Green

Channel	3	23	36	46	55
30' Quality	4.1	Y	4.0	-	3.9

NE Stations - Yellow

Channel	9	18	42	58
30' Quality	3.4	3.4	-	2.9

VHF elements did little to help ghosts, improvement noticed on Channel 18

A Corner reflector with a bow tie feed was tried at this location, but produced no material improvement in ghosting.

4.4 Location 4

Location 4 was 40.8 miles from the digital transmitter. This location was expected to be, and was, a clean receive location from both transmitter sites. The distance from the Northwest stations required the use of a medium directional antenna with preamp.

Antenna #1: Medium size directional antenna (With Preamp)

NW Stations - **Blue**

Channel	3	23	36	46	55
30' Quality	4.0*	Y**	4.1	4.0	4.1

* Channel 3 had 60 Hz EMI

** Digital worked at ground level as well (12')

NE Stations - Yellow (Near Red)

Channel	9	18	42	58
30' Quality	4.1	4.1	4.0	4.1

Antenna #2: Unamplified multidirectional Disk

NW Stations - Blue

Channel	3	23	36	46	55
30' Quality	3.9	Y	3.8	3.7	3.7

Noisy with this antenna

NE Stations - **Yellow (Near Red)**

Channel	9	18	42	58
30' Quality	3.9	3.8	3.6	4.0

4.5 Location 5

This site in Harrisburg was a fairly clean site amidst the Northeast group of stations and 30.2 miles from the digital transmitter. The test location was a church parking lot in a residential neighborhood.

Antenna #1: Unamplified Omnidirectional Disk

NW Stations - Red

Channel	3	23	36	46	55
30' Quality	3.8	Y	3.3	4.0	3.8

A little noisy

NE Stations - **Yellow**

Channel	9	18	42	58
30' Quality	4.1	3.8	3.9	3.7

Antenna #2: Medium size directional antenna

NW Stations - **Red**

Channel	3	23	36	46	55
30' Quality	4.0	Y	4.1	4.1	4.1

NE Stations - Yellow

Channel	9	18	42	58
30' Quality	4.1	4.0	4.1	4.0

4.6 Location 6

Location 6 was a residential neighborhood 13.3 miles from the digital transmitter with large brick homes in an area that was expected to perform well for both groups of transmitters.

Antenna #1: Unamplified Omnidirectional Disk

	NW Stations - **Yellow**				
Channel	3	23	36	46	55
12' Quality	3.8	N*	3.4	3.9	3.7
30' Quality	3.9	Y	3.8	3.5	3.6

Analog channels were a little noisy at 12'
* The digital signal was in and out at 12'. With careful positioning the digital signal could be received but had a large notch in it's spectrum.

	NE Stations - **Yellow**			
Channel	9	18	42	58
12' Quality	3.4	3.0	3.9	--
30' Quality	3.9	4.0	3.0	3.9

Antenna #2: Small UHF Yagi

	NW Stations - Yellow				
Channel	3	23	36	46	55
12' Quality	-	Y*	3.9	4.0	3.9

*Despite the stations being in the same direction 70° to 80° difference in antenna pointing was noted for Channel 23 compared to the other stations. A UHF corner reflector was tested and provided the same results. This pointing anomaly disappeared quickly as the antenna was raised and disappeared by the time 30' was reached.

This turned out to be an unusual location apparently due to the higher RF reflectivity of brick structures compared to other residential construction. There seemed to be a very complex interference field at ground level that moderated very quickly as the antenna was raised on the trucks pneumatic mast. By the time the antenna reached 30' the area behaved as expected. At ground level optimum antenna pointing was frequently not in the direction of the station.

4.7 Location 7

Location 7 was only a few blocks from location 6, but now instead of being towards the high ground of the neighborhood it was towards the low ground along a river. It was 13.6 Miles from the digital transmitter.

Antenna #1: Unamplified Omnidirectional Disk

	NW Stations - Yellow				
Channel	3	23	36	46	55
30' Quality	3.9	Y*	0.7	2.7	2.9

*The digital receiver worked at 12' with careful antenna positioning.

	NE Stations - Red			
Channel	9	18	42	58
30' Quality	3.4	3.0	3.8	-

Antenna #2: Medium size Directional Antenna

	NW Stations - **Yellow**				
Channel	3	23	36	46	55
30' Quality	4.1	Y	3.7	3.9	3.8

	NE Stations - **Red**			
Channel	9	18	42	58
30' Quality	4.0	3.9	4.1	4.0

While the map for Northwest stations showed this location to be a yellow it behaved as a red. This was another area where terrain detail smaller than the map's resolution must be taken into account by the consumer. Since this was a low area compared to neighborhood terrain, a consumer following the Antenna Selector Guide instructions would select the medium directional antenna.

As expected from a terrain blocked area, performance of multidirectional antennas was marginal at this site. The exception was that digital worked at ground level (12') with positioning of the multidirectional antenna, and worked very solidly at 30' with a multidirectional antenna despite the poor analog performance. A medium sized directional antenna produced acceptable but not perfect pictures on all channels along with very solid digital performance.

4.8 Location 8

Location 8 at Lincolnton was selected as a site near the northwest group of transmitters including WBTV-DT (9 miles) that would be subject to high signal levels and would be a challenging reception location for the Northeast group of stations.

Antenna #1: Unamplified Omnidirectional Disk

NW Stations - **Yellow**

Channel	3	23	36	46	55
12' Quality	4.1	Y	4.0	3.9	4.0
30' Quality	4.1	Y	4.1	3.9	3.9

NE Stations - Blue

Channel	9	18	42	58
12' Quality	3.8	3.5	--	--
30' Quality	3.8	--	--	3.8

All NE stations were noisy with this antenna

Antenna #2: Medium size directional
(With preamp)

NE Stations - **Blue**

Channel	9	18	42	58
30' Quality	4.1	3.8	4.0	3.9

All NE stations showed some preamp overload symptoms

The appropriate antenna for the northeast group of stations required a preamplifier with which it produced very acceptable pictures. But the presence of the preamp in an area only a few miles from the other group of transmitters caused the preamp to overload which was evidenced by subtle artifacts in the pictures from the northeast group of transmitters. From this location the two groups of transmitters were roughly in line.

4.9 Location 9

Iron Station, 5.1 miles from the digital transmitter, is on the North side of the NW group of transmitters

Antenna #1: Unamplified Omnidirectional Disk

NW Stations - **Yellow**

Channel	3	23	36	46	55
12' Quality	3.9	Y	3.8	3.8	3.9
30' Quality	4.0	Y	4.0	3.6	4.0

NE Stations - Blue

Channel	9	18	42	58
12' Quality	3.7	3.7	3.0	--
30' Quality	4.0	3.8	3.0	3.7

All NE stations were noisy with this antenna

Antenna #2: Medium size directional
(With Preamp)

NE Stations - **Blue**

Channel	9	18	42	58
30' Quality	4.0	4.1	4.1	4.1

Like location 8 this site was expected to be relatively clean with the northeast group of stations in need of a preamp. The prediction was accurate and a preamp provided excellent results.

4.10 Location 10

At Paw Creek, 16 miles from the digital transmitter a quick test was performed with a multidirectional antenna in an area that was near the high end of terrain in the neighborhood. This location was predicted to be good for both transmitter locations, even though the distance to the northeast group of stations varied from 12 to 19.8 miles. This area behaved as expected.

Antenna: Unamplified Omnidirectional Disk

NW Stations - **Yellow**

Channel	3	23	36	46	55
12' Quality	3.9	Y	4.0	4.1	4.0
30' Quality	--	Y	3.9	3.8	4.1

NE Stations - **Yellow**

Channel	9	18	42	58
12' Quality	3.6	4.0	3.9	3.9
30' Quality	3.9	3.7	3.7	3.4

In this case the quality at 30' was worse than at 12'. This effect emphasizes the importance of carefully positioning a multidirectional antenna for desired channel reception before selecting a mounting location.

4.11 Location 11

Rockingham, NC started the fourth, and final, day of testing. Day four was devoted to testing from distant receive locations. The first stop was at Rockingham, NC 85.9 miles from the digital transmitter. Rockingham would prove to be one the more interesting test sites.

Antenna: Large directional antenna
(With Preamp)

NW Stations - **Violet**

Channel	3	23	36	46	55
30' Quality	3.9[1]	N[2]	4.0	3.8	3.5

All NW channels were visibly noisy and 55 was quite noisy

[1] Picture was degraded by interference. Without the interference the picture would have been ghost free and mildly noisy.

[2] A strong local translator on channel 24 was 48 dB stronger than the digital signal. A spurious audio carrier from the channel 24 translator was visible through the digital carrier on the spectrum analyzer as shown below in figure 2.

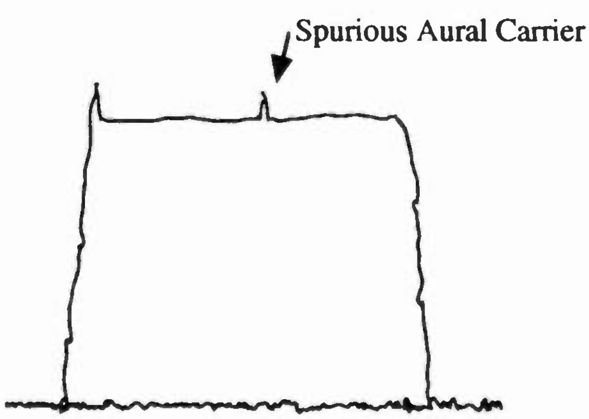


Figure 2

NE Stations - **Violet**

Channel	9	18	42	58
30' Quality	4.0	4.0	4.0	--

Rockingham is in hilly country and the test location was at an average terrain level that was short of being the highest but well above being the lowest terrain. The test was conducted on a clear sunny day in late October 1998. At this site scintillation of the UHF signals of 6 - 8 dB was noticed when the antenna was at lower (just above the truck approximately 12') elevations. As the antenna was raised the scintillation was reduced to 3 - 4 dB at 40'. Signals from the northwest transmitters were all quite acceptable except that the WBTV Channel 3 signal was seriously degraded by interference that appeared to be 60Hz related, otherwise the signal would have been free of ghosting and would be mildly noisy with subtle co-channel interference.

Our tests of the digital signal were thwarted due to a nearby Channel 24 translator that had a signal that was 48 dB above the digital signal (as viewed on the spectrum analyzer with a 300kHz resolution bandwidth, thus the actual ratio of carrier powers would be less than 48dB) and was producing a spurious aural carrier signal in channel 23 due to the intermodulation of the aural carrier and the visual carrier. (See Figure 2) The channel 23 digital signal had a clean waveform with 3 - 4 dB of scintillation and was of adequate signal level to work had the spur not been present.

This location was clearly in the violet map contour. The violet map contour transitions from the CEMA defined picture quality at the inner contour boundary to the FCC defined no signal picture quality at the outer boundary.

Absent the interference on channel 3 all of the Charlotte stations at this location had readily discernible noise but were ghost free and quite viewable.

Due to the presence of the adjacent channel translator interference to channel 23 at the Rockingham test location, a short test stop next to the highway at a range of 78 miles was made. The truck's normal channel 23 test antenna was employed with the truck's normal preamp. The antenna was positioned to put the translator in one of the antenna's nulls. This allowed reception of the digital signal at 78 miles with solid results. The test team then moved on to location 12.

4.12 Location 12

Location 12 was 72.1 miles from the digital transmitter at the parking lot of the Savanna AME Zion Church. This site was on terrain that was towards the high end of the average terrain in the area and closer to the station in order to avoid the Channel 24 interference.

This location is where the test team took a break from work to watch the high definition broadcast of the space shuttle Discovery launch with John Glenn. Due to unresolved compatibility issues between the digital receiver and WBTV's transmitter the receiver was not producing audio at the time. Audio from the totally different broadcast that appeared on WBTV's Channel 3 transmission was employed.

All channels performed well and the digital signal was solid throughout. The only draw back to this location was a small amount of visible co-channel interference on Channel 3 from a distant Channel 3 signal.

Antenna: Large Directional Antenna
(With Preamp)

NW Stations - **Violet**

Channel	3	23	36	46	55
30' Quality	4.0	Y	4.0	3.9	3.9

NE Stations - Blue

Channel	9	18	42	58
30' Quality	4.1	4.1	4.1	4.0

4.13 Location 13

Location 13 was at Wadesboro, NC, 69.3 miles from the digital transmitter. Wadesboro was a little closer to the digital transmitter, but was selected to intentionally be a site that's closer to the low ground in the immediate area.

Antenna: Large Directional Antenna
(With Preamp)

NW Stations - **Violet**

Channel	3	23	36	46	55
30' Quality	4.0	Y	4.0	4.0	4.0

NE Stations - **Violet**

Channel	9	18	42	58
30' Quality	4.1	4.0	4.1	4.1

Even before the large directional antenna with a preamp was raised, WBTV's Channel 23 digital signal was receivable and quite steady at an antenna height of 12'. The signal improved and showed greater stability as the antenna elevation increased. All channels produced good results except for Channel 3 which suffered from quite a lot of 60Hz related interference. Without the interference the picture would have been quite acceptable. The rating above reflects its ghosting only.

At this distance from Charlotte the test van was within range of other TV markets and took the following readings on other stations in the area.

Antenna: Large Directional Antenna
(With Preamp, aimed for each channel)

Channel	13	15	21	31
30' Quality	4.1	4.1	4.0	4.0

The large directional antenna provided very good pictures from these other cities.

4.14 Location 14

Location 14 was intentionally selected to be a troublesome location. It was 22.2 miles from the digital transmitter and closer to the Northeast group of transmitters. It is in the shadow of downtown when looking toward the Northwest transmitters. While not really in a downtown building canyon the site was in the fringe of downtown with three and four story buildings in the immediate area. The taller buildings of downtown were between the test site and the digital signal. It was expected that this would be an exceptionally bad reception area. Reception with the multidirectional antennas was not even attempted at this location. Testing began with the best medium sized directional antenna. (the antenna that has consistently delivered the best ghost reducing capabilities)

Antenna #1: Medium Size Directional

NW Stations - **Blue**

Channel	3	23	36	46	55
12' Quality	3.0	Y	2.0	1.8	2.3
30' Quality	3.0	Y	0.4	3.2	2.3

Signals were noisy, a preamp was called for here and should have been used.

NE Stations - **Red**				
Channel	9	18	42	58
12' Quality	3.7	1.2	2.8	3.9
30' Quality	3.8	2.8	2.8	3.9

Antenna #2: Large UHF Yagi

NW Stations - Blue					
Channel	3	23	36	46	55
12' Quality	--	Y	2.1	1.8	1.8
30' Quality	--	Y	1.7	3.2	1.9

NE Stations - Red				
Channel	9	18	42	58
12' Quality	--	2.4	2.9	3.6
30' Quality	--	3.8	4.0	3.8

As can be seen from above, even with good front to back ratios and good directivity, two good antennas provided very poor analog performance due to the high multipath environment, and yet the digital worked in all cases. The digital signal performed as well with the antennas at 12' as it did at 30'. This site is one of the most powerful examples of the real value of digital. By using antennas with a good front to back ratio the digital signal worked in an environment where there was no real analog solution for the northwest group of transmitters when shadowed by downtown.

5.0 CONCLUSION

This series of tests was designed to answer a simple question related to the CEMA Antenna Mapping Project, namely, does the map work for digital. The answer is: yes. The R-5 committee's assumption that the digital link is more robust than the analog link appears to be true based on the testing done in Charlotte.

Once our digital receive problems were resolved digital worked everywhere that an acceptable analog picture was received, as well as locations where an acceptable analog picture was not received. Charlotte provided the opportunity to test where there was a full height full power digital signal co-located with other analog VHF and UHF stations where a valid comparison could be made between analog and digital transmission.

It is clear that if an antenna produces satisfactory analog reception, it will also produce satisfactory digital reception for transmitters that are co-located.

Throughout the week of testing the test team had the opportunity of discussing the findings with the WBTV staff. Both the WBTV staff and the CEMA test team came to the same conclusion; namely that important factors in digital television reception were:

1. Placing the antenna outside is better.
2. Good antenna front to back ratio
3. Correct antenna aiming

According to the WBTV staff it did not seem to be terribly important whether the antenna was particularly high so long as it had a good front to back ratio and was pointed approximately in the right direction.

Thus it appears that some of the previous studies that have raised concerns over the effectiveness of indoor antennas, and nondirectional antennas in general, for digital have been flawed in that they did not use transmitting plants that were directly comparable with analog or that were even full service facilities.

In many television markets all the digital stations will be UHF. This will allow for small, aesthetically acceptable, antennas to be employed. This will allow a higher percentage of the TV marketplace to be served by antennas, allowing broadcasters to retain more control over their signal distribution. Thus it appears that the migration to digital transmission and digital TV will be easier and have a far brighter future than most people presently believe.

The CEMA TV Antenna Selector Map project will help simplify the consumer antenna selection process in the retail environment thus promoting a speedier transition to digital television broadcasting.

Comparing Panel and Slot Antennas in Shared Digital and Analog Service.

By Robin Blair and Tony Magris,
Radio Frequency Systems Ltd.,
Melbourne, Australia.

Abstract

During the forthcoming period of simultaneous transmission of DTV and NTSC services, many operators will find it convenient and economic to transmit both services from the one antenna if they are allocated frequencies in the same band. At UHF, broadband panel arrays are eminently suitable for this application. However, if the DTV and NTSC channels occupy adjacent or near adjacent channels, they may permit the use of traditional slot arrays, even though the bandwidth of these antennas is quite limited. We show in this paper, however, that the radiation pattern instability of these antennas across more than one channel can lead to considerable deficiencies of coverage in the service area. In this respect, the paper analyses the differences between slot and panel arrays and shows that the panel arrays are generally to be preferred.

Introduction

As the terrestrial television service evolves from analog to digital, existing transmitter operators find themselves faced with carrying dual transmissions of their existing NTSC and new digital service until at least the year 2006. This has major implications for the capability of existing plant, but those with a DTV frequency allocation in the same band as their NTSC channel have the advantage that it should be possible to carry both services in the same transmitting antenna. Wideband panel arrays are eminently suitable for these applications, but if the services are on UHF, and sufficiently close in frequency, they might be accommodated in a traditional high gain slot array in some circumstances.

Slot arrays are known to have very narrow impedance bandwidths, inasmuch as an acceptable impedance match to the feeder cables can be achieved over only one or two TV channels. Hence they would be useable only where the DTV and NTSC allocations specify adjacent channels, or sometimes, separation by one vacant channel.

This limitation on slot arrays arises largely because of the nature of the feed system, which is restricted by the small mechanical cross section of the antenna to having the radiating elements connected in sequence or series. By contrast, panel arrays with their greater mechanical cross sections accommodating complex sets of cables can exploit parallel feeding and associated impedance compensation techniques. It seems not often realized that the two different feed techniques also have a major influence on the susceptibility of the radiation pattern to vary with frequency, as we shall now show.

Antenna Feed Systems

There are three methods of arranging the power distribution to the large number of levels of radiating elements found in a high gain arrays. These are shown diagrammatically in Fig.1 and comprise:

(a) End feeding - whereby the feeder enters the bottom of the antenna and power to the radiators is tapped off the cable as it ascends the antenna column:

(b) Center feeding - here the main cable enters the array at its mid-height and splits into one descending and one ascending feeder within the column from which power is tapped off to the radiators in sequence as before.

(c) Parallel feeding - whereby the feed cables split regularly and form a fan-like arrangement spreading over the height of the column. It is readily seen that all the radiating elements are thus fed in parallel, or more importantly, the electrical path lengths to the radiators from the common feeder are all essentially equal.

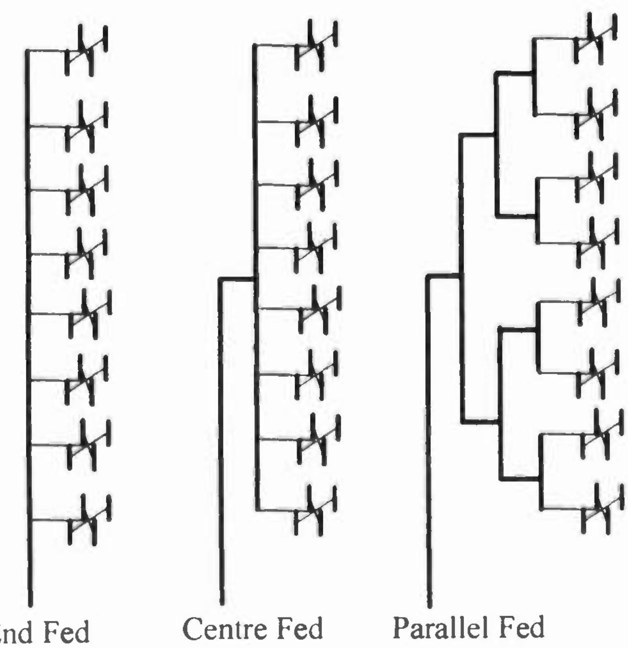


Fig.1 - Antenna Array Feed Methods

In end and center feeding, the taps off the feed lines are arranged to be at intervals of one wavelength at the design frequency so that all the radiators are correctly phased at this frequency.

As mentioned above end and center feeding are necessarily used in slot arrays to minimize mechanical size, whereas parallel feeding is almost always exploited in broadband panel arrays.

Mechanically, the arrangement of the feed systems in panel arrays is relatively straightforward, as the space allows the use of semi-flexible coaxial feeders formed into cable harnesses. In slot arrays, the slotted external cylinder commonly forms the outer of a transmission line, with the slots being inductively coupled to an internal tubular conductor located coaxially within the outer cylinder and forming the inner of the line. For center feeding, there may be a third conductor inside the lower half of the inner conductor, forming a triaxial line connecting to the common feed point at the center of the antenna.

Vertical Radiation Patterns

In looking at Fig.1 it is intuitively obvious that the relative phases to the radiating elements will vary with frequency more in the end fed arrangement than in the center fed, and this in turn will vary more than in the parallel feed system. In the end fed array the shortest and longest cables from the common feed point differ in length by an amount equal to the length of the antenna column. In high gain antennas this may be equal to some 25 to 30 wavelengths. Obviously, the phase difference over this length will change markedly with frequency, and cause a frequency dependent variation of the vertical pattern.

By contrast, in the parallel array, the lengths of all the cables from the common feed point are all essentially equal and no relative phase changes are to be expected as the frequency is varied. The center fed array has a characteristics somewhere between the two.

Fig.2 shows the extent that these phase changes can cause the vertical pattern to vary with frequency..

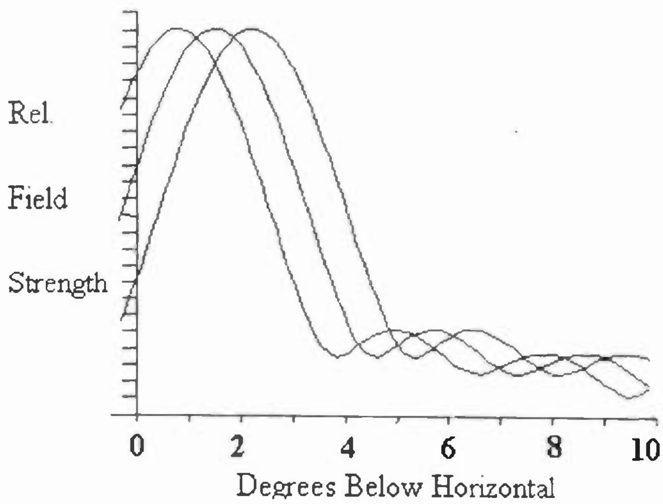


Fig.2 (a) VP Variation - End Fed Array

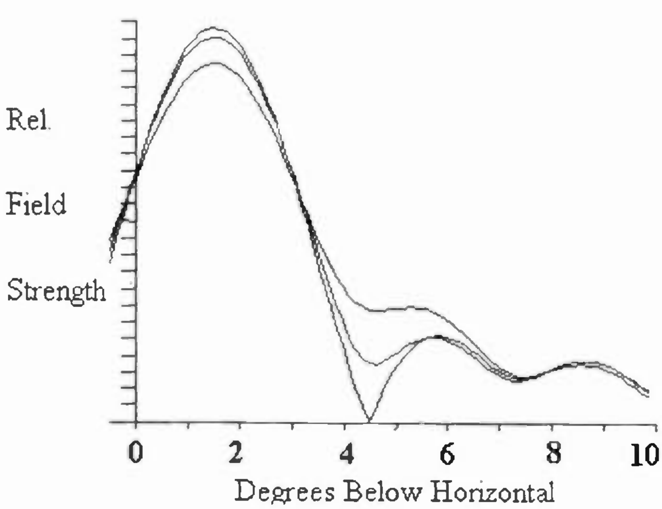


Fig. 2(b) VP Variation - Centre Fed Array

Fig.2(a) is for the end fed case, 2(b) for the center fed case, and 2(c) for the parallel fed array. In each case the design frequency of the antenna is 600MHz and the plots are taken at this frequency and +/- 6MHz. The antenna mid band gain is 15dB; typical of many applications

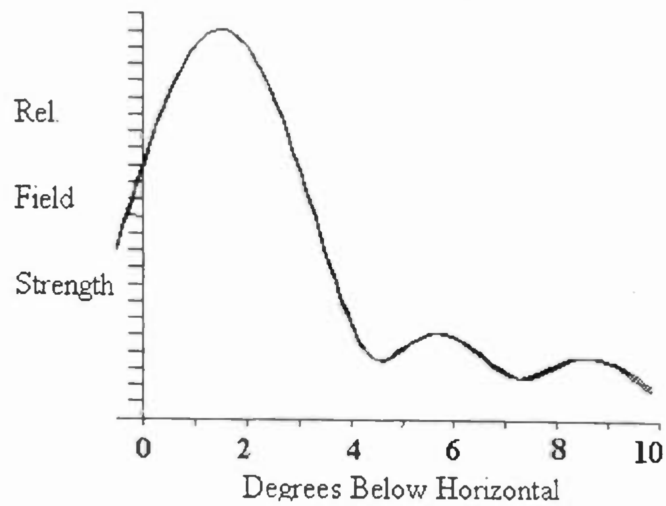


Fig.2(c) VP Variation - Parallel Fed Array

Notice that the variations are quite large for the end fed and center fed arrays, but essentially negligible for the parallel fed array. The reasons are easy to see from the cabling diagrams of Fig.1. Notice that, when the frequency changes there will be a progressive phase change over the aperture of the end fed array, and this will cause a tilting up or down of the main beam. In the center fed array, the top half will tend to tilt in one direction, and the bottom half in the opposite direction. Hence, there is no nett beam tilt, but what we do see is a type of splitting of the main beam. The principle effect of this is to cause distortion around the first null, even to the extent of destroying the null fill as seen here.

The conclusion must be that the elevation patterns for both end fed and center fed slot arrays are quite unstable with frequency, whilst that of broadband panel arrays are relatively quite stable. It follows that the latter are much more suitable for multi-channel operation, even with adjacent or closely spaced channels.

Horizontal Radiation Patterns

Whilst the relatively wide cross section of the panel array does allow the use of a comparatively stable feed system, it does bring with it the disadvantage that the horizontal radiation pattern is less smooth than that of comparable slot arrays, caused by the greater mechanical cross section. Consider a test receiver traversing a distant circular path around the array. The phases of the fields arising from the individual radiators change according to the changes in distance from each radiator to the test receiver. Of course, if all radiators were at the center of the test path circle, there would be no change in distance to each, and no variation in the signal received. As the array cross section increases, however, the radiators are moved further from the ideal center, and the variation in signal level increases. The comparative effect for typical slot and panel arrays is shown in Fig.3.

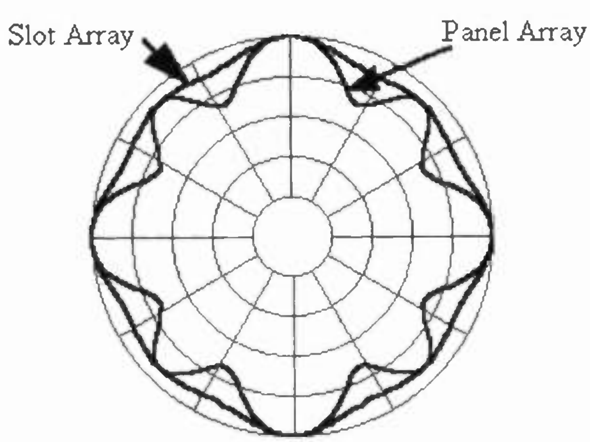


Fig. 3 Comparative Horizontal Patterns

The physical electrical distances here are small when compared to those over the vertical extent of the antennas as previously considered. Hence, variations with frequency are quite small for both slot and panel arrays and are not a factor in choosing between them.

Since the FCC has mandated the maximum signal strength that a digital transmission may radiate towards the horizon, it is often argued that the comparatively deeper dips in the panel array pattern represents a lowering of the average signal level, and a consequent loss of viewers. We shall show that it is not that simple.

Performance Summary

We can now make two statements about the comparative radiating characteristics of slot and broadband panel arrays:

(a) Slot arrays exhibit marked variation with frequency of the field strengths laid down in all parts of the service area, whereas the panel array performance is very stable.

(b) Dips in the horizontal patterns of the panel arrays are deeper than those exhibited by slot arrays of comparable gain.

The first of these characteristics is undesirable if one is attempting to replicate the coverage of existing services, or to transmit two comparable services into the same service area. The second may represent a loss of signal strength to distant customers. So which should be chosen? To look at that we must take into account the statistics of propagation.

Propagation Effects

A simple examination of the radiation patterns in Fig.2 shows that slot arrays exhibit their major variations at fairly steep angles below the horizon. Hence their effects are likely to most effect the near-in viewers living relatively close to the transmitter. On the other hand, the panel array may give less than the mandated maximum signal on the horizon, as shown in Fig.3.

In terms solely of their radiation performance then, the choice between slot and panel arrays comes down to a choice of whether it is better to sacrifice some signal strength in the near-in service area or far out near the radio horizon. Recall that this is not simply answered by assuming the near-in field strength is much higher. It is not, by virtue of the shape of the vertical radiation pattern giving progressively lower signal levels as the angle below the horizon increases. Good antenna design attempts to replicate an inverse cosine shape which, it can be shown, ideally gives the same field strength on the ground at all distances up to the horizon.

Rather the question can be answered by looking at the statistics of propagation. To this end, and keeping in mind that DTV picture is either perfect or non-existent, let us assume that a customers service passes from satisfactory to unsatisfactory if the percentage of time for which he can receive a picture falls from 99% to 90%. These values are not rigorous, but realistic, and do serve to quantify the argument.

It is well known and only to be expected that the time variability of the received signal increases with distance from the transmitter. Fig.4 taken from standard propagation curves[1] illustrates this. It shows the difference in dB between the curves for 50% and 10% of the time [i.e. F(50,50)-F(50,10)], plotted against distance for a 1000 foot antenna height above average terrain. This difference is seen to increase strongly with distance. For lower antenna heights the general shape is the same, but the increase occurs at shorter distances.

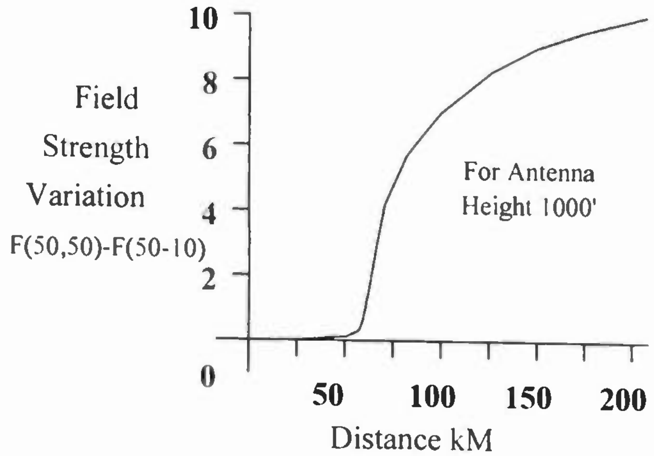


Fig.4 - Variability of Field Strength v's Distance

The other variable we need to consider is the percentage of locations which receive a satisfactory service at a given distance. Recent field tests[2] have confirmed that variations in field strength due to terrain irregularities, buildings trees and the like, tend to show a log-normal distribution.

Putting these two factors together, we can draw a type of operating characteristic for some fixed distance, as illustrated in Fig.5. Here we plot the percentage of locations receiving a particular signal level for various percentages of time, with the curves anchored by an applicable F(50,50) value. The three curves apply for 90%, 50% and 10% of the time respectively. Fig.5 shows the general shape of such curves. If an operator wishes to ensure that say 95% of locations receive a satisfactory service for say 90% of the time (point A), he must design his transmitter system to lay down the F(50,50) field strength given by point B. These curves will be more or less widely spaced horizontally depending on the time variance of the received signal. As we saw from Fig.4, this variance increases greatly, as therefore does the spread of the curves, as the distance from the transmitter increases.

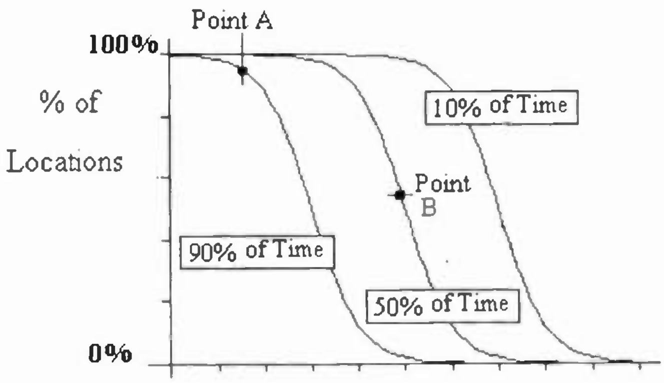


Fig. 5 - Service Availability Operating Characteristic

Now consider Figs.6(a) and (b), which show 90 and 99% of time curves for near-in and distant parts of the service areas respectively.

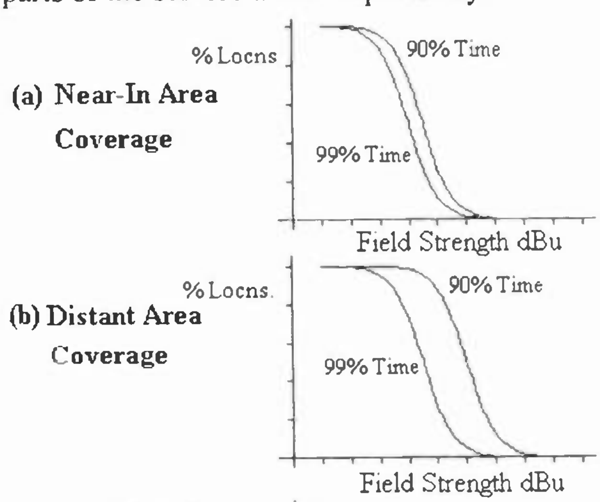


Fig. 6 - Near and Far Coverage

It is readily evident that at any given percentage of locations the fall in field strength to take the service from satisfactory to unsatisfactory (i.e. from 99% to 90% of time) is much smaller near-in than far out. Similarly, if two channels are being transmitted and the signals level differ, the situation of one being received satisfactorily and the other not is more likely to occur near-in in the service area than far out.

It follows from all this that, if in choosing between antennas some field strength must be sacrificed either close in or far out, it is better to choose the latter. Hence, judging by the discussion of the comparative radiation pattern performance of slot and panel arrays earlier in this paper, panel arrays are evidently the better choice.

There is a more heuristic argument to make this same point. Let us assume that viewers generally equip themselves to obtain good reception at he signal level available for 99% of the time in their locality. If the transmitted signal level should then fall by some small amount, the proportion of time for which near-in viewers will then get that necessary field strength may fall from say 99% to 90% of the time. For the far-viewer, however, it may still exist for 97 or 98% of the time, and he will have much less cause for complaint.

Conclusions

There is no perfect transmitting antenna. We have shown in this paper that traditional slot arrays exhibit fairly severe variations of signal strength with frequency, which mitigates against their use for multi-channel or combined NTSC and DTV working. The problem is particularly severe in the near-in parts of the service area. Broadband panel arrays have excellent characteristics in this respect, but their horizontal radiation patterns are less smooth than those of slot arrays, and the FCC mandated signal strength may not be realized at the outer edges of the service area. However, when the statistical exigencies of propagation are taken into account, it is found that the loss of service due to this cause is much less severe than would occur near-in with the slot arrays. Accordingly, the panel array should almost always be the antenna of choice for multi-channel and NTSC/DTV operation.

References:

1. "Propagation in non-Ionised Media", Recommendations and Reports of the CCIR, XVIth Plenary Assembly, Volume V, Dubrovnic, 1986, Rec. 370-5.

2. Chris Weck, "VALIDATE Field Trials of Digital Terrestrial Television", NAB 1998 Broadcast Engineering Conference Proceedings, pp.162-169.

DTV TABOO CHANNEL INTERFERENCE INTO NTSC AT HIGH POWER LEVELS

Stanley J. Salamon, Charles W. Rhodes, and Charles W. Einolf, Jr.
Advanced Television Technology Center
Alexandria, Virginia

ABSTRACT

The Advanced Television Technology Center (ATTC) has recently completed an investigation of the interference that may arise from Digital Television (DTV) signals on the NTSC taboo channels. This study supplements the results that were obtained during the testing of the Grand Alliance system. In particular, the study concentrates on the interference effects that occur under strong signal conditions. The ATTC RF Test Bed was modified to produce the necessary power. The results of the tests show that interference may occur at significantly lower power conditions than what is predicted by the FCC Planning Factors. This paper discusses the results of this study and provides DTV system planners with design parameters to avoid interference conditions.

1. Introduction

For the past three decades, the FCC has been collecting data relative to the interference rejection characteristics of NTSC receivers under different taboo channel relationships and signal level conditions. The FCC, however, did not collect information relating to the front-end overload characteristics of these receivers. The original FCC NTSC table of allocation was designed to ensure that these receivers were not exposed to extremely high undesired signal levels.

Digital television (DTV) tests of the original proponents and the Grand Alliance did not measure the Threshold of Visibility (TOV) for the case of Strong Desired NTSC in the presence of Taboo Interference since the ATTC RF Test Bed was not designed to output such high undesired powers[1]. It was believed by the Advanced Television Systems Committee (ATSC) that DTV signals would be broadcast at Effective Radiated Powers 12 dB lower than those used for NTSC. Consequently, the ATTC RF Test Bed was constructed to deliver these power levels. As a result, the DTV power levels were not sufficient to cause interference on all UHF taboo channels from DTV into NTSC on more than 50 percent of the NTSC receivers (the median of 24 receivers) under the "strong NTSC power level" of –15 dBm. Data was obtained at the "weak" (-55 dBm) and "moderate" (-35 dBm) power levels. The FCC Planning Factors, in the Sixth Report and Order[2], are based on the ATTC results of taboo tests performed with weak (-55 dBm) NTSC power levels.

It is useful to understand the mechanism of taboo channel interference. A functional block diagram of a typical NTSC television receiver tuner section is shown in **Figure 1**. The antenna is coupled to a tuned circuit, which is nominally tuned to the Desired frequency. A Radio Frequency (RF) amplifier follows where a DC voltage from the Delayed RF AGC circuit controls the gain. A double tuned circuit between the RF amplifier and the signal input of the mixer is also nominally tuned to the Desired frequency. The Local Oscillator operates at a frequency 45.75 MHz above the visual carrier frequency of the Desired channel. The IF output from the mixer is filtered by a Surface Acoustic Wave (SAW) filter whose mid-frequency is about 44 MHz. The signal is then passed on to an IF Amplifier and Video Demodulator.

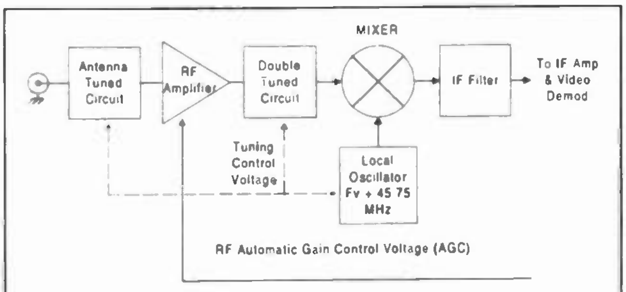


Figure 1 – UHF Tuner typical of NTSC receivers. The RF Amplifier, mixer, and local oscillator are key portions of the circuit for introducing Taboo Channel Interference.

Taboo Channel interference results when Undesired signals reach the mixer in **Figure 1**. It is important to note that the tuned circuits before the signal input to the mixer are of a low Q. Thus, these tuned circuits provide very little attenuation for channels near the Desired channel. In fact, there can be appreciable power, from the non-linear taboos within the first N+/- 8 channels, which can reach the mixer and overload it, causing intermodulation. The N+4 taboo is unique because it generates two interference mechanisms, the 2nd harmonic may beat with the second harmonic of the local oscillator as well as intermodulation. N+14 and N+15 are linear taboos which affect the aural and visual carriers, respectively, since they represent a signal coinciding with the image response of the tuner. Non-linear interference results when the Undesired signal levels exceed the linear dynamic range of the mixer or RF amplifier. Furthermore, interference from UHF Taboos cannot be removed by the IF Filter, as the interference is now within the IF Bandpass and hence able to reach the video demodulator.

These tests used the same procedures and test materials as the original Grand Alliance tests except for a minor modification of the RF Test Bed to permit the introduction of higher power levels. The tests included the verification of the test procedure repeatability compared with previous data; collection of the supplemental data at higher power levels; and the evaluation of cumulative interference from DTV signals on multiple taboo channels.

2. Test Equipment Setups

The test procedure, in general, follows the Grand Alliance System Test Procedures, SSWP2-1306, for UHF Taboo Channel Interference[3]. The Desired NTSC channel is observed while DTV power is increased on one or more of the Undesired Taboo Channels. The power of each signal is recorded at the point when the impairment is just perceptible. A range of Taboo Channels and Desired power levels was investigated.

2.1. Single Interferor Test Setup

The test was conducted using Channel 23 (524-530 MHz) as the Desired NTSC Channel (see **Figure 2**). The NTSC Visual to Aural Carrier power ratio was 13 dB. The NTSC program material was M14, "Texas Sign Dude," which has been used in previous tests of interference into NTSC.

The Undesired DTV Channel was set to one of the Taboo Channels. The DTV program material was the PN23 sequence. The PN23 sequence is a 23-bit pseudo-random number that is generated by a well-defined

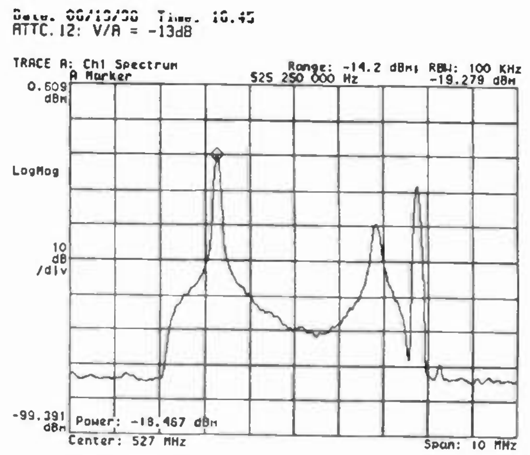


Figure 2 – The spectral response of the Desired NTSC signal at Strong peak power of -15 dBm used to evaluate DTV interference into NTSC.

polynomial. In this case, it is simply used to fill the DTV channel with data. A Channel 23 Band-Stop Filter was patched into the Undesired DTV signal path to ensure the absence of spurious signals within the Desired Channel.

The taboo channels tested are as follows:

Taboo	Mechanism	Channel
n-8	IF Beat	15
n-3	Intermod	20
n-2	Intermod	21
n+2	Intermod	25
n+3	Intermod	26
n+4	Intermod & Half IF	27
n+8	IF Beat	31
n+14	Aural Image	37
n+15	Visual Image	38

An additional amplifier was added to the output of the RF Test Bed (see **Figure 3**) to ensure that the Undesired Signal was strong enough to cause impairment to all receivers when the Desired NTSC peak power was -15 dBm. The gain was restricted to keep the sideband splatter below 45 dB down from the average signal power. This modification provided sufficient undesired power to perform all of the tests.

The power was measured in the viewing room using an HP89441 Vector Signal Analyzer. The HP VSA was controlled by a software program written by Zenith Electronics Corp. The power delivered to each of 24 receivers varied by no more than ± 0.3 dBm among

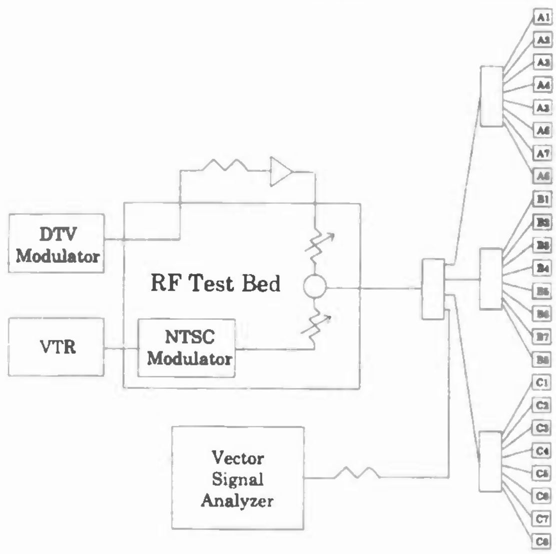


Figure 3 – Block Diagram of Test Setup showing the addition of an amplifier to the RF Test Bed in order to produce higher powers which will simulate the DTV signal for distribution to 24 receivers.

them. For -15 dBm desired power level, the power was measured at the 75 ohm input to one of the receivers. For all other desired power levels, the power was measured at the 4th tap of the 4-way splitter and attenuated by 13 dB to make the reading equivalent to the power at the input to the receiver within 0.2 dBm. **Figure 4** illustrates the typical DTV Undesired Signal.

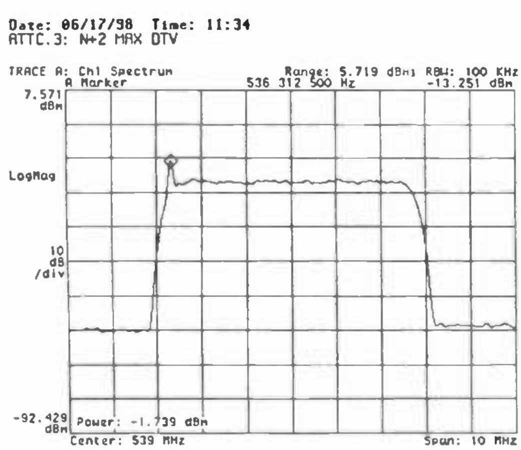


Figure 4 – The spectral response of a typical DTV Undesired Signal at maximum power on a Taboo Channel. This signal is used to evaluate DTV interference into NTSC.

2.2. Multiple Interferor Test Setup

In order to test the effect of multiple interferors, the RF Test Bed was modified to include a second Undesired DTV signal on another UHF channel. This modification is depicted in **Figure 5**. The DTV signal at IF frequency was taken from an unused tap of a splitter. A time delay (20.47 μs), much greater than the symbol period (93 ns), was introduced to ensure that the Undesired signals are not correlated. The upconversion was done externally to the RF Test Bed. This second Undesired path was made similar to the primary Undesired path. These two Undesired paths were then combined before reintroduction to the normal path through the RF Test Bed. The relative powers were carefully adjusted so that the output powers were always equal to each other.

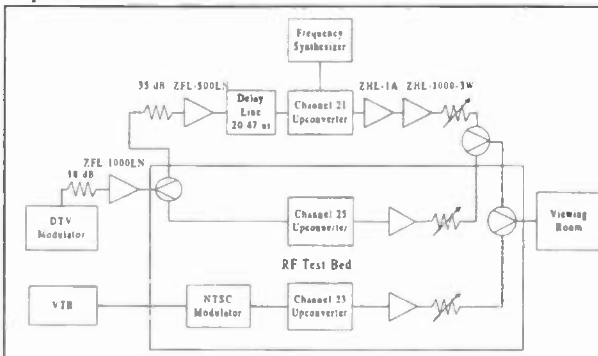


Figure 5 – Block diagram of the Multiple Interferor Test Setup showing the injection of an additional channel into the ATTC RF Test Bed.

3. TOV Test Procedures

3.1. Single Interferor TOV Test Procedure

The Desired NTSC signal was set to Channel 23. A panel of three expert observers viewing a bank of 24 NTSC receivers determined the TOV levels. First the unimpaired video was viewed on all sets. Then the undesired DTV signal level was increased until the impairment was just perceptible on the first TV receiver. The undesired signal was toggled on and off with a 2 second period to enhance perceptibility of the impairments caused by the interferor. This technique has been used previously during Grand Alliance testing. Then the undesired signal level was increased in 1 dB increments until all the receivers showed the impairment or maximum power was reached. The power level at which TOV occurred was noted for each receiver. Any special observations were also noted.

Once the TOV values were known, the TOA check was performed. Each receiver was checked individually and the volume of the other receivers was set to zero. The audio material was selections from a classical music CD. First, the unimpaired audio was heard and the volume set to a comfortable listening level. Then the undesired signal was introduced at the TOV level for that receiver. If significant impairments were heard at this level, then TOA would have been found.

3.2. Multiple Interferor TOV Test Procedure

The test procedure was the same as for the other Taboo tests, except that the additional attenuator and frequency synthesizer needed to be operated manually. When measuring DTV power, only one signal was measured. Since both DTV signals have equal power, the total power is obtained by adding 3 dB to the power of the single signal. **Figure 6** is a plot of a Strong NTSC Signal flanked by N-2 and N+2 Taboos at maximum power.

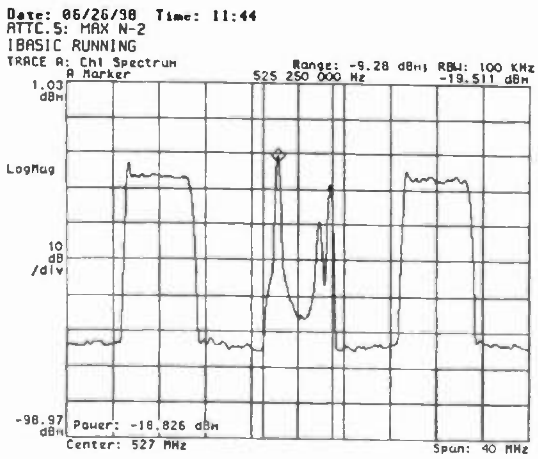


Figure 6 – Spectral response showing N-2 and N+2 Taboo Interference into and NTSC channel (N) used to test multiple Taboo Interference.

4. Taboo Channel Interference Test Results

4.1 Verification of Test Procedure

A comparison was made between the data collected for this report and the data reported from the Grand Alliance tests to verify that the RF Test Bed and NTSC test receivers were operating properly. This comparison was made for the taboo channel N+2 at the weak and moderate power levels. The Desired (NTSC) to Undesired (DTV) ratio at TOV for the median receiver at weak and moderate power levels was determined to be -28.40 dB and -17.63 dB, respectively, compared with -27.93 dB and -17.46 dB for the Grand Alliance tests. These results demonstrated that the ATTC Test Facility continues to operate within acceptable limits.

4.2 Single Interferor Test Results

The Threshold of Visibility (TOV) of DTV interference on several taboo channels was measured over a range of NTSC power levels with an emphasis on the strong levels. **Table 1** shows the results of these tests. The results are presented in terms of Desired (NTSC)-to-Undesired (DTV) Power Ratio at the Median TOV as a function of Desired NTSC Power Level and Taboo Channel.

A receiver-by-receiver account of voting by three expert observers was collected. This data shows a distribution of threshold levels by receiver. The median TOV, by definition, is a state where 50 percent of the receivers show visible impairment. However, it should be noted that some of the receivers had reached the point of unusability due to front-end overload at the median TOV level.

Table 1. Desired-to-Undesired Power Ratio at Median TOV (dB) shown for new data which supplements the original Grand Alliance test results. The emphasis is on strong Desired (NTSC) power levels.

Taboo Channel	Desired NTSC Power Level (dBm)				
	-55	-45	-35	-25	-15
N – 8					-6.90
N – 3					-1.73
N – 2		-13.33		-7.57	-1.43
N + 2	-28.40	-19.40	-17.63	-11.63	-3.80
N + 3					-5.55
N + 4					-5.60
N + 8			-22.97		-9.77
N + 14					-8.40
N + 15					1.28

4.3 Multiple Interferor Test Results

The situation where multiple DTV channels may cause interference was tested. The two worst case Taboos for the case of Strong Desired power level (- 15 dBm) were N+2 and N-2. These were the two taboos used for the multiple interferor test. The undesired power for each Taboo signal at the median TOV was measured to be -16.01 dBm. Therefore, the combined power is -13.01

dBm. The median undesired power at TOV for N-2 alone was -13.57 dBm and for N+2 alone was -11.20 dBm. The combined effect of multiple interferors causes NTSC receivers to fail at a lower Undesired power than they would have individually.

4.4 Audio Impairment Test Results

An audio impairment check on Taboo Channels N+2 and N+14 showed that the video fails before the audio. No more than 2 receivers showed evidence of audio impairment at the video TOV for those receivers. This observation is consistent with results obtained during Grand Alliance testing. It is clear that the audio is not a dominant interference mechanism.

5. Analysis of Test Results

Data gathered in this experiment, coupled with previous Grand Alliance data, provides a more complete picture of the effect of DTV taboo channel interference into existing NTSC receivers as well as a better understanding of the selectivity of NTSC receivers and their susceptibility to interference.

5.1 Taboo Channel Interference for Strong Power Conditions

Figure 7 illustrates a plot of the Undesired (DTV) power as a function of Desired (NTSC) power for TOV interference from the taboo channels N+2 and N-2. The plot demonstrates the operation of a typical NTSC receiver. The curve is flat between -55 and -45 dBm. There is a steep slope between -45 and -35 dBm. This is the region where most of the tuner AGC action takes place. Beyond -35 dBm the curve seems to approach an asymptotic limit. This is the region where the front end of the receiver is experiencing an overload condition as a result of the high input powers, even though the RF gain is reduced. For comparison, the FCC Planning Factors, which are based on the measurements at –55 dBm, are also included. The plot highlights the difference between the measured values and the FCC Planning Factors.

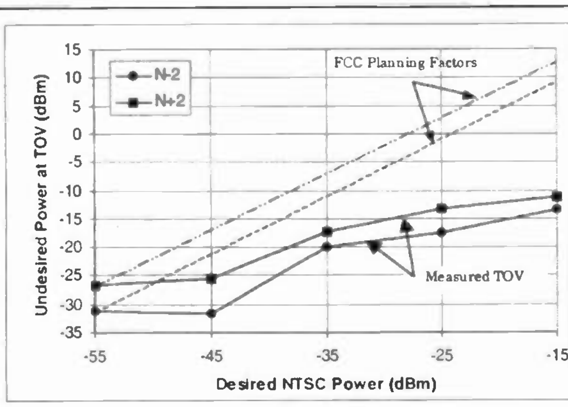


Figure 7 – Undesired Power measured at TOV as a function of Desired Power showing increased susceptibility at high NTSC powers. FCC Planning Factors, based on the measurements at – 55 dBm (shown by the dotted lines), underestimate the effects of Taboo Interference.

5.2 Selectivity of NTSC Receivers

Table 2 is a composite of original Grand Alliance test data and the supplemental data reported here. It should be noted that the FCC has adopted, as its DTV Planning Factors, the values for the case where the Desired NTSC Power is – 55 dBm.

Table 2. Composite data for various Desired NTSC Power Levels from the original Grand Alliance tests (shown with an *asterisk*) and the new supplement data (shown in **BOLD**) for the Desired (NTSC)-to-Undesired (DTV) Power Ratio in dB at the Median Threshold of Visibility (TOV)

Taboo Channel	Desired NTSC Power Level (dBm)				
	-55	-45	-35	-25	-15
N – 8	*-31.62**		*-16.11**		**-6.90**
N – 3	*-29.73**		*-18.28**		**-1.73**
N – 2	*-23.73**	**-13.33**	*-15.00**	**-7.57**	**-1.43**
N + 2	**-28.40**	**-19.40**	**-17.63**	**-11.63**	**-3.80**
N + 3	*-34.13**		*-19.79**		**-5.55**
N + 4	*-24.96**		*-18.21**		**-5.60**
N + 8	*-43.22**		**-22.97**		**-9.77**
N + 14	*-29.55**		*-22.24**		**-8.40**
N + 15	*-17.58**		*-14.53**		**1.28**

The Undesired DTV Power at TOV as a function of Taboo Channel describes the characteristic of the NTSC receiver RF selectivity and does not reflect IF selectivity (see **Figure 8**). The selectivity curve is a varying function that shows more susceptibility to interference as the taboo channel approaches the Desired Channel N. Taboos N-8 to N+8 are all due to inter-modulation and /or cross-modulation products being generated in the tuner as a result of exceeding the dynamic range of the mixer. There are a few exceptions: N+4, N+14, N+15. These cases have a different interference mechanism than the rest of the taboos. N+4 is a special case in which the second harmonic of the local oscillator beats with the second harmonic of the Undesired signal on channel N+4. The beat frequency falls within the 44 MHz IF of the NTSC receiver. This mechanism serves to lower the tolerance to the Undesired DTV power and lowers the TOV at N+4 below the trend of the selectivity curve. N+14 and N+15 are linear mechanisms that produce a sound and picture image, respectively.

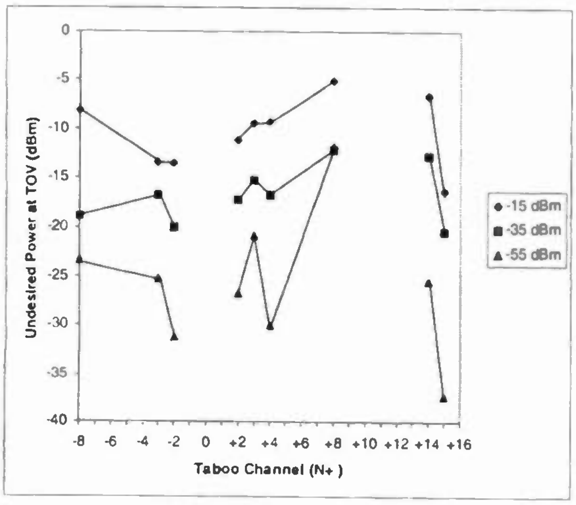


Figure 8 – Undesired Power at TOV measured as a function of Taboo Channel showing the selectivity of NTSC receivers and susceptibility to interference. Sensitivity to interference from signals on Taboo Channels N-8 to N+8 varies because of the RF selectivity. N+4 is a special case as explained in the text. N+14 and N+15, unlike the other Taboo Channels, are linear interference mechanisms.

6. Conclusion

The Planning Factors that the FCC have used in the 6^{th} Report and Order are based on the Taboo Interference results at the Weak (-55 dBm) Desired NTSC power level. Recent tests suggest that the susceptibility of NTSC television receivers extrapolated from Weak Desired NTSC power levels is not representative of the susceptibility of NTSC television receivers at stronger Desired NTSC power levels. Furthermore, the combined effect of multiple taboo interferors, not considered in the FCC model, would only exacerbate the problem.

The results of this study illustrate that the total power reaching a receiver is critical. Extremely strong Undesired DTV signals on Taboo channels would cause unexpected interference to NTSC reception even though the D/U ratios are within FCC permitted values. For cases where strong Undesired signal level conditions are predicted, the planning of a DTV station should consider these new test results. Specifically, the planning should consider a maximum Undesired DTV power level as a function of the expected Desired NTSC power level.

7. Acknowledgements

The ATTC thanks Dennis Wallace for his evaluation of the ATTC RF Test Bed and the test setup.

References

[1] *digital* HDTV Grand Alliance System, Record of Test Results, Federal Communications Commission, Advisory Committee on Advanced Television Service, October 1995.
[2] Sixth Report and Order, adopted April 3, 1997, FCC 97-115.
[3] Grand Alliance System Test Procedures - Part I: Transmission and Objective Tests, FCC Advisory Committee on Advanced Television Service, SSWP2-1306.

Digital Radio Production

Tuesday, April 20, 1999

9:00 am - 12:00 pm

Chairperson: Tom McGinley
WPGC-FM, Greenbelt, MD

*9:00 am **Implementing the Uncompressed Digital Air Chain**
Jerry Brown
CBSI/Custom Business Systems
Reedsport, OR

9:30 am **State of the Art Speech Processing for Broadcasting**
Martin Wolters
Cutting Edge Technologies, Inc
Cleveland, OH

10:00 am **Codec After Codec After Codec...**
Christer Grewin
Swedish Broadcasting Corporation
Stockholm, Sweden

10:30 am **How Many Bits Do You Hear?**
J.B. Brown, P.E.
Auditronics, Inc.
Memphis, TN

11:00 am **Universal 'cart' file interface for audio production/delivery systems**
Richard Pierce
Orban, Inc.
Hanover, MA

*11:30 am **Digital Workstations - Which One Is Right for You?**
Alan Peterson
Radio World
Falls Church, VA

*Papers not available at the time of publication

State of the Art Speech Processing for Broadcasting

Martin Wolters
Cutting Edge
Cleveland, Ohio

ABSTRACT

Many algorithms for processing speech have been developed over the past few decades including compression, automatic gain control, de-essing and equalization. Today, equipment is available providing only a subset of these functions (compressor, de-esser) or providing a combination of many functions (microphone processor). Sometimes, the same devices are used by engineers in recording and broadcasting, although there are different objectives in each application and different considerations have to be taken into account. Historically, limitations of analog equipment and limited budgets often led to workarounds and very inefficient use of available algorithms. This paper explores state-of-the-art processing of speech signals in a broadcast environment. The advantages of digital processing are described taking into account the interaction between speech processing and commonly used audio processing. Finally, different ways of integrating a digital microphone processor into a broadcast studio are illustrated.

INTRODUCTION

Creating the "sound of the station" has become an important issue in the broadcast industry over the past decades. A number of factors make the aesthetics of sound a key point in a station's format and success. These include, for example, increased competition, improved quality of alternative transmission systems (e.g. cable, DAB), the high quality of new receivers and stereo systems (even for car radios) and the higher expectations of listeners for good sound quality. Using purpose-built audio processing equipment — usually inserted at the very end of the audio chain — is a common technique to create the specific "sound of the station". Most of the time, this audio processing is optimized to improve the sound of the radio station's music format — obviously a very important, sometimes the most important part of the program. Since the music within a format and therefore the sound of the different songs within a program tends to be quite consistent, one can find that the application of certain processing parameters suffices to establish a station's on-air sound. In this case the raw material fed into a sound processor consists of more or less carefully produced recordings with a certain standard of quality with regard to leveling and equalization.

But announcers', talents' and DJs' voices are also an essential component of most formats. Much of this raw material is produced live, and very often there is no way to maintain the same standard that you can find in the above mentioned recordings. A specific processing of speech becomes necessary and is part of most modern broadcasting facilities. Nevertheless, it seems that the development of speech processing specifically for broadcasting has been neglected during the past decades. The result is little knowledge about how to use digital signal processing most effectively for such applications.

Based on knowledge about broadcasting and sound processing, combined with new scientific approaches about the properties of speech signals and the utilization of digital signal processing, new investigations toward the development of microphone processing products have recently been made. Some of the results are presented in this paper.

ALGORITHMS AND FUNCTIONS

The algorithms used in processing speech are automatic gain control (AGC), equalization (EQ), dynamic range control (DRC), de-essing, phase rotation (PR) and reverberation. Each of these algorithms has a specific task and the order in which these functions are arranged should not be arbitrary. Figure 1 shows an optimized signal path.[1]

Each function will be discussed in the following paragraphs focusing on the specific requirements in a broadcast studio, the advantages of digital signal processing and the benefits of combining these functions into a single unit.

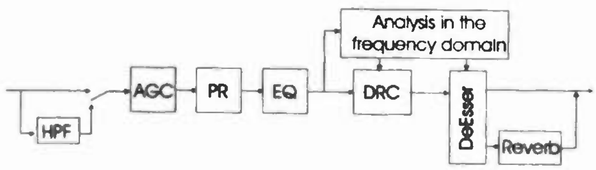


FIGURE 1

Automatic Gain Control (AGC)

One important issue in processing speech signals is level control. In a recording studio, the sound engineer usually takes care of the correct gain settings. The necessary gain is dependent on the room, the choice and position of the microphone and, of course, the person's voice. In a broadcast environment, the same room and the same microphone is used most of the time, so one could adjust the gain by taking into account these two factors. But the person's voice and the person's position might change. This is why an AGC is necessary. From an engineer's point of view, an AGC is a gain controller with a slow attack and release time. Another point of view is to consider the AGC as a replacement for the sound engineer. This latter concept might be more helpful, because one can visualize some important issues:

1. A good sound engineer carefully monitors the input level to make sure that it is nominally at 0 dB (the reference level). A compressor does a similar task; however it changes the level more frequently and more quickly.

2. A sound engineer can "detect" if a person is speaking or not and tries to maintain the desired level of operation when the person is speaking. When the person is not speaking, the sound engineer "freezes" the last gain setting. Therefore, the AGC must take into account the operation of a noise gate[2] which detects voice activity.

3. A sound engineer not only watches the input level, but also watches the compressor's activity. Adjusting the parameters of a compressor and adjusting the input gain are not independent and, therefore, the AGC and the compressor should interact.

The AGC should be one of the first stages within the signal path. Only a high-pass filter with a very low cutoff frequency (often referred to as a "rumble filter") should be placed before this algorithm. This high-pass filter reduces a possible DC offset introduced by the analog input circuit and filters unwanted noises such as hum, low frequency disturbance from touching the microphone (stand), etc. These signals would otherwise affect the operation of an AGC.

Dynamic Range Control (DRC)

In a superficial view, AGC and DRC appear similar in some ways. This is the reason that compressors — one part of DRC — are sometimes used as AGCs by adjusting the threshold very low so that the compressor provides an almost constant output level, independent of the input level. The result is very poor gain control, since none of the above mentioned issues are taken into account.

There are three new issues addressed by DRC:

1. DRC is used to "optimally use the full amplitude range of a recording system"[3]. Unfortunately (from a sound engineer's point of view) there are a few high level peaks in speech which reduce the available headroom of a recording. These peaks do not increase the perceptual loudness of a signal because this is affected more by an average value[4]. Hence, reducing the peaks does not decrease the loudness, but increases the available headroom and allows additional gain, resulting in an overall increase in loudness. This is sometimes referred to as peak control and is a more technical aspect of DRC, especially compression/limiting. Carefully chosen parameters lead to inaudible compression, up to a certain amount of gain reduction.

2. Beyond this certain level of gain reduction, compression becomes audible. Fortunately, this "sound" imparted by a compressor — the increased density of the speech signal — can be considered pleasant and is sometimes used to create a specific "sound". This is the more art-related aspect of

---

[1] A discussion about the benefits of the particular order is beyond the scope of this paper.

[2] See the next section about dynamic range control that includes a description of noise gates.

[3] From *Digital Audio Signal Processing* [1], page 207

[4] See *Psychoacoustics* [2], page 471

compression; the compressor as a tool for creating the "sound of the station".

3. DRC consists of more than just compression/limiting. A second, lower threshold can be utilized to further reduce all signals below this value. This is called an expander and, if the ratio of the reduction is almost infinity, signals below that threshold are muted and would be referred to as a "noise gate". The idea is that signals below a certain threshold are generally non-speech signals (e.g. background noise, paper shuffling and so forth) and should be reduced. This is especially important during interviews with studio guests or in the case of multiple announcers where a person's microphone is open but that person is not speaking. New research in the field of speech detection (e.g. for applications like mobile phones) led to "intelligent" noise gate algorithms.[3] Rather than just monitoring the energy of a signal, these algorithms utilize zero crossing rate and analysis in the frequency domain to determine if a valid speech signal or a disturbing background noise is present. Digital audio processing allows the implementation of some of these ideas into a microphone processor, resulting in a more accurate noise gate.

De-essing

In the past DRC — especially compression — and de-essing were integrated. De-essing was an extension to a compressor. Since research during the last two years resulted in new information about the properties of sibilants and led to development of new algorithms based on psychoacoustic evaluations, the connection between de-essing and DRC needs to be re-evaluated. An overview of the algorithms and concepts used in the past and an overview about sibilants and the problems in recorded speech introduced by these sounds can be found in a research study from 1998 [4].

To summarize the new information, one should distinguish between detection and reduction of sibilants. Investigations on speech recordings in four different languages showed that a very good and reliable detector for unpleasant sibilants is the psychoacoustic unit sharpness.[5] Figure 2 shows the mean and standard deviation of sharpness calculated for 50 test sentences and for 141 sibilants within these test sentences which were marked as disturbing by at least three of four test persons (experts from the recording industry). A value of 1.2 acum can be utilized to safely detect unpleasant sounding sibilants. Based on a frequency analysis related to the human hearing system, sharpness can be calculated in today's digital signal processors (DSPs).[6]

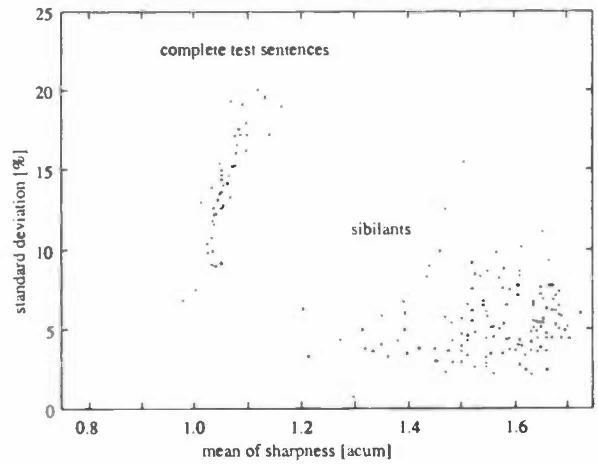


FIGURE 2

A very effective algorithm for the reduction of sibilants without many artifacts can be implemented using a combination of spectral subtraction with a time varying band-pass filter and broadband compression. Time varying means the band-pass adapts to the spectral properties of a specific sibilant. In addition the use of a small amount of broadband compression reduces the so called lisp-effect.[4]

Equalization

There are three different types of filters used in audio and speech processing:

1. High-pass/low-pass filters: As already mentioned, a high-pass filter with a very low cutoff frequency can be used to reduce a possible DC offset, low frequency hum, and background noise. Similarly, low-pass filters can eliminate high frequency noise. In general, these filters are used to limit the audio spectrum. They are less important in controlling the "sound of a station".

2. Shelving filters: These filters are used to weight (boost or cut) certain frequencies, in particular high frequencies above the cutoff frequency and low frequencies below the cutoff frequency respectively. One can create a specific sound of a station using these filters. But it may not be necessary to carefully adjust the parameters for each individual person. A more general approach (maybe separate for male and female announcers) can lead to a successful, good sounding timbre.

3. Peak filters: These filters allow very detailed changes within the frequency spectrum. They also allow changes of any desired frequency. Full parametric peak filters provide control of the center frequency, Q-factor and gain. Used with a low Q-factor, peak filters can be used as a general cut or

boost of the midrange, similar to the effect of shelving filters on high and low frequencies. Peak filters can be used for more detailed changes as well by utilizing a high Q-factor. But these kind of adjustments need to be made for each individual speaker.

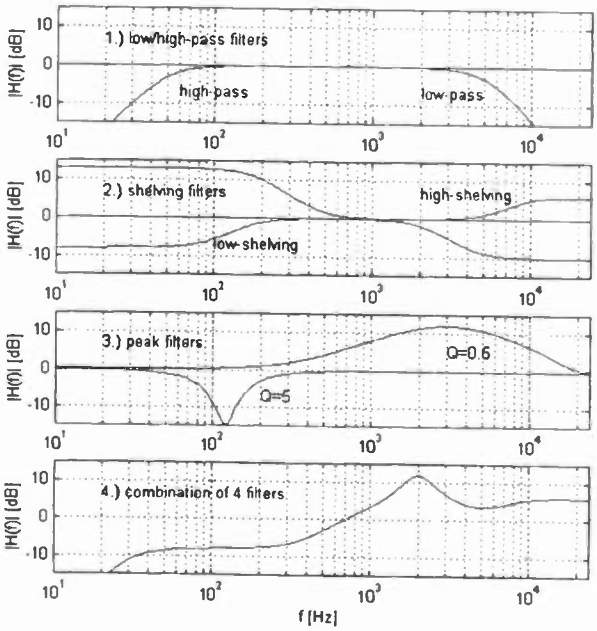


FIGURE 3

Figure 3 shows examples of the described audio filters. The graphic depicts second order low- and high-pass filters, four different shelving filters, peak filters with low and high Q-factors, and a combination of a high-pass, a low-shelving, a peak and a high-shelving filter. This combination simulates a possible EQ-chain in a microphone processor.

These filters, in general, are probably the best known sound processing tools. Rather than review the fundamentals of filters, there are two properties of digital signal processing related to the implementation of filters that will be discussed:

1. Without explaining the reasons and effects in detail, one should know that if DSPs with fixed point arithmetic are used, the quality of filters with low cutoff frequencies can be inferior. Even if this is a problem of fixed point arithmetic in general, there are good sounding, low noise algorithms available. These problems are less prevalent in a DSP with a floating point arithmetic, but even this approach may yield poor filter performance. This means a digital microphone processor that uses floating point arithmetic does not necessarily sound better than a unit that uses fixed point arithmetic; the best way to test such units is to tune them to low cutoff frequencies.[5]

2. The number of algorithms that can be used at the same time within a digital processor is limited by the computational power of the DSP used. This means, for example, that the number of filters that can be used at the same time is limited. Traditionally, three filters have been a reasonable number for a broadcast microphone processor. However, there are no restrictions to the number of types of filters. Since there is no drawback, a digital microphone processor allows one to use all of the above mentioned types of filters in any combination. Assuming there are three filters available the following combinations could be useful: a) Three peak filters (this combination might need careful adjustments on a per person basis); b) an adjustable high-pass filter followed by peak and/or shelving filters; c) a low shelving, a peak and a high shelving filter[6] or any other combination.

Artificial Reverberation

Although digital signal processing makes high quality reverberation possible there are still huge differences in the quality of artificial reverberation. This depends significantly on the computational power available — more than any other function described in this paper — and therefore directly impacts the price of a unit. In broadcasting, where artificial reverberation is infrequently used, the highest quality products are not required. For example a detailed adjustment of reverb parameters — such as the kind of surface, size of a room or absorption of higher frequencies — might not be necessary.

However, for broadcasters desiring artificial reverberation, there are two significant advantages in integrating artificial reverberation into a broadcast microphone processor:

1. A specific microphone preset (e.g. the personal preset of an announcer) would contain <u>all</u> parameters, including the settings for reverberation. Anticipating a later discussion in this presentation, it should be mentioned that it is very important to be able to restore settings of all parameters in a quick and easy way. It seems not to be very applicable to store

---

[5] See *Digital Audio Signal Processing* [1] for a discussion of these topics.

[6] In case one uses the peak filter with a low Q-factor this combination might be a good starting point for a general approach.

parameters for a separate reverb processor within a microphone processor. Integrating microphone processing into a broadcast facility includes controlling of reverb parameters and can be accomplished more easily by a built-in reverberation algorithm.

2. The combination of de-essing and reverberation increases sound quality. Some sound engineers in recording studios realized that the unpleasant sound of sibilants in recorded speech is significantly increased by artificial reverberation. They discovered that the problem could be mitigated by the use of two different de-essing units: One that controls the sibilants of the main signal and a second one that controls the sibilants of the signal used by the reverb processor. Integrating a de-esser and a reverb processor into a single microphone processor allows the use of this idea without increasing the cost of the unit. Since the detector for sibilants has to be implemented only once, a specific, advanced reduction of sibilants in the signal used by the reverb algorithm does not require much more computational power.[7]

Phase Rotation (PR)

A function unique to microphone processing for broadcasting is phase rotation. It was invented a couple decades ago to minimize artifacts created by general sound processing in broadcast facilities, especially during clipping. The reason for these artifacts is the asymmetric nature of some human voices.

Figure 4 shows, in the upper left corner, the waveform of a typical asymmetric voice. Although the average over time of this signal is zero (meaning that there is no DC offset), one can see that the peak values above zero are much smaller than the peak values below zero. A clipper limits a signal to an absolute value. The dashed lines in Figure 4 indicate a possible clip threshold. Clipping would affect the two halves of the signal differently. In such a case, clipping produces a more disturbing sound than clipping of a symmetric signal.

The reason for the asymmetry of a voice signal can be found by observing the relation in time of the different formants of a specific phoneme.[8] The two bottom plots on the left of Figure 4 show the major frequency components resulting in the asymmetric

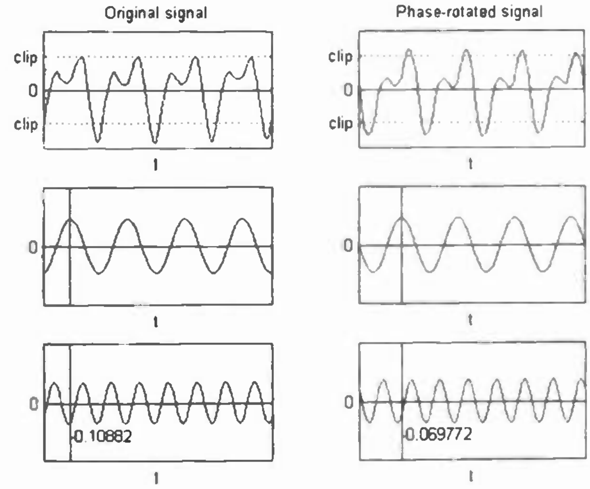


FIGURE 4

waveform. This relation in time is formed by the human vocal tract which can be modeled as an acoustical system of tubes with different lengths and sizes. The dimensions of these tubes are different for different individuals and different phonemes. This can cause an "unfavorable" phase of the frequency components resulting in an asymmetric signal.

Changing the phase of these signals more or less randomly (with an "all-pass" filter) is called "phase rotation" and results in reestablishing a symmetric signal. The right side of Figure 4 shows the processed signal and the changed relation in time of the two formants. Figure 5 clarifies the effect of an all-pass filter on the phase. One can see that the amplitudes of the signal are not affected. These changes of phase are usually not audible, except in the case of very

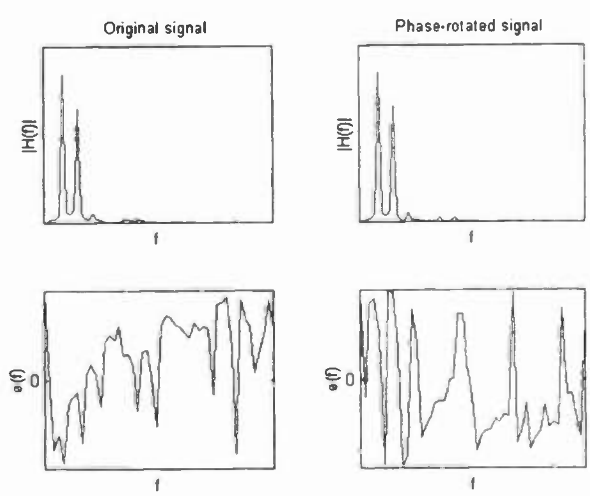


FIGURE 5

---

[7] See Figure 1 also.

[8] The relation in time can be studied by watching the realtionship of the phase of each freqency too.

transient signals. The best solution is to implement a phase rotator in a microphone processor and adjust this function for each person individually. In this way, the music programming is not affected and the phase rotator is only used when desired.

The challenge for the sound engineer is how to determine whether to use a phase rotator for a specific person and, if so, how much phase rotation is necessary. One could simply trust his ears. In that case, limiting the voice using a clipper can aid in adjustment. This might not be very accurate but in the end is the most important detector. One could add an oscilloscope to visualize the signal making asymmetric voices easier to analyze. The most accurate and elegant method would be an indicator within the microphone processor. By measuring the peak-to-average level of the positive and negative signal values and comparing these values, a simple but highly effective indicator would help the sound engineer to adjust phase rotation for a specific person.

USER INTERFACE AND INTEGRATION OF A MICROPHONE PROCESSOR IN A BROADCAST FACILITY

Requirements

There are some important requirements on how to integrate a microphone processor in a broadcast facility which affect the user interface of such a device and which are different from requirements in a recording studio. Besides the differing algorithms and functions described in the first part of this paper, the requirements of the user interface are an important reason to design specific microphone processors for broadcast facilities:

- There are generally several on-air and production studios within a broadcast facility. Once the parameters are adjusted for a specific person, it should be possible to use these settings in every studio.

- There is often no technician available. Selecting the correct preset must be very simple so non-technical persons can perform that task.

- Radio stations take their sound very seriously. In most cases the talent should not have access to change parameters capriciously.

- The unit should assist a technician in troubleshooting. Live broadcast requires reliability and, in case of technical problems, a quick way to detect and fix problems.

- A microphone processor should be able to be integrated in an on-air scheduler. That way the selection of correct presets can be automated.

- The microphone processor can be inserted as an effects processor into a mixing console or can be used as a microphone preamplifier as the first component within the audio chain. In the case of a digital studio, the AES/EBU outputs should be able to be synchronized.

An Elegant Solution

Based on the premise that most radio stations are already equipped with a computer network, the following system was designed:

1. The microphone processor itself has a very easy-to-use user interface. The simplest design is appropriate — meaning that the user can only change the preset of the unit but no other parameter. He chooses from a list that is sorted by preset number, preset name or the most recently used presets, allowing a convenient and fast way to find a specific preset.

2. In the case where a fixed preset is required (e.g. guest microphone), the preset can be locked.

3. There are level meters and status LEDs to assist in case of technical problems.

4. A headphone jack allows monitoring without additional hardware (e.g. at a workstation) and assists during troubleshooting.

5. Parameters and presets can be edited using remote software running on a computer. A more sophisticated user interface on this remote application assists the sound engineer when adjusting parameters much better than a necessarily smaller display on the front panel of the unit. The remote software can use different physical connections to the microphone processor such as TCP/IP networks, RS232 ports or other serial connections.

6. In a broadcast facility with more than one microphone processor, the units are connected to the network. A preset management system integrated into the remote application allows for easy distribution of a new or changed preset to each unit. Bigger radio networks can administer microphone processors in different studios from a single place. A security system allows only certain people to change presets and protects the units against unauthorized tampering.

Figure 6 gives an example how the different units are connected to control parameters and presets. Whereas a computer in the production studio might primarily be used to adjust parameters for a specific person, another computer (e.g. in the office of a station engineer) could run an application for the preset management and other administrative tasks.

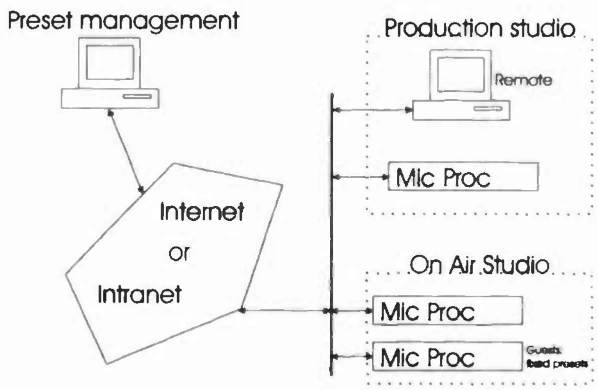


FIGURE 6

CONCLUSION

The algorithms and functions used in state of the art speech processing for broadcasting were summarized. Where digital signal processing can improve these functions, the necessary technical information was given. The benefits of combining all speech processing in a single unit were listed. A summary of the properties of speech signals was added where they explain the goals and reasons of a specific processing function. An overview of the requirements for the integration of a microphone processor into a broadcast facility led to a new approach for a specialized user interface.

REFERENCES

[1] U. Zoelzer: *Digital Audio Signal Processing*, John Wiley & Sons Ltd, Chichester, 1997

[2] E. Zwicker: *Psychoacoustics*, Springer Verlag, Berlin, 1990

[3] R.J. Santiago: *A Noise Robust Method for Detection of Endpoints of Speech Utterances*, Master Thesis, Marquette University, 1997

[4] M. Wolters: *The Acoustical Properties of Sibilance and New Basic Approaches for De-essing Recorded Speech*, paper at the 20th Tonmeistertagung, Karlsruhe, Nov. 1998

[5] M. Sapp, M. Wolters, J. Becker-Schweitzer: *Reducing Sibilants in Recorded Speech Using Psychoacoustic Models*, paper at the ICA/ASA-Meeting, Seattle, 1998

[6] M. Wolters, M. Sapp, J. Becker-Schweitzer: *Adaptive Algorithm for Detecting and Reducing Sibilants in Recorded Speech*, 104th Convention of the AES, Amsterdam 1998, Preprint 4677

Codec after Codec after Codec .....

by

Christer Grewin
Swedish Broadcasting Corporation
Stockholm

Abstract

Low bit-rate audio codecs have today found their way to many different applications in a broadcast or production chain. This has inevitably led to cascading of codecs. This paper discusses the consequences on audio quality when perceptual audio codecs are cascaded. Results from both subjective assessments and objective measurements are presented.

1. Introduction

Perceptual audio codecs have made an extremely rapid entry on the market. Only six or seven years ago there were virtually no low bit-rate audio codecs in operational use. It is also less than sseven years since the MPEG 1 audio standard [1] was established. Although there were other coding schemes before that, it was with this standard that perceptual audio coding exploded into the market. Today the broadcasting, telecommunications and computer industries have invested and are still investing heavily in advanced audio codecs. Thousands of codecs are already in every day use in a wide variety of applications.

Even if it seems to be dominant, MPEG is not the only coding algorithm in wide use. For example ATRAC [2] from Sony is used in MiniDisc, Dolby AC2 and AC3 [3] have found a large number of applications and PAC [4] from AT&T may find an application in the US DAR[1] systems. But the development goes on and a new coding algorithm, AAC [5], which allows for lower bit-rates for equal quality, is on its way to enter the market.

As MPEG is the coding method probably most widely used today, results presented in this paper are related to this standard. However, measurements and subjective assessments of other algorithms show that these behave similarly when cascaded.

2. Evaluation Methods

The introduction of these codecs has confronted users with new problems. The codecs are based on psychoacoustics and are inherently non-linear. Consequently, conventional measurement techniques, which are designed to measure small deviations from linear behaviour in a piece of audio equipment will tell very little about the audible effects of the coding algorithm, which uses properties of the human auditory system for reducing the bit-rate.

The only way to assess an implementation or a signal path, which contains a perceptual codec, has, until recently, been to arrange formal subjective listening tests.

2.1. Subjective assessments

The establishment of existing standards [1, 6] and Recommendations [7] for low bit-rate audio coding has been based on results from numerous and extensive subjective assessments performed by organisations around the globe. All these tests were performed with a methodology originally developed by the Swedish Broadcasting Corporation [8]. During the last 6-7 years the method has been refined [9] to a point where extremely small differences between a coded and a reference signal can be detected by a skilled panel of subjects. The method is also recommended by ITU-R [10], to be used when small impairments are expected.

The basic principle of the test method can be briefly described as follows: the listener can select between three sources "A", "B" and "C". Source "A" is always the known Reference Signal, i e the uncoded signal. A hidden Reference Signal and the Signal

[1] Digital Audio Radio

Under Test[2] are simultaneously available but are "randomly" assigned to "B" and "C". The listener is asked to assess the impairments on "B" compared to "A", and "C" compared to "A", according to a five-grade impairment scale. The grading scale shall be treated as continuous with "anchors" derived from the ITU-R five-grade impairment scale [11]. See Figure 1.

5.0	Imperceptible
4.0	Perceptible but not annoying
3.0	Slightly annoying
2.0	Annoying
1.0	Very annoying

Figure 1
The ITU-R five-grade impairment scale

As one of the sources, "B" or "C", is identical with the Reference Signal "A", it must by definition be given grade 5.0 = imperceptible difference from the reference. The other source may contain artefacts and may be given any grade on the scale. All detected differences between the Reference and the other source must be regarded as impairments.

As the subjects have to grade both the hidden Reference Signal and the Signal Under Test, this method allows for evaluation of the ability of each subject to hear and judge impairments.

The analysis of the result from a subjective listening test is generally based on the Subjective Difference Grade (SDG) defined as:

$$SDG = Grade_{\text{Signal Under Test}} - Grade_{\text{Reference Signal}}$$ [3]

The SDG values range from 0 to -4, where 0 corresponds to an imperceptible impairment and -4 to an impairment judged as very annoying. Together with the SDG, also the confidence interval at the 95% level is often calculated and presented.

[2] Signal Under Test is defined as the signal which has passed through the equipment or the signal path which is tested.

[3] In an earlier convention of presenting results from subjective assessments, grades for the Reference and the Signal Under Test were calculated and presented separately.

2.2. *Objective measurements*

Although it may be argued that the human ear is the best and most sensitive tool for evaluation of audio quality it would of course be desirable if a listening panel could be replaced by a measurement tool which gave a similar output value, i.e. a direct mapping to perceived audio quality.

In 1994 ITU-R established Task Group 10-4 with the aim to propose a draft Recommendation for a perceptual objective measurement method for assessing low bit-rate coding systems. After having defined applications and requirements the Task Group issued a call for proposals for such a measurement method.

A total of 6 organisations proposed models to participate in the evaluation process that took place in the beginning of 1996.

KPN	Netherlands	PAQM
FhG/ Deutsche Telecom	Germany	NMR
University Berlin/ Deutsche Telecom	Germany	Disturbance Index
CCETT	France	POM
CRC	Canada	PERCEVAL
IRT	Germany	IRT "Tool Box"

Although some of the models showed promising results, the Task Group came to the conclusion that **none of the models *__at this stage__* fully met the requirements for being submitted as a Recommendation.**

In order to reach the objective to have a draft Recommendation ready during 1998, Task Group 10-4 encouraged the model proponents to collaborate in a joint effort with the goal to develop an improved measurement method. All model proponents did agree to this and since mid 1996 further development has been done jointly by the proponents.

A draft Recommendation was completed by Task Group 10-4 in March 1998 and later the same year it was approved as ITU-R Recommendation BS.1387. The objective measurement method has now got the name PEAQ, Perceptual Evaluation of Audio Quality. It contains two versions, a Basic Version which allows for real-time implementations and an Advanced Version in which high accuracy is more important than real-time capability. A first implementation of the method is now available on the market through a German Company, Opticom in Erlangen.

As the first implementation of the PEAQ method was made available to customers only very recently practically no measurements have been made outside the development group. However, some of the models originally proposed to Task Group 10-4 have been available. Thanks to a co-operation with KPN Netherlands the Swedish Broadcasting Corporation (SR) has had access to the model proposed by them, Perceptual Audio Quality Measure or PAQM [12]. All objective measurements presented in this paper have been made using the PAQM method.

3. DAB and Choice of Bit-Rate

A strong incentive for broadcasters and others to introduce perceptual codecs has, of course, been the economical factor. If the need for transmission bandwidth or storage capacity became less, costs would become lower. However, during the development phase of the coding methods, the strongest force may have been the wish to make a digital broadcasting system feasible.

During the work of the MPEG Audio group in the late 80s and beginning of the 90s, it became clear that the Eureka 147, DAB, project would adopt the MPEG standard when it was completed. Layer 2 of the MPEG1 and MPEG2 standards [1, 6] is the standardised source coding algorithm in DAB.

The DAB standard is flexible and allows for bit-rates from 32 to 384 kbit/s with 48 kHz sampling and down to 8 kbit/s with 24 kHz sampling. The broadcaster is free to choose the bit-rate and thereby also the audio quality for each audio channel.

Some guidance for quality as a function of bit-rate can be found in the results from the second subjective assessment for MPEG/Audio in May 1991 [13]. Figure 2 shows the result at 256 kbit/s, stereo. On the X-axis are the 10 programme items used in the test (**All** is the mean value for the 10 items) and on the Y-axis are the SDG-values together with the 95% confidence intervals. The interpretation is that when a "bar" crosses the zero-line there is no statistically significant difference between the reference and the coded signal. The item is subjectively transparent. It should, however, be noted that there is no such thing as true transparency in perceptual audio coding. The coding process modifies the original signal, removes information and adds noise, and the original can **not** be re-generated. It is justified to talk about "lossy coding".

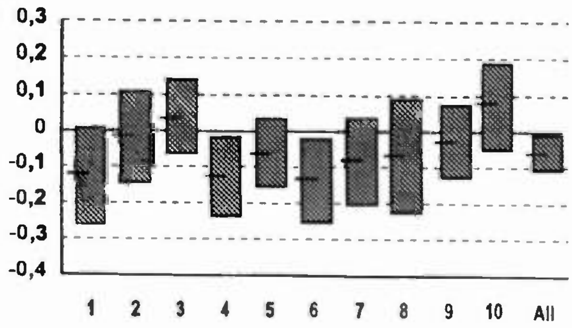


Figure 2
MPEG/Audio, Subjective Assessments, May 1991
Layer 2, 256 kbit/s (stereo)

From the result in Figure 2 it seems that 256 kbit/s would be a realistic choice for a high quality audio channel. Remember that DAB is often marketed as a system capable of "CD-quality", and this is what many consumers expect. It is easy to understand those broadcasters who have settled for 256 or even 224 kbit/s (which gives a similar quality).

However, it is much more difficult to understand broadcasters who have chosen 192 kbit/s as their basic bit-rate. Figure 3 shows the result for Layer 2 at this bit-rate. This result is also taken from the MPEG 1991 subjective assessment [13]. Here 9 of the 10 programme items are below the transparency level, with one item at –1.25 grades.

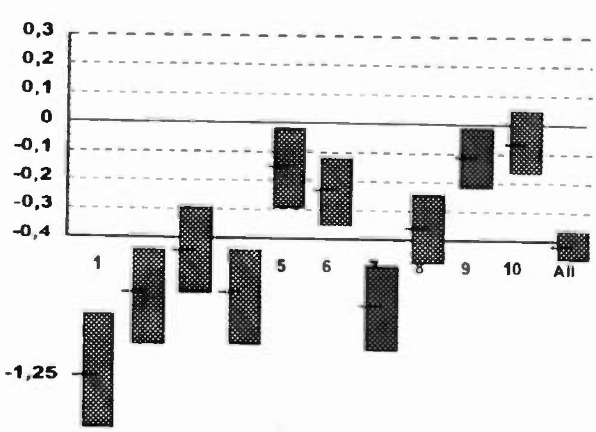


Figure 3
MPEG/Audio, Subjective Assessments, May 1991
Layer 2, 192 kbit/s (stereo)

Furthermore, most broadcasters plan for data services transmitted as Programme Associated Data, PAD, which is transmitted as a part of the audio bit-stream and takes bits from the audio and results in even lower audio quality.

Broadcasters decision on 192 kbit/s may have been taken with the hope that encoders would be developed further, and audio quality thereby improved. This would be possible as the MPEG standard only describes the syntax of the bit-stream and a decoder that can interpret it. But has quality improved? Unfortunately the answer seems to be "no". Figure 4 is an excerpt from the results in the subjective assessments performed in 1995 by CRC in Canada in the evaluation process for DAR systems in the US [14].

Five of the items are downgraded 1.0 to 1.5 grades on the ITU-R 5-grade impairment scale. An outcome that is poorer rather than better compared to the 1991 MPEG test. Although results from two subjective assessments can not be compared directly, the conclusion must be that audio quality has not been improved.

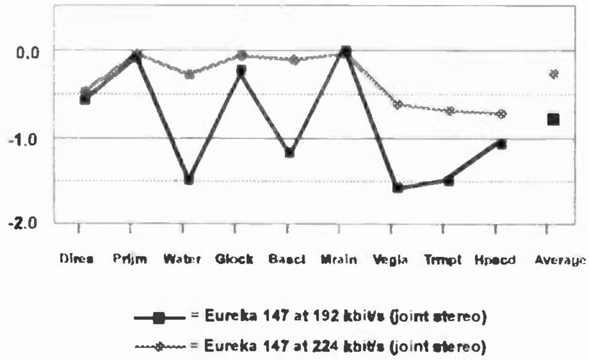


Figure 4
Excerpt from the US DAR subjective assessments

The quality at 192 kbit/s is below what is required by for instance the EBU[4] and also below what my company believes is acceptable. A naive listener may not discover the quality degradation at a first listening but the learning effect is strong. What an expert listener discovers during a listening test will be noticed at a later stage also by the ordinary listener.

4. Other applications

Another application that attracted broadcasters early was programme contribution and distribution over digital networks. These networks are based on the telecom hierarchy and allow normally only bit-rates in multiples of 64 kbit/s.

Furthermore, many broadcasters started to introduce hard disc based editing systems in the early 90s. Although most of these used linear PCM, systems to be used on local area networks, LAN, were often based on data-reduction. In the Swedish Broadcasting Corporation approximately 40% of the 800 workstations use MPEG Layer 2.

[4] European Broadcasting Union

Other applications were digital cart machines for jingles and play-out systems for music. Broadcasters suddenly found that their material passed through some coding stages before it was finally delivered to the listeners. And the last stage could, of course, also include a low bit-rate codec, e g DAB or ADR[5].

The quality degradation caused by cascading became obvious for those who had chosen bit-rates as if each application had been isolated. When SR began to introduce bit-rate reduced hard disc based editors we demanded from the manufacturer that the system should operate at 384 kbit/s (MPEG Layer 2); a requirement hitherto unheard of by that manufacturer. For the national digital network, which became operational in the same period we chose MPEG Layer 2 coding at the same bit-rate, 384 kbit/s. As shall be seen in the following this proved to be wise decisions and we have so far been spared most negative effects of cascading.

5. Cascading

When the Swedish Broadcasting Corporation evaluated audio codecs for a national network in 1994 an objective perceptual measurement model, PAQM [12], was used as a first step in the audio quality evaluation. This work is described in [15]. At that time there was no time to also perform a formal listening test for validation of the objective measurement results.

Later the same year measurements on cascaded codecs were performed. These were complemented with a listening test on some of the measured conditions. A commercially available MPEG 1, Layer 2 codec was used at 256 and 384 kbit/s (stereo). The programme items were most of the items used in the MPEG 1 and in the ITU TG 10-2 tests, a total of 27. The original (16 bit linear PCM at 48 kHz) was played from a hard disc through the codec and all odd numbers of cascading were recorded and used for the measurement and for the listening test. At 256 kbit/s 1, 3, 5, 7, and 9 stages were measured and at 384 kbit/s also 11 stages. Detailed results are presented in [16].

A summary of the objective measurements is shown in figures 5 and 6. The mean value for all 27 items is shown together with the best and the two worst items.

[5] Astra Digital Radio

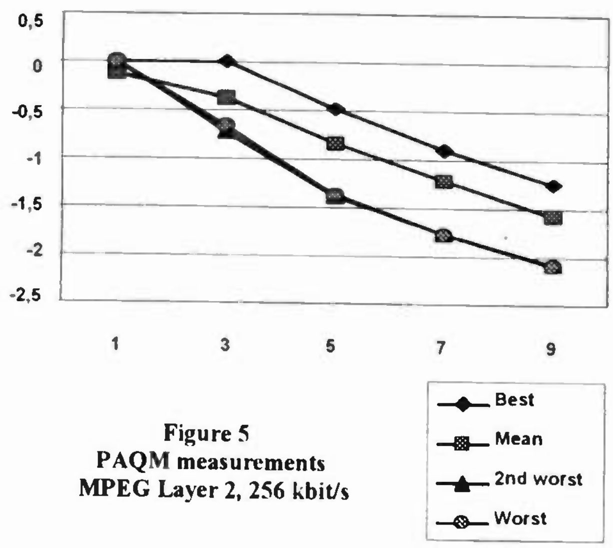


Figure 5
PAQM measurements
MPEG Layer 2, 256 kbit/s

At 256 kbit/s it can be noted that degradation of some items start immediately and that there is an almost constant difference between the best and the worst item at each coding stage.

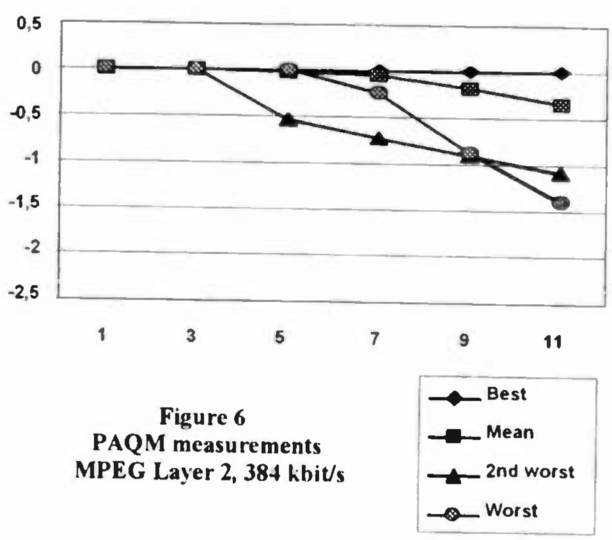


Figure 6
PAQM measurements
MPEG Layer 2, 384 kbit/s

At 384 kbit/s all items proved to be "transparent" after 3 stages and only the item "glockenspiel" was significantly degraded after 5 stages. After 11 stages of coding about 50% of the items did not reveal any significant artefacts.

As mentioned before a subjective assessment was performed on a sub-set of the conditions. A total of 20 test conditions were chosen for this test. The result revealed a rather god correlation between the objective measurement and the subjective assessments. It was shown that certain items were "transparent" after 11 stages of cascading at 384 kbit/s and it was also shown that in all cases were PAQM had indicated a degradation this was confirmed by the subjective assessment. Although the subjective assessment only included a sub-set of the test conditions the over-all behaviour of the codec shown by the PAQM measurement is likely to be very near the truth

Another listening test with the same codec at 256 kbit/s gave a result, which is remarkably close to the result with the objective measurement method. Figure 7 shows the result from a listening test performed within the framework of ITU-R TG 10-4 with a total of 4 programme items. Note how close the mean values are compared to the objective measurement, figure 5.

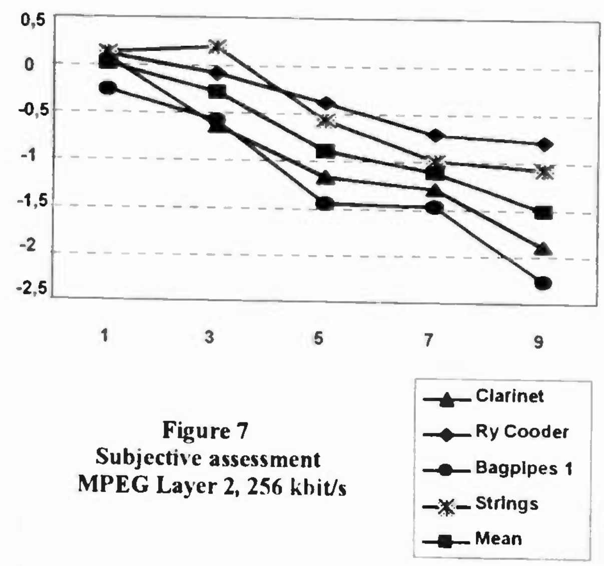


Figure 7
Subjective assessment
MPEG Layer 2, 256 kbit/s

It has sometimes been argued that if the 24 ms audio frames were time aligned between coding stages the negative effect of cascading would be eliminated, or at least considerably reduced. The IRT in Germany has made some tests with cascaded codecs [17]. Two different chains with cascaded MPEG Layer 2 codecs were tested with and without frame alignment.

Chain 1:	384 – 256 – 384 – 256	kbit/s
Chain 2:	384 – 256 – 384 – 192 (js[6])	kbit/s

Four test items were used and 16 subjects participated in the test. Figure 8 shows the result. For each chain and each programme item, the left bar is the unframed case and the right one is the framed. As can be seen there is a statistically significant improvement for three items when chain 1 is frame aligned. But for two of the items, "strawinsky" and "harpsichord", there is still a noticeable degradation. For the item "dire straits" there is of course no improvement to be expected as it was "transparent" already with the unframed chain.

[6] Joint Stereo Coding

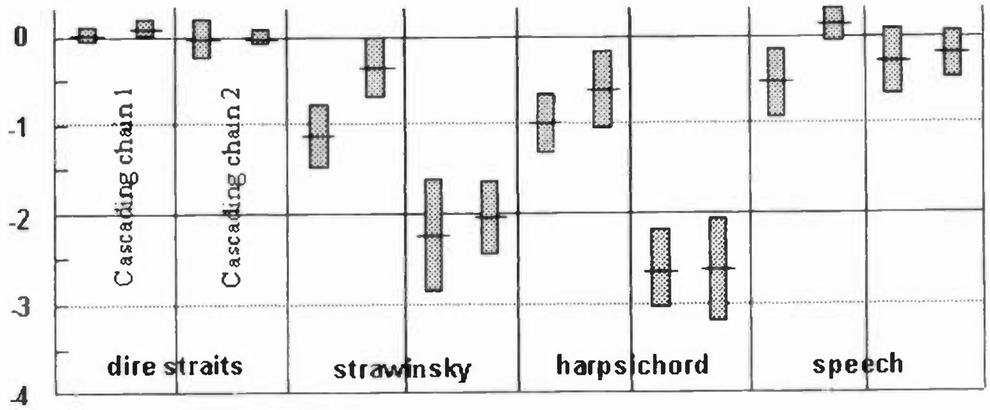


Figure 8
Comparison between framed and unframed codecs

However, for chain 2 there is no difference if the codecs are frame aligned or not. Obviously the rather large degradation in the last stage (192 kbit/s) has a much stronger impact on the overall quality then the small improvement of the framing.

In the same test at the IRT, considerable differences were found between different codec implementations. This is of course due to the fact that the encoder is not standardised, which leaves a lot of freedom for variations in implementations.

6. Conclusions

All cascading of audio codecs has a negative influence on audio quality. And the negative effect of a low bit-rate can never be compensated for later. The highest bit-rates in the MPEG standard allow for some coding stages without serious degradation of quality. But 256 kbit/s may introduce noticeable artefacts already after 2 or 3 stages. The degree of quality loss is also closely related to the programme material.

As mentioned earlier alignment of the audio frames has some positive effect but is not sufficient to allow for several coding stages. A further step is taken within the ACTS ATLANTIC project. A method for transporting control data between codecs has been defined. The method, which is called Audio Mole[7], is also proposed as an SMPTE Standard. Although the method has a potential for preserving audio quality after cascading it remains to be seen if this is a practical approach.

[7] The term "Mole" is a registered trade mark of one of the partners in the ATLANTIC consortium

During a number of years we have seen how storage capacity has increased without a similar increase in cost. That is to say that the cost per stored bit is much lower today than it was a couple of years ago.

And the capacity is there. Any laptop computer today has a hard disc that can store some GByte while the standard only three or four years ago was some MByte. And when SR introduced an archiving system based on data tape (QIC) in the early 90s [18], the capacity per cassette was just over 2 GByte. This allowed us to store 24 hours of bit-rate reduced audio per cassette. Today the capacity is 26 GByte per cassette at an only slightly higher price. In fact this allows us to archive in a linear format instead of a bit-reduced without any increase in cost.

The same development can be seen in the field of transmission. Capacity has increased tremendously only during the last few years. Network operators in every country install fibres with large transmission capacity and this will lead to lower prices.

The question is if this development towards higher storage and transmission capacity will make low bit-rate audio codecs obsolete for production, contribution and distribution in the near future. It can of course be argued that storage or transmission of fewer bits will always cost less. However, if the difference is only marginal or if the minimum tariff is for a rather high bit-rate the argument is of no importance. It might be that low bit-rate audio codecs will make sense only in the last link to the consumer, over air or via some recording media.

But currently we have a large number of perceptual codecs in use and there are likely to be more also in the coming years, so it is justified to give the following advice to anybody who intends to introduce a codec:

- Consider the entire environment, i e other possible codecs in the chain, before deciding on the introduction of a low bit-rate audio codec for a new application.

- Choose a higher bit-rate than what is required for the single application.

Finally, maybe all perceptual audio codecs should have a text of warning similar to what we find on cigarettes.........

Cascading is Dangerous and May Cause Distortion

7. References

[1] ISO/IEC 11172-3, "Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to 1.5 Mbit/s – Audio Part", November 1992

[2] K. Tsutsui et al "ATRAC: Adaptive Transform Acoustic Coding for MiniDisc", AES Collected Papers on Digital Audio Bit-Rate Reduction, pp. 95-101, 1996

[3] L. D. Fielder et al, "AC-2 and AC-3: Low Complexity Transform-Based Audio Coding", AES Collected Papers on Digital Audio Bit-Rate Reduction, pp. 54-72, 1996

[4] J. D. Johnston et al, "AT&T Perceptual Audio Coding (PAC)", AES Collected Papers on Digital Audio Bit-Rate Reduction, pp. 73-82, 1996

[5] K. Brandenburg and M. Bosi, "ISO/IEC MPEG-2 Advanced Audio Coding: Overview and Applications", AES Preprint 4641, New York 1997

[6] Coding of Moving Pictures and Associated Audio (MPEG-2) ISO/IEC International Standard 13818-3 (1994)

[7] ITU-R, Rec. BS.1115 (1994) Low Bit-Rate Audio Coding

[8] C. Grewin, S. Bergman and O. Kejving, "A Listening Test System for Evaluation of Audio Equipment", 80th Convention of the AES, Montreux 1986, preprint 2335

[9] T. Grusec, L. Thibault and R. J.Beaton, "Sensitive Methodologies for the Subjective Evaluation of High Quality Audio Codng Systems", presented at the AES UK DSP Conference, London, September 1992

[10] ITU-R, Rec. BS.1116, "Methods for the Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems", 1994

[11] ITU-R, Recommendation BS.562, "Subjective Assessment of Sound Quality", 1994

[12] Beerends, J. G. And Stemerdink, J. A. A Perceptual Audio Quality Measure Based on Psychoacoustic Sound Repesentation J. AES, Vol. 40 (1992), No. 12, December, pp. 963-978

[13] S. Bergman, C. Grewin and T. Rydén, "The SR Report on the MPEG/Audio Subjective Listening Test", Swedish Broadcasting Corp., ISO/IEC/JTCSC2/WG11 (MPEG91/010), June 1991

[14] T. Grusec et al, Communications Research Centre, Ottawa, "EIA/NRSC DAR Systems, Subjective Test of Audio Quality and Transmission Impairments", Final Report, July 1995

[15] C. Grewin, "Evaluation of Low Bit-Rate Audio Codecs for a National Network", 98th Convention of the AES, Paris 1995, Preprint 3944

[16] C. Grewin, "Can Objective Methods Replace Subjective Assessments?", 99th Convention of the AES, New York 1995, preprint 4067

[17] S. Ritscher and U. Felderhoff, "Cascading of Different Audio Codecs", 100th Convention of the AES, Copenhagen 1996, preprint 4174

[18] E. Hardeng, "Multi-User Digital Audio Storage and Archiving Using Data Reduction", Proc. AES UK Digital Audio Interchange Conference, London 1993, pp. 113-115

How Many Bits Do You Hear?

J. B. Brown, PE
Auditronics, Inc.
Memphis, Tennessee

ABSTRACT

Current digital audio systems range from 16-bit to 24-bit, however, some bits may be pure noise, having no contribution to audio quality. Several comparisons are given relating digital audio quality to analog audio specifications (Dynamic Range, THD+N, SNR), and the equivalent number of non-noise bits. These comparisons are then used to evaluate the latest analog-to-digital (A/D) and digital-to-analog (D/A) converters.

WHERE AUDIO QUALITY IS DEGRADED

First we must discuss the parameters that affect the audio quality for a digital processing system. In general, the transport of digital audio information into, out of, and within a digital processing system is assumed to be error free, so long as appropriate bit width is maintained (16 bits for CD quality, 24 bits AES/EBU maximum) and rate conversion is not required. Within the heart of a digital system, processing (Compressing/Limiting, Compression/Decompression, EQ, Mixing, Gain) may reduce the effective bit width. However, as far as straight recording and playback are concerned (unity gain), bit width is maintained. Thus, the interface to the analog world has the most limitations. In this paper we will focus on the performance of A/D and D/A converters.

DIGITAL SPECIFICATIONS

The main specifications quoted to compare A/D and D/A converters are Resolution, Signal-to-Noise Ratio (SNR), Dynamic Range, and Total Harmonic Distortion plus Noise (THD+N). A brief discussion of how these parameters relate to A/D and D/A converters is helpful.

Resolution: the maximum bit width of the digital word output by the A/D converter or input to the D/A converter.

Dynamic Range: the ratio (in dB) of the device output when driven by a maximum input to the device output when the input is grounded (for an A/D converter) or zero (for a D/A converter).

SNR: the ratio (in dB) of the power in the notch-filtered signal peak of the device output when driven by a maximum input to the power in the remaining signal.

THD+N: the inverse ratio (in -dB or percent) of the power in the notch-filtered (on frequency) signal peak of the device output when driven by a maximum input to the power in the remaining signal.

Of the above parameters (Resolution, SNR, Dynamic Range, THD+N), THD+N provides the most accurate indication of audio quality (indicates maximum number of non-noise bits).

BITS TO dB AND %

Each bit represents an additional power of 2: n bits provides a dynamic range of 2^n-1. Therefore, a 16-bit representation has a dynamic range of $2^{16}-1 = 65535$. The formula to convert to dB is dB=20*log(x). If x corresponds to one bit, an additional power of 2, then

$$20*\log(2) = 6.02\text{dB}. \quad (1)$$

Therefore, each bit corresponds to about 6dB and vice-versa. $2^{16}-1 = 65535$ goes to 20*log(65535) = 96.3dB which is about 6dB*16 bits. For a percent measurement, the percent value is converted to the equivalent fraction before computation: 0.05% goes to 20*log(0.05%/100) = -66dB, which is 11 bits, which is $1/(2^{11}-1) = 0.05\%$. Putting the above together yields the following equations for converting between bits, dB, and percent:

$$\text{dB} = 20*\log(2^n-1) \quad (2)$$
$$\text{dB} = 20*\log(\%/100) \quad (3)$$
$$n = \text{dB}/6, \quad \text{approximation} \quad (4)$$
$$n = 20*\log(\%/100)/6, \quad \text{approximation} \quad (5)$$
$$\% = 100/(2^n-1) \quad (6)$$
$$\% = 10^{2-\text{dB}/20}. \quad (7)$$

Note that Equations (4) and (5) are approximations based on Equation (1).

COMPARISON OF ANALOG AND DIGITAL SPECIFICATIONS

Equations (2) and (6) are used to generate Table 1, which provides a basis for comparing analog and digital audio specifications. As mentioned above, Dynamic Range and SNR are usually quoted in dB, while THD+N may be quoted in either dB or %. The dB values in Table 1 are expressed as ± as the conversion works for Dynamic Range and SNR (positive values), as well as THD+N (negative values). Note that 9-bit audio could be acceptable for some consumer applications. From a THD+N standpoint, 12-bit and greater meets professional audio standards. Perfect 16-bit audio has an exceptional -96dB THD+N (0.0015%). Achieving true 24-bit performance will yield an astounding -144dB THD+N (0.000006%)!

Table 1. Relation of Bits to dB and % Provides Basis for Comparing Audio Specifications

Bits	dB	%
9	±54	0.195695%
10	±60	0.097752%
11	±66	0.048852%
12	±72	0.024420%
13	±78	0.012209%
14	±84	0.006104%
15	±90	0.003052%
16	±96	0.001526%
17	±102	0.000763%
18	±108	0.000381%
19	±114	0.000191%
20	±120	0.000095%
21	±126	0.000048%
22	±132	0.000024%
23	±138	0.000012%
24	±144	0.000006%

STATE-OF-THE-ART A/D AND D/A CONVERTERS

Table 2 shows manufacturer's performance data (AKM[1], Analog Devices[2], and Crystal[3]) for state-of-the-art A/D converters, while Table 3 shows the data for D/A converters. Competitive pressures in the consumer electronics market have driven the design and manufacture of a variety of converters from 16-bit to 24-bit. An additional column in each table indicates the number of non-noise bits, computed based on the manufacturer's specifications, using Equation (4). In general, the converters do not meet theoretical maximum performance. State-of-the-art A/D converters deliver up to 18-bit performance,

Table 2. A/D Converters Deliver Up To 18-Bit Performance

Model Number	Manufacturer	Resolution (Bits)	SNR (dB)	Dynamic Range (dB)	THD+N (dB)	Number Non-Noise Bits
AD1878	Analog Devices	16	95	97	-98	16
AD1879	Analog Devices	18	98	103	-98	16
CS5330A/1A	Crystal	18	94	94	-84	14
AK5352	AKM	20	104	104	-97	16
CS5334	Crystal	20	100	100	-90	15
AK5393	AKM	24	117	117	-105	18
CS5397	Crystal	24	120	120	-105	18

Table 3. D/A Converters Deliver Up To 16-Bit Performance

Model Number	Manufacturer	Resolution (Bits)	SNR (dB)	Dynamic Range (dB)	THD+N (dB)	Number Non-Noise Bits
AK4309	AKM	16	*	92	-86	14
AD1858	Analog Devices	16	*	94	-90	15
AK4319A	AKM	18	*	92	-87	14
AD1861	Analog Devices	18	110	108	(0.004%) -88	15
CS4330/1/3	Crystal	18	94	94	-86	14
AK4323	AKM	20	100	100	-90	15
AD1862	Analog Devices	20	119	102	(0.0016%) -96	16
CS4327	Crystal	20	108	100	-90	15
AD1855	Analog Devices	24	113	113	-97	16
CS4390	Crystal	24	115	106	-97	16

* Not Listed

while D/A converters deliver up to 16-bit performance. Increasing the number of bits beyond 18 currently adds more noise bits, but there is no further increase in audio quality.

Note that THD+N is generally the limiting performance specification.

WHAT IS CD QUALITY?

CDs are currently 16-bit audio recordings. Just how good is CD quality? The quality depends on the quality of the recording, the quality of the playback, and the nominal signal level of the recording. Let us assume that the A/D and D/A converter interface circuits are designed so that overall circuit performance equals the performance indicated for the best 16-bit converters in Tables 2 and 3 (perfect interface circuitry). Then recording and playback will incur a penalty of a minimum of 1 bit (6dB) total. Now, a typical analog recording has a nominal level 20dB below maximum (20dB headroom). However, at Auditronics, we have found that assuming 12dB headroom (2 bits) on a CD recording (AES/EBU output) provides a better level match with analog format material (issues involving analog and digital level matching will be discussed in a future paper). Therefore we have 16-1-2 = 13 bits = 78dB THD+N theoretical maximum audio quality on a CD. Much material is and will continue to be CD quality (remember that tape cartridge machines are still in use).

CONCLUSIONS

While the AES/EBU standard allows for transport of up to 24-bit digital audio, current analog interfaces do not provide 24-bit performance for conversion to and from digital. Some bits are pure noise, having no contribution to audio quality. State-of-the-art D/A converters deliver only up to 16-bit performance (16 non-noise bits), while A/D converters deliver up to 18-bit performance (18 non-noise bits). Even so, much broadcast material is, and will continue to be mastered at CD quality: 16 non-noise bits maximum (-96dB THD+N). Within these 16 bits A/D and D/A interface losses must be accounted for (6dB THD+N minimum total, not including the effects of converter interface circuitry), as well as headroom (12dB for CDs). Therefore, current broadcast systems deliver up to 13-bit audio performance (-78dB THD+N maximum) at nominal signal levels. As such, specifying current A/D and D/A converters greater than 18-bit provides minimal improvement in audio quality.

REFERENCES

[1]AKM website, www.akm.com, February 1, 1999.
[2]Analog Devices website, www.analog.com, February 1, 1999.
[3]Crystal website, www.crystal.com, February 1, 1999.

CART/Audio Delivery Extension to the EBU Broadcast WAVE Format

Dick Pierce, Principle Software Developer
Orban Audicy Digital Audio Workstation
Orban, Inc. (A Harman International Company)
San Leandro, CA USA

ABSTRACT

This paper proposes a new RIFF WAVE data type specifically for use by broadcast audio production and on air delivery systems. It's purpose is to allow the export of audio data in the form of WAVE files along with needed scheduling, traffic[1] or continuity information from one system to another. Our intent is to allow systems from disparate manufacturers to participate in an integrated production/delivery network, greatly benefiting the broadcast user by preserving their product choices, and simplifying the task of integrating their systems.

This proposed extension is based on the well-established EBU Broadcast Extension WAVE file format (BEXT), the MPEG extension, and supporting documentation.

1 THE INDUSTRY'S NEED FOR THE 'CART' EXTENSION

Different on-air delivery and production systems use incompatible databases, audio file types and access methods, yet the scheduling, continuity or traffic information they use share many common attributes. As broadcasters leverage technology, audio becomes increasingly file-based, and the Wave file has become a de facto standard interchange format[2]. But the Wave file is somewhat brain-dead, in that it lacks the critical "label" information which readily identifies it for broadcast use. To simplify the integration of different systems, in this case, audio production and on-air delivery systems, a "digital cart label" for representation of continuity/traffic information, attached to and transmitted with an audio data file, will greatly benefit broadcasters using systems from multiple manufacturers. Efficiencies will be gained as users of one system export audio files and continuity data faster than real-time from one system to another using industry standard networking topologies.

The RIFF WAVE format has emerged as a dominant audio representation, and supports a wide variety of audio formats (linear PCM, MPEG and others), samples rates, and bit rates. The RIFF conventions allow the arbitrary addition of other data without impacting the ability of diverse RIFF-compliant[3] applications from reading and interpreting needed data. Thus, adding an extension to a WAVE file allows inclusion of needed continuity/traffic data to a widely accepted standard representation.

---

[1] As a point of clarification, "traffic" in the context of this proposal is intended to mean radio station traffic management, as in play scheduling and the like, and not "road traffic" information.

[2] In many ways, WAVE files have come to represent the same sort of "universal audio media" that two-track ¼" tape was up till recently. The computer revolution in audio has further emancipated audio by severing the once inexorable link between physical media and format representation. WAVE files can easily be stored and transmitted over nearly any media with no loss in fidelity.

[3] The RIFF specification requires all readers to be able to read all compliant RIFF files. When such an application encounters data that it is not prepared to handle, it can simply ignore the data and move on. There, indeed, exist some RIFF WAVE consumer applications that are intolerant of new and unknown chunks. For this reason alone, these applications are not RIFF-compliant. They may be front-ended by so-called "chunk stripper" utilities, the combination of which are, then, RIFF-compliant.

By utilizing a standard audio file format (WAVE and EBU/BEXT) and incorporating the common cart information into a specialized chunk within the file itself, the burden of linking multiple systems is reduced to producer applications writing a single file, and the consumer applications reading it. The destination application can extract the needed information and insert it into the native database application as needed. Communication between a production/delivery system is thus reduced to a simple, purely passive link that allows the production application to write the properly formatted WAVE file in a standard "drop box" location, where the delivery system, periodically polling the drop-box for new additions, finds the file, opens it, and uses it's own native access methods for adding this information to its database.

The result is that both production/editing systems and on-air delivery systems can communicate readily without the need for implementation-specific intelligence or design.

We certainly don't intend to suggest that the contents we describe here represent the only relevant data for all cart systems. Rather, we surveyed a variety of systems from several manufacturers and found that a large subset of the data is common to all. We thus based the design on this common set of fields.

As to the audio contents, the recommendation has been made elsewhere of the importance on standardizing on a common exchange format, and WAVE, especially in the form of the EBU Broadcast Extension standard. We endorse this recommendation, though this does not necessarily require the use of EBU/BEXT files as each system's native format.

2 BROADCAST WAVE WITH CART EXTENSION

The relation of the new 'cart' chunk to other WAVE file components is illustrated schematically in Figure 1[4].

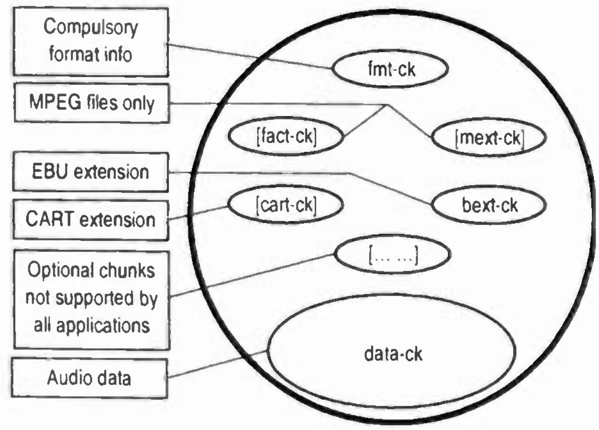


Fig 1: Broadcast WAVE file with CART extension

2.1 Contents of Proposed BEXT with CART extension

We are proposing using the standard EBU/BEXT wave format file with the addition of a new optional chunk type. Such a file would have the following chunks at a minimum:

```
<WAVE-form>  ->
  RIFF('WAVE'
    <fmt-ck>              // mandatory for WAVE files
    [<fact-ck>]           // required for MPEG files
    <bext-ck>             // BEXT information
    [<mpeg-ck>]           // required for MPEG files
    [<cart-ck>]           // optional cart information
    <data-ck>             // audio data
  )
```

Note: *Any additional chunk types that are present in the file are considered private. Applications are not required to interpret or make use of these chunks.*

[4] The ordering of chunks within a WAVE file is somewhat arbitrary. However, the prevailing opinion is that the 'fmt' chunk should *always* be first and the 'data' chunk should *always* be last in the file with no chunks following. This avoids the inclusion of so-called 'end chunks' which some applications can't handle. Within these constraints, though, the ordering could be arbitrary. We feel these conventions are prudent.

```
typedef struct cartchunk_tag
{
  DWORD ckID;             // chunk ID: "cart"
  DWORD ckLen;            // chunk data length
  BYTE ckData[ckLen];     // chunk data
}

typedef struct cart_extension_tag {
  CHAR Title[64];         // ASCII cart sequence title
  CHAR Artist[64];        // ASCII artist, creator name
  CHAR CutNum[8];         // ASCII 8 digit cut number
  CHAR Category[16];      // ASCII Category ID
  CHAR OutCue[64];        // ASCII out cue text
  CHAR StartDate[10];     // ASCII yyyy/mm/dd
  CHAR StartTime[8];      // ASCII hh:mm:ss
  CHAR EndDate[10];       // ASCII yyyy/mm/dd
  CHAR EndTime[8];        // ASCII hh:mm:ss
  DWORD dwTimers[4];      // 4 time markers after head
  CHAR UserDef[64];       // User defined text
  DWORD dw0level;         // Sample value for 0 dB
  WORD Version            // Version of the cart
  CHAR Reserved[718];     // Reserved for future
                          // expansion
  CHAR TagText[];         // Free form tag text
} CART_EXTENSION;
```

Field	Description
Title	A 64-character long ASCII string for the title of the cut. This differs in use from the EBU BEXT <description> field. The title is normally viewable on the cart or delivery applications and can be used as an entry into a table of contents or a key in an indexing or search system[5]. If the title occupies less than 64 characters, the last valid character is followed by a NULL byte '\0'. Some applications may not support a 64-character title and may truncate it as needed.
Artist	A 64-character long ASCII string holding the artist or creator name. This is different than the <originator> field in the EBU BEXT chunk in that it can be used to describe the original artist of a song, for example, while the <originator> field would be more appropriate as the producer of the specific audio file. If the string is less than 64 characters, it is terminated with a NULL byte.
CutNum	A decimal number string made up of ASCII digits '0'-'9', ranging in number from 1 to 99999999, representing the cut number, or unique cut key. The string should be left justified and terminated, if less than 8 digits long, by a NULL byte. Using a cut number of 0 signals the destination delivery system to auto-assign a cut number.
Category	This holds a category name up to 16 characters long. The category name is somewhat application dependent. It's advisable, though, to use common or standard category names, such as "PSA" or "NEWS", etc.[6].
OutCue	A 64-character ASCII string holding the optional outcue phrase to be displayed when the cut is being played. This is a user readable cue string. If less than 64 characters in length, it's terminated with a NULL byte.
StartDate	An ASCII date string of the form yyyy/mm/dd, such as 1998/12/25, holding the start date. Any valid date can be used. To signify an immediate start date, use a date and time significantly earlier than the present, for example 1900/01/01 Year is defined as 0000 to 9999. Month is defined as 1 to 12 (or 01 to 12). Day is defined as 1 (or 01) to 28, 29 30 or 31, as applicable.

[5] It is probably reasonable, though, to use the BEXT Description field for more detail about the cart.

[6] A list of category names might be appropriate for standardization.

Separator between fields is normally "/", but can be any of "-", "_", ":" " " (space) or ".".

StartTime — An ASCII time string of the form hh:mm:ss, such as 12:31:45, representing the 24 hour time-of-day for the start time on the assigned <StartDate>. There is no default for this field.

Hour is defined as 0 (or 00) to 23.

Minutes and seconds are defined as 0 (or 00) to 59.

Separator between fields is normally ":", but can be any of "-", "_","." " (space) or ".".

EndDate — As above in start date and time, but indicating the end date[7]. This the final air date after which the sequence will no longer be active. If the sequence is to run forever, use an impossible date such as 9999/12/31. There is no default for this field.

EndTime — Indicates the time of day on the appointed end date. There is no default for this field.

dwTimers — Four 32-bit integers representing a time mark offset from the start of the cut. Time units are in sample periods at the cut's current sample rate[8]. These timers can be thought of as secondary or tertiary markers or "trip tones", used to activate events in the cart system. If a timer is not used (unset), it's value is set to 0xFFFFFFFF.

UserDef — A 64-character ASCII string whose use and contents are defined by the user of the system. If the user text is less than 32 characters long it is terminated with a NULL byte.

dw0Level — A 32-bit signed (two's complement) integer word holding the sample value of the 0 dB reference level for the originating system. This is to facilitate scaling and metering consistency across disparate systems. As an example, a 16 bit linear PCM system that has it's meters calibrated as 0 corresponding to maximum signed digital value will have the value set to 32768 (0x00080000). A similar system with the peak value set to +6 dB will have this value set to 16384 (0x0004000), since the 0 dB meter reference level is 6 dB below, or ½ have the value of, saturation.

Version — An unsigned binary number giving the version of the cart, particularly the contents and usage of the Reserved area.

Reserved — Area reserved for future expansion of the standard.

TagText — Non restricted ASCII characters containing a collection of strings each terminated by CR/LF. This text can be system or even user defined descriptive text for the sound, such as script information, instructions, notes on an artist, or a live tag for a commercial[9].

2.2 Other Relevant Information

All the other information regarding WAVE audio characteristics can be found in the mandatory "fmt" chunk. This includes sample rate, number of tracks, sample width and sample format. For other than

[7] Often referred as the "kill" date. The term "end" date is to be preferred because of the confusion with the concept of "killing" the entry altogether, purging it from the system, as opposed to simply deactivating it.

[8] The timer range is 2^{32} or 4,294,967,295 sample periods. This allows timer ranges at a sample rate of 48 kHz, for example, to extend beyond 24 hours (24:51:18).

[9] An alternative use for this area is discussed later in this document.

PCM format, the "fact" chunk and the EBU "mpeg" chunk will contain further information. Refer to the relevant documentation for information on these data. Information regarding the exact format of the audio data representation is also to be found in the reference documentation.

2.2.1 'bext' chunk usage suggestions

The BEXT chunk already has fields that could be of use in a cart application without violating the intentions of the BEXT standard. Some of them are described here with some suggestions for usage in a cart/delivery system context. Refer to EBU 3285 for more specifics on the "official" usage of these fields.

Field	Description
Description	A 256-character ASCII field that can hold a free form descriptive text for the program. Several cart/delivery systems utilize a long description field, and this could be used for such a purpose.
Originator	A 32-character ASCII field holding the originator or creator of the sound.
OriginatorReference	A 32-character ASCII field holding, as 3285 describes, "a non-ambiguous reference allocated by the originating organization."
OriginationDate	A 10-character ASCII for date of creation
OriginationTime	An 8-character ASCII for time of creation
TimeReference	A 64-bit time code of sound sequence
Version	Unsigned short integer version number.
Reserved	A 254-byte reserved area for future expansion
CodingHistory	Non-restricted ASCII characters containing a history the sound's coding process.

3 APPLICATION-SPECIFIC INFORMATION

It is certain that some radio scheduling applications can make use of or require data not included above. The RIFF chunk file format allows addition of arbitrary new chunks in which such information could be placed for private exchange between applications that can understand the information. Such chunks, however, will fall under the category of "private" chunks and can be ignored by applications conforming to the standard definitions. As such, they are beyond the scope of this document.

Some correspondents have raised objections to the fixed-field layout of the cart chunk, preferring instead a free-format keyword-oriented approach. The latter, while providing a high degree of flexibility and expandability, suffers from a great increase in complexity. A fixed field format has the advantage of not requiring parsing and interpretation. Further, it runs a serious risk of violating the very notion of a standard by allowing essentially indeterminate content to be added to the chunk. One must remember that the objective here is to provide a commonly agreed upon means of communicating known and well-characterized data between diverse applications. The goal is to allow as many such diverse applications to connect as possible. It's also to be noted that none of the other relevant chunks (fmt, fact, bext, mext, etc.) employ such a free-format layout.

The proposed standard addresses these issues in two ways. First, as in the EBU BWF standard, an area has been set aside for future standardization. This area would continue expansion of fixed size fields in a manner as currently done. Secondly, the area at the end of the chunk (currently designated as the TagText member) is available for arbitrary expansion. If desired, this area could be used for such keyword-oriented data.

Assuming the latter path is chosen, it is the recommendation of the authors that a tagged format, essentially the same as the basic RIFF chunk format, be used in this area. This retains the advantage of keyword-orientation while also retaining the full backwards and forwards compatibility of the basic tagged methods. Diverse applications can easily navigate this area without the requirement of understanding all possible members. It is further recommended that such an approach be subject to the standardization process.

And, ultimately, there is a third approach, as mentioned above: further chunks can be defined as seen necessary.

Not to belabor the point, but again the goal is to provide a common means of communicating specific information for a reasonably well defined and constrained application. This standard makes no attempts to get it 100% correct, such is impossible. The user base is far better off with 95% correct now, instead of 100% correct never.

4 REAL WORLD USE

The 'cart' extension is no pie-in-the-sky vaporware. A version of the Orban Audicy system that fully implements networked cart file production has been developed and is now in being tested at several customer sites. We have been working closely with several manufacturing partners on fully integrating our systems with theirs via 'cart' WAVE files. Among these partners are ENCO, Prophet Systems, David GMBH, Studer, and Scott. Discussion is now underway with several more manufacturers. Most importantly, users have responded enthusiastically to the idea of the 'cart' WAVE file and the resulting ease of integration of their Orban editors and other manufacturers' on-air systems.

5 CONCLUSION

No single manufacturer can claim to address the needs and budgets of all possible users with workable integrated production and on-air delivery systems. Simple resource limitations dictate that any such system can never be everything to everyone. Moreover, broadcasters want choice in terms of what solutions they buy. Manufacturers need to hear what our market is saying, in short, "make it work together." To provide the maximum utility to the broadcast market as a whole, it makes sense to design interchange standards that allow the users to select production systems and delivery systems that suit their unique needs, integrating them with minimum effort. Our proposed 'cart' standard is intended for just that purpose, and we are hoping that its adoption will provide significant operating benefit to all parties.

6 ACKNOWLEDGEMENT

Geoff Steadman, Product Manager for the Orban Audicy workstation, can justifiably take credit as being the "father" of 'cart' junk idea. It was through his discussions with many Audicy users that he came to develop the basic germ of the idea and pushed the original formulation ahead. I would also like to acknowledge the support of Dr. Barry Blesser, Barry Demchak and others at Orban, who helped develop the idea into a viable, concrete concept. I would also like to recognize our many industry partners who are helping to make the idea of integrating all these systems a reality for the users.

7 REFERENCES

- Microsoft Resource Interchange File Format, RIFF
- Microsoft Software Developers Kit Multimedia Standards Update, rev 3.0 15 April 1994
- "Specification of the Broadcast Wave Format," EBU Tech. Doc. 3285, EBU Publications, Ancienne Route 17A, CH-1218 Grand Saconnex (Geneva) Switzerland, July 1997
- "Specification of the Broadcast Wave Format, Supplement 1: MPEG Audio" EBU Tech. Doc. 3285A, EBU Publications, July 1997

Radio Consolidation: Real World Experiences

Tuesday, April 20, 1999

1:00 PM - 5:00 pm

Chairperson: Milford Smith
Greater Media Inc, East Brunswick, NJ

1:00 pm Smart Radio Studio Construction Techniques in the Age of Consolidation
G. Michael Patton
Michael Patton & Associates
Baton Rouge, LA

1:30 pm Automating the Consolidated Station Operation
Jim Dalke
Medium Limited
Seattle, WA

2:00 pm 2.4 GHz Spread Spectrum Radio and T1 Technology Combine for a Wireless T1 STL/TSL Solution
David Agnew
Harris Corporation/Broadcast Division
Quincy, IL

***2:30 pm Consolidation of Five Major Market FM Stations into a Single Physical Location**
Paul Shulins
WROR-FM
Boston, MA

***3:00 pm Building a Group of Mickey Mouse Radio Stations**
Bert Goldman
ABC Radio Networks
Dallas, TX

***3:30 pm TechTrek: A View from the Bus**
Frank McCoy
GulfStar
Austin, TX

***4:00 pm Options of Studio Construction: Stick Built Versus Prefab**
Tony Gervasi
WPST-FM
Princeton, NJ

*Papers not available at the time of publication

SMART RADIO STUDIO CONSTRUCTION TECHNIQUES IN THE AGE OF CONSOLIDATION

G. Michael Patton
Michael Patton & Associates
Baton Rouge, Louisiana

ABSTRACT

In the current climate of mass consolidation, when half a dozen or more radio stations often must be relocated to a common facility, usually by fewer qualified engineering personnel and in a desperately short time frame, the conventional hand-wired techniques of terminating audio and control cabling in intra- and inter-studio wiring assemblies are so cumbersome and labor-intensive that they no longer serve the industry as a paradigm for efficient interconnection of audio equipment.

This paper discusses the author's history of experience and frustration with conventional methods, his search for alternatives, his struggles to separate the truly innovative from the merely different, and the considerably success he has had so far using this new paradigm for studio construction. The text will include discussions of wire types and uses, termination components, tools, and techniques, as well as audio performance data for real-world studios constructed in this manner.

STUDIO BASICS

Let us review briefly the way the typical radio station studio is built. The nerve center of any studio is, of course, the audio console, or board. Although there are ways to build a studio without a console, such studios are rare, and even then, something (usually a computer) must perform the functions of the console, i.e., controlling the audio switching and levels from various sources, issuing and receiving start/stop and other control tallies from those sources, and routing the audio to the various console outputs/destinations. Broadly, radio station studios can be split into three categories: control rooms, which serve as on-air studios feeding one or more transmitters; production rooms, where commercials are produced, recorded, and/or dubbed onto various storage media; and newsrooms, where newscasts are produced and recorded. All studios are, however, just variations on a theme, and can be built using the same basic techniques.

Studio Wiring

The console in any studio is connected to the source and destination equipment through a wiring harness. These harnesses can easily become complex enough to be confusing, and quickly turn into a tangled nightmare. Care must be taken at every step of the process to maintain strict "discipline" in routing, tagging, and allowing sufficient slack for the various wires. Compounding this are the inevitable changes in style and competence level between engineers. Additionally, it is often a different engineer, operating under a different set of pressures and priorities, who is the one to install additional or upgraded equipment during the life of the studio. It is no wonder then that most studios come to resemble rats' nests within a few years or even months, and that's assuming that they didn't start out that way.

Wiring Techniques

Traditionally, studios have been built along one of two lines: either the wires were run point-to-point, from the console to and from each other piece of equipment, or some form of terminal blocks were used as a central "point of exchange".

Point-to-Point Wiring

Point to point wiring is cheaper and takes less time to install. Because each wire is run from a specific console input or output to a specific piece of equipment, the entire wire must generally be replaced any time the outboard equipment is to be replaced or upgraded. Of course, the entire wiring harness must be replaced if the console itself is changed.

Central Point Harnesses

Using some form of terminal blocks as a central wiring point can be a much more sophisticated and flexible form of construction, although it is also quite a bit more complicated and requires substantial additional investments in time and supplies. Over the years, patch panels and what were called "Christmas Trees" have been used as terminal blocks, although the standard today is the (Siemens Type 66) punch block, which came to us from the telephone industry, where it is universal.
In a good punch block installation, wires are run from each console input and output to a set of punch blocks located at some reasonably accessible point in the studio (often located under a panel in the furniture), and wires are also run from each input and output of every other piece of equipment to adjacent punch blocks. Then, what are called "cross-connect wires" are punched down to interconnect the various sources to console inputs, and console outputs to destinations, plus any other interconnections that may be needed. In a Cadillac installation, punch blocks are also used for remote control and tally functions.

Wiring Between Studios

In a multi-studio situation, the gold standard is to make the interstudio wiring much like a large version of the wiring in each studio, using a central set of punch blocks with cables running to each studio and the STL, remote pickup, air monitor, phone, and other I/O equipment. Cross-connects are made between the various punch blocks to interconnect studios to each other and to I/O points. This is called a "star" system, with cables radiating out, figuratively if not literally, from the wiring "center" to each branch of the installation.

Cables & Techniques

The wiring used in a "conventional" installation is usually multiple-pair, individually-shielded audio ("snake") cables such as the Belden 877x series for interstudio wiring, individually shielded pairs like Belden 8451 or 8723 (and sometimes snakes, especially for consoles) for intrastudio wiring, and control cables such as Belden 8457 (12-conductor unshielded). Each of these cables was terminated by hand at each end, usually to punch blocks at one end and to the console or source equipment at the other end. Good techniques called for the use of heat-shrink tubing to cover exposed shields, especially on snakes, where the shields were susceptible to unraveling once the outer jacket was removed.

Studio Construction Standards

For the purposes of discussion, in this paper I will assume when referring to "conventional" studio wiring techniques that what I'm referring to is either point-to-point wiring, or the use of punch blocks, with cross-connects, including control wiring, using the proper techniques as I've just described.

AUTHOR'S EXPERIENCE

I have been involved in the construction of over 75 studios, mostly as the crew or project chief. Over the years, I have been involved with projects that range from single-studio radio stations that were literally assembled in one day by two people, to a six-station multiplex that took a crew of 3 full-timers and 2 part-time helpers a total of two months to complete. Up until two years ago, all of these installations were "traditional", using the wires and techniques described above.

Studio Construction Time

After looking at my experience in many stations, and also comparing notes with other engineers, we agreed that it takes about 150 man-hours of labor to construct a quality studio in the conventional fashion. It is easy to see where this number comes from, when to terminate one side of one punch block took 2 to 3 hours (not to mention the time spent on the other end of those same wires), and there are usually more than 12 punch block sides (using split blocks) in each studio. This was a large number of hours, but I was determined first and foremost to provide my clients with high quality studios.

Unsatisfactory Time Requirements

Over the years, I had become increasingly disenchanted with the efficiency of these conventional techniques, feeling that we in the broadcast industry were lagging years behind other fields in adopting modern, more time-efficient techniques for constructing studios. This came to a head when I was approached to move a five-radio-station group into a new building from two different old locations, all in a timely manner. Management also wanted to have provisions made for a sixth control room, a central interview studio, and 4 production rooms. (In retrospect, this many production rooms were not needed, but at the time, this was new territory for all of us, engineers and managers alike, and four production rooms were what I was asked to bid on.) This many studios called for a "master control room" with all the central wiring and I/O equipment in it, which in itself was comparable in labor required to build it to an additional studio. Assuming that the half-built sixth control room along with the interview studio added up to another studio's worth of labor, that totaled 11 studios to build; six control rooms, 4 production rooms, and master control. At 150 man-hours per, this would be 1650 man-hours, or almost three months full-time work for four people. Well, since I didn't have four full-time people or three months' time, I realized that it was time to develop some improved techniques.

Time to Innovate

While we in the broadcast industry have been living in the dark ages, the telecommunications people along with the computer industry have been making tremendous strides in reliable wiring that is assembly-efficient. They commonly use mass-termination connectors that can be connected in one operation, along with ribbon cable that has become quite sophisticated, including ribbon coax cables, twisted pair ribbon, and shielded ribbon cables. Out of desperation to devise a plan to build this giant facility in a reasonable time frame, what I did was to adapt technology used in these more progressive industries to the needs of radio studios.

Studio Construction Legends

There are a number of myths and legends in the audio business, many of which I have discovered to be hollow icons from a bygone day. For example, conventional wisdom in studio construction calls for each wire carrying audio at both mic and line level to be individually shielded, ostensibly to prevent hum and crosstalk. Through the course of several experiments, I determined, however, that only mic level signals need shielding for hum prevention, and that line level signals do not need to be shielded for crosstalk (in a properly-designed cable and when all lines are driven and terminated in a balanced configuration) and not at all for hum. This discovery, coupled with my finding a suitable twisted-pair ribbon in the Belden catalog, sent me down the road to developing my new set of construction techniques.

The Times, They Are a'Changin'

Another myth, or more of an unspoken assumption, is that high quality audio can only be carried through *audio* cables and *audio* connectors. This is patently untrue when you stop to think about it, as audio is not very different from telephone signals, or some computer data, or many other signals common in today's hi-tech office environment. Consider how much the times have changed: the standard "audio" cable was developed in an age where the cable's level of shielding, capacitance, etc., was necessary to provide sufficient isolation and high frequency response for the high-impedance output drivers and terminating inputs that were state-of-the-art for the tube-type studio audio gear in the early parts of this century, when these cables were developed. In today's studio environment, where almost every output is driven by a high-current op-amp with a typical output impedance of 40 ohms, and almost all inputs are balanced bridging, the advantages, indeed, even the necessity of individually shielded cables for professional line level (+4 dBm) signals becomes moot. Yet, in studios across America, and indeed the world, the wiring used is a clear descendant of these same early heavily shielded cables designed to protect the weak audio signals of the earliest microphones and studio gear. As for connectors, it is again plain that any connector that will provide a reliable low-resistance path will carry audio just as well as many other signals. The continued insistence on XLR and other low-density, high installation-time but venerable *audio* connectors (the ultimate example: the RCA phono plug, which has to be a candidate for the worse audio connector ever devised yet, against all logic, is defiantly still in common usage) was yet another refusal to keep up with the technological Joneses, one which was costing the broadcast engineering industry dearly.

A Journey of Discovery

Armed with my iconoclastic insights about cables and connectors, I set out to the local supply stores for telephone and computer installers and spend many hours browsing the isles, bugging the countermen, perusing the catalogs, and attempting to pick the brains of the professionals in that field. I was shocked to learn that while most of these "professional" installers would be lost immediately if asked to make even a simply repair to the inside of a piece of audio equipment, they were light-years ahead of the so-called sophisticated standards of the broadcast business when it came to wiring infrastructure. On my travels, I saw cable and connector technology that made our use of XLR connectors and convention snake cables seem like stone knives and bearskins against an F-16. I finally managed to quit acting like an awestruck tourist in the city and was able to come up with a set of cables, connectors, and accessories that integrated well into the modern radio studio.

THE ROAD TO PARADIGM

Before my supply of insight had quite been depleted, I had developed quite a set of new supplies, tools, and techniques, useful for many different aspects of wiring studios, but all designed to reduce the amount of time, especially skilled engineering time, needed to build a sophisticated, flexible, reliable studio or set of studios.

The Players

The basic innovative building blocks are :

- Punch blocks: Siemens (and others) make blocks that are commonly available that have 50-pin Amphenol "blue-ribbon" connectors attached to each side, with every punch terminal brought down to a pin in the connector.

- Wire: Belden (and others) makes ribbon cable with 25 twisted pairs, each pair of which is twisted opposite of the ones on each side of it to reduce crosstalk. Every 24" there is a 2" untwisted area to facilitate the installation of connectors. These ribbons are also available rolled up into a round jacket, for strength in conduit installations, and either with or without an overall shield. The overall shield makes the cable suitable for use in RF fields, and the round overall jacket makes it very tough and easy to pull through long runs without concern for the cable's integrity. Even the fanciest of these cables (the shielded and jacketed version) are smaller and less expensive than conventional snake cables. Another advantage of this cable is that its characteristic impedance is 110 Ohms, thus making it suitable for AES/EBU digital audio signals. Conventional audio cables have widely different impedances, and the most common ones can carry AES/EBU signals over only short distances without errors. On the cusp of the all-digital studio world, it only makes sense to use digital-ready wire and wiring systems, regardless of your current requirements.

- Connectors: Amphenol (and others) make IDC connectors that match those on the punch blocks and which can be crimped onto the ribbon cable in a single operation.

- Accessories: Siemens (and others) make devices called "harmonicas" (because of their physical resemblance), which have a 50-pin Amphenol connector on one side and RJ phone jacks on the other side. These harmonicas are available with RJ-14s (4 conductor) and RJ-45s (8 conductor). I use these harmonicas as breakouts on the source-equipment end of the ribbon cables coming from the punch-blocks.

Note: Terminating the ribbons with harmonicas may seem like an extra and unnecessary step, especially to engineers who have been drilled to avoid too many connectors (and potential poor connections) in any audio wiring run. Indeed, the first studio I build this way made little use of harmonicas; but what I found was that the use of pigtails has several advantages, including one that was not clear to me until I stumbled onto it, late in my first installation. See the detailed explanation below.

The Basic Installation

Here is the typical studio built using my techniques: The punch blocks all have the built-in 50-pin connectors, and there is twisted-pair ribbon connected to each of them with a matching one-operation-crimped connector. The console's ribbon is conventionally wired at the console end, unless the console has DB or other mass-termination-ready connectors, which make its installation even easier. Harmonicas are mounted in each equipment rack or pod or simply under the counter in the vicinity of equipment on the top side and a ribbon cable is run from the various harmonicas back to punch blocks adjacent to those going to the console. The audio source gear is connected to the harmonicas using "pigtails" of either shielded cable, such as Belden 8723, or simply good-quality stranded telephone cord, with the RJ connector crimped on one end and the audio connector soldered on the other. In fact, XLRs and some other common audio connectors are available in Insulation Displacement Connector (IDC) styles, which make it even easier. In some cases, the source equipment has DB connectors, which can be crimped directly onto the ribbon cables (which are easily split into segments of wires), thus avoiding another need for a soldered connection.

Multi-studio installations

Interstudio cables (which are also ribbons) go from the central wiring area in each studio to yet more punch blocks in a main central wiring area, usually in master control or in a dedicated cable-spreading room. At each punch block group, both in each studio and in the main central wiring area, telco-type cross-connect wires are used to actually interconnect the various console ins and outs with the source and destination equipment, and to interconnect the various studios and connect them to the I/O equipment in master control.

Advantages of this system

- Mating a cable onto a punch block takes 10 minutes, and one piece of heat shrink, and one shield wire to deal with. Conventionally, it takes 1-2 hours to terminate a snake to one side of a punch block.

- A set of pigtails can be made up at one time for an entire studio, or even an entire installation, all of the same length, with different connectors on the end away from the harmonica, and all with interchangeable RJ connectors on the other end. The obvious time savings here lie in only having to crimp an RJ connector on one end of each pigtail (a 30-second procedure involving no soldering), and in being able to cut and make all the pigtails at one time, and thus to become efficient at it. However, as mentioned earlier, there is another advantage hiding here: the level of flexibility the pigtails give the engineer in changing equipment....if a new piece of equipment has different connectors from the old one, simply change the pigtail, too, *and leave the rest of the cross-connect alone*. This makes for an incredible time savings when swapping out equipment and extreme flexibility of use of different source equipment.

- These connectors and wiring accessories have been designed to require no soldering and to be user-friendly in many ways that the tried-and-true methods will never be. Unlike the broadcast industry, the telecommunications and computer industries have clearly recognized that more and more technical work is having to be done by fewer, less-qualified people, and also that it is silly to waste true brainpower on tedious and repetitive procedures which nevertheless take an inordinate amount of skill and technical expertise. Clearly, it is time for us in this industry to cast aside our blinders of tradition and convention and embrace these more efficient methods.

- Another, more subtle advantage of an innovative system such as this one is that it tends to keep the wiring relatively clean over the life of the installation. For several reasons, the engineer is not as tempted to jury-rig the installation of a new piece of gear which he or she has to install before the start of the next newscast; most of the infrastructure to wire any new piece of gear is already in place. To wit: all of the console ins and outs are already brought to the punch blocks; because the harmonicas don't have conventional audio connectors on them, the engineer more or less has to use a pigtail for each piece of equipment; and one can make and keep in stock a set of spare pigtails with various connectors on them. With a little foresight in providing extra cross-connects from unused board inputs to unused harmonica ports, an engineer won't even need anything more complicated than a screwdriver to install almost any new or replacement outboard gear in a studio, not even a punch tool!

REAL-WORLD TIME SAVINGS

I have now overseen the construction of 19 studios using these innovative techniques. This is enough to determine an average time savings with some confidence in the accuracy and repeatability of the answer: It has been my experience that using these techniques gives the studio installer a savings of approximately 50 man-hours, or 1/3 or the former total, per studio. Not only that, but the man-hours saved involved mostly skilled but tedious labor, the kind of thing (e.g., punch block termination) that one could not reasonably expect to teach a "grunt" to do without months of background and practice, but which makes a competent engineer feel like he is working way below his level. And, of course, this skilled labor is the scarcest, and most expensive, component of studio construction. Not only can apprentices be taught how to competently strip and crimp IDC connectors much more easily than conventional punch block termination, but much of the labor left in the studio is simply the mechanics of running the wires, drilling the holes, mounting punch blocks, and other tasks that any person who is reasonably familiar with hand and small power tools can do without any further training.

REAL-WORLD PERFORMANCE

My main concern about these innovative studios has been compromised audio performance; specifically, increased crosstalk and induced hum and noise. Having made tests using an Audio Precision Portable One, I was pleasantly surprised to find the crosstalk was greater than 85 dB below +4 dBm on adjacent pairs in my longest run of ribbon cable (about 70 ft.), and there was no measurable crosstalk between non-adjacent pairs. Any trace of induced hum was absent from a line level signal, as well. Indeed, the hum was so low that in one control room, we accidentally ran the main on-air mic-level audio through an unshielded ribbon (due to a mic processor with incorrect settings), and it was almost *three weeks* after moving into the new studio that anyone noticed any hum in the microphone! And this was a *mic-level signal!* The only trace of noticeable crosstalk came from the sync signal of the studio master clock, which is a TTL-level (about 4 V P-P) unbalanced square-wave at an apparent frequency of about 1 kHz. This extremely-high-level signal produced noticeable crosstalk in the adjacent 3 pairs (we put it at one end of a ribbon, for this reason). We found that if we installed balancing transformers at each end, most of the crosstalk went away (only one adjacent pair poisoned), but this distorted the clock signal to where it could not be decoded by the clock slaves at the other end. Based on this, I'm sure that a clock system designed to use a balanced signal like RS-422 would have caused no problems. Short of that, what we eventually did to solve the problem was to move the signal to a ribbon that carried control and tally signals, instead of audio.

CONCLUSIONS

The savings in labor costs are somewhat offset by the increased cost of items such as the connectorized punch blocks, but my experience is that the labor savings are much greater than the additional supply costs. For example, in a recent 4-station bid I provided to a client, the supply costs were increased by about $1200 over the cost of conventional supplies, but the labor costs were reduced by over $6000 from the estimates for a conventional installation, and the time frame was reduced by several weeks. This savings greatly pleased my client, and because of it he was able to afford a much better installation than he had previously thought possible.

For me, the bottom line is that I can hire less-technically-qualified (and therefore cheaper and more easily located) assistants for a studio installation, that I can oversee more of them and still be productive myself, and that even the relatively untrained workers can be considerably more productive using these new, user-friendly tools and procedures than using conventional methods. This translates into lower, more competitive bid prices, fewer headaches, and more profits for my firm, and more reliable and cheaper installations which are more likely than most to stand the test of time for my clients.

Automating the Consolidated Station Operation

Jim Dalke
Mediun Limited
Seattle WA

ABSTRACT

Advances in computer technology can be effectively used to streamline today's consolidated radio station operation. This paper describes implementation of digital technology in a five-station consolidated operation. Digital program automation is networked with traffic operations and digital audio workstations for production and editing. Transmitter remote control and logging have been integrated into the system. The system includes a sophisticated Internet server for streaming audio for each station and provides Internet access to program schedules and other station information.

1.0 PROGRAM AUTOMATION

The core of the radio automation operation is the digital audio program system. While they have been around for a number of years, until recently they have been expensive, cumbersome DOS based systems using proprietary hardware and non-standard software. When design for this system was begun, the primary requirement was for a Windows based software design, using standard off-the-shelf hardware.

After evaluating the possibility of developing the software in house using Windows development tools, the decision was made to use a software package developed by Ron Burley of Broadcast Software International. Benefits of this software package include the familiar graphical look and feel of Windows, and the use of standard hardware interfaces for existing products. This makes the system easy to learn and use. This also means the system uses readily available and cost effective computer hardware.

Other considerations for the system include the ability to import existing computerized traffic logs. This particular system uses CBSI traffic program logs, which require an import program routine to efficiently transfer the log data from the CBSI traffic system to the BSI automation system. The traffic import process requires the use of a standard template for programming day-to-day information with appropriate time slots for specific spots filled in from the CBSI traffic information.

The automation software has provision for external audio routing for satellite delivered programs. There is also a hardware interface provided for utilizing the digital cue signals from the program providers for local breakaways.

The automation system uses standard full duplex PC audio cards. This allows overlap and segues of audio files. All spot announcements and promos, as well as all prerecorded programs are stored on the systems hard drive. These sources are seamlessly integrated with live studio programs and satellite network programs.

The system hardware is based on an Intel Pentium II processor running at 300 MHz with 128 Megabyte RAM and 10 gigabyte Hard Drives.

The system can be operated in a walk-away, fully automatic mode, or with manual live assist.

2.0 DIGITAL PRODUCTION

The digital audio workstation is the input port for the digital radio station. Spot announcements and programs are recorded and produced in an all-digital environment. Windows PC's are used as digital audio workstation, with each production room equipped with a Windows 98 operating system and PC sound cards. The functions of the workstation include editing and routing, recording on hard disk, and processing. With the addition of digital mixing functions, the digital production studio allows most of the production tasks to be handled solely in the digital domain.

The digital audio workstation hardware includes the Windows based computer with both analog and digital I/O. The simplest interface device is the common Sound-Blaster type sound card. A sophisticated interface can include a professional sound card such as the Antex SX-36 or the Digital Audio Labs CardD. The introduction of Creative Labs SoundBlaster Live brings professional PC sound at consumer prices.

Each of the cards has advantages and disadvantages. The SoundBlaster Live performance seems to rival the much more expensive cards in performance and the only drawback is the lack of balanced analog I/O that is standard on the other professional cards. The Antex has an on board DSP for hardware compression and decompression. The CardD and the SoundBlaster Live cards have digital I/O facility, which allows direct digital transfer of sound files from DAT, CD, and Minidisc.

There is an impressive array of software available for the Windows based digital workstation. The software packages used in the system take advantage of the intuitive Windows graphical interface, providing easy to learn, efficient interactive operation. The primary function of the editor software includes cut and paste operations, time compression and stretching, level normalization, and file format conversion.

Which editor is used depends on the complexity of the task. For quick edits that may include cut and paste, trimming beginning and end, and normalizing, an audio editor program from Minitonka Software called FastEdit is used. FastEdit has a useful feature, which allows limited editing of MPEG files.

For more sophisticated editing for production purposes, Sound Forge is a full Featured editing software package from Sonic Foundry which includes some very sophisticated production tools including time compression which can change a 36-second spot announcement to 30-seconds without apparent artifacts. Sound Forge has a sophisticated Graphic and Multi-Band Dynamics process that can add apparent fullness to an entire file. Sound Forge also allows file sample rate conversion, as well as stereo-mono conversions.

Another relatively low cost, audio edit program is CoolEdit from Syntrillium Software. Among its useful features is the ability to compress and stretch, provide noise reduction as well as file type and sample conversion.

The system sound files normally use a standard 32-kilobyte per second sample rate with 16 bits per sample in the Microsoft .wav format. For long program storage, MPEG compression allows more efficient use of hard drive space. The problem using MPEG compression on all files is the difficulty in editing and using the segue feature of the program automation software.

The final transition into the all-digital radio station will be the addition of the digital mixer. The digital mixer is a complex device requiring a great deal of digital processing power. The concept of converting microphone signals and other analog signals to a digital format at the source reduces the signal degradation that occurs in the analog mixing world, including noise and distortion. It is now possible to efficiently perform the basic mixing console functions of attenuating, mixing and signal routing all in the digital domain.

The digital audio workstation has been installed on the desktop PC in several production staff offices. This allows audio recording to hard disk and editing on the desktop, freeing the production facilities for live recording requiring microphones. The desktops are equipped with cassette players, DATS and minidisks as well as the internal desktop CD player, allowing recording and editing of spots and programs as needed.

3.0 NETWORKING

All of the station computers are linked through a local area network or LAN. The decision to use Ethernet was based on availability and economy. Ethernet was developed for the transfer of simple data and not large audio files streaming audio. To overcome some of these problems, the system uses switched 100 MHz Ethernet to link the critical paths in the operations area to provide as much bandwidth as possible. The switched network guarantees a path between two nodes of the network allowing the full 100 MBPS bandwidth for audio file transfer from computer to computer.

The administrative portion of the LAN, including management, traffic, and sales uses the more conventional and economical 10 MHz hub type network.

Audio servers have been implemented in the network to allow the automation systems to access all of the commercial and program files. The server uses Windows NT Server, Enterprise edition with clustering capabilities. The clustering technology is based on industry standard hardware providing incremental growth and uninterrupted operation. The system provides load balancing or sharing between computers while maintaining duplicate files for reliability.

The LAN has been extended to link the transmitter sites with Telco provided T1 links. By placing routers on the studio network hub, and at the transmitter site, the Ethernet has been extended. Audio can be streamed either direction for program delivery as well as satellite audio feeds from the transmitter sites. Transmitter control and monitoring uses the same network path.

4.0 TRANSMITTER OPERATIONS

Transmitter control and monitoring has been integrated in to the automated system. The design objective of the control and monitoring system was to meet unattended transmitter operation requirements of the FCC.

The transmitter remote control uses the ARC16, manufactured by Burk Technology. The studio unit is linked to the transmitter via the T1 circuit provided by the local Telco. The studio unit is connected through a serial link to the transmitter control PC. Burk's AutoPilot for Windows has been installed on the computer, which in turn is connected to the station Ethernet network. This allows remote monitor and control of each transmitter from any workstation on the network.

The AutoPilot control software allows automatic day-night antenna changes as well as parameter monitoring, with alarm notification for alerting station personnel in case of out-of-tolerance operation.

5.0 LOGGING AND ARCHIVING

The system provides extensive digital archiving of program logs, event logs, and transmitter logs. The system also includes provision for archiving actual off-the-air audio. The system uses the latest DVD optical storage technology for archiving.

Once a day, the system automatically retrieves the program log and event log from the automation computer. (The event log is a list of events and time they occur.)

The system also retrieves the transmitter operating logs from the transmitter control computer for archiving.

The system records audio from the air monitor using a compression algorithm with brings the stream data rate to about 6kbps. A 24 hour recorded audio file is written to the hard disk each day. Once each month the files stored on the hard disk (totaling about 2 gigabytes) are transferred to the DVD recorder. This allows anyone with appropriate access to the LAN to listen to actual recorded air audio at any time during the current month from the network, and by retrieving from the DVD any audio aired since the system was implemented.

6.0 INTERNET

To provide Internet connectivity, the system includes a Windows NT Server with Microsoft's Internet Information Server software. The server is connected by T1 for Internet access, and to the LAN.

It provides all the station PCs and workstations full time Internet access and allows Web services to be provided for station promotion as well as for client advertising.

This also allows streaming each stations air signal to the Internet. A separate workstation provides a platform for the audio encoder. The workstation has a sound card installed for each station. An off-the-air audio signal is provided to each sound card and NetShow encoder session is started for each encoder.

The Web Server then redirects the stream to the Internet user requesting a feed through the station Web site.

The Web server provides Web services for each station with information about the station as well as a program guide. Microsoft's FrontPage editor provides web page editing.

The combination of the Windows NT server with the Microsoft Internet Information Server delivers a well-managed, high-speed, secure information publishing solution for the stations.

A separate PC has been established on the network as a Mail Server. This allows all employees, and well as selected clients to have E-mail services.

7.0 FUTURE

The system in place today is first step into the future. The system was designed to accommodate changes as rapidly as technology allows.

It will be advantageous to have the entire audio path from microphone to transmitter all digital. More satellite network feeds are using digital program distribution. Commercials are being distributed on ISDN and Internet. Other content is provided on CDs and DATs. Two of the transmitters being fed audio from the automation system are Harris DX series AM transmitters, which use an A-D converter for digital modulation. The entire path can remain all digital to deliver the best signal possible signal to the listening audience.

The old hybrid telephone system (PBX) will be replaced soon with a all digital Network PBX. This digital telephone system is based on a Windows NT server, and will provide services well beyond cumbersome existing system. The NT server is connected to telephone sets and incoming digital trunks through specialized interfaces. The system emulates the traditional telephone with all of the button pushing managed by a Windows graphic interface on the workstation monitor.

All of the features of the traditional PBX such as Auto Attendant, Voice Mail, Call forwarding will be handled digitally.

This digital PBX will be a particular advantage for the call-in talk show. The digital trunks from the Phone Company inherently have separate send and receive audio signals eliminating the need for hybrids in putting the caller on the air.

ATM technology will supplement the traditional TCP/IP packet oriented network, making it far more efficient to move digital audio files on the network.

8.0 CONCLUSION

It is a fast pace technological world we live in today. The power and speed of the computer processor is increasing at a staggering rate while costs plummet. Mass storage capacity, whether it is random access memory chips or hard disk drive is increasing geometrically and the price per megabit decreases. We can do things in electronic media today we could not dream of yesterday. By implementing this technology in our radio world of communication and entertainment we can provide an increasingly attractive product at a very effective cost.

This approach provides a gateway to today's internet broadcasting and tommorrow's over-the-air digital broadcasting.

2.4GHz Spread Spectrum Radio and T1 Technology Combine For A Wireless T1 STL/TSL Solution

By

David W. Agnew
FM Applications Engineer
Harris Corporation, Broadcast Systems Division
3200 Wismann Lane
Quincy, Illinois 62301
Phone: (217) 221-7399
Fax: (217) 224-7480
E-Mail: dagnew@harris.com

Abstract

Spread spectrum technology dates back to World War II. The US military has used Spread spectrum technology for the last 25 years. Today spread spectrum is used in cellular, PCS and other types of data networks. Spread spectrum radios in the 2.4GHz range provide reliable point-to-point wireless links without FCC license requirements.

Spread spectrum and T1 STL/TSL have combined to make possible bi-directional T1 radio links of up to 25 miles. These systems are capable of transmitting uncompressed AES3 digital audio, data for remote control systems, LAN interconnection and voice channels for PBX extensions.

This paper covers applications methods, path reliability analysis and cost benefits for commercial radio broadcast facilities.

Introduction

Spread spectrum technology originally used by the military is now being used in many commercial applications such as point-to-point wireless communications. When operating under Part 47 Section 15.247 of the FCC Rules and Regulations licensing of these radios is not required. This makes the implementation of these systems elementary and cost effective. In part, licensing is not required due to the characteristics of the spread spectrum signal. The low power density of the transmitted signal prevents interference to narrow band systems and the de-spreading characteristic of the receivers spreads a narrow band interference signal, preventing destructive data degradation. The receiver inherently can reject co-channel spread spectrum signals due to specific coding of the transmitted signal. The receiver will simply reject any data that is not coded correctly.

Affordable spread spectrum radios operating under Part 15.247 in the 2.4GHz range are now available and capable of transmitting full duplex T1 data formats. Wired T1 STL/TSL systems have been in use by AM and FM broadcasters since 1980. Many times these systems were chosen over a typical 950MHz STL due to their ability to transport larger amounts of data, the capability to be bi-directional and the capability of transporting uncompressed AES3 audio. Now all of these benefits and more are available in a wireless radio link without ongoing T1 toll charges. This becomes particularly attractive in larger markets where the 950MHz band is increasingly more congested and frequency coordination is more difficult.

History Of Spread Spectrum

Spread spectrum radios are relatively new in STL/TSL applications in the commercial broadcast industry. However, spread spectrum techniques date back to use by the United Sates Military in World War II. Our allies also experimented with spread spectrum techniques as well.

Initial applications of spread spectrum were used to provide countermeasures to radar, navigation communication and regular communications. An example of an early application of spread spectrum is the Magnavox USC-25 modem. Taking up at least three six foot racks, It had a maximum data rate of 64Kbps and occupied 60MHz of bandwidth.

In 1942, Hedy Lamar (married name Hedy Kiesler Markey) the Hollywood celebrity from the 1930's to the 1950's, along with co-inventor George Antheil were awarded a U.S. patent #2,292,387 for a "Secret Communication System". This spread spectrum communications system was based on frequency hopping techniques. The invention was given to the U.S. Government at no charge. A piano roll synch scheme was used to synchronize the transmitter and receiver frequency hopping sequence for secure radio control torpedo guidance systems.

In 1957, Sylvania replaced these types of systems with early computer processor based systems. This led to the development of new secret military communications systems. Spread spectrum has since provided the military with reliable, secure, jamming resistant communications. The U.S. military has continued to develop and make use of spread spectrum for the last 25 years.

Regulatory Issues

In 1948, the 2.4GHz band was made available for use by industrial, scientific and medical (ISM) services. As one of eight original ISM bands, the 2.4GHz band was used for such things as microwave ovens and other special RF devices. These devices were FCC type accepted but did not require licensing by the end user. Microwave radio links were not one of these applications.

Before 1985, spread spectrum use was limited almost entirely to the military and ISM services. In 1985, the FCC modified the rules to permit commercial use of spread spectrum techniques. The same attributes that caused the military to use spread spectrum is now available to the broadcast industry. Today, spread spectrum is used for personal communication systems (PCS), personal communication networks (PCN), wide area networks (WAN), local area networks (LAN) and unlicensed point-to-point radio links.

The FCC Rules and Regulations that apply to the unlicensed 2.4GHz band can be found in:

- Spread Spectrum Systems, Part 15.247
- ISM Equipment, Part 18
- Amateur Radio Service Part 97

The FCC Code of Federal Regulations Part 47 Section 15.247 sets the maximum power output of a spread spectrum transmitter at 1 watt (+30dBm) for frequency hopping systems in the 2.4GHz and 5.7GHz bands and for all direct sequence systems. On April 4th, 1997 Part 15.247 was amended to set a practical limit on the effective isotropic radiated power (EIRP) by imposing a new 3-for-1 rule. This rule is exclusive to 2.4GHz spread spectrum point-to-point links. An application of this rule will be demonstrated in the RF Link section of this paper.

The regulations of wireless radio links in the 2.4GHz band varies widely from one country to another. It is therefore, important to investigate the local rules and regulations related to the 2.4GHz wireless radio emissions before implementing a point-to-point radio link. Bandwidth, emission restrictions, maximum transmitter power output, maximum effective radiated power and licensing requirements should all be considered.

Spread Spectrum Primer

Spread spectrum is not a conventional type of modulation as in AM or FM broadcasting. Spread spectrum is a modulation technique that spreads the information data and power density over a wide bandwidth and is then de-spread and demodulated in a spread spectrum receiver.

One way of viewing spread spectrum is, that it trades a wider bandwidth for an increased signal to noise ratio. In order to be considered a spread spectrum system at least two technical specifications must be satisfied: First the total occupied bandwidth must be wider than the bandwidth of the information being transmitted; and second some type of coding data of a specific pattern other than the data being transmitted must be added to the information data.

The most basic and indispensable element of spread spectrum is the craft of spreading the information over a wide bandwidth, transmitting the expanded or spread signal, then receiving this signal and de-spreading the information back into its original form. This method of modulation can be equated to another method that trades a higher bandwidth for a lower bit-error-rate. This technique is known as forward error correction (FEC). In FEC redundant bits are added to a digital signal before it is transmitted and then stripped or used to replace corrupted data when demodulated. FEC nearly meets the requirements for spread spectrum and therefore could be considered a form of spread spectrum. The difference being the coding bits in spread spectrum are totally independent of the information data. Wideband FM broadcasting can be considered a form of spread spectrum in that the bandwidth required is wider than the information being transmitted.

Many spread spectrum systems use special pseudo-random noise (PN) codes to spread the transmitted signal. The length of these codes will ultimately determine the final bandwidth of the signal.

Today's commercial spread spectrum systems use bandwidths 10 to 100 times that of the information bandwidth. Military systems have used spectrum bandwidths from 1,000 to 1,000,000 times the information bandwidth.

A spread spectrum receiver will de-spread the received signal using a locally generated replica PN code and a receiver correlator allowing the separation of only the desired coded information from all possible signals. A correlator can be thought of as a special matched filter. This filter will only respond to signals that are encoded with a PN code that matches the locally generated code. The correlator will not respond to other types of noise or interference. Most commercial spread spectrum radios have several different code sequences to select from, allowing selective addressing. The use of different code sequences results in low cross correlation between receivers. This helps ensure that receivers operating with different codes will only be reached by a transmitter using an identical code sequence. The low power density of the spread spectrum signal aids in preventing interference to narrow band systems. And in spread spectrum receivers, rejection of narrow band signals is accomplished by the de-spreading process.

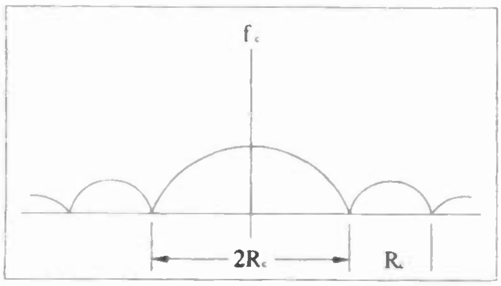
Several methods are used to accomplish the spreading and de-spreading of the RF carrier such as:

- Direct Sequence Spread Spectrum (DSSS)
- Frequency Hopping Spread Spectrum (FHSS)
- Pulsed FM (Chirp)
- Time Hopping

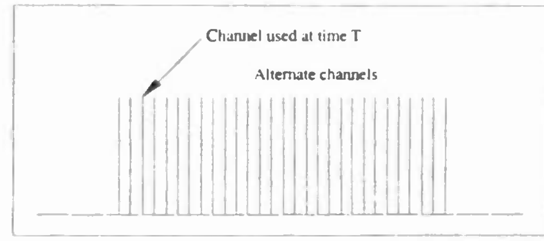
The two most common methods used in commercial transceivers are direct sequence spread spectrum (DSSS) and frequency hopping spread spectrum (FHSS). Of these two DSSS tends to be the most popular. Pulsed FM (chirp) systems are used for radar applications and time hopping techniques are reserved for military use.

In DSSS a narrow band carrier is modulated using a pseudo-random noise or PN code mixed with, but independent of the information data. This code has a fixed length and continuously repeats. The carrier phase of the transmitted signal is changed in accordance with the PN code sequence. The bandwidth of a DSSS system is directly related to the chip rate. The chip rate (Rc) is normally measured in Mega chips per second (Mc/s) and in some systems is the product of the QPSK RF symbol rate and the number of chips in a code sequence. For example in T1, an RF symbol rate of .772 Mbaud and a code length of 15 chips would have a chip rate of 11.58Mc/s. The chip rate is also used to determine other specifications such as the processing gain of a particular system. In a DSSS receiver the received code synchronizes with an identical locally generated PN code, de-spreads the RF carrier and recovers the information data.

The diagram in figure 1 is an example of an ideal DSSS power spectrum. The null-to-null bandwidth is equal to two times the chip rate which is determined by the type of PN code being used. These codes and code rates vary widely from system to system.



(Fig. 1)



(Fig. 2)

In FHSS systems the carrier frequency of the transmitter changes or "hops" in accordance with the PN code used. Frequency hopping modulation could be more precisely termed "multiple-frequency, code-selected, frequency shift keying". It is simply FSK (frequency shift keying) except that the set of frequency choices is greatly expanded. Simple FSK uses only two frequencies, f1 is sent signifying a "mark," and f2 to signify a "space". Frequency hopping systems can have thousands of frequencies available.

In the FHSS receiver the locally generated PN code is synchronized with the received PN code to produce a constant IF signal. The IF then passes through a filter and is demodulated. If the codes do not match the IF filter will reject the signal and demodulation will not occur.

The diagram in figure 2 is an example of an ideal FHSS power spectrum.

Hybrid systems using a combination of DSSS and FHSS also exist and make use of the advantages of both systems.

An attractive feature of spread spectrum systems is their ability to reject or resist interference and jamming. Two common parameters specified to measure this ability are Processing Gain and Jamming Margin. Processing Gain is defined as the gain or signal-to-noise (S/N) improvement by a spread spectrum system due to the coherent band spreading and de-spreading of the desired signal. Processing gain can be estimated by the ratio of the RF emmision bandwidth to the information symbol rate of the information data being transmitted. Processing gain can be found by the following:

$$Gp = 10 \log BW_{RF} / Rinfo$$

Where:

Gp = Processing gain in dB
BW_{RF} = Emmission bandwidth
R info = Symbol rate

The processing gain is not a measure of the amount of interference a particular spread spectrum receiver can tolerate. Another specification known as jamming margin is used for this purpose. Processing gain will directly affect the jamming margin and the jamming margin is always less than the processing gain.

All radio receivers specify some type of jamming margin. In most cases jamming margin is specified as selectivity or adjacent channel rejection. Only receivers that are expected to operate when the unwanted signal is on the same frequency is the term jamming margin used.

Jamming margin is defined as the amount of interference a system is able to withstand while producing the required output signal-to-noise ratio or bit-error rate. Jamming margin can be found by the following:

$$Mj = Gp - [\, Lsys + (S/N)out \,]$$

Where:
Gp = Processing gain in dB
Lsys = System losses in dB
(S/N)out = Signal-to-noise ratio
at the information output in dB

For example a system with 18dB of processing gain, a minimum signal to noise ratio of 13dB and 2dB of system losses, would have a 3dB jamming margin. Therefore it could be assumed the receiver would not operate at a specified BER with an interfering signal greater than 3dB above the desired signal.

These specifications as well as others will give some idea of the systems robustness and are also used to ensure FCC compliance as specified in Title 47 Section 15.247 of the Code Of Federal Regulations.

Some advantages and appealing characteristics of spread spectrum systems are:

- Selective addressing capability
- Co-channel spread spectrum Interference rejection due to coding techniques
- Low power density
- Rejection of narrow band interference

Disadvantages include:

- Wide occupied bandwidth
- Long acquisition times with long PN codes

Spread Spectrum RF Links

The propagation characteristics of a 2.4GHz spread spectrum point-to-point radio link do not differ a great deal from 950MHz links. As shown spread spectrum systems have the ability to reject interference better than a narrow band system. Also spread spectrum links can tolerate more multipath than narrow band systems due to the wide bandwidth and coding techniques used. The .6F1(first Fresnel zone radius) with a K factor of 4/3rds earth curvature are used in spread spectrum link studies.

Typically low loss foam coax such as Andrew LDF series or equivalent 1/2 inch to 2 1/4 inch having losses from 3.5dB to 1dB / 100 ft. are adequate. FSJ series cable is not recommended due to the high losses in the 2.4GHz region. Elliptical wave guide can be used with losses less than 0.5dB per 100 feet however the cost of this type of cable and connectors is prohibitive and can cost as much as 5-7 times more than LDF series cable and connectors.

Antennas for 2.4GHz spread spectrum systems are very similar to those used in 950MHz STL systems. These grid antennas vary in size from 3 to 15 feet with specified gains from 24dBi to 39dBi. There are also some appealing planar array antennas available, which have specified gains from 20dBi to 28dBi. These antennas can be used where aesthetics and zoning restrictions are an issue.

Free space attenuation at 2.4GHz is slightly greater than 950MHz systems. This will restrict a reliable path length to between 25 and 30 miles. The 3 for 1 rule that is specified in part 47 section 15.247 applies to 2.4GHz spread spectrum point-to-point links. This can add some flexibility to the design of a spread spectrum point-to-point link.

The following example uses an existing, operating, 950MHz RF link and replaces it with a Harris Aurora 2400 transceiver. The Aurora 2400 is a DSSS transceiver with a maximum power output of +26dBm. This example illustrates how the 3 for 1 rule can be used to increase the EIRP and increase the flat fade margin. Note that we must change both the transmit and receive antennas because the system is bi-directional. The 3 for 1 rule states, that for each 3dB the antenna gain exceeds 6dBi the power output of the transmitter into the antenna must be reduced by 1dB.

The fade margin of the Harris Aurora 2400 is recommended to be 25dB or about 1dB per mile in order to provide an acceptable reliability percentage at a specified BER of 10E-3. In example 1, the system is using an antenna with 24dBi of gain. Under the 3 for 1 rule, taking into account the 2dB of transmission line loss, the transmitter can operate at +26dBm. The following formula illustrates the 3 for 1 rule calculations. Note that 2dB of power was added to make up for the transmission line losses.

$$Pt = (Pa - (Ga - 6)/3) + Lt$$

where: Pt = Transmitter output power
Pa = Max. antenna input power (30dBm)
Ga = Antenna gain in dBi
Lt = Transmission line loss

$$Pt = (30 - (24 - 6)/3) + 2 = 26\text{dBm}$$

Microwave Radio Path Calculations

Customer: WXYZ-FM			Project: NAB99
Equipment Type: Aurora 2400			Capacity: 1 T1
Protection Type: Non Protected			Mid Frequency: 2.442GHz
Transmit Power: High (+26dBm)			
Site Name:	**STUDIO**		**TRANSMITTER**
Latitude (N):	40 ° 49 ' 12.0"		40 ° 31 ' 6.0"
Longitude (W):	96 ° 39 ' 29.0"		96 ° 46 ' 7.0"
Ground Elevation:	1247 ft		1499 ft
Path Length:		21.6 miles	
Azimuth:	195.63 °		15.56 °
Free Space Loss:		131.0dB	
Absorption Loss:		0.2dB	
Feeder Type:	1/2 Foam		1/2 Foam
Feeder Length:	50.0 ft		300.0 ft
Feeder Loss:	1.9dB		11.5dB
Miscellaneous Loss:	0.0dB		0.0dB
TX ACU Loss:		0.0dB	
RX ACU Loss:		0.0dB	
TOTAL LOSSES:		**144.6dB**	
Antenna Size:	3.0 ft		3.0 ft
Antenna Centerline:	0.0 ft		0.0 ft
Antenna Gain:	23.5dB		23.5dB
Transmit Power:	26.0dBm		26.0dB
TOTAL GAINS:		**73.0dB**	
Unfaded Receive Signal Level:		-71.6dBm	
Threshold:		-91.0dBm at 10E-3 BER	
Flat Fade Margin:		**19.4dB**	
Reliability (%):		**99.9559%**	
c Factor: 1.0			
Avg. Annual Temp.: 50C			

Example 1

In example 1, we see that the recommended fade margin of 25dB was not met using an output power of 26dBm with an antenna gain of 24dBi.

The next example increases the gain of the transmit and receive dish from 24dBi to 33.1dBi. The following calculations are used to determine the correct transmitter power for the studio site transmitter. The transmitter site transmitter would be left at its maximum power of 26dBm due to the 11.5dB of line loss.

$$Pt = (Pa - (Ga - 6)/3) + Lt$$

where: Pt = Transmitter output power
Pa = Max. antenna input power (30dBm)
Ga = Antenna gain in dBi
Lt = Transmission line loss

$$Pt = (30 - (33.1 - 6)/3) + 2 = 23 \text{ dBm}$$

Microwave Radio Path Calculations

Customer: WXYZ-FM		Project: NAB99	
Equipment Type: Aurora 2400		Capacity: 1 T1	
Protection Type: Non Protected		Mid Frequency: 2.442GHz	
Transmit Power: High (+26dBm)			
Site Name:	**STUDIO**		**TRANSMITTER**
Latitude (N):	40 ° 49 ' 12.0"		40 ° 31 ' 6.0"
Longitude (W):	96 ° 39 ' 29.0"		96 ° 46 ' 7.0"
Ground Elevation:	1247 ft		1499 ft
Path Length:		21.6 miles	
Azimuth:	195.63 °		15.56 °
Free Space Loss:		131.0dB	
Absorption Loss:		0.2dB	
Feeder Type:	1/2 Foam		1/2 Foam
Feeder Length:	50.0 ft		300.0 ft
Feeder Loss:	1.9dB		11.5dB
Miscellaneous Loss:	0.0dB		0.0dB
TX ACU Loss:		0.0dB	
RX ACU Loss:		0.0dB	
TOTAL LOSSES:		**144.6dB**	
Antenna Size:	8.0 ft		8.0 ft
Antenna Centerline:	0.0 ft		0.0 ft
Antenna Gain:	33.1dB		33.1dB
Transmit Power:	23.0dBm		26.0dBm
TOTAL GAINS:		**89.2dB**	
Unfaded Receive Signal Level:			-55.4dBm
Threshold:			-91.0dBm at 10E-3 BER
Flat Fade Margin:			**35.6dB**
Reliability(%):			**99.9989%**
c Factor: 1.0			
Avg. Annual Temp.: 50C			

Example 2

In this second example we now exceed the 25dB fade margin goal and achieve an availability of 99.9989% which surpasses most T1 service specifications. Typical T1 leased line reliability is specified at 99.9800%. The transmitter site transmitter power will again remain at 26dBm due to the 11.5dB line loss. This results in a 38.6dB fade margin for the return path.

Some commercially available spread spectrum transceivers can be configured with a PC via an RS-232 port, allowing PN code selection, power level adjustments and other software selectable features. Spread spectrum communications does offer a certain amount of increased robustness over narrow band systems. However, this does not replace good engineering practice when implementing these systems. The link should still be studied for proper path clearance, signal levels and fade margins. Interfering sources should also be identified and steps taken to prevent threshold degradation. In most cases, cross polarizing, using larger antennas with narrower beam widths and higher front to back ratios and selection of PN codes will help resolve most site problems.

Combining a 2.4GHz Spread Spectrum Radio and T1

Initially, radio STL systems were limited to transmitting analog left/right, stereo composite and one or two SCA channels. Recently Digital STL systems have allowed the transmission of compressed and uncompressed AES3 digital audio, low bandwidth auxiliary audio channels and RS-232 data channels.

Since about 1980 STL/TSL systems have been utilizing T1 phone lines. However these systems have come at great expense due to installation fees and ongoing monthly toll charges. While the expenses have been coming down in cost, it is still an ongoing operating cost that may be unnecessary. The T1 STL/TSL systems allow bi-directional transmission of 1.544Mbps (T1) of data. T1 STL systems have enjoyed greater versatility and increased capacity over 950MHz STL's. If a line of sight path (up to 30 miles) is available, these TI based STL's may be coupled to 2.4GHz spread spectrum radios, eliminating monthly T1 lease charges. A typical return on investment of 3-5 years could be expected. This of course depends on what hardware is needed and the cost of the final configuration of the STL/TSL system.

Many different configurations are currently available and offer great flexibility and choice of configuration. Some typical options available from popular T1 STL/TSL systems are:

- Compressed AES3 digital audio
- Uncompressed AES3 digital audio
- Compressed 7.5/15kHz analog audio
- Uncompressed 7.5/15kHz analog audio
- Voice modules for PBX extensions
- LAN/WAN interface modules
- Synchronous data channels
- Asynchronous data channels

The block diagram in figure 3 is an example of some of the capabilities of a Harris Intraplex system.

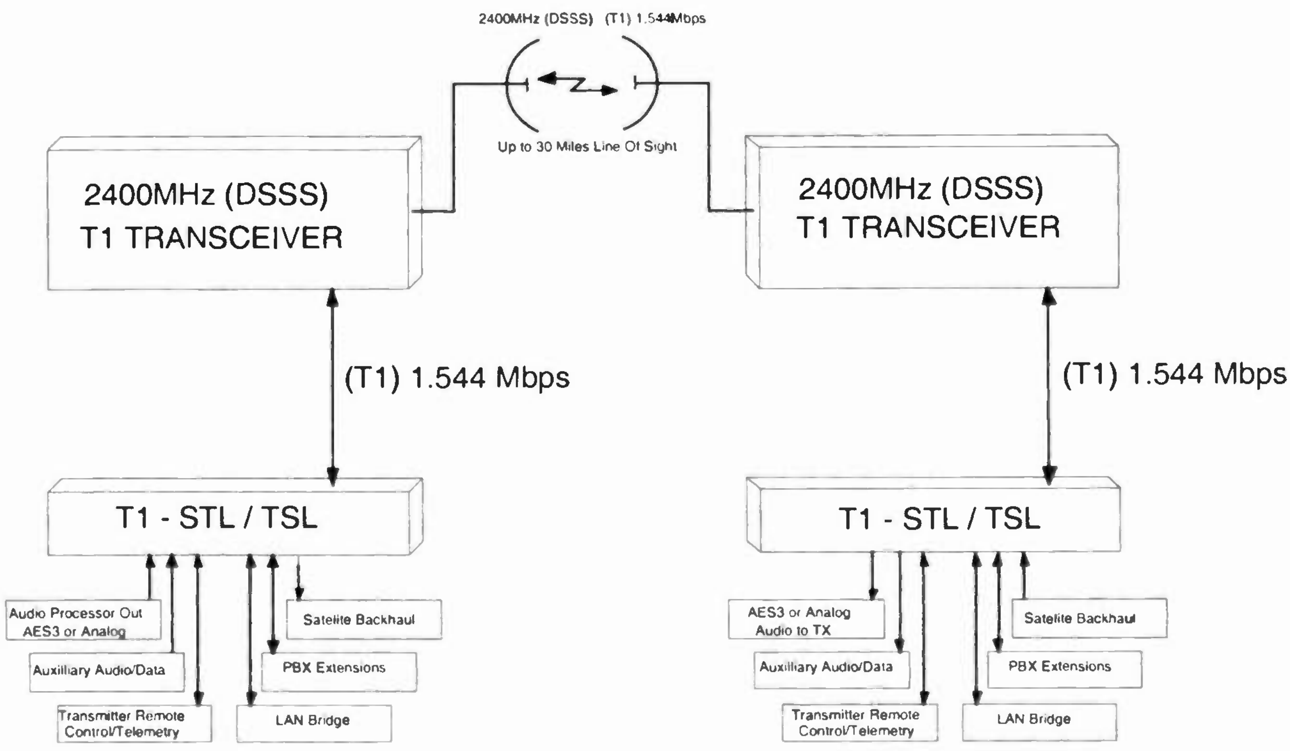


Figure 3

Another application of 2.4GHz spread spectrum radios is for dedicated LAN/WAN connections between two intercity studio sites or offices. A few stations are using this configuration to allow the transfer of hard disk based digital audio data, as well as traffic, billing and payroll information between co-owned stations. Spread spectrum is ideal for this application due to the secure nature of the signal. It is very difficult to intercept spread spectrum due to the wide bandwidth, low RF power level, and coding techniques. It is also possible to use these systems for remote broadcasts, where a line of site or near line of site path is available. This opens up a number of excellent options for remote broadcasting.

Conclusions

- Spread spectrum techniques have a proven track record as a reliable and secure method of transporting digital data in the U.S. military for the past 50+ years.

- Spread spectrum radios operating in the 2.4GHz band are now available for dedicated point-to-point links of up to 30 miles line of sight, and do not require FCC licensing.

- These new 2.4GHz spread spectrum radios are capable of transporting T1 data and are compatible with current T1 based STL/TSL systems.

- Coupling these 2.4GHz radios with T1 based STL/TSL systems can replace costly T1 line leases, and can also be used as back-up to an existing T1 system.

- 2.4GHz spread spectrum radios are also capable of other dedicated connections such as LAN/WAN, as well as temporary installations such as remote broadcasts where line of sight or near line of site links are possible.

Acknowledgments

The author thanks:

Daryl Buechting, FM Product Line Manager, Harris Broadcast Systems Division

Richard Fry, Senior FM Applications Engineer, Harris Broadcast Systems Division

Tim W. Dittmer, Manager, Development RF Power and Digital, Harris Broadcast Systems Division

Mauro Di Lorenzo, Systems & Transmission Engineer, Harris Corporation, Microwave Communications Division

References

1. Dixon, Robert, C., "Spread Spectrum Systems with Commercial Applications", Third Edition, 1994.

2. Laine, R.U., Lunan A.R., Shaw, Walt, "Aurora 2400 Spread Spectrum Digital Radio For Unlicensed T1/E1 Transport in the 2.4GHz Band Applications Note", Harris Corporation, Microwave Communications Division.

3. RF Path Analysis Calculations by: Harris Corporation, Microwave Communications Division, Starlink© Microwave Path Analysis Software.

4. Jachimcyk, Witold, "Spread Spectrum"

DTV Transmission Systems

Tuesday, April 20, 1999

1:00 PM - 5:35 pm

Chairperson: Jerry Whitaker
Technical Press, Morgan Hill, CA

1:00 pm Strategic Development of New DTV Facilities and RF Systems
William Booth
Quality Engineering, P.A.
Gastonia, NC
Robert Denny
Denny & Associates, P.C.
Washington, DC

1:25 pm WBTV Field Experience, The First Full Power 1,000,000 DTV ERP Station
Andy Whiteside
Comark Division of Thomcast Communications, Inc.
Southwick, MA

***1:50 pm DTV/NTSC Co-locate or Not to Co-locate**
Dennis Heymans
Micro Communications Inc.
Manchester, NH

2:15 pm Stabile High Power Filters for the FCC DTV Emissions Mask
Derek Small
Passive Power Products
Gray, ME

2:40 pm Bandpass Filter and Linearization Requirements for the New FCC Mask
Robert Plonka
Harris Corporation/Broadcast Division
Quincy, IL

***3:05 pm Field Results of a New High Power DTV Emissions Mask Filter**
Kerry Cozad
Dielectric Communications
Raymond, ME

3:30 pm On-Channel Repeaters for Digital Television - Implementation and Field Testing
Walt Husak
Advanced Television Technology Center
Alexandria, VA

3:55 pm A Low Beam Voltage, High Power, TV Broadcast IOT Amplifier System, with an Optional Multistage Depressed Collector
Timothy Crompton
EEV, Ltd.
Chelmsford, England, United Kingdom

4:20 pm Architecture of a DSP Based Dual-Mode ATSC/NTSC Television Exciter and Transmitter
David Hershberger
Continental Electronics, Inc.
Grass Valley, CA

4:45 pm Performance of a Multi-Program MPEG Encoding System in Constant Bit Rate and Statistical Multiplexing Modes
Elliot Linzer
DiviCom
White Plains, NY

5:10 pm Modulation Schemes for Transmission of DTV Signals Over Satellite
Girish Chandran
Tiernan Communications, Inc.
San Diego, CA

*Papers not available at the time of publication

FACILITY DESIGN AND MODELING FOR DTV TRANSMITTER SYSTEMS

William I. Booth, P.E.
Quality Engineering, P.A.
Charlotte, NC

Robert W. Denny, Jr., P.E.
Denny & Associates, P.C.
Washington, DC

INTRODUCTION

Virtually every television station in the United States now faces the challenge of constructing a new DTV transmitter plant. Few engineers have the luxury of starting this process from scratch. Most have to decide how to install a new DTV transmitter plant at, in, and around their existing NTSC transmission facility. DTV deployment problems arising from inadequate building space, electrical power distribution system capacity, available tower aperture, and experienced technical staff often can be overcome through careful planning and use of modern computer-assisted design (CAD) tools. In this paper, we explain our design philosophy and present three case studies which demonstrate the efficiencies afforded by its use.

DESIGN PHILOSOPHY

Over two years ago, when we started preparing the standard architectural and engineering permit documents for our first DTV facility that was to be collocated with an existing NTSC facility, we did not plan to use three-dimensional (3-D) CAD techniques. Early in the design development process, the number of drawing layers used in the floor plan grew to over 40. We knew we had a problem before we added the first piece of RF transmission line to the plans.

The design was further complicated by the standard layout of the interior RF system. A 40-foot wide RF system had to fit into a 36-foot wide space. In addition, the RF system and the DTV transmitter had to fit into the back corner of the existing transmitter room already cluttered with existing HVAC system components, electrical conduits, and NTSC transmission system components. The DTV RF system had to be installed overhead, so a steel supporting structure also needed to fit in the same limited space. Do any of these dilemmas sound familiar to you?

In short course, it became painfully obvious to us that standard floor plans and elevations were not sufficient to ensure that, for example, a DTV combiner reject load would not bang into the NTSC aural feed to the diplexer.

We learned two important things from our early experiences. First, coordination of all disciplines, including electrical, mechanical, structural, <u>and</u> RF transmission, is absolutely imperative at the design phase to avoid serious conflicts later during the construction phase. Second, 3-D modeling can be used successfully to assure that the new DTV RF system can be installed at the transmitter site as it was designed, fabricated, tested, and approved by the manufacturer at the factory. This second factor is extremely important given the performance requirements mandated by the FCC, particularly with respect to the DTV emission mask specification. While an RF system usually can be reconfigured in the field, it is a costly, time-consuming process that can be avoided by adopting our design philosophy.

In the three case studies that follow, we discuss the addition of new DTV transmitting facilities at existing NTSC sites. All three cases clearly fall into the category of stuffing 10 pounds of transmitter into a 5-pound transmitter building. We will show some of the more complex problems we have encountered, and how, in each case, careful planning and engineering have resulted in the best possible DTV system given the station's budget and deployment schedule.

WBTV-DT

The WBTV-DT RF system is a four-tube design that includes three magic-T combiners and two mask filters. The RF system which includes the transmitter, the combining system, and the mask filters is physically quite large, even when compared to the size of a conventional high-power UHF analog system with the same number of tubes.

In laying out the DTV portion of the WBTV transmitter room, traditional two-dimensional architectural drawings like that of Figure 1 soon became confusing making it difficult, if not impossible, to detect potential space conflicts. To remedy this problem, we built a 3-D library of waveguide devices and imported them into our CAD system. This enabled us to generate 3-D views of the space including the new DTV transmission system. We were then able to move system components around in 3-D space to find the optimal layout of the system comprised of the components dictated by the RF design criteria. We are now able to identify conflicts readily within the RF system like the one shown in Figure 2. Once the electrical design of the RF system is complete, we place it into 3-D space and verify and correct any conflicts between the RF system and the other disciplines such as HVAC or structural as shown in Figure 3.

The final design of the WBTV-DT RF system is shown in Figure 4. This and other similar documents were provided to the manufacturer and the station. Drawings like that of Figure 4 enabled the manufacturer to fabricate the system at the factory in the exact configuration the equipment would be installed in the field. These drawings also dramatically reduced the amount of time that the station's engineering staff spent at the transmitter site checking and double checking fabrication drawings prior to approval.

The single most important benefit the station receives from this kind of design process is that the RF system may be installed exactly as it was match-marked and tested in the factory. This ensures that the RF system performs just as the manufacturer intends, and the need for slug tuning of transmission line or waveguide components in the field is minimized. In fact, in the WBTV-DT system, no field installation of slug tuners was required. The standard system load and antenna tuning sections were field tuned in a few hours and no other adjustments were necessary.

MOUNT SUTRO

Mount Sutro was a prime candidate for 3-D design. Here the challenge was to design a transmission line system for 10 new DTV stations within a facility that already housed 10 operating NTSC facilities.

Again, conventional plan view and elevation view drawings proved inadequate to locate and identify all potential conflicts between the new transmission lines and walls, pipes, conduits, and other items. It was necessary to coordinate 10 transmission line runs from 10 transmitters to 4 combiners and also to coordinate the 4 transmission line runs from the combiner outputs to the tower. Figure 5 is an example of one of the many drawings produced for this project.

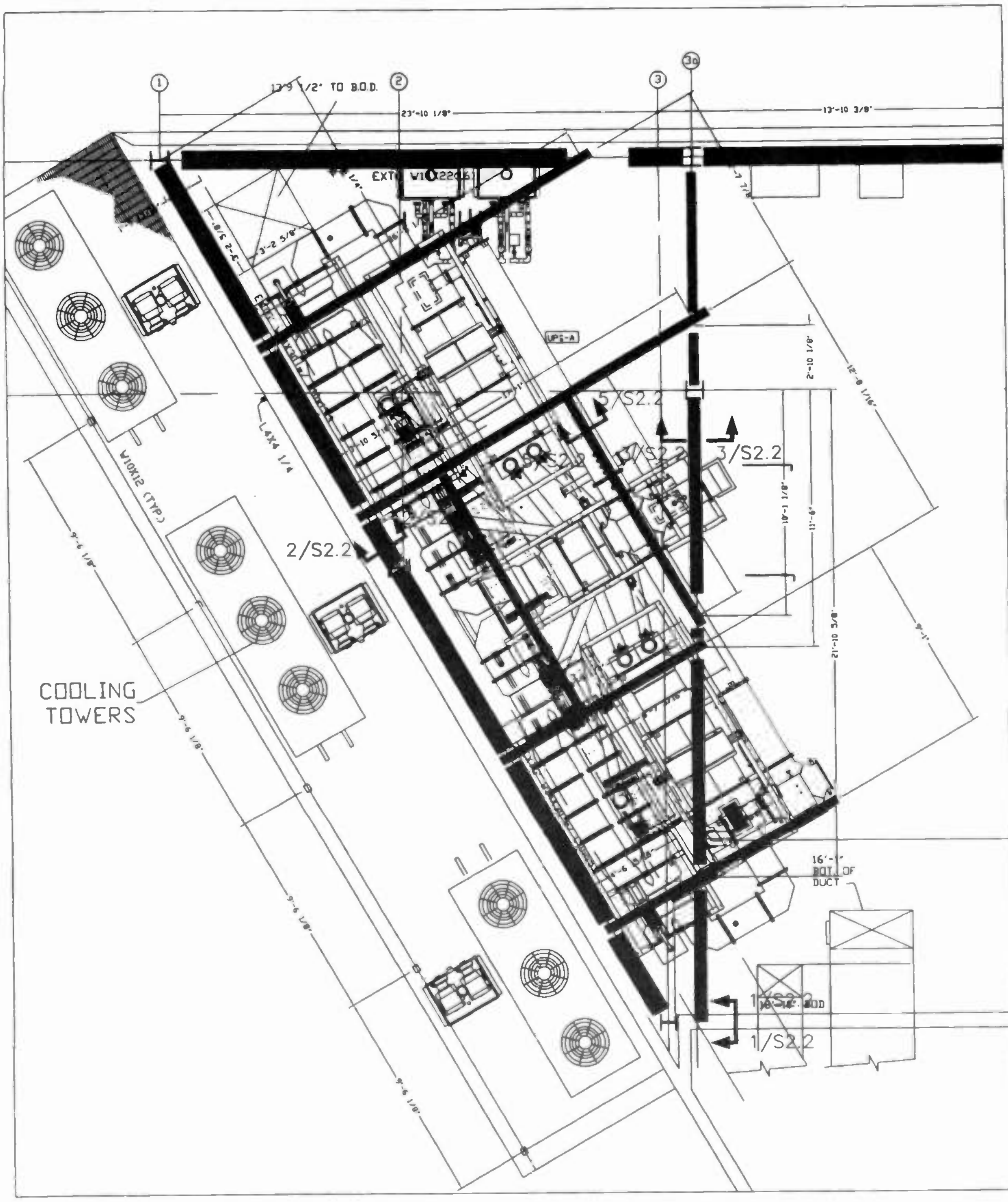


Figure 1: Standard architectural type drawing

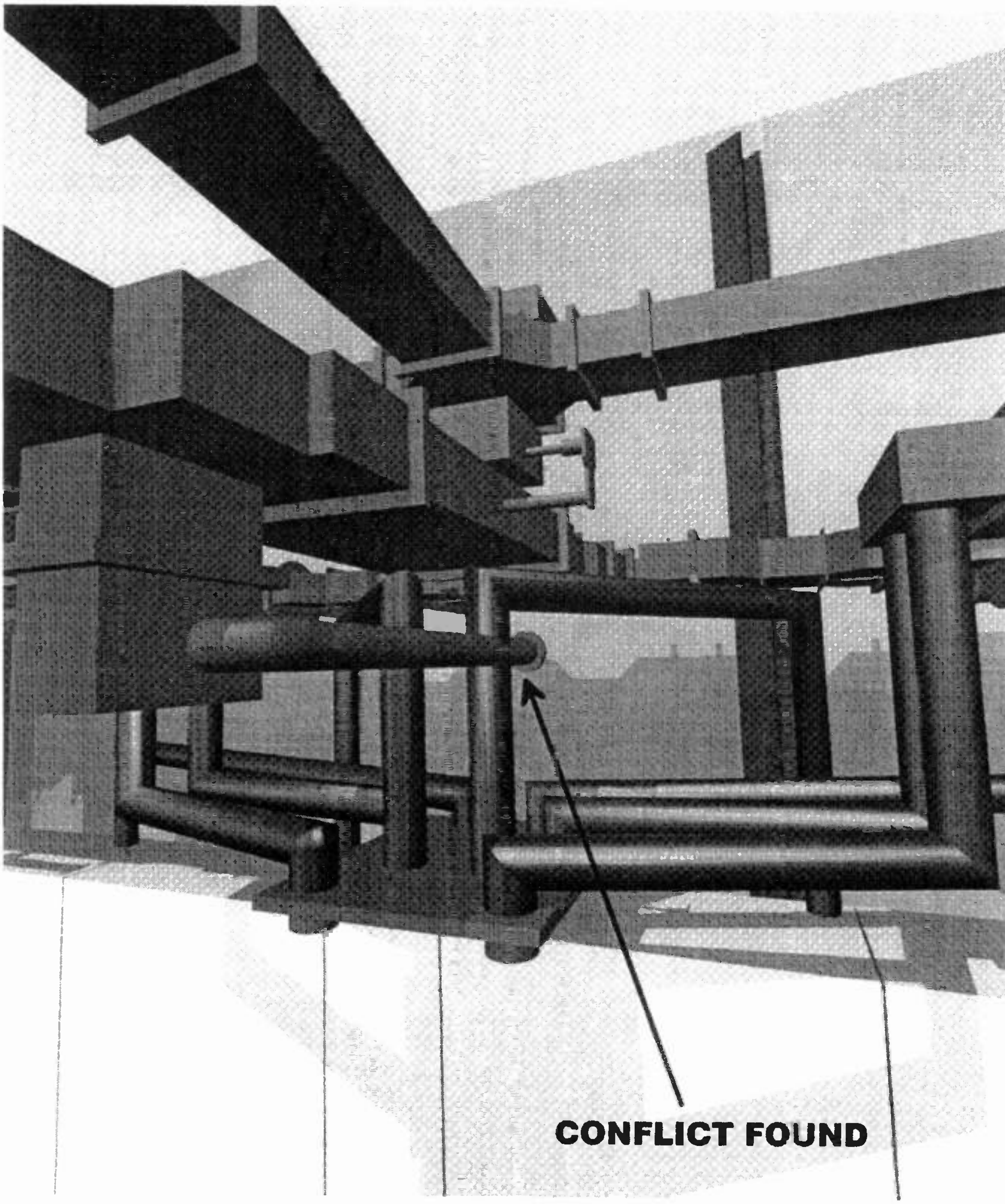


Figure 2: 3D drawing showing conflict of RF system

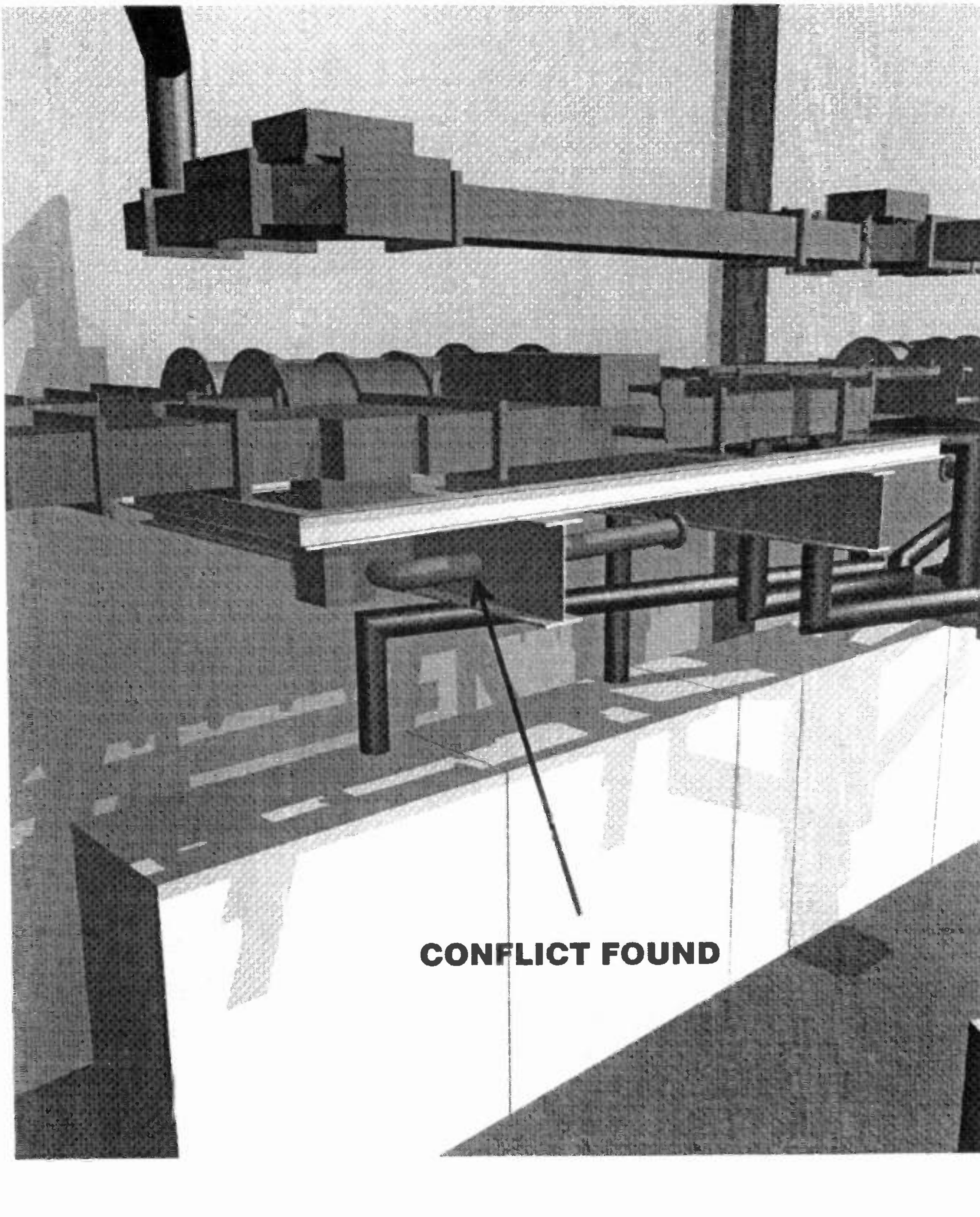


Figure 3: 3D drawing showing conflict of steel with reject lead

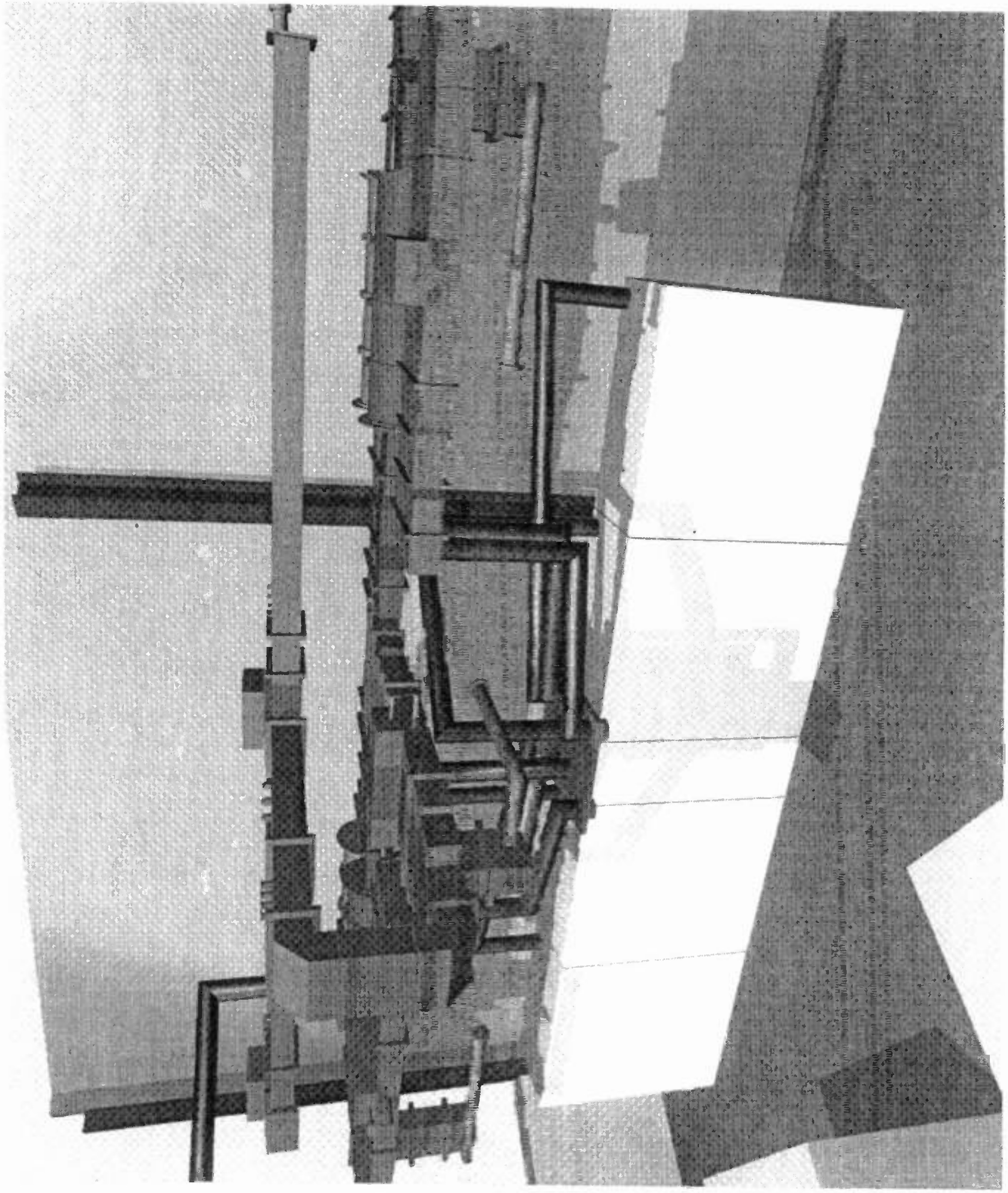


Figure 4: Final 3D rendering of WBTV-DT

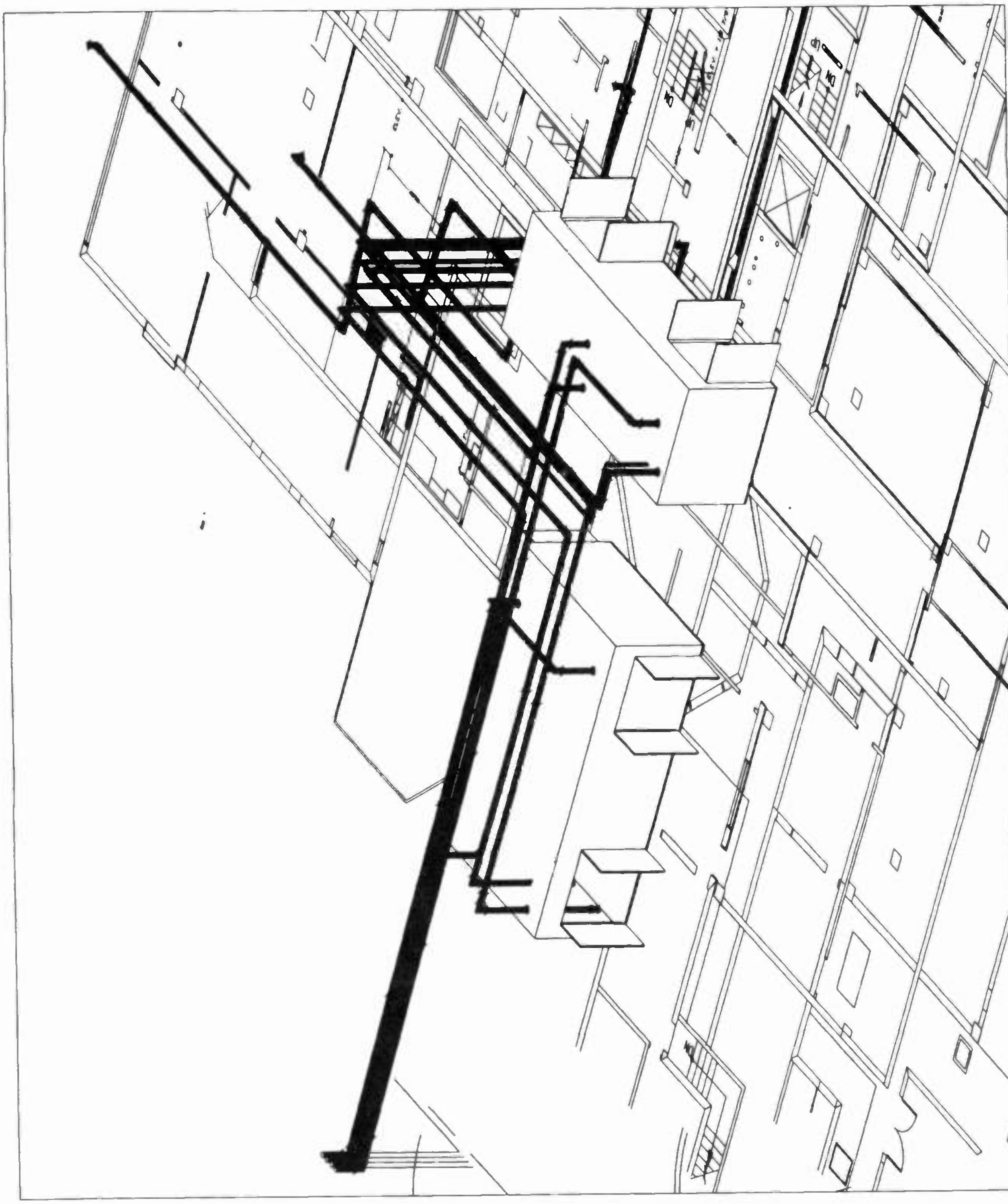


Figure 5: Mount Sutro transmission lines

The most difficult aspect of this project was locating critical walls and floors exactly in 3-D space. The most important lesson learned was to verify all dimensions from one unique point in 3D space. Once the vertical shaft and combiner buildings were located, we were able to move lines around to gain clearances needed for flanges and hangers. This allowed us to generate an isometric of each station's RF line as shown in Figure 6 and have all but the most critical custom cut line sections fabricated at the factory, thereby saving money and time in the field.

BOSTON

One of the major design challenges at the WBZ facility was to layout the RF system for six new DTV transmitters given an existing tower that is loaded with numerous NTSC, FM, and other communication antennas and transmission lines. Specifically, there are three 8-3/16-inch lines from the new DTV transmitter building to the new antenna system at the top of the tower. The building and tower are modeled to scale in 3D space to exacting standards based on surveyed data. An important design criterion for this project was the use of the minimum number of elbows necessary between the combiner and the antenna. In the broadband systems necessary for DTV transmission, it is not an acceptable practice to wait until the line is on the tower and the ice bridge is installed and then stand at the base of the tower with a pile of elbows and field flanges and try and figure out how to circumvent the diagonal that was overlooked in the design phase.

Figure 7 is a sketch from February 1999 showing the DTV transmission line scenario for the base of the WBZ tower. The final design may differ from this, but the drawing does effectively communicate the proposal at this stage of the design development to the client, the tower rigger, the ice bridge designer, and the manufacturers. This type of visual representation avoids miscommunication and results in a better final design.

After this design concept is approved, we will model minimum and maximum thermal expansion of the transmission lines and make sure that transmission line components do not hit a tower diagonal or other obstruction. Once the final clearances have been determined, then a complete isometric will be drawn showing the exact transmission line lengths required. All transmission lines including custom length sections will be fabricated at the factory, thereby facilitating the installation process, limiting the time existing stations are required to operate using their auxiliary antennas, and minimizing off-air time necessary to prevent the overexposure of workers on the tower.

CONCLUSION

Installation of a new DTV transmitter within the confines of an existing facility is a complicated endeavor that requires careful planning. Engineering design that includes the use of 3-D CAD tools to help visualize the final installation and identify potential conflicts at the design stage save the broadcaster both time and money in deployment of a new DTV transmission system. 3-D modeling prior to fabrication assures that the factory-optimized RF system can be installed in place as designed without costly field modification and tuning.

Description	Label	Length	Note	Loc.
9"-ELBOW	24.01	N/A	TYPICAL 90d ELBOW	2nd Floor
9"-ELBOW	24.02	N/A	TYPICAL 90d ELBOW	2nd Floor
6.125" LINE	24.03	146.55"	HORZ.-	2nd Floor
GAS BARRIER	24.04	2"	TRANSMITTER	2nd Floor
9"-ELBOW	24.05	N/A	TYPICAL 90d ELBOW	2nd Floor
6.125" LINE	24.06	146.375"	HORZ.	2nd Floor
9"-ELBOW	24.07	N/A	TYPICAL 90d ELBOW	2nd Floor
6.125" LINE	24.08	65.50"	HORZ.-	2nd Floor Shaft
9"-ELBOW	24.09	N/A	TYPICAL 90d ELBOW	2nd Floor Shaft
6.125" LINE	24.10	240.00"	VERT.-	Shaft
6.125" LINE	24.11	131.75"	VERT. -	Shaft
9"-ELBOW	24.12	N/A	TYPICAL 90d ELBOW	Roof
6.125" LINE	24.13	240.00"	HORZ.-	Roof
6.125" LINE	24.14	240.00"	HORZ.-	Roof
6.125" LINE	24.15	70.25"	HORZ-	Roof
9"-ELBOW	24.16	N/A	TYPICAL 90d ELBOW	Roof
6.125" LINE	24.17	240.00"	HORZ.-	Roof
6.125" LINE	24.18	240.00"	HORZ.-	Roof
6.125" LINE	24.19	12.75"	HORZ.-	Roof
9"-ELBOW	24.20	N/A	TYPICAL 90d ELBOW	Roof
6.125" LINE	24.21	131.00"	VERT.-	Roof
GAS BARRIER	24.22	2"	COMBINER ROOM	Roof
6.125" LINE	24.23	8"	VERT.-	Roof

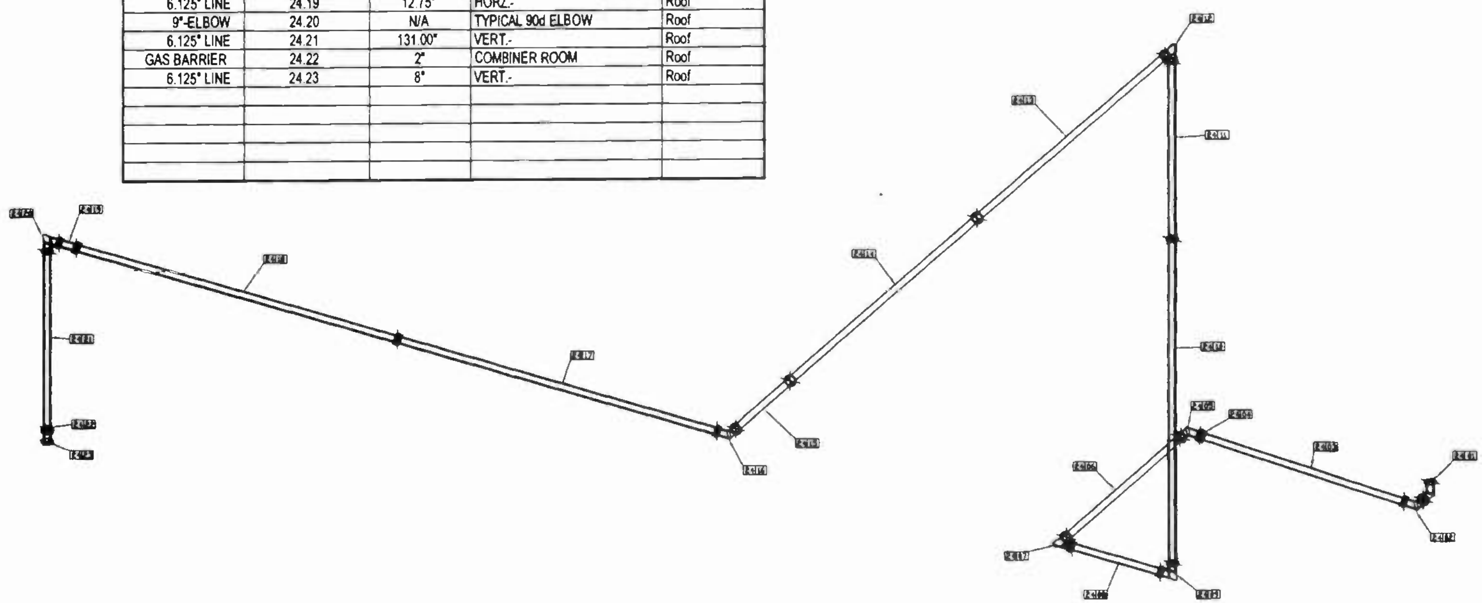


Figure 6: Transmission line isometric

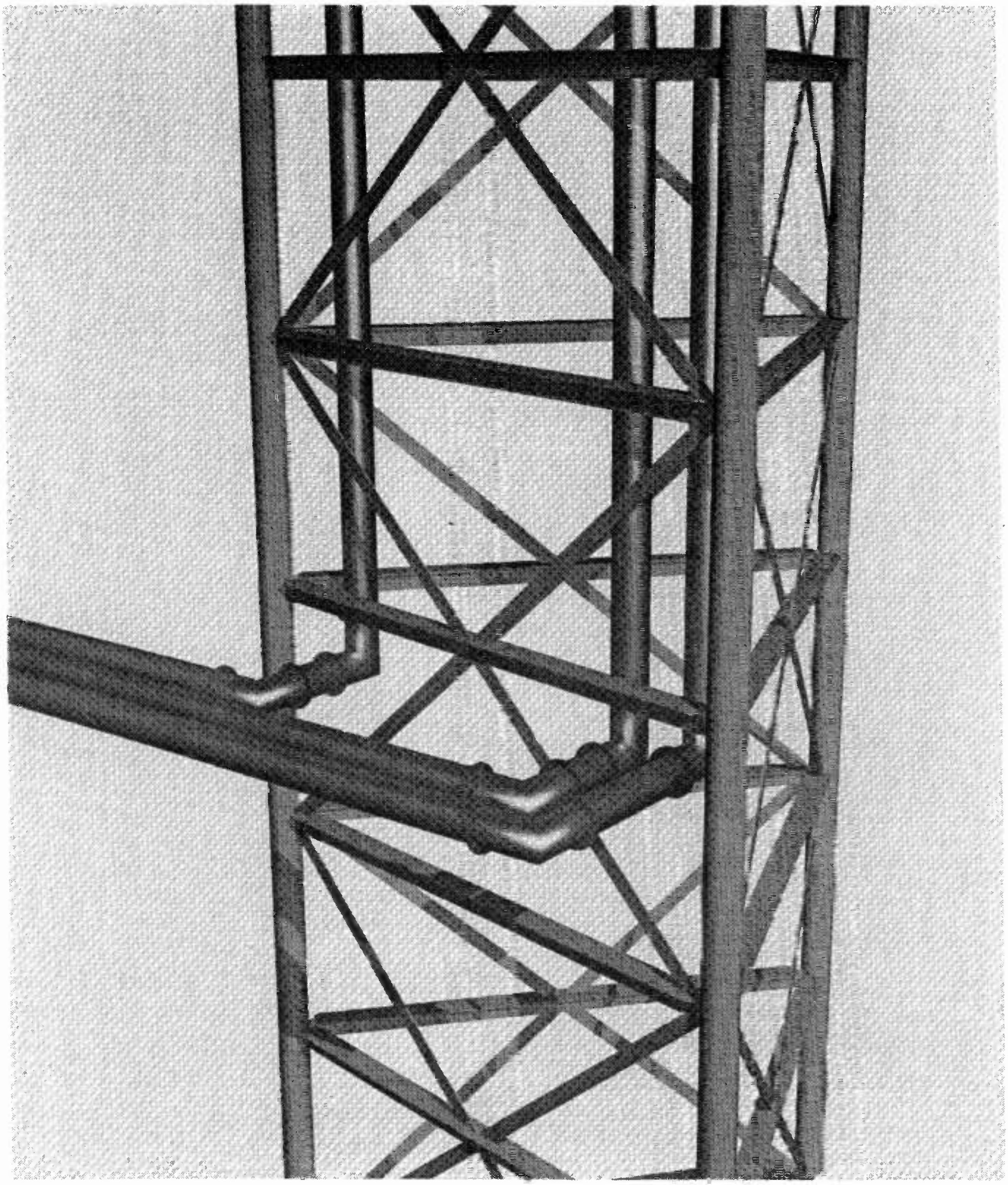


Figure 7: WBZ tower base elbow complex

WBTV FIELD EXPERIENCE
The First Full Power 1,000,000 Watt DTV ERP Station

Andrew H. Whiteside
Comark Division of Thomcast Communications Inc
Southwick, MA, USA

Abstract

This paper will present information and discuss the design and installation of the first full power, 1MW ERP, DTV station in the US – WBTV D23, Charlotte, NC. Measurement of the DTV performance of the transmitter was expected to be a challenge – and it was! Performance results are included in the form of a "Proof of Performance" which can be considered a template for future DTV installations.

Introduction

WBTV, the Jefferson-Pilot Communications station in Charlotte, NC, was granted a Construction Permit for a DTV station in October 1997. Some 10 months later the station began on-air testing at the full, authorized power of 1,000,000 Watts ERP with a new DTV transmitter and a new DTV Antenna atop the 2,000 foot WBTV tower.

Needless to say, a lot happened in those 10 short months. The activity level focussed on this "first of its kind" project remained intense throughout the 10 month period, both at WBTV and Comark. The FCC was busy too!!! – the final adoption of the technical requirements of the FCC's 5th and 6th Report and Order created considerable excitement at WBTV and Comark - especially considering the fact that this occurred some 3 to 4 months after the project started.

The final technical requirements of the FCC sparked heated debate in our industry – in particular, the requirements placed upon the spectral composition of the transmitted signal, the FCC Mask, led to a move to petition the FCC for a relaxation of the requirements on the grounds that compliance would prove financially and technically onerous. Comark's position, stated in writing to the Commission, was that such relaxation was unnecessary. This position was based upon the capabilities of the correction circuits in the 8-VSB DAP^{tm} exciter (DAP = Digital Adaptive Precorrection). This new technology promised enhanced signal performance as well as quantum improvements in user interface and operational flexibility (reference 1). Coupled with conventional High Power Filter technology, we were confident that DAP would allow full compliance with the FCC requirements, particularly in the critical "shoulder" region close to the signal of interest where high power filtering is virtually impossible and certainly ill-advised.

Nevertheless, it was a long road to proving such compliance, and we all breathed a collective sigh of relief on September 4 when we did demonstrate compliant performance at full power, albeit with a pilot version of DAP.

The structure of the following paper is intended to allow the reader to follow the overall system design, installation and performance verification processes of this significant achievement.

It is worth pointing out that much of the success for this project is due to the vision, thoroughness, technical astuteness and open cooperation of the WBTV engineering staff at all phases of the project.

WBTV Goals

From the very onset it was the goal of WBTV to "do DTV the right way". Not only did that mean putting together a transmission system to operate at the full rated ERP of 1MW, in full compliance with FCC requirements, it also meant making a commitment to DTV as the future of broadcast TV by placing the new DTV antenna on top of the 2,000 foot tower. This represented the best opportunity for achieving optimum signal coverage with this new format TV signal. Remember that, at the time this station was in the design phase, there was little in the way of concrete 8-VSB field performance data.

WBTV is a very successful VHF broadcaster. The FCC awarded WBTV the UHF assignment of Channel 23 for DTV operation. In addition to the goals outlined in the paragraph above, WBTV insisted on the same levels of redundancy and signal reliability routinely provided by their parallel tube VHF transmitter. Early in the project, Bill Napier, then VP of Engineering for WBTV, was fond of quoting the off-air record of the WBTV VHF signal. Suffice it to say, it was an impeccable record, and we all really appreciated Bill reminding us so often!

Station Description

As is usual in designing a transmission system, the most important factor is the actual required ERP. Antenna height and gain, coupled with the choice of transmission line, then yield the power needed at the output of the transmitter system. In the case of WBTV D23, the Antenna selected was a Dielectric TFU-24GTH-RO4. This antenna was chosen because WBTV believed that the lower gain would provide better overall coverage DTV signal. The WBTV tower had been built in 1984 with the knowledge that it would have to support DTV Transmission Line and Antenna in the near future. 8" rigid co-axial line was selected for the DTV application

The required transmitter output power was calculated from the equation below:

$$TPO = \frac{(ERP / Gain) * 100}{efficiency}$$

where,

TPO	= Transmitter power output in watts average
ERP	= Effective radiated power in watts average
Gain	= Antenna gain (times)
Efficiency	= transmission line efficiency in %

Substituting the predicted values for the WBTV Antenna and transmission line yields:

TPO = 67.1 kW

Note that this is the average power of the transmitted 8-VSB signal, not the peak power. As has been described in numerous papers the typical peak to average ratio of the transmitted signal can range from about 6dB to over 8dB. Whilst the actual peak power is of great interest to the equipment designer it is of little interest or meaning to the end-user, since the FCC is only interested in the station's performance at the rated ERP, which is the average ERP of the signal.

At this power level and frequency there is really only one practical choice for transmitter type – the IOT. Both the final amplifier tube type and transmitter are available from numerous reputable manufacturers and, in the decade since inception, a healthy population of tubes has reached the milestone of 50,000 hours life. In addition, the operational efficiency of this type of transmitter is outstanding, leading to very low overall power consumption.

Once the output power is determined, the next most significant factor in equipment design becomes the level of redundancy required by the end-user. In general, high power UHF NTSC stations will tolerate a drop in power of up to 3dB for first or second level equipment failures at the transmitter site. It is not common to find stations designed for full power redundancy due primarily to the cost of achieving such redundancy but also due to the relative robustness of the NTSC analogue signal. A short-term drop of 3dB will not necessarily lead to a loss of viewers. The VHF world is different – many stations are designed in Main/Standby configurations with a complete back-up transmitter. This expectation of redundancy plus the unknown impact of the so-called DTV "Cliff-effect" led WBTV to demand previously unheard of redundancy in this UHF transmitter plant design.

Transmitter Design

At the time of contract negotiation and order placement, most transmitter and IOT manufacturers were estimating that one IOT could deliver up to 25kW fully corrected average 8-VSB power at the IOT output flange. Although well within the maximum average power handling capacity of the tube and circuit, the peak power associated with this average power was thought to be the primary limiting factor. As it turned out, this was not the complete story.

Nevertheless, transmitter design was based upon an individual HPA output of 25kW average. When typical RF system losses of 10% are factored in, the resultant contribution of any one IOT at the system output is no more than 22.5kW. Therefore it is clear that 3 IOT HPAs (High Power Amplifiers) would be capable of providing the required power of 67.1kW.

However one configures the RF system with 3 HPAs, the maximum power available if one amplifier chain is down becomes 66%, or 44kW in this particular case. This did not satisfy the redundancy criteria mandated by WBTV. It was not practical, from an economic or space consideration, to provide a complete standby transmitter thereby duplicating the existing VHF configuration.

The solution to this problem lay in the unique capabilities of the DAP exciter and the enhanced control capabilities of the Comark ADVANTAGE transmitter system selected by WBTV.

The basic concept of DAP is very simple – see Figure 1. A sample of the output signal is downconverted in a precision demodulator to baseband I and Q components for comparison to the "clean" input components.

The DAP then applies an equal and opposite amount of precorrection to the input signal components thus yielding a final output substantially free of distortion. The precorrection is applied for both linear and non-linear distortion products.

The uniqueness of DAP lies in the capability to precisely "adapt" the amount of precorrection to the actual amount of distortion present in the return sample, without the need of applying a separate test signal. This adaptation can be realized in several different ways e.g. constant closed loop at regular intervals or event driven by an operator, the Transmitter control system or when certain distortion thresholds are reached, depending on particular system or end-user preferences.

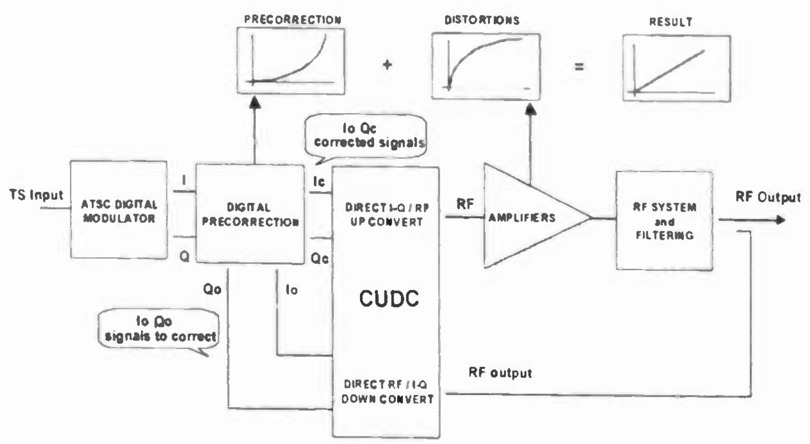


Figure 1 – DAP Overview

It is this capability which offers the possibility of providing enhanced redundancy in the high power IOT transmitter.

The final configuration chosen by WBTV included 4 IOT power amplifiers normally operating at 75% of rated HPA power, or 18.75kW. The high power RF system was configured with 3 Magic-T switchless combiners to allow any 3 HPAs to be directly combined into the final output –see RF flow diagram Figure 2. Loss of signal in any amplifier chain can then be overcome by removing that amplifier chain from the output configuration, reconfiguring to combine the remaining 3 amplifiers, raising the output power of the 3 remaining amplifiers to 100% of rated HPA power and, the coup de grace, allowing DAP to re-correct at the new power level. This process is entirely reversible when the fourth HPA is returned to service.

The additional capability of the Advantage Transmitter Control system offered the possibility of making this whole process automatic, should the end-user so desire.

The additional features of the transmitter system shown in the RF Flow diagram met another of the specific WBTV requirements – the capability to carry out troubleshooting and maintenance at the HPA level during the normal broadcast day whilst retaining full power operation. This was realized by providing a 9 pole patch-panel at the HPA output and a standby 8-VSB exciter with appropriate low-level switching interlocked to the control system. Also shown in Figure 2 are the two high power BandPass Filters. These filters are Constant Impedance devices similar to those used in Common Amplification NTSC systems. Chebyschev filters were chosen due to their inherent lack of spectral regrowth outside the first nulls in response. Two filters were used to ensure adequate power handling. Basically this transmitter is capable of delivering a total of 100kW average. The peak powers associated with this average power were thought to exceed the ratings of any individual BPF hence the decision to include 2 identical filters.

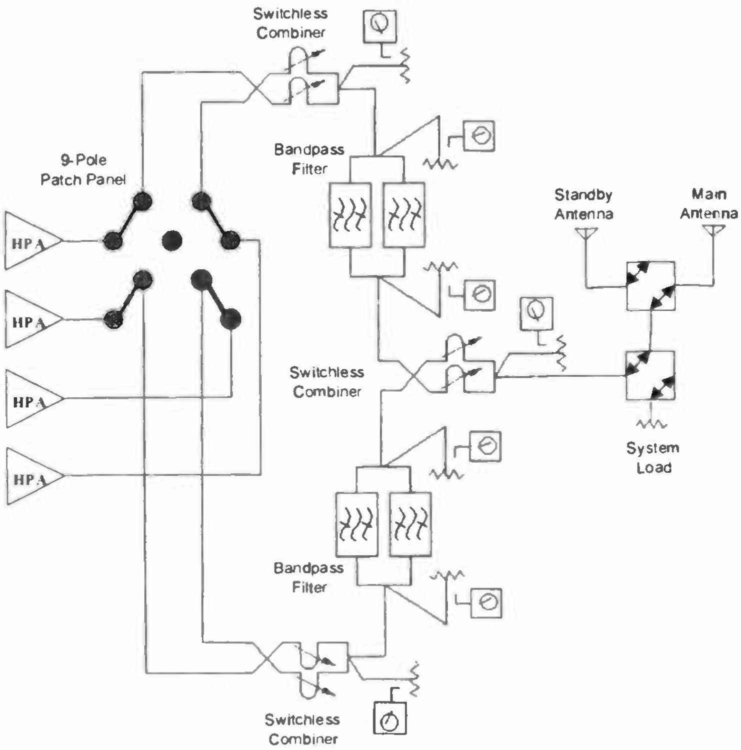


Figure 2 – RF System Overview

Not content with building in redundancy at only the RF level, WBTV engineers pushed for a "bullet-proof" cooling system. The main intent with this design was to provide automatic back-up in the event of pump failure and to reduce the likelihood of IOT contamination by separating the glycol from the IOT collector.

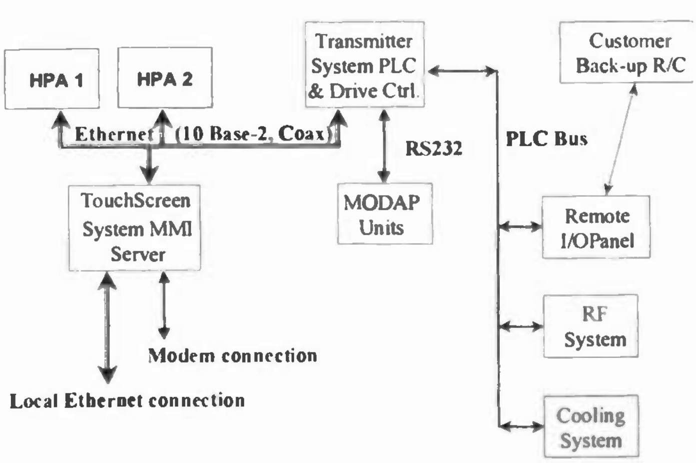


Figure 3 – Control System Overview

The overall Transmitter Control system design – see Figure 3 - was another reason for the selection of the Comark product. As well as providing enhanced user-interface and flexible networking, the system was designed with 2 layers of redundancy. The normal mode of user-interface, whether on-site or from one or more remote locations, is via the System MMI (Man-Machine Interface) in the Exciter Cabinet. This MMI provides access to all HPA and system parameters and features. If the computer operating this MMI fails, or the remote link is lost, back-up is provided via the Remote I/O interface and the Transmitter PLC. Typically this Remote I/O would be interfaced to a third-party system such as Moseley or Gentner. Finally, should the PLC itself fail during operation, emergency Drive and Cooling system control can be provided via the back-up path.

DTV Transmitter Performance

It is worth pointing out just what are the important parameters of the transmitter performance. There are, in fact, only a few parameters of critical importance to DTV transmitters. Once the power output has been achieved, the in-band signal quality – typically measured in dB as SNR or in % as EVM - and the out-of-band spurious signals – typically measured in dB below the desired signal – are the two key parameters determined by the transmitter design and set-up of signal amplifiers, RF components and correction circuits. Frequency of operation, stability and set-ability of same, are also important factors but these are determined exclusively by exciter design criteria and not affected by choice of final amplifier or type of correction.

There are two prevalent sources of distortion in a transmitter: linear distortion typically due to tuned circuits and non-linear distortion due to non-linear transfer curves in amplifying devices.

In-band signal quality is affected by both distortions whereas out-of-band signal quality is affected only by the intermodulation products of the non-linear distortion.

In general, the most important criteria for transmitter set-up becomes the overall corrected non-linearity since this will determine the actual level of the "shoulders" of the transmitted signal. As noted earlier, further reduction of these shoulders by high level filtering is impractical.

The corrected non-linearity is, in turn, dependent on the non-linearity of the amplifier chain and the correction capability of the exciter. The amplifier non-linearity is a function of the operating point on the transfer curve which also determines the operational efficiency of the transmitter.

In short, the out-of-band transmitter performance will depend very heavily on the capability of the correction circuits and the IOT. If these are well matched at the rated power, any further distortions due to linear effects can easily be corrected.

IOT Performance

A major focus of the transmitter design became the realization of 25kW IOT output power with shoulder performance exceeding the FCC requirements of 47dB below in-band signal. The resultant linearity of such performance would allow the requirements of the Bandpass filter to be readily achievable with conventional materials and techniques.

Testing of individual IOT amplifiers with DAP at Comark's facility in the early Summer of 1998 indicated that the achievement of fully compliant performance at 25kW per IOT was indeed going to be a challenge. (Maybe those other guys were right to petition the FCC after all......).

Work was under way to determine the final RF drive configuration for the EEV and CPI IOT systems. (At that time TTE did not have a suitably rated IOT system although formal qualification was subsequently completed in February 1999). The 8-VSB test results were not at all as initially expected. Although capable of providing sufficient average and peak drive power the observed limitation on corrected power output did not appear to be directly related to either. To cut a long story short, the initial limitation was ultimately shown to be a function of the characteristics of the IOT circuit assembly and not the tube itself. DAP proved itself to have adequate raw correction capability but the circuit assembly characteristics limited the actual achievable performance. Representatives from both IOT manufacturers worked closely with Comark engineers to improve the situation and it was finally decided to use the CPI K2D110W IOT system as CPI's solution was the first to be successfully demonstrated in the Comark facility – since that time successful demonstrations of all three IOT systems have been performed. Power levels in excess of 25kW with corrected shoulders of better than –48dB were demonstrated to WBTV engineers in July 1998 – shortly after the transmitter was shipped without tubes!!

Technical Results

The transmitter installation was started in June of 1998. Actual turn-on of transmitter cabinets was not started until the IOTs arrived on site in late July. Basic commissioning continued throughout August until we were ready to combine all four IOT HPAs at full rated power in early September. There were a number of hiccups along the way. Typically engineers like to introduce one major change at a time thus allowing attention to be focussed on the results of that change. Not this time......this was the first full system installation of the newly introduced Advantage Transmitter with the first DTV application of the CPI IOT and the first field implementation of DAP!! Things went surprisingly well considering the magnitude of the challenge.

It should be noted that neither the Advantage Transmitter nor the DAP exciter were installed with their final control interfaces. This led to significant ingenuity being required to provide a "system" capable of operating in a safe mode. Development has continued and upgrades of primarily control software, yielding the full system capability, are expected to continue into early Summer 1999.

The remainder of this paper will focus on the method of DTV performance measurements as well as the actual results attained. The format of the presentation will closely follow the organization of Comark's proposed DTV Proof of Performance.

Transmitter Performance Measurements

The following measurements were deemed to be necessary and sufficient to demonstrate full compliance of the transmitter system with the FCC requirements:

- Power output
- Pilot Frequency
- Signal to Noise ratio
- Spectrum Mask
  - Shoulders - plot
  - Out-of-Band
    - 20MHz span - derived
    - 100MHz span – derived
    - 20MHz span – plot
    - $2F_o$
    - $3F_o$

In addition the following information was taken or provided for reference:

- Swept response of RF system
- Test Equipment set-up
- Basic Test methods
- Transmitter meter readings

The next section of this paper includes data taken directly from the actual WBTV Proof of Performance.

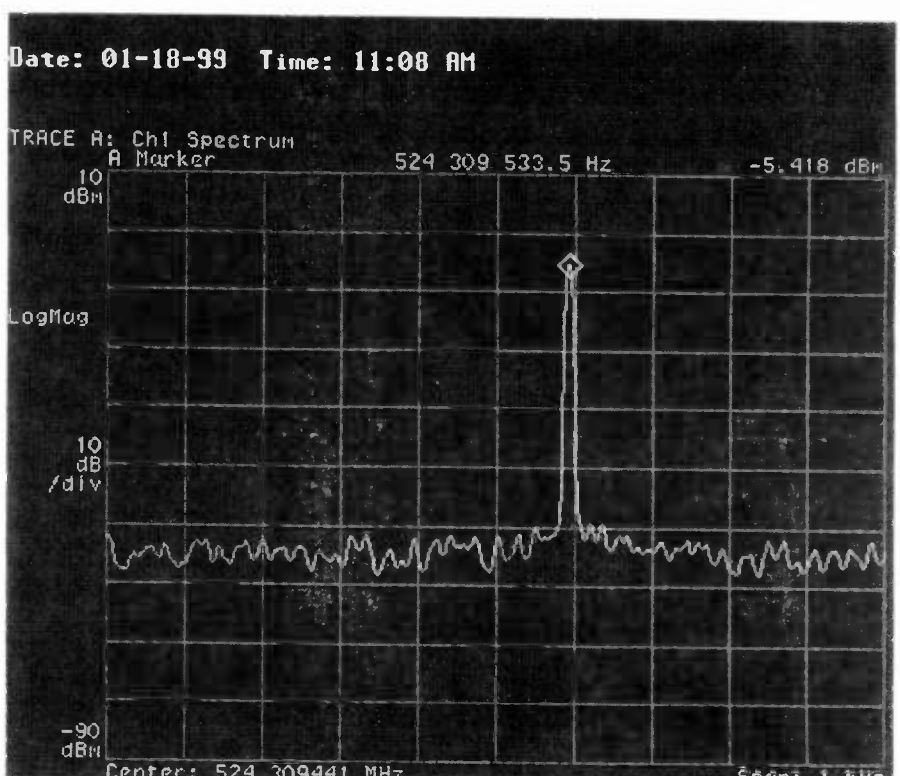
PILOT FREQUENCY

REQUIREMENTS:

The pilot frequency for this station is:	524.309441	Mhz	±	1000 Hz
The Band Center frequency is:	527.000	Mhz	±	1000 Hz
Transport stream data rate is:	19.392658	Mhz	±	Hz

MEASUREMENTS:

			Δ Hz
LO Frequency:	527.000092	Mhz	+ 92.0
Pilot Frequency:	524.309533	Mhz	+ 92.0



PILOT FREQUENCY

See next page for information on Pilot Frequency determination and measurement

PILOT FREQUENCY DETERMINATION

Two categories exist for determining the pilot frequency of a DTV transmitter, those with offset requirements, and those without.

The following systems do not require offset pilots and use the following calculations to arrive at the pilot frequency:

No Adjacent Channel, Adjacent Channel DTV, Upper Adjacent NTSC (N-1).

Pilot frequency = Lower Band Edge frequency + 0.309441Mhz.
The frequency tolerance is ± 1000Hz.

The following systems require offset pilots and use the following calculations to arrive at the pilot frequency:

Co-Channel DTV
Pilot frequency = $F_{p(DTV)}$ ± 19,403Hz or Lower Band Edge frequency + 0.309441Mhz ± 19,403Hz.
The frequency tolerance is ± 10Hz.

Co-Channel NTSC
Pilot frequency = $F_{V(NTSC)}$ − 911,944Hz or $F_{p(DTV)}$ + 28,615Hz
The frequency tolerance is ± 1000Hz.

Lower Adjacent NTSC (N+1)
Pilot frequency = $F_{V(NTSC)}$ + 5.082138Mhz or $F_{p(DTV)}$ + 22,697Hz
The frequency tolerance is ± 3Hz.

To determine Band Center frequency use the following calculation:

Band Center Frequency = Pilot frequency + 2.690559Mhz.

Note: Band Center frequency is also the Upconversion LO frequency in the Comark DTV Exciter.

PILOT FREQUENCY MEASUREMENT

The pilot frequency is measured using the VSA test set. The unit is locked to a GPS 10Mhz-reference source common to the modulator in systems where an offset pilot frequency is mandatory, or when a common 10Mhz source is available.

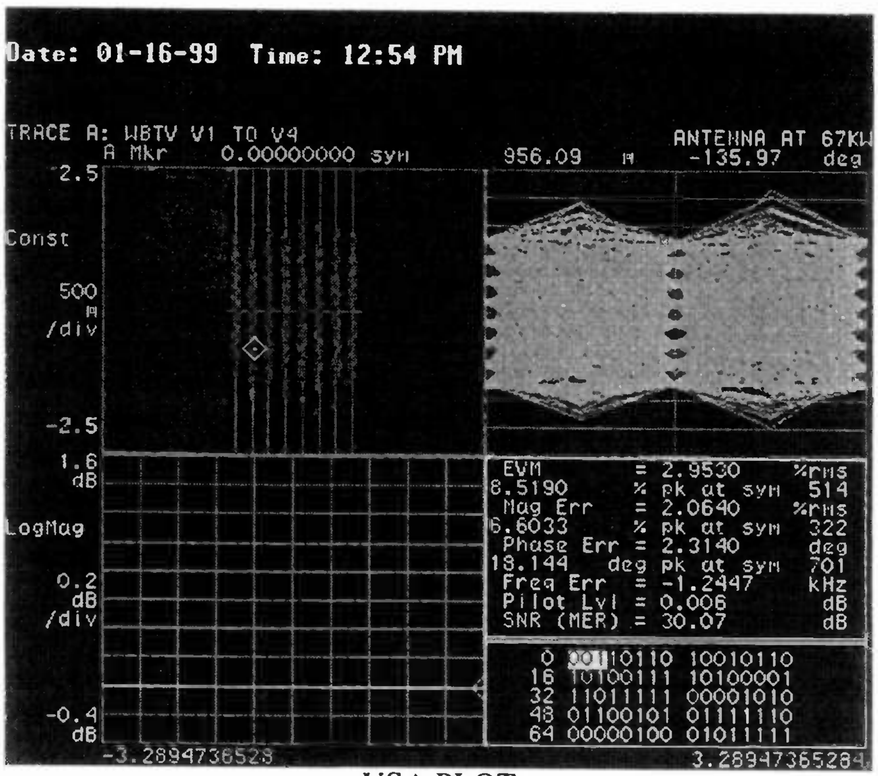
Measurement accuracy of the pilot frequency is dependent on the resolution bandwidth setting of the VSA. A resolution bandwidth of 10Hz is used for these measurements.

A marker is placed at the peak of the pilot and the frequency is displayed on the screen.

SIGNAL TO NOISE RATIO

The present industry recommended minimum signal to noise ratio is 27dB. This should be measured at the final system output after all filters without receiver linear equalization.

Measured Signal to Noise Ratio: 30.07 dB



VSA PLOT

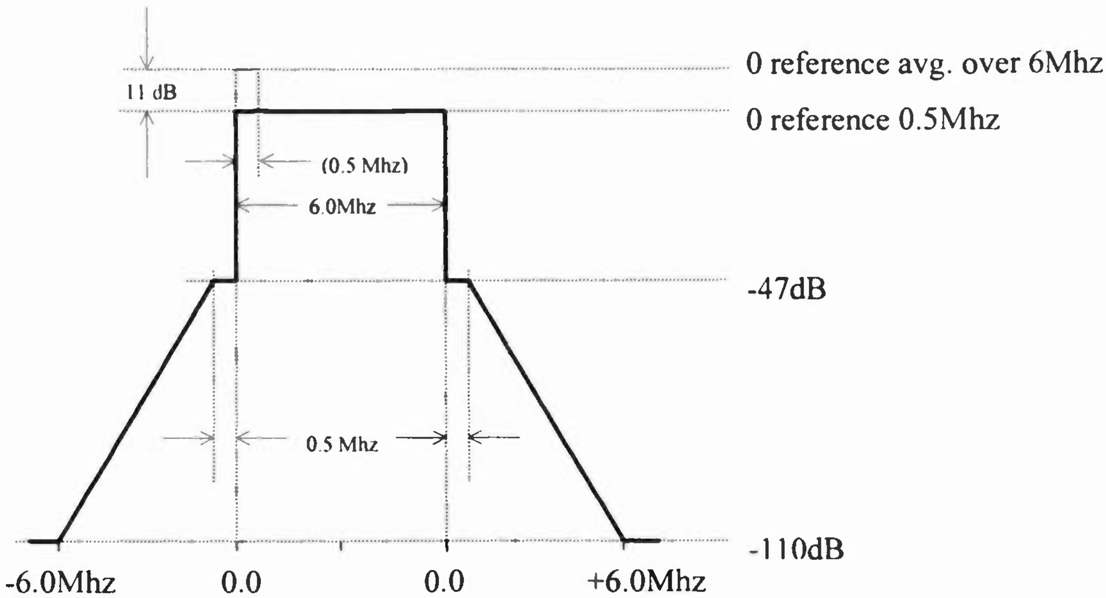
SIGNAL TO NOISE MEASUREMENT METHOD

This measurement, using a VSA or 8VSB test set, will be made without receiver linear equalization. This method demonstrates total transmitter system performance.

Additionally, since SNR is by definition an average measurement, the averaging mode of the instrument is activated in order to assure accurate measurements.

FCC SPECTRUM MASK

Specification: see chart below



Comments:

Measurement of the out of band levels is dependent on the bandwidth of the measurement. The present FCC specification refers to measurements made with a 500kHz bandwidth. When referred to the full 6MHz-bandwidth signal a correction factor of approximately 11dB should be added to the measured level as shown on the chart above.

REQUIREMENTS:

1. Out of band signals within 0.5Mhz of band edges must be attenuated at least 47dB below the average transmitted power.
2. Out of band signals greater than 6Mhz from band edges must be attenuated at least −110dB below the average transmitted power.
3. Out of band signals between 0.5Mhz and 6.0Mhz from band edge must be attenuated below the average transmitted power by a level determined by the following formula:

Attenuation in dB = -11.5(Δf +3.6); where: Δf = frequency difference in Mhz from the edge of the channel.

Note: Attenuation limits are based on a measurement bandwidth of 500 kHz.

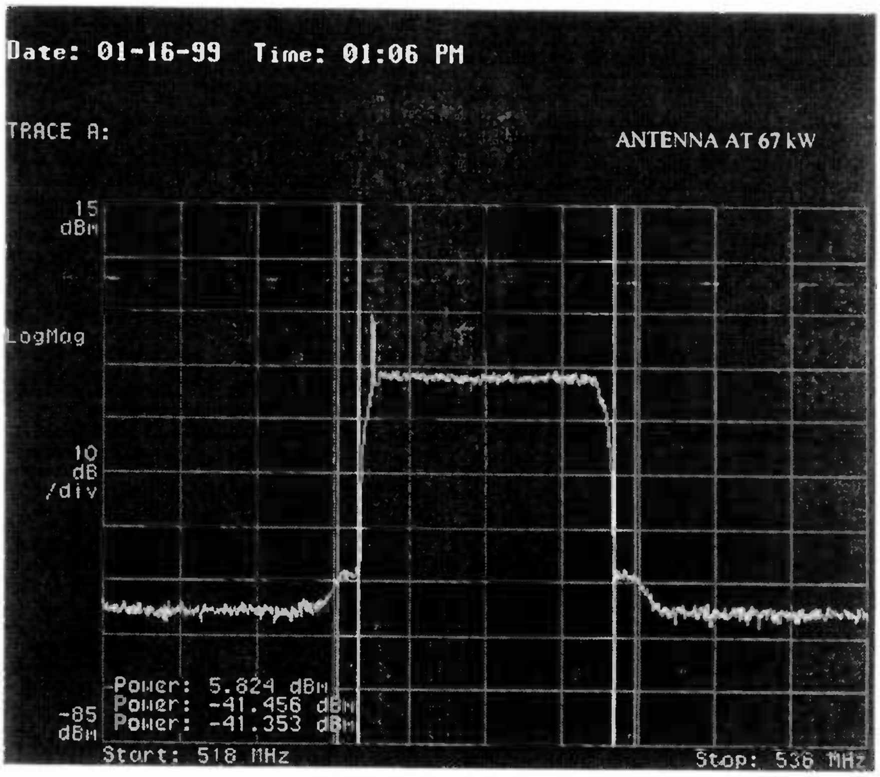
OUT-OF-BAND MEASUREMENTS

SHOULDERS

Specification: - 47 dB

Lower: -47.28 dB

Upper: -47.177 dB



SHOULDER PLOT

Out of Band Performance Measurement Method

Mask measurement is accomplished using a combination of various techniques. The shoulders versus in-band level measurement is straightforward. Difficulties arise however when attempting to view spurious levels that are at 110dB (99dB) below the in-band signal. Readily available instruments will have a tendency to have input overload and those units also have noise floor limitations. A band-stop filter at the spectrum analyzer input will prevent overload of the input circuits.

It was decided to employ a measurement and calculation method to prove the performance. Careful characterization of the filter and the couplers is mandatory. These characterizations must cover the range from frequencies well below the channel to the 3rd harmonic. Also, it is required that numerous points be taken into consideration over the range. This is necessary because of the possibility of sharp "spikes" in the component responses.

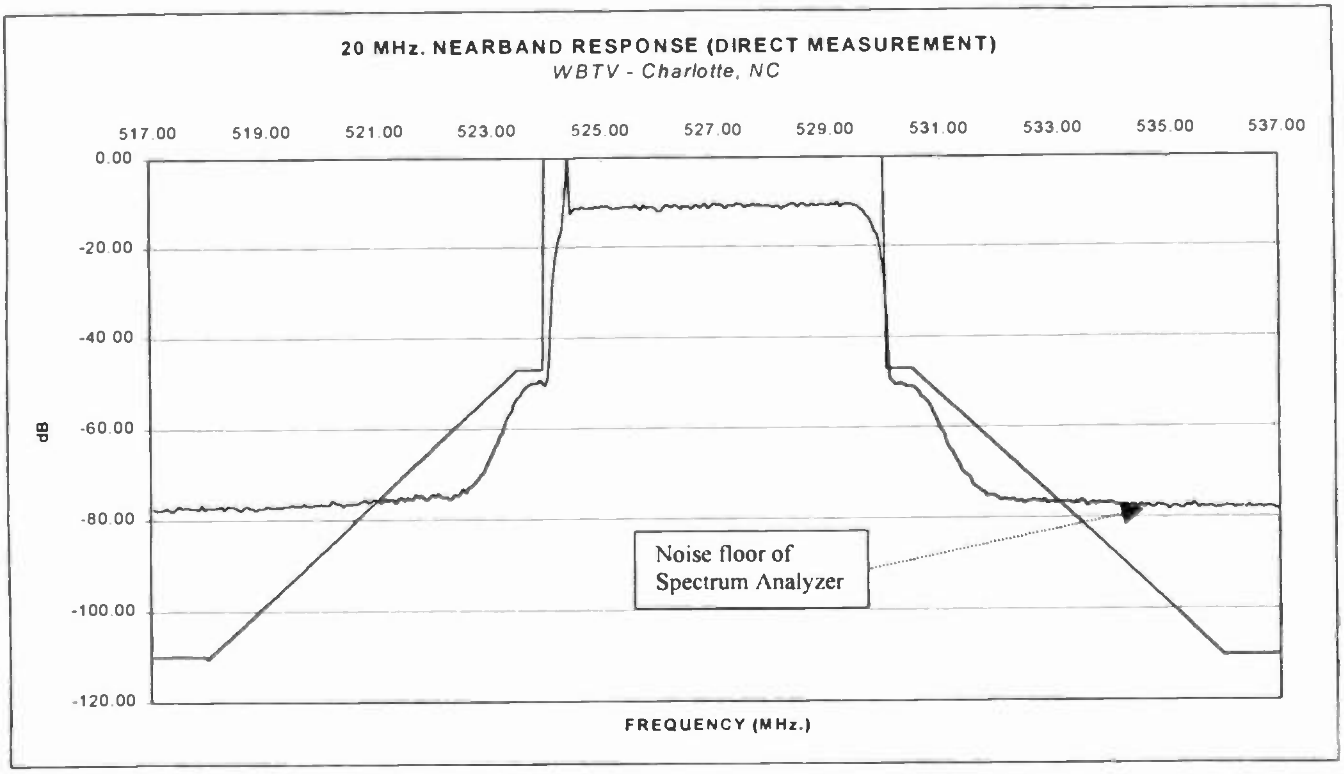
OUT-OF-BAND MEASUREMENTS

FCC MASK COMPLIANCE

+/- 2 MHz from Band Edge

The VSA response shown on the previous page gives a direct measurement of the actual power in the 500kHz region at the band edge – the shoulders. A further "direct" measurement is necessary to show the response in the region up to about 2MHz from the band edges. Beyond this frequency the dynamic range of the typical field spectrum analyzer will not allow verification of compliance with the FCC mask.

The next plot is taken from the spectrum analyzer at the system output. The data points are then fed into a spreadsheet with the FCC mask to show the comparison. This plot shows that the system output response is well within the FCC requirements in the region close to the band edges – there is at least 5 to 10 dB headroom visible before the spectrum analyzer noise floor limits measurement. It is important to use this plot in conjunction with the actual VSA display and two further plots, explained below, to demonstrate full compliance with the FCC spectrum requirements.



OUT-OF-BAND MEASUREMENTS

FCC MASK COMPLIANCE

> +/- 2MHz from Band Edge

Measurement Method

In order to demonstrate compliance with the FCC Mask beyond the +/-2MHz region it is necessary to derive the overall system response from several precise measurements. The following two plots were derived from three separate measurements. First, the attenuation characteristic of the High Power Bandpass Filter(s) was measured with a network analyzer. Next the directional coupler prior to the Bandpass Filter(s) was characterized. The transmitter output spectrum was then measured prior to the Bandpass Filter(s) using a spectrum analyzer. These three characteristics are then fed into a simple spreadsheet to derive the final system responses shown below.

The top trace on the plots is the FCC required output response – the so-called FCC Mask, the lower trace represents the "derived" system response. The reason this method cannot be used for the measurement of shoulders and within +/- 2MHz is that the fact that the RF system response is plotted when the system is cold. The filters are generally designed to have the desired response when operating at normal temperature with full power through them.

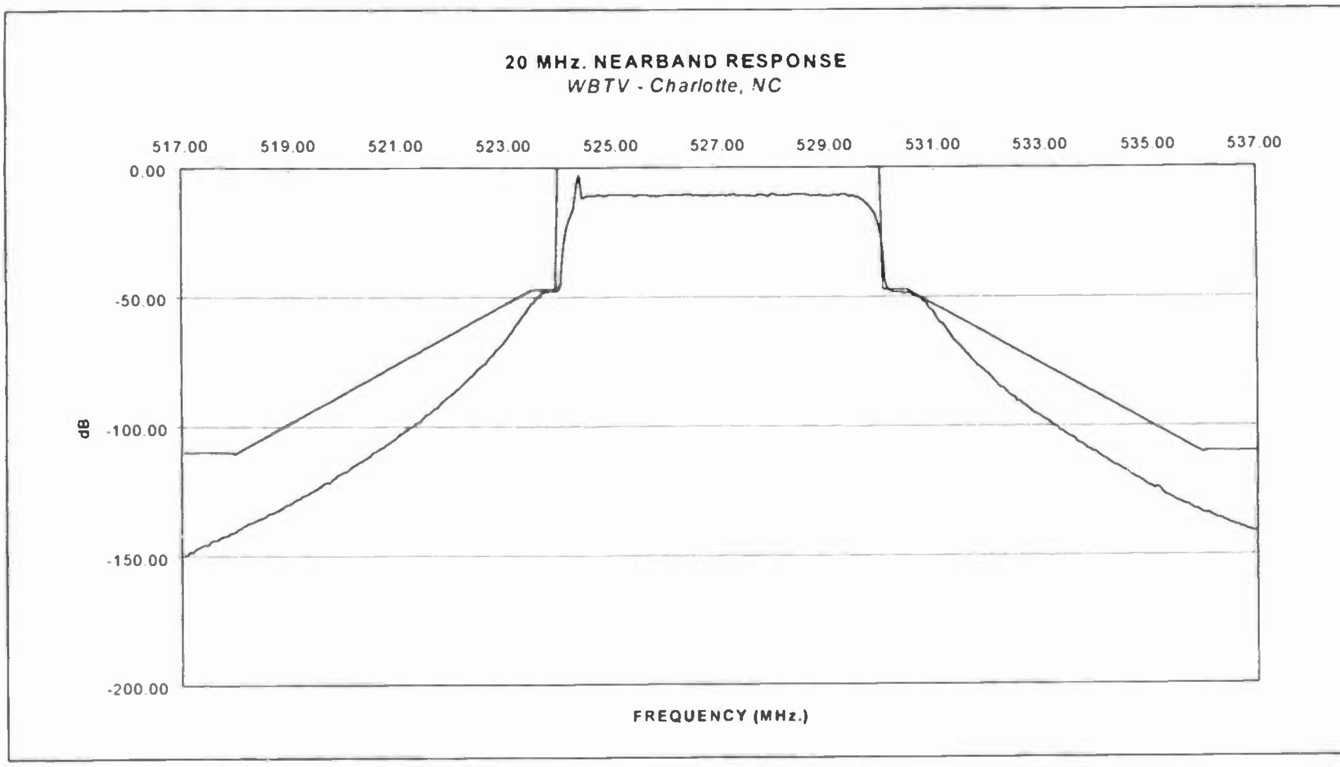
As can be seen from both plots, it appears as if the shoulder response is out of specification, when in fact we know the opposite to be true from the actual "hot" system output shoulder plot directly from the VSA. Part of the difficulty in determining compliance in this shoulder region is in actually applying the FCC specification strictly to the typically observed display of a spectrum analyzer - the specification calls for the total shoulder power within the 500kHz measurement bandwidth to be better than 47dB below the in-band signal. The specification is being met even if the spectrum analyzer display appears to show non-compliance right at the band edge. As noted above, a "true" indication of performance in this region is provided by the VSA which accurately measures the noise power in the required measurement bandwidth.

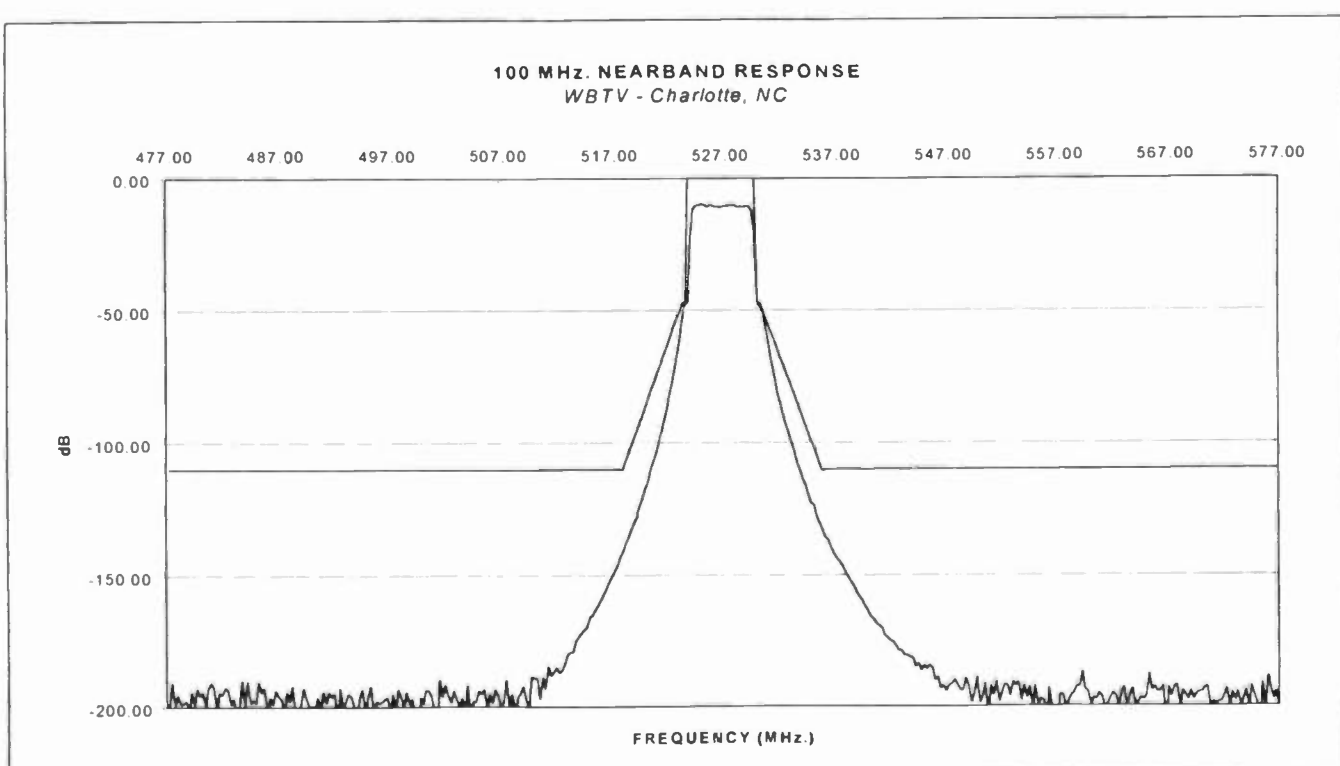
Therefore, the usefulness of these "derived" plots lies not in shoulder evaluation, or response evaluation within +/- 2Mhz of the channel edges, but in the verification of compliance with the FCC mask beyond +/- 2MHz.

OUT-OF-BAND MEASUREMENTS

FCC MASK COMPLIANCE

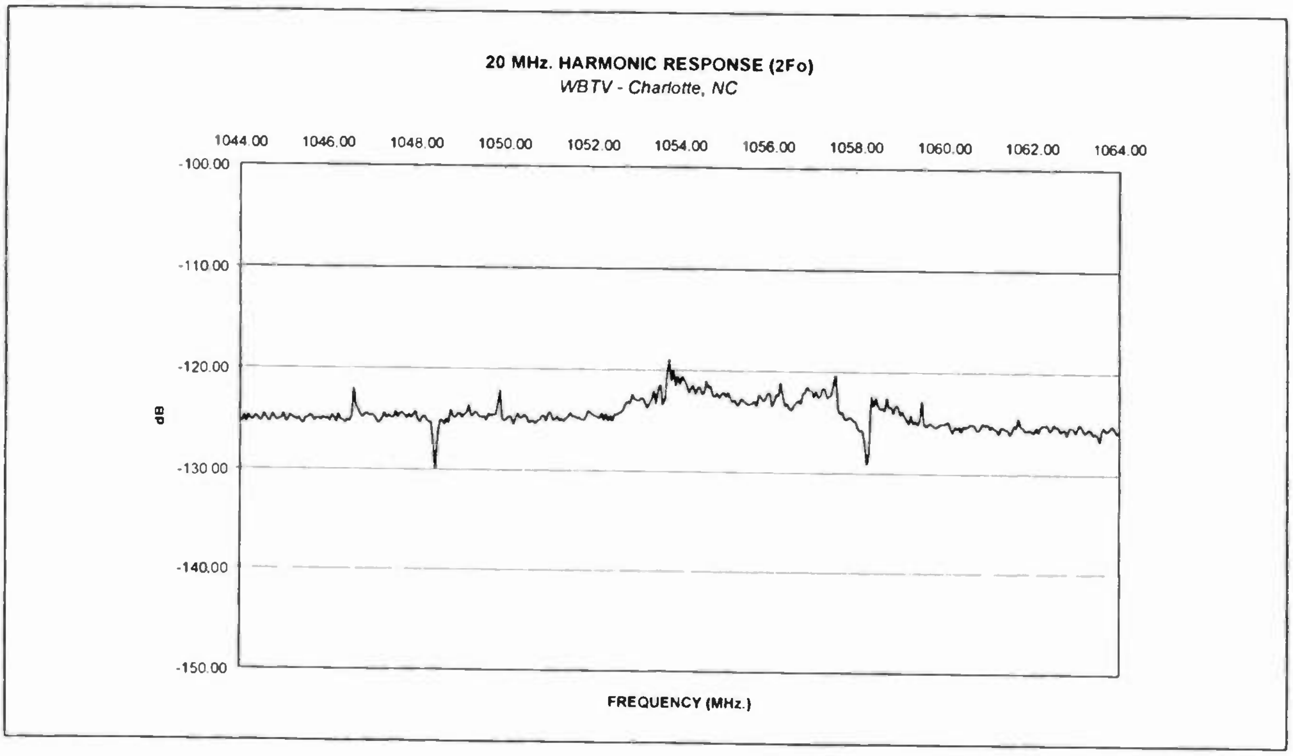
> +/- 2MHz from Band Edge

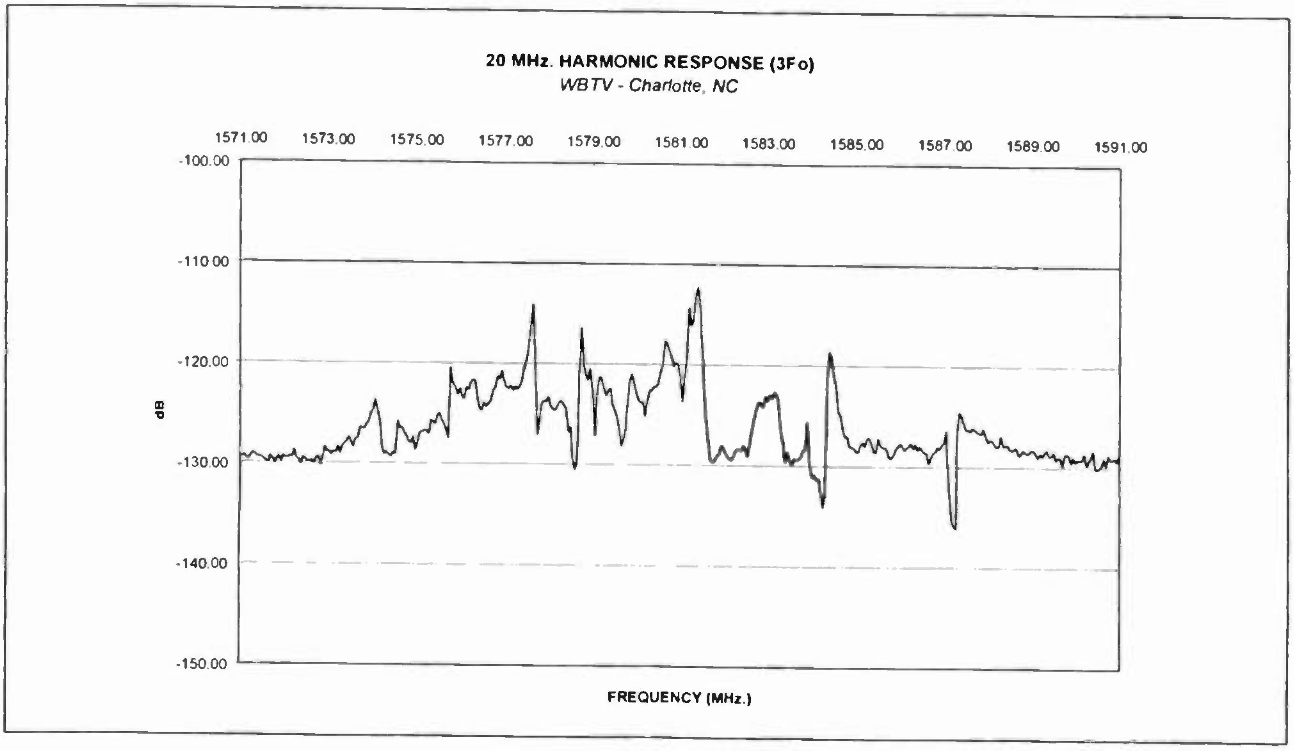




OUT-OF-BAND MEASUREMENTS

HARMONICS





Harmonic Measurement Comments

Accurate measurement of Harmonics is always difficult in UHF systems. The likelihood of higher modes being generated in the transmission line is very real. The presence of such modes can make nonsense out of the value coupled out of a directional coupler, the apparent value of harmonic levels varying dependent on where the coupler is located.

In order to combat this measurement difficulty and to demonstrate compliance with the admittedly tight FCC requirement, several precautions were taken. First of all, additional stages were added to the typical coaxial Low Pass filter present on the output of each IOT system. Then careful characterization of the behaviour of the designated system directional couplers was accomplished at both the fundamental and harmonic frequencies using the network analyzer and tapered waveguide transitions. A reference value was recorded at the fundamental using a spectrum analyzer. A highpass filter, which was also characterized with a network analyzer, was inserted between the coupler and the spectrum analyzer. Measurements were then made at both two and three times the fundamental frequencies with the spectrum analyzer. Again a spreadsheet is used to derive the response seen at the output of the system.

As can be seen in the two plots, both harmonic measurements (2Fo and 3Fo) exceed the FCC requirement. This was achieved without the use of an extra harmonic filter, such as a waveguide waffle filter, on the system output.

General Performance Comments

Overall, the installation process at WTBV was a very successful venture. The transmitter performance results met or exceeded the FCC requirements and the capability to meet requirements with three amplifiers was also demonstrated. This actually occurred earlier than planned due to the very inconvenient infancy failure of one of the 4 IOT amplifiers!! Comark engineers were able to demonstrate the redundancy design features of the system in a manual mode using the basic control system then present on site. Reconfiguration of the RF system was easily achieved, the IOT power was raised manually and subsequent initiation of DAP correction yielded fully compliant performance at full power with three amplifiers, each producing 25kW.

System performance since installation has been very reliable, especially considering the new design and application of much of the equipment. Equipment failures have been infrequent and all such occurrences have been analyzed for root cause with subsequent corrective action initiated.

Further work is planned at WBTV in the Spring of 1999. Much of this work is linked to completion of the overall Transmitter Control System and the added features and operator flexibility that this is designed to bring. There are also further improvements to the capability of DAP in progress. Tests carried out in Southwick have confirmed that appropriate manipulation of the digital 8-VSB signal prior to the actual DAP correction circuits can enhance the output performance in a number of ways dependent on site selectable criteria. For example, it is possible to increase the shoulder attenuation at a given power level or, conversely, increase the power level for a given shoulder attenuation. In both cases there will be a minor decrease in output SNR as a direct consequence of the signal manipulation. Ultimately, the final system performance will be dependent upon the end-user's particular preferences.

Conclusions

The broadcast of a full power 8-VSB test signal from the WBTV tower on September 4, 1998 was the culmination of many months of planning, design, discussion, changes, crises, successes and, above all, communication. All parties involved contributed to the achievement of several notable "firsts"……A fully redundant DTV transmission system was placed in operation…the first at a 1 MW ERP level …the first introduction of Digital Adaptive Pre-correction (DAP)…the first field demonstration of 25 kW from each IOT. The complete system met all of the objectives and all within specifications. WBTV not only has a "showpiece" DTV transmitter system but the actual results and the methods of measurement will, no doubt, set the standard for upcoming installations by others in the industry.

Acknowledgements

The experience, enthusiasm and cooperation of the entire Engineering Staff at WBTV, led at the time by Bill Napier, was instrumental in the implementation of this project. In addition, invaluable contributions were made by Bill Booth of Quality Engineering and Dean Sargent.

The following individuals at Comark provided valuable input to the preparation of this paper: Henry Fries, Fred Stefanik, Brett Jenkins, Gordon Gummelt, Mark Aitken and Ray Kiesel. Thank you.

References

1. Brett Jenkins, " Digital Adaptive Precorrection – A must in Digital Television Transmitters" presented at the NAB Engineering Conference, 1998.

STABILE HIGH POWER FILTERS for the FCC DTV EMISSIONS MASK

Derek J Small
Passive Power Products
Gray, Maine

Abstract- High power filters at the output of a DTV transmitter differ significantly from their analog counterparts. The FCC DTV emissions mask is very restrictive for power radiated in the adjacent channels and out to GPS frequencies. Transmitter out-of-band emissions and the output filter shape define the filter stability requirements to remain compliant with the DTV emissions mask. A computer program is developed to combine the two responses, adjust the filter shape, and show the allowable drift.

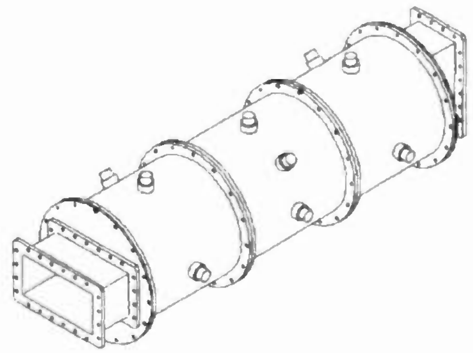
Construction technique and high power RF heating significantly affect filter stability. A filter may be compliant at low level testing, but non-compliant under high power. Cost and performance of high power filters manufactured of invar versus aluminum is discussed.

The integration of bandpass and lowpass filters to provide the necessary RF suppression out to GPS frequencies is also discussed.

Introduction- High power dual-mode filters were first utilized for terrestrial TV broadcast applications in 1993[1]. The filter was developed for common amplification transmitters and allowed for the design of wide passband widths with steep rejections to clean up visual/aural IMD products. Today, the output of nearly all high power analog common amplification transmitters use these filters due to their low loss and ability to realize complex filter functions. Most applications utilized six section pseudo-elliptic function filters with a passband width of 7.3MHz exhibiting a 1.05 vswr for a six-megahertz channel. A typical filter is shown in figure 1.

Possibly the greatest asset of dual-mode filter design for analog TV signals is the flexibility to realize wider pass bandwidths with fewer sections to compensate for frequency drift due to RF heating and ambient temperatures changes. Therefore, filters are manufactured of aluminum, a low cost lightweight material.

With the advent of Digital Television, terrestrial broadcast RF system requirements have changed, particularly the filtering.



Typical six-section pseudo-elliptic function dual-mode filter used for high power terrestrial broadcast applications.

Figure 1

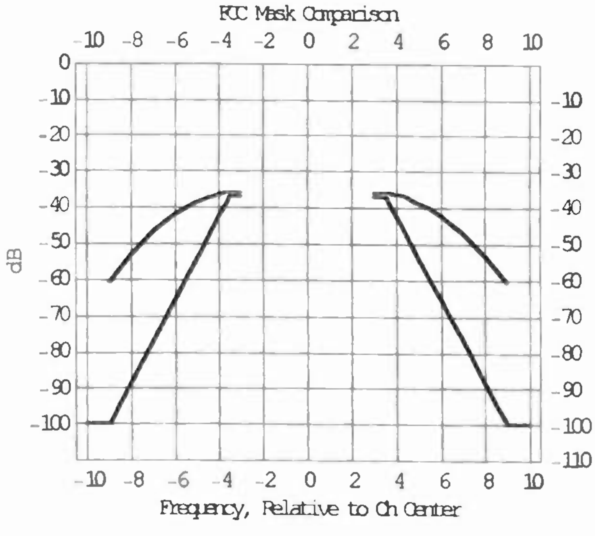
The FCC emissions mask for DTV broadcast stations is very restrictive for power radiated in the adjacent channels relative to the channel power, or from here on referred to as the adjacent channel power ratio ACPR. Consequently, stability, power handling, construction materials, and manufacturing techniques have changed. Regardless of current practical transmitter emission levels, remaining mask compliant requires a temperature stabile filter. The problem is slightly relaxed if only emission compliance in 500kHz measurement bandwidths is adhered to. Lower cost higher performance filters could be used if the mask is completely ignored but total integrated emission level in the adjacent channels guaranteed at 44dB.

The Mask- The initial FCC DTV emissions mask as reported in the "*Sixth Report and Order*" was quite moderate from a filter design stand-point, especially now, as transmitter manufacturers have improved their non-linear correction. This mask and the current DTV emissions mask as reported in the "*Memorandum Opinion and Order on Reconsideration of the Sixth Report and Order*" are shown in figure 2. Each mask is referenced to the total DTV power (power over a 5.38 MHz bandwidth). Since the FCC specifies a 500 kHz measurement bandwidth be used, the mask is normalized to the channel bandwidth by a factor of:

$$10 \text{ Log } (5.38\text{MHz}/500\text{kHz}) = 10.3\text{dB}$$

Commonly referred to as the power spectral density (PSD) normalization factor, it's used so that main channel and measurement bandwidths are normalized to the same bandwidth.

Due to the gradual roll-off, filters for the old mask could be designed wide, which allowed for considerable frequency drift. A wide filter design results in less delay variation over the channel. A transmitter manufacturer *may* not mind large delay variations, but, as pointed out later in this paper, larger delay variations result in larger voltages and increased dissipation in the filter. The new mask straddles the channel tightly and forces narrow filter designs which increases delay variations and dissipation for a given filter geometry.



Previous and current FCC masks with amplitudes scaled

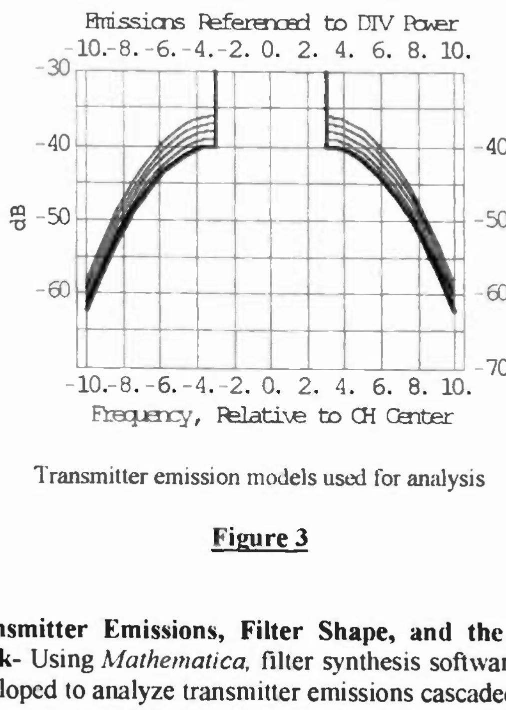
Figure 2

Each FCC mask was programmed into *Mathematica* to verify the adjacent channel power ratio. The ACPR was 44.6 dB and 39.5 dB for the new and old masks respectively. Perhaps the difference between these calculated levels and levels calculated by others is due to the pilot or actual channel bandwidth chosen.

Transmitter Emissions Model- The mathematical model used to define a transmitter emissions signal is similar to the equation used to describe the initial FCC emissions mask;

$$- L+[\Delta f^2/R] \text{ dB}$$

Where L sets the shoulder level at band edge, and R adjusts the attenuation rate in f^2. Herein, when analyzing cascaded emission and filter responses, L is varied between 36 and 40 with R set to 2.2. Figure 3 depicts the emission responses used for analysis.



Transmitter emission models used for analysis

Figure 3

Transmitter Emissions, Filter Shape, and the FCC Mask- Using *Mathematica*, filter synthesis software was developed to analyze transmitter emissions cascaded with a filter response. Filter synthesis is based on insertion loss theory in the z-plane so that numerical accuracy is maintained[2]. The filter approximation problem is solved in terms of rational functions, not in the characteristic function squared, and therefore, all phase data is retained so that the group delay can be calculated. Attenuation poles may be placed arbitrarily for a prescribed attenuation function.

The emission/filter attenuation is summed and integrated over the adjacent channel to determine the ACPR. The cascaded response is plotted with the FCC mask for further analysis. Six section filter responses were used throughout the analysis realizing that eight sections could be used in some cases to provide slightly more drift.

1. Mask Compliant Response The emission levels in figure 3 are cascaded with a filter response that produces mask compliance and therefore, emission compliance. A filter response that provides the maximum drift while maintaining mask compliance was chosen. The criterion was that attenuation at channel edge should be no greater than .5dB (which in most cases resulted in a pilot attenuation of about .2 dB) when the filter is at a maximum frequency offset ΔF. Given this, filter bandwidth, group delay variation Δτ, and allowable drift ΔF is stated for mask compliance. In addition, Δτ is given with the filter at ΔF. Results are shown in table 1.

TX Shoulder (dB)	**MASK COMPLIANCE (n=6)**				
	Filter BW (MHz)	Δτ (nsec)	ACPR (dB)	ΔF (kHz)	Δτ @ ΔF (nsec)
36	5.65	165	48.8	30	175
37	6.20	102	46.9	200	140
38	6.27	97	47.6	240	141
39	6.34	92	48.3	275	139
40	6.40	88	49.1	320	141

Table 1

2. 500kHz Emission Compliant Response Each emission level in figure 3 is cascaded with a filter response that produces mask non-compliance, but, FCC emission compliant. Here, the filter bandwidth is increased until the combined integrated power in **successive** 500kHz increments is at, or lower than the FCC mask integrated power in successive 500kHz increments. As stated in Dr. Smith's, "*Understanding the FCC's DTV Emissions Requirement*"[3], he shows how a combined response may be emission compliant but mask non-compliant in the first 500kHz. Here, it is taken another step further by breaking the **entire** FCC mask into 500 kHz increments, integrating to find the emission levels, and performing the same to the cascaded filter/transmitter response and ensuring emissions are below FCC 500kHz emission levels. There is actually very little gained by doing this as compared to Smith's analysis, but, performed to obtain what little gains were available.

Using the same criterion as for the mask compliant response, filter bandwidth, group delay variation, and allowable drift are again stated so that FCC emission compliance in 500kHz increments is maintained. Table 2 summarizes. Response plots for the mask-noncompliance emission compliant scenarios are shown in Figure 4. All plots are relative to average DTV power. The poles of attenuation are actually not needed, but they allow for more drift by creating a sharper response. Note there is essentially no gain in filter bandwidth and modest gain in the available drift. Obviously this is because the mask roll-off is so sharp and tightly straddling the DTV channel.

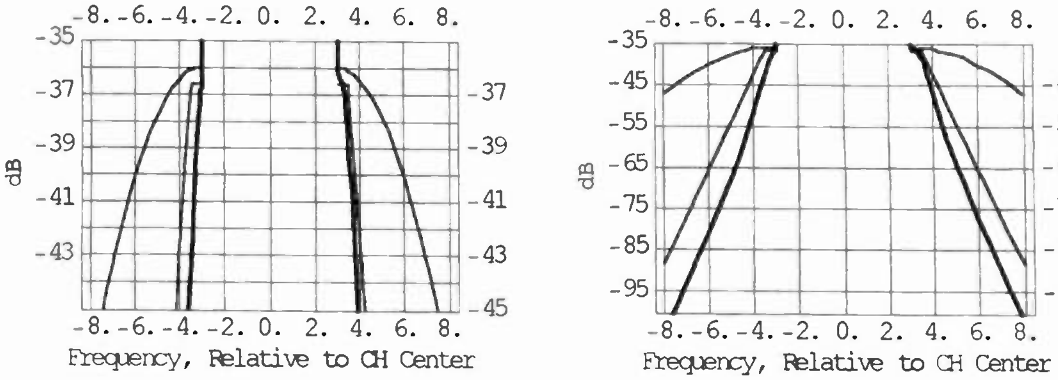
TX Shoulder (dB)	**500kHz EMISSION COMPLIANT (n=6)**				
	Filter BW (MHz)	Δτ (nsec)	ACPR (dB)	ΔF (kHz)	Δτ @ ΔF (nsec)
36	5.70	105	46.7	130	134
37	6.22	97	46.6	230	138
38	6.30	91	47.3	300	144
39	6.45	85	47.9	370	145
40	6.53	80	48.7	400	143

Table 2

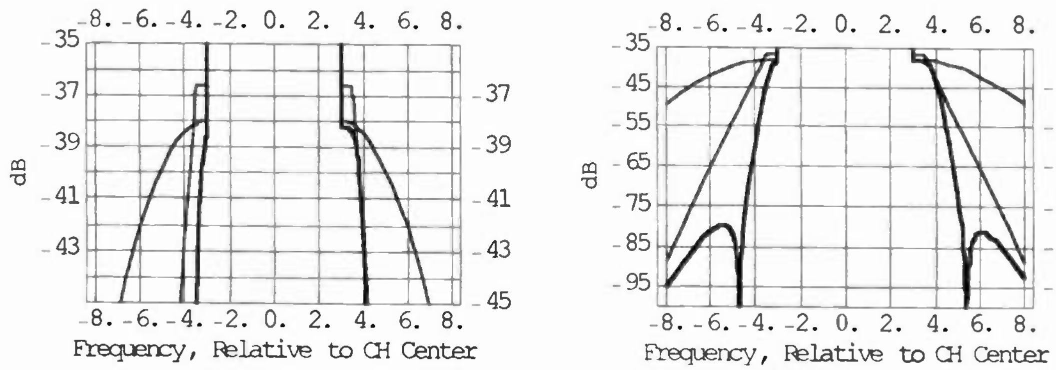
3. Adjacent Channel Emission Compliant The mask is completely ignored but the ACPR of 44dB is guaranteed over the adjacent 6MHz for a maximum offset, Δf. The idea here is to widen the passband to allow for more drift and attenuate sharply to increase ACPR. The impact of the spectral content at the N+1 vision carrier and N-1 aural can not be ignored as explained in "*The Development of a High Definition Television (HDTV) Terrestrial Broadcasting Emission Mask*" by Carl Eilers[4]. Here, there certainly is a trade off between drift, bandwidth, and keeping emissions below levels described by Eilers. This analysis was performed to illustrate the effects of wider bandwidths and ACPR. The filter parameters are again summarized and cascaded plots shown in figure 5. Note the lower de[illegible]y variations due to increased bandwidth.

TX Shoulder (dB)	**-44dB ACPR COMPLIANCE (n=6)**				
	Filter BW (MHz)	Δτ (nsec)	ACPR (dB)	ΔF (kHz)	Δτ @ ΔF (nsec)
36	6.46	87	46.1	310	146
37	6.67	73	46.2	420	142
38	6.97	60	46.3	540	145
39	7.30	48	46.5	710	139
40	7.80	36	46.5	1000	140

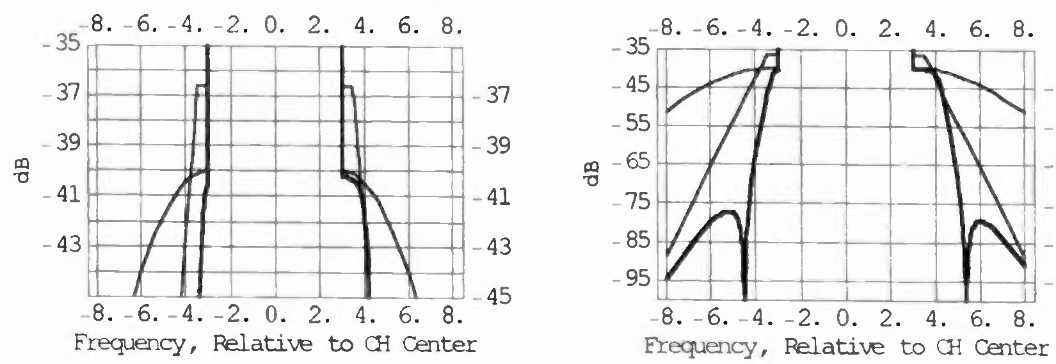
Table 3



a) -36dB shoulder, filter shifted up by 130kHz, emission compliant in 500kHz increments



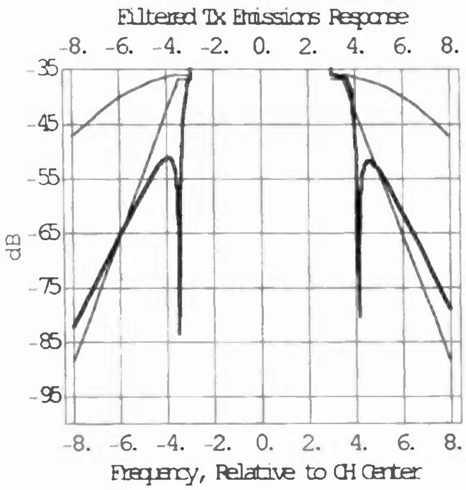
b) -38dB shoulder, filter shifted up by 300kHz, emission compliant in 500kHz increments



c) -40dB shoulder, filter shifted up by 400kHz, emission compliant in 500kHz increments

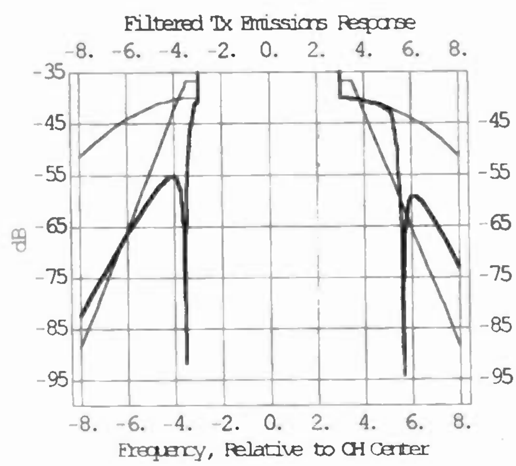
Responses emission compliant in 500 kHz increments for transmitter shoulder levels of 36, 38, and 40dB

Figure 4



a) 36dB shoulder, ΔF = 310kHz

b) 38dB shoulder, ΔF = 540kHz

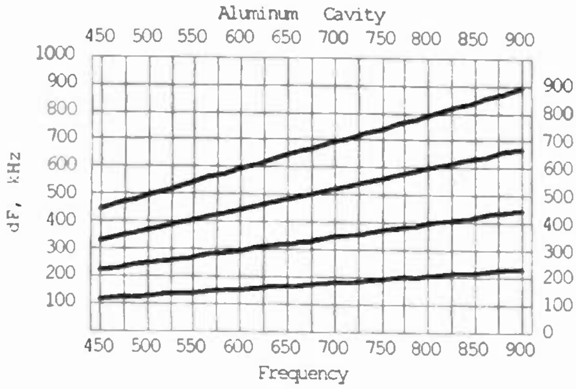


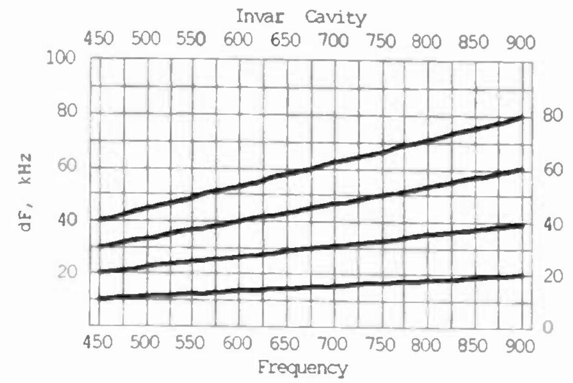
c) 40dB shoulder, ΔF = 1000kHz

ACPR Compliant responses (mask ignored) for transmitter shoulder levels of 36, 38,and 40dB

Figure 5

Drift- The question is...how much will an aluminum, steel, or invar filter drift. Since high power filters will see heat differentials on the order of 25 to 45 degrees C from start-up to operating temperature (DTV powers of 24kW to 50kW), one has to be concerned about the drift. To minimize center frequency drift with temperature changes, whether due to changes in ambient or RF power, dimensional changes in the cavities must be minimized. Figure 6 illustrates theoretical changes in center frequency due to temperature differentials of 10, 20, 30, and 40 degrees C for aluminum and invar cavities. These curves are base on linear coefficients of thermal expansion for aluminum and invar being 24.7×10^{-6}in. /in. /C and 1.6×10^{-6}in. /in. /C respectively. Based on the analysis above, silver-plated invar cavities, or some other temperature compensated filter is required for FCC compliant DTV broadcast applications.





TE11 cavity center frequency shift, dF vs. frequency for 10, 20, 30, and 40-degree C temperature change.

Figure 6

Construction technique plays a very important role in filter stability. Silver plated invar will keep most filters stabile in center frequency; the difficulty lies in maintaining good vswr with temperature change. When compared to other metals, invar is dramatically more stabile with temperature but also a very poor thermal conductor, therefore care must be taken when using other materials within the filter. For example, cavities manufactured of invar and iris plates manufactured of aluminum will result in uncontrollable elevated vswr during warm-up. The aluminum iris plate will want to grow over 15 times faster than the cavity and has the tendency to either bow or deform the cavities thereby shifting the cavity resonant frequencies in an unpredictable manner. This problem is compounded by the fact that most of the dissipated heat is generated at the iris so it's growing while the relatively cool cavity made of invar is not.

The vswr problem is easily solved by replacing the aluminum iris with an invar iris. However, by doing so, another problem is created. Since invar is a poor thermal conductor, the heat generated on the iris has a tendency to stay on the iris and limits the average power handling of the filter.

Heat- Average and peak power capability of a filter is dependent on the filter function and construction. Examining the effect of the filter function first, the greatest difference between an analog and DTV filter is the bandwidth, 7.3MHz and approximately 6.2MHz (depending on shoulder level of transmitter emissions) respectively. Table 4 denotes significant parameters that effect the dissipation loss, or heat build-up, of DTV and analog filters. Both filters are six sections with a vswr of 1.06 across their respective passband. Filter center frequency was chosen high in the UHF band to amplify the results (i.e. insertion loss is greater and percentage bandwidth is lower). Cavity Q_u was arbitrarily chosen to be 20000, but closely resembles the Q_u of aluminum dual-mode TE111 cavity at this frequency. The mid-band insertion loss is calculated from the cavity Q_u, n, vswr, and **bandwidth**. Note the increase in mid-band loss for the narrower DTV filter. Because of the increased mid-band loss and narrower bandwidth, the roll-off at band edge is also significantly greater for the DTV filter. The insertion loss was integrated across the channel to arrive at an average. This is particularly important for the DTV signal where power is evenly distributed across the channel. Given 24kW of average DTV power, P_{Tx}, the dissipated power, $P_{Dis.}$ is calculated. There is approximately 25% increase in insertion loss for the DTV filter, which results in hotter running filters constructed of the same TE111 cavity.

	TE111 FILTER	
	Analog	DTV
n	6	6
Vswr	1.06	1.06
Bw (MHz)	7.3	6.2
f_o (MHz)	755	755
Q_u	20,000	20,000
mid-band loss (dB)	.156	.186
band edge loss (dB)	.214	.283
Average Loss	.173	.218
P_{Tx} (kW)	24	24
P_{dis} (W)	937	1175

Table 4

A similar analysis will show that peak power handling of a narrow DTV filter is reduced by approximately 25% of the wider bandwidth analog filter[5]. However, this is not a significant problem because there is plenty of peak voltage headroom in properly constructed cavities.

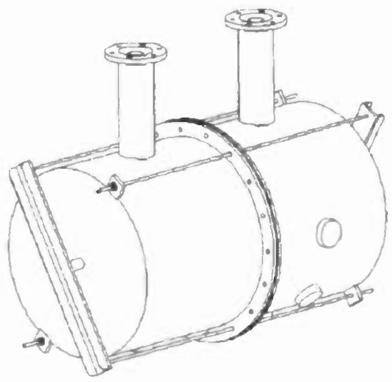
Voltage (and heat) problems occur with cavities that are designed too short and have excessive tuning probe penetration. Voltage breakdown may occur between probes within the same cavity, or between the probe and iris plate. Engineers will often design cavities too short due to lack of knowledge, or for insurance that a second operation that consists of cutting the cavity to its proper length will not occur. Shims are sometimes added to short cavities to increase the cavity length, but they have a tendency to reduce the cavity Q and further degrade insertion loss and increases heat.

Out-of-Band Attenuation- Spurious modes that lie relatively close to the passband of a waveguide filter are quite easy to predict, the difficulty lies in predicting the level of suppression for each mode. Spurious frequencies are frequencies at which modes other than the dominant mode propagate. The dominant mode is the lowest propagating mode for a given transmission line size. Filters are generally designed in the dominant mode and any discontinuity in a waveguide filter will tend to "light up" the spurious modes. Spurious mode attenuation depends on the mode propagation constant and length of guide in which it travels. For the case of a dual-mode filter, spurious frequencies can be predicted by using a mode chart or simply calculated in the order in which they occur.

Spurious frequencies are eliminated by cascading the waveguide filter with a mode-free lowpass filter. Lowpass filters can be manufactured in coax or waveguide and are very geometry dependent for good mode-free performance out to GPS frequencies. The number of sections in the lowpass filter is dependent on the level and location of the spurious modes requiring attenuation.

Invar vs. Aluminum, Cost Comparison- Several factors influence the cost of an invar filter. The cost of raw material is several times that of aluminum. Processes used to manufacture defect free invar cavities complete with flanges and tuning bosses require significantly more quality control over the processes used to manufacture aluminum cavities. The cavities also require silver plating and an aggressive quality control program that ensures plating adherence and proper thickness. The result is a filter that costs about 70% more than an aluminum filter.

A Mixed-Mode Temperature Compensated Filter- The cost of manufacturing invar filters, and the heat problems associated with them would inspire any engineer to design a temperature compensated filter manufactured in aluminum. One such filter is shown in figure 7. The bandpass filter illustrated uses the deformation of cavity surfaces in response to thermal changes to compensate for the resonant frequency shifting effects of thermal expansion. Additionally, a mixed-mode operation helps relieve lowpass filter requirements compared to a straight TE11 cavity design. The filter is designed for a single IOT output.



Temperature Compensated Mask Compliant Aluminum Filter

Figure 7

Summary- It has been shown that a FCC mask compliant response at the filter output requires the use of a temperature compensated filter regardless of current practical transmitter emissions. A temperature compensated filter is also required for emission compliance that's calculated in successive 500kHz increments.

When using temperature stabile cavities constructed of Invar, other materials within the filter must be chosen carefully to ensure stabile vswr during temperature change due to high power RF heating and ambient. Construction and cavity design significantly effect the average power handling capability of the filter. For optimal performing invar filters, precise cavity lengths and an aggressive quality control program is required.

Lower cost temperature stabile filters manufactured in aluminum are currently available for power levels up to 30kW in a constant impedance configuration. Spurious suppression is reduced using mixed mode operation and thereby reducing lowpass filter requirements.

References

1. D.J. Small, "High Power Multimode Filters for ATV, DAB, and Common Amplification TV System", NAB 1996 Broadcast Engineering Conference Proceedings.

2. H.J. Orchard and G.C. Temes, "Filter design using transformed variables", IEEE Trans. Circuit Theory, CT-15, No. 4, 385-408, Dec. 1968.

3. P.D. Smith, "Understanding the FCC's DTV Emissions Requirement", Micro Communications, Inc. report No.6157, Mar 1998.

4. C.G. Eilers, "The Development of a High Definition Television (HDTV) Terrestrial Broadcasting Emission Mask", IEEE Trans. Broadcasting, Vol. 41, No. 4, Dec 1995.

5. G.L. Matthaei, L. Young, and E.M.T. Jones, "Microwave Filters, Impedance Matching Networks and Coupling Structures", p.p. 910-920, McGraw-Hill, N.Y. 1964.

Bandpass Filter and Linearization Requirements for the New FCC Mask

Robert J. Plonka, Principal Engineer
Harris Corporation, Broadcast Systems
Quincy, Illinois

Abstract

This paper will discuss the cost and performance trade-offs between selecting a sharp cutoff DTV filter versus providing significant amounts of linearization in a DTV transmitter in an attempt to meet the new FCC out-of-band suppression requirements.

The engineering issue is to answer the question: Which is better, to provide a temperature compensated bandpass filter for IMD sideband suppression or depend on linearization techniques to eliminate the filter? Maybe the best course of action would be to use a combination of both techniques for a cost effective approach.

To help answer these questions, this paper will provide measured data on bandpass filter performance and linearization techniques to compare the two methods for determining the best solution. This paper will also provide comments on various methods of interference reduction to NTSC using a DTV filter.

The FCC Mask.

The current FCC mask is the key point in beginning any discussion on DTV filter requirements. The mask, however, has been the subject of many heated discussions by various industry officials ever since its issuance by the FCC in Feb. 27, 1998, which also includes this author. The paramount feature of the new mask is its exceptionally sharp cut off characteristics and the deep level of out-of-band emissions of being −110 dB down from reference power. There has never been such a fierce requirement imposed on the industry before, barring a few land mobile protection situations. The new mask requirement applies to all broadcasters with a DTV license. Here is what the new mask looks like. Refer to Figure 1.

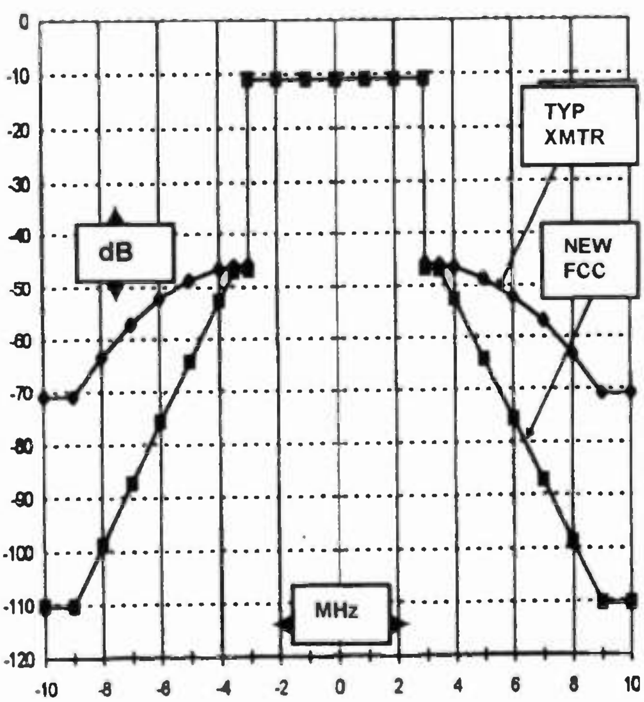


Figure 1. FCC mask requirement.

In an attempt to discuss the various issues concerning the mask and to determine what filter and linearization requirements are to be met it, it is worth the effort to revisit the wording of the new rules. Here is a verbatim transcript of the rules. FCC 98-24, section 92.

The DTV out-of-band "emissions mask" require that:

1. in the first 500 kHz from the authorized channel edge, transmitter emissions must be attenuated no less than 47 dB below the average transmitted power;

2. more than 6 MHz from the channel edge, emissions must be attenuated no less than 110 dB below average transmitted power,

3. and any frequency between .5 and 6 MHz from the channel edge, emissions must be attenuated no less than the value determined by the following formula: attenuation in dB=-11.5(Δf+3.6); where Δf = frequency difference in MHz from the edge of the channel.

Note in the above, there are three sections to the rules, and it is assumed in this paper the Commission will require that all sections are to met simultaneously. This is in spite of some industry comments that if a broadcaster reduced his of out band emissions by 5 dB below the old mask level, in the adjacent 6 MHz slots, then mask requirements have been fulfilled.

This view can provide some interesting filter strategies in setting up a filter response to just meet the 5 dB out-of band reduction. The additional 5 dB reduction, when added to the original 39 suppression level, will now require adjacent channel spectral components to be -44 dB.

There is, however, potential interference to adjacent NTSC operations. More on this later.

Refer to Figure 2 and 3 to see what a mask compliant system looks like.

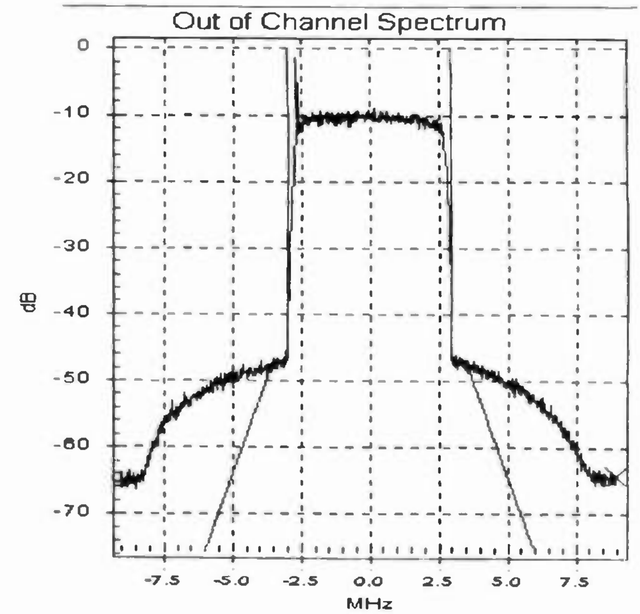


Figure 2. Typical transmitter output with linearity set to put IMD shoulder levels at -37 dB.

Note in Figure 2 , the IMD levels spill over the FCC attenuation limits (shown as sloping lines) which means an aggressive filter system will have to be used. This shown in Figure 3

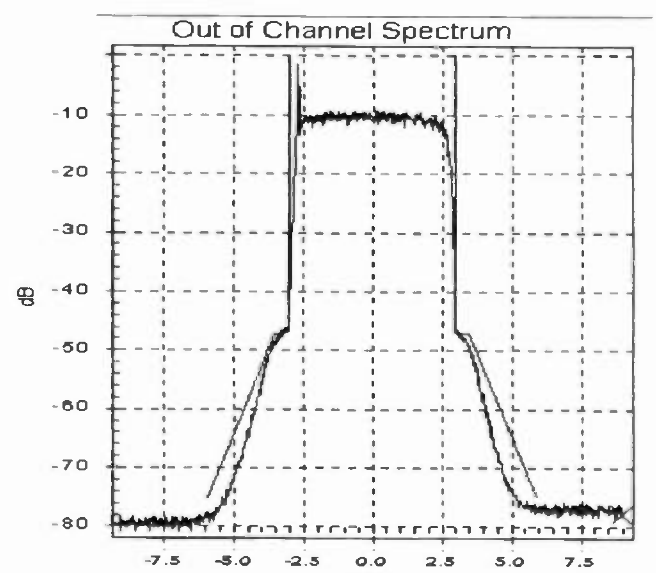


Figure 3. DTV mask filter on transmitter output to be fully mask compliant.

Notice in Figure 3 . the IMD spectral response is contained within the FCC mask limits. It appears from Figure 3, the FCC out of band emission issue is solved, however, there are a number of practical issues to be considered first.

Mask Filter Issues.

1. The response shown in Figure 3 is expensive to build.

2. The sharp tuned response shown in Figure 3 calls for elaborate temperature compensation to keep the transmitted band edges from being distorted during temperature drifting.

3. The filter can be configured as a single ended unit, lowest cost, or a more desirable constant impedance unit which is more expensive, approximately 2.5X the cost of a single ended unit.

4. Can another filter shape be substituted for the response shape shown in Figure 3 that is mask compliant and low cost?

Item number 4 above is the main thrust of this paper, to examine various filter shapes and to investigate the adjacent channel areas for potential interference that could be a major problem if left unaccounted for, particularly if the filter is not appropriate for lowest interference levels. It should be noted, the interference potential will, in most cases, be inflicted on the NTSC parent station, a cause for concern that could lose viewership.

The major test items in this paper for mask selection.

1. Examine the linearity solution and determine if it is an effective tool to be used as a stand alone item to suppress IMD products or if it is more appropriate that it be used in conjunction with the filter solution.

2. Investigate a sharp tuned mask compliant filter shape versus a wider shape that could lower filter costs but still be FCC compliant

3. Find out any potential interference to the parent NTSC station. when the filter shape is varied from a narrow mask compliant system to a wider, lower cost unit.

The Linearization Tool.

Figure 4 shows the effect of linearization. As the transmitter linearization system is adjusted for best performance, through the use of a manual or adaptive system, the out of band IMD components can be reduced.

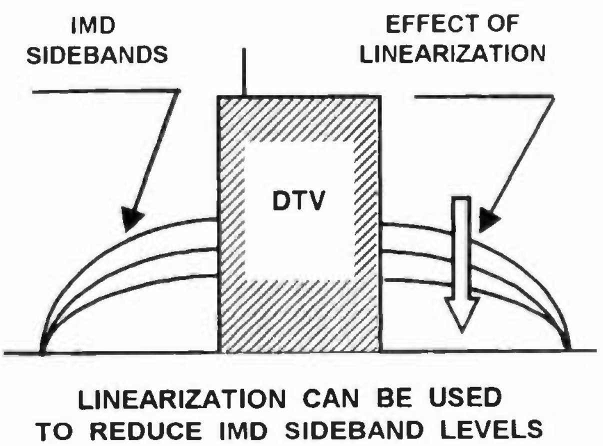


Figure 4. Effect of linearization.

The question, of course, is how much? It was first envised that linearization techniques could be used to reduce the IMD products below FCC requirements.. At least this was the thinking after the first FCC mask issuance., requiring -35 dB suppression at the shoulders and -39 dB in the adjacent 6 MHz slots.. Then the second mask requirement was issued asking for -37 dB at the shoulders and suggesting -44 suppression in the adjacent 6 MHz slots but further requiring a steep slope of attenuation from .5 MHz to 6 MHz in the adjacent slots to a level of -110 dB. This has certainly changed thinking about using a linearization technique to meet IMD suppression levels.

At the present time, a practical strategy is to use enough linearization to reduce the IMD shoulder level to be under the current mask requirement and to use a band pass filter to be fully mask compliant. The next issue is; what about the filter shape, since the shape factor is a cost item.

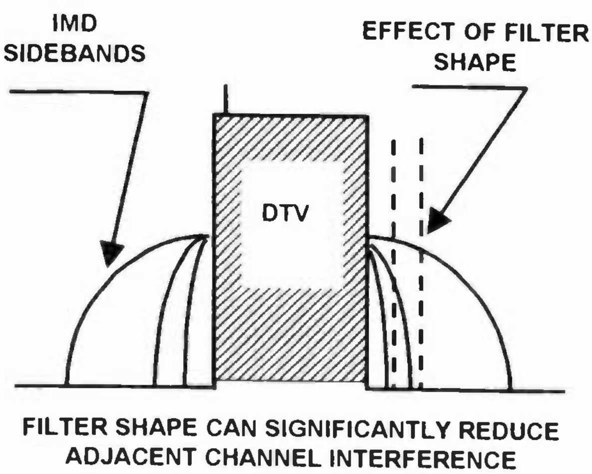


Figure 5. Using the filter shape to suppress out of band products.

In Figure 5, the dotted lines indicate two possible filter shapes for reducing the IMD components, a narrow shape and a some what wider shape. The scale is not indicated in Figure 5 but the narrow shape would fit under the current FCC mask out to .5 MHz from each channel edge and the wider shape would be stretched out another .5 MHz. These two shapes are next to be examined , along with linearization, as a method to meet the FCC mask

The gauge of measurement is the FCC mask shown in Figure 6. Only the lower half of the mask is shown for clarity and will also be used to set the level of linearization to see the effects of both methods.

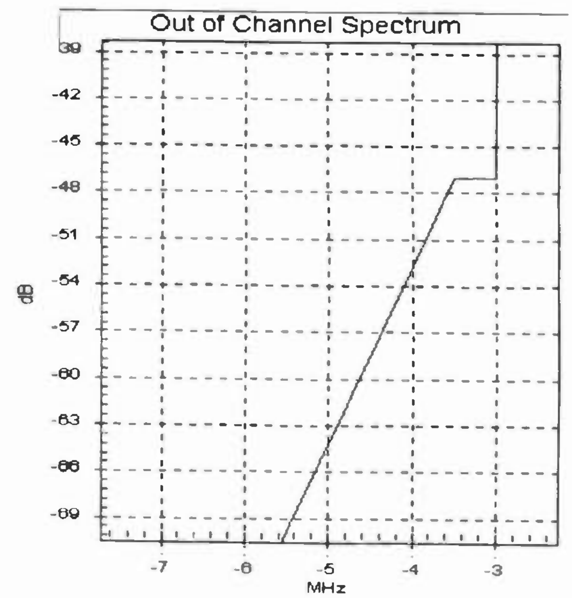


Figure 6. FCC mask limit, with expanded view of lower sideband edge. Note .5MHz line is at -47 dB on graph which is 11 dB below top of scale.

To begin the analysis, the transmitter output was adjusted for an IMD level to be at the mask .5MHz line. This is shown in Figure 7.

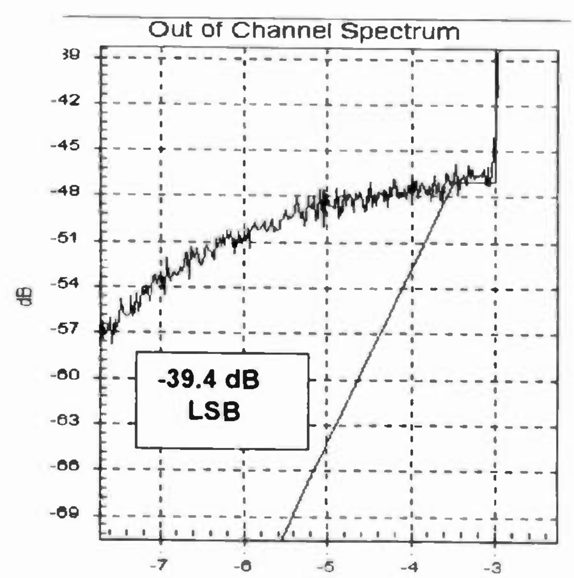


Figure 7. Spectral spread of a transmitter at the mask point resulting in a power level of -39.4 dB in the lower sideband (LSB) 6 MHz slot.

The spectral spread level shown in Figure 7 is a good starting point to see the effect of linearization. In this case , the transmitter llinearity circuits were adjusted to reduce the sideband IMD power to the new FCC suggested level of -44 dB. This is the total power in the lower 6 MHz slot as referenced to the total transmitted power in the authorized channel.

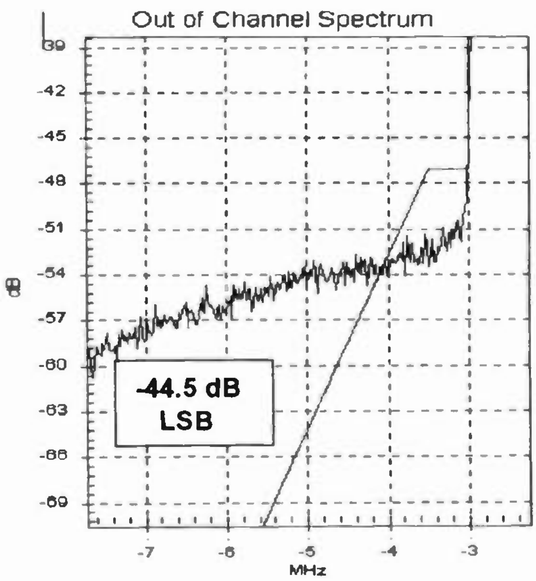


Figure 8. Reduction of out of band IMD sidebands through linearization which meets the -44 dB criteria, however, the lower 5.0 MHz area is over the FCC curve.

Since the amount of out of band energy has been reduced from -39.4 to -44.5 the -44 dB criteria has been met. Note, however, the IMD sidebands still spill over the FCC mask limit a little further into the lower 6Mhz zone. This says that even with about 5 dB of linearzation improvement, as shown in Figure 8 ,(the shoulder levels), the mask has not been met.in the 5.0 MHz area below the channel. A little later in this paper, it will be shown that this area is a potential interference zone to NTSC.

The additional 5 dB of linerization improvement required of the PA means the current operating value of -37 must be set to -42 dB which is at the practical limit of achieving stable linearization. Note, these numbers are referenced to the center of the channel while the Figures showing the LSB values are referenced to total power about 10 dB higher.

This shows that linearization is a weak tool in trying to achieve full mask compliance as a single item but it can be very helpful to set the shoulder levels to -37dB or less and in conjunction with a filter, provide a more secure mask compliant system .

The next question is; what about other filter shapes? Figure 9 shows a wider filter response which is less costly.

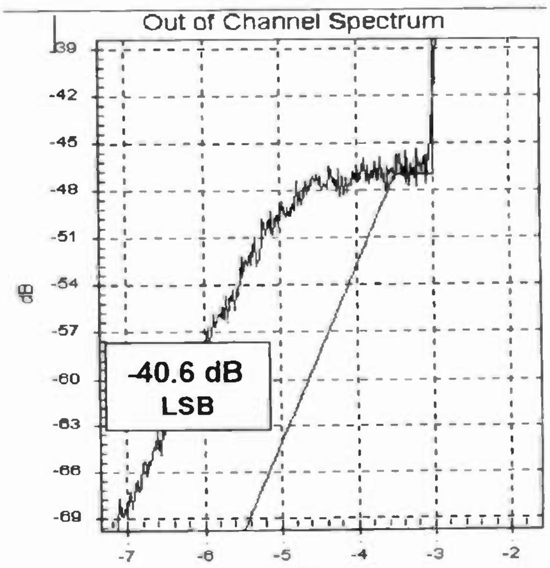


Figure 9. A wider filter response resulted in lower sideband power (LSB) of -40.6 dB, which is out of spec.

The wider filter characteristic as shown in Figure 9 does not meet the -44 dB criteria so an attempt to correct this was applied using more linearization to move the IMD shoulder level down another 4 dB. This is shown in Figure 10.

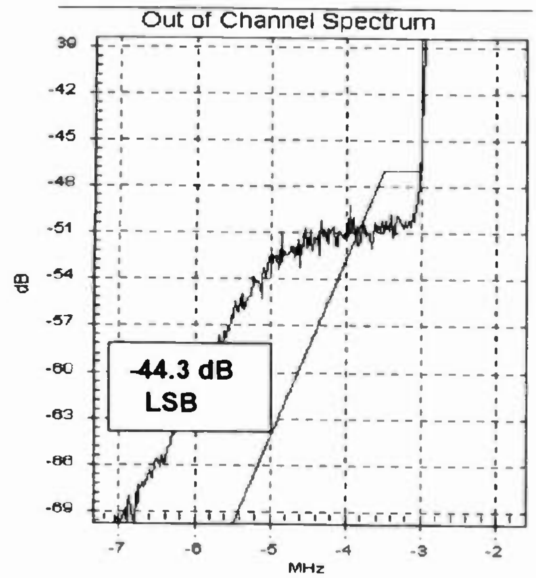


Figure 10. A wider filter shape with an additional 4 dB of linerization resulted in meeting the -44 dB criteria but is still over the FCC limit in the lower 5.0 MHz zone.

The effect of the wider filter was to increase the linearization requirement and still pose a problem in meeting the FCC mask in the lower 5.0 MHz zone. The wider filter approach does not appear to be an effective solution in spite of its lower cost. The main issue here is the potential interference in the LSB zone where NTSC stations are located. Also the upper sideband area.

A fully mask compliant system as previously shown in Figure 3, where the linearization was set to a practical level at the FCC mask .5 MHz line, can fulfill the mask requirement and provide efficient transmitter operation with acceptable, but less demanding linearization requirements. This also unburdens the adaptive equalization system where its primary mission is to iron out the system variations that occur as a function of temperature.

NTSC Interference Issues.

It is clear the FCC has attempted to remove interference problems by changing the radiation mask to its current demanding state. Refer to Figure 1.

In general, the FCC has accomplished its interference reduction task with the mask but it should be noted this applies only when other transmission parameters are also met and the principal one is the 12 dB NTSC peak of sync to DTV rms power ratio. If this ratio varies due to NTSC and DTV antenna patterns not matching in either elevation or azimuth planes, then the potential for interference from the DTV IMD sidebands spilling over into the NTSC channel can occur. Also, over lapping coverage zones that have different signal levels and various local areas where reflections can also alternate the 12 dB ratio. This section will examine this potential interference a little further.

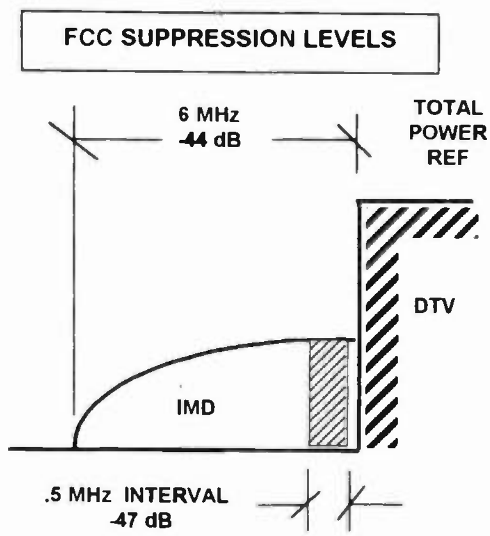


Figure 11. The FCC has indicated it is looking for at least -44 dB additional sideband suppression in the 6MHz zone and requiring -47dB suppression in the .5 MHz zone.

The interference suppression of -44 dB in the lower adjacent sideband zone shown in Figure 11 may not be enough for satisfactory operation and the issue of a wide filter response versus a narrow filter response will come back into consideration. This of course also applies to the upper sideband zone. In this regard, the sensitive areas of NTSC are to be identified next.

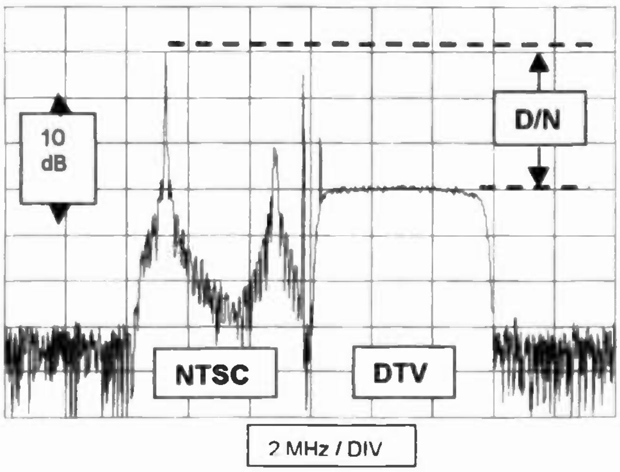


Figure 12. Upper adjacent N+1 DTV next to NTSC at the correct -12 dB ratio shown as D/N. Note RBW =30kHz.

In Figure 12 the D/N ratio (DTV/NTSC) was set using a power meter to accurately measure the total 6 MHz power in the DTV channel and then set the NTSC peak of sync 12 dB above it. This was done using black picture power (rms) and applying the usual NTSC 1.68 (2.25 dB) form factor.

It should be noted the nomenclature D/N ratio will be used in this paper as a less ambiguous term by stating N stands for NTSC and D stands for DTV in place of D/U ratios where U, the undesired signal, maybe either NTSC or DTV depending on which one is the designated interfering signal.

The 12 dB ratio, however, does not appear as such in Figure 12 because of the resolution bandwidth (RBW) setting of the spectrum analyzer, in this case 30 kHz. Many DTV and NTSC comparison observations on spectrum analyzers will display this characteristic which is a consequence of using practical RBW settings with various video filtering functions that produce the results as displayed in Figure 12.

Figure 13 illustrates a situation where the 12 dB planning ratio can vary by a considerable amount, due to differences in the DTV and NTSC antenna patterns, which can lead to a potential NTSC interference condition. A fully mask compliant filter system can help reduce this interference potential. The following graphs and pictures can help explain why.

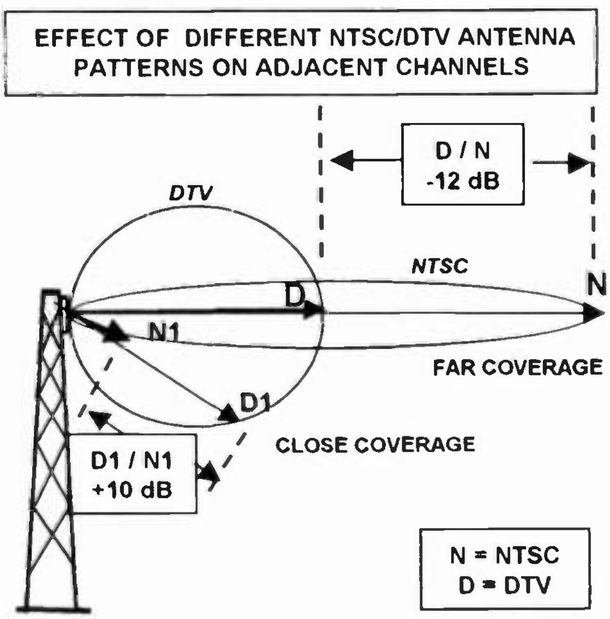


Figure 13. A low gain DTV pattern is compared with a typical high gain NTSC pattern to show how the 12dB ratio can be changed. The result is potential interference to NTSC.

The interference comes about from the spill over of DTV IMD products into the sensitive zones of NTSC. Refer to Figure 14.

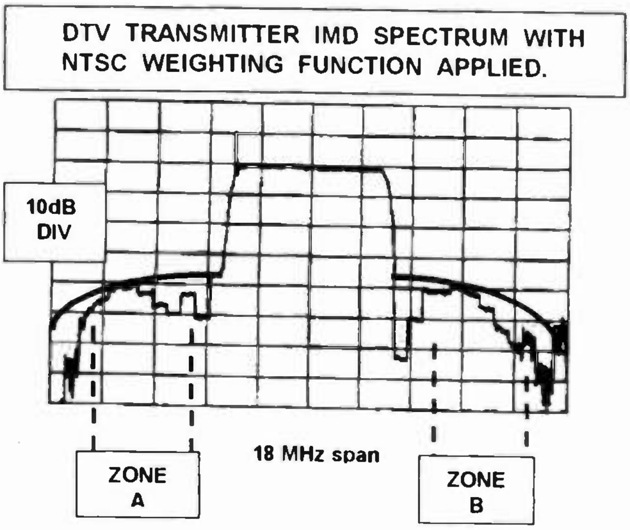


Figure 14. Zone A and B locate the NTSC sensitive areas. These zones include most of the weighting function points.

Figure 14 brings back into consideration the NTSC weighting function to help identify the NTSC sensitive zones. From this, a test setup was created that used an adjustable DTV mask filter, a DTV and NTSC exciter set on adjacent channels with the RF outputs combined in an isolated hybrid to observe the IMD noise in NTSC.

The monitoring system was set up using a Tektronix 1450 demod tuned to the NTSC channel. A Tektronix VM700 video analyzer was set to measure the resulting adjacent channel IMD noise on a quiet line in the vertical interval. The VM700 also provided the NTC7 video weighting function that correlates, to a reasonable extent, the observed picture impairments for noise.

The measured noise data was plotted with and without a DTV mask filter. This was an attempt to see how effective the mask filter is in reducing potential NTSC interference when the D/N ratio varies. Both a fully mask compliant filter and a wider, low cost filter were used with the characteristics as previously shown, Figure 3 and 9 respectively. Figure 15 shows the results of the measurements.

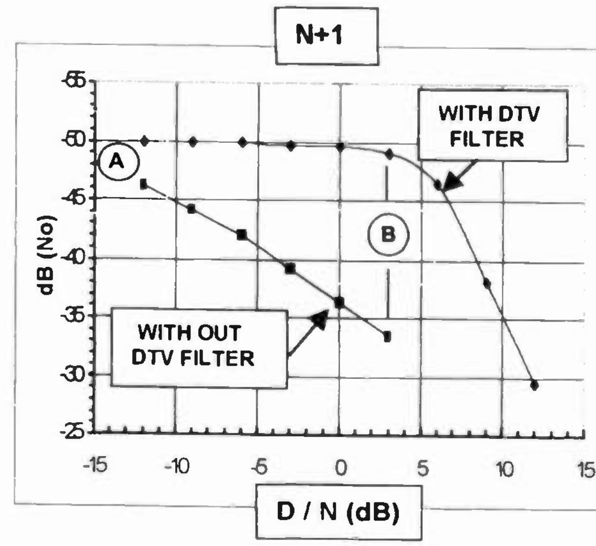


Figure 15. Video noise NTC7 weighted versus DTV to NTSC power ratio D/N for the N+1 case.

Figure 15 is interesting because it shows that with the DTV filter, the video noise remained nearly constant at -50 dB over the DTV/NTSC power range from -12 to nearly + 5 for minimum interference.

On the other hand, Figure 15 shows that without a DTV filter, the video noise was increasing over the same range to unacceptable values.

This is the main thrust of this analysis to show the beneficial effects of using a fully mask compliant filter as a means of protection against NTSC interference and other facilities should the D/N ratio vary in the coverage area.

The amount of improvement can be seen by noting the demodulated NTSC wavefoms in Figures 16 and 17 that show the resulting video performance at point B on the curves (Figure 15). Point A is the reference for the tests which represents the current planning D/N ratio of -12 dB. Point B identifies the location on the upper and lower curves (Figure 15) where the following pictures were taken

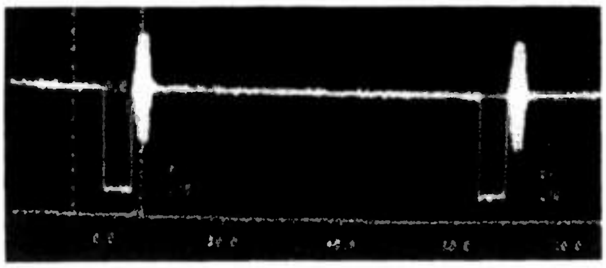


Figure 16. Demodulated NTSC sync at point B on the curve with the DTV mask filter in for D/N=+3.

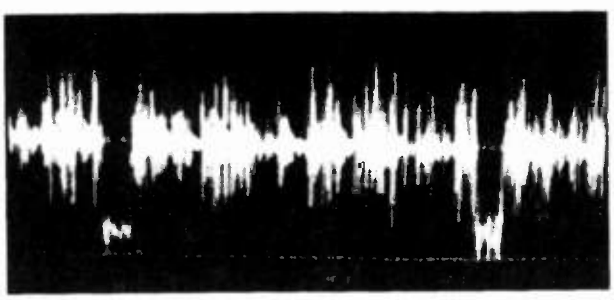


Figure 17. Demodulated NTSC sync at point B on the curve without the DTV filter just before the system lost sync and crashed for D/N=+3.

In spite of comments about the new FCC mask, Figures 16 and 17 show how effective it is in removing NTSC interference from DTV IMD sidebands when the D/N ratio changes. The starting point at A (Figure 15) shows the effect when the IMD shoulder level is at -37 dB for the conditions with and without a mask compliant filter. This is close to earlier test results. What is new here is the same test was repeated with a mask compliant filter that was not available in the earlier tests.

The Commission, in the latest rules, stated the threshold of interference for upper DTV into analog TV was -17 dB, which would be +17 on the scale in Figure 15.

Figure 15 further shows the crash point at the end of the upper curve at an D/N ratio of +12 dB (-12 FCC scale). This is more pessimistic than the FCC adjacent D/U ratios.

The object of this test, however, was to determine the effectiveness of the new mask filter and secondarily to observe the video threshold effects with high D/N ratios. For more accurate adjacent channel threshold values for interference planning, a new test setup would be recommended.

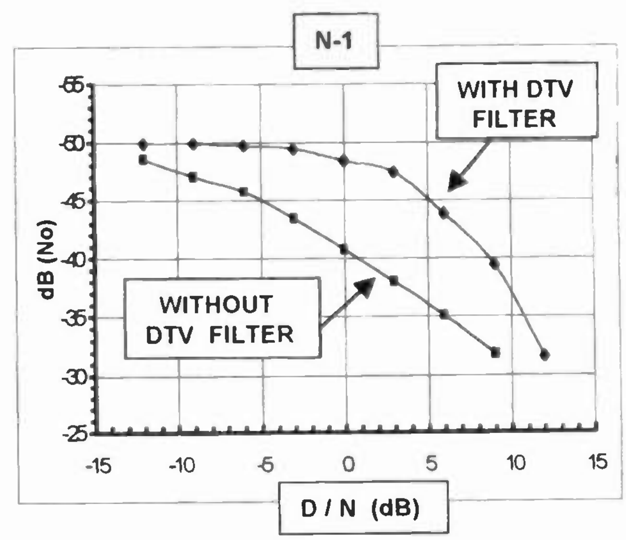


Figure 18. N-1 case showing video in the NTSC channel as a function of D/N power ratio .

Figure 18 shows N-1 case is a little more sensitive to DTV IMD spill over than N+1. This is in keeping with previous tests and the FCC threshold values for adjacent channel operation.

Continuing on with the adjacent channel tests, the following data plots use the same corresponding set up conditions but instead put in a wide filter response as shown in Figure 9.

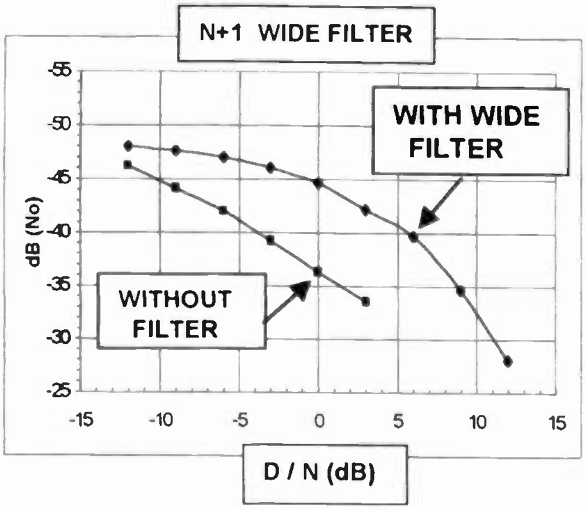


Figure 19. N+1 case showing video noise with a wide filter response.

The filter response case for N+1, Figure 19, does show some improvement but not as much as that shown in Figure 15 for a full mask compliant filter.

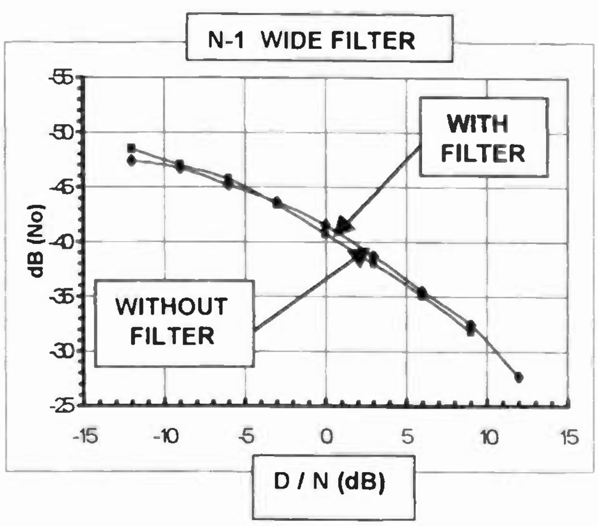


Figure 20. N-1 case showing little effect with or without a wide filter.

Figure 20 is surprising in that it shows no benefit from using a wide filter response approximately .5 MHz wider than a full mask compliant filter. A probable cause for this maybe the IMD sidebands are extending into the NTSC carrier for the N-1 case with sufficient amplitude to raise the detected video noise level. The full mask compliant filter removes this set of sidebands before reaching the visual carrier.

Wrap up comments.

The first part of this paper dealt with attempts to linearize the IMD components below the FCC requirements but it was noted that even if the linearization attempt dropped the IMD component below -44 MHz in the adjacent channel slots , the potential for interference still exists.

Using a wide filter strategy to help reduce DTV filter costs, increases the potential interference to adjacent NTSC services when antenna pattern matching is not practical to keep the 12 dB planning ratio in tact or if localized reflections change the NTSC to DTV received power ratio.

Conclusions.

1. It is highly recommended to use a fully mask compliant filter on DTV transmitter operations to reduce interference to a minimum.

References.

1. FCC Sixth Report and Order, 98-24, February 23, 1998, Section 92 and appendix E.-37.

2. ATSC VSB Transmission System Tutorial, Gary Sgrignoli, Staff Consulting Engineer, Zenith Electronics Corporation., pages 59 to 96.

On-Channel Repeaters for Digital Television Implementation and Field Testing

Walt Husak, Charles Einolf, and Stan Salamon
Advanced Television Technology Center
Alexandria, Virginia

Abstract

The Advanced Television Technology Center successfully developed and demonstrated an On-Channel Repeater (OCR) for Digital Television (DTV). The demonstration repeated WETA-HD's primary signal beyond the Blue Ridge Mountains into Charles Town, West Virginia. OCRs provide local broadcasters with a means to replicate or extend their current NTSC coverage while having no impact on the DTV allocation table. The OCR allows rebroadcast of a DTV signal, without frequency shifting, into an area previously unable to receive the originally transmitted signal. This paper describes the design, construction, and field testing of the ATTC On Channel Repeater.

Introduction

The ATSC Digital Television (DTV) standard through the use of the 8-VSB modulation technique, provides sufficient performance margins to allow for the practical introduction of an On-Channel Repeaters (OCRs). 8-VSB receivers have been shown to be immune from the effects of noise up to a 15.2 dB C/N ratio.[1] In addition, 8-VSB receivers can tolerate ghosts as large as -3 dB relative to the primary signal. Therefore, 8-VSB should permit the application of OCRs for DTV, even though additional noise and ghosts may be introduced. The Advanced Television Technology Center (ATTC) has undertaken a project to develop and test the OCR concept.

The project was divided into three phases. The first phase was a paper study to determine whether the system is feasible. A repeater would be constructed and tested in the second phase. The third phase tests the interaction between the repeater and the main transmitter, also known as mutual interference. The phases are sequential however, a new phase has been added extending the study of high power terrain isolated OCRs. This report summarizes the project up to the end of the second phase.

The paper study focused on issues such as the type of repeating action, repeater and receiver expected performance, and critical elements in the system.[2] The results of the study showed a repeater was possible and could augment a digital broadcaster's audience. The repeating action can be either regenerative, providing error correction prior to retransmission, or non-regenerative which provides no error correction. The study also highlighted antenna isolation and the repeater's main signal received power as being the two most critical elements contributing to the repeater's retransmit power.

In order to focus on the antenna design, the non-regenerative repeating action was selected as the design approach in the second phase. This would allow a reflection or antenna mutual coupling to manifest itself as a measurable ghost at the repeater's receive antenna. Experiments were conducted which tested antenna isolation both on and off the tower in order to understand the types of reflections and interferences that occur on a functioning tower. A terrain shielded site was chosen to isolate the repeater's transmission pattern from the main transmitter's signal.

Background

WETA-HD provided the primary DTV signal for the OCR implementation. The transmitter is located in Arlington, VA near Washington, DC. During the period of their experimental license, WETA-HD was broadcasting on UHF channel 34 with an ERP of 100 kW (+20 dBk or +80 dBm). The antenna had a cardioid pattern pointed toward the east. Charles Town and Harpers Ferry, WV could not receive channel 34 due to the Blue Ridge Mountains located between the transmitter and the target communities. An existing railroad microwave relay facility was selected for the repeater's location. The tower has a line-of-sight view of the WETA-HD transmitter and provided a panoramic view of the Shenandoah Valley.

Two antennas were selected for the repeater's receiver and transmitter. Antenna placement on the tower was chosen and appropriate wind and tower loading analysis performed. Interference and radiation studies were done in preparation for modifying the existing DTV license to include the repeater. No significant interference condition was discovered and an extension to WETA's license was granted.

System Design

Figure 1 illustrates the operation of the repeater. The system design includes a main transmitter, the repeater, and one or more receivers. The top of the ridge is approximately 1000 feet above the transmit and receive sites. This provided an excellent opportunity for line-of-sight to both locations. The resulting received power at both sites could be reasonably estimated using the free-space loss formula.

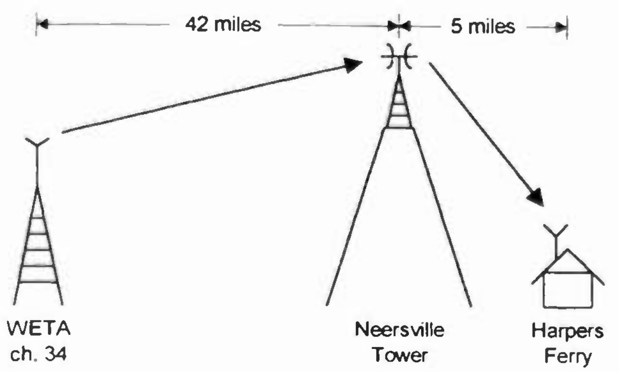


Figure 1

System topology of the OCR showing the relative location of all major sections.

Free space loss can be calculated using Equation 1.[3] The loss for the WETA to repeater leg is 125 dB using 68 km (42 miles) and 593 MHz (chan. 34).

$$L = 32.45 + 20Log(f) + 20\log(d)$$

Where :

L = Loss in dB
f = Frequency in MHz
d = Distance in km

Equation 2

Free space loss formula.

The received power formula is shown in Equation 2.[3] The estimated receive power at the repeater site is –50 dBm. An adjustment of 5 dB needed to be included to account for being on the backside of the cardioid.

$$P_{rx} = ERP - L - A_{pat}$$

Where :

P_{rx} = Repeater receive power at antenna
ERP = Primary transmitter ERP in dBm
L = Free space loss
A_{pat} = Cardioid adjustment

Equation 1

Repeater receive power formula.

The initial estimate of isolation between antennas was 110 dB. The coupling margin for proper operation of the final receivers and to prevent repeater oscillation was determined to be 20 dB. This reduces the useful isolation estimate to 90 dB. The projected repeater transmit power can be calculated using Equation 3. It was believed 10 dB of further isolation could be added once implementation was underway. This meant the target transmit power would be between 10 and 100 watts.

$$P_{tx} = P_{rx} + \mu_c - M$$

Where :

P_{tx} = ERP in dBm
P_{rx} = Received power in dBm
μ_c = mutual coupling (dB)
M = coupling margin (20 dB)

Equation 3

Repeater ERP formula

Using Equation 1 for the repeater to receiver leg yielded a loss of 106 dB. Retransmit power of +50 dBm would result in a receive power of –56 dBm. By applying the 131 dBuV/dBm dipole factor, a field strength 75 dBuV/m would be measured at Charles Town. This is 34 dB higher than the 41 dBuV/m set by the FCC.[4]

Tower Design

The design of the tower is important due to the interaction of the antennas. As much vertical and horizontal separation as possible was employed in order to reduce the amount of mutual coupling between antennas. The receive antenna was located above the transmit antenna because of a slight rise behind the tower which would have blocked WETA's signal.

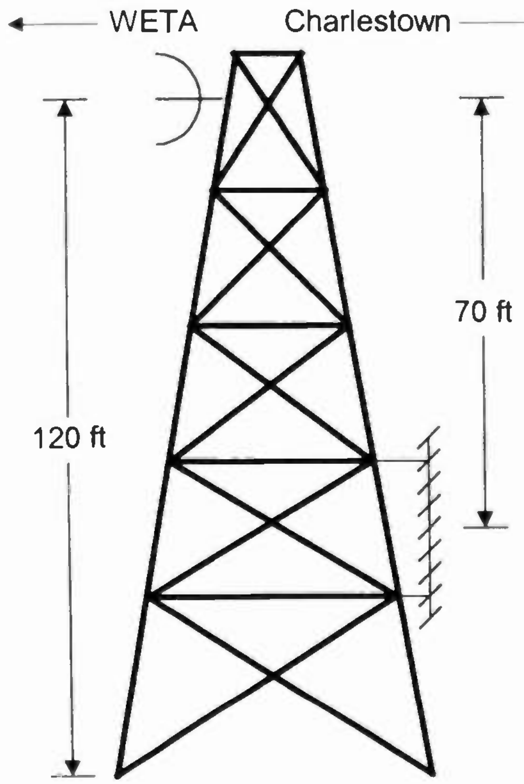


Figure 2
Antenna placement on the tower.

Figure 2 shows the general configuration of the tower. The position and orientation of the antennas on the tower is optimized for best reception and retransmission characteristics while while keeping the mutual coupling between the antennas to a minimum. In general, greater separation between the two antennas results in greater antenna isolation. The relative orientation of the antennas can be adjusted to further reduce the mutual coupling by making use of the nulls in each antenna pattern.

The receive antenna is a wide band, narrow beam, high gain antenna with a high front-to-back ratio in the horizontal plane. The transmit antenna also has a high front-to-back ratio with a wider beamwidth than the receive antenna. Both antennas have a narrow beamwidth in the vertical plane. This is accomplished by using a parabolic antenna for the receive antenna and a multiple stacked element transmit antenna.

Figure 3 illustrates the interaction between the transmit and receive antennas. The main transmitter signal is received at the node labeled P_{rx}. The signal is increased by the gain of the receive antenna and applied to the input of the repeater's electronics. It is then filtered and amplified. The processed signal is then further increased by the gain of the transmit antenna and radiated. A portion of the signal, represented by the chain of attenuators, feeds back into the receive antenna.

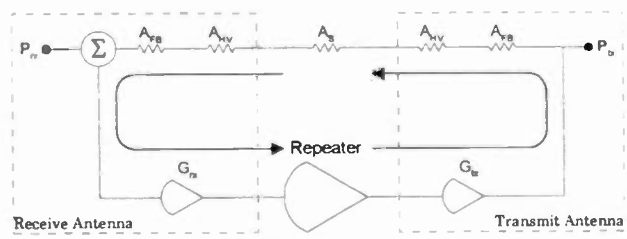


Figure 3
Simplified Schematic of the antennas.

The feedback is the result of mutual coupling in the antenna system. It affects three things, maximum retransmit power, performance of the receiver, and the oscillatory characteristics of the repeater. ATTC chose 20 dB of margin for the amount of feedback relative to the received primary signal, known as coupling margin.

Repeater Design

The repeater electronic design chosen for this phase was the non-regenerative analog design. Figure 4 is a simplified schematic of the design. The design consists of a preamplifier, a channel filter, and a power amplifier. A 6 MHz SAW filter, centered at 44 MHz, isolates the channel of interest from adjacent and near adjacent channels.

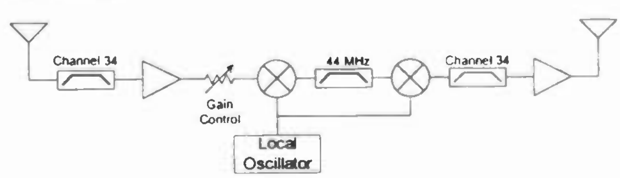


Figure 4
Simplified schematic of the repeater's electronics

The channel of interest is down converted to IF and then up converted back to the original channel by using the same local oscillator in order to ensure small differences in frequency will not manifest itself as dynamic multipath. The effects of mutual coupling appear as multipath due to the feedback being correlated with respect to the primary signal.

The preamplifier stage has a channel 34 bandpass filter in order to eliminate the chance of any N+14 and N+15 taboos. The heterodyne process uses high side injection eliminating N-14 and N-15 taboos. As an additional benefit, channels beyond N±5 are attenuated by more than 30 dB, reducing the possibility of front-end overload. Two notch filters for the visual and aural carriers of channel 32, N-2, were installed to further reduce the possibility of overload and intermodulation. The gain control ensures none of the subsequent amplifiers will either compress or cause excessive intermodulation products.

The IF stage begins with a single heterodyne stage feeding into a SAW filter. Not shown is an amplifier to mitigate the insertion loss of the SAW filter. Also not shown is a 100 MHz low pass filter used to eliminate the residual on channel signal from the IF stage. Once the adjacent and near adjacent signals are filtered, the signal is upconverted back to the original channel.

The power amplifier stage begins with another channel 34 bandpass filter. The purpose of this filter is to attenuate any out-of-channel energy such as the IF signal, the image, and the LO. The amplifier shown is a representation of multiple cascaded stages with a final small RF power amplifier.

Implementation

Once the required leases and licenses were obtained, construction of the shelter began and utilities were installed. The biggest unknown was what the final mutual coupling was going to be. As stated earlier, the performance of several key parameters is directly related to mutual coupling.

The first set of experiments characterized mutual coupling. WETA turned off their transmitter allowing ATTC to transmit a locally generated signal at the repeater without the primary signal being present. This allowed ATTC engineers to do two independent tests to confirm what the mutual coupling would be.

The first test performed was a locally generated signal with an ERP of +42.2 dBm. Figure 5 shows the results of the test. The measured receive power can be approximated as –90 dBm. After adjusting for channel bandwidth, downlead loss, and antenna gain, the power at the receive antenna is –80.6 dBm. The difference, 122.8 dB, is the total isolation between the co-located transmit and receive antennas.

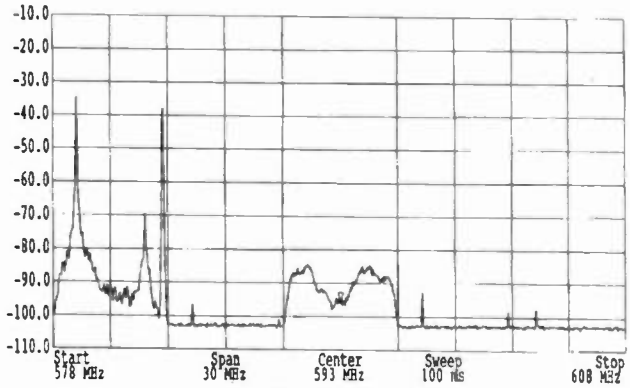


Figure 5

Mutual coupling experiment using a single ended transmission

A two port network analysis was done in order to confirm the isolation in the previous experiment. The results from this experiment are shown in figure 6. Note the similarity of the general shape of the frequency response within the channel between the two methods.

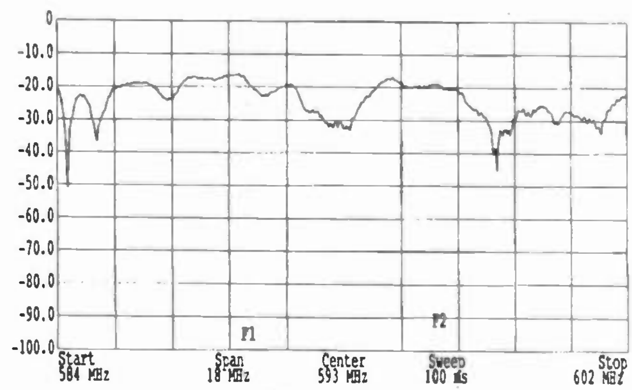


Figure 6

Network analysis of the antenna isolation.

Due to the limited dynamic range of the network analyzer, an RF amplifier was calibrated into the test setup to increase the input sensitivity. A 70 dB pad was also calibrated into the system and removed from the measurement, creating a 70 dB offset. The result is a 25 dB reading. After adjusting for the antenna gains, feedline losses, and the attenuator, the calculated isolation is 123.7 dB. The average of the two measurement techniques is 123 dB.

The isolation determines the repeater's ERP. The ERP is calculated by the formula given in Equation 3. Figure 7 is the spectrum plot taken at the downlead of the repeater's receive antenna of WETA-HD's spectrum. Adjusting the measurement for channel bandwidth, downlead losses, and antenna gain results in a receive power of –52.4 dBm at the antenna. By applying the receive power, mutual coupling, and the coupling margin to Equation 3, the final ERP for the repeater is +50.8 dBm or 120 watts.

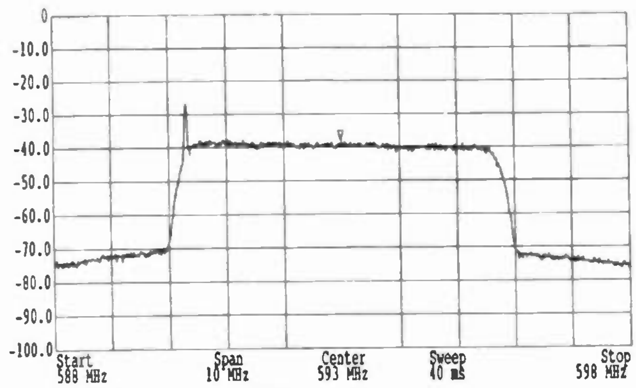


Figure 7

Receive spectrum of WETA-HD at the repeater.

The repeater was then optimized for the new ERP and a final transmit plot was generated. Figure 7 and 8 are the before and after plots respectively of the repeater with and without the repeating action. It is interesting to note the ripple in the spectrum due to multipath formed at the receive antenna because of mutual coupling. The nulls are spaced approximately 250 kHz apart, corresponding to an approximately 4 µsec delay.

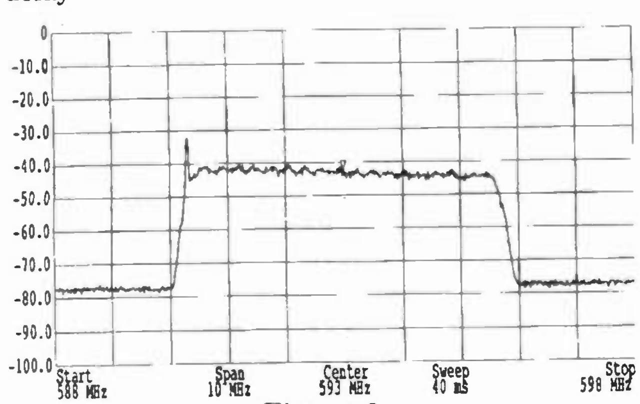


Figure 8

Spectrum plot of the repeating action

A further confirmation that the repeating action is occurring correctly, is shown in Figure 9. This is a graph of the dynamic equalizer tap energies from the ATTC/Wuppertal professional demodulator receiving the repeated signal. The horizontal axis is delay time and the vertical axis is relative tap energy verses the main signal.

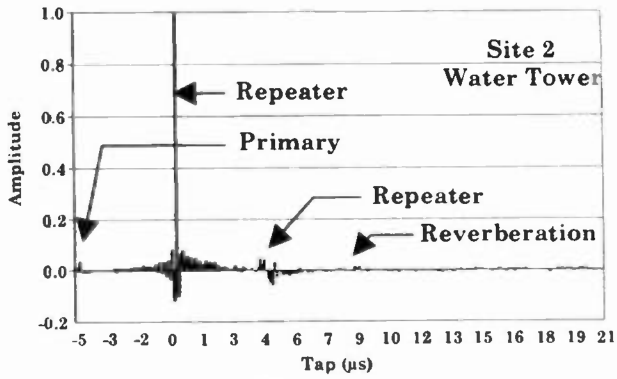


Figure 9

Receiver Dynamic Equalizer tap weights relative to the main signal

Three ghosts of interest can be seen in the plot located at –4.8 µsec, 4.4 µsec, and at 8.8 µsec. The first ghost is the weak primary signal directly from WETA. It appears as a leading ghost and is not delayed by the repeater. The repeating action inherently causes the second ghost. The third ghost is the result of the reverberation within the feedback loop of the repeater.

The definitive experiment involved using a consumer DTV set-top box (STB) to actually receive video and audio. The set-top box and monitor were provided by Panasonic. The receive site was located at the intersection of US Rte. 340 and WV Rte. 9 ("Burger King"), 5.2 miles from the repeater. The site is boresight with the repeater's transmit antenna which offers line-of-site to the tower. The video and audio appeared without any errors as if we were receiving the primary signal without being repeated.

A sensitivity margin test was also conducted. The input signal was attenuated until the picture exhibited blocking artifacts. The receiver was able to withstand nearly 16 dB of additional attenuation before losing lock. Factoring in the power split for additional test equipment and the extra downlead losses, the total margin was 21 dB.

ATTC arranged a demonstration and technical briefing to show the OCR works in a real world environment. The Charles Town Racetrack was selected as the venue because of the location and the ability to see the repeater's tower from the room where the demonstration would take place. A Radio Shack double bowtie antenna was used as the receive antenna and the receiver was the same Panasonic STB/Monitor combination used previously. The entire system was indoors including the antenna.

The demonstration was conducted without reception errors. ATSC provided the video programming. Representatives from ATSC, NAB, CEMA, WETA, and PBS were the first to see an 8-VSB signal repeated without frequency translation.

Field Tests

MSTV provided the use of their field truck to do field tests just prior to the "Great Channel Swap" in the Washington, D.C. Market. On November 1, 1998, the experimental channels used by WETA-HD and WHD-TV were exchanged for permanent licenses on different channels. This gave us several days to execute a preliminary field test based on the original repeater ERP.

A subset of eleven sites from the fifty total sites was performed with satisfactory results. The eleven sites, shown in Figure 10, included five in an urban cluster, five along an arc creating radials, and one behind the repeater to measure the effects of mutual interference with the primary transmitter. There was no measurable mutual interference due to the antenna

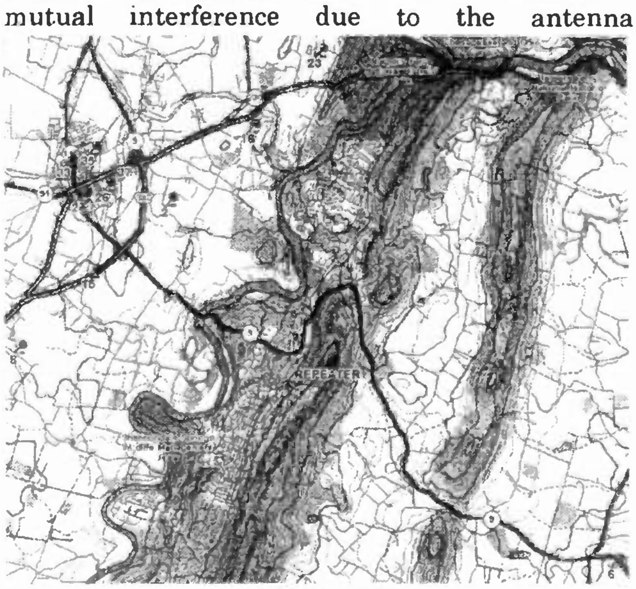


Figure 10
Preliminary Field Test sites.

orientation and terrain shielding.

The final field tests are scheduled for the first two weeks of February 1999. Figure 11 shows the fifty sites selected. The new owner of channel 34, WUSA, has agreed to continue supporting the OCR project.

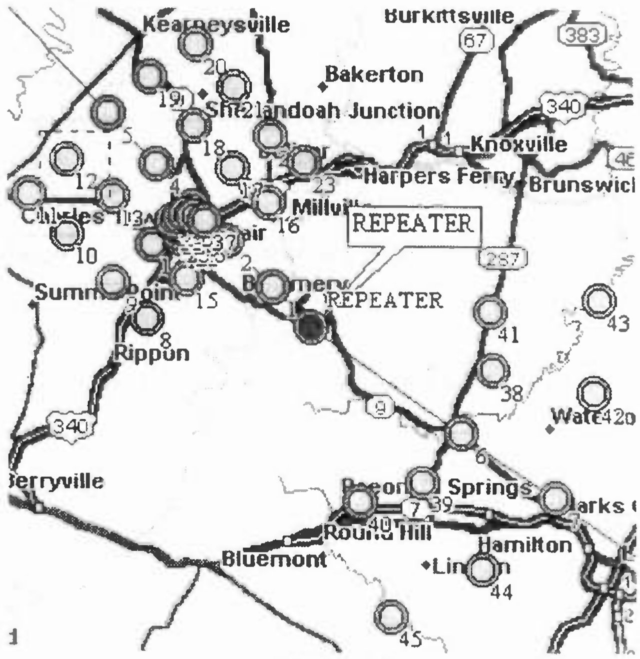


Figure 11
Field Test sites for the high power OCR tests.

WUSA is broadcasting +88 dBm (632 kW) with an omnidirectional antenna. The expected receive power at the repeater will be –37 dBm using Equations 1 and 2. The repeater's maximum ERP can be +66.2 dBm (4.2 kW) by using Equation 3. The tests will be conducted using +60.8 dBm (1.2 kW) for easy correlation to the previous field test results.

The increase in primary transmitter ERP will result in a higher primary signal received beyond the ridge. This will increase the mutual interference measured in the valley. The repeater and the receivers will still operate correctly due to all signals being scaled by the same amount. The repeater's signal, the primary signal, and the interaction between the two (mutual interference) will be measured at all sites. The results will be presented at NAB99.

Conclusions

The On-Channel Repeater works in a terrain isolated scenario. A consumer DTV receiver can demodulate and decode an 8-VSB signal repeated into an area shielded from the primary signal. The primary interference mechanism is multipath created by co-locating receive and transmit antennas. This phenomenon is known as mutual coupling.

Off-the-shelf antennas that are mounted on a tower in a conventional manner can achieve as much as 123 dB of antenna isolation. Given a design similar to the ATTC OCR and using a 500 kW UHF station as the primary signal, the repeater's ERP can as much as 3 kW.

Since the repeater extends coverage into a new area, the same interference issues as the primary broadcast exists, e.g. adjacent channel, co-channel, and taboos. Furthermore, great care must be exercised in the design and implementation in order for the repeater to effectively retransmit a DTV signal.

In the spring of 1999, ATTC will begin the third phase of the OCR project. This phase will address the interaction between the primary signal and the repeated signal. The regenerative electronic design will be also tested allowing the use of error correction at the repeater site.

---

[1] *digital* HDTV Grand Alliance System, Record of Test Results, Federal Communications Commission, Advisory Committee on Advanced Television Service, October 1995.

[2] Charles Einolf and Walt Husak, "On Channel Repeaters for Digital Television," Proceedings 1998 Broadcast Engineering Conference, April 1998.

[3] Donald G. Fink and Donald Christiansen, "Electronics Engineer's Handbook 2nd Edition," McGraw-Hill, New York, NY, 1982, pp. 18-57.

[4] FCC, Fifth and Sixth Reports and Orders, MM Docket 87-268,both adopted April 3, 1997, amended by the Memorandum Opinion and Order on Reconsideration of the Sixth Report and Order, adopted February 17, 1998.

A High Power, Low Beam Voltage, TV Broadcast IOT Amplifier, with an Optional Multistage Depressed Collector

Tim Crompton, Steve Aitken, Steve Bardell, Roy Heppinstall and Geoff Clayworth
EEV Limited, Chelmsford, England.

ABSTRACT

A new high power Inductive Output Tube (IOT) has been designed, with versions suitable for transmission of either analogue or digital television. The tube can produce up to 35 kW average output power in 8-VSB digital television transmitters. The analogue version of the tube can produce up to 77 kW of peak sync. vision power and 7.7 kW of aural power at its output flange.

A multistage depressed collector for this tube has also been designed to combine increased efficiency with reliability and simplicity and thus provide the best overall cost-of-ownership profile.

Some novel features of the design are described along with the computer modelling techniques used. Experimental results of both analogue and digital performance are also presented.

INTRODUCTION

1998 saw the start of commercial digital television transmission in the USA. Most of the stations on air use an IOT as the final amplifier in the transmitter.

EEV introduced the Inductive Output Tube (IOT) to service in 1991. The EEV IOT is now recognised, used worldwide and is the market leader as the final amplifier for television transmitters. Transfer characteristics of this type of tube are inherently linear which means that it can be operated as a vision-only amplifier or as a common mode amplifier for analogue systems and also it is particularly suitable for digital TV services. EEV IOTs are being used in the majority of modern, high power television transmitters in the USA. They have already achieved individual lives in excess of 50,000 hours, and a total life of over 15,000,000 hours. The range of IOTs available in terms of power rating goes from 22 + 2.2 kW to 77 + 7.7 kW for analogue operation (Table 1) and from 10.5 to 35 kW average digital output power (Table 2 overleaf).

	Transmitter Rating (kW)	Tube Rating (kW)	Circuit Type
Liquid Cooled			
IOT9707	70 + 7	77 + 7.7	Plug-in
IOT8707	70 + 7	77 + 7.7	Build-up
IOT8505	50 + 5	55 + 5.5	"
IOT8404	40 + 4	44 + 4.4	"
IOT8303	30 + 3	33 + 3.3	"
IOT9303W	30 + 3	33 + 3.3	Plug-in
IOT9202W	20 + 2	22 + 2.2	Plug-in
Air Cooled			
IOT8303R	30 + 3	33 + 3.3	Build-up
IOT9303R	30 + 3	33 + 3.3	Plug-in
IOT9202R	20 + 2	22 + 2.2	Plug-in

Table 1. The EEV Range of IOTs, for use in Common Amplification Analogue TV Transmitters.

In the design of a new IOT, the performance as a digital signal amplifier has become very much the key parameter. Digital terrestrial television signals necessitate a much larger dynamic range than that of an analogue televison signal. This is because digital TV signals, like noise, have an approximately gaussian amplitude distribution that means they have high peak-to-mean power ratios. To accommodate the larger dynamic amplitude range of the digital signals without distortion the IOTs are run at lower average power compared with their analogue TV rating, to ensure they are operating linearly and ensure sufficient headroom for the high power peaks. Therefore a tube is required that will deliver very high peak powers while maintaining good efficiency at low average powers.

	Transmitter Power Rating (kW)	Tube Rating (kW)		Circuit Type
	Peak	Peak	Av.	
Liquid Cooled				
IOTD3130	120	135	35	Plug-in
IOTD2130	120	135	35	Build-up
IOTD2100	100	110	30	"
IOTD270	70	75	19	"
IOTD150W	50	55	14	Plug-in
IOTD140W	40	42	11	Plug-in
Air Cooled				
IOTD150R	50	55	14	Plug-in
IOTD140R	40	42	11	Plug-in

Table 2. Range of Digital IOTs

The latest addition to the EEV range of digital IOTs is the IOTD3130 (Figure 1), which has a peak digital output power capability of 135 kW, suitable for a transmitter capable of a digital peak output power of at least 120 kW. The RF input has a good match over a wide bandwidth and low levels of regenerative feedback, which make it highly suitable for automatic precorrection of the digital signal.

The design of this tube, IOTD3130 and its analogue equivalent IOT9707 is the subject of this paper.

DESIGN CONSIDERATIONS

The design of this tube followed the concept of the low power plug-in IOTD150W described in previous papers[1][2]. This is now a well-proven IOT that is rugged, compact and less complex than previous designs. Some modifications were made to the existing design to fit the new electron gun structure and simplify the external RF contacts. In addition, concentric ring contacts were adopted for the high voltage, grid bias and heater connections. Familiar established technologies were used for the output circuit of the system.

Electron Gun

From the start the maximum beam voltage was kept below 38 kV to reduce any problems in power supply design and facilitate the future use of high frequency compact power supplies. This implies a lower beam voltage for a given power level than all previous IOT designs. For instance, for a peak output power of 95 kW, the existing IOTD2100 requires a beam voltage of 35 kV, whereas the IOTD3130 requires a beam voltage of only 30 kV. Designing an electron gun to meet these requirements is a decidedly non-trivial exercise. The electrodes must maintain the required beam current and profile while keeping the electric fields to an acceptable level to prevent voltage breakdown between electrodes. The grid must be designed to be far enough from the cathode to ensure that the sensitivity to cathode position is minimised while ensuring that the grid does not need a positive voltage with respect to cathode at the peak of the RF cycle.

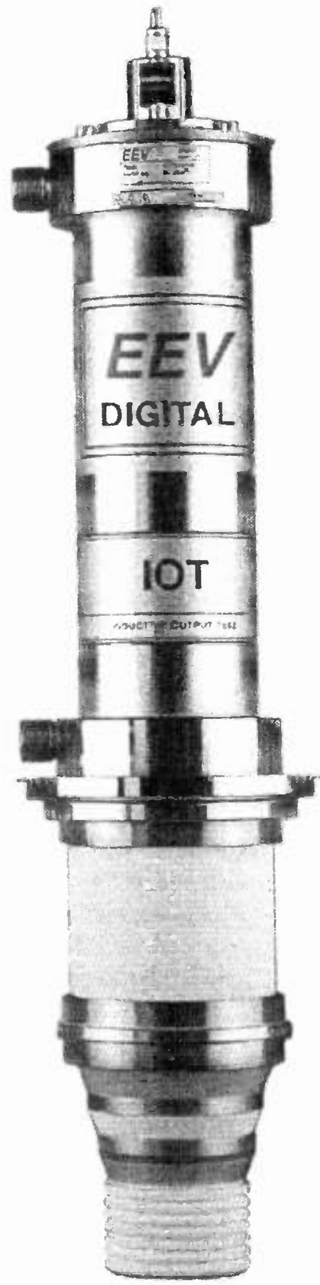


Figure 1. IOTD3130

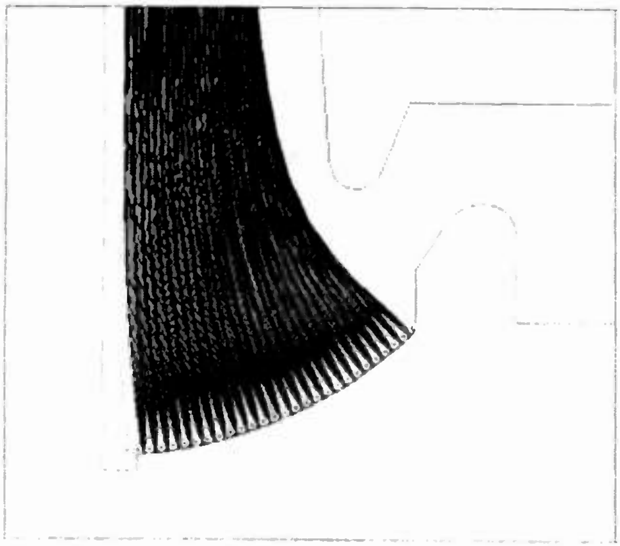


Figure 2. OPERA-2D Predicted Electron Trajectories

Initially, the effect of the electric and magnetic field on the electron beam was modelled using a finite element analysis program OPERA-2D[3]. This program was used to model the detail of the grid geometry and predict the electron trajectories through the whole tube from cathode to collector. Figure 2 shows the predicted electron trajectories near the cathode. If the grid voltage is varied with respect to the cathode voltage, then a grid transfer characteristic can be plotted. The amplification factor μ can be calculated and a prediction made for the gain of the tube, by comparison with IOTs of known gain and μ. In addition, the required grid bias voltage for a particular idle current can be estimated. The calculated values for voltage and gain were $V_{bias} = -200$ to -170 V, depending upon the idle current, and a 20 to 21 dB gain. These predictions are in excellent agreement with analogue TV experimental results.

The mechanical design of the electron gun and the RF input system is similar to that used on the low power plug-in tube IOTD150W. The cathode, grid and focus electrode are assembled inside a single ceramic cylinder, which is also part of the vacuum envelope of the tube. The high voltage is isolated from the input cavity by capacitive couplings formed by metallising the inner and outer surfaces of the envelope ceramic. In this tube, the ceramic has several conical sections (Figure 3), so that a larger cathode can be introduced while retaining the same basic design of input cavity as the low power tube. The input cavity RF connections directly contact the metallised outer surface of the capacitive coupler without the need for contact rings to be brazed to the outer surface of the ceramic.

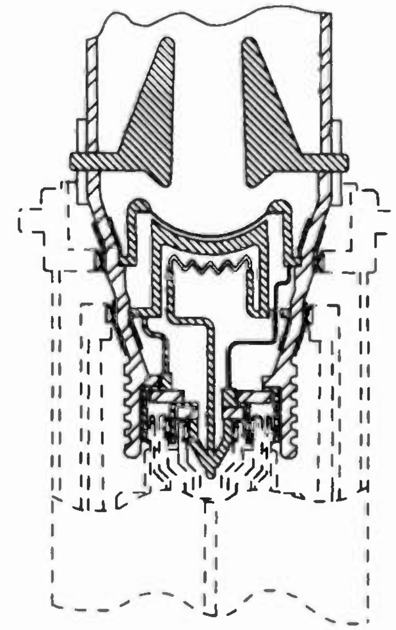


Figure 3. Schematic of Input System

This design has the advantage that all the circuit external to the vacuum envelope is at earth potential. Other IOT types have the high voltage isolation of the input cavity external to the vacuum envelope. This can lead to complex structures if voltage breakdown problems are to be avoided. The new tube is also a true plug-in system. The circuit is shipped coarse tuned to a given channel and installed as a complete unit with no on-site assembly and if required can be mounted in a standard 19-inch rack.

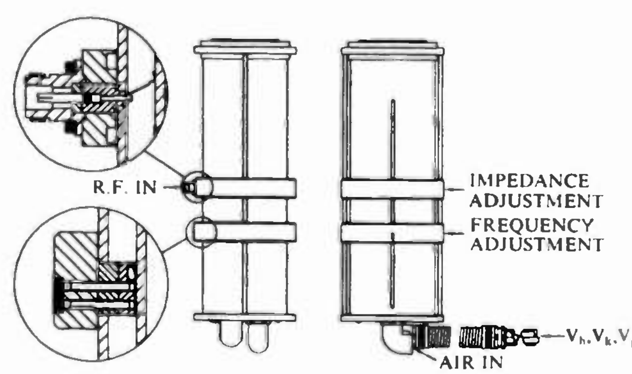


Figure 4. Slotted Input Cavity Design

Input Cavity

Because the high voltage isolation is integrated with the tube vacuum envelope, the input cavity can be a very simple design. It is a slotted, cylindrical cavity in which the tuning is effected by manually adjusting the lower ring, which is attached to the tuning door, to the required position (Figure 4). The RF input connection is attached to the upper ring and fed to a probe that makes contact with the inner conductor of the cavity. The upper ring is positioned for an optimum match. No external slug or stub tuner is required. With this system, the whole frequency range (470 to 810 MHz) can be covered by a single mode. The design leads to an excellent match over a wide bandwidth with reduced regenerative feedback, which in turn means that the tube has excellent digital performance characteristics.

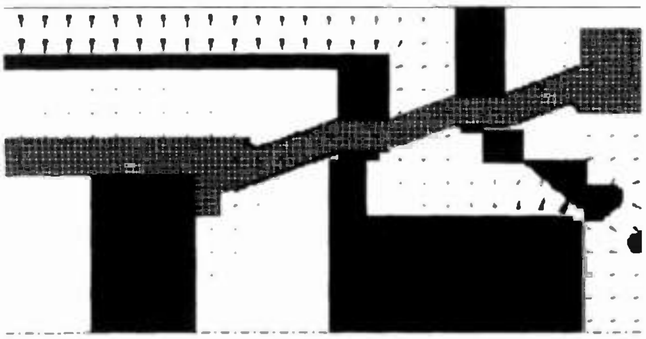


Figure 5. MAFIA Plot of the RF Electric Field

The optimum cavity dimensions were calculated analytically and by using a lumped circuit model. The performance was analysed for frequency response and the presence of spurious modes using a finite element integration code MAFIA[4]. Figure 5 is a typical MAFIA output plot. The triangles represent the magnitude and direction of the electric field at a particular RF frequency. The RF input probe is not shown in this plot. Several simulations have been carried out showing how the fundamental frequency varies with door position and if there is likely to be any interaction with the septate (dipole) modes of the cavity. Also the resonant frequency of the anode-grid space can be found and the effectiveness of any lossy material in damping this resonance can be modelled. Septate modes and anode resonances were found not to be a problem with this cavity because of the small cavity diameter.

Tube Body

In designing the body of an IOT the magnetic focusing field must be considered, to ensure the electron beam is not intercepted on the tunnel walls, and the RF interaction gap that has to be designed for optimum conversion efficiency. The shape and velocity profile of the electron beam after it passes the output cavity gap is highly dependent upon its interaction with the RF fields and the static electric and magnetic fields.

Modules from the MAFIA finite integration suite of software were used to model the interaction of the electron beam with the static and dynamic fields. In particular module S (Static solver) calculates the static magnetic and electric fields. Module E (Eigenmode solver) calculates the dynamic RF fields in the cavity. Module TS2 (Particle solver) models the interaction of the previously calculated fields with bunches of electrons emitted from the cathode.

In practice the magnetic circuit is modelled independently and the calculated magnetic field is interpolated onto the mesh of the tube model. This is done to speed up the particle calculations, because the magnetic field is modelled over a much larger volume than the vacuum envelope of the tube. Also different magnetic circuits can be modelled and their effect on the electron beam analysed using the same tube model each time. Figure 6 shows a MAFIA plot of the magnetic flux.

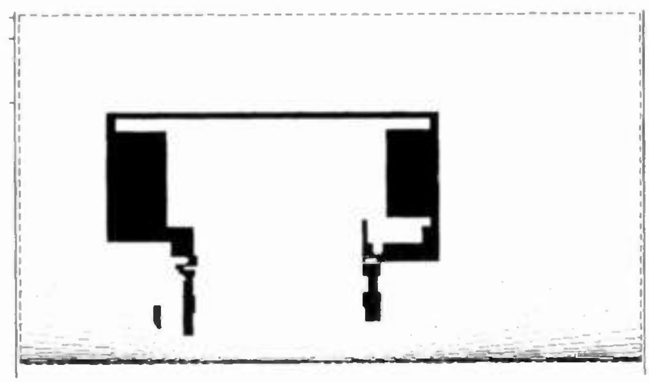


Figure 6. MAFIA Plot of the Magnetic Flux

The electron gun, interaction gap, output cavity and collector were modelled as one unit in two dimensional axi-symmetric geometry. MAFIA can model fields and beams in 3D but this is not necessary here. The output cavity was modelled as a cylinder into which some loss was introduced, to reduce the Q value to the known measured Q of an IOT cavity. The resonant frequency, Q, R/Q, and the complex dynamic field pattern were computed using the Eigenmode solver. This dynamic field is initiated at the correct phase and approximate amplitude at the start of the particle solution.

The input cavity and grid were not represented in the tube model. Instead a pulse shape and mean current emitted from the cathode were assumed. Theoretically the pulse shape should follow the function $(\sin x)^{1/2}$ with the amplitude offset so that at $x = 0$ and $x = \pi$ the value of the function is equal to the idle current. (The idle current is the beam current flowing when no RF signal is applied to the tube). In practice the exact pulse shape may vary depending on the shape of the grid characteristic around the operating point. Fortunately, reasonable variations in pulse shape do not significantly affect the calculated parameters. Bunches of negatively charged particles, each carrying several million electron charges, are initiated at the cathode surface. The charge density is made to vary in time and space to form the correct bunch shape at the chosen mean beam current. The initial thermal velocity and angular velocity imparted by the grid wires are also taken into account. Each bunch contains a few thousand particles and each of these particles is tracked as it interacts with the static and dynamic electromagnetic fields and the space charge of the surrounding electrons. The amplitude of the RF field in the output cavity changes from its initial value to a final value when steady state is reached. Depending upon how good the initial approximation is, this usually occurs after between five and twelve bunches have passed the output gap. At this point the simulation is ended.

The output power, efficiency, intercepted energy, charge and many other parameters can be calculated using MAFIA's post-processor. Figure 7 is a plot of particle bunches travelling through the tube.

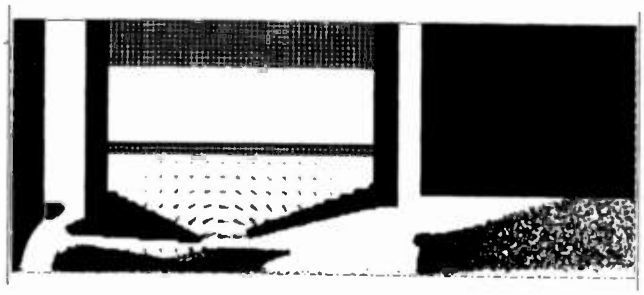


Figure 7. MAFIA Plot of Particle Bunches

Collector

The IOTD3130 uses a single-stage, scrolled, liquid cooled collector of conventional design but it is known that the use of a multistage depressed collector will further reduce the energy input required by the tube, for a given output power level.

THE MULTISTAGE DEPRESSED COLLECTOR

The idea of using a depressed collector on an inductive output tube is not new. Indeed, Andrew Haeff, the inventor of the IOT, considered it almost essential[5]. A multistage depressed collector has one or more electrodes set at potentials between ground and cathode potential. The electrodes set nearer the cathode potential collect high velocity electrons and those at or near ground potential collect the low velocity electrons that have lost most energy to the RF field. With more electrode stages electrons can be collected closer to their appropriate potential, when they just

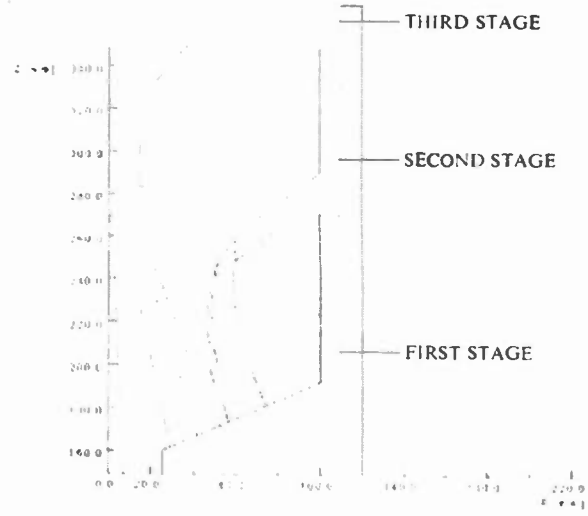


Figure 8. OPERA-2D Plot of Secondary Electron Paths

intercept the electrode with zero velocity. This results in a very efficient tube. However, there is a problem with secondary electrons in depressed collectors. Secondary electrons are produced when electrons strike a surface. Typically one or more electrons are emitted for each electron that hits the surface, more if the electron does not strike vertically. These secondary electrons can be either accelerated into the previous electrode stage or back down the beam tunnel (Figure 8). In either situation, they can seriously affect the efficiency of the IOT. The best way of reducing the detrimental effects of secondary emission is to establish a suppressing field at the surface of each collector electrode. This can be achieved by arranging each electrode to provide the suppressing field for the previous electrode and by having a spike at cathode potential protruding down the centre of the collector[5].

When designing the depressed collector for the IOTD3130 collectors were modelled with between three and five stages. Four- and five-stage collector designs provided small increases in efficiency over the three stage designs but they needed to be a much larger radius to work effectively. It was decided that the three-stage design provides the best compromise between efficiency, complexity and cost.

Figure 9. MAFIA Plot of Equipotentials in a Typical Three-stage Collector

The proposed collector geometry was modelled using MAFIA. First the static electric field was computed. Figure 9 above shows the equipotentials at 1000 V intervals in a typical three-stage collector. The geometry is designed so that there is a suppressing electric field, over the electrode surfaces, where most electrons are intercepted.

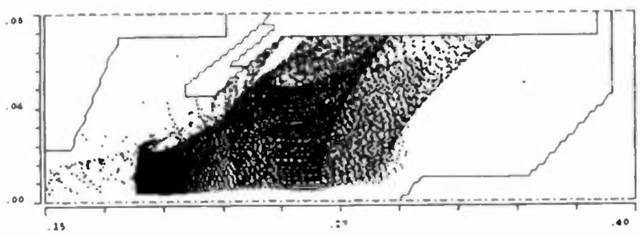


Figure 10. MAFIA Plot of Charged Particles in a Three-stage Collector

The electron beam data, from the tube model, was saved at a z-plane near the collector mouth, over one RF period. This data was used to initiate the electron pulses entering the collector. Figure 10 shows a plot of charged particles in a three-stage collector. The energy and charge intercepted, by each electrode, over a predetermined time (usually one RF period) are computed. From these values, making allowances for secondary electron emission, the efficiency of the collector can be calculated. Typical efficiency results are shown in Table 3. In this case the operating parameters are as follows:

Beam Voltage = 36 kV; Mean Beam Current = 2.4 A;
Stage 1 Voltage = 0 V; Stage 2 Voltage = 12 kV;
Stage 3 Voltage = 36 kV.

The secondary emission ratio is assumed to be 1. The secondaries emitted from the unsuppressed area of stage 2 are assumed to be accelerated into stage 1 and all the current intercepted by stage 3 (the cathode spike) is re-emitted as secondary current intercepted by stage 2. This set of results shows very little secondary interception and no effect on the resultant efficiency. The overall tube efficiency is increased from 48% to 65.5% in this case. For example, the energy input to the beam supply for a tube with 25 kW average digital output would be reduced from 52 kW to 38 kW.

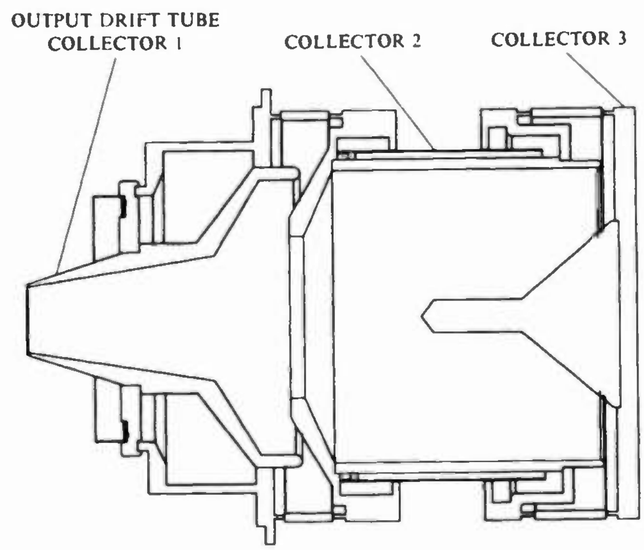


Figure 11. Schematic of Three-stage Collector

Figure 11 is a schematic of the three-stage collector design. The first stage is integral with the tube body and is separately water cooled. The second stage is also liquid cooled but with deionised water running through insulating hoses, using techniques established with EEV multistage collector klystrons. The third stage (cathode spike) intercepts very little energy; to reduce voltage hold-off problems this stage is air cooled. The complete three-stage depressed collector, including the outer can, is small enough not to need any modifications to the circuit assembly used on the single collector IOTD3130.

	Beam	Body	Stage 1	Stage 2	Stage 2 unsuppressed area	Stage 3	Total Collector
Voltage (kV)	36	0.00	0.00	12.00	12.0	36.00	
Power (kW)	43.9	0.21	2.37	19.24	0.139	0	21.96
Power + Secondary Power		0.21	3	19.26	0.139	0	22.61
Power (%)		0.9	13.4	85.0	0.6	0.0	
Current (A)		0.02	0.20	2.05	0.06	0.00	2.33
Electronic Efficiency (%)	48.0						
Overall Efficiency (%)	65.4						

Table 3. Computed Efficiency for a Three-Stage Collector

Frequency (MHz)	471.25	595.25	703.25
Heater voltage (V)	7.5	7.5	7.5
Beam voltage (kV)	35	35	35
Peak current (A)*	4	3.85	3.9
Grid voltage (V)	191	180	186
Focus current (A)	24.1	24	24.2
Picture content	Mid-grey		
Beam current (A)	2.3	2.0	2.1
Grid current (mA)	0	0	0
Peak vision output (kW)	77	77	77
Aural output (kW)	7.7	7.7	7.7
Peak vision drive (W)	427	509	525
Aural drive (W)	40	52	51
Drive sync. height (%)†	31	31.5	34
Sync efficiency (%)	55	57	56
FOM (%)	105	120	115
Vision gain (dB)	22.6	21.8	21.7
Aural gain (dB)	22.8	21.7	21.8
Bandwidth (MHz)	7	7.25	7
LF linearity (%)	5	9	11
ICPM (°)	1	0.5	0.5
Differential phase (°)	6	5	5
Differential gain (%)	2	4	5
Intermodulation (−dB)§	50	50	50

Table 4. Analogue Test Results of a New High Power IOT

Notes:

* indicates peak sync. vision-only current

† nominal sync. height is 27%

§ intermodulation amplitude with a 250 mV modulated ramp at 4.43 MHz.

EXPERIMENTAL RESULTS

The analogue and digital performance of a number of the newly-developed IOTs has been established. Typical measured results when operated in common amplification conditions are shown in Table 4 at three different channels. Excellent Figures of Merit were obtained at a peak sync. power of 77 kW and aural power of 7 kW. The linearity performance of the tube is good, as evidenced by the measured values of LF linearity, ICPM, differential gain, differential phase and the excellent Intermodulation Products obtained. The gain was also good, comfortably meeting a 20 dB minimum specification set during the design phase. The beam voltage was 35 kV, the same value as is needed by tubes of earlier design to deliver 66 kW peak sync. vision plus 6.6 kW aural power. Therefore the IOTD3130 and IOT9707 represent a significant step forward in IOT power capability with no increase in beam voltage requirement – a great benefit to the transmitter manufacturer.

Frequency (MHz)	473	575	689
Beam voltage (kV)	35	35	35
Uncorrected			
Beam Current (A)	2.55	2.1	2.5
Average Power (kW)	34	31	28
Efficiency (%)	38	42	32
Maximum Shoulder (dB)	32	32	31
EVM (%)	5.2	4.2	4
Signal/Noise Ratio (dB)	25	27	27
Corrected (using a digital pre-corrector)			
Average Power (kW)	29	26	23
Peak/Average Ratio (dB)	7.3	7.5	6.6
Peak Power (kW)	155	146	105
Maximum Shoulder (dB)	41	38	38
EVM (%)	2.9	2.0	2.1
Signal/Noise Ratio (dB)	30	34	33.4

Table 5. Digital Characteristics of a New High Power IOT

Table 5 shows typical uncorrected and corrected 8-VSB digital performance at three channels, on a tube containing the high perveance electron gun. In this case, it is particularly important to note that the tube after correction meets the specified FCC values for Shoulder Height, Error Vector Magnitude (EVM) and Signal/Noise Ratio. A peak output power of about 150 kW was obtained, again demonstrating the very high power capability of this new tube. For these experiments, a digital pre-corrector was used to correct for the IOT's non-linearities. The corrected efficiency of the tube was 32% at 473 MHz and 35% at 575 MHz, well in line with expectations based on experience gained, both in the laboratory and in the field, on previous IOTs covering a wide power range.

CONCLUSIONS

The use of state-of-the-art scientific computing codes has facilitated the design of a new IOT electron gun which provides significantly greater beam currents for a given applied beam voltage than established IOT electron guns. The application of this electron gun to an uprated version of the IOT body and RF input circuit used with a proven design of a lower power IOT has led to the construction and evaluation of plug-in IOT systems, both analogue and digital, which have much greater output powers than previously available **but which do not require increased beam voltages.**

The new tube with a single collector has the same high level of energy efficiency typical of existing generations of IOTs already in long-term service. However, the electrical energy input to the transmitter required to provide a given level of output power available can be reduced by at least 25% by the use of the same IOT electron gun, RF circuit and body but fitted with a new three-stage depressed collector.

Cost of ownership calculations show that the new tube with the three-stage depressed collector will save the broadcaster 20% of the cost of operating the transmitter, compared with the use of the same IOT with a single collector, even allowing for the cost of the relatively minor transmitter modifications needed to use the depressed collector version.

ACKNOWLEDGEMENTS

The authors wish to thank their many colleagues at EEV Limited and EEV Inc. in the USA for their valuable technical contributions to the work described.

The views expressed are those of the authors and not necessarily those of the General Electric Company of England.

INTELLECTUAL PROPERTY RIGHTS

This product is the subject of numerous granted patents in Great Britain and corresponding patents in other countries, including:

US5548245 US5239272 US5536992

REFERENCES

1. "An Expansion of the Output Power Range of Analogue Television IOT Systems and their Application to Digital ATV", G.T.Clayworth et al., NAB Broadcasting Engineering Conference, Las Vegas, April 1995.

2. "An Inductive Output Tube and Circuit Optimised For Digital ATV", G. T. Clayworth et al., NAB Broadcasting Engineering Conference, Las Vegas, April 1996.

3. "OPERA-2D Version 7.0 User Manual", Vector Fields Ltd., Oxford, England.

4. "MAFIA Release 4 User Manual", CST, Gesellschaft fur Computer-Simulationstechnik m.b.H., May 1997.

5. "A Wide-Band Inductive-Output Amplifier", A. V. Haeff et al., Proc. I.R.E., vol. 28, March 1940.

Architecture of a DSP Based Dual-Mode ATSC/NTSC Television Exciter and Transmitter

David L. Hershberger
Continental Electronics, Inc.
Grass Valley, California

I. Introduction

The FCC has mandated a period during which ATSC and NTSC signals will be simulcast. After the transition period, the NTSC signal will be discontinued. In some cases, at the end of the transition period, the ATSC signal will move to the NTSC channel. Transmission equipment that is suitable for both NTSC and ATSC broadcasts will provide the greatest degree of flexibility for the broadcaster.

When designing hardware to produce ATSC signals, a little extra thought, and a little extra hardware will allow use of the same platform to produce both ATSC and NTSC signals. Novel combinations of modern Field-Programmable Gate Arrays, (FPGAs), general purpose digital signal processors, and programmable digital filter parts make possible the development of a flexible signal generation system that can produce either analog or digital TV IF signals.

II. System Requirements

A digital television (DTV) transmitter includes many subsystems, including:

1. Digital line receiver (SMPTE 310M)
2. Data randomizer
3. Reed-Solomon encoder
4. Data interleaver
5. Trellis coder
6. Sync inserter
7. 8-VSB modulator
8. Up converter
9. RF power amplifier
10. Adaptive linear equalizer
11. Adaptive nonlinear equalizer
12. Channel filter
13. Power supply
14. Control system
15. Embedded software

Although all of the above functions are included in a DTV transmitter, the emphasis of this paper is on the vestigial sideband modulator.

Before we look at ways that might be used to produce TV signals in the digital domain, we will review one of the most important parts of a DTV exciter: the filter or set of filters that produce the vestigial sideband shape.

To ensure that the receiver has a maximum eye opening, minimum error vector magnitude (EVM), and makes most efficient use of the available bandwidth, the overall frequency response shape from the transmitter to the receiver has been defined as a flat amplitude channel with raised cosine shaped band edges. However, both the transmitter and the receiver must have bandpass filters. The transmitter needs a bandpass filter to limit and shape its emitted spectrum. The receiver needs a bandpass filter to select the desired received signal and reject others.

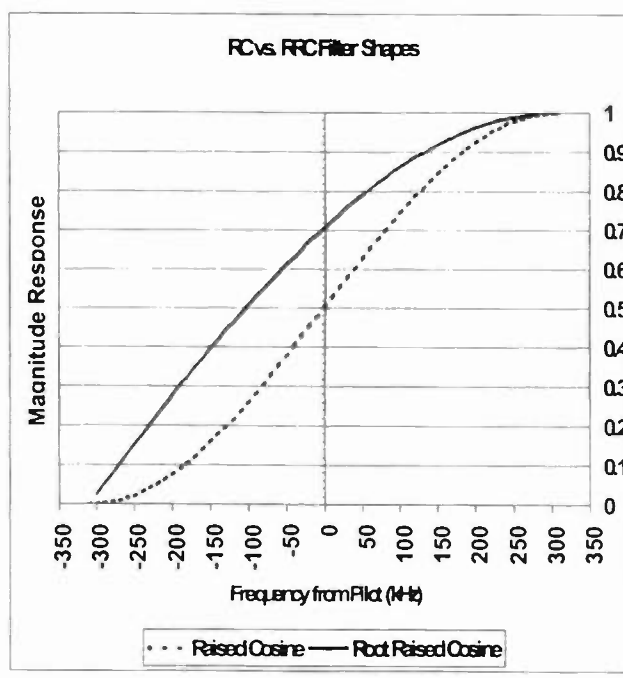
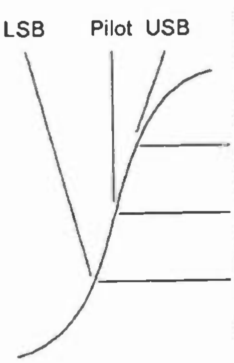


Figure 1 - Raised Cosine vs. Root Raised Cosine Filtering

To obtain the desired raised cosine response, the transmitter and receiver responses are multiplied. If both filters were raised cosine, then the overall shape would be raised cosine squared. But if both filters have a response that is the square root of the desired response, then the combined response will be the desired response. This is why the transmitter (and receiver) must have a [square] root raised cosine response in their transition bands. The difference between a raised cosine and a root raised cosine spectral shape is shown in Figure 1.

Most bandpass filters and lowpass filters have a passband (ripple amplitude) specification and a stopband (out of band attenuation) specification. The shape of the filter response in the transition band, between these two passband and stopband regions, is usually not specified, and the particular shape ends up being whatever allows the passband and stopband areas to be optimized. But in ATSC, the transition band is also specified (as root raised cosine), which makes the filter more complicated because its response must meet a certain shape *everywhere*, not just in the passband and stopband. As a result, the root raised cosine filter generally requires more coefficients than conventional filters that leave the transition band unspecified.



Raised Cosine	Root Raised Cosine	Gain Factor
-3 dB	-1.5 dB	1.5 dB
-6 dB	-3 dB	3 dB
-10.7 dB	-5.3 dB	5.3 dB

Figure 2 – Root Raised Cosine Equalization

Figure 2 shows the difference in the transition bands between a raised cosine (RC) response and a root raised cosine (RRC) response. The numbers in the first column show the response of a RC filter at various points. The second column shows the response of a RRC filter at those same points. The third column shows the gain that must be applied to a RC filter to turn it into a RRC filter. The pilot is 6 dB down on a RC filter but only 3 dB down on a RRC filter The point of this figure is that the gain difference between RC and RRC filtering is not symmetrical about the pilot.

There are several additional requirements and considerations for a digital TV exciter. These include:

1. Linear adaptive filtering. This allows the transmitter to flatten its frequency response automatically, to correct for amplifier mistuning, load variations, etc.
2. Nonlinear adaptive equalization. Nonlinear correction allows power amplifiers to be more efficient for a given distortion level, reduces sideband shoulders, and improves EVM performance.
3. Frequency or phase locking to external references. Where precision frequency control is needed between co-channel stations, GPS locking may be needed.
4. Making RF frequencies independent of input serial bitstream frequency errors. A SMPTE 310M input stream may be off frequency as much as 2.8 ppm. The RF or IF output frequency, however, should not include this error.
5. Channel offsets. To reduce co-channel and/or adjacent channel interference, slight adjustments in the pilot or carrier frequency are sometimes required.
6. Transmission of sync signals in the absence of an input signal. If the SMPTE 310M input signal is momentarily interrupted or lost, sync should continue to be transmitted, to allow receivers to recover faster when the input signal is restored.

A seventh consideration is the subject of this paper:

7. Dual mode operation (NTSC/ATSC)

III. Signal Generation Methods

There are several ways to generate vestigial sideband signals using DSP.

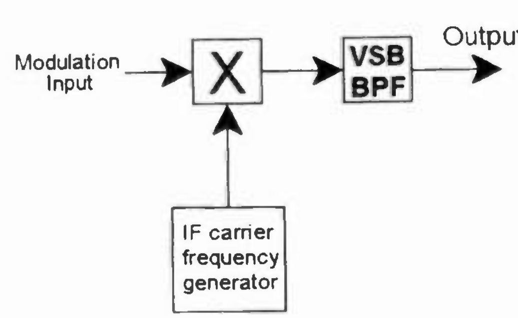


Figure 3 - Simple VSB Modulator

The brute force approach would in digital circuitry replicate what is typically done in the analog domain.

This is shown in Figure 3. First, the analog video or digital multilevel baseband signal modulates an IF carrier, producing a double sideband (DSB) signal. Then, the DSB signal is filtered to vestigial sideband (VSB). This simple approach is feasible in DSP but is quite inefficient and costly. Although straightforward, this approach would require a large number of filter taps (in the thousands) at a high sampling rate (40 MHz or more). Clearly, a little finesse is required to do the job efficiently in the digital domain.

Another method is to produce the in-phase and quadrature (I and Q) components of a VSB signal, using the Hilbert transform, as shown in Figure 4.

Hilbert transform modulation produces two baseband signals, each with a bandwidth of approximately 5.69 MHz. As the Nyquist frequency is approximately 5.38 MHz, an interpolation to a higher sampling rate is required to do the RRC shaping of the upper sideband edge without aliasing. Figure 4 shows an interpolation to twice the symbol rate (approximately 21.5 MHz). The blocks with the up-arrows are interpolators, which double the sampling rate.

Hilbert transform VSB modulation is feasible, but the necessary equalizers are rather complex and costly. Also, proper USB spectral shaping requires processing at a sampling rate higher than the symbol rate.

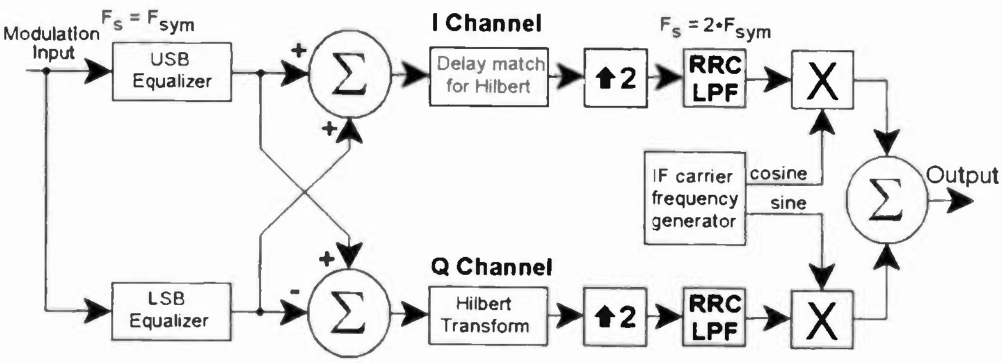


Figure 4 – Hilbert Transform VSB Modulator with RRC Equalization

This method is more practical. A Hilbert transform shifts the phase of frequencies within its passband by 90 degrees. However, since it is impossible to phase shift DC by 90 degrees, there is a practical low frequency limit to the Hilbert transform and it always has zero response at DC. In a quadrature modulator system, the low frequency rolloff and suppression of DC will result in the pilot being down 6 dB instead of the 3 dB required in a RRC spectral shape. So in the case of ATSC signals, special equalizers are required to convert the sideband shape in the vicinity of the pilot (carrier) to a root-raised-cosine response. Strict adherence to the ATSC spectral tail shapes requires the equalizers to have unequal responses for positive and negative frequencies, because the gain function which converts raised cosine to root raised cosine is not symmetrical about the ATSC pilot tone (as shown in Figure 2 above). Figure 4 shows one way to accomplish the equalization, using matrixed equalizers to apply different amounts of gain to the upper and lower sidebands in the vicinity of the pilot.

IV. Weaver Modulation

Another modulation method, known variously as Weaver modulation and as the so-called "third method" of single-sideband (SSB) generation, can be modified to produce vestigial sideband signals, with any arbitrary sideband shape including root raised cosine.

Weaver modulation was not often used in analog circuits because it requires two accurately matched filters and signal paths. However, in DSP, it is trivial to obtain accurate matching.

Traditional SSB Weaver modulation (shown in Figure 5) begins by multiplying the modulating signal by a pair of quadrature phased sinusoids. When generating SSB, the frequency of the sinusoids is usually the arithmetic mean frequency of the modulating bandwidth. The frequency of these two sinusoids is called the "folding frequency."

This multiplication produces a pair of orthogonal baseband signals. (Figure 5 also shows spectra that exist at various points in the system.) A lowpass filter after each modulator restricts the bandwidth of each output to half the bandwidth of the original modulating signal. At this point, the modulating signal has been "folded," such that the folding frequency is translated to DC while both the upper and lower band edges are translated to the

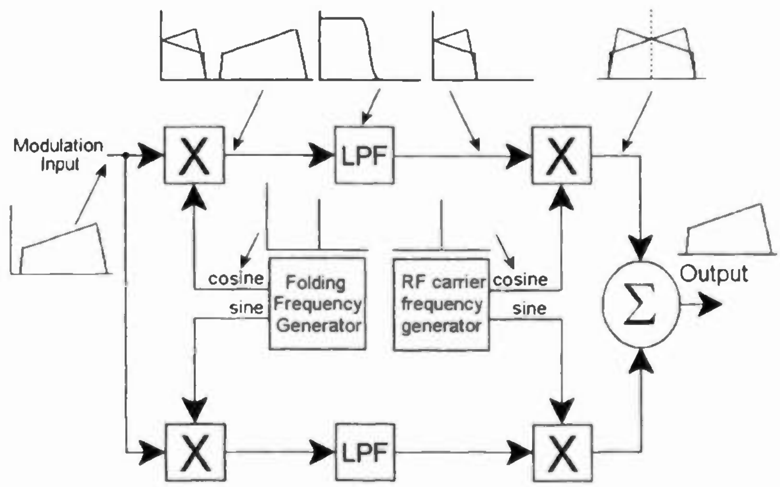


Figure 5 - Basic Weaver SSB Modulator

highest frequencies in the folded spectrum. Although two different frequencies map to a single frequency, the (orthogonal) phase relationship between the two baseband signals conveys the information necessary to recreate the modulating signal as SSB.

The two baseband signals are then applied to a pair of mixers driven with quadrature phased versions of an IF or RF signal. If the quadrature phasing is accurate, and if the two lowpass filters are matched, and if the gain, phase, and delay of the two signal paths are matched, then the sum of the two mixers is a SSB signal. The two baseband signals are I and Q signals, although they are offset in frequency from the I and Q signals found in the Hilbert method.

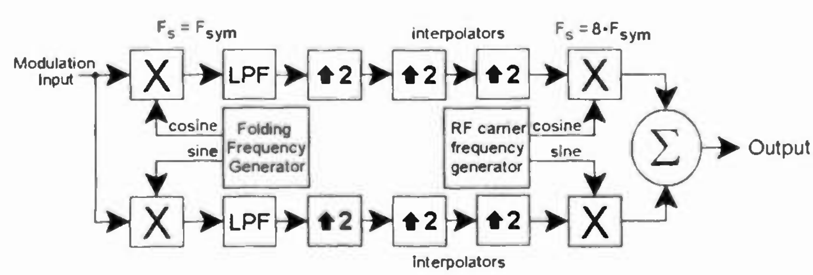


Figure 6 - 8-VSB Modulator Using the Weaver Method

Figure 6 shows how Weaver modulation may be applied to generate 8-VSB. To modify the method to create vestigial sideband instead of single sideband, the folding frequency is decreased enough so that a certain bandwidth of negative frequencies will produce folded frequencies within the bandwidth of the lowpass filters.

The bandwidth of the baseband signals produced by Weaver VSB modulation is only 3 MHz, only about half what it is in the Hilbert transform case. This means that the relatively low symbol rate (approximately 10.76 MHz) may be used to do the RRC filtering, without aliasing.

In the ATSC mode, the 10.76 MHz sampled data is multiplied by orthogonal sinusoids at one quarter of the symbol rate or approximately 2.69 MHz. (The multiplication is relatively trivial, because the folding frequencies may be represented by simple four sample 1,1,-1,-1 sequences.) This produces a pair of folded spectra, each having a bandwidth of just half the channel width, or 3 MHz. Root raised cosine lowpass filtering (as opposed to bandpass filtering) is applied to each folded spectrum. Each of the two lowpass filters has a flat response to approximately 2.38 MHz, and a root raised cosine response rolloff between 2.38 and 3 MHz.

In a Weaver implementation, the root raised cosine bandpass filter is actually a pair of lowpass filters. Two 256 tap filters, operating at the 10.76 MHz sampling rate, are approximately equivalent to a brute force bandpass filter of 2048 taps, operating on a DSB signal at 4 times the symbol rate. Brute force VSB generation requires approximately 88.2 gigataps per second, while using the Weaver method only requires 11 gigataps per second - just one eighth of the computational "horsepower."

Use of Weaver modulation in the generation of ATSC signals has the following advantages:

1. Complex root-raised cosine filtering may be done at the lowest sampling rate of the system.

2. Since the sampling rate where the bulk of the filtering is done is low, digital filters may be multiplexed to conserve hardware.

3. Operation of finite impulse response (FIR) filters at the lowest possible sampling rate generally means that the order of the filter may be lower for a given

performance level. So, the filter can be shorter, simpler, slower, yet at the same time better than other implementations.

4. Because the bandwidth of the baseband signals is only 3 MHz instead of 6 MHz, subsequent interpolation filters will be considerably simplified over other methods. (The bandwidth of the interpolation filters is less.)

V. Implementation

A digital television exciter is a multirate sampling system. This means that some functions are performed at low sampling rates while others are done at high sampling rates. Generally, basic filtering operations including adaptive linear equalization and vestigial sideband filtering are done at the lowest sampling rates. Nonlinear predistortion, where new frequencies are produced and the bandwidth is increased, is performed at higher sampling rates to avoid aliasing. Finally, digital intermediate frequency (IF) modulation is done at the final highest sampling rate.

For dual-mode operation, it is advantageous to use approximately the same highest sampling rate, in order to use the same analog reconstruction filter and analog up converter for both NTSC and ATSC. Fortunately, there is a convenient close convergence of NTSC and ATSC related clock frequencies at 86 MHz.

SMPTE 310 input data is transmitted at a symbol rate of approximately 10.76 MHz. SMPTE 244M or SMPTE 259M digital NTSC data is transmitted at a sampling rate of four times subcarrier which is approximately 14.318 MHz.

Eight times the ATSC symbol rate is approximately 86.098 MHz. This is close to 85.909 MHz, which is approximately 24 times subcarrier.

Use of a 86 MHz output sampling rate for both NTSC and ATSC allows use of a relatively high digital IF frequency centered at 21.5 MHz, approximately one quarter of the sampling rate, or one half of the Nyquist rate. Using a 21.5 MHz IF comfortably allows nonlinear predistortion sidebands to be generated which are at least three channels wide, or 18 MHz of bandwidth.

In the ATSC mode, baseband linear equalization and Weaver modulation are performed at the symbol rate. From there the signal is interpolated to four times the symbol rate. At this higher sampling rate, nonlinear predistortion is applied to correct for power amplifier nonlinearity. At this point, the 3 MHz wide signal may increase to 9 MHz of bandwidth or more (corresponding to an IF bandwidth of 18 MHz or more).

Finally, the sampling rate of the orthogonal signals is increased to 86 MHz where they modulate orthogonal IF carriers, to produce the IF signal centered at 21.5 MHz. Figure 7 shows the various sampling rates and the nonlinear equalization applied at four times the symbol rate.

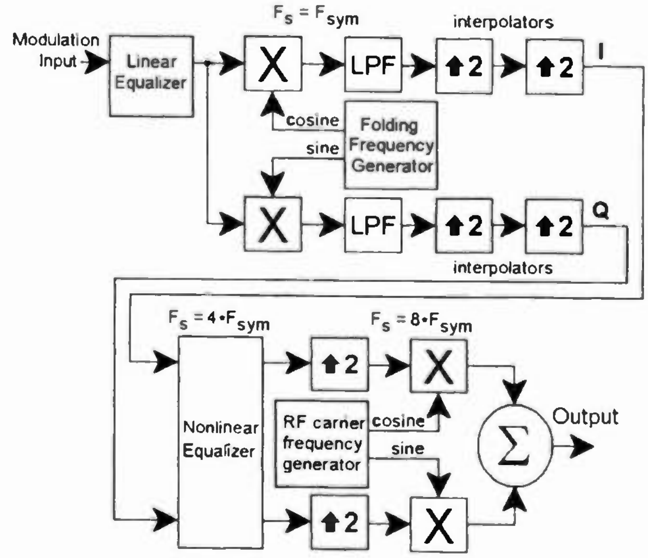


Figure 7 - Nonlinear Equalization at 4*Fsym

One quarter of the output clock is 21.52447552... MHz, which is close to 21.5 MHz. But it is desirable to produce an IF signal centered on exactly 21.5 MHz, because this makes the analog up converter easier to design. Furthermore, it is desirable to make the IF frequency independent of the input clock frequency. Also, it is necessary for some stations to produce channel offset frequencies to minimize interference. To achieve these objectives, the Weaver VSB modulator includes a frequency shifter. The frequency shifter is used to (1) center the output frequency on exactly 21.5 MHz, (2) compensate for input clock frequency errors, and (3) produce channel offset frequencies where required.

The frequency offset generator, which operates on the Weaver baseband signals, is shown together with a Weaver modulator in Figure 8.

To determine the offset frequency, the frequency of the SMPTE 310 input clock is measured against an accurate internal 10 MHz reference. The 10 MHz reference may be locked to GPS or to a rubidium standard. (The SMPTE 310 input frequency may vary as much as +/-

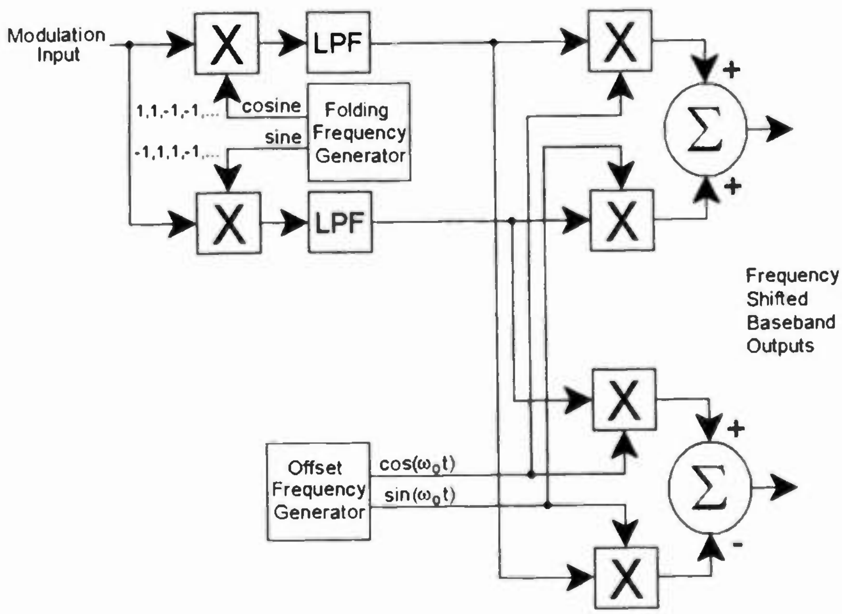


Figure 8 - Frequency Shifter

2.8 ppm, which would cause an IF frequency error of up to 60 Hz if uncorrected.) Using a 32 bit DDS, frequency precision is approximately 0.0025 Hz.

The basic Weaver VSB modulator (shown without linear and nonlinear equalizers) is shown in Figure 9.

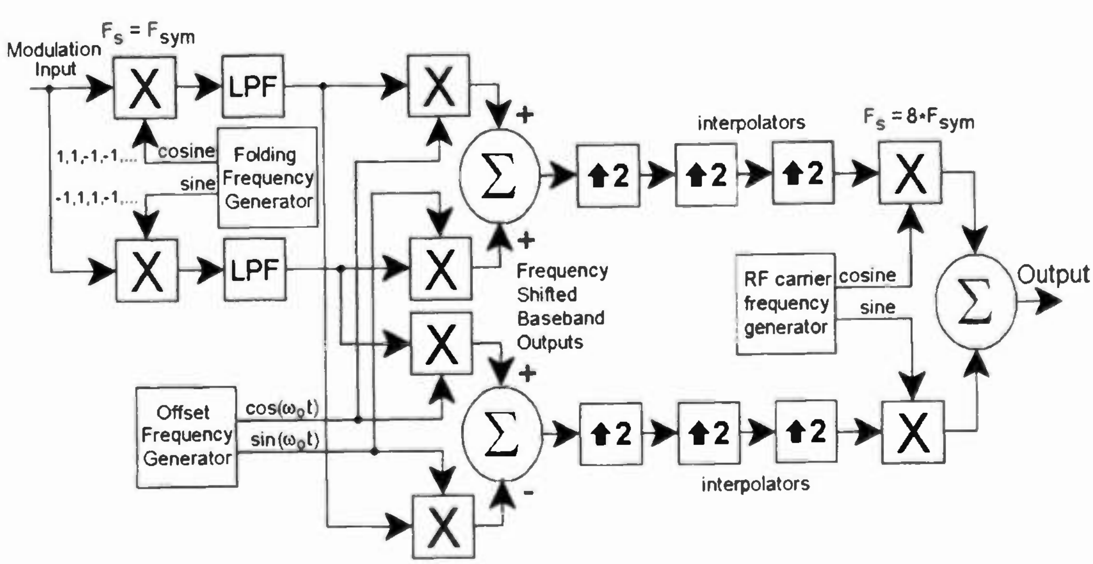


Figure 9 - Basic Weaver 8-VSB Modulator with Frequency Shifter

VI. NTSC Mode

Digital NTSC signals are described by SMPTE 244M (parallel) and SMPTE 259M (serial). A digital NTSC exciter should accept the SMPTE 259M signal as its input.

Although the generation of NTSC signals requires less severe filtering than the ATSC case, there are several considerations when using this same Weaver modulator architecture to produce NTSC signals. These issues are:

1. Unlike ATSC, the NTSC signal does not have symmetrical sideband shapes.
2. NTSC includes an aural carrier signal.
3. NTSC includes a receiver delay equalizer.

If we want to produce a frequency inverted (lower sideband) NTSC IF signal centered at 21.5 MHz, Figure 10 illustrates what the various frequencies will be.

If the transmitter uses common aural and visual amplification, the aural carrier will appear in the visual exciter's output. Transmitters with a separate aural power amplifier will have a separate aural exciter.

Using the frequency scheme shown in Figure 10, the highest Weaver baseband signals will be produced by the upper IF sideband. Therefore, it will be the vestigial upper IF sideband that is shaped by the Weaver lowpass filters. The Weaver lowpass filters will be flat to 2.5227... MHz and will cut off by 3.0227... MHz. The sharp cutoff of the lower IF sideband is independently shaped at baseband by a simple video lowpass filter.

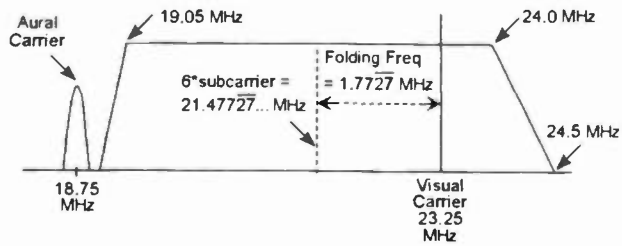


Figure 10 - NTSC IF Spectrum Showing Weaver Folding Frequency

The aural carrier, if introduced at baseband, would appear in the folded baseband signals at 2.727... MHz, which is within the transition band of the NTSC Weaver lowpass filter. The filter's slope would introduce undesirable incidental envelope modulation onto the aural carrier. Therefore, when this frequency scheme is used, the NTSC aural carrier, if amplified in common with the visual carrier, must be introduced subsequent to the Weaver modulation. Fortunately, this is easy to do because common direct digital synthesizer (DDS) chips generally produce quadrature sinusoidal (sine and cosine) outputs.

BTSC signals are still produced in the analog domain today. Therefore, the aural modulation input will accept a BTSC composite baseband signal. The BTSC signal, which has a bandwidth of some 120 kHz, is digitized at a submultiple of subcarrier. Then its sampling rate is increased to four times subcarrier or 14.318... MHz. At this point the BTSC signal frequency modulates a DDS.

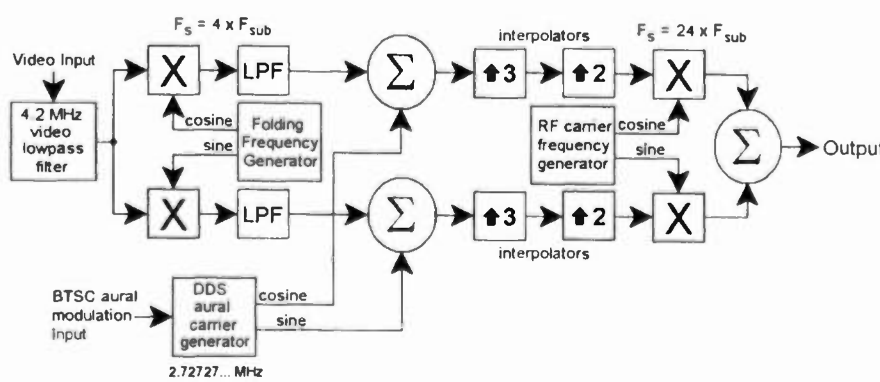


Figure 11 - Weaver NTSC VSB Modulator

The DDS produces frequency modulated quadrature sinewaves at 2.727... MHz, which are simply added to the Weaver baseband signals after lowpass filtering. This is shown in Figure 11.

Unlike the ATSC implementation, where the Weaver folding multipliers are trivial (multiplying only by +1 and −1), the NTSC folding frequency (1.7727... MHz) is not a round number. Therefore, that folding frequency is produced by another DDS and the multipliers that perform the modulation are of a nontrivial, general purpose type. However, they only need to operate at the four times subcarrier video sampling frequency (14.318... MHz). With a 32 bit DDS, frequency precision is approximately 0.003 Hz.

As is done in the ATSC case, channel offsets and compensation for frequency errors in the incoming SMPTE 259M serial bitstream are accomplished by adjusting the frequency of the folding frequency generator DDS.

VII. The Transmitter System

A complete transmitter system includes an exciter, a visual power amplifier system, and an adaptive linear and nonlinear equalization subsystem. Often in the case of NTSC, a separate aural power amplifier system and diplexer is included. ATSC transmitters will also include channel filters.

Power amplifier requirements differ between NTSC and ATSC service. In ATSC service, the signal is nearly stationary in the statistical sense. The transmitted signal resembles a band limited noise signal with a peak to average power ratio of approximately four to one. Average power remains nearly constant. Dissipation also remains nearly constant. The RF envelope, which determines the power supply currents that must be decoupled, is a wideband noise-like function.

For NTSC service, the average power varies, the peak power remains constant, and therefore the peak to average ratio varies. The power supply must be capable of delivering constant voltage in the presence of large low frequency video currents, related to vertical sync pulses and other components. Dissipation varies as a function of average picture level. So in some ways, NTSC service has more severe requirements for an RF power amplifier. These include variable dissipation and power supply dynamic performance.

ATSC service has the more severe requirement when it comes to linearity. In-band nonlinearities affect EVM performance, while out of band nonlinearities affect the transmitter's spectral mask.

Adaptive equalization, both linear and nonlinear, can be performed using a non real time demodulation process to sample the transmitter's output. First, the transmitter's output is down-converted to the 21.5 MHz IF frequency. Next, a high speed A/D converter takes "snapshots" of the transmitter output. A floating point digital signal processor demodulates the snapshot in non real time. Finally, the results from the numerical demodulation are used to update linear and nonlinear equalizers, based on designs used in DAB transmitters.

For transmitters that use a separate aural power amplifier and notch diplexer in the NTSC mode, ATSC operation will require either bypassing of the diplexer or detuning of its notch resonators such that the diplexer bandwidth is extended to the entire 6 MHz channel width.

VIII. Conclusion

Careful system architecture, use of reprogrammable logic devices, and the proper choice of system clock frequencies make it possible to produce a system that can transmit either NTSC or ATSC television signals. A reconfigurable transmission system offers more options and flexibility to broadcasters during the transition period to digital TV broadcasting.

Performance of a Multi-program MPEG Encoding System in Constant Bit rate and Statistical Multiplexing Modes

Paul Haskell, Elliot Linzer and Humphrey Liu
DiviCom
Milpitas, CA

Abstract

We compare the perceptual video quality of a multi-program MPEG-2 encoding system in constant bit rate operation and using statistical multiplexing. The video quality is measured using a pycsho-visual distortion score for each frame. The distortion scores are aggregated over time to compute an over all quality score.

1. Introduction

An MPEG-2 transport stream is a time-division-multiplexed combination of one or more "programs". Each program may comprise one video stream, one or more audio streams, and private data streams. Multiple programs are multiplexed into a single transport stream by using packet headers to specify which data belong to which program, so the bit rates of individual programs may vary over time even if the transport stream itself is constant bit rate.

As is well known, "statistical multiplexing" can be used to dynamically allocate bit rate among the programs in a single MPEG transport stream. More bits are allocated to programs that are more "complex", i.e., those with more spatial or temporal variation, so that "easy" programs do not "waste" bits and "hard" programs do not suffer from severe artifacts.

Our goal is to answer the question "how much gain in channel capacity can be achieved by using statistical multiplexing in an MPEG video compression system?" More specifically, what average rate A is needed when statistically multiplexing video programs to get approximately the same video quality as using a constant bit rate B for each program? The answer depends on many factors:

1. How is "quality" measured?
2. How are quality scores aggregated?
3. What kinds of video sources are used for testing?
4. What rate control "algorithms" are used for constant bit rate and variable bit rate control?

The best quality measures are based on human viewing, but to conduct extensive experiments based on human viewing is impractical. Further, results often are unrepeatable. In this paper, we use a measure of compression-induced distortion that is based on the quantisation step-size used to compress each video frame.

The second question is particularly tricky, and has a significant effect on how much gain statistical multiplexing achieves. If quality is measured as worst case quality over long periods of time then statistical multiplexing will show an implausibly large improvement over constant bit rate, but if quality is averaged in a way which rates lots of very minor distortions as equal to a few large distortions then statistical multiplexing would seem to have little advantage over constant bit rate. In practice, humans seem to judge video sequences based on their worst periods of distortion, even if the high-distortion periods are brief.

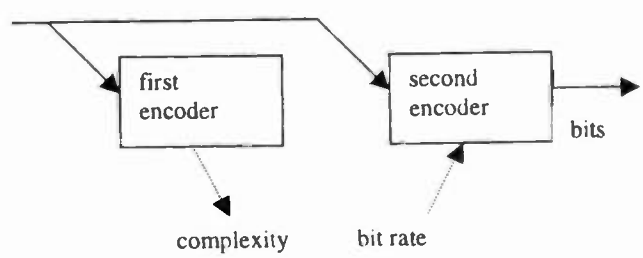
To conduct a fair comparison of VBR vs. CBR video compression, we must use several different types of video material, such as sports and movies. Further, each video clip must contain a realistic amount of variation—constant bitrate encoding performs excellently when coding video with no significant changes in content. Very little video material lasts for more than 10 seconds without a scene cut!

The fourth question should not be ignored. Indeed, first generation statistical multiplexors simply did not perform as well as they might have, and, for example, limited bit rate variations to fairly modest ranges. But the quality of statistical multiplexors has certainly improved over time. In these experiments, we use the "latest and greatest" rate allocation and control algorithms that we are developing at DiviCom.

The bottom line is that we do not claim that our answers will be very precise, but we do believe that we can give a realistic estimate of the expected gain.

2. Rate allocation and distortion measures

The statistical multiplexing algorithm that we use in the following experiments is based on a two-pass encoding algorithm. The video is encoded twice, and the bit stream transmitted is the result of the second encoding. There is a delay of about ¾ of a second between the first and second encoding. The rate allocation algorithm allocates bit rates to each program based on the results of the first encoding (for pictures that have not yet been encoded a second time), with corrections from the second encoding. The algorithm seeks to equalize a proprietary measure of visual quantisation distortion across programs.



Because we use a two-pass encoding approach for rate allocation, we are able to use very low transmission rates for easy material and still to increase bit rates rapidly to maintain quality through rapid transitions to much harder material.

For the purposes of these video quality experiments, we modified the DiviCom encoder so that during the second encoding, each encoder computes a distortion score for each picture, measuring the visual quantisation distortion for that picture. It is this score that we will use measure the effectiveness of statistical multiplexing versus constant bit rate encoding.

After computing a score for each picture, we aggregated the scores by computing percentiles. The *x* percentile score is the score that for x percent of the video had a higher score (more distortion) and 100-*x* percent had a lower score (less distortion). Aggregating scores by using percentiles allows us perform a meaningful comparison between CBR and VBR encoding. The whole point of using VBR encoding is to allocate a higher rate to a program that is more complex and to take away that rate from programs that are less complex. Thus, although the complexity of an individual program may vary quite a bit over time, we expect the distortion introduced by compression to vary within a narrow range when VBR encoding is used.

Using a very high percentile score, e.g., the 99.5th percentile, as an aggregate of video compression distortion, will isolate a small fraction (1/200) of the "worst case" frames. But even though the fraction is small, the score is quite meaningful; a severe drop in quality for one second every three minutes twenty seconds (1/200th of the time) will certainly be objectionable. As lower percentiles are used, the measure looks at more video put penalizes occasional peaks in distortion less.

A very low percentile score, e.g., the 10th percentile, is not a useful measure. Occasional drops in the distortion caused by video compression will not improve the overall perceived video quality.

3. Experimental results

We ran our experiments with 6 sources, as listed in the table below. For CBR, we choose bit rates appropriate for each genre of material, not for the specific content; thus, all movies got the same rate, but the video feeds received higher rates than the movies. Using equal rates for all sources would unfairly penalize the CBR encoding approach, as broadcasters can and do vary bit rate by program content even when using CBR. Manually varying rates based on scene content (e.g., more bits for an action scene) is unrealistic in a real broadcast environment and would make the CBR results look unfairly good.

TABLE 1. LIST OF SEQUENCES AND BIT RATES USED FOR CBR EXPERIMENTS

Program	*Description*	*CBR rate*
Sports	Basketball game and interviews with players	5 Mbits/sec
News	Talking heads, computer generated text/graphics, on-location footage	3
Mov 1	Terminator II	2.5
Mov 2	The Man in the Iron Mask	2.5
Mov 3	Jumanji	2.5
Mov 4	Pinocchio	2.5
All		18

As mentioned in the previous section, we compute a quantisation distortion score for each encoded picture and then aggregated the scores by computing percentiles. Because it is not clear which percentile score is the best method for aggregating scores, we computed a range of percentiles, specifically the 99.5th, 99th, 98th, 90th and 80th percentile distortion scores. The following table gives the aggregated scores using our proprietary quantisation distortion measure:

TABLE 2. PERCENTILES FOR PERCEPTUAL QUALITY DISTORTION FOR CBR ENCODING WITH A TOTAL RATE OF 18MBITS/SEC

Percentile *Program*	*99.5th*	*99th*	*98th*	*90th*	*80th*
Sports	19.6	19.3	17.0	12.5	10.2
News	8.5	8.4	8.1	7.4	7.1
Mov 1	11.0	10.0	9.4	6.0	4.6
Mov 2	15.7	14.0	12.4	10.0	7.1
Mov 3	18.4	16.0	14.0	10.4	8.0
Mov 4	18.9	15.8	14.0	10.0	8.0
All	18.0	15.5	13.9	10.0	7.5

We repeated the above experiments using VBR (statmux) encoding at a total "pool" rate of 18 Mbits/second, which is equal to the sum of the rates used for the CBR encoding. The results are shown in table 3. Particularly if one concentrates on a high percentile, i.e., on only the worst small fraction of frames, using VBR over CBR sees a large improvement.

TABLE 3. PERCENTILES FOR PERCEPTUAL QUALITY DISTORTION FOR VBR ENCODING WITH A "POOL" RATE OF 18 MBITS/SEC

Percentile *Program*	*99.5th*	*99th*	*98th*	*90th*	*80th*
Sports	10.0	10.0	9.5	8.4	7.9
News	9.0	9.0	8.0	8.0	7.0
Mov 1	10.1	10.0	9.9	8.6	8.0
Mov 2	10.3	10.0	9.9	8.5	8.0
Mov 3	10.1	10.0	10.0	8,7	8.0
Mov 4	10.0	10.0	9.9	8.6	8.0
All	10.0	10.0	9.7	8.4	8.0

Even if a particular percentile of scores were a perfect measure of overall quality, Tables 2 and 3 still do not provide enough information to determine bit rate savings for VBR versus CBR. To make that comparison we need to run the VBR or CBR experiments (or both) at different rates. In Tables 4 and 5 we show the experimental results for encoding the same video using total pool rates of 16.2, 14,4, and 12.6 Mbits/sec; i.e., at 10%, 20%, and 30% less than the CBR rates.

As can be seen from these tables, if the 90th percentile is used, VBR encoding needs about 10% fewer bits than CBR encoding (VBR at 16.2 Mbits/sec gets about the same score as CBR at 18 Mbits/sec.) When a higher percentile is used; i.e., when we place a greater emphasis on "worst case", the gain for VBR grows. Indeed, using the 98th percentile VBR at 14.4 Mbits/second scores about the same as CBR at 18 Mbits/sec -- VBR needs 20% fewer bits than CBR. If we use the 99.5th percentile, the gain is greater than 30%; the distortion score for VBR at 12.6 Mbits/sec is lower than the distortion score for CBR at 18 Mbits/second

TABLE 4. PERCENTILES FOR PERCEPTUAL QUALITY DISTORTION FOR VBR ENCODING WITH A "POOL" RATE OF 16.2 MBITS/SEC

Percentile *Program*	*99.5th*	*99th*	*98th*	*90th*	*80th*
Sports	12.0	11.7	11.0	10.0	10.0
News	11.9	11.1	11.0	10.3	10.0
Mov 1	12.0	12.0	11.4	10.4	10.0
Mov 2	12.0	12.0	11.0	10.0	10.0
Mov 3	12.0	12.0	11.3	10.1	10.0
Mov 4	12.0	11.5	11.0	10.2	10.0
All	12.0	11.8	11.0	10.1	10.0

TABLE 5. PERCENTILES FOR PERCEPTUAL QUALITY DISTORTION FOR VBR ENCODING WITH A "POOL" RATE OF 14.4 MBITS/SEC

Percentile *Program*	*99.5th*	*99th*	*98th*	*90th*	*80th*
Sports	13.9	13.4	13.0	11.8	11.0
News	13.0	13.0	12.2	11.9	11.0
Mov 1	14.1	14.0	13.8	12.0	11.0
Mov 2	14.3	14.0	14.0	12.0	10.9
Mov 3	14.5	14.0	14.0	12.0	11.0
Mov 4	14.4	14.2	14.0	12.0	11.0
All	14.2	14.0	13.7	12.0	11.0

TABLE 6. PERCENTILES FOR PERCEPTUAL QUALITY DISTORTION FOR VBR ENCODING WITH A "POOL" RATE OF 12.6 MBITS/SEC

Percentile / *Program*	*99.5th*	*99th*	*98th*	*90th*	*80th*
Sports	16.0	15.6	15.0	13.0	12.5
News	15.0	14.0	14.0	12.9	12.0
Mov 1	16.0	16.0	15.3	14.0	13.0
Mov 2	16.1	15.9	15.1	14.0	13.9
Mov 3	16.0	16.0	15.0	13.8	12.7
Mov 4	16.6	16.0	15.5	13.6	12.8
All	16.0	15.9	15.0	13.4	12.5

4. Conclusions

Statistical multiplexing can dramatically reduce the fraction of video that suffers from severe compression artifacts. This can be seen from the numerical results obtained in the last section: the 99.5th percentile of the distortion measure for statistical multiplexing is about the same as the 90th percentile for CBR at the same total rate. Moreover, reduction in severe coding artifacts is easily verified visually; we must run at a much a higher CBR rate than average VBR rate to get the same frequency of video that shows severe artifacts (such as bad blockiness). In fact, a significant reduction in blockiness is noticed at a given overall bit rate when even two programs are statistically multiplexed together.

Statistical multiplexing does a far better job than CBR of maintaining equal distortions across all video encoders in a pool. This greatly simplifies the "tuning" task that a broadcaster performs to get maximum usage out of his available transmission bandwidth. Similarly, for a given video channel, statistical multiplexing significantly reduces time variations in distortion. Possibly the biggest benefit of statmux is that it can allocate very high bitrates for very short periods of time during scene cuts, explosions, dissolves, and other video effects that almost always cause annoying artifacts with CBR encoders.

But how much bit savings results from using statmux? Traditionally, one would seek an answer using source modeling, information theory, and psycho-visual modeling. But these disciplines do not provide a complete answer, because none can answer the question "how *important* are occasional drops in quality when compared to more frequent but less severe drops?" If they are very important, then statmux saves a lot, and if they are not so important, the savings are not great. The question of importance ultimately boils down to issues of psychology and taste. A precise answer will be elusive.

Moreover, a judgement of video quality made by an observer, even a non-expert observer, who is purposefully evaluating video quality, may differ from the observations of someone who is simply watching a program. The casual observer would probably not be bothered by an artifact that occurs commonly but would be bothered by an unusual event. An example is flicker in 50Hz video that is quite annoying to someone used to watching 60Hz video but that quickly becomes less noticeable as time goes on. Similarly, someone used to watching VBR video might object strongly to CBR at the same bit rates because the occasional blockiness would stand out. On the other hand, someone used to

CBR video who watched VBR at a lower rate (but one that still maintained the same or a lower fraction of very noticeable artifacts) might object to the constant presence of more minor artifacts.

Having said that the bandwidth gain from using statmux is not knowable, we can nonetheless say that the experimental results of the last section are in line with our (unscientific) observations of VBR versus CBR. Namely, that when VBR encoding is used the frequency of significant distortions is greatly reduced, and the frequency of subtler but nonetheless noticeable distortions is reduced also. It is our admittedly subjective opinion that when VBR encoding is used transmission rates can be reduced by 10-20% compared to CBR while maintaining the same *overall* subjective video quality.

But we cannot emphasize enough that while the overall video quality might be the same, the nature of any distortions will be different. If one views CBR video at a given rate and VBR video at an average rate 10-20% lower, then the CBR video will look better on low complexity scenes whereas the VBR video will look better in more complex scenes and through transitions. Deciding which is "better" is a matter of taste.

Modulation Schemes for Transmission of DTV Signals Over Satellites

Girish Chandran, Ph.D.
Tiernan Communications, Inc.
San Diego, California

Abstract

Traditionally QPSK has been the modulation scheme of choice in transmitting DTV signals over satellite. However there are modulation schemes that allow higher throughputs and are more spectrally efficient than QPSK. This allows high bandwidth signals like contribution quality high definition video to be squeezed into existing satellite transponders. Spectral efficiency also becomes an important consideration with multichannel SDTV. However, there is a price to be paid for these spectral efficient schemes in power and the ability to stack multiple carriers. In recent experiments 8PSK and 16QAM were used to transmit multiplexed high definition and standard definition signals. In light of the results from these experiments, this paper examines the trade-off parameters between QPSK, 8PSK and 16QAM.

1. Introduction

The infrastructure of collecting and distributing audio, video and data is now migrating to digital from analog. Many broadcasters have already started broadcasting in digital high definition and standard definition formats. This migration to digital production and distribution allows content and bandwidth to be managed more efficiently. Managing bandwidth is a way of maximizing the use of available bandwidth. As an illustrative example, therefore, it is useful to examine the bandwidths required for digital standard definition and high definition television.

MPEG-2 Main Profile @ Main Level (MP@MPL) video was optimized for 4Mbps. Since resolution of high definition (MP@HL) MPEG-2 video used in ATSC A/53 is about four to six times that of standard definition, a bandwidth of about 18Mbps is used for high definition video. For high quality contribution and distribution applications, DS-3 rates (44.736 Mbps) are commonly used. For multichannel standard definition applications (widely used in direct-to-home broadcast over satellite) 27Mbps is a popular data rate. Satellite transponders have fixed, finite bandwidths. Satellite time is also expensive. There are two ways of accommodating bandwidth guzzling applications like HDTV in multichannel applications. To broadcast a fewer number of channels with attendant loss of revenue or, preferably, to find a way to increase the throughput on expensive satellite resources by using bandwidth efficient modulation schemes.

8PSK (Phase Shift Keying) and 16QAM (Quadrature Amplitude Modulation) are two bandwidth efficient modulation schemes that have recently been used with considerable success for transmitting digital television signals via satellite. These provide higher throughput than the traditional QPSK scheme that has been used for the past several years. A theoretical examination of three modulation schemes (QPSK, 8PSK and 16QAM) under different metrics is presented in Section 2. Section 3 examines some of the deviations that arise in practical systems and Section 4 presents results of experiments performed with transmission of digital television signals with 8PSK and 16QAM modulation via satellite.

2. Comparison of Modulation Schemes

2.1. 8PSK and 16QAM or 8QAM and 16PSK?

Digital content that is presented for modulation is a stream of bits that, without loss of generality, can be scaled by $\sqrt{\frac{2E}{T}}$ such that the energy in a bit interval T is E. This input is used to modulate the amplitude or phase, or amplitude and phase of a signal.

The modulated signal can be written as

$$s(t) = I_m \cos[2\pi f_c t] - Q_m \sin[2\pi f_c t]; \qquad m = 1,\ldots,M$$

where M is the number of signals in the constellation.[1] This modulated signal is then upconverted, amplified and broadcast over satellite.

Noise on the satellite link contaminates the signal constellation before it gets to the receiver. When that happens, the signal dots in the constellation (refer to Figure 1) become fuzzy, with the amount and shape of fuzziness depending on the amount and shape of the noise added. For example, when the noise added is "white" Gaussian, the signal dots become fuzzy balls. The receiver needs to make decisions as to which signal dot is being transmitted based on the fuzzy dots that it demodulated. This clearly depends on how far apart these fuzzy balls are, which depends on signal energy (spacing between balls) and noise energy (the size of individual balls). For a given signal energy, the performance of the modulation scheme is very closely tied to the minimum Euclidean distance between the signal points in the constellation.

It can be shown that for the PSK signals with M points in the constellation, the minimum Euclidean distance between the signals is

$$d_{min} = \sqrt{2E}\sin(\frac{\pi}{M}) \approx \sqrt{2E}\frac{\pi}{M}; \quad \text{for large } M$$

where M is 4 and 8 for QPSK and 8PSK respectively. As M increases the constellation becomes closer together if the power remains the same. To keep the distance between the signals constant as M increases, the radius of the circle must increase, implying that the power must increase as $(M/\pi)^2$. It can be shown that for a QAM constellation with M points, to maintain the same distance between points as M increases, the power needs to increase as 2(M-1)/3 [1]. Thus there is an advantage of using QAM to PSK as M increases as shown below in *Table 2-1*.

M	Advantage (dB)
8	1.43
16	4.14
32	9.95

Table 2-1: Advantage of M-ary QAM over M-ary PSK. For M = 4, the QAM and PSK constellations are identical.

[1] $I_m = \sqrt{\frac{2E}{T}}\cos[\frac{2p}{M}(m-1)+\frac{\pi}{4}]$ and

$Q_m = \sqrt{\frac{2E}{T}}\sin[\frac{2p}{M}(m-1)+\frac{\pi}{4}]$

for PSK (M = 4 for QPSK and M = 8 for 8PSK) signals.

For M = 8, there are advantages offered by PSK systems that outweigh the 1.43dB advantage of 8-QAM. This is discussed in Section 3. When M = 16, the advantage offered by QAM is significant. Hence 8PSK and 16QAM are used instead of 8-QAM and 16-PSK.

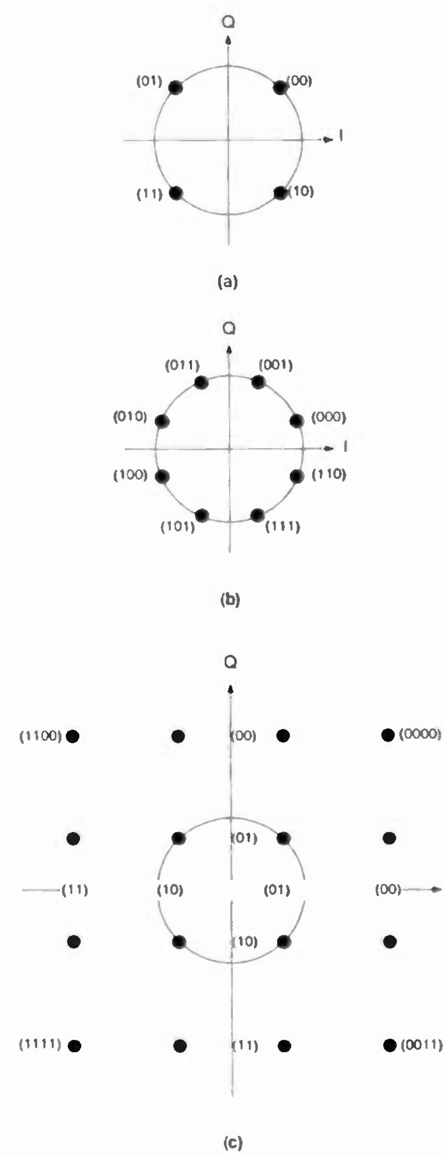


Figure 1: Gray-encoded signal constellations per pr EN 300 210, ATSC T3 Doc. T3-497: (a) QPSK (b) 8PSK (c) 16QAM. The black dots represent the modulated signal, where the modulation is governed by the binary digital content that is presented. This binary input is indicated alongside the black dots. In (a) each signal represent two bits, in (b) three bits and in (c) four bits.

2.2. Bit Error Rates

While all bits are created equal, some are more equal than others.

With digital television with video, audio and data, it is important that the digital *video* bit error rate be kept small, since the human eye is very sensitive to visual errors. This drives the bit error rate requirement for the satellite link used for DTV broadcasts. At a digital video rate of about 27Mbps (a typical multi-channel rate) operating on a link with a bit error rate (BER) of 10^{-11}, we get a probability of about 1 bit error per hour. At the ATSC rate of 19.39 Mbps a BER of 10^{-11} corresponds to 2 bit errors in 3 hours.

The following theoretical evaluation of three modulation schemes (QPSK, 8PSK and 16QAM) assumes that the primary additive disturbance in a satellite channel is "white" Gaussian noise. It is "white" in the sense that the disturbance in one bit time interval is independent to the next interval.

It can be shown that the probability of making an incorrect decision about a signal point in the constellation can be closely approximated as

$$P_s \approx erfc(\sqrt{\frac{E_b}{N_o}(\log_2 M)}\sin(\frac{\pi}{M}))$$

where Eb/No is the signal-to-noise ratio per bit for M-ary PSK signals and *erfc* is the complementary error function defined as

$$erfc(x) = \frac{2}{\sqrt{\pi}}\int_x^{\infty} e^{-t^2}\,dt$$

For M-ary QAM the worst-case probability of making an incorrect decision (upper bound) is to a tight approximation

$$P_s \le 2erfc(\sqrt{\frac{E_b}{N_o}(\log_2 M)(\frac{3}{2(M-1)})})$$

These two symbol error probabilities are plotted below in Figure 2.

It is possible to estimate bit error rates from these symbol error probabilities. These error rates are several orders of magnitude larger than the 10^{-11} BER operating point discussed above. There are several ways to achieve this BER. Increase transmitter power, increase antenna

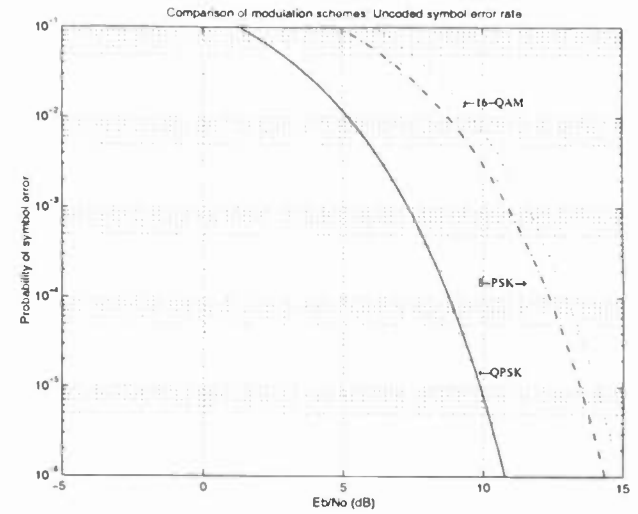


Figure 2: Uncoded symbol error rates of QPSK, 8PSK and 16QAM modulation schemes using equations derived above.

sizes/gains, or use the gain afforded by adding error correction coding.

Convolutional codes have long been used in "white" Gaussian noise channels to provide error correction. Sequential and Viterbi decoding have been developed to decode convolutional codes. The latter method has gained tremendous popularity because of its simplicity and optimality [2, 3]. Figure 3 compares the three modulation schemes of interest after convolutional coding and Viterbi decoding for a convolutional code of constraint length 7.

This error correction comes at a price, namely bandwidth. It can be seen that for a given modulation scheme, 8PSK in Figure 3, higher code rate (2/3) implies more error protection bits for each data bit, hence better error protection but it then requires higher bandwidth. For a given bit error rate, say 2×10^{-4}, it can be seen that for a QPSK modulation scheme with convolutional code rate of 1/2 the required signal-to-noise ratio (SNR) per bit, E_b/N_o, is 3.2 dB, for 8PSK modulation scheme with a convolutional code rate of 5/6 an E_b/N_o of 7.1dB is required and for 16QAM modulation scheme with a convolutional code rate of 7/8 an E_b/N_o of 8.2 dB is required. This is better than uncoded transmissions but not enough to get a 10^{-11} bit error rate without huge antenna sizes/gains or extremely high transmitter powers.

Viterbi decoders themselves occasionally generate error bursts. By breaking up bursts by distributing them among a longer sequence of bits and using a Reed Solomon code to correct block errors, the problem of error bursts out of the Viterbi decoder can be mitigated [4]. This is the approach used by the DVB standard used worldwide and the proposed

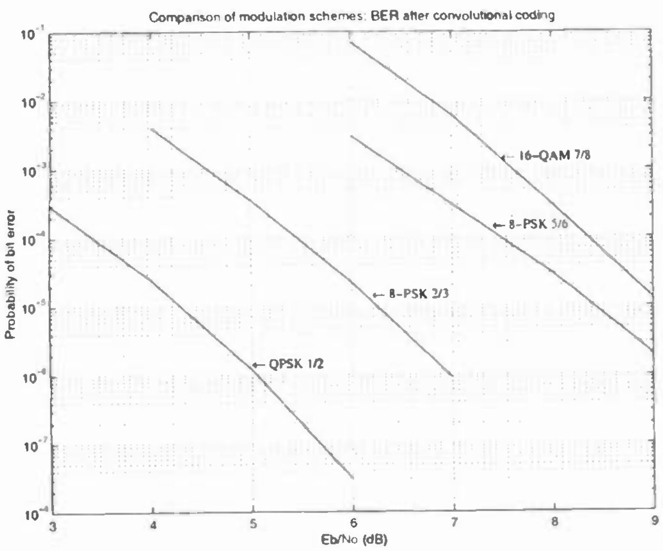


Figure 3: Bit error rates after convolutional coding. The convolutional code rate is indicated alongside the modulation scheme. These simulation results assume perfect timing and phase synchronization. 16QAM ¾ and 8PSK 5/6 error curves are within a couple of tenths of a dB within one another and hence not indicated to preserve clarity. The QPSK ½ result is from [1].

ATSC standard for contribution and distribution [5, 6, 7]. DVB specifies the following table of required Eb/No for quasi-error free (QEF) operation (BER of 2×10^{-4} after the Viterbi decoder or a BER of 10^{-11} after the Reed-Solomon decoder).

Modulation	Conv. Code rate	Required Eb/No for BER for QEF (dB)
QPSK	1/2	4.5
	2/3	5.0
	3/4	5.5
	5/6	6.0
	7/8	6.4
8PSK	2/3	6.9
	5/6	8.9
	8/9	9.4
16QAM	3/4	9.0
	7/8	10.7

Table 2-2: Comparison of BER performance requirements of QPSK, 8PSK and 16QAM per EN 300 42, pr EN 300 210, ATSC A/7x performance requirement [5,6,7].

Table 2-2 above assumes an "implementation margin" of 0.8dB for QPSK, 1 to 1.5 dB 8PSK and 1.5dB for 16QAM ¾ and 2.1dB for 16QAM 7/8 for modem imperfections. This table shows the minimum E_b/N_0 required at the input to the receiver to guarantee the BER requirement for the three modulation schemes at various convolutional or "inner" code rates. The Reed-Solomon or the "outer"code used is (204,188).

2.3. Bandwidth Efficiency

Higher signal power is required for higher order modulation schemes in order to keep the signal points in the constellation well separated. However it is important to bear in mind the bandwidth efficiencies that higher order modulation methods offer. For a given bandwidth, 16QAM is able to squeeze more bits through than 8PSK or QPSK or, put another way, for a given data rate, 16QAM will require a smaller bandwidth than 8PSK or QPSK. In fact a more meaningful metric for comparison of these modulation schemes is not based on BER vs E_b/N_0 but on bandwidth efficiency vs E_b/N_0. *Table 2-3* illustrates this.

Modulation Type	Convolutional Code rate	Bit rate possible in a 36 MHz transponder (Mbps)
QPSK	1/2	24.5752
	2/3	32.7669
	3/4	36.8627
	5/6	40.9586
	7/8	43.0065
8PSK	2/3	49.1503
	5/6	61.4379
	8/9	65.5338
16QAM	3/4	73.7255
	7/8	86.0131

Table 2-3: Comparison of QPSK, 8PSK and 16QAM efficiencies

In an SCPC application there are many more channels that can be accommodated in a transponder when higher order modulation schemes are used. For example, in a 36 MHz transponder, four channels each at 8.448Mbps can be accommodated when the modulation used is QPSK with convolutional code rate of ¾. In the same transponder, eight channels each at 8.448Mbps can be accommodated when the modulation scheme used is 16QAM with convolutional code rate of ¾. There are however other consideration that arise when stacking higher order modulated carriers on a transponder. These will be examined in Section 3.

A metric that is often used in comparing modulation schemes is the bandwidth efficiency which is defined as data rate/ bandwidth (R/W bits/sec/Hz). Figure 4 shows the trade-off involved in selecting modulation schemes in terms of this metric.

It can be seen that QPSK 7/8 and 8PSK 2/3 have comparable power efficiency but the latter has better spectral efficiency. Hence it would appear that 8PSK 2/3 is preferable to QPSK 7/8. When there are imperfections in the system like phase noise, the choice is no longer obvious. Similarly, it is not obvious that 16QAM ¾ is better than 8PSK 5/6 because 16QAM is sensitive to both phase and amplitude nonlinearities. This is discussed next.

3. Perturbations To Be Considered

All the results in previous sections assumed that the satellite channel was linear. There are several perturbations which in practice cause deviations from this assumption. Some of these are filter distortion, timing errors, phase noise. Higher order constellations exhibit more sensitivity to these deviations from perfect behavior.

3.1. Filter Distortions

There are filters at various points along the transmit and receive chain. For example, modulators, upconverters, IF amplifiers, diplexer filters, transponders, downconverters etc. have filters in them. These filters are characterized by an amplitude and phase or group delay response. While the design goal often is to approximate "brick wall" type filters with flat amplitude spectrum and linear phase or constant group delay practical considerations, cost being one, make the implementation of these filters non-ideal. The filter amplitude and phase nonlinearities can then distort the signal which, in turn, degrades the modem performance. There have been some studies that have evaluated the effects of these filter imperfections on performance [9]. Typically, in these studies, the filter response is characterized in terms of simple mathematical functions. For example, for QPSK modulation with phase nonlinearity modeled as a parabolic function, a phase deviation of 10° will cause a 0.75dB degradation in E_b/N_o at a BER of 10^{-6}. With the phase nonlinearity modeled as a cubic function, the same phase deviation of 10° will cause a 0.6dB degradation. At 20° the parabolic function degrades the signal more than the cubic function.

The effects of these filter distortions not only depend on the filter distortion but also on the position of the filter within the link. When the filter bandwidth becomes comparable to the bandwidth/symbol rate of the signal being filtered, the effect on performance becomes very significant. For example, for a Chebyshev transmit filter (5 pole equi-ripple), the degradation is 1 dB in E_b/N_o at a BER of 10^{-6} if the ratio of filter bandwidth to symbol rate is 2:1 [9].

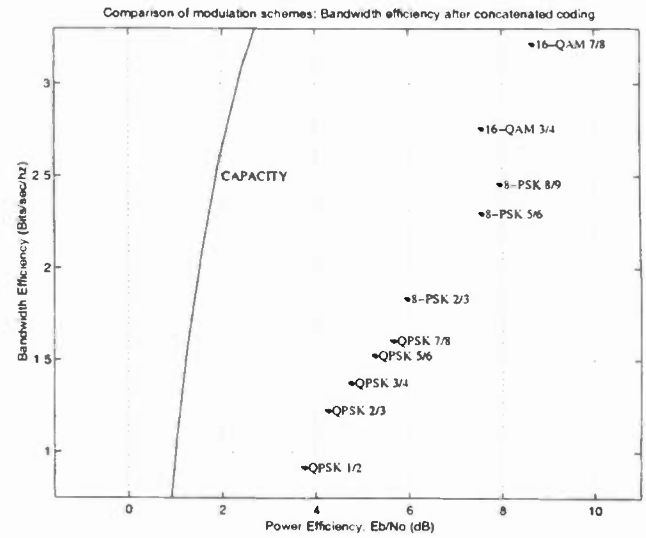


Figure 4: Comparison of various modulation schemes with concatenated coding at a BER of 10^{-11}. The Reed-Solomon code used in all cases is (204,188). Modem implementation margins have not been considered [7]. Channel capacity is the best figure of merit that one can hope to achieve.

The degradation mechanism is very complex, and depends on filter nonlinearities along various points in the chain. So instead of trying to characterize the effect of individual filters in a link, DVB and ATSC specify an amplitude and a group delay mask representing the cumulative effect of multiple sources [5, 6, 7].

When amplifiers are driven to saturation amplitude and phase linearities overwhelm other forms of distortion. Non linearities in amplitude response can also affect phase (AM/PM). This AM/PM distortion becomes dominant when the drive level to an amplifier input is low. For PSK systems, the information is carried only in the phase, but in QAM systems, both amplitude and phase carry information. While Figure 4 shows that 16QAM ¾ is better than 8PSK 5/6, QAM systems are vulnerable to non-linearities more so than PSK systems. In fact there has been a study that indicates that trellis-coded 16PSK and uncoded 8PSK perform better than 16QAM in nonlinear satellite channels. By using pre-distortion filters or equalizers, these nonlinear effects can be overcome.

3.2. Phase Noise, Bit Timing Errors

The bit timing recovery mechanism in the receiver will have timing errors (jitter) due to oscillator phase noise and frequency drifts. Frequency synthesizers used in various modules in the satellite link introduce phase noise and frequency drifts. Phase noise and frequency drifts also influence the carrier recovery in the demodulator. The effect of phase noise can be characterized by the SNR in a phase locked loop in the carrier recovery system in the receiver. If $p(\phi)$ is the phase noise probability density function, then the influence of phase noise on the symbol error rate can be calculated as follows

$$\overline{P_s} = \int p(\phi)\, P_s\, d\phi$$

where P_s is the symbol error probability without phase noise. For example, in a QPSK system with high loop SNRs (>5Eb/No) in the carrier recovery phase-locked loop, the BER degradation is less than 0.2dB. For small loop SNR (< Eb/No) the effect starts to become pronounced (>0.5dB degradation in BER performance) [10]. The same method is used to estimate the influence of timing error in the bit synchronizing loop in the receiver.

3.3. Co-Channel Interference

Co-channel interference introduced by cross-polarization, intermodulation products or sidelobe interference, also causes additional BER degradation. At a BER of 10^{-6} and 20dB of signal power to co-channel interference power ratio, the degradation is about 1dB for QPSK and 2dB for 8PSK[10]. In general, as the number of signal points in the constellation increases, the susceptibility to various types of noise increases.

4. Experimental Results

Experiments were performed over a live satellite segment to see how the QPSK, 8PSK and 16QAM systems would stack up in real-life. Two data sources were used:

1. A multiplexed transport stream at 44.736Mbps with a high definition ATSC compliant MP@HL stream at 19.3 Mbps, a standard definition MP@ML stream at 8.448Mbps and the rest of the bandwidth being filled with null transport packets. This author believes that this was the first time ever that multiplexed HD/SD bit streams were created.
2. A high definition encoder operating alone at ATSC rates.

The bit stream from the sources was fed to a modulator. The signal was then modulated, upconverted, amplified, uplinked and then downlinked, downconverted and demodulated.

4.1. BER Tests

For this test, the multiplexed source was used. The satellite transponder was initially driven to saturation. The Carrier-to-Noise (C/N) at this saturation point was 22dB. The power was then backed-off until the demodulator indicated a BER of 10^{-10}.

HD rate Mbps	SD Rate Mbps	Aggregate Mbps	Modulation	Conv Code Rate	Symbol Rate MSps	E_b/N_o (dB)
19.3	8.448	44.736	QPSK	5/6	29.12	~6
19.3	8.448	44.736	8PSK	5/6	19.52	8
19.3	8.448	44.736	16QAM	7/8	13.87	10

Table 4-1. Eb/No at a BER of 10^{-10} for QPSK 5/6, 8PSK 5/6 and 16QAM 7/8 concatenated coded signals.

This was done to test the modem performance. A comparison of *Table 4-1* with *Table 2-2* illustrates that the modem performance is good, albeit the Eb/No's are at slightly different BERs.

4.2. Fade Test

For this test, the second source was used. The satellite transponder was initially driven to saturation. The power was then dropped to simulate a fade. The amplifiers were now operating in their linear region. This test was conducted to examine at what C/N the demodulator would lose lock. This figure would indicate the cumulative effect of Gaussian and phase noise in the system. The results are consistent with the previous experiment.

HD rate Mbps	Modulation	Conv Code Rate	Symbol Rate MSps	C/N at loss of lock (dB)	E_b/N_o at loss of lock (dB)
19.3	8PSK	5/6	8.46	11	7.4
19.3	16QAM	7/8	6.01	14	8.9

Table 4-2. Eb/No at a loss of lock for8PSK 5/6 and 16QAM 7/8 concatenated coded signals indicating the cumulative effect of Gaussian and phase noise.

4.3. Intermodulation Test

To test the effect of intermodulation effects due to amplifier nonlinearities, two HDTV signals each occupying 6MHz were placed on carriers that were 9MHz (1.5xsymbol rate) apart. While QPSK signals can be stacked 1.3 symbol rates apart with a small degradation in performance, the intermodulation effects with QAM signals were severe and the demodulators would not lock when the carriers were placed 1.3 symbol rates apart. The C/N at saturation was the same as before 22dB.

HD rate Mbps	Modula -tion	Conv. Code rate	Symbol Rate MSps	C/N at loss of lock (dB)	Eb/No at loss of lock (dB)
19.3	16QAM	7/8	6.01	15.5 both carriers	10.4

Table 4-3. Eb/No at a loss of lock of two16QAM 7/8 concatenated coded signals with carriers separated by 9MHz.

Comparing the Eb/No when the demodulator lost lock between the 16QAM signals in *Table 4-2* and *Table 4-3* it can be seen that with both carriers present there is a performance degradation of 1.5dB because of the intermodulation products generated by each carrier. Comparable tests with QPSK signals indicate that the degradation is less than 0.5dB.

5. Conclusion

A basis for comparison of three higher order modulation schemes was established. The bandwidth efficiency of the three modulation schemes was examined. The cumulative effect of imperfections in the design of various components in the link give rise to non-ideal, nonlinear satellite channel models that are difficult to theoretically predict with any degree of precision. As a result, using margins for these imperfections is the method of choice when calculating link budgets. In general, the higher the order of modulation, the higher the susceptibility to various types of imperfections in the link. The feasibility of higher order modulation for digital television signals over satellite was tested with a set of experiments.

6. Acknowledgments

The author would like to thank Dr. Jack Ma for the simulation results in Figure 3.

7. References

[1] Proakis, J.G., "Digital Communications", McGraw-Hill, 2nd ed., 1989

[2] Viterbi, A.J., "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm", IEEE Trans. Inform. Theory, vol. IT 13, Apr 1967.

[3] Heller, J. A., and Jacobs, I.M., "Viterbi decoding for satellite and space communication", IEEE Trans. Comm. Tech., vol COM-19, Oct 1971

[4] Odenwalder, J.P., "Optimal decoding of convolutional codes", Ph.D. thesis, UCLA, 1970

[5] EN 300 421, "Digital broadcasting systems for television, sound and data services; Framing structure, channel coding and modulation for 11/12 GHz satellite services", Dec 1994

[6] ATSC T3 Doc. T3-497, "Modulation and coding requirements for digital television (DTV) applications over satellite", Draft standard, 11 Aug, 1998

[7] pr EN 300 210, "Digital Video Broadcasting (DVB): Framing structure, channel coding and modulation for digital satellite news gathering (DSNG) and other contribution applications by satellite", Draft March 1998

[8] Biglieri, E., "High-Level Modulation and Coding for Nonlinear Satellite Channels", IEEE Trans. Comm., vol COM-32, May 1984

[9] Jones, J.J, "Filter distortion and intersymbol interference effects on PSK signals", IEEE Trans. Comm. Tech., vol COM-19, Apr 1971.

[10] Spilker, J.J., "Digital Communications by Satellite", Prentice-Hall, 1977.

DTV Audio Workshop

Wednesday, April 21, 1999

9:00 am - 12:00 pm

Chairperson: Andy Butler
PBS, Alexandria, VA

9:00 am The ATSC Digital Audio System
Craig Todd
Dolby Laboratories Inc.
San Francisco, CA

***9:30 am Audio Processing for DTV**
Robert Orban
Orban
San Leandro, CA

***10:00 am Microphone Technique and Monitoring for Digital Television**
Tomlinson Holman
TMH Corporation
Los Angeles, CA

***10:30 am A 12 Channel Digital Audio Interface**
Birney Dayton
NVISION, Inc.
Grass Valley, CA

11:00 am A Multichannel Audio Infrastructure Based on Dolby E Coding
Stephen Lyman
Dolby Laboratories Inc.
San Francisco, CA

*Papers not available at the time of publication

The ATSC Digital Audio System

Stephen B. Lyman
and
Craig Todd
Dolby Laboratories Inc.
San Francisco, CA

ABSTRACT

This paper is an overview of the Audio system that has been selected by the Advanced Television Systems Committee (ATSC) for use in the US and Canadian High Definition Television systems, and has been selected as the preferred method of carrying the audio in the Australian HDTV system.

It starts by start by examining why producers would want to use multichannel sound for television, what the requirements for a multichannel audio system are, and perhaps most importantly, the operational changes that will have to make in going from the one or two channel system used with NTSC television to a multiple channel sound system.

The paper continues by examining some of the features and service options that are provided by the ATSC specification, then examines a method of enabling the rest of the broadcast chain to handle multichannel sound and the metadata that DTV receivers depend on.

1.0 The Antecedents of Multichannel Sound Systems

Television started with a single channel of sound. In many places in the world, television still uses monophonic sound, often with limited bandwidth. It may seem strange, with all the emphasis on quality these days, but the appeal of monophonic sound is that it is simple. It requires the least amount of equipment in the plant, the transmission system, and in the home receiver. The main advantage of monophonic sound is that the listener can sit almost anywhere in the room and hear all of the program material. The source of the sound also stays with the picture, no matter where in the room the listener is sitting. It does not, however, give the production staff very much freedom to create a realistic sound field or to emphasize aspects of the visual image.

Stereo or two channel sound was an attempt to give to program producers more freedom to create a realistic sound field. Most receivers have speakers that are mounted on the sides of the set or are placed some distance to the sides of the receiver's screen. This means that the sound that is closely associated with the image, such as the dialog, has to be produced by a "phantom center" sound image. The placement of this "phantom" sound source depends on the balance between the two stereo speakers . If the relative left and right channel balance has not been changed, and if the listener is sitting on the center line of the two speakers, the phantom sound image is placed in the center of the screen, as it should be. If the listener moves, or if the left to right channel balance is disturbed, the image shifts, which is annoying. Stereo reproduction also limits the sound designer to aural events that happen in front of the listener. Stereo is more flexible than mono, but has its own problems and does not provide enough flexibility to create realistic sound fields.

A great deal of experimentation backed up by experience in the film industry has shown that five independent channels of sound are the minimum number required to be capable of producing realistic sound fields. Unlike the stereo situation, the Center channel speaker provides a solid link between the visual and aural images. The listening area is consequently much larger than for the stereo case; almost as large as with mono sound. Because all the channels are independent, the sound designers can either create an enveloping ambient sound field, or can use the individual channels to create specific effects, placing program elements precisely in the sound field.
A sixth speaker, a subwoofer, is used for the ".1" or Low Frequency Effects channel. The bandwidth of this channel is restricted to 3 Hertz

to 120 Hertz, or about 10% of the bandwidth of the other channels (hence the term ".1"). This channel has been used in film presentations to handle loud low frequency special effects, so has 10 dB more headroom than the other main channels to accommodate these levels. It can be used for the same purpose in television practice, but it may also be used to supplement the low frequency response of the five main speakers, to allow the use of smaller, more aesthetically acceptable speakers in the home. Because low frequency material provides very few directional cues, the low frequency content of the main channels can be combined and reproduced by the sub woofer speaker without harming the desired aural image.

2.0 Rate Reduction

One of the main differences between analog and digital television broadcasting is that the data rate of both the audio and video signals must be drastically reduced if the combined signal is going to fit the available spectrum. The audio data rate must be reduced from about 4.8 Megabits per second to 384 kilobits per second, without loosing subjective quality.

The following explanation of how digital audio rater reduction works only touches on the most basic aspects of the process. It is intended to give the reader an intuitive feel for the process, rather than being an exhaustive analysis of the process. Rate reduction is all about managing quantizing noise. In linear systems, 16 bit resolution is considered to be about the practical minimum number of bits to use to keep the quantizing noise down to an acceptable level (in this case about 96 dB below the maximum signal level). If we want to use fewer bits to represent the signal, we have to find a way of dealing with the increased level of quantizing noise. Fortunately, the human hearing process provides several mechanisms to do this.

The first is the basic threshold of hearing. Our ears tend to be less sensitive at low and high frequencies than they do at mid frequencies. The second characteristic of the ear that makes rate reduction possible can be understood by considering the structure of the inner ear. The cochlea is a spiral, tapering passage with the basilar membrane stretched more or less across the diameter along its length. Sound is conducted from the outer ear to the fluid in the cochlea where it travels the length of the basilar membrane. Different frequency components of a sound wiggle the hair cells at different locations along the membrane, stimulating the auditory nerves. The frequency dependent movement of the hair cells makes the ear act like a spectrum analyzer. A high level frequency component will not only wiggle the hair cells at the location sensitive to that specific frequency, but some of the adjacent hair cells as well. This "spreading" of the response to a specific frequency can override or mask the response to other lower level, nearby frequency components. The ability of relatively loud sounds to mask lower level ones is usually described by sets of frequency and level dependent "Masking Curves" .

If the quantizing noise produced by a coarse quantizer can be confined to the spectral region near to the signal component being quantized (or encoded) and if that noise is low enough to fall below the masking curve of the signal being coded, then the listener will not be able to hear the quantizing noise.

Complex program signals are transformed into the frequency domain, and the masking curves for the different signal components computed. The masking and hearing threshold curves (and other similar phenomena) are superimposed on the spectrum of the program signal. This determines the limits on the level of quantizing noise that can be "hidden" by the program signal. The encoder can then make decisions about the coarseness of the quantizer, or the number of bits that will be assigned to each of the frequency components of the program signal.

The recovered program now no longer has the uniform low level noise floor of a PCM (linearly) coded signal, but a dynamically changing, program material dependent noise floor that is part of the program signal. The rate reduction process thus leaves its "signature" on an audio signal, in the same way that the NTSC process leaves its "chroma crawl" signature on a video signal.

An encoder fed with a previously encoded and decoded signal will make its decisions about the amount of quantizing noise that can be concealed by that signal. The noise added by the second and subsequent rate reduction processes will add to that created by previous generations, and will rise towards the masking curve limit. At some point the demand for bits will exceed the supply, and no matter what efforts the encoder makes to avoid it (such as limiting the high frequency content) the noise will exceed the capability of the signal to mask it, and the listener will hear "coding artifacts". At this point, we can say that the process has run out of coding margin.

In very general terms, rate reduction systems that operate at low data rates do not cascade or tandem very well because they have to operate at low coding margins to achieve the low rates. Coders intended to be tandemed must operate at higher coding margins, and all other things being equal, must operate at higher data rates.

3.0 Downmixing

Television currently has to deal with one or two channel program material. Depending on the receiver, the program is either presented as such or the channels are combined for a mono presentation. Digital television is quite different in that every program will be seen by many different home receivers, each capable of presenting anywhere from one to six channels of sound, depending on the desires of the listener. The type of program and the desires of the producer will determine if the audio will be produced with either one, two, four, five or six channels. The broadcaster has no choice other than to transmit as many audio channels as are supplied by the program, and as will be shown later, with the full original dynamic range. The DTV audio system must be able to fulfill all these requirements simultaneously. This is a big change from the conventional practice of creating and transmitting a "one size fits all" program.

The key to being able to do this is to transmit some information about the audio program signal, or metadata, to the receiver. This metadata, in combination with information supplied by the listener about the number of reproduction channels available, allows the receiver to downmix a multichannel program to the number of channels available.

4.0 Control of Loudness

The current TV audio practice is to try to provide a "one size fits all" kind of signal. The mono or stereo program material is produced with a relatively restricted dynamic range that "fits into" the approximately 20 dB of headroom provided by most current systems.

Changes in loudness from program to program have always been a problem. Currently, the only way of trying to normalize the subjective loudness of programs has been to further (and sometimes drastically) reduce the dynamic range of the program material, increase the average level until all the programs occupy the top part of the dynamic range available, and are thus roughly the same loudness. This necessitates limiting or clipping the peaks to avoid over modulating the transmitter and leaves very little, if any, of the original carefully constructed program dynamics.

The ATSC sound system uses another form of metadata to provide uniform loudness to the listener. Each program style, if not each program, will have specific headroom requirements that dictate where in the available dynamic range the "average level" or loudness of the material falls. This point can be identified by the "dialnorm" metadata parameter. If the dialnorm is transmitted to the receiver along with the program, the receiver can reproduce all program material at a common loudness level. In the case of ATSC compliant receivers, the program material is attenuated by the difference between –31 dB and the dialnorm parameter. If the dialnorm value is correct, then all the program material will be reproduced 31 dB below the clipping level, and will be presented at (ideally) the same loudness.

Since there is currently no universally accepted method of measuring loudness, the process is subjective. This makes the value of dialnorm a judgement call, but also permits different styles of programs to have different loudness, as they should.

5.0 Dynamic Range Control

Listeners do, of course, need some control of the program dynamic range. Feature films, for instance, tend to have large changes in loudness which may be totally unsuited for late night listening. The best solution would be to give each listener control of the program dynamics, rather than force all listeners to make do with the same restricted dynamics as present practice does.

The choice of dynamics is also made possible with the help of metadata. The system establishes a band around the average program loudness (as defined by the dialnorm value) where no processing is done. Levels above the deadband can be reduced, and those below it can be brought up independently. This process leaves the loudness of the most important parts of the program (usually the dialog) unaffected. The listener has control of how much compression the receiver will apply so can listen to a heavily compressed program, or to the entire original dynamic range, depending on their individual desires.

The dynamic range control metadata is generated by the Dolby Digital encoder at the end of the signal chain, according to one of several compression profiles selected by the production crew. This allows selection of an artistically appropriate method of compression, rather than the one size fits all technique used today.

6.0 Sound Quality

The requirement for high sound quality is obvious. The difficulty is that there is no way to specify objectively what acceptable sound quality is. Because the data rate reduction systems used for audio are based on how the human ear perceives sound, the only way to measure the sound quality is with a very laborious and time consuming series of subjective tests. Subjective measurements are made by panels of listeners assigning ratings to samples of program material, referenced to the original material.

The restricted amount of public spectrum demands that the audio use as little data rate as possible, and yet that the rate reduction system deliver high quality sound. The ATSC specification limits each "Complete Main" multichannel audio service to 384 kilobits per second. This is a slightly higher rate than the film industry uses (324 kb/s) and a bit lower than the 448 kb/s used for multichannel audio on DVD.

MPEG 2, Layer II, BC (for backwards compatible) is another audio rate reduction system. It was originally selected to handle multichannel sound for European DVDs (but has since been dropped). MPEG 1 was designed as a two channel system, and later extended (hence the MPEG 2 label) to multiple channels through the use of a "compatibility matrix" which demands that additional data be transmitted. This partially redundant information reduces the overall coding efficiency of the system, and thus may demand higher data rates to maintain the quality of the recovered signal.

The only recent subjective quality tests that have been done for multichannel audio codecs were completed in February of 1998 by the European Broadcasting Union. These compared Dolby Digital and MPEG 2, Layer II, BC codecs, both operating at 384 kb/s and then the Dolby Digital codec at 384 kb/s to the MPEG 2, Layer II, BC codec running at 512 kb/s. The testing group felt that the 512 kb/s data rate was justified and that the second test would be more representative of actual operations because of the additional (helper) data required by the MPEG 2, Layer II, BC algorithm when used for multichannel audio.

The results showed that with both codecs operating at 384 kb/s, the Dolby Digital codec delivered better quality on six of the eight test items. When the MPEG 2, Layer II, BC codec data rate was increased to 512 kb/s, the score changed (as expected) with the Dolby Digital codec delivering better quality on five of the eight test items.

7.0 ATSC Audio Services

The ATSC audio system specification includes provisions for several different types of audio services. Implementation of the Associated services depends on the receiver manufacturers' willingness to supply the second audio decoder. The paper describes these services for sake of completeness, but the reader should be aware that it may not be possible to provide some of the Associated services if receivers do not include the second decoder. Dual decoder receivers for special audiences may become available, but broadcasters should not count on being able to supply these additional services universally, at least in the near term.

The services types are defined as follows:

<u>Main Services</u>
* A Complete Main service has all the elements (music, effects and dialog) of a normal complete audio program.
* The Music and Effects Main service lacks only the dialog elements. (It may also be called international sound).

<u>Associated Services</u>
* The Visually Impaired service can be either just a narrative description of the image, or a complete mix of all the program elements.
* The Hearing Impaired service can also be supplied as dialog processed for better intelligibility or a complete mix of all the program elements.
* The Dialog service carries one or more channels intended to be mixed into the M&E service to provide a choice of languages.
* The Commentary channel can be thought of as a dialog channel containing optional rather than

necessary program contents. It may be a single channel decoded along with the Complete Main service, or may be a complete service itself.
* The Emergency service is a single channel that overrides any other service(s) that may be in use when it is transmitted.
* The Voice Over service is a single channel that is decoded and added into the center channel.

Each program item has an individual program identification code that can be selected by the listener. The Associated services are implemented by selecting either the complete or music and effects Main service and mixing the Associated Service form the output of the second decoder with the output of the main decoder. The reader is referred to section 6.6 of ATSC document A/54 for a complete description of the different service types. (See the ATSC web site at www.atsc.org).

8.0 Getting Started with ATSC Audio Programming

This paper does not cover any of the issues associated with producing audio for DTV broadcast except to say that multichannel audio techniques will probably initially be borrowed from film techniques. These methods will serve, and probably continue to serve the more critical productions well. Eventually, when enough practical experience has been accumulated, multichannel techniques will be developed for rapid turnaround of day to day TV productions.

There is a way to get multichannel programs from the point of origination through the rest of the broadcast system to the transmitter and to get the metadata needed by domestic receivers there in sync with it. Dolby E is a new form of rate reduction that is tailored specifically to the needs of program distribution. It forms a multiplex of anywhere from one to six or eight channels of audio and the metadata associated with the program(s) which is carried through the contribution and distribution portions of the broadcast system. It can be edited and encoded and decoded enough times, without loosing subjective quality, to meet the needs of typical contribution and distribution applications. It is a digital signal, and thus requires at one AES/EBU signal path in the broadcast plant.

Some broadcasters may elect not to upgrade their existing analog stereo audio facilities immediately. They have had a "surround Presence" in the NTSC market and will be able to keep it in the DTV market through the use of Dolby Surround Pro Logic matrix encoding. The Dolby Surround system has only one band limited surround channel, and some production restrictions, but it can operate in a two channel (analog or digital) infrastructure. The main problem is that there is still no signal path for the metadata that the DTV system needs.

9.0 Signal Formats for Two Channel Systems

The approach to audio in the ATSC system is, as was mentioned earlier, is to transmit the maximum number of channels available with their full original dynamic range intact, then let the individual receivers create the desired downmix and present it with the desired dynamic range.

Many broadcasters are planning to adapt or "upconvert" existing video material for use with DTV; line doublers and aspect ratio conversion are some of the ways being considered to make current equipment and program archives merge with new HD productions.

Single or two channel audio can't be "upconverted" to a multichannel source; the spatial information was either lost in the original downmix or was missing in the first place. One idea in the industry was to decode Dolby Surround material to Left, Right, Canter and (mono) Surround, split the Surround channel to the Left Surround and Right Surround inputs of the Dolby Digital encoder then to label it as 5 channel material, with the intent of lighting the "Multichannel" indicator on the receiver. This is a very good example of what not to do.

Assume that the false 5 channel program is received by a listener who wants a monophonic presentation. The original Dolby Surround signal should be "mono compatible" meaning that the Left total and Right total (Lt and Rt) signals will sum to a signal that still has some of the Surround channel material present. This requires that a portion of the Surround elements also be present in the front channels so it doesn't cancel and disappear in the mono downmix.

A Pro Logic decoder adds some delay to the Surround channel to create a feeling of being enveloped by the Surround material. If the fabricated five channel signal is mixed down to a mono presentation in the home receiver, the

Surround elements will suffer from comb filtering effects because of the delay and the fact that they are present in both the Surround and Front channels. Stereo or Dolby Surround encoded two channel should be presented as such to avoid problems downstream.

10.0 Preset Metadata Parameter Values and Signal Formats in Existing Systems

There are no signal paths for metadata in current broadcast plants, yet the DTV receiver will not function as it should without metadata. The solution is to pick a set of parameters that can be loaded into a preset of the DP 569 Dolby Digital encoder at the transport stream multiplexer input that will allow Dolby Surround, Stereo or Mono programs to be reproduced gracefully in either four, two or single channel receivers.

If the metadata labels Dolby Surround, stereo or mono signals as being two channel (with the encoder in 2/0 mode) material that is Dolby Surround encoded, receivers will provide the following presentations for the different programs.

If the program signal is Dolby Surround encoded, a receiver set to reproduce four channels will decode the Lt, Rt signal to Left, Center, Right and Surround signals, as expected. If the program is in stereo, the receiver will attempt to do the same thing. Any components common to Left and Right channels and in phase will be reproduced by the Center speaker. Out of phase or non correlated components will be reproduced in the Surround while signals originally in the Left and Right channels will be reproduced there. Stereo signals are thus decoded quite gracefully.

A mono program should be lowered in level and split equally between the Left and Right channels. It will be reproduced as a Center channel signal, with nothing in the other three channels, exactly as it would be desired.

A receiver set for two channel reproduction will not decode Dolby Surround Lt, Rt or a plain Stereo program, but will present them as a stereo signals, as desired. Mono signals should again be dropped in level and split equally to Left and Right. They will be reproduced as a phantom Center signal, with the same material in both speakers.

A receiver set for single channel reproduction will sum both channels of the Lt, Rt, Stereo or split Mono program. Any Dolby Surround encoded material that has been mixed for good mono compatibility will sum to a mono presentation well, as will the Stereo signal. The split mono signal simply returns to its original form.

If the broadcaster cannot change to Dialnorm parameter to reflect the average loudness of each piece of program material, then falling back on the current practice of using some form of multiband signal compression as a way of trying to control loudness variations from program to program is the only option. The Dialnorm parameter, however, still has to be specified.

The type of processing done to the station's normal program material will determine the effective loudness and hence the appropriate value for the Dialnorm parameter. The value can be determined by subjectively comparing the program material to similar material with a known Dialnorm value. Any gain changes required to equalize the loudness of the two programs can be added to or subtracted from the Dialnorm of the reference material to arrive at a value for the processed programming. Dolby will soon make available a Compact Disc with selections of programs whose Dialnorm values are known.

The caveat in this case is that the usual function of the Dialnorm parameter has been abdicated to the compression and limiting processing. The loudness from program to program will not be as well controlled, particularly when the original version of the program material had quite a wide dynamic range.

The question of dynamic range gets to the next metadata parameter. The DP569 offers five different dynamic range reduction "profiles" that can be applied to different classes of programs. The "light" profiles have a deadband in which no signal processing is done that is plus or minus 10 dB wide and centered on the specified Dialnorm value. If the program has been processed to try to create a relatively uniform loudness, then it is unlikely that the light dynamic range reduction profiles will have any effect at all. The deadband of the "standard" profiles is plus and minus 5 dB around the Dialnorm value, so these may cause some additional dynamic range reduction at the receiver. The results will clearly depend on the type and amount of preprocessing done to try to control loudness variations. Dynamic range control profiles can only be selected experimentally.

Center and Surround downmix parameter values determine how much of the center and surround channel material will be mixed into the Left and Right channels of two channel presentations. These parameters clearly affect the sound of the program (in fact the surround can be turned off). The best option here is probably to pick values that suit local production practices and use these for all programming.

11.0 Preset Availability

The next problem is how to get sets of metadata parameters into the system and how to change sets of parameters when the program material changes.

There are four different sets of user defined metadata parameter and operating mode that are easily available. Operationally, there are many choices of how to use and trigger these presets. In the early days of DTV, there will probably only be a few sources of material for the DTV service. Many of these may well have quite similar characteristics (most film originated Lt, Rt material will have the same Dialnorm value for instance) so will be able to share the same preset. Another two presets could be used for the bulk of the locally originated programming. Some networks are and will provide true multichannel feeds on a more and more regular basis. In these cases, a preset can be dedicated to multichannel material.

The desired preset can be triggered or recalled by pulling one of the four rear panel GP I/O Inputs to ground (the input is triggered by a low going transition, not a low level). The GP I/O Outputs include tallies for each of the four available presets. The trigger signal may be provided by a contact closure from an on air automation system or even from a tally signal associated with a specific source on the Master Control presentation switcher. There are obviously many possibilities that depend on the specific facilities.

If more than four presets are required, new sets of values can be defined or existing presets recalled via commands sent to the RS-422 Remote Control port on the rear panel. Command strings can be sent to the port from an automation or similar system, depending on what station facilities are available.

12.0 Summary

The ATSC Digital Audio system presents the broadcaster and listener with a much wider choice of presentation formats and listening conditions than does the existing analog television audio system. The ATSC approach does depend on metadata support. Ideally, the metadata will originate with the program material and be carried through the Contribution and Distribution systems (probably in the multiplex provided by Dolby E type codecs) to the Dolby Digital encoder at the transport stream multiplexer. If this path for metadata does not exist, there are a variety of ways to access or create metadata sets at the Dolby Digital encoder, ensuring that the ATSC audio system will operate satisfactorily, even in the stages of its introduction.

A Multichannel Audio Infrastructure Based on Dolby E Coding

Stephen Lyman
Dolby Laboratories Inc.
San Francisco

ABSTRACT

Dolby E Coding was developed to expand the capacity of an existing two channel AES/EBU digital audio infrastructure to make it capable of carrying up to eight channels of audio plus the metadata required by the Dolby Digital coders used in the ATSC and other DTV transmission systems. This allows the existing digital video tape recorders, routing switchers and other TV plant equipment, as well as satellite and telco facilities used in program Contribution and Distribution systems to handle multichannel audio. The coding system has been designed to provide broadcast quality even when decoded and re-encoded many times and to provide clean transitions when switching between programs.

1.0 INTRODUCTION

The ATSC audio system uses Dolby Digital bit rate reduced audio (based on AC3 rate reduction technology) to carry anywhere from one to 5.1 channels of sound from the DTV transmitter to the home receiver. Some listeners want to hear programs transmitted as 5.1 channels in that format, others want a stereo or even a mono presentation. Some listeners want the full original dynamic range of the material; others want a lower dynamic range version that doesn't disturb the light sleepers in the family. It is impossible to transmit a "one size fits all" signal anymore.

Metadata is the answer to providing listeners with the choices they need. Metadata is data transmitted to the home along with the audio data that describes the program material ("data about the data") to the receiver so individual receivers can produce the format each listener wants. The ATSC audio system is built around the metadata concept; it doesn't provide the listener with all the desired features without it.

Dolby Digital is designed to carry the audio and metadata to the receiver at the lowest possible data rate to conserve as much of the precious over the air spectrum as possible for video data and other DTV services. It is not intended to be decoded and re-encoded many times as is required in the extensive series of Contribution and Distribution circuits that make up the bulk of a broadcast system.

Dolby E is a new multichannel audio bit rate reduction system that is designed to tolerate the multiple encoding and decoding operations required by Contribution and Distribution systems. It encodes up to eight audio channels plus the necessary metadata and inserts this information into the "payload space" of a single AES digital audio pair. Because the AES protocol is used as the transport mechanism for the Dolby E encoded audio, Digital VTRs, routing switchers, DAs and all the other existing digital audio equipment from the point of production through to the Dolby Digital encoder at the transmitter can now handle multichannel programming. It's possible to do insert or assemble edits on tape or to make audio follow video cuts between programs because the Dolby E data is synchronized with the accompanying video. The metadata is multiplexed into the compressed audio, so it is switched with and stays in sync with the audio.

2.0 The Existing Infrastructure

Broadcast plants are not generally equipped to handle multichannel sound. The vast majority of plants in North America are stereo analog plants which, though clearly inadequate for multichannel programming, is not the disaster it seems to be at first glance. All of these facilities can, and usually do, have a "surround" presence in the market through the use of Dolby Surround Pro Logic encoded programs. Dolby Surround is a matrix encoding process that allows a 2 channel facility to handle four channel programs. The two channel matrix encoded signal (often called Lt, Rt or Left total and Right total) is fully compatible with stereo analog signal paths and storage devices, as long as the channel to channel balance and phase response is consistent. Listeners equipped with stereo television receivers and a Pro Logic decoder get a Left, Center, Right and mono Surround presentation of Dolby Surround encoded programs. This allows just about any television station to have a "Surround presence" in their market.

A straightforward transition from two channel (analog or digital) audio to multichannel audio is not an easy one. The obvious path is to upgrade existing facilities to six channels (the ".1" channel of a 5.1 channel audio path has a bandwidth of 120 Hz, but is really just another channel from the signal distribution point of view) and add a data path for the metadata. The most unimaginative thing to do in plant would be to upgrade the audio routing system to three layers of AES/EBU capability (to provide 6 audio channels) and a data layer for the metadata. The cost of this would be a lot more cabling to install, more rack space for the switchers (and jackfield) and all the operational problems of operating an additional three layers of routing in parallel.

That approach could possibly work in the plant, but a broadcast system consists of Contribution circuits that are used to bring programs or program segments together and Distribution circuits that are used to send the finished program streams to the individual stations for transmission. The majority of these Contribution and Distribution circuits are limited to something less than six channels (usually two), and do not have an associated metadata channel. The digital video tape recorders and many other storage devices form another serious bottleneck. None of the DVTRs in common use have more than four channel (two AES/EBU pairs) capability. This is a limitation of the tape formats themselves, so is not easy to overcome. Nor do they have space to store the metadata information, so cannot be used for multichannel programming.

The Contribution and Distribution circuits are a bit less limited, in that there is no mechanical media format to limit the data rates, but there are practical limitations to the bandwidth or datarate available. Satellite circuits use public spectrum and the common carriers sell "bits per second per mile" to the broadcaster. These factors suggest the use of some form of audio data rate reduction system to conserve spectrum or to reduce the cost of program distribution.

A little thought about the programming requirements makes it clear that six channels of audio is not enough. For the next several years at least, broadcasters will have to supply both the DTV and existing analog television transmitters. DTV services require from one to six channels of audio plus metadata; the analog TV service needs one or two channels of audio. The two channel "analog TV" sound tracks will probably be supplied as a Dolby Surround or Lt, Rt signals to allow these broadcasters to have a "Surround presence" in their markets, as mentioned above. Some DTV stations may also use the 2 channel signal because of a (temporary) lack of multichannel facilities in their plants. Programmers will thus have to produce and distribute both soundtracks to service both markets. The Contribution and Distribution systems have to handle up to eight audio streams, plus the associated metadata.

3.0 Requirements for a DTV Multichannel Audio Infrastructure

It is clear from the preceding comments that the existing signal distribution infrastructure does not meet the needs of a digital television broadcasting system. Study of these requirements has shown that the Contribution and Distribution systems have to handle up to eight channels of audio and several streams of associated metadata at a total data rate of approximately two megabits per second. There are several data rate reduction systems available, but normal broadcast operations impose several additional constraints. Program material has to be encoded for transmission, then decoded to "sweeten" or combine incoming items with locally produced items. This has to be done several[1] times as the program makes its way from the original point of production to the input to the Emission encoder. Repeated encode – decode cycles and concatenation with another type of rate reduction system (Dolby Digital in

[1] A typical "worst case" number is between 6 and 9 or 10 tandems

the case of DTV) usually leads to some loss of signal quality.

The most common operation in the Contribution and Distribution process is switching from one program feed to another. This may be done by switching between "live" signals, or by assembling different program segments on tape or some other storage medium. In some cases, individual items, such as an advertisement, have to be "dropped into" gaps intentionally left for them in longer program segments. This can either be done live or with insert edits done on a storage device. In many cases, the transitions between different programs are made by fading to silence, making the switch, then fading back up to the new program (a "V" fade) to eliminate any disturbing clicks or pops during the transition. Crossfades between programs or voiceovers are other types of transition, but these are generally done as part of the sweetening process or in the Master Control (Presentation) suite. Because most of the transitions made by a Master Control Switcher are tied to time of day and automated, transitions either occur or begin and end at specific time codes and hence on specific video frame boundaries.

4.0 A Rate Reduced Multichannel Audio Infrastructure

As alluded to previously, it is very difficult to replace the existing audio infrastructure with something that can handle the required number of channels and metadata. It might be possible to enlarge the audio router, but other plant equipment cannot be expanded to the necessary six or eight channels. The same idea applies to interfacility links, except in these cases, it may well be impossible to expand their capacity because of limited availability of spectrum or cost.

Some form of data rate reduction is needed to get the data rates down to practical levels for connections between and within plant facilities. Data rates must also be compatible with digital VTRs and other storage devices so that six or eight channels of audio can be recorded on these existing devices. The system selected must be able to be concatenated both with itself and with Dolby Digital without suffering any subjective quality loss. The data stream must be switchable on video frame boundaries and should produce clean (noiseless) transitions between sources. The transport stream produced by the rate reduction system should be compatible with the existing equipment found in the Contribution and Distribution chain.

4.1 Selection of a rate reduction system for Contribution and Distribution

The first candidate for this application might be the Dolby Digital system itself. It produces low data rate 5.1 channel audio, carries the metadata needed by consumer decoders, and puts all the data into an AES/EBU transport stream, so that it is compatible with all the in plant digital audio equipment. The problem with this approach is that Dolby Digital was designed as a method of delivering multichannel programs at a very low data rate. It was not intended to be concatenated, so while the degree of quality loss after several generations depends on how the program material interacts with the rate reduction algorithm, it is not possible to guarantee that some sort of coding artifact won't appear after the number of generations typically encountered in a Contribution and Distribution chain.

Dolby Digital encoders produce a complete block of rate reduced audio and metadata every 32 msec (assuming a 48 kHz audio sampling rate). This unfortunately does not match the duration a video frame (in any television system) so any audio follow video transitions that are made on video frame boundaries will probably occur during a block of Dolby Digital data, so will corrupt the information. This results in a short mute (1 or 2 blocks long) at the output of the next decoder in the signal path. While this is not a disaster, it is not a good way to operate a broadcast system. Dolby Digital encoders and decoders are thus not the ideal choice for Contribution and Distribution applications.

4.2 The Dolby E rate reduction system

Dolby has designed a new audio rate reduction system for Contribution and Distribution applications. It can be cascaded several times[2],

[2] At 1.92 Mb/s, Dolby E carrying a 5.1 channel program can be cascaded at least 50 times. Adding a stereo program or a pair of mono programs to the 5.1 channel program in the E type stream reduces the allowable number of cascades to 10. Note that these

produces clean audio follow video switches and caries up to eight channels of audio and the associated metadata. The design goals, outlined in the following sections, for the Dolby E system are quite different from those for the Dolby Digital (Emission) system.

4.3 Multigeneration performance

The main problem in designing rate reduction systems for multiple generations is to keep "coding artifacts" from appearing in the recovered audio after several generations. The coding artifacts are caused by a buildup of noise during successive encoding and decoding cycles, so the key to good multigeneration performance is to manage the noise optimally.

The noise is caused by the rate reduction process itself. Digitizing or quantizing a signal leads to an error signal that appears in the recovered signal as a broadband noise. The smaller to quantizer steps (ie. the more resolution or bits used) to quantize the signal, the lower the noise will be. This "quantizing noise" is related to the signal, but becomes "whiter" as the quantizer resolution rises. With resolutions less than about 5 or 6 bits and no dither, the quantizing noise is clearly related to the program material.

Bit rate reduction systems try to squeeze the data rates down to the equivalent of a few bits (or less) per sample and thus should create quantizing noise in quite prodigious quantities. The key to recovering signals that are subjectively indistinguishable from the original signals, or in which the quantizing noise is inaudible, is in allocating the available bits to the program signal components in a way that takes advantage of the ear's natural ability to mask low level signals with higher level ones.

The masking effect of the ear can be understood by imagining the spectrum of a segment of a simple program signal consisting of a strong frequency component at, for example, 4 kHz. Add lower level signal components at 3.5 kHz and about 6 kHz. The relatively high level signal component at 4 kHz stimulates an area on the basilar membrane in the ear that is not confined to the location of the membrane that is most sensitive to 4 kHz, but tends to spread along the membrane. The two lower level signals will also stimulate the basilar membrane, but their stimulus may be overcome, or masked, by the higher level signal, thus rendering the them inaudible[3]. Now if the quantizing noise associated with each program signal component can be confined to the region of the spectrum that is masked by that component, and if the noise level is not allowed to rise above the "masking threshold" it will be present, but inaudible in the recovered signal.

The rate reduction encoder sends information about the frequency spectrum of the program signal to the decoder. The set of reconstruction filters in the decoder confines the quantizing noise produced by the bit allocation process in the encoder to the bandwidth of those filters. This allows the system designer to keep the noise (ideally) below the masking thresholds produced by the program signal. The whole process of allocating different numbers of bits to different program signal components (or of quantizing them at different resolutions) creates a noise floor that is related to the program signal and to the rate reduction algorithm used. The key to doing this is to have an accurate model of the masking characteristics of the ear, and in allocating the available bits to each signal component so that the masking threshold is not exceeded.

When a program is decoded then re-encoded, the re-encoding process (and any subsequent ones) adds its noise to the noise already present. Eventually the noise present in some part of the spectrum will build up to the point where it becomes audible, or exceeds the allowable "coding margin". A codec designed for minimum data rate has to use lower coding margins (or more aggressive bit allocation strategies) than one intended to produce high quality signals after many generations

The design strategy for a multigeneration rate reduction system, such as one used for Dolby E, is therefore quite different than that of a minimum data rate codec intended for program Emission applications.

limits are highly dependent on program material, and are derived from current subjective test results.

[3] Note that there is no one single "ear model" that is universally accepted, and that an algorithm designer may have refined the various published models, based on their experience in designing and testing rate reduction systems.

4.4 Switching

The most common operation in the Contribution and Distribution of television signals is a simple cut transition between program segments. Other operations, like V fades (fade to silence then back to unity) are placed around a cut transition to ensure that here are no transients at the transition. The majority of audio cut transitions are made at the vertical interval switch point of the video signal because it is convenient to slave the audio switcher to the video switcher and because the transition point can easily be labeled with video time code. Unfortunately, as was pointed out in the case of Dolby Digital coding, the block structure of rate reduced audio systems does not match the video frame structure. Audio follow video switches almost inevitably corrupt the blocks of data and cause some sort of interruption in the recovered audio.

The situation for baseband (or PCM) audio is not much better. The majority of digital audio switchers are simple "crash" switchers that make the cut as soon as they receive a command, so can corrupt the audio sample structure. Synchronous switchers wait until the beginning of the next audio sample pair (AES frame) before making the cut, so do preserve the data integrity, eliminating one source of transients. Unfortunately, audio transitions may fall at a time when a large peak of one polarity in the first signal matches a peak of the opposite polarity in the other signal. This produces a sharp transient click in the resulting signal, so even a synchronous switcher cannot guarantee clean transitions between audio signals. V fades that make the transition during the silent period are the usual cure for this problem.

The Dolby E system design assumes that audio transitions will take place at vertical interval switching points[4], so aligns the blocks of rate reduced audio data with these points. The rate reduction algorithm sacrifices a small amount of coding efficiency so that the decoder can make short cross fades between the end of one block and the beginning of the next block of audio information. This eliminates transients at switching points, even if the "positive peak to negative peak" problem is present at the transition point and produces reliably clean transitions. The switching capabilities of the Dolby E system were demonstrated during the 1998 NAB Convention. Atendees were able to switch freely between program streams from three different digital VTRs while listening to the result on a high quality multichannel monitoring system.

5.0 Metadata

As mentioned in the introduction, metadata is an essential ingredient of the ATSC audio system. The unfortunate part of the existing signal distribution system in existing plants is that there is no signal path for the metadata. The Dolby E system carries metadata in a multiplex with the rate reduced audio, thus providing a way of moving metadata through the Contribution and Distribution links.

Dolby E also caries up to eight channels of audio. Any one of the eight channels, or any combination of up to eight channels can be defined as a program, and thus have a group of metadata parameters associated with it. The most common combination for DTV applications will probably be a group of six channels (for the main 5.1 channel program) and a left, right pair carrying a Dolby Surround (Lt, Rt) signal for the associated analog TV service. In this case, there would be two groups of metadata in the multiplex, or one for each program service.

Most of the metadata carried by the Dolby Digital Emission system is intended to allow individual listeners to tailor the audio presentation to their needs, so is referred to as consumer metadata. The Dolby E system also carries Professional Metadata that can be used by the broadcaster to resynchronize, monitor and modify the level of the decoded audio, again on a program by program basis. The professional metadata is only used in broadcast operations, and is never sent to the home DTV receiver

5.1 Time stamping

Time stamps are an important part of the Dolby E data stream. Many operations require treating the audio and video portions of the program individually; it would be a pity if there were no

[4] The switch point and video frame duration are different for different video systems, so the E type encoder uses the video reference signal ("color black" or its equivalent) of the video system associated with the audio to generate the cock and timing information it needs.

convenient way of reliably laying the audio back, in sync with the video. SMPTE Time Code is fed to the E type encoder and multiplexed into the data stream so that it can be recovered by the decoder. It is intended to be a time stamp, rather than a way of keeping track of time of day, so the recovered time code is identical to the time code that occurred when the audio was being encoded, and does not take into account any encoding or decoding delays. The Drop Frame flag and user bits are also carried.

5.2 Monitoring program signal levels

A common operation, particularly in the Distribution part of the signal chain, is to monitor the level of the program signal. In areas where many programs are present simultaneously, it might be too confusing to reproduce each program from its own set of loudspeakers, or not timely enough to switch one set of speakers between the various programs. Level monitoring can provide some level of confidence that the program material is still present, and can be done for many feeds at the same time without confusing the operator.

It would be a pity to have to decode the Dolby E data stream just to drive a set of meters to indicate that there was some activity in each of the channels of a program group. Part of the professional metadata is metering information. The individual channel signal levels are measured during the encoding process and carried in the professional metadata. Measurements are of the peak and RMS signal levels over the entire block duration (or during one frame period of the associated video reference signal). The amplitude resolution of the measurements is approximately 0.1 dB. There is no attempt to provide a combined signal level for each program group.

5.3 Changing levels

The beauty of digital audio is that levels stay the same. Unlike the analog days when the common carriers guaranteed signal presence but not its level, digital data is not expected to change between the transmission and reception points; the green tweaker that hung on the equipment racks can be retired. But there are still some good reasons for being able to trim a signal level, so the Dolby E professional metadata carries gain words that can instruct a decoder to change the level of a received signal. Each block of data carries two gain words, one applicable to the beginning of the block, and one that applies to the end of the block. If they are different, the decoder interpolates a linear ramp over the duration of the block (or over 1 frame of the associated video) to avoid "zipper noise" as the level changes. The gain range is from +6 dB to minus infinity and is applied to all channels in the program group equally.

The ability to change the level of the recovered program signal is probably more useful as a way of doing fades without loosing a generation than it is as a gain adjustment. As mentioned earlier, the V fade is a very common transition between programs and because it is used on air so often, is usually initiated by a presentation automation system that specifies times and event durations in units of time code. Metadata gain words have the same temporal resolution, so the whole concept fits very easily into normal operational practices. Gain words within a few frames of the end of one program segment instruct the decoder to ramp the level of the program down to silence, the switcher cuts to the next program segment whose gain words cause the decoder to ramp the level back to unity during another few frames.

6.0 The Transport Mechanism

The key to making the E type rate reduction concept practical is to make it easy to integrate with the existing broadcast plant infrastructure. It is a relatively high rate digital signal, so will not integrate easily with an analog plant. As discussed in the Existing Infrastructure section however, DTV requires at least digital VTRs, all of which are currently limited to recording one or two AES/EBU digital audio signal pairs. If at least one layer of AES/EBU signal distribution and routing capability can be added to an existing plant (which may well already have this capability) and if the Dolby E signal can be transported by the AES/EBU mechanism, then that plant immediately becomes capable of doing multichannel programming.

The AES/EBU signal carries two audio sample words, each in its own audio subframe, during each sample period. The two subframes start with a 4 bit Preamble or sync word which is followed by 24 bits of audio data payload space. The subframes end with four additional bits (one for each of the Validity, Channel Status, User and Parity bits) for a total of 64 bits for both subframes. Bit 1 of Byte 0 of the Channel Status information can be set to indicate that the information carried in the audio payload space is not an audio signal. We are thus free to put as much non audio information as we care to in the payload space.

The most sensible choice for the time being would be to use the 20 MSBs[5] of this space for Dolby E data, as this is the maximum number of bits that can be recorded on most studio level digital VTRs. Fortunately this produces a data rate of (20 + 20) * 48 kHz = 1.92 Mb/s which is quite sufficient for Dolby E data. Note that the number of bits used to transport the Dolby E signal has nothing to do with the dynamic range of the audio signal carried by the Dolby E system. The current specification of audio program dynamic range is 110 dB, or the equivalent of 18 bits.

The other advantage of using the AES/EBU signal as a transport mechanism is that the Dolby E signal immediately becomes compatible with the rest of the digital audio equipment in the plant. It can be switched, recorded, edited (cuts or insert and assemble edits) just like any other digital audio signal, as long as some basic precautions are observed. The data must not be changed by any part of the system it passes through as this would destroy the coded audio information (and the metadata). Specifically:
1) Any gain controls must have a unity gain position or bypass function that ensures that the data recorded by or passing through the system is an exact duplicate of the input data.
2) A system must not change the word length of a non-audio signal by truncating it. If the data is being carried in as AES/EBU data stream, the channel status information (byte 2, bits 0 to 5) should be set to indicate the intended word length. The channel status information should either be carried through the system or set appropriately at the system output.
3) Recording systems, switchers, editing systems and similar devices must be able to make butt splices in data streams. The switching points must happen during the vertical interval switching period of the video used as the sync reference to avoid destroying the encoded audio data.
4) Any cross fades, fades to or from silence, sample rate conversions or other process that are intended to modify the data (including rounding or dithering) must be bypassed when handling non-linearly coded data
5) Dolby E rate reduction systems assume an error free channel between the encoder and decoder(s). Any channel coding intended to protect the E type data from error prone channels must be provided by the communications channel in use.

[5] If the entire Contribution / Distribution system was transparent to 24 bits, the increased data rate could be used to improve the mutigeneration performance of Dolby E. 16 bit wide data paths are also acceptable, but may limit the number of channels available, or the number of artifact free generations that can be expected.

7.0 Application of Dolby E type codecs to Program Contribution and Distribution

The previous sections have covered the basic requirements of a rate reduction system intended for Contribution and Distribution applications. Figure 1 shows a few of the areas where the Dolby E codec can be used.

Mobile or Outside Broadcast vans are often used where no leased lines are available and have to rely on a microwave link. Digital radios provide data rates from about 20 to 50 Mb/s, depending on what type of modulation is used. There are claims that "there will be a need for a variety of digital modems and multiplexers" to suit the marry the different video and audio sources to the microwave system in use. Total data rates like these will allow space for approximately 2 Mb/s of Dolby E encoded audio without squeezing the video channel seriously. In some cases, digital radio systems also provide a "wayside T1" connection that provides a 1.544 Mb/s connection that could be used to carry the E type data at 1.536 Mb/s, which is the data rate of the E type information when transported in the 16 MSBs of the AES/EBU stream. This is also the version of the data that will fit on the 16 bit digital audio tracks of some ENG cameras, turning them into multichannel capable devices.

Remote studio operations are generally more entrenched, so would tend to use leased facilities to haul the program signals back to a studio center or a network origination point. In this case, the audio and video will probably share a 45 Mb/s DS3 service. The baseband audio data rates are in the order of 5.5 Mb/s which would take too much of the channel away from the video signal, but at about 2 Mb/s, the Dolby E data can comfortable share the channel with the video signal.

Satellite distribution facilities can also provide data rates from about 45 Mb/s to 60 Mb/s, depending on the transponder bandwidth. The same ideas apply here as for the remote studio situation.

Post production facilities have to find a new release format for finished high definition, multichannel programs. The emergence of mezzanine level video rate reduction codecs allows existing tape formats to carry HD material, but the tape format is unchanged and limited to four audio channels (or two AES/EBU pairs). E type encoders could allow these tapes to carry up to 16 channels of audio and metadata, but in practice, probably only one track pair will be used for E type data. It will carry a 5.1 channel program (with its metadata) and an Lt, Rt Dolby Surround version of the program, with metadata, on the remaining two channels. The second AES/EBU track pair will be used to carry another Lt, Rt version of the program that has been mixed for NTSC release. This is particularly convenient for stations that have no DTV service or digital infrastructure, as they will be able to take the analog output of the "NTSC" track pair and operate as usual. This "multiservice" release format is particularly well suited to advertisements and programs with high production values, as it allows to program producer to tailor the audio to the intended service, but uses only one common tape format to carry them all.

Dolby E thus provides a "point of production to transmitter" path for multichannel audio and he necessary metadata, using the same codecs operating at the same data rates serve all the applications.

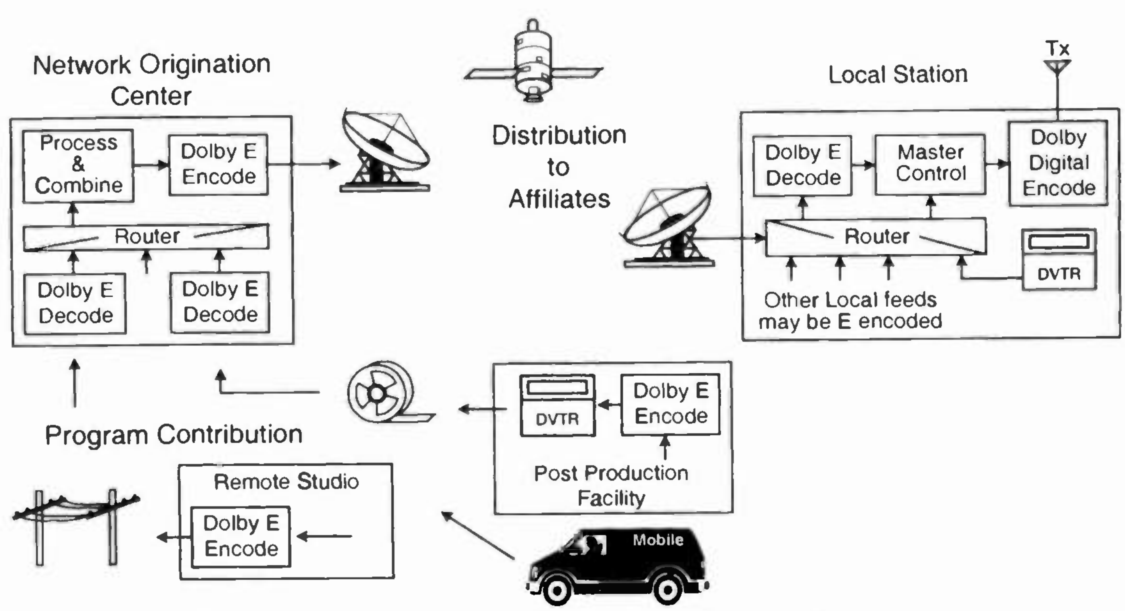


Figure 1 - Application of Dolby E codecs to Program Contribution and Distribution

Technical Regulatory Hot Topics for Broadcasters

Wednesday, April 21, 1999

9:00 am - 12:00 pm

Chairperson: Dane Ericksen, P.E.
Hammett & Edison, Inc., San Francisco, CA

9:00 am FCC RFR Guidelines: Is Your Facility Really in Compliance?
Stanley Salek and Robert Weller
Hammett & Edison, Inc.
San Francisco, CA

***9:30 am Recent FCC Enforcement Activity**
Harold Hallikainen
Hallikainen & Friends
San Luis Obispo, CA

***10:00 am Solutions to Critical Problems Broadcasters are Facing with EAS.**
Richard Rudman
KFWB-AM
Hollywood, CA
Tim McClung
National Weather Service Forecast Office
Oxnard, CA

***10:30 am Electronic Filing of Broadcast Applications**
Barry Umansky
Vorys, Sater, Seymour and Pease
Washington, DC

***11:00 am The Operation of Wireless Microphone Systems in the New RF Environment**
Gary Stansfield, Vega, El Monte, CA; Craig Blakeley, Enterprise Law Group, Vienna, VA; Bob Tamburri, Sony Electronics, Inc., San Jose, CA; Uwe Sattler, Sennheiser Electronic Corp., Old Lyme, CT; Gordon Moore, Lectrosonics, Inc., Rio Rancho, NM; Kevin Mikes, Shure Brothers, Inc., Evanston, IL and Bruce Franca, Federal Communications Commission, Washington, DC

---

*Papers not available at the time of publication

FCC RFR Guidelines: Is Your Facility *Really* in Compliance?

Robert D. Weller, P.E., and Stanley Salek, P.E.
Hammett & Edison, Inc.
Consulting Engineers
San Francisco

ABSTRACT

The FCC adopted new two-tier guidelines limiting human exposure to RF energy, effective October 15, 1997. Over the past one and a half years, some of the implications of the new guidelines have been seen and new policies implemented. Various methods for conducting spatial measurement averaging are considered, along with studied techniques that may be useful in converting previous peak measurement values to average values. Applications for new DTV facilities have triggered accelerated RFR compliance obligations for stations who had certified compliance under the prior guidelines. The FCC has, in some cases, taken a newly active role in the field measurement and verification of compliance with the new guidelines. Further, the FCC has held up the license renewals of some stations that complied with the previous requirements but that now find it difficult or impossible to comply with the present requirements. Potential mitigation measures are presented, based on experiences with recent FCC policies and measurement techniques.

BACKGROUND

RFR Standards History. Anecdotal and pseudo-scientific reports of biological effects of radio frequency radiation were first reported shortly after the commercial development of radio. Outrageous claims of the curative properties of intense exposure made RF energy the "snake oil" of early 20th century. Although many of these claims proved false, RF energy clearly can affect biological tissues and, at sufficiently high exposure levels, the effects may be adverse to human health. Reports of eye damage and other adverse effects resulted in the 1950s in the establishment of the first exposure guidelines to protect humans exposed to RF energy.[1]

As research has identified the conditions under which RF exposure is harmful, these guidelines have been refined. Today, there is broad agreement worldwide among scientists having appropriate expertise that established safety standards are adequate to protect health both in the workplace and among the public. In the U.S., there are two exposure guidelines that are considered prevailing. These are ANSI/IEEE Standard C95.1-1992 and NCRP Report No. 86 (1986). The FCC adopted the latter standard in 1996,[2] concluding its Docket ET 93-62 proceeding to update its former standard.

Docket 93-62 History. Under the National Environmental Policy Act of 1968, the FCC has an obligation to ensure that its actions do not have a significant impact on the human environment. Since 1985, the FCC has had codified regulations to limit the potential for human exposure to RF energy. In 1993, following the adoption by ANSI of IEEE Standard C95.1-1991 (now properly called ANSI/IEEE Standard C95.1-1992), the FCC in ET Docket No. 93-62 proposed adopting the new ANSI/IEEE joint standard. Because compliance difficulties were anticipated with the new standard, the FCC proposal generated a large number of comments from industry, which effectively delayed the proceeding.

Meanwhile, to accommodate substantial growth in cellular telephone subscribership and to permit the competitive deployment of nascent Personal Communications Services, the nationwide wireless infrastructure had to be rapidly enlarged. The proposed construction in residential neighborhoods of many "radio transmission towers" led to heightened public awareness of and concern about RF safety. These concerns frequently led to delays in the construction of new cell sites and caught Congressional attention. In 1996, Congressional pressure to preempt local regulation of RFR resulted in a mandate that the FCC complete work on Docket 93-62 within 120 days, and a new *de facto* nationwide standard was adopted by FCC Report and Order (R&O) on August 1 of that year.

Section 705 preemption. When Congress ordered the FCC to adopt an RF exposure standard, it simultaneously established a limited federal preemption over state and local regulation of RF exposure. The preemption, contained in Section 705 of the Telecommunications Act, is limited, because it applies only to "personal wireless services."[3] Broadcast services were *not* included in the Congressional preemption, meaning that local authorities are free to adopt whatever exposure standards they wish for broadcast sites.[4] The lack of a Federal preemption in this area has already resulted in the required preparation at great expense of a detailed Environmental Impact Report for a site where worst-case power density levels at ground level were calculated to be less than 1% of the applicable public limit.

Although most jurisdictions have been reasonable in their regulation of RF exposures, the presence of multiple standards can sometimes create difficulties. For example, although included in both standards, the FCC does not routinely require evaluations of contact (shock/burn) currents at broadcast sites. It is imprudent, however, to simply ignore this provision of the standards. Apart from this evaluation possibly being required by a local ordinance, intentional disregard of an established safety standard might lead to a successful tort lawsuit being brought by an allegedly injured party.

More Surprises. The R&O contained a number of surprises, not the least of which was the adoption of the guidelines published as Report No. 86 by the National Council on Radiation Protection and Measurements (NCRP), rather than the more recent ANSI standard. Following the release of follow-up orders, it was decided that compliance with the NCRP guidelines for new stations would be required effective October 15, 1997, with licensees filing for new facilities, renewals, and modifications having to "bring their facilities into compliance,"[5] and *all* existing facilities having to comply with the new guidelines no later than September 1, 2000. Based upon these statements, many broadcasters who had renewed their licenses just prior to the adoption of the new guidelines thought they had until their next renewal (or September 1, 2000) to bring their facilities into compliance with the new guidelines. Largely unnoticed was a statement in the R&O "... that if a transmitter at a multiple-transmitter site is approved under one set of guidelines but, later, another transmitter locates at the site and, as is required, operates under the new exposure criteria, then the new criteria must be used to evaluate the entire site."[6] With this statement and the pending authorization of some 1,600 Digital Television (DTV) stations, the FCC created a time bomb.

Exempted versus excluded. When a broadcast station applies to modify its transmitting facilities or renews its license, it must certify that its operations are in compliance with the prevailing standard that the FCC has adopted. Note that the obligation to comply always exists; the periodic requirement to certify compliance does not relieve a station from the obligation to continuously comply. Similarly, although a station may fall under one of the "categorical exclusion" provisions contained in the FCC rules, the obligation to comply remains. An *exclusion* from the requirement to routinely demonstrate compliance does not imply an *exemption* from compliance.

SPATIAL AVERAGING

Traditionally, RF exposure conditions at broadcast sites have been reported using spatial peak values. That is, for a given point on the ground, the measurement probe would be moved vertically up to a height of perhaps 2 meters, with the maximum value encountered over this range being recorded. While the prevailing standards specify both whole-body and partial-body exposure limits, the ANSI/IEEE standard has been unclear, at least with regard to reliance upon whole-body spatial averaging to achieve compliance at broadcast sites. The unclear language was identified as early as 1993.[7] There being only cumbersome equipment available at the time to conduct spatially averaged measurements, the several attempts at clarification failed due in large part to lack of interest. Finally, in response to a 1998 petition from Hammett & Edison, the applicable IEEE Standards Coordinating Committee (SCC-28) Interpretations Working Group released a finding "... that C95.1-1991 requires the spatial averaging of measured uniform or non-uniform fields..."[8]

Description of techniques. While both standards now clearly require assessment of "whole body SAR," difficulties have been encountered in defining just what constitutes a "body." An adult male, a child in a wheelchair, a fetus? Because the underlying data upon which the standards are based include a variety of human body types (not limited to the 60 kg, 1.75 m "standard man"), there is some inherent safety factor built in. Measurement techniques based upon the "standard man" are, however, consistent with past recommendations of SCC-28.[9] Compliance with the field limits is based strictly on a "go/no go" criteria, so the variation encountered in the different techniques becomes important. Attempts at determining blanket compliance with the standards in a given area are fully successful only if there is a prior knowledge of the "bodies" that will be in that area. Little has been published on the subject of spatial averaging.[10,11] No U.S. regulatory or standards-setting body has endorsed a particular measurement technique for spatial averaging, although work on this topic is presently underway within IEEE. Discussions with persons who routinely conduct such measurements suggest that two different techniques are presently in common use at broadcast sites.

Vertical line method. In this method, depicted in Figure 1.a), the probe is swept with uniform velocity from the ground to a height of about 1.8 meters (6 feet), with the average power density over the line being calculated automatically by the instrument. The authors' firm has used a PVC pole with base to ensure that measurements are taken along a vertical line and to ensure that repeated measurements are at the same location.

Planar equivalent methods. In this method, the probe scans a planar area approximating the adult trunk, as shown in Figure 1.b). This method, which is defined in Canadian Safety Code 6 (SC-6),[12] was originally designed for use with meters having no automated averaging capability. A total of nine discrete measurements are made over a 1.0 by 0.25-meter rectangle as shown in Figure 1.c), with the result being the average value. With the availability of commercial equipment having automated spatial averaging capabilities, Hammett & Edison has explored several possible means of adapting the SC-6 method to current measurement technologies. All of these methods involve continuous collection of data within the measurement rectangle. Two of these alternative methods are called here the "zig-zag" and "two-pass" methods and are depicted graphically in Figures 1.d) and 1.e), respectively.

Probe orientation. Regardless of the method used, it is recommended that the measurement probe be placed at an angle of about 90° relative to the dominant RF source, as shown in Figure 2. This orientation minimizes the effect of the operator's body on the measurement, and typically provides the most conservative (highest) reading. As is seen in Table 1, readings can vary by almost 2 dB, depending upon the orienta-

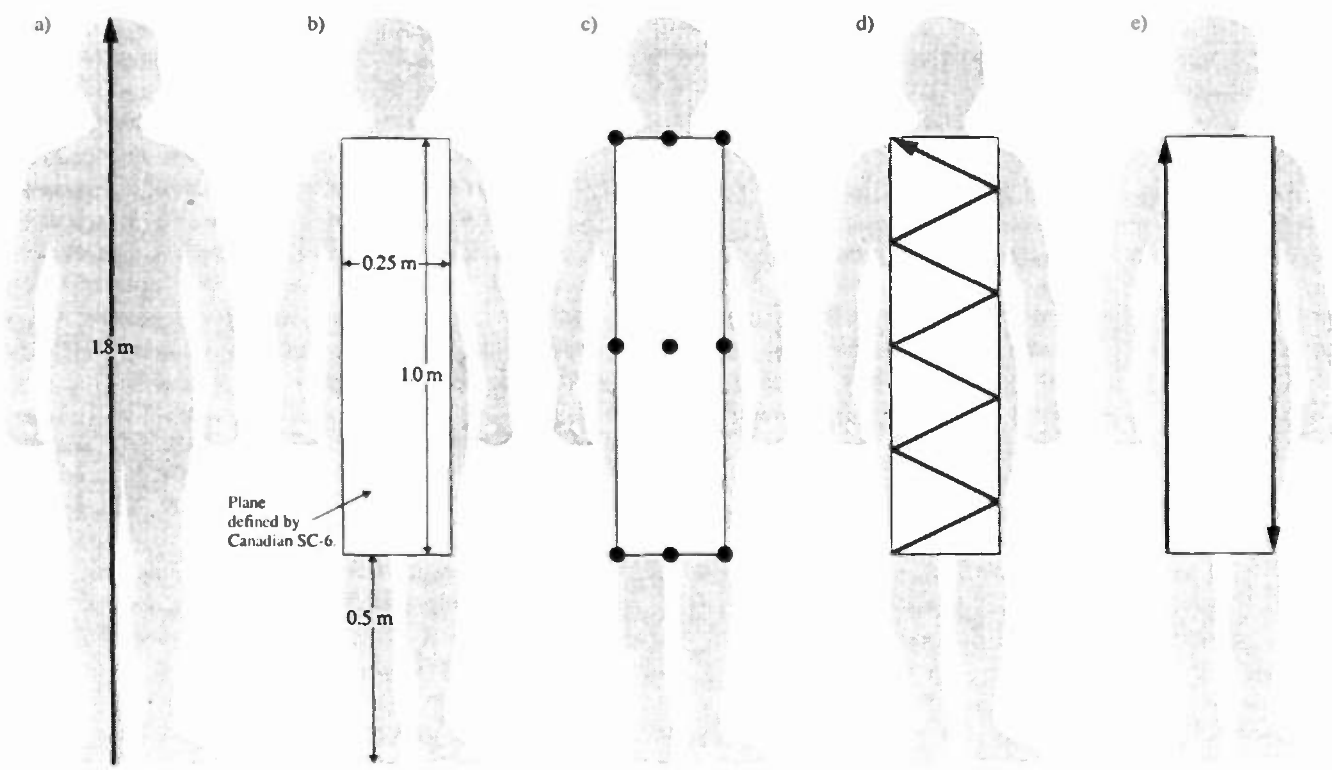


Figure 1. *The commonly used spatial averaging techniques are divided into two types as discussed in the text. The Vertical Line technique a) involves sweeping the probe from near the ground up to some height, typically 1.8 meters. The planar equivalent techniques involve moving the probe over an area approximating the trunk of the "standard man," as shown in b). Four different techniques (a, c, d, and e) were evaluated.*

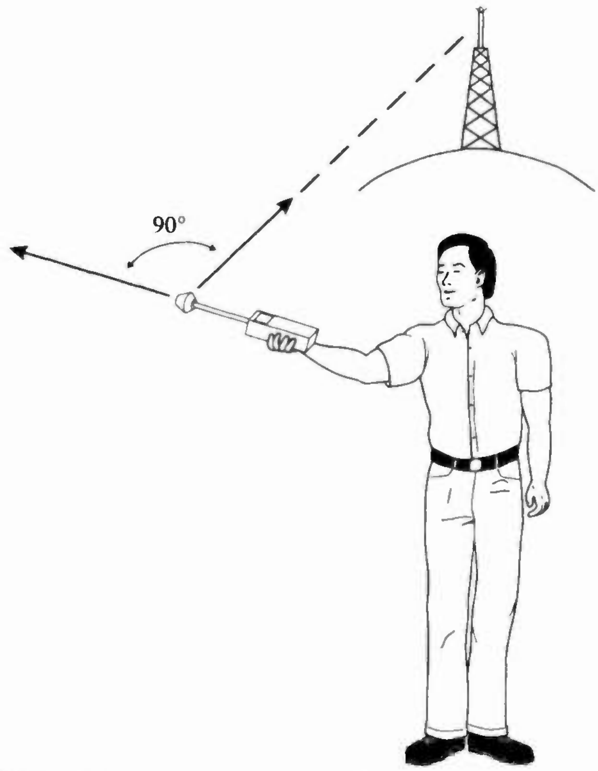


Figure 2. *Proper orientation of the measurement probe.*

Point	Measured Power density (% of limit) Orientation with Respect to Source 0°	90°	180°
1	10.7%	14.0%	10.6%
2	11.1	15.8	11.5
3	9.3	11.4	7.5
4	5.2	7.4	5.9
5	9.7	14.7	9.6
6	8.6	11.3	8.4
7	14.5	19.2	—
8	14.3	18.6	—
9	10.7	16.3	—
10	7.2	9.5	—
11	13.5	19.9	—
12	11.0	15.3	—
Average Difference from 90°	-1.4 dB	0.0 dB	-1.4 dB

Table 1. *Variation with probe orientation of spatially averaged measurements. To ensure a conservative measurement, the probe should be oriented approximately 90° with respect to the dominant RF source. This orientation is contrary to some manufacturer's recommendations. Due to interaction with the operator and instrument, however, other orientations may produce results that are too low.*

tion of the probe and operator with respect to the source. If there is no dominant source, or if the location of the dominant source is uncertain, it is recommended that four readings be taken at each point with the operator moving around the point 90° between each sweep; the highest of the four readings would be recorded for that location.

Comparison of methods. Table 2 compares statistically the vertical line and three planar equivalent methods. The data suggest that all of the planar equivalent methods produce similar results, which are all conservative (higher) compared to the vertical line method. Because of the small number of points studied (N=8), the data do not support a clear conclusion that any of the three planar equivalent methods is superior to another, but certain methods are inherently more practical. For example, the 9-point SC-6 method requires considerably more time than either of the other two.

Method	Average Difference with Respect to 2-Pass Method	Standard Deviation
Vertical Line	-0.9 dB	0.7 dB
9-Point	-0.1 dB	—
Zig-zag	0.2 dB	0.6 dB
Two-Pass	0.0 dB	0.5 dB

Table 2. *Comparison of four different spatial averaging techniques. Although the small number of points considered (N=8) does not support a conclusion that any method is clearly superior, it appears that the vertical line method gives results that are low compared to the other (planar-equivalent) methods. The standard deviations, a measure of the repeatability of several measurements at the same point, are similar for all of the methods.*

Spatial average-to-peak comparisons. Because measurements of RF exposure conditions at broadcast sites historically have been based upon spatial peak readings, it is of interest to know how such peak data might compare with spatial average readings. Although the spatial average/peak ratio is theoretically dependent upon several factors (see *e.g.*, Effect of Ground discussion below), considerable measurement data has been collected and analyzed by the authors to determine whether a "rule of thumb" conversion factor value might be applicable. As shown in Figure 3, the distribution of ratios follows an almost linear distribution over a range of 0.4–0.9 and has a mean value of 0.6. While this suggests that a value of 0.6 could be applied as an estimate, it does not eliminate the requirement to measure sites that are calculated to exceed the applicable limit using standard spatial peak calculation methods (*i.e.*, OET-65).

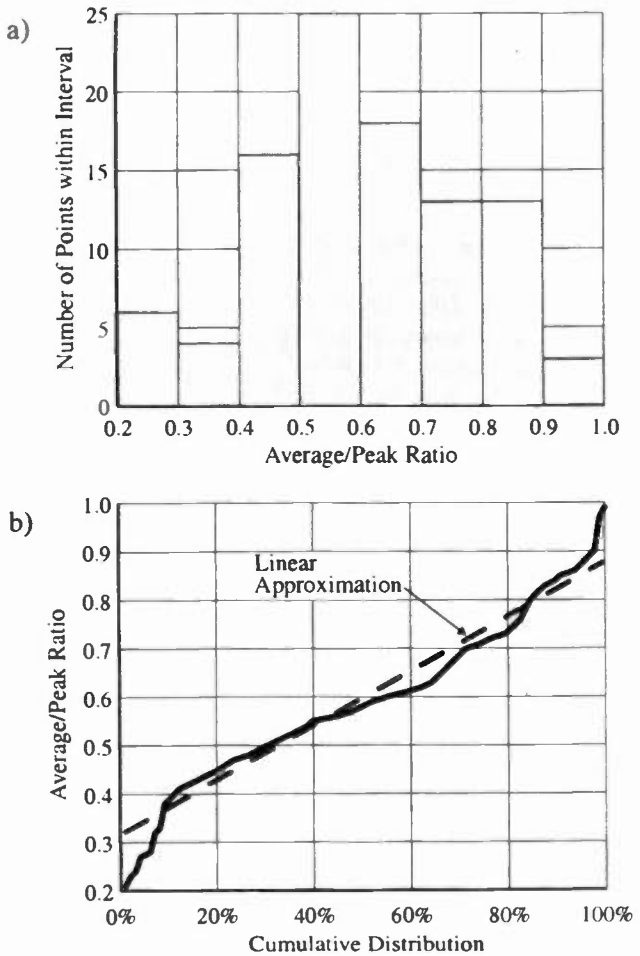


Figure 3. *The histogram a) and cumulative distribution function b) show the distribution of measured spatial average/peak ratios for a number of points (N=98). The mean and median values are 0.6, but the data are almost linearly distributed over the range 0.4–0.9. While the mean value may be used for estimating purposes, sites that calculate close to the limit using standard techniques would still have to be surveyed.*

ACCURACY ISSUES

Unlike many physical quantities that can be measured with great accuracy, the uncertainty associated with the measurement of RF fields at broadcast sites is relatively high. This uncertainty arises from meter and probe limitations, site variability, and interaction of the electromagnetic fields with the instruments and operator. Because uncertainties may be on the high or low side, because of the large safety factors included in the standards themselves, because most RF "hot spots" are small in extent, and because of site variability, it is customary not to apply calculated uncertainties to RFR measurements. Despite this, a recognition of measurement limitations is important to the analysis.

Instrument accuracy. In comparing specifications for RF survey meters, it is easy to become "spec happy." While one may be tempted to simply add up all of the uncertainties listed in the specifications, doing so is not representative of typical operating conditions. Overall accuracies may be analyzed in two ways: worst case and RSS. The worst case uncertainty comes about if all the possible sources of error were at their extreme values and in such a direction as to add together constructively. It is much more realistic to combine the uncertainties using the root-sum-of-the-squares (RSS) method. The RSS uncertainty is based on the fact that most of the measurement errors, although systematic and not random, are independent of each other. Since they are independent, they are random with respect to each other and combine like random variables:

$$RSS = \sqrt{(U_1^2 + U_2^2 + U_3^2 + \ldots)}$$

where each U_i represents an uncertainty (expressed as a fraction or percentage). The worst case and RSS values for several popular instruments are shown in Table 3.

One means of increasing confidence in RFR measurements is the use of several meters (preferably of different manufacture and technology) to survey each

Meter Probe	Older analog (diode detector)	Digital #1A (diode/shaped)	Digital #1B (diode/shaped)	Digital #2A (diode detector)	Digital #2B (diode/shaped)
Worst case max	1.9 dB	2.8 dB	2.6 dB	3.5 dB	3.5 dB
Worst case min	-2.7	-4.6	-4.3	-7.9	-8.8
RSS max	1.2	2.1	1.8	2.5	2.4
RSS min	-1.3	-2.3	-2.0	-2.9	-2.9

Table 3. *Calculated worst case and RSS uncertainties for several popular RFR survey instruments. When comparing specifications, the RSS uncertainty is a better indicator of overall performance.*

location. The readings from the several meters can then be averaged together (a technique called *ensemble averaging*). Since it is unlikely that several meters would have the same accuracy profile for a given measurement situation, the uncertainty associated with the average of several meters' readings will generally be less than the uncertainty of any individual meter. Good agreement between several meters is a strong indication of an accurate measurement.

Most RF exposure measurement equipment is based upon traditional power meter designs, with an antenna substituted for the transmission line input. Consequently, diodes and thermocouples are the most common types of detectors used.

Diode detector issues. Detectors using metal-barrier or Schottky diodes are perhaps the most common type used at broadcast sites. They offer broad frequency response (0.1–4,000 MHz or greater), flat frequency response (±1 dB or better), and large dynamic range (1–1,000 V/m or about 0.1%–5,000% of the standards). Diodes do have a major limitation when used at typical broadcast sites: often, they do not respond properly in a multiple station environment. Multiple source and frequency (MSF) errors have been examined in some detail both theoretically[13] and empirically.[14] If a diode-type detector is not operating in the square-law region, the measurements will not be valid. The response will be the result of squaring the sums of the voltages rather than the summing of the squares. Although MSF errors can theoretically be positive or negative, *ad hoc* tests conducted by the authors appear to confirm earlier findings that the errors are always positive (*i.e.*, at multi-user sites, the measured results may be too high). Our results showed errors of 0–3 dB, with an average error of about 1.2 dB. Meters that include circuitry to compensate for this effect appear to be less susceptible to MSF errors.

Site accuracy issues. The question is sometimes asked how repeatable RF exposure measurements are at a given site. Unfortunately, RFR compliance measurements are usually conducted only once at a given site; additional measurements are conducted only if there is a change (such as an antenna replacement or when a new station is added). It is therefore difficult to calculate the repeatability of such measurements. For a variety of reasons, the electromagnetic compatibility (EMC) industry has historically been concerned about measurement repeatability and has published data that may be of use in estimating the repeatability of RFR measurements at broadcast sites.

For example, Kolb[15] reports that for five sites, the typical standard deviation at one site was 0.6 dB. Assuming a normal distribution, this would mean that 95% of the time readings would vary less than ±1.2 dB. In calculating this value, measurements were conducted five times at each site, usually on different days. EMC measurements are conducted using measurement equipment that is located with a fixed spatial relationship to the source. RFR measurement protocols call for spatial averaging, which tends to reduce variation at specific points. So day-to-day variation at broadcast sites could be expected to be somewhat better than ±1.2 dB.

Effect of ground. A single reflection from a perfectly conducting earth can double the electric field strength at certain locations in space, thus quadrupling the power density. However, the earth is not perfectly conducting, and the efficiency with which it reflects radio waves, called the coefficient of reflection, ρ_r, is always less than 1. It may be calculated[16] from knowledge of the grazing angle (the angle between flat earth and the antenna), ground conductivity, and dielectric constants of the site. In its Bulletin OET-65, the FCC recommends that power density calculations near broadcast sites assume $\rho_r = 0.6$.

Considerable field experience has shown the recommended value for ρ_r, which increases the power density by a factor of 2.56 above free space, to be conservative over a wide variety of ground types and measurement heights. This result is expected, because at the typically large grazing angles (>30°) encountered at broadcast sites, ρr would theoretically be expected to lie in the range 0.2–0.85, depending upon conductivity and dielectric constant, for horizontally polarized waves and 0.05–0.85 for vertically polarized waves.[17,18] Examination of the mean values of these two ranges suggests that the 0.6 assumption lies on the conservative (high) side. Furthermore, at locations near the tower on which a source is mounted, the reflected wave is typically dominated by a single polarization. So the common practice of multiplying *both* the horizontally and vertically polarized components simultaneously makes this calculation additionally conservative.

FCC COMPLIANCE ISSUES

The authors' firm has studied a number of broadcast sites that comply under the old FCC rules but that fail to meet the revised rules. The addition of public exposure limits that are generally five times more restrictive than occupational limits can create areas of noncompliance that were formerly well within the limits of the previous requirements.

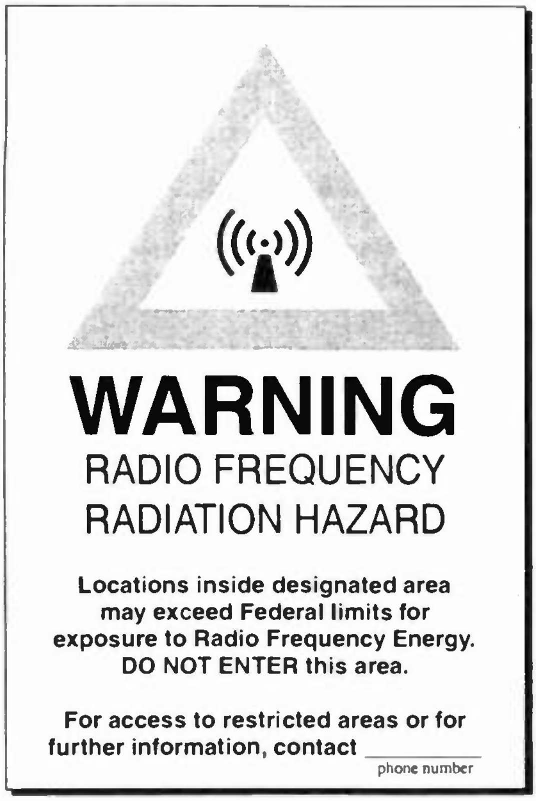


Figure 4. *Example of a warning sign appropriate for OSHA-mandated (controlled/occupational) exposure situations. Such signs should be placed on towers and at the entrances to controlled-access broadcast sites, as discussed in the text.*

Within fenced compounds. Areas within the fenced perimeter of a transmitting site usually would be considered subject to the occupational exposure limits, since only persons authorized for access within the fence would be present in those areas. One or more signs, as shown by the example of Figure 4, should be posted at entry points and other approaches to the fenced area.[19,20] OSHA-mandated warning signs should include four pieces of information:[21,22]

- The universal warning symbol
- The words "Radio Frequency Radiation Hazard"
- Specific site information describing the location of the hazardous area and how to avoid it
- A telephone number for further information.

Such signs provide some measure of assurance that persons unfamiliar with the site may seek guidance about which areas within the fence, if any, should be avoided with respect to prolonged exposure. Except for the differences noted earlier, the requirements for protection of workers and other authorized persons in a compound having RF emitters is essentially unchanged from the earlier FCC policies.

An important issue related to occupational exposure is on-tower access. The FCC generally requires that licensees at a common site cooperate such that occupational exposure limits are not exceeded for access on any tower at the site. For some sites, compliance will require systematic power cutbacks by stations to allow a worker to access a given transmitting tower. Sites should have in place[23] an "occupational exposure guide" (OEG) that specifies each station's responsibilities for access to any part of the site, including towers. It is important that such guides be updated regularly to account for changes and additions to transmitting facilities.

Outside fenced compounds. The vast majority of compliance difficulties at existing sites are related to exceeding continuous public exposure limits outside a site fence, *i.e.*, in publicly accessible areas. Before adoption of the FCC Docket 93-62 rules, there was no distinction between occupational and public exposure requirements. For FM and/or TV/DTV transmitter sites, allowable continuous public exposure is five times more restrictive than for occupational exposures. Especially for mountaintop sites having high-power transmitting antennas mounted on relatively short towers, measured exposure levels outside the typical site fence may approach occupational levels. In some cases, the FCC has required licensees to reduce power, such that measured exposure levels can be kept below public limits in public areas.

At multi-user sites, it is not uncommon to have one or two "culprit" stations whose contributions account for the majority of the excessive fields measured in public areas. It is important to note that the FCC generally requires all stations identified as significant contributors,[24] either by calculation or by measurement, to cooperate in formulating a solution to the identified problem. The grant of an FCC license renewal for one contributing facility does not absolve it from responsibility. The FCC has, in the past, threatened revocation of licenses for quarreling contributors, but, to date, no actual actions of that type are believed to have been taken.

Solutions that may be employed to remedy excessive exposure in public areas could include one or several of the following:

- Expansion of the site fence to encompass areas exceeding public exposure limits
- Individually fencing areas exceeding public exposure limits
- Changing one or more transmitting antennas, usually FM broadcast transmitting antennas, to

types exhibiting lesser radiation at greater depression angles

- Relocating transmitting antennas of the "culprit" stations further inside the compound and/or increasing their radiation center height(s)
- Consolidating multiple stations onto one or two transmitting antennas that are located at greater height and/or use low-RFR antenna technologies
- For larger sites, relocating one or more contributing stations to other towers well away (several hundred meters or more) from the existing installation(s)
- Relocating stations to entirely different transmitting sites.

It is noted that all of these methods employ techniques that result in no need for direct involvement of an unsuspecting member of the general public. That is, it may not be assumed that the public can read and interpret a posted warning sign, and it should not be assumed that a readily accessible area exceeding the public limit is in compliance because it is in a usual "transitory location" where prolonged exposure would not be anticipated. That is, time averaging generally should not be relied upon to achieve compliance in publicly-accessible areas. Positive means provide the best assurance in preventing exposures exceeding public limits.

DTV facility construction. As mentioned previously, another aspect that has triggered the need for formerly compliant stations to take a proactive stance in RF radiation compliance involves the implementation of DTV facilities. Multi-user transmitting sites, especially those with several existing NTSC TV stations, likely will be subject to considerable change in RF radiation exposure characteristics. This will trigger, in numerous cases, an FCC requirement that existing stations make changes to reduce RF radiation characteristics in both occupational and public areas.

CONCLUSION

The key to avoiding possible problems related to RF radiation exposure at a given transmitting site is research, planning, and assumption of a proactive stance. Existing sites, if not already evaluated for compliance with present FCC requirements, should be studied without delay. Any proposed changes to an existing site should be carefully reviewed with respect to potential RF radiation exposure scenarios before project work begins. Any deficiencies noted at existing or planned facilities should be immediately addressed and corrected. Finally, spatially-averaged measurements used to establish compliance with the prevailing standards should use one of the planar equivalent methods.

REFERENCES

1 William F. Hammett, *Radio Frequency Radiation* (New York: McGraw-Hill, 1997), p. 45ff.

2 Report and Order, ET Docket 93-62, released August 1, 1996, FCC 96-326, ¶28.

3 Personal wireless services are licensed under FCC Rule Parts 22 (cellular and paging), 24 (PCS), and 90 (SMR).

4 See *e.g.*, Gray Frierson Haertig, "Compliance with Local Radio Frequency Radiation Regulations," *Proc. 47th Annual Broadcast Engineering Conference*, (Washington, DC: NAB, 1993), pp. 487-498.

5 Second Memorandum Opinion and Order, ET Docket 93-62, released August 25, 1997, FCC 97-303, ¶113.

6 *Ibid.* ¶75.

7 Letter of Jules Cohen to Eleanor R. Adair and Om P. Gandhi, May 8, 1993.

8 Letter of James B. Hatfield, Chairman, SC-4, SCC-28, IWG to Dane E. Ericksen, Hammett & Edison, October 30, 1998.

9 See *e.g.*, Letter of Om. P. Gandhi, Co-Chair, SCC-28 to Linda A. Gargiulo, March 18, 1992.

10 NCRP Report No. 119, *A Practical Guide to the Determination of Human Exposure to Radiofrequency Fields*, (Bethesda, MD: NCRP, 1993), p. 83.

11 See *e.g.*, Richard A. Tell, "Real-time averaging for determining human exposure," *Proc. 40th Annual Broadcast Engineering Conference*, (Washington, DC: NAB, 1986), pp. 388-394.

12 Health Canada, Safety Code 6, *Limits of Exposure to Radiofrequency Fields at Frequencies from 10 kHz–300 GHz*, (Ottawa: Canada Communication Group, 1993).

13 J. Randa and Motohisa Kanda, "Multiple-Source, Multiple-Frequency Error of an Electric Field Meter," *IEEE Trans. on Antennas and Propagation*, Vol. AP-33, No. 1 (Jan. 1985), pp. 2-9.

14 R.R. Bowman, "Some recent developments in the characterization and measurement of hazardous electromagnetic fields," cited in NCRP Report No. 119, *Op. cit.*, p. 201.

15 Lowell E. Kolb, "Statistical Comparison of Site-to-Site Measurement Reproducibility," *Proc. IEEE 1996 International Symposium on Electromagnetic Compatibility*, (Piscataway, NJ: IEEE, 1996), pp. 241-244.

[16] Ramo, Whinnery, and Van Duzer, *Fields and Waves in Communication Electronics*, (New York: John Wiley & Sons, 1965), pp. 344-347.

[17] Shigedazu Shibuya, *A Basic Atlas of Radio-wave Propagation*, (New York: John Wiley & Sons, 1987), pp. 166-167.

[18] US Environmental Protection Agency, *An Engineering Assessment of the Potential Impact of Federal Radiation Protection Guidance on the AM, FM, and TV Broadcast Services*, (n.p.: April 1985).

[19] FCC Bulletin OET-69, Appendix B.

[20] 29 CFR §1910.268(p).

[21] 29 CFR §1910.97

[22] Robert Curtis, OSHA, Personal Communication.

[23] 29 CFR §1910.147

[24] A significant contributor is generally defined as any station whose individual contribution at the point in question exceeds 5% of the applicable limit. See 47 CFR §1.1307(b)(3).

Digital Audio for Broadcast Engineers

Wednesday, April 21, 1999

9:00 am - 4:00 pm

Chairperson: Richard Farquhar
RAF Associates Inc., Canal Winchester, OH

***9:00 am The 2000 Engineer**
Paul McLane
Radio World
Falls Church, VA

9:15 am AES Audio for Broadcast Systems
Gary Stephens
Leitch Incorporated
Chesapeake, VA

10:30 am Computers and Radio, The Networking Connection
Chriss Scherer
BE Radio magazine
Overland Park, KS

***2:00 pm Technology in Transition**
Chriss Scherer
BE Radio magazine
Overland Park, KS

2:15 pm Building a Digital Station
David Baden
Radio Free Asia
Washington, DC

*Papers not available at the time of publication

AES AUDIO FOR BROADCAST SYSTEMS

Gary Stephens
Leitch Incorporated
Chesapeake, VA

Abstract

Digital audio offers the broadcaster many advantages over analog audio. With digital audio we can avoid problems associated with analog audio such as hum, noise, level shifts and other effects of an analog distribution system. We also have the ability to create the perfect distribution system, creating perfect copies of our work without generation loss. This paper will cover highlights of the AES/EBU audio system from generation to distribution of the signal.

What Is AES Audio?

Digital audio, also know as AES/EBU audio, is jointly defined by the Audio Engineering Society (A.E.S.) and the European Broadcasting Union (E.B.U.) The AES/EBU serial digital audio format is a method of transmitting two channels of linearly coded digital audio, which may be stereo or two completely different mono programs multiplexed in a single serial digital signal. The format, specified in AES3-1992, accommodates digital audio word lengths of up to 24 bits. While AES3 specifies no sample rate, a separate document, AES5, recommends a sample rate of 48 KHz for broadcast and recognizes the compact disc sample rate of 44.1 KHz.

AES Frame Structure

The AES frame is made up of two AES subframes. These subframes are paired as either Channel A/Channel B or Left/Right. Each subframe carries one audio sample and contains 32 bits of information. Out of these 32 bits, 4 bits are used for the preamble, 4 bits for auxiliary data and 20 bits for the audio sample. The remaining 4 bits are the V,U,C and P bits.

The 4 bit preamble is used to synchronize the AES receiver to the AES source. It also differentiates the Channel A subframe from the Channel B subframe.

The auxiliary data area is made up of 4 bits. This user-defined area can be used to carry data (such as an extra voice channel) or to extend the audio sample to 24 bits.

The next section of the AES subframe is the actual audio sample. 20 bits are used for the audio sample.

The V bit is the validity bit. If we have a valid sample, this bit is a 1. An invalid sample is represented by a 0.

The U bit is the user bit. It can be used for data. The U bit of several subframes can collectively carry low data rate messages or data.

The C bit is also known as the audio channel status bit. Here, all of the C bits from one AES block (192 frames) are accumulated and grouped into 24 status bytes. These bytes are used to convey a variety of information about the audio channel, including sample rate, channel mode, consumer vs. professional use, and emphasis on/off.

The final bit is the P bit or parity bit. It is set to provide even parity for the subframe and provides a level of error detection.

Finally, 192 frames are combined to form an AES block. If we assume a 48 KHz sampling rate, each block represents 4 mS of dual-channel audio.

AES Sampling Rates

The AES standard supports sampling rates from 32 KHz To 50 KHz. The 32 KHz rate is typically used by NICAM and DAB. Consumer audio products and CD players are at 44.1 KHz, and 48 KHz is the sampling rate of importance to the broadcast engineer. 48 KHz was selected because it yields integer number of audio samples to an integer number of video frames, allowing audio and video clocks to be locked together so we can obtain deterministic distribution of audio samples.

Calculation of the data rate for the AES signal is relatively straightforward. It is the samples per second times the number of bits per sample times two channels. This yields the following for a 48 KHz sampling rate:

$$48{,}000\ \frac{\text{samples}}{\text{second}} \times 32\ \frac{\text{bits}}{\text{sample}} \times 2\ \text{channels} = 3.07\ \frac{\text{Mbits}}{\text{second}}$$

The AES standard supports audio word lengths of up to 24 bits. The number of bits per sample is determined by A to D and D to A converters. The routing and distribution system of a broadcast facility will pass the data stream without regards to the number of bits present in the audio word. While audio word rates of 18 bits are supported, the trend is to 20- and 24-bit audio words. 24 bits is normally used for high-end applications such as recording and mastering, and 20-bit audio words are dominant in TV environments.

What are some of the advantages and disadvantages of using more bits in the audio word? The additional bits provide more resolution to each sample, improve the dynamic range and the signal to noise ratio. The down side to higher bit rates is that the equipment becomes more expensive. Depending upon the application, the higher bit rate may not be warranted due to the increased cost versus performance trade off.

Nominal Operating Level And FSD

In analog devices we were concerned with a nominal operating range of the audio signal plus some additional head room for the peak signal. The additional head room keeps the audio signal from clipping. These same concerns are also valid in digital audio due to the nature of A to D and D to A converters. In the digital domain, we try to select a nominal operating level for the signal and set it for a much lower value than our Full Scale Digital (FSD) level set on the A to D converter.

FSD is the level of the analog signal that will result in the largest possible digital value; i.e., a signal of all *1s*. It is also refered to as 0 dBFS. A similar concept in the analog domain would be a peak-to-peak signal just below the clipping threshold. If the analog input level is higher than the FSD set on the A to D converter, then the A to D converter will not accurately represent the signal. A to D and D to A converters typically include an adjustment for determining the FSD used in the conversion.

Analog audio levels are often specified relative to the FSD level set on the A to D converter. If we were to set the FSD on the A to D converter to 24 dBu, then the following relationship would apply:

dBu	dBFS
24 dBu	0 dBFS
8 dBu	-16 dBFS
4 dBu	-20 dBFS

Whatever value a broadcast facility selects for its FSD level is a matter of preference. The important thing to keep in mind is that all devices in the facility should be set to the same FSD level. If we have mismatched FSD levels within a facility or material from another facility using a different FSD level, then loss or gain through the D to A conversion is possible.

Distribution Options

We have two options to chose from for the distribution of the AES signal. One is sending the AES stream over a 110-Ohm balanced twisted pair interface. The second method utilizes the unbalanced approach of 75-Ohm coaxial cable. Both types have their advantages, disadvantages and markets.

110-Ohm balanced AES is defined by the specification AES3-1992. It is a 110 Ohm balanced interface using standard shielded twisted pair audio cable. The XLR connector is specified, but many routing switchers utilize DB 25 or other high-density connectors. The voltage level is typically 3 to 7 volts. 110-Ohm balanced is commonly used in audio only installations and many audio for video installations in Europe.

75-Ohm unbalanced AES is defined by the SMPTE specification 276 and the AES-3id-1995 specification. It is a 75-Ohm unbalanced interface using 75-Ohm coaxial cable. The BNC connector is the choice for this application. The voltage level is specified at 1 Volt peak-to-peak maximum. The higher cost of the coax cable is offset by the lower costs of the BNC connectors and their installation. 75 Ohm unbalanced is commonly used by North American broadcasters and is just now gaining acceptance in Europe.

The conversion from coax AES to balanced AES and vice versa is accomplished using passive transformer devices. These feature an XLR connector on one end and a BNC connector on the other end. These devices are are available for about $50 each from several manufacturers.

Synchronous vs. Asynchronous Routing

When it comes to routing the AES signal around the broadcast facility, we have to consider two types of routers. These are the synchronous and asynchronous routers. Each has its place in the broadcast environmen, and both have their advantages and disadvantages.

The simplest form of router is the asynchronous router. It is very cost effective and works with both balanced and coax AES. It does not do any reclocking or reframing and works with any data rate. It is the work horse for a broadcast facility in getting a signal that will not be switched live on the air from point A to point B. This type of router is not concerned with precision switching of the data packs and may cause pops and clicks during the switching transition.

The synchronous switcher is designed for switching applications where prevention of data corruption is important or where the signal is being switched live over the air. This router also works with both balanced 110 Ohm systems and AES coax systems. Each input of the router is reclocked and retimed so that switching occurs on frame boundaries, reducing the occurrence of pops and clicks. Because this is a synchronous router, it will work at only a single data rate. Typically synchronous routers are about 50% more expensive than asynchronous routers.

Reference Generators

There are three types of reference signals commonly used to synchronize the AES router. The most common is an actual AES audio signal carrying program samples or silence. This signal is often called DARS (Digital Audio Reference Signal). Another type of reference signal is the TTL level word clock. It is a square wave pulse at the sampling frequency. The least common type of reference signal is an analog video signal. This signal is normally used in television-related broadcast facilities where the video switching must also be synchronous with the audio switching. In this application, the AES device uses a multiple of the video frequency to produce an audio sampling clock.

Audio to video relationships must also be considered when switching both signals together. In the 525/60 video format we find that the audio samples line up with the video signal every 5 frames. This gives us 8008 audio samples for every 5 video frames. The 625/50 video signal has a much cleaner arrangement with the EBU digital audio signal. Here the audio and video samples line up every frame, producing 1,920 audio samples per frame.

The importance of synchronization cannot be understated. For proper operation all AES sources must be synchronized in both sampling and phase. Failure to lock AES samples will result in intermittent pops and cracks as sample values are corrupted. Failure to lock on the proper frame boundary can result in analog phase errors when converting back to an analog format.

A/D and D/A Converters

We are living in a digital world but our senses are still analog devices. The importance of selecting the proper A to D and D to A converter for our facility is very critical. A good converter should offer switch-selectable sampling rates, at least 20-bit conversion, adjustments for FSD levels and an external reference input to synchronize its output to the system. Critical specifications to be considered are frequency response, THD + N, signal-to-noise ratio, and interchannel cross talk.

Distribution Amplifiers

AES provides the operator with three types of distribution amplifiers to choose from. These amplifiers include fan-out AES DAs with or without equalization and reclocking, standard analog video DAs or AES specific digital DAs when using coax.

Fan out DAs provide the function of creating multiple copies of the signal. Limited looping of the signal is possible but sometimes awkward due to bridging problems with the audio cable. For cable lengths of greater than 1,000 feet, a DA with automatic equalization and reclocking is required.

Analog video DAs also make an excellent choice for distributing the AES signal over coax cable. In this type of DA, the AES data stream is treated like an analog signal. Limited equalization is possible, but reclocking of the signal is not performed. The bandwidth of the DA must be greater than 11 MHz to prevent slew rate from distorting the edges.

The final choice for a DA is an AES-specific digital DA for distributing the signal over coax. These DAs typically provide differential inputs, automatic equalization and reclocking, and jumpers for selecting approximate cable length. The automatic equalization circuit restores the signal to its proper level and fills out its shape of data pulses. The reclocking circuitry is composed of slicer sampler and a phase-locked loop. The result is a jitter-free signal that is a perfect copy of the original signal.

Conclusion

Although digital audio has been available for some time, it has only recently become the standard for broadcasters. The equipment for supporting a digital audio system has matured and become cost effective. The launch of DTV has also helped promote interest in digital audio. This development, combined with the fact that manufacturers are designing more and more digital products (and fewer analog products), has made the transition to digital audio equipment and its operating requirements an inevitable part of our broadcast future.

COMPUTERS AND RADIO: THE NETWORKING CONNECTION

Chriss Scherer, CSRE
editor, BE Radio magazine
Overland Park, KS

Abstract

Computers are used in radio stations for storing, manipulating and transferring audio. With consolidation, the need to tie individual systems together becomes an important issue to complete the integration. Basic networks allow for file sharing between two computers. More involved systems can tie complete networks together. While most radio engineers have skills in audio and RF, computers and networks are a new area of expertise that is being added to the job requirements.

Filling a need

Network advantages

When computers were first introduced into radio stations they were being used strictly for business functions. Even then, there were efforts made to have machines controlling the source material of the station. These early efforts had considerable shortcomings. You may remember an early automation system and that it likely caused more problems than it solved.

Computers have become more a part of our daily business lives. Digitizing audio is a simple task. Manipulating, storing and transferring digital information is what computers are designed to do. A single computer can handle the manipulation and storage just fine, but transferring takes two or more.

The scope of an engineer's duties in the typical station has been expanded to include not only those networked PCs used for the storage and delivery of audio-video information, but to networked PCs in the general non-technical business environment, as well. They are not even called computers anymore - now they are *clients* or *servers*. Trying to keep pace with this technology is difficult to say the least.

Network beginnings

There are two terms that form the basis of any computer network or network model:

Message. This is the digitized information that is sent from one computer to another, or in other words, the basic unit of information transmitted across the network.

Protocols. Simply stated, the rules that define the proper method to package the transmitted message so that it can be understood and processed on the destination computer.

These two basic elements allow us to build a functioning network.

Network architectures

The ultimate goal of the network architecture is to allow interconnectivity of various similar or different computer systems. The specific objectives for a network architecture are:

1. Provide seamless connectivity between all computers on a network.
2. Simplify the task of building the network by use of modular hardware and software components.
3. Support reliable error-free communication.
4. Easy implementation that subscribes to a set of standards.

There have been several network architectures created since 1974 when IBM first introduced its Systems Network Architecture (SNA). Digital Equipment created the Digital Network Architecture (DNA). As is the case with some others, these network architectures are proprietary and designed to allow only computer systems of the respective brands work on a common network.

Network architectures are designed using a "layered" approach where the network is organized as a series of layers with each building on its predecessor. Each layer performs a specific task and makes available (or advertises) the results to adjacent layers.

The International Organization for Standardization (ISO) has been the primary independent body that is responsible for the establishment of international data communications standards. It also provides the dominant framework for network architectures known as the *Open Systems Interconnect* model or the OSI reference model. The specific data communications standards compliant with the OSI reference model are established by the Institute for Electronic and Electrical Engineers (IEEE) through subcommittees known as the 802 family. The OSI reference model is not only concerned with the transfer of information, but also with the interconnection of "systems."

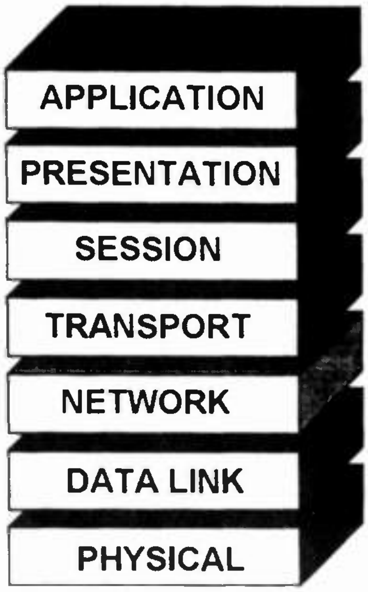
Data communications primer

Data that is intended to be transmitted over a network is broken into "packets." These packets are packaged between several bytes of new data known as the "header." The header contains information about the routing of each packet of data in order for it to be passed to the appropriate destination.

Application process

We all use applications on our computers every time we start a program. If that program is loaded on and running only on a single PC, the total processing of information takes place exclusively on that PC. In a networked environment, however, the processing may still take place on that single PC, but with information derived from another source located on a network. The term application process in this context refers to the application that relies on other network resources to complete a task.

The OSI reference model is based on seven protocol layers. The layers are typically viewed in a vertical stack because each subsequent layer builds on the previous layer.



The OSI model

1. Layer 7, the application layer. This specifies the communications interface with the user and manages the communication among the various computer applications.
2. Layer 6, the presentation layer. This should probably be called the "translation" layer because that's what it does. Computer code is converted from different formats to be usable on the destination computer. For example, this layer makes it possible for an Apple to communicate with an IBM PC or for a PC to communicate with a mainframe.
3. Layer 5, the session layer. Establishes, synchronizes and manages dialog between communicating applications.
4. Layer 4, the transport layer. Maintains the integrity of the data communications. Data flow regulation and error recovery take place at this layer. The data is also segmented into units that will be passed on to the network layer.
5. Layer 3, the network layer. Takes care of "routing" information across networks comprised of multiple segments. Network level addressing is implemented here.
6. Layer 2, the data link layer. Organizes the data into logical frames of information (or packets). Also provides low-level error detection and data recovery. Hardware level addressing takes place here.
7. Layer 1, the physical layer. Refers to the mechanical and electrical specifications for the cabling, network interface cards (NIC) and other items required to physically create the network.

Keep in mind that the OSI reference model is just that - a model for developers and manufacturers to use when releasing a product that is intended to be used in an "open" or non-proprietary environment. Network architectures may not require the use of all seven layers.

Some architectures, such as TCP/IP (used on the Internet), use only four of them. In fact, even those architectures that are considered "proprietary" generally have the ability to map to the OSI reference model and thus be able to communicate with other systems.

For the most part, having a complete understanding of the specifications for each of these layers is not critical. However, you should at least be familiar with them.

Implementing a network

Channel access methods

The amount of expected traffic on your LAN will be the foremost consideration in the design phase of a network; i.e., how much data, how fast and how many users? Passing real-time digital audio and video signals on a LAN would be much more demanding than running a simple business application. Increasing the amount of simultaneous users on the LAN will also encumber the performance of the system. Let's consider how data is passed on each of the three topologies described above. It would be impossible to maintain the integrity of data on the network if all of the devices were permitted to broadcast constant amounts of data. Each device connected to the network must be able to access the network in a somewhat orderly fashion. Any one of three methods may be used to accomplish this. These are called channel access methods. They are:

1. **Contention**. All devices can transmit at any time. If two or more devices transmit at the same time, each device will wait a random interval of time and try to rebroadcast the data. This will be repeated until a reliable delivery is made. This method is used in bus and star configurations.
2. **Polling**. The server initiates queries to other devices on the network in a predetermined order. Devices attached to that network may only respond when queried.

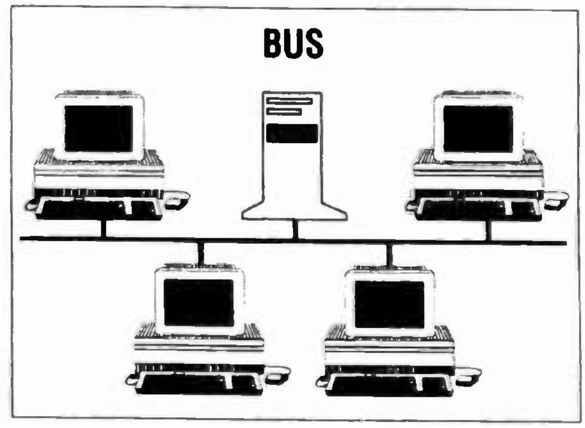
3. **Token passing**. Imagine a token signal circulating around a ring (the cable): the token is passed in an orderly fashion around that ring. A device on the network can only transmit when it is in possession of the token; once finished, the token is released. This protocol is used exclusively in ring-type network configurations.

Because of its speed and inherent reliability, the token ring network has been a popular choice for larger networks and, in fact, is the only method used by systems using fiber-optic backbones, such as the Fiber Data Distribution Interface (FDDI). Systems using FDDI will operate at speeds in excess of 100Mb/s. In the past, most businesses found the cost to deploy FDDI prohibitive and opted to use the copper-based Ethernet protocol, which operates at the slower speed of approximately 10Mb/s. Recent advances in Ethernet technologies have brought the speed in line with FDDI, with the advantage of using the less-expensive UTP cable.

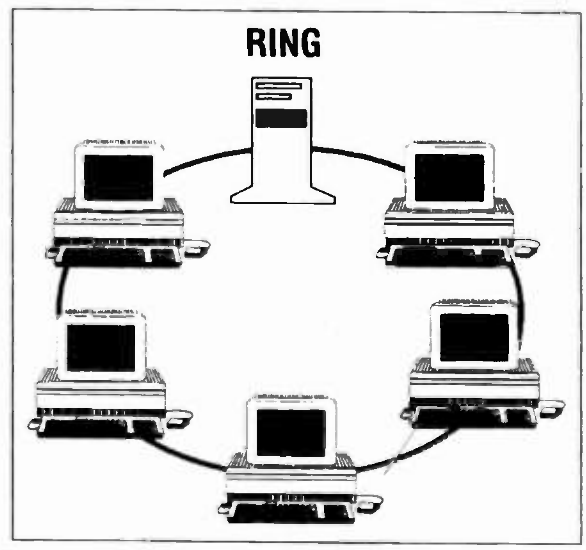
Topologies

Of all the layers defined by the OSI, we should be most interested in the physical layer. This is the level that we, as engineers, will have the most involvement with. The physical layer deals with the cabling, connectors and interfaces required to physically attach the computers on the network. The types of topologies that you will most likely run across are:

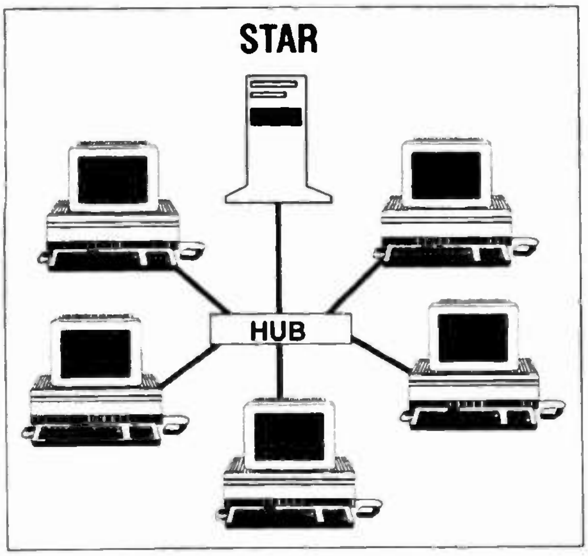
Bus. Consists of a linear section of media, such as coaxial cable, terminated on each end and are connected through a "tee" connector.



Ring. Similar to the "bus," except the ends are joined, hence the name. In practice, the ring configuration actually consists of dual counter-rotating concentric rings. Devices attached to a ring network will have "in" and "out" ports.



Star. More commonly referred to as a hub configuration. The signals are carried from each device on the network to a central point (the hub). These hubs can be passive or active.



Network protocols

By definition, network protocol refers to the rules that govern how any two or more elements on a network communicate. The OSI is defined by seven protocol layers. Working committees of the IEEE are responsible for defining and ultimately standardizing the various LAN protocols at the datalink and physical layers. There are many elements that are considered in the protocol standards, but for our purposes, the primary differences of these standards relate to data throughput, transmission media (cabling) and channel access method.

Currently, the 100Mb/s Ethernet protocol has become the LAN protocol of choice for most new installations. Keep in mind that standards are currently in development for 1,000Mb/s (Gigabit) Ethernet.

Three types of media have been specified to transmit 100Mb/s Ethernet signals:

1. 100BASE-TX (Fast Ethernet). Essentially a faster "contention"-type access method.
2. 100BASE-T4 (100VG-AnyLAN). Uses a new approach to access the network called demand priority. This system supports token passing. One of the real advantages of this method is that it uses four pairs of telephone voice-grade cable. The VG stands for voice grade.
3. 100BASE-FX. Fiber-based "Fast Ethernet" similar to FDDI.

Here are a few more high-speed protocols you should be aware of. These are not very common for station audio operations, but may be used for business operations.

Asynchronous transfer mode (ATM). A technology based on high-speed packet switching, ideal for multimedia and other complex applications. ATM is capable of data rates from 51Mb/s to 622Mb/s.

Copper data distribution interface (CDDI). FDDI using copper-based media instead of fiber.

Fibre Channel. Developed jointly between Hewlett-Packard, IBM and Sun Microsystems. Fibre Channel requires the use of fiber cabling and is capable of passing data at a rate from 266Mb/s up to several Gigabits/s. Fibre channel is being commonly used in video applications.

These are some of the high-performance protocols available. All of these offer data throughputs of 100Mb/s or more, however, most of the PC-based networks in existence today are still using protocols that pass data at approximately 10Mb/s. In terms of performance, these will work just fine for many business-type applications, but generally are not suitable for installations requiring high throughputs or passing real-time information. Some of the 10Mb/s Ethernet standards are:

10BaseT. Uses UTP cabling in a hub arrangement.

10Base2. Uses "thin" coax (RG-58), arranged in a bus configuration.

10Base5. "Thick Ethernet" uses heavier coax (RG-8), allowing longer line distances.

Although not considered a high-speed protocol, wireless Ethernet is becoming a popular media for connecting portable workstations (laptops) to a LAN or for connecting LANs between different buildings where cabling is impractical. The FCC has permitted unlicensed operation of these systems at 902-928MHz and 2.4-2.483GHz using spread-spectrum modulation. Most of these systems will provide data rates from 2Mb/s to 10Mb/s, at distances of approximately 30 miles with directional antennas.

For high data rates via a wireless media, consider using a standard point-to-point T1 microwave system. This approach is typically more expensive and requires FCC licensing.

The LAN infrastructure

Proper planning and installation of the LAN infrastructure is the single most important task in the building of a reliable and upwardly expandable LAN. Given the current pace of technology, this is the one area that probably won't need to be replaced in the next five years if the wiring and power systems are designed and constructed properly. Strangely, many companies (particularly broadcasters) skimp on this phase and, consequently, spend a great deal of money and time resolving related problems.

Cable and connectors

Coaxial Media

Coaxial cable is more expensive than twisted pair cable, but because of its increased immunity to outside EMI, can be a good choice for some applications.

Coaxial cables, once the de facto standard to connect devices on a LAN, are becoming less common. Unshielded twisted pair (UTP) and fiber cables are currently the most popular transmission media. Depending on the size of your LAN and the distance it must cover, the cabling and installation costs will represent a significant portion of the project. Each type of media has pros and cons.

The Ethernet specification provides for the use of two different types of coaxial cable, using a bus (or daisy chain) topology. Thin-net or 10Base2 uses 50 ohm RG-58/u or equivalent cable terminated with BNC connectors. The cables are strung between each network device using a BNC *Tee* connector and each end is terminated with a 52-ohm termination resistor. The maximum length of the string cannot exceed 185 meters or about 607 feet. Thick Net or 10base5 uses 75 ohm RG/6 or equivalent cable. It is similar to the thin-net approach, however uses special *taps* which are clamped over a drilled hole in the cable. Thick Net should not be used for new installations, but does offer the ability to increase the total cable length to 500 meters or 1640 feet. Optical fiber cable would be a more cost-effective approach.

Optical Fiber Media

If you need raw bandwidth and absolute immunity to EMI, Optical fiber cabling provides the highest performance of all network media types. Two types of fiber optic media can be used. Multimode fiber provides several paths for light to pass through a cable, while single mode has only a single path. The light source and wavelength used to transmit over these cables also varies. Light Emitting Diodes (LED) operating at 850 to 1300 nanometers (nm) are used for multimode cables and lasers operating at 1310 and 1551nm are used for single mode. Fiber cables are constructed either as a *tight* buffer, where the cable is tightly encased in its sheathing, or a *loose tube* design, where the cables are suspended in a moisture resistant gel for outdoor use.

The conductor(s) in optical fiber cable are made from glass, and thus subject to a great deal of handling and installation requirements. Extreme care should be exercised in attaching connectors to, splicing and bending the cable.

The typical fiber media used for LANs consists of a graded index multimode fiber-optic cable with a 62.5 micron fiber-optic core and 125 micron outer cladding. Single-mode fiber is used in applications requiring wider bandwidths and longer distances.

Twisted Pair

Most currently designed Ethernet networks operate over a cable made up of four individually twisted pair wires. Also called category 5 cable, these are tested and certified to pass data at speeds of at least 10Mb/s at distances of up to 328 feet. There are new specifications called "extended" category 5, which allow data speeds beyond 100Mb/s up to Gigabit Ethernet. The point here is to purchase cabling that is rated for as much data bandwidth as possible. Token passing networks require two signal paths: signal in and signal out. For this reason, either a special coaxial cable utilizing two inner conductors -- called Twinax -- or dual fiber optic media are used. In any case, you need to exercise care in deploying the cables, as there are total length limits for runs of cables that depend on the network topology selected.

The nuts and bolts of designing an infrastructure that will deliver maximum performance and reliability -- not only within a building, but to the outside world. Ethernet networks are by far the most popular and will most likely end up as your choice; however, you could have both Ethernet and Token passing networks working together.

That's a summary of just a few of the cables you will need to be aware of in your installation. Remember, choosing the proper cable is not a substitute for proper installation techniques.

What is Ethernet?

In 1977, Xerox was granted a patent for a "Multipoint data communication system with collision detection." Collision detection is the key element that distinguishes the "Ethernet" network. The formal acronym for this type of access method is CSMA/CD or "Carrier Sense Multiple Access/Collision Detection. Here's how it works: 1) A computer can send a "message" over a network as long as it doesn't sense another computer also sending one, this is the "carrier sense" 2) When the message is finally sent, it is distributed over the network and "listened" to by all other devices attached to that network -- the "multiple access" 3) If two or more computers send a message at the same time a "collision" occurs, when the collision is detected, each computer "backs off" or waits a random period of time then retransmits -- the "collision detection." The "access method" is only one part of several individual standards which have been established by various working groups within the IEEE.

How fast?

The most common Ethernet networks currently in place are operating at 10Mb/s and typically connected by a specially rated cable which contain 4-twisted pair wires, commonly called category 5 (or cat 5). This type of network is also called *10BaseT* (10 = 10Mbit/s rated throughput, Base = Baseband (signal) and the T = twisted pair cable). Standards now exist for networks operating at throughputs of 100 Mb/s and known as *100BaseT* or *Fast Ethernet*. The price for deploying a 100Mb/s network is rapidly falling and even if you presently have an investment in a 10Mb/s system, it still may make sense to implement the higher speed technology as you add new equipment, because it is possible for the two to coexist with proper planning. Network interface cards (which plug into one of the slots on the PC) that support *both* 10- and 100BaseT networks can be purchased. A *Gigabit Ethernet* standard is new standard that is beginning to appear. Gigabit Ethernet will operate over existing Category 5-EX cabling, as well as coaxial and fiber optic cables. It's unlikely that radio broadcasters will be implementing this anytime soon, but you can expect that it will grow in popularity over the next couple of years.

Putting it together

Most network problems relate directly to cabling problems. Specifications exist for every aspect of the physical implementation of your network, especially cabling. Most Ethernet networks in use are connected using category 5 cable which is terminated with modular RJ-45 connectors. The span (typically cabling) between devices connected to your network are known as "segments." The physical length of each segment is defined in that specification. These lengths were derived based on the propagation factor of the transmission media and the inherent signal delay that would result. Care should be exercised when installing any network cabling as bends, kinks, poorly made connectors and placement in close proximity to interfering sources (i.e. electrical wiring) will degrade network performance. It's possible to slow the performance of your entire system -- significantly -- with a defective cable in just ONE segment.

The transmission media itself is defined within the same specification. For example Category 5 twisted pair cable is graded to reliably pass 100MHz at 328 feet, which is more than acceptable for a 10BaseT (10Mbit/s) network. If you're designing a 100BaseT network (100Mbit/s), you should consider using a higher performance grade of cable such as Category 5-EX, which is rated to about 300MHz for equivalent lengths.

Additional hardware

The term "star" in its' real world form refers to a network in where all of the computers (nodes) are connected to one or more common device called a *hub*. Hubs are multiport devices which, simply repeat data packets received from any one node to all of the attached nodes. Hubs are a CRITICAL link to consider in your network design, because all of the data must pass through at least one of them. Hubs can also be connected to other hubs forming a *cascading star* topology. There are several choices of hubs available, so you should carefully research your options. These choices include: basic units which support only a single throughput speed, hubs that will adapt to 10 or 100Mb/s (either by autosensing or manual selection) and even hubs that can mix and match different media i.e. twisted pair wire and fiber optic cable. For obvious reasons, hubs that adapt to more than one speed must be able to "buffer" data transmitted from the higher speed segment (100Mb/s) to a 10Mb/s segment. This creates two problems from a performance standpoint 1) the hub must have a sufficiently large enough memory to temporarily store the faster received data 2) if the network has a large amount of traffic flowing through it *and* the hub cannot retransmit data from the high speed segment to the lower segment in a timely fashion (on the order of microseconds), the Ethernet protocol takes over and assumes the data was lost and calls for the originating device to retransmit. This can cause a serious performance problem! A new variety of hub uses intelligent "switching" techniques to route data directly to its destination, not unlike a telephone central office routes a telephone call. The technology is known as *Layer 3 switching*, of which there are about 12 different vendor specific choices. The price for these vary with application, but are becoming much more affordable. Standards also exist for these devices, however, many switches on the market currently utilize some form of proprietary technologies. Many of these switches offer the ability to operate in a "full duplex" mode which eliminates the speed limitations imposed by the CSMA/CD access method by establishing discreet paths for the transmitted and received data. Network performance typically doubles -- a good reason to spend a few dollars more!

Because the signal propagation delay is critical on an Ethernet network, you must limit the not only the length of the segment, but also the amount of "hops" from end-to-end. In this case "hops" means any network device that "retransmits" data such as a hub. The rule of thumb for 10Mbit/s networks is:
No more than 5 segments in series
No more than 4 Hubs
No more than 3 populated segments (applies to coaxial based networks only)
This is known as the "5-4-3 rule." The rule is pretty much the same for the 100Mbit/s networks, except they tend to limit the amount of hops allowed, because the total system signal propagation delay must be decreased, in order to sustain the higher speeds.

Connecting to the outside
The growing need for network users to be connected to the outside world (i.e. telecommuters, Internet access and interconnecting networks between various facilities) has caused an explosion in the number of affordable remote access solutions. These devices are usually called *routers*. In its simplest form, the term router defines a device that connects two or more NETWORKS using the most efficient path available. What a router is really useful for is it's ability to pass only network traffic (in one network) that is destined for a different network. The Internet is a good example of how routers work. When you specify a particular web site on your browser, for instance, that name is forwarded to a Domain Name Server which associates the name with a specific numeric address -- the path to that address is established through one or more routers located throughout the world. Most current routers are extremely configurable in terms of how and what traffic is allowed to pass through.

Other types of routers are used to connect the network to private or public networks. Routers such as these can be purchased to interface with virtually any type of outside network in existence, these include: POTS, ISDN, switched 56, full (or fractional) T1/T3, frame relay and ATM. Most of these routers are intelligent and can be programmed to establish connections with dial-up services from any authorized user(s) connected to the network as well as allow authorized user(s) to have direct access to certain network resources.

Performance Issues
Current digital based broadcast facilities offer an interesting challenge to network design. Unlike most other businesses where the networks are used to access files, share resources or even accessing the Web -- Broadcasters have the additional requirement of passing various forms digitized audio (and/or video) streaming data through it. Recall that in an Ethernet network, any network device may transmit at any time, collisions due to simultaneous transmissions will need to be retransmitted at some random future time. Networks that are designed to carry streaming data are typically sending huge amounts of traffic for long periods of time, thus tying up the systems for other devices. Networks intended to

carry streaming data must be properly *segmented* in order to maintain network performance throughout the system. In context, segmenting a network means that only certain data is allowed to pass through to another network, while the remainder of traffic stays within the respective network. Most radio stations use networks to perform several functions -- for example: traffic, business, E-mail and perhaps Internet access. Consider that this station also uses an audio storage/playback system (for several stations) which is connected to a few digital audio workstations -- all through your network. Assuming these functions were carried on a single network (or unsegmented multiple networks), data congestion would certainly be a problem. Ethernet networks are prone to several conditions where excessive traffic can cause problems ranging from poor network performance to outright network failure.

Currently in the majority of stations in which this scenario exist, the engineers have opted to keep the network carrying the streaming data separate from any other network. In those cases where a traffic system interfaces with an On-Air storage/automation system, it would be helpful to transfer daily traffic information directly to the audio system. Many stations use a simple bridge (using a standalone PC) or just carry a disk between the systems (sneakernet). This type of problem can be solved by implementing a proper routing structure in your design.

One of the most important goals in designing a flexible/high performance network should be to understand and contain collisions within separate network segments, called *Collision Segregated Domains*. The deployment of computer networks within most radio facilities are installed to address the specific needs of particular vendor equipment (i.e. Audio storage/workstations, traffic systems, general accounting systems), which has created a great deal of discontinuity within the facility. The problem is further compounded by the vendors themselves, who typically specify minimal (read-spend more with us, not on the network) network infrastructures that are designed to provide specific types of connectivity in order to accommodate their equipment.

Special considerations

Another important cable specification that you should be aware of when using cables which contain 2 or more twisted pair cables, particularly those cables where the pairs are NOT individually shielded (also called unshielded twisted pair or UTP) is Near-End Cross-talk or *NEXT*. NEXT occurs when the signal from one pair crosses to another pair. Here's the problem: certain data communication applications are bi-directional and use separate pairs to transmit and receive; one example is Ethernet networking. When the cross talk occurs at the "Near-End" of the cable, the potential exists for the signal of one cable (intended for the receiver) to mix with the transmitted signal on the other cable, thus causing a jumbled data stream. NEXT typically occurs with cable runs of 60 feet or less. NEXT can also be caused by improper terminations.

NEXT can also be a factor in specialized cables called *shared sheath*, which is a multipair cable intended to carry a variety of different signals, such as telephone, high speed data, etc.

In areas where cables are subject to high levels of electromagnetic interference (EMI), such as that produced from fluorescent lamps, motors or radio frequency fields, shielded twisted pair or *STP* cable should be used. STP cable utilizes a foil or copper braid shield surrounding each pair of cables within the sheath. In some cases, shielding may be applied to both the individual wire pairs and the outer sheath or to just the sheath. STP cabling is also an effective method to control NEXT.

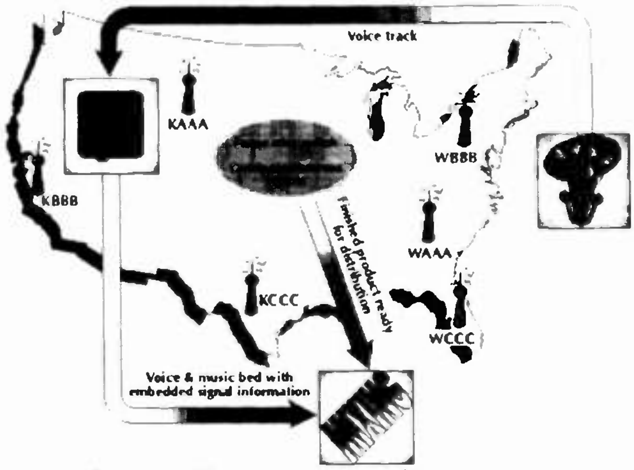
Consolidation of production

With the dust finally settling on the consolidation landscape, owners are looking for methods to leverage their resources in order to realize the true economies of scale. At least one group has already made a substantial commitment to link on-air programming material from their various facilities via PC-based servers connected to a Wide Area Network (WAN), It is likely that more groups will use this approach as well. Implementing such a network to carry real-time programming is still expensive, due to the nature of the leased, dedicated, high-speed data circuits required to reliably deliver the programming to its destination, with the key word being reliably. Sharing tasks that are not real time in nature, such as custom production or certain programming elements, make our choices a whole lot easier (and cheaper) due to the availability of several less-reliable, and bandwidth-intensive, delivery methods. For this reason, consolidating production tasks among various stations within a group is a practical solution.

Two roads

Production facility consolidation can take one of two possible forms:

1. Using fewer production resources (people) and concentrating capital spending on the minimum number of production studios necessary to create and distribute materials to the other facilities.
2. Creating a collaborative workgroup environment by using two or more resources, presumably with specialized skills, but located in different facilities, or even states. An example of this would be a production talent with a unique voice who is located in one city, and who shares a voice track with another production person (located in another city) who then, in turn, creates custom music beds. Then perhaps the resultant tracks are sent to yet a third person who can mix and create the final product.



Sharing files over a WAN to maximize production resources.

The first model is the obvious choice for most owners; however, the second model allows a company to leverage its talent pool in a unique way in order to create a very different sound for each of its stations.

You need to be aware of certain limitations with respect to the specific file formats used by various manufacturers and what data transport methods are available to send the material around. So let's review some of the more universal technical issues.

Data formats

Before we delve into this area, you need to be aware that most, if not all, manufacturers work with proprietary data formats. These formats typically contain not only the audio, but other types of embedded data that are used to provide specific information with regard to where (which tracks) audio will appear, or even to control faders, equalizers, etc. Having this said, unless all of your digital audio workstations are from a common manufacturer, it is unlikely you will be able to use any of these embedded features. Some manufacturers do provide a means to convert their proprietary formats to another, either internally or externally, but this is more the exception than the rule. Many systems provide a means to transfer these formats to other manufacturers' digital storage devices; however, you will need to verify this with the respective companies.

Although there are several good digital audio data formats (such as MPEG), most digital workstations will only allow import and export of the traditional waveform audio format (.WAV) files, normally found on all Windows operating systems since version 3.1. The .WAV format supports multiple data encoding and compression methods and follow the Rich Information File Format (RIFF) specification; the .WAV format can also work across other platforms, i.e. PC to/from Mac.

Although not a specification for sampled digital audio, The Music Instrument Digital Interface (MIDI) is a serial data format widely used to permit electronic-based musical instruments, processing and other devices to communicate how and when the various devices produce sounds. This format can be valuable for use in the collaborative environment.

In a perfect world, you should have your choice of production tools or storage systems without worrying about compatibility. The *Open Media Framework* (OMF) file format addresses this problem. The OMF is an open standard that permits a seamless interchange of digital program material (audio and/or video) and all related formatting, effect, timing and layering information between digital-based equipment from different manufacturers. To date, only a handful of broadcast digital audio manufacturers have included OMF compatibility in their products.

Transporting the data

Once an appropriate file format has been established, you have several choices for getting these data files from point A to point B. In the opening paragraph, I used the term "reliable" to describe a method to get data to its destination. In data communications lingo, a reliable data transport protocol describes a method in which the destination provides confirmation to the originating source that data has indeed been delivered. (An example of this would be a point-to-point T-1 data circuit.) In contrast, a non-reliable transport protocol does not perform this level of confirmation. Frame relay and even Internet protocol are considered non-reliable.

The reason for making this point is that audio that is delivered in real-time should generally use reliable data transport protocols, in part because the high bandwidth demands of such circuits require the data streams arrive in sequence so that they can be decoded properly.

Applications utilizing a non-reliable protocol perform some method of error checking. However, this takes place within the application and can cause slow or out of sequence data to be transmitted between the source and destination. A data file transmitted over a non-reliable data transport protocol doesn't need to arrive at the destination in sequence, because the application will ultimately handle the task of reconstructing the file in its original form. This process doesn't lend itself to real-time file delivery.

The obvious choice for sending your files between facilities is as a simple e-mail attachment through the Internet. Many stations now have dedicated full-time high-speed Internet access, but even if you don't have that luxury, most dial-up connections will give you more than enough speed for most transfers.

Serious work

If you are serious about connecting your facilities for collaborative production, you may consider a dedicated connection to a frame relay network. The frame network can be a private or public network service that is leased from one of the many telephone companies. The cost of frame relay service varies with the speed required (in 64kb/s blocks up to 1.544Mb/s) and, in some cases, the actual amount of data traffic used, but expect to pay at least $350 per month (excluding installation charges) for the lower speed service. Another option to consider is called a Virtual Private Network or VPN. Your Internet Service Provider (ISP) establishes the VPN and it works like this: the ISP creates a virtual point-to-point TCP/IP connection between any two or more Internet addresses (usually addresses hosted by the particular ISP.) The connection performs much like a direct connection and the cost varies, but it can be cheaper than using frame relay.

Of course, there are several other methods to connect and consolidate broadcast facilities, but these should provide a reasonable starting point for your consolidation.

Maintaining your Computers

Computers have become so commonplace in our lives that most of us forget they require maintenance. After all, they usually boot up when you turn the power on, right? Improvements in the manufacturing process of computer components combined with stringent quality control have made the PC and the associated peripherals, even the cheap ones, reasonably reliable.

With prices for PC components lower than ever and still dropping, it is easy to adopt the "I'll just replace it when it fails" attitude. But keep in mind, the real cost of a system failure could be the time and expense that results from the lost data or airtime.

The maintenance requirements for computer equipment are fairly basic. They can be divided into two general categories: physical and data.

Physical Maintenance

The amount of physical maintenance is minimal and determined by the location and specific environment in which the PC hardware will reside. As computers become more powerful and clock speeds increase, so does the duty cycle of the CPU and associated memory ICs. The net result is that these chips operate at increasingly higher temperatures with each new product release. The fan(s) on a computer are designed to blow air out the back, which causes air to be drawn in through open vents in the front. This creates airflow in and around the motherboard and the other components inside the cabinet. Dust accumulating on these vents and inside the unit restricts airflow and consequently the life of the components. Maintaining good airflow through the PCs cabinet is essential. Keeping your equipment in a temperature controlled "clean room" environment is ideal. Few of us have that luxury; however, it should be considered if you are building a new facility.

Canned air, which can be purchased in various forms at most computer and electronic stores, is the most effective method to remove dust and other airborne particles from inside the PC cabinet. The air can also be used on keyboards, floppy/CD-ROM/tape drives, monitors and other external peripherals. Also, make sure all unused card slots and I/O ports on the back of the computer are covered. If you experience intermittent failures, try re-seating any plug-in cards and any connectors attached to each drive.

Hard drives are sealed units and are not normally affected by the environment, however the drive can be damaged by abrupt "power downs." An un-interruptible power supply (UPS) is an effective means to prevent it.

UPSs are not only more affordable, but smarter. Most UPS systems on the market today provide a means to shut down your computer "softly" in the event of a power failure, minimizing data loss and avoiding drive damage. If you own a UPS, don't forget to check the battery(s) periodically.

Data Maintenance

The first line of defense in reducing your chances of data loss is to create a schedule for backing up all of your critical data. Several options are now available, including tape, high density disks, optical and magneto-optical drives. Which one you use depends on your specific needs and budget. If you have PCs attached to a network, consider purchasing a RAID (redundant array of independent disks) multiple drive system which uses various methods to store data on multiple disks. If that is not an option, backing up data to another drive on the network can also be effective.

You can reduce your chances of losing data on your machine through the regular use of assorted hard drive utilities on the market. There are two forms of data maintenance that you need to use: hard drive and virus checking utilities. Current operating systems include utilities which not only format, but also identify and repair errors on your hard drive(s). Both the Windows 95 and Windows NT operating systems include two excellent hard drive utilities.

The process of reading, writing and erasing data may cause files to "fragment" across your hard drive. Fragmented files can cause data loss and waste disk space. Using the disk defragmenter utility will reconnect those files and eliminate remaining fragments of previously erased files. ScanDisk is an included utility that checks the integrity of your drive and fixes many types of errors. Both of these utilities should be run on a regular basis and can be set to operate automatically at predetermined intervals.

Hard drive viruses come in several forms and new types are appearing every week. The damage that these viruses cause ranges from harmless nuisances to permanent damage to your drive and/or system. Of course, your risk is minimized if access to your system is limited to yourself and people you trust; however, if you load files from a disk or download them from the Internet, your chances of getting infected rise dramatically.

There are many good programs that not only find and eliminate existing viruses, but can remain resident in your system and intercept viruses from external sources before they have a chance to infect your system.

Remember, hard drives are mechanical devices and have a finite life. Keep an ear out for drives that are beginning to wear. These tend to exhibit typical "bearing" wear noises, such as whining.

Updating Your Software

Software and their subsequent upgrades are being released at ever decreasing intervals. Have you noticed that the software you buy off-the-shelf is out of date the day you buy it? Many times a new "fix" or "patch" needs to be downloaded in order keep the software current. Most major software manufacturers are releasing these fixes on the Internet. There are even programs available that analyze the current version of programs on your drive and download the appropriate updates automatically.

Network Maintenance

For those of you that are the network "gurus," here are a few items that will save you a maintenance call.

Cabling can directly affect the performance of your network, something particularly true if you send digital audio/video through the system. Just one faulty cable segment in your network can cause performance problems everywhere. If you own or have access to a network cable analyzer/time domain reflectometer (TDR), check the integrity of each run of cabling in your facility. If you have not tested the cabling previously, you may find a situation that could be causing network performance problems. If you have previous documentation of your cabling, re-testing it may reveal changes that can also impact present performance or allow you the chance to correct a future problem.

If you have the proper supervisor rights to the network, you should consider "cleaning" up your server(s). First, compare the authorized user list in the server with the names of those who should have access. It's not unusual to find several previously authorized users that no longer work for the company. Also, make sure that the remaining users have access to only those portions of the network that they need. All current network operating systems have utilities for managing users and their specific levels of access.

Second, check all of the network drives for files that are no longer in use. For example, if the sales department is connected to the server, you may find old proposals, presentations and schedules. If your system stores music and/or spots, there may be a lot of old data that can go! While you are at it, establish a limit on the amount of space that each user can use to save data on the network. This can also be set through the utilities.

BUILDING A DIGITAL STATION

David M. Baden,
Radio Free Asia,
Washington, DC

Abstract

The move to a digital platform for a radio station is no longer an option. With the gradual elimination of analog recording equipment and analog tape medium it is no longer a question of if a facility will migrate to an all digital platform but when they will migrate.

With digital equipment there is no point in time where technology will be static enough to plan around obsolescence. In most cases, especially on a PC workstation platform, equipment bought today will be obsolete as soon as it is received. You can be 100% sure the digital audio equipment you will buy today will be viewed as slow and costly by tomorrow's standards.

The modern broadcast technical manager now shares the same dilemma that Information System (IS) managers have faced for years:

- How do I make intelligent decisions today with rapid growth technology?
- How can I get the most for my money in a world where RAM is traded as a commodity?

With the myriad of digital systems available there is no universal standard with the exception of the AES/EBU interface. There are only systems that are compatible or share at least one common compatible interface.

The challenge is to focus on your facilities needs and plan the best system you can within budgetary limitations.

This seminar will try to navigate through the myths and promises of digital technology in order to make the platform migration as painless as possible. Also examined will be the many unexpected pitfalls of building a digital facility that most technicians fall prey to.

1.0 Conceptual Design Understanding Digital Selling the Project.

Conceptual Design stage of a digital facility design should also be referred to as the "selling phase". This is the point of a project where an overall facility vision is conceived. This vision should include how the facility will function and a general idea of how the various digital systems will be used as a whole. Most importantly it is also the point of the project where the goals of new facility are stated and sold to management. Attitudes need to be readjusted before the project starts.

At this stage budgets are generally set as ballpark figures. It is important everyone involved in the decision making process realistically knows what to expected from a digital facility.

1.1 Digital Realities

There should be a clear understanding as to how a digital facility will effect the technical staff's operational and support efforts. The division between administrative computers users and broadcast systems computer users will become blurred as network, Internet access, system file servers and system storage devices start to become shared resources. Analog station's engineering support staff will slowly start to resemble an IS (information systems) support staffs typical of large organizations.

As the technical support staff gradually move to more of an IS support role they must be made aware of certain operational realities. Realities that IS staffs have been dealing with for many years which will now also become digital broadcast realities. A new operational mindset will be required in regards to all future technical decision making processes. All members of non-technical management must also adopt during the conception phase of the project an IS attitude towards their broadcast facility.

One IS fact is the acceptance of a short life cycle for computer based equipment. Especially in comparison to analog broadcast equipment. In the IS world no matter what you purchase or how well you have planned most computer equipment will be "last version" before it makes it out of the box.

As an example, Radio Free Asia came into being in mid-1996. Since this time the Intel processor platform has upgraded seven times. In 1996 there was the Pentium processor as the high end CPU; next came the Pentium Pro, then the Pentium w\MMX, the Celeron, the Pentium II, the Pentium II Xenon and the most current high-end processors the Pentium III. At the same time processor speed has gone from 133 MHz to 450 MHz with 1GHZ and better planned for release by the year 2,000. This does not even take into account processor releases from Cyrix, AMD or the various Intel mobile processor releases.

The processor has not been the only increase in PC workstation performance. For example: motherboard buss speed has gone from 60 MHz to 100 MHz, RAM from 60 to 10 Ns, SCSI hard drive throughput from 40 MB to 80 MB (not counting RAID improvements), USB ports are new, fire-wire is new, AGP is new, etc, etc.

The OS (operating system) is also an important factor that drives the need to constantly upgrade. Windows for Workgroups the standard OS until Windows '95 was a 16-bit operating system. Windows '95 was a 32 or 16 bit operating system. Windows NT and Windows '98 are 32 bit operating systems only. The expected summer 1999 release of Windows NT-5 will be a 64-bit operating system.

The translation of the OS structure is that 16-bit systems will run on all machines from the Intel 286, 386, 486 to the Pentiums. 32 bit systems will only run efficiently on Pentium processors or better. The new 64 bit systems will only run on Pentium II processors. While 16 and 32 bit operating systems will run on machines that are 64-bit capable over all performance and system stability is sacrificed.

OS drives application software. As the operating systems improve all the application software improves. Newer versions of software will not run on older operating systems and therefore will not run on older PC workstations.

If you are working only with typical office applications (i.e. word processors, small databases and spreadsheets) these performance improvements while noticeable are not necessarily relevant to work. You will never work faster on a document then you can type. This is not true for digital audio and other multi-media applications that are machine and OS performance dependent. Any performance improvement in either OS or machine will yield noticeable improvements.

When dealing with digital audio it is important to remember that all digital audio manufactures take advantage of these PC improvements because they do translate into noticeable performance improvements. In today's highly competitive broadcast equipment business environment all digital system manufactures must upgrade constantly in order to maintain market parity with their competitors. Companies that do not keep in step with today's technology soon go out of business.

While a digital audio system could be installed and let run "as is" for as long as your are willing to accept a static level of performance. In doing so you will reach a point where both the hardware and software will not be supported.

Most maintenance agreements with software providers make version upgrades mandatory. If you do not upgrade as required the maintenance support agreement might eventually be dropped.

This is why, even though digital broadcast equipment may be less expensive on initial purchase, over the long term you may see no great savings. In the pre-digital world of broadcast the life expectancy for analog equipment is much longer then that of it's digital counterpart.

An analog ¼" tape deck retails for approximately $15,000. In an analog recording studio it would be typical to have three or four ¼" tape decks in order to do complicated productions. This is an estimated cost of $45,000 to $60,000 per studio with an equipment life expectancy of at least ten years. In a digital studio the replacement for multiple ¼" tape decks would be one digital audio production workstation.

A digital audio production workstation is capable of multi-track editing and would cost an estimated $20,000 to $30,000. The life expectancy for the digital audio workstation would be no longer then five years. Therefore the digital audio production workstation would be less expensive to implement then analog. Since there is a shorter life cycle with digital equipment the real digital cost over ten years would be $40,000 to $60,000 compared to an analog cost of $45,000 to $60,000.

This is the reality of the digital broadcast age. Corporate IS managers have been dealing with this reality for years. It is a new mindset that a digital broadcast facility must operate and plan with. The true savings of digital audio are the gains realized in efficiency of operations. The price for these operational efficiencies will be a faster, continual facility upgrade path.

1.2 Real Digital Benefits

Operational efficiencies gained in moving from a linear to a non-liner work process are the major benefit of a digital broadcast facility.

The best analogy for linear versus a non-linear work process that can be easily explained to all is the comparison of the typewriter (analog) to a word-processor (digital). When using a typewriter you must know from the beginning to the end of a document, word for word what is to be written. Once the typewriter key is struck the letter is committed to the page. The typewriter is an example of a "linear" work process.

In the typewriter example of a linear work process the first sentence must be written first and the last sentence written is the last sentence. Practically any change to the text generally requires a total rewrite. The word processor functions "cut and paste" in the typewriter analog domain literally translate to cut a paragraph from a piece of paper and paste it on another piece of paper as a correction before the copy process takes place. The number one accessory for a typewriter is the wastebasket.

To produce a written document in a non-linear work process we would use the word-processor. Using a word-processor the starting sentence entered could be the end, the middle or the beginning sentence. Sequence of events can easily be changed, corrections can be made, the document can be formatted and other source material can be merged in seamlessly.

Any technology that allows you to work in a non-linear fashion will allow you to work faster and in most cases yield a better product.

Anyone who has used both a typewriter and a word processor will understand the efficiencies they gained in going from a linear to non-linear work process. If management and the other budgetary powers that be understand this analogy, planning for a facility migration to digital will go smoother. Especially if the first question asked is why do we need to invest in a digital facility.

The migration to from a linear to a non-linear work process is the major advantage of going digital. Creating this flexibility and efficiency in the work environment is the only justification that will hold true over the long term. You never hear an office discuss the possibility of returning to typewriters even though computers are far costlier.

Other Benefits that are commonly sighted as a reason to build a digital facility are also arguments that can be used to sell a project but one must use caution. Some promises may not be 100% deliverable. Some of these common gray area promises are as follows:

1.2.1 Elimination of Tape

The promise of a totally tapeless facility can for the most part be realized. Eliminating daily work tape, storage of tape and the handling cost associated with tape can be one savings and selling point for the move to digital.

The caution is to truly achieve a tape-less facility you must first have a clear understanding of the complete facility workflow. How are programs created and how do they move around? There must be a clear understanding of the broadcast air chain and the facility design must ensure that this chain is duplicated 100% in the digital domain without interruption.

If there is a break in the workflow of the air chain that break will be bridged by "sneaker-net" and a new piece of media. "Sneaker-net" is a term used to designate a transfer of digital information by copying it to a storage media (Jaz, zip, etc.) and physically walking the media around a facility.

The sneaker-net replacement media in a digital facility will be an expensive large storage format digital media (i.e. Jaz, portable hard drive, Zip, DAT, etc). It will be a large format media because digital audio files are generally large and will not fit on an inexpensive floppy diskette. This digital media is more costly then analog media. Three minutes of program on a $0.40 cassette or a ¼" analog tape (with a $0.10 plastic reel) is far cheaper then storing the same three-minute program on a $90 Jaz cartridge.

There will always be an archive cost with either a digital or analog facility. It still holds true that there is a direct relationship between the cost of archiving and the level of archive quality you are willing to accept. But with digital if the archive is not available on line then you may also find producers keeping a favorite program on expensive media as a personal archive. Were as in an analog facility a producer keeping a five-minute program on a $0.40 90-minute cassette is no big deal, 5-minute programs stored on a $5 minidisk start to add up.

There is a minimal 10:1 cost ratio between digital audio storage and analog audio storage. Aggressive media inventory control must be in place to realize media savings in a digital facility.

Before media cost savings are promised, in keeping with our typewriter/word-processor analogy, ask yourself if you have every truly seen a paperless office.

1.2.2 Staff Reductions

The promise of staff reductions as a benefit of migrating to a digital facility is another gray area. While you may be able to realize a staff reduction for tape handling and archiving, you will pickup staff positions needed to maintain the digital systems.

While moving to a non-linear work process will allow the production staff to work faster, it is questionable if this will translate to automatic staff reductions. It is possible that they will be putting in the same amount of time to produce a program but the quality will significantly approve.

There is also a common myth that the migration to digital will allow some talent to start producing themselves. This can be accomplished only with an extensive training program. Part of this training will be in standard audio production techniques. It is important to stress that one must know audio production before they can make a radio program, analog or digital. In keeping with the typewriter analogy, one has to learn to type before they can use a computer. Remember too that learning to use a computer does not make you a writer.

With a digital facility you also eliminate the maintenance task associated with analog equipment (i.e. aligning decks, cleaning heads, etc.). On the other hand this does not translate to the elimination of facility maintenance. Maintenance technicians are still needed to perform maintenance on the digital equipment. The maintenance required for a digital facility again will resemble more industry standard IS (Information System) support.

This will be a detriment as IS technical positions are paid more then audio technicians because they are in higher demand. Salaries for technical maintenance staff members may go up.

In general the staff savings in a digital facility will in reality be less then what many be expected. What you will instead be faced with is a transformation of job duties, which will involve an intensive retraining program.

1.2.3 Automation of Scheduling and Facility Management

Most facilities will already have some level of automation and management. A well designed digital system will allow for the linking of facility automation and tracking together, such as air schedules to, billing, text data files and audio data files. While possibly making life simpler, in the long run it is important to remember that databases require maintenance and support as well as data entry.

1.3 Incremental Migration or 100% Change Over

Years ago I have seen office upgrades where computers were introduced and typewriters were left as a backup for system crashes. In these cases there were some staff members that always opted to use the typewriter. I have witnessed the same phenomenon in radio facilities that tried to mix analog and digital equipment. Some people always used the analog. In addition it complicated the production workflow of the facility by trying to share radio programs across platforms. This caused multiple dubbing of programs from the analog domain to the digital domain.

I am a proponent of a 100% digital facility as opposed to a mixed analog/digital facility. This should always be the goal when funding permits. Once in the digital domain the audio signal should stay digital (as much as your budget permits).

In my view it should be an either or choice, stay analog or go 100% digital. If this means putting off migration to a digital facility for a year until you can afford to do the entire facility then my choice would be to wait. After the year passes, due to the fast life cycle of digital systems, you will have better faster systems to chose from.

Expect a systems performance increase gain from delaying implementation. Do not expect any dollar savings. While IS products are dynamic and constantly changing, which is to your advantage, the prices are relatively static. A top of the line desktop PC now cost about $3,000. This has remained true since the introduction of the 286 chip.

1.4 Pre-Design Budget Estimate

At this stage in the project you should base any budget estimates on corporate financial realities. This is how much money there is to spend. The available dollar amount will dictate the decisions on facility design more then any other factor.

2 Initial Design

Assume now that we are proceeding with eyes wide open as to what life will be like in a digital facility. There should also be at least an idea of budgetary limits. A facility design parameter should now be set. A design parameter is a system performance goal that the digital facility will operate under.

For example at Radio Free Asia our three main design parameters were as follows:

1- *To achieve an integrated desktop PC workstation*. This includes all applications, available to the entire user base at every desk. Applications include, multi-lingual word processing, spreadsheets, database access, house audio monitoring, CATV television monitoring, internet browsing, e-mail, news agency access, network file sharing, network print sharing and digital audio editing. This goal was accomplished.

2- *To maintain a 100% digital audio signal path*. To avoid the all-unnecessary conversion between the digital and analog domain. This was accomplished except in places where equipment did not permit.

3- *To be able to share digital audio files across systems without format conversion*. This was not 100% accomplished but is still in development.

Before you establish a design parameter you must have a clear understanding of the facilities broadcast chain. Your number one question in the selection process for the central digital audio system even before "what can I afford" should be "can this system move the audio files to where they are needed in a timely fashion".

2.1 Digital Systems Overview

In the rapidly changing field of digital audio electronics there is no right choice or one dominant standard. The ultimate test in choosing equipment should be your own personal confidence level in the system. The following are some observations and suggestions to consider when selecting a system.

2.1.1 Digital Audio Systems Server Based

Most large-scale digital audio server based systems are designed to basically to be a sophisticated replacement for the cart machine. They generally record and allow basic editing functions. There are many systems available each offering like and unique features.

In selecting a system I would recommend basing your choice on the "to air" interface that best duplicates current operations. It is my belief that practically any system can be engineered to work for any facility. A basic understanding of the inherent shortcomings of any digital systems and the exact requirements of the facility is needed to engineer a system correctly.

Systems that have DSP (digital signal processor) cards located on the workstation and store audio files on central shared servers require a high network volume to move files from location to location. The files originate and reside, during the editing process, on the local workstation.

When the workstation is finished with an audio file it is upload to a server for shared system access. Workstation DSP based systems will require careful attention to be paid to the facility LAN (local area network) design and segmentation.

Systems in which the DSP card resides on the servers place little or no burden on the LAN. The digital audio file never moves off the server even during the editing process at the workstation. These systems generally require additional cabling to carrying the audio signal between the server and workstations.

Multiple server systems no matter where the DSP card resides require careful planning to make sure that audio files can be reached from all workstations that need to access the file.

On a DSP workstation based system this is accomplished primarily through network segmentation or utilizing high speed switching networks. Make sure that all workstations that need to access data from a given server are on the fastest path possible to hit that server.

On a server DSP based system there is generally a correlation of the number of workstations that can actually connect to any given server at any one time. This numbers is usually limited to the number of DSP cards that can be installed in the server. Files are generally moved between servers, when required, over the network via a transfer program. The ideal would be to have sufficient storage space on all the servers to allow a mirroring across servers of all audio data files. Unfortunately this would defeat the purpose for selecting a server based DSP system because it would overburden the LAN with massive audio file transfers.

Radio Free Asia chose the BE AudioVault system which is a server DSP based system. Network file transfer is kept to a minimum to avoid overburdening the network. There is a physical AES/EBU cable connection to every workstation.

To allow access to audio files across servers a function called "dynamic assignment" was engineered by Broadcast Electronics and the manufacturer of our digital router, Lighthouse. What dynamic assignment does is allow the AudioVault system to access a database in order to find an audio file. Once the file is found an automated connection is made from the workstation to the target server.

This connection is twofold. One connection is via the LAN for workstation control of the DSP card and the Lighthouse router makes the second connection. The Lighthouse router makes the AES/EBU digital audio connection between the server and the workstation. This is one example of engineering a system to work for your specific facility application.

2.1.2 Digital Audio Systems Production Workstations

Most server-based systems offer fairly simple audio editing functions. As servers based digital audio systems replace the cart machine function the production workstation replaces the multi-track tape deck or the multiple single track tape decks. A digital production workstation is a digital audio workstation that allows advance-editing features such as virtual multi-track mixing. A digital audio workstation generally does not have "to air" automation as the server based digital audio system. Unfortunately these digital production workstations are usually more expensive per unit then an edit station on a server based system.

Ideally you should select a digital production workstation based on function and compatibility. Choose several workstations within the projected budget. Narrow down the selection based on the workstations that can move data files to the server based audio system in the most expeditious fashion. Take this selection and allow the user base to pick the workstation they feel most comfortable with.

At Radio Free Asia the Orban Audicy was selected. The users found it to be one of the most intuitive to use. The Audicy also offered network audio program transfer to the BE AudioVault system. In addition the Audicy has recently become a "networkable" system allowing the storage of files on a shared file server.

2.1.3 Digital Consoles

Three years ago when Radio Free Asia was being designed the selection of digital audio consoles was limited. This is an area that has expanded greatly in the past two years. There are several digital consoles on the market that will fit almost any budget and application.

This is another example where you should make your selection based on what is affordable with the features needed. Again I would recommend getting input from the users before the final selection is made.

2.1.4 Digital Routers

When selecting a digital audio router the one should again carefully examine the control program offered on the router. Choose what will be compatible with the existing facility switching automation.

Most digital audio routers are in reality time division multiplexors. This is device that instead of making the connection from inputs to outputs through a switching matrix, as an analog audio router, transmits the digital audio data across a common shared data buss.

For ideal flexibility, select a router that has optional analog input and output modules. This is usually accomplished by having DA/AD (digital to analog/analog to digital) converters on the module. The analog audio passes through the switch in the digital format and allows analog inputs to be switched to digital outputs and conversely from digital inputs to analog outputs.

2.2 Vaporware

Avoid the pitfall of "vaporware" at all cost. As a rule never buy a piece of equipment unless it is already being shipped. Do not trust operational demo units. Shipping next month may translate to shipping sometime in the next six months (or shipping never). This is especially true if you are on a time sensitive dead line.

As a normal general rule of thumb, during any project, if there is a possibility of any piece of equipment being late it will be late. This is true to the power of ten with equipment "shipping next month".

2.3 Real Software Development Time

The IS world has know for years that the phrase "its only software" is in reality password to the fourth dimension where time stands still. Expect all beta copies to be full of bugs. Expect any bug to reveal itself at the moment that it will knock you off the air. Expect the next rev (revision) upgrade that will fix the problems you no have not to have the fix but some new features that you can't even imagine how anyone would use.

Remember a new feature is a bug in the last rev that could not be fixed. To top it off it will always arrive two months after you expected it.

2.4 Factoring In Upward Migration

Figure in obsolescence from the start and plan to deal with it. Unlike analog broadcast gear some digital equipment can be enhanced by upgrading the software in EPROM's (called firmware). All this generally requires is exchanging the EPROM chips.

Ask the manufactures for all broadcast digital gear when there was the last firmware upgrade and when the next one is scheduled. Likewise some upgrades can be accomplished by installing plug-in processor boards. Again ask the manufactures what upgrades are planned for the future and find out if they are included in the original purchase price.

For broadcast equipment that is PC based make sure that the motherboard can be upgraded to faster CPU's and additional RAM. Make sure that the PC case can accommodate additional drives or PCI cards.

2.5 Budget Estimate

By now you should have a pretty good idea of the "anchor" systems for your facility and you should be ready to do a preliminary budget estimate.

My recommendation is to use a spread sheet and:

1- List all the audio equipment that you have selected, use the recommended retail price.

2- Figure in $5,000 per for PC workstations if the total number is over 20 ($7,000 if under 20). This should cover the cost of the local PC workstation, all software (even though you do not now know what you will be required to purchase) and the workstation's share of network services (printers, file servers) cost.

3- Figure $100 per LAN port and this cost should cover all network routers, switches and cable management systems.

4- Figure up the total cable runs estimated in feet and estimate $1.50 per foot as the installed cost.

5- Total the items 1 through 4 and add 25%.

This may not be an exact number but your actual dollar figure when the project is complete should be close and hopefully lower.

3 Detailed Design

3.1 Documentation

When you start to commit the facility design to CAD I strongly recommend that you take the extra time and use a program that is 3-D capable. While the initial drawing process will take a little longer on the back end you will have more accurate and flexible drawings that can be displayed in multiple modes.

To draw a table in 2-D every view has to be drawn as a separate drawing. In 3-D only the object has to be drawn. Once the object is drawn it can be viewed in any angle. It can also now be assigned true to life materials and rendered as a "photo realistic" picture.

A good CAD program, such as AutoCAD, is in real-life scale. An inch in the drawing equals an inch in any drawing. This allows drawings to be inserted as external references or elements inside of other drawings.

3-D cad drawings allow for ergonomic studies and virtual walkthroughs. While this may seem frivolous on the surface there is a large community, sometimes in management, that has a very hard time making a spatial correlation from a 2-D drawing to reality. The closer to life your documentation can be the larger audience you can share the facility vision with.

3.2 Hidden Systems Glitches, Considerations & Cost

3.2.1 Network WAN/LAN

Since an earlier portion of this session covered networks I will not cover this in depth. I will only express that a true IS network is more then a bunch of cables passively passing information. IS network devices such as switches, routers and bridges are active intelligent devices and should not be equated with analog distribution devices such as DAs switches and signal processors. In large facilities the importance of at least one staff member that is a network analyst can not be over emphasized.

Remember three other golden IS rules:

1. *There is no such thing as fast enough (CPU clock, network speed, modem speed, disk access, etc).*

2. *There is no such thing as too much (disk space, RAM, etc).*

3. *More is always better.*

3.2.2 Disk Storage

The future in mass high-speed data storage is SAN (storage area networks) over transmission connections such as fiber channel and fiber optic. Unfortunately this technology is currently expensive, with the average cost of $2,000 per user. Until SAN technology cost come down RAID drives, with high speed SCSI disk, is still the data storage system of choice.

RAID is the acronym for redundant array of independent (inexpensive) disk. It is a way of storing the same data in different places (thus, redundantly) on multiple hard disks. By placing data on multiple disks, I/O operations can overlap in a balanced way, improving performance. Since multiple disks increases the mean time between failure (MTBF), storing data redundantly increases fault-tolerance. A RAID appears to the operating system to be a single logical hard disk. RAID employs the technique called striping, which involves partitioning each drive's storage space into units.

Any disk storage system used will not be failsafe. It can not be expressed strongly enough to include a dependable backup system in any facility design and make it a habit to backup often.

3.2.3 Digital Audio Clock Sync

Pops clicks and digital jitters can occur when playing a digital feed between systems that are not in sync. A flexible master facility sync clock as a reference point is needed in any facility that plans to mix digital audio signals from different sources. This master clock is commonly a word clock and/or an AES/EBU sync clock.

Since it is not possible to run word clock over cable run's of greater than 20 feet AES/EBU sync code is used wherever long distances are involved. This is then converted back to word clock format at the destination.

Digital distribution amplifiers are used to port this clock to wherever it is needed. These DA's are non re-clocking; meaning that what goes in comes out. Therefore there is no inherent processing delay.

3.2.4 Digital Sample Rates

Since an earlier portion of this session covered this topic it will not be covered in detail in this paper. It is ideal to maintain a standard sample rate throughout an entire facility. Excessive sample rate conversions can degrade the digital audio quality.

As a rule of thumb the sampling rate should be set to the highest rate that all digital audio equipment can accommodate.

3.2.5 Digital Signal Levels

It is normal for digital equipment to show a different meter readings when fed by the same test tone. For an all-digital solution it is necessary to establish a facility reference operating level. This reference level must provide enough gain for intelligibility of broadcast while maintaining sufficient headroom throughout the entire signal chain without distortion.

Traditionally with analogue audio equipment the VU meter is the standard device employed for level display. The VU or Voltage Unit is an absolute scale; meaning that any two pieces of equipment with VU meters will display the same level when fed the same test tone. Simple and effective. This is because the VU meter displays an average signal level (R.M.S) - which is directly proportional to the way we perceive apparent loudness.

The digital age has seen the emergence of the peak level meter (ppm). This is because when we record onto a digital format the most important thing to know is "where is full scale" or in other word's "how much further can I go before distortion". Since full scale varies from digital equipment to equipment a standard operating reference must be established.

This standard facility wide reference level will be set after full scale is known for all digital audio equipment.

3.2.6 Cabling

The AES/EBU specification has an impedance tolerance allowing for the use of cables with impedance from 88 ohms to 132 ohms. 110

ohms is the optimal impedance for AES/EBU cable.

It is true that digital audio cable can be used for analog since the low capacitance of digital cables makes them a superior analog cable.

Conversely Analog audio cables should not be used for digital audio since the cable impedance of most analog cables is between 40 to 70 ohms. If analog cable is used distances of 50 ft. should not be exceeded or bit errors may occur.

Shielded Category five (Cat-5) data cable has impedance of 100 ohm and may be considered as a low cost altemative to AES/EBU audio cable for installations.

3.3 Installation Timetable

It must be established if there is an absolute "finish by" or "drop dead" date. If they is such a date, for whatever reason (i.e. facility relocation), then this date will drive the project. A drop-dead date will influence the type of equipment purchased based on availability in the needed time frame.

If there is no firm deadline then one should be set. This date should be realistic but also should be treated as a concrete target. Only allow slippage due to circumstances absolutely beyond your control.

3.4 Equipment Ordering and Delivery Dates

Using a multiple vendor biding process will generally result in the best prices. Do not be married to the lowest bib. Remember before awarding a purchase contract such variables as local product support and delivery dates.

Saving 5% is no real savings on a purchase if there is no timely support offered or your must wait a month, or longer, to receive the item.

3.5 Lining UP the Installation Team

Unless you have contracted a systems integrator you will need to assemble an installation team. Make sure when putting together this team that you have expertise in all the areas needed. This should include at least one expert in networking and one expert in the various operating systems that will be installed.

If current technical support staff exists to do the installation now is the time to start intensive specialized training. Crash two-week courses, offered at technical leaming centers, might be helpful. Make sure to get all training and technical manuals on all major systems purchased as quickly as possible. Take advantage of any manufacturing training programs that are available.

If you are assembling a new team for the installation it is a good idea to get them started as early as possible. There is no such thing as too much preparatory work such as, starting to prepare operational manuals for the facility user training.

3.6 Budget Check

As orders are being placed the real cost for equipment and systems can now be figured in. You will also find that the list of equipment, software, hardware, test equipment and tools that need to be purchased will keep growing. Over estimates in the initial budget calculation should be enough to cover these additional items.

4 Implementation

The acceptance of any extreme change in the work environment will always be meet with an initial resistance and resentment by the staff members that are effected. Most people prefer not to change their working habits without a good reason. Even with a good reason the natural response to any change that will require an extra effort on their part will be translated into to criticism of the new system.

4.1 Training

The best way to minimize the complaining period post installation is with an aggressive training program. This training should be interactive and hands on. Pilot training systems should be built as soon as possible.

Training should be a planned path to proficiency in computer and production skills. Training before the installation period should be mandatory for all system users. The installation team are the first staff members that should be trained. A sharp competent team will train themselves and develop the training program for the general user base.

I believe that all members of the installation team should be tasked with training the user population. The constant review of system functions and operations will prove useful during the system installation. No one can ever be too close to the system.

4.2 Pre-installation Work

As much equipment should be assembled and tested as possible before installation. This will generally require a staging area if the new systems are replacing existing systems. Every piece of equipment received should be tested and certified before installation.

It is not safe to make an assumption on the performance of any equipment. I have personally witnessed one hundred computers being received and thirty-two disk drives being bad. This is an astronomical failure rate for a device that is generally rated for a 60,000 hours life expectancy. There is not only such a thing as a factory produced "lemon" there is also "lemon production runs".

All new cabling should be installed and standard IS practices should be adopted and all cabling should be tested and certified. This includes all digital and analog audio cable.

4.2 Facility Change Over

If you are changing a facility over from analog to digital then the change over period should be as fast as possible in one crossover if possible. With proper preparation (i.e. training) the user base should be ready for an overnight system change.

To allow the old and new system to operate together invites a backslide to the old system use. The first week is always the most troublesome, this is where all the unseen problems shake out. Why invite unfavorable comparisons to an online stable older system

4.3 Budget (Hard Figures)

No matter how well a project is planned towards the end you start to bleed money. This is unavoidable. It is not humanly possible to anticipate every detail and complication that will arise. This is why you must over estimate your budget projections at the beginning of the project.

Looking back at our initial budget formula:

6- List all the audio equipment that you have selected to use the recommended retail price.

   The audio equipment probably came in at least 20% under suggested retail. But there are more likely 20% more items on your list.

7- Figure in $5,000 per for PC workstations if the total number is over 20 ($7,000 if under 20). This should cover the cost of the local PC workstation, all software (even what you do not know now that you will be required to purchase) and the workstation's share of network services cost.

   Real cost was probably $1,700 per PC. There were probably 15% more PCs ordered then expected and more unexpected software. If you include server *and network OS cost it is probably pretty close to even.*

8- Figure $100 per LAN port and this cost should cover all network routers, switches and cable management systems.

Network Ports probably came in at $50 per. There should be at lest 15% spare ports for backup and possible expansion. There should also be some unplanned network test equipment.

9- Figure up the total cable runs estimated in feet and estimate $1.50 per foot as the installed cost.

Runs were about $1.10 per foot. There are most likely 10% more cables run the expected.

10- Total the items and add 25%.

This overage was the project safety net. It should almost be gone. It functions as a philosophical insurance policy to make sure the proper equipment and materials are ordered. If this figure was not there many items that were not considered in earlier planing stages would not have been ordered, no matter how seriously the were needed.

Interactive Technologies for Television

Wednesday, April 21, 1999

2:00 PM - 5:30 pm

Chairperson: Jerry Butler
PBS/Public Broadcasting Service, Alexandria, VA

2:00 pm **Going Beyond TV on the Web ⇔ Making Real-Time Web Based Searchable TV a Reality**
Mark Juliano
ISLIP Media, Inc.
Pittsburgh, PA

*2:30 pm **From the Internet on Television... To Interactive TV**
John Matheny
WebTV Networks, Inc.
Mountain View, CA

*3:00 pm **The Future of Digital Television**
Mark Porter
Oracle Corporation
Redwood Shores, CA

3:30 pm **The Convergence of TV and PC - An Arranged Marriage**
Jurij Paraszczak
IBM T.J. Watson Research Center
Hawthorne, NY

*4:00 pm **YourStation.com: Broadcasting for the Broader Broadband Audience**
Adam Sharp
Assets New Media Corporation
Evanston, IL

4:30 pm **So You've Gone Digital.. Now What?**
Dr. Dov Rubin
NDS America
Newport Beach, CA

5:00 pm **ATSC Digital Television Data Broadcast and Interactive Services**
Aninda DasGupta
Philips Electronics
Briarcliff Manor, NY

---

*Papers not available at the time of publication

Going Beyond TV on the Web -- Making Real-Time Web-Based *Searchable TV* a Reality

by

Mark Juliano
President & CEO
ISLIP Media, Inc.
Pittsburgh, PA

Abstract

Technology currently exists for hosting TV programming (news, sports, features, etc.) on the Internet. However, to date, the primary focus has been given to streaming and broadcasting. While these technologies are prerequisites, they do not enable the core time-honored Web functionality of searching. Creating truly 'Searchable TV' requires a new set of tools for cataloging, indexing, browsing and searching of video. This paper discusses the current status of these tools and techniques as well as the core benefits of real-time Web-based Searchable TV.

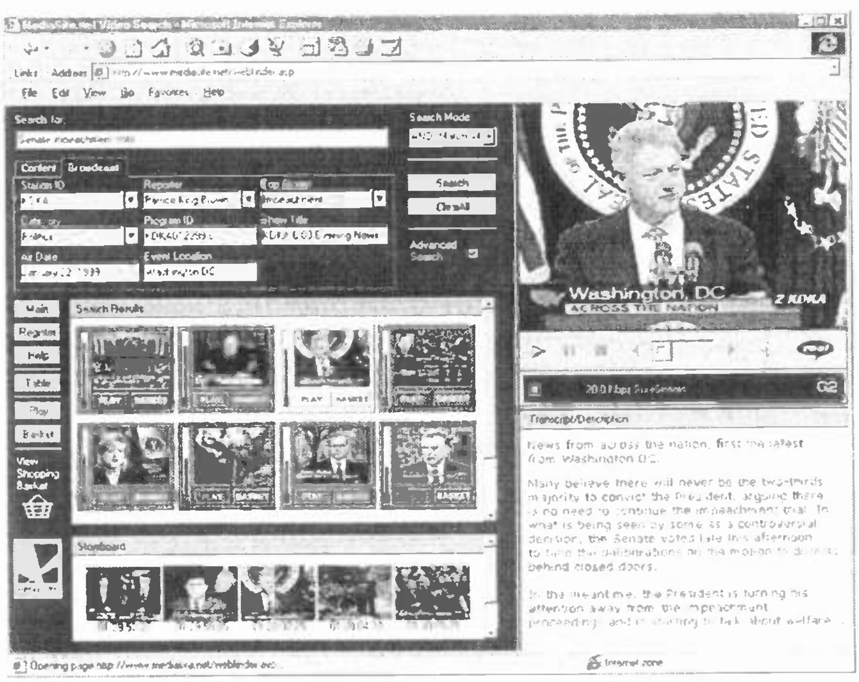
Introduction

With the increasing use of TV and cable as an information medium, a wealth of video information exists. This video could be tapped for a variety of uses if only people could easily and instantly search the millions of hours of existing, archived video content, as well as the new video-based information generated every day.

The concept of interactive television has long been discussed, but has not been implemented on a wide scale. With the growth of streaming video on the Internet and today's video cataloging, search and retrieval technologies, Internet based *Searchable TV*, which goes far beyond interactive TV, is now a technical reality, and users can search broadcasts as quickly as one hour after airing. Tools and technology are also in development to make real-time *Searchable TV* possible.

The real-time *Searchable TV* technology, discussed herein, makes applications such as custom news, on-demand time-independent TV viewing, and TV-based video research, a reality. Real-time *Searchable TV* would allow any wired home or business user to search all channels currently broadcasting, as well as stored programming for topics of interest.

In essence, today's TV broadcasting is a time-based, 'push technology' event. If a



viewer is not tuned into that particular channel, at that particular moment in time, the viewer will not see the program. Videotaping a program does not adequately solve this problem since no individuals could tape all the channels, at one time. Nor could specific programming on a specific topic be found within all of the resulting tapes. Also, when programming is only broadcast over specific geographical area, location becomes a limiting factor. If the viewer is out of broadcast range, they do not have access to programming.

Real-time *Searchable TV* is a classic one-to-one 'pull technology' based upon individual viewer's desires. As with the text and image based Internet of the 1990's, search engines and subsequent portals became the key software enabler to creating an effective Internet. The following examples illustrate this point applied in the video domain.

- Enter *"show me all current programming aired within the past 24 hours, which discussed John Glenn and the space shuttle Discovery"*. This query would yield a number of short segments (not whole programs) from news, talk shows, feature programming, and even NASA-produced video which specifically discuss this topic. Segments would be put in a bin where the user could watch them in any sequence from a video-enabled PC.

- Enter *"what has Allan Greenspan said about interest rates this past week?"* The segments returned might be short clips from CNN Headline News, national news stations (i.e. ABC, NBC, CBS, Fox), 20/20, Crossfire, CSPAN, etc.

- Enter *"show me touchdown catches by Jerry Rice of the 49ers"*. The segments returned would be from games played that day or that season, highlights on ESPN, highlights from local and national sports news (during a longer news program), or an interview with Jerry Rice after the game.

- The viewer could develop *custom news* by creating a "profile" of desired programming. For example, *"local sports, finance, and cooking." Searchable TV* would then collect segments of programming during the day which, for example, a person could watch that night. The clips would only match that viewer's programming profile. This would save countless hours and eliminate undesirable programming.

The core enabling technology of real-time *Searchable TV* is a "real-time" version of existing video cataloging, indexing, search and retrieval software to allow viewers to search programming within seconds of its broadcast. For pre-recorded programming, searching could even occur before the program aired (if the broadcaster desired). This is similar to current newspapers putting select stories on the Internet before their next printed edition.

Overview of Today's Near-Real Time Web-Based *Searchable TV* Capabilities

Technology has just emerged to enable near-real-time *Searchable TV* (as quickly as one hour after airing) from companies such as ISLIP Media. The tools and technologies which are integrated to make this occur include the following.

Encoding Tools digitally capture and encode incoming video in real-time and use advanced compression techniques for use in low-bandwidth (i.e. RealNetworks and Microsoft NetShow) or in MPEG format.

Cataloging Tools perform sophisticated content and time-based image and motion analysis to automatically identify key scene changes with associated time-codes, and build a storyboard. Catalogers can then add attributes to video segments including description, title, location, and other user customizable information via spoken annotation (using speech recognition) or typed input to quickly capture the maximum amount of information possible.

Automated Post-Processing Tools integrate speech recognition, image recognition, language understanding, and global positioning system (GPS) technologies, which enable video to be searched with a high degree of accuracy.

Video Streaming Software performs the 'Web-casting' function on the Internet. While video streaming technology (i.e. RealNetworks and Microsoft NetShow) are prerequisites for video-on-the-Web, they are not sufficient for truly *Searchable TV*.

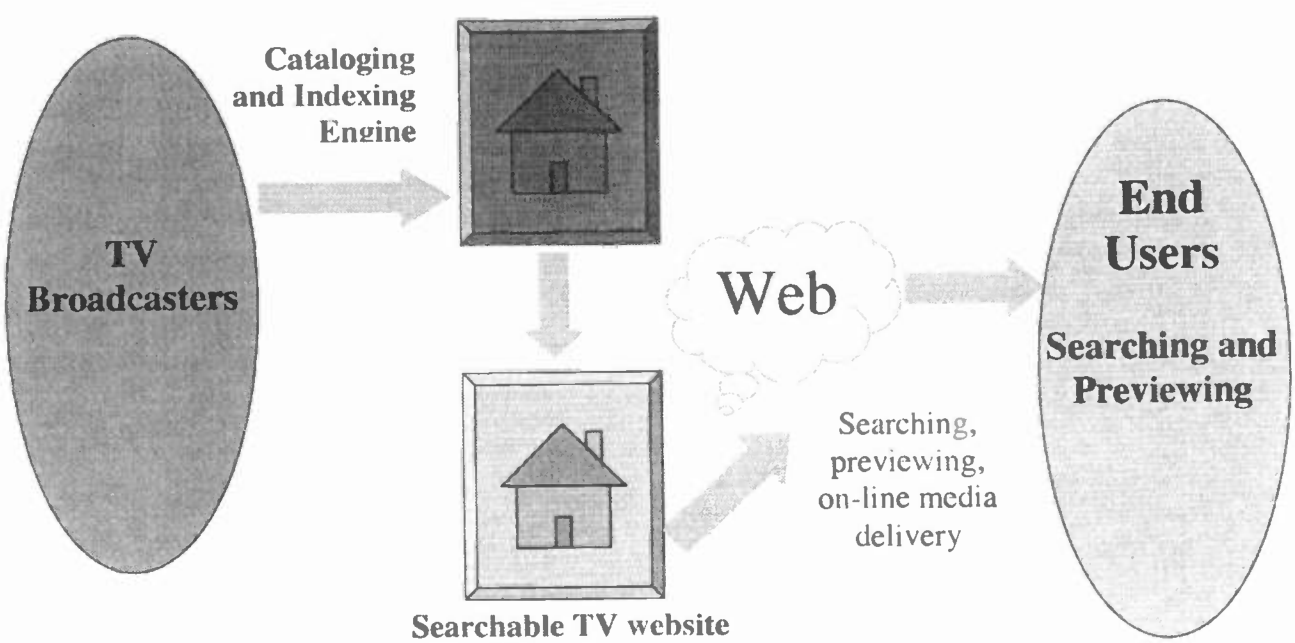
Web Based Search and Retrieval Tools support a range of online access speeds, from low-bandwidth (28.8kbps) connections to high-speed T1, and in-house Intranet connectivity with full-motion video playback. Navigation techniques allow browsing and searching 'in the fourth dimension' using speech-to-text scrolling and alignment, video skims and summaries, and storyboards to watch video *faster than real time.*

Asset Management Tools for managing a range of content, including text, image, video and audio, and for managing usage rights.

Server hardware and software necessary to manage, retrieve and play video and audio technologies over the Internet. These include

> *Database Servers* for managing metadata, which is key information extracted from video and audio.

Searchable TV Architecture



Video Servers include various streaming media technologies giving users a choice depending upon preference and bandwidth.

Web Servers contain Web, Search and Retrieval and *V-Commerce*™ (video E-commerce) hardware and software for Web connectivity and E-commerce capabilities.

V-commerce (video E-commerce) software which allows short clips (a few seconds) and long programs to be viewed and purchased on a pay-per-view, pay-per-clip, or pay-per-videotape basis.

Process for Creating Near Real-Time *Searchable TV* Today

Near real-time *Searchable TV* is a reality today. The following steps are performed during video cataloging by a coordinated suite of programs:

1. Digitize video and audio data into low – bandwidth (RealNetworks) and/or MPEG-1 compression format. Using relatively inexpensive off-the-shelf PC-based hardware, video and audio can be compressed to about 150 Mbytes/hr of video (RealNetworks) or 520 Mbytes/hr of video in MPEG-1 format.

2. Create a time-aligned transcript (from speech recognition and/or closed-captioning). The audio portion may be fed through the speech recognition routines, which produces a searchable transcript of the spoken text. A number of broadcast programs also have closed-captioning text available.

3. Closed-captioned data may lag up to 25 seconds behind the actual words spoken. To create a time-aligned transcript, the time-accurate speech recognizer output is aligned against the closed-captioning.

4. Segment "story" boundaries. To allow efficient access to the relevant content of the news and programming, video broadcasts are broken up into small pieces or *segments*. The speech signal is analyzed for low energy sections that indicate acoustic "paragraph" breaks through silence. If a closed-captioned text transcript is available, structural markers are used to identify stories. Segments can also be determined using speaker change detection. An alternative is to perform manual segmentation.

5. Image analysis methods are used to describe image features such as color histograms and hues. Face matching, a variant of image analysis can be used to detect and classify human faces for later search and retrieval.

6. Index all segments. The segments and their corresponding transcript, keywords, images and faces are indexed in the video catalog. An inverted index is then created.

7. Abstract and summarize the content. By analyzing the words in the audio track for each video paragraph with TF*IDF weightings, the system is able to determine probabilistically the theme and subject area of the narrative. This understanding can be used to automatically generate headlines or summaries of each video segment for icon labeling. Using data extraction technology developed in the DARPA TIPSTER and MUC research programs, the system can attempt to identify names of people, places, companies, organizations and other entities mentioned in the sound track. This

allows the user to find all references to a particular entity with a single query.

What is Currently Available to Broadcasters Today?

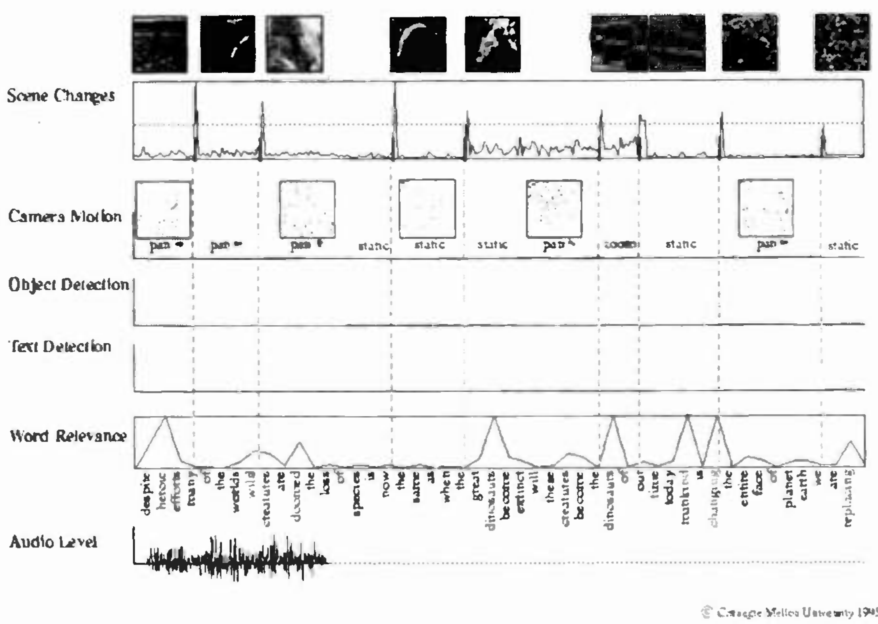
Broadcasters of news, sports, and general programming currently have several alternatives to deploying *Searchable TV* over the Internet. The first fundamental question is a make or buy decision. Broadcasters can work with a service provider (such as MediaSite.net) who can catalog, index and host *Searchable TV* . This option is a turnkey solution that puts the technical components in the hands of the technical experts.

Alternatively, broadcasters who possess the expertise, or desire full control of the site can purchase hardware, software and systems to build and manage their own *Searchable TV* Web site.

As with all programming, 'content is still king'. Broadcasters must be sure to host content that is valuable, not only that which is popular and generates Web hits and advertising. The Internet has clearly demonstrated that high-value, specialized content is in equal demand as entertainment-oriented content.

Approach for Processing Broadcast Content in Real-Time

The combined real-time video cataloging, indexing and search are in development at ISLIP Media in conjunction with Carnegie Mellon University, and a contract with NIST. These real-time techniques include: 1) acoustical and environmental robustness for speech recognition, 2) sufficient semantic understanding for indexing and retrieval, and 3) reduction to real-time capture and analysis.



Real-time Speech analysis

Multimedia indexing and browsing does not require perfection in speech accuracy. Even relatively poor recognition accuracy (measured by average word accuracy) often delivers useful multimedia document retrieval. Studies show that even at 50% word error rates in the speech recognition generated transcript, retrieval rates are 80%.

- Typically, content words (noun, verbs, adjectives) are longer and better articulated than short function words (articles, pronouns, etc.) and generally result in better recognition accuracy. These content words carry most of the meaning, and good retrieval accuracy at the segment/clip level can often be maintained.

- Retrieval accuracy does not depend on single words but on joint evidence from several words in a scene or segment. Since errors are generally semantically meaningless, correlation between the semantically meaningful and pertinent words can be exploited to effectively describe the content of a story.

- The product of the multimedia information retrieval is a set of ranked ordered documents or document segments. The final judge of relevance and usefulness is the human being, who can quickly discard irrelevant material.

Nevertheless, it is clear that improved performance at the recognition level is necessary to further enhance speed and relevance of indexed documents. These particular activities aim to improve segmentation of news programs into differing acoustic segments (correspondent, announcer, background music, telephone, talk show, etc.) and aim at raising recognition accuracy within each segment. This background will provide evaluation methodology and basic recognition strategies and algorithms for this project.

Research in language models indicate twenty to thirty percent improvement in accuracy may be realized by dynamically adapting the vocabulary based on words that have recently been observed in prior utterances. Often there are documents related to the meeting at hand which can be used for system training by adding specific vocabulary while limiting its extent. In combination, these resources can provide valuable additions to dictionaries used by the recognizer.

Image analysis

The application of image analysis has already been developed and is being continually modified to: (1) detect scene changes when people or subjects shift, (2) object tracking through optical flow analysis, (3) face detection and matching, (4) name-to-face correlation, (5) video-OCR of regular text on whiteboards, and (6) video-handwriting recognition. Considerable research must be conducted in these areas with respect to their algorithms.

In order to reduce their indexing time and space requirements, their sensitivity, or performance, needs to be understood as a function of the number of features or vectors with which they deal. This requires both a formal analysis and a large set of experimental studies to evaluate and validate. This analysis will augment the companion work in algorithmic parallelization to be conducted by Lawrence Livermore Lab computer scientists.

New research needs to be applied to the interpretation of hand-written text and graphics. Handwritten character recognition has been an active area of research for several decades. Using taxonomy of local, regional, and global temporal clues, which are often found in hand-written samples, provides a comprehensive understanding of the handwriting signal and a detailed analysis of stroke and sub-stroke properties [Doermann91].

Conclusion

The tools and technologies exist today for making near real-time *Searchable TV* over the Internet a reality. Indexing, cataloging, browsing and searching tools are the core engines which enable this reality. Truly Real-time *Searchable TV* is on the near-term horizon via advancements to these core engines. These engines will enable instantaneous, searchable access to currently airing broadcast programming as well as a continually-generating historical broadcast history/archive.

Acknowledgments and Further Reading

For further information, please refer to the ISLIP Media Web Site located at www.islip.com or contact Mark Juliano by phone at (412) 288-9910 or by fax at (412) 288-9905, or email at mjuliano@islip.com. Additional information can be obtained at the Carnegie Mellon University Informedia Web site.

ISLIP Media provides and operates the world's first truly searchable TV site at www.mediasite.net, as well as systems, software and technologies to create searchable Media Sites (video and audio focused Web sites). ISLIP serves customers in the broadcast, film, advertising, corporate, government and stock footage industries.

Author Biography

Mark Juliano is the president and chief executive officer of ISLIP Media, with over 14 years of start-up, marketing, sales, and general management experience in the high technology industry. Prior to joining ISLIP, Mr. Juliano held executive positions at AVIDIA Systems and FORE Systems, industry leaders in high-speed networking. Prior to FORE, Mr. Juliano held a variety of management positions at N.E.T., T.T.C., and IBM/ROLM. He holds an undergraduate degree in engineering from Princeton University and an MBA from Stanford University. Mr. Juliano has authored numerous articles and papers for publication in the SMPTE Journal, NAB Proceedings, Content Watch, Byte Magazine, Telecommunications Magazine, among others.

The Convergence of TV and PC - An Arranged Marriage

Jurij Paraszczak, Frank Schaffa, David Seidman,
Richard Chernock, Paolo Dettori, Steven Depp
IBM T.J. Watson Research Center
Yorktown Heights, NY 10598

Abstract

Although engineers have invented a plethora of technologies which awkwardly attempt to fuse television components to PC systems, it is the Authors' belief that this marriage is unlikely to proceed smoothly without considerable "Sturm und Drang". Today's television viewers expect immediacy of response, unobtrusiveness of interaction and audio which is ubiquitous. Although the next generation of viewers appears to be using this medium in a much more interactive way, with a shorter attention span which can be distributed over a number of simultaneous sources , we believe that the PC and the TV will maintain their distance as separate elements but will evolve with each other.

We will review the current state of the art in the merging of these technologies and describe some of the work at IBM Research which is focussed on providing an interactive yet unobtrusive experience for the TV viewer. We will also propose how we see the evolution of the PC and TV in a highly connected environment and try to predict how PC's and TV's will be used in the home and business in 2005.

Introduction

Ever since the late 1980's and early 1990's, Interactive Television has been a goal of many businesses, engineers and financiers, looking to reap the rewards of delivering content and services to users on demand. Much hyperbole in this marketplace owes its origin to the trials resulting from these goals and although none of those efforts generated a financially sustainable business, they started the present effort to converge the PC and TV by a large number of different interests. It is instructive to list these different interest groups and their goals to ascertain where this convergence is headed and what the likely outcome is going to be. Before doing so, it is useful to review some pertinent demographic data which may help explain some of the background of these motivations.

The total number of Worldwide consumers connected to some form of cable, terrestrial digital TV or digital satellite network is around 180 million or so. Most of these consumers are not connected on a 2 way network through these systems, but they do enjoy high speed, full motion video supported by synchronised and omnipresent audio. They think nothing of punching a few keys on a remote and obtaining the instant gratification of seeing another channel almost instantaneously. Some of them also have the capability of paying for a specific event which has a specified broadcast time.

With the exception of the commercially available video on demand service in Hong Kong, very few of these users yet possess the capability of interacting with the content that is displayed, without accessing the Internet on the television through a variety of means, most of which rely on a slow telephone line for the Internet connection, or through low bandwidth data transmitted through the Vertical Blanking Interval (VBI). In Europe, several companies have started offering a range of interactive services delivered over satellite networks. These include TPS and BSkyB.

Today, the number of people using the "Internet" is a number which is highly debatable, since it is difficult to determine the exact number, but pundits place it at anywhere from 60 to 100 million worldwide. These users can access any kind of content that they are allowed to, but they all need a PC, with its attendant hardware and software maintenance issues, a slow connection which cannot support even 2% of the bandwidth of a Television signal and almost total lack of audio. At the same time, backbone providers are increasing the network bandwidth by factors of 2X and 3X per year to try to support these users, but it is still not sufficient to enable the direct delivery of high bandwidth, full motion and audio content on demand. Clearly there is much to be said for a medium which can provide the virtues of both characteristics and this is one of the reasons that there is a push to create a converged platform.

In our view, the chief driving forces which are pushing to create some level of convergence were rooted in the data industry, which sees these 180 Million "unconnected" users as a potential source of revenue for interactive services and the sale of hardware and software. Today, they are also joined by the media industry and the distributors of this content who understand the value of adding interactivity and are working hard to define this interactive marketplace.

Thus a cursory list of the interested parties includes:

(a) Engineers

Many engineers innately see the integration of function as virtuous and the notion of creating a novel device which possesses both interactivity and full motion video as a technical goal. This notion has been difficult to substantiate since the business cases are not clear for this emerging opportunity. Many of the earlier trials were technology extravaganzas, which provided an understanding of the customer response to the services, but were far from profitable. Nonetheless they did seed the initial players with ideas and directions which are only just coming to fruition.

(b) PC Software Companies

Although PC unit shipments are continuing to rise the home penetration rate is slowing to about 45%-50% of homes and slowly moving toward 60% in the US European penetration rates are much smaller, only about 10%. Vendors who make their profits from the sale of operating systems and applications running on these PC's see the 180 Million users, who exist with nary a single copy of their products, as a source of future revenue. They are frustrated by the slow emergence of 2 way high bandwidth networks and this frustration has manifested itself in the formation of many efforts to support 2 way cable through modems and set top boxes(STB's), through xDSL efforts and through acquisitions and alliances destined to push the operating system and software into these new devices.

(c) PC Hardware Companies

Most companies in the PC Hardware business operate on very thin margins with the exception of Intel which gains the lion's share of the profit from each PC sold. The consumer electronics companies operate on margins of around 4-10% NEBT compared to around 25% for Intel. Clearly use of the existing high performance PC processor within a converged PC/TV device which will spread to the 180 millions users is an appealing target if the profit margins can be maintained. However the profit margins for these Set top Boxes are somewhat different and may well change the financial model for this area.

(d) IT Network Providers

Here, both the companies which own the underlying data networks and the companies which sell into these networks stand to gain the most in first delivering interactive Internet content and then delivering broad band content in full motion. Although it is true to say that Telcos are capable of providing this transport, many of them are either too slow to respond to the marketplace, restricted from it by the FCC or simply wondering what to do about their current deregulatory issues. It is likely that emerging companies such as Qwest, Level 3 and also existing companies such as

AT&T who are focussed in this area will become the transport providers for this data network.

(e) Media Owners and distributors
Although the efforts to create a converged platform have really been driven from the above groups, the media owners and their cable subsidiaries have quickly understood the value of providing an interactive service which could generate additional revenue. Unfortunately, the applications which generate the additional revenue have not been explicitly defined and so there is considerable exploratory activity looking at the impact of possible services upon existing business models. One of the most compelling at the present time is the capability of creating an instantaneous electronic purchase by the consumer predicated upon the content which attracted the consumer in the first place. Others such as video and data on demand are also being considered.

(f) Consumer Electronics Companies
Amidst this change, companies that manufacture televisions and PC's are looking at the potential new revenue generated by unit sales, through the emergence of the potential converged platform. As digital television evolves the market forces which try to drive down the cost of the end devices will be supported by the increasing capability of the technology (see below) to provide a whole new set of services and applications which can generate revenue. Many set top box providers and related companies maintain a proprietary approach to the systems they sell into the marketplace. The mandate by the FCC to allow consumers to hook up a set top box they purchased anywhere in the US into a cable network, begins to erode some of this proprietary nature and may ultimately commoditize the market for these boxes and the networks that support them. Nonetheless some companies are hanging on to their encryption technologies (the so called conditional access) as a control point into these networks. To-date this has been an effective approach to keeping out competition. Convergence will confuse this position.

Convergence
We believe that convergence of functionality, in some form and to some extent, will occur. In our opinion the most likely form it will take is the incorporation of interactive and data-related functions into the television and video viewing environment. In the following we will describe many of the factors that are expected to contribute to the likelihood of convergence in the foreseeable future. The most significant of these factors are:

1. The advance of technology contributing to more affordable digital consumer electronic devices
2. Technological advances in digital video and television including improvements in resolution, functionality and connectivity
3. The growth of e-business, on the Internet and elsewhere
4. The adaptation of content to these converged platforms in a manner that will minimise users confusion with the enriched content
5. The rollout of the digital broadband networks

We will briefly describe some of the problems which may affect convergence, some existing examples of convergence, and elaborate on one of the projects at IBM Research which supplies some convergent functions to digital video.

Technologies – PC's, TV's and Set Top Boxes

Perhaps one of the most significant issues which will enable the creation of a converged platform is the base technology itself. Disk storage, memory and processing power have advanced considerably since the early 1990's. For example, in 1990, the average density of transistors on a microprocessor was around 1 million for a microprocessor device. The standard RAM chip was around 4 Mbits. By the year 2001, a standard leading edge microprocessor will contain about 25 million transistors for the same price as was paid in 1990 for 1 million transistors and the leading memory technology which will be available will contain 1 Gigabit of storage on a chip. Over the same period of time a ~$200 disk drive has increased in capacity from ~200 Mbytes to a projected 30 Gbytes - enough to store 4-5 standard definition MPEG2 digital movies. Clearly this technology is enabling a wide range of functions at a consumer price point (~$300 for the whole set top box), which now allows the decoding and storage of interactive data at reasonable cost. Above all, this technology will allow companies to embed more function at the end user device, provide more intelligence in the search, indexing and storage systems that deliver this data, and ultimately it is likely to make the boundaries between the proprietary systems which deliver the content to the end device (which today contain proprietary hardware and software) and open networks, paper thin. However, access to content will probably continue to rule the business case for the services delivered to this end device.

It should be borne in mind that the most impressive technologies will only turn a profit for their providers if the general public is interested in their utilization. Many of the interactive television trials have shown that users will respond positively to interactive, PC-like features only if these features

- Add only marginally to device cost,
- Are not difficult to learn or operate,
- Are not obtrusive or distracting,
- Allow the user a high degree of control over what appears on his/her screen
- Are integrated with, and provide an enhancement of, the viewing experience.

Digital Television

Broadcast digital video is an emerging technology. With the exception of the DirecTV® satellite system, the digital broadcast systems being implemented follow a variety of standards including:

- SMPTE (Society for Motion Picture and Television Engineers) (http://www.smpte.org)
- MPEG (Motion Picture Experts Group): Video & Audio encoding and Transport (http://drogo.cselt.stet.it/mpeg/)
- DVB: Digital Video Broadcast (initially a European standard for digital video) (http://www.dvb.org)
- ATSC: Advanced Television Systems Committee (initially an American standard for HDTV) (http://www.atsc.org)
- SCTE: Society for Cable and Television Engineers (http://www.scte.org)
- DAVIC: Digital Audio/Video Council (http://www.davic.org)

These standards bodies are specifying audio and video compression, transport, data encapsulation and interactivity.

The hardware required for receiving digital television in these standards' formats is significantly more complicated than that for analog television, but at the same time offers more capabilities for interactivity. New components include those necessary for dealing with the compressed digital bitstream: demodulators, MPEG2 transport demultiplexors, Video and Audio decompressors, microprocessor, graphics generator and NTSC or PAL encoder. The microprocessor and graphics generator offer the capability to execute local applications on the receiving device, which can interpret data in the broadcast stream and provide interactivity between the user and the receiving hardware. In addition to the above, fee services (such as"pay-per-view") will require smart cards and conditional access/decryption hardware.

Software is required on the receiving device – at a minimum to parse the MPEG transport stream for Program Specific Information, set up the demultiplexer and decoders and provide a useable interface for the viewer (typically an electronic program guide (EPG)). Existing digital television systems have a wide range of software complexity – from the simple systems as described above to more "computer-like" systems comprising RTOS (real-time operating system) and middleware layers.

A number of television-related data broadcast systems have appeared for the analog broadcast domain including a service which has been offered in Italy by IBM for 6 years which uses the terrestrial TV signal supplied by RAI to distribute data securely to businesses called DataVideo. DataVideo operates today on RAI-1 and RAI-2 public channels and provide small businesses equipped with a PC and a DataVideo receiver card with a huge amount of data from more then 15 Information Providers. The system utilizes VBI technology to deliver data over the analog TV signal; furthermore IBM has implemented a full protocol stack that handle multicast, single user addressing and encryption.
Other examples of embedded data include the close captioning system which uses a scan line during the vertical blanking interval (VBI) to transport data, or the teletext system in Europe (again using VBI).

Many broadcasters and cable system operators have seen the potential of digital television to open a market in advanced TV services. The data bandwidth available in a digital TV system is orders of magnitude greater than that available with VBI, and this suggests the possibility of data-intensive digital applications to be run entirely on the users' equipment or in interaction with head-end equipment. One example of such a service is an "edit once, view many" application for content blocking. The digital video to be broadcast, in such an application, is edited with annotations marking words or images in the video which may be objectionable to some viewers, and a degree of control is assigned to each. These annotations are broadcast synchronized with the video. At the user's STB, the viewer's profile specifies the control level which the viewer wishes to block, and this is performed automatically by the STB, independent of the levels chosen by all other viewers of the broadcast. Another example is an interactive advertisement, in which a file (for instance a music video clip) can be embedded into a commercial and stored in the STB, if the viewer so desires. It is currently unclear exactly what form these new applications will take, and how much users (or advertisers) will be willing to pay for them.

Examples of Converged Systems

There are a range of approaches which result in the creation of a converged device. These include:

1. PC's which receive video over IP networks and decode it locally on the screen using a software or hardware decoder. The user interface is PC centric, namely it requires the use of a keyboard and is not focussed on delivering entertainment per se. Ancilliary data is delivered over the same IP transport. This is the most common approach used today. The definition of a PC here is a device which runs an operating system used to accomplish a wide range of tasks, from word processing to multimedia display, and whose main focus is not dealing with MPEG2 transport based content. In this case, even though video may be encapsulated as MPEG packets, it is transported over IP networks.

2. PC's which contain some kind of RF tuner card which delivers video to the display adaptor or to an external TV and whose interface consists of a tuner interface akin to a TV set. Data is delivered over either the Vertical Blanking Interval for analogue networks, or embedded in the MPEG packets as private data. Examples include Intel's Intercast and DirecPC. In the case of the former, ancilliary data is presented to the user, being extracted from the VBI and used for a variety of purposes.

3. TV's which use TV/PC monitor as a display device for a PC connected to it. All TV functionality is enabled and there is only a mechanical integration of the TV and PC functions. The Gateway Destination XTV® is an example of such a design.

4. STB's which connect to a TV through which a separate IP packet based channel delivers Internet based content to the video display device in the Set Top Box. Any Internet site can be reached and has to be transcoded to adapt to the restrictions of a TV. Examples of this approach include WebTV. The relationship between the TV content and the

	Television	PC
Usability	Viewing occurs at a distance	Viewing occurs at close range
	Viewing experience often shared	Viewing experience dedicated to one person
	Input device simple	Input device complex
	Device associated with entertainment and relaxation	Device associated with work, communication and information gathering
	User has no notion of an operating system	User forced to deal with operating system level functions
	No customisation of experience available	User controls appearance of environment
	Predictable content and content access	Unpredictable and chaotic access to content
Device Functionality	Access to network content instant (1-2 seconds delay)	Access to networked functions slow and dependent on function
	Interactivity limited to a range of functions	Interactivity spectrum very rich
	Sound is always available and intrinsic to device	Sound is an afterthought and limited
	Very high bandwidth available instantaneously	Bandwidth limited to network connection (usually dial up modem)
	Content shared across millions of observers	Content used individually
	Focus on analogue or MPEG2 based content using MPEG or analogue transport	Focus on digital content within device and digital content delivered over IP
	Limited storage (today)	Storage freely avaialable
Network Functionality	Network delivery optimised for broadcast environment	Network delivery optimised for individual content access
	User response can only be tracked indirectly	User response can be tracked individually
	Network managed centrally and focussed on broadcast center	Network management system from server to end user
Business Case	Supported by advertising and subscription and in limited cases, by government license fee. Hardware sales totally separate from network delivery	Supported by sales of software, hardware and applications. Network access supported by subscription. Advertising currently a small component.

Internet content can be coupled, although this relies on separate agreements between the content provider and the broadcaster for maximum benefit. Again the user interface is very PC centric and relies on a keyboard for maximum interactivity.

5. STB's and TV's which use data embedded in the transport stream (either analogue or digital) to enhance the user experience on the Television. Examples include allowing the user to respond to an unobtrusive icon to gain access to supplementary information. Depending on the sophistication of the network and the capability of adding a return channel to a central processing center, the content lead a specific user to an interaction which is localised to just that user. For example, the user may complete an electronic purchase, directly from the Television. The broadcast channel contains the ancilliary data, the return channel provide a capability to customise the response.

Clearly a range of hybridisation can be achieved which draws upon the different characteristics of both the Television and the PC. The following table attempts to make the specific differences explicit and reflects the current state of the art.

Based on an analysis of this table we believe that there are several key trends emerging which are attempting to hybridise the TV and PC. We list them in 2 categories, those which we believe will occur in the next 2 to 3 years and those which will not.

Trends which we believe will result in real business opportunities:

- The massive amount of advertising revenue being spent on broadcast will continue to be mainly be focussed on participants who engage content which is broadcast. This does not mean that the user cannot be targeted with more specific content, but it will mainly be delivered over a broadcast environment. This approach will enable the current stemming of the tide of decreasing broadcast advertising revenue, since users will be able to receive content which is more relevant to their interests Any attempts to segment this revenue into individual "on demand" channels, such as that available on the Internet will take a lot of time to prove themselves to the advertisers because of the huge infrastructure requirements. Today's primary example of a commercial video on demand service, in Hong Kong, is taking time to prove its commercial value in the marketplace.
- Systems which efficiently manage the creation and accurately handle indeed storage of content, rights management, payment and monitoring of user response will enable the "broadcasters of the future" to maximise revenue and enterprises to efficiently manage their advertisement spending. The present inefficiencies, which have resulted in broadcasters' disputes with rating agencies and in delays of payment and verification of advertising have stifled the progress in building new markets for content owners and distributors and have driven consumers to the exigencies of modem connected to an unstructured Internet, where although they are in control, they wait and cannot easily find what they want.
- The ability of a user to gain instant gratification for the purchase of the object of their desire will be dramatically enabled by the inclusion of electronic commerce capabilities linked to the broadcast content. This will be the killer application for the converged device which we believe will use the capability of broadcast content to reach millions coupled with the pervasive protocols and programmers currently using the Internet to develop new businesses. This approach is likely to disintermediate the current purchasing channels such as mail order and retail, but many of these parties have already understood this issues and are rapidly making their catalogues and services available on line.
- PC like devices which allow individual users to interact using a keyboard and access thematic

purchasing channels will emerge. The ability to rapidly find and purchase articles will be key. Thus the broadband networks which are being deployed to the end user, by companies such as @Home or AT&T are vital in this regard and will enable them to cut special deals with their advertisers to directly tap into the high bandwidth of their network.

- The delivery of electronic media, such as video, audio, games or software in a secure process which assures the content owner that payment is always made for services rendered while assuring retailers and intermediaries of their own revenue will emerge as a new business opportunity. Consumer acceptance will be determined by the usability and seamless nature of the interface.

- The use of disk storage on the TV, which can be controlled by both the advertiser and the end user will provide two critical levels of functionality including the ability of the advertiser to deliver supplementary and compelling information and the ability of the user to automatically store content which is broadcast as supplementary information which reflects the end users interest.

- The difference between a TV and PC will end up being determined by the degree of privacy that a user expects in using the device. The shared device, exemplified by today's TV will have many of the same components as an interactive terminal exemplified by today's PC, but will typically not be expected to be used by an individual. who expects privacy during its use.

- The use of entertainment delivered over IP networks, utilising pervasive standards such as Java, HTML and open interfaces is an interesting issue which will probably develop rapidly once the individual bandwidth to and from the consumer or business exceeds about 1.5 Mbits/second. We will see devices akin to TV's emerging using these standards but hiding all of the complexity of the operating system in the manner that a television set top box does today

Trends which we believe will not result in real business opportunities:

- The notion that a user sitting on a couch using a keyboard perched on the edge of their knees using the TV as an output device for today's PC is unlikely to create either a new business or to appeal to a broad audience. We believe that this will remain a niche marketplace for as long as the providers are unable to reduce the chaos of the Internet for the average TV viewer and leverage the entertainment that is delivered over the broadcast network. We believe that users use these shared viewing devices for entertainment and will continue to do so.

- Devices which try to deliver entertainment content to a PC which manages both generalised computing and provides a connection to a display are unlikely to succeed.

Development of converged applications and systems

We have been working to develop MPEG based systems which allow users to interact with embedded content within a video stream. Our focus is to develop applications which are of perceived users value which take advantage of user interactivity with and without a return channel. In both cases, data is embedded in the private data stream of an MPEG2 transport and delivered to an intelligent MPEG capable device which allows the user to interact with the content in a natural manner, such as that enabled with a simple remote control. We have attempted to classify this data to help clarify for ourselves the different business and technical opportunities presented in the different domains. Figure 1 below, shows this classification

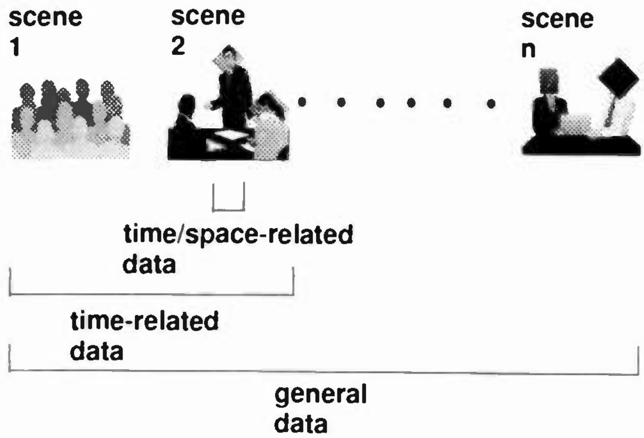


Figure 1 - Classification of Multimedia Objects

As can be seen from figure 1, several different classes of data can be created which are related to the specific content including:

- General Data which contains information pertaining to the whole video segment
- Time related data which pertains to specific portions of the video segment
- Space related data which pertains to specific objects within the video segment

Clearly these classes of data can be combined, for example information about a specific actor (space related data) may appear not just in a specific scene, but will persist throughout the video if called to use. At IBM Research, we are using several classes of internally developed tools to identify scenes, objects within scenes and the speech within the scene to help classify the content for later indexing. These tools enable us to more rapidly author the content for distribution. At the same time, we are using IBM's **HotVideo**(http://www.software.ibm.com/net.media) tools, developed for the hyperlinking of video objects for Internet Browser applications enabling the more rapid repurposing of content from Internet to Television. Combining the **HotVideo** authoring tools and our internally developed MPEG2 multiplexing technologies as shown in figure 2 allows us to rapidly develop interactive applications for the TV environment. The MSC card referred to in the figure is IBM's internally developed MPEG2 multiplexor capable of 100 Mbits/second output.

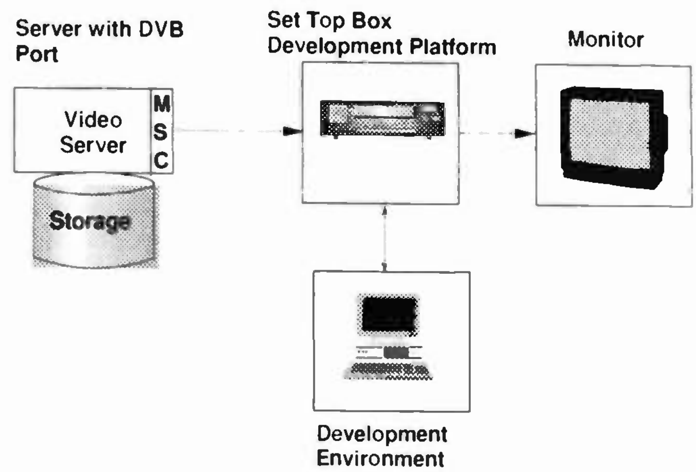


Figure 2 Interactive Application Development Environment

The development environment has been used with a variety of different Set Top Box technologies and allows us to develop a wide range of different applications, for example those which allow users to get access to supplementary information such as video-catalogues, those which allow users to choose different modes of content display, for example filtering out objectionable speech and replacing it with less pernicious content and applications which use a return channel to enable a transaction such as purchasing.

At the 1998 Western Cable Show in Anaheim, CA, IBM and Scientific Atlanta demonstrated a complete system based on the Scientific Atlanta Explorer 2000 Set Top Box and the IBM Net.Commerce suite to enable users to securely purchase content at the click of a remote button over cable networks. We have demonstrated a similar capability based on IBM's set top box chipset where both purchasing and the user's current stock portfolio overlayed on the existing video stream showed how content could be personalised using a telephony back channel.

We are continuing work in this area to leverage the ubiquitous protocols and programming prowess of many individuals focussing on the Internet for the management of the interactive applications, while using the compelling nature of the entertainment content to invite the customer to participate.

Conclusion

Although it is unlikely, in our opinion, that today's TV's will merge with PC's producing a device which requires a keyboard for interactivity and a television monitor for display, we believe that the two devices will remain largely separate, being defined by the degree of privacy and ease of use that the customer expects. There will doubtless be convergence of the protocols which are propelling today's 2 way interactivity over the Internet with the entertainment value of the Television. As data networks increase in bandwidth to the point where they can support MPEG2 video rates for interactive purposes, it is likely that IP will begin to not only deliver user specific interactivity, but will also be put into service to deliver video encoded material in the format which provides the least expensive decoding at the client device. At this point (which we envisage to be about 3-4 years from today), it is likely that local servers delivering video, network access and other services such as telephony will be available for installation at the customer premises.

We also believe that the main driver for this convergence will be electronic commerce, which will allow users to satisfy their purchasing whims immediately. Technologies which enable this process, such as smartcards, which allow stored customer information to be rapidly transmitted, and storage devices, which allow the user to review the details of an item to be purchased will be instrumental in growing this marketplace.

So You've Gone Digital ... Now What?

Dr. Dov Rubin, NDS Americas
Newport Beach, California

ABSTRACT

The future of broadcasting is digital. The benefits are immense. Consumers will, no doubt, wish to benefit from the latest technological digital TV wave. Who will actually pay for all this technology? Will consumers get what they want? Consumers have been brought up on the concept that "TV is Free" (or, at least, nearly free). Without the revenue to fund the programs which incorporate digital TV technology, broadcasters will be reluctant to invest in producing state-of-the-art material. Viewers, in turn, will have a hard time finding anything worth watching. It is going to quickly fall upon the shoulders of the broadcasters, content owners, and advertisers to come up with creative ways of generating revenue to break this chicken-and-egg cycle. This discussion presents a myriad of enabling technologies and creative ideas to show how digital TV can pay for itself which, hopefully is both what the consumer wants and what the consumer needs.

In true digitally compressed tradition, I'd like to take these few minutes to give you some digital "food-for-thought". Let's examine some of the many revenue enhancing opportunities. I'll start with a few we all know.

1. The first basic opportunity is **Digital Compression** which enables a proliferation of channels, offering new revenue opportunities such as NVOD, re-packaging of popular back catalogue programming, niche programming—such as educational, and health & fitness programming, etc.

2. For the first time, **terrestrial broadcasters** will be able to have a relationship with their customers. Digital gives all broadcasters the chance to interact through a set top box and offer such things as cross marketing opportunities, pay-to-view for special events, season tickets for premium programming, along with a highly valuable opportunity to capture precise data on the viewing habits of their customers.

3. Digitization can make the scheduling and day-to-day operation of a **multi-channel broadcast operation** more efficient and operate with a lower cost-base. Technologies such as MPEG Splicing along with Store & Forward also make the merging of localized programming into the national network much more efficient and manageable, leading to revenue earning opportunities from local advertising.

4. The complete digitization of **program production** from the camera through to edit room makes it possible to produce programs more cheaply. This means that it will be possible to economically produce and broadcast niche programming. Such programming will accrue higher per viewer advertising rates as it delivers a much more targeted audience and will encourage TV advertising niche interest companies that would not have been able to justify TV advertising rates.

5. Another area is one of **data and multimedia services.** I expect to see the wired home of the future purchase much of its entertainment/information content (music, computer games, electronic newspapers, etc.) via data broadcasting services offered by TV broadcasters.

6. Digital gives broadcasters the opportunity to develop truly **interactive programming and interactive advertising**.

7. Digitization allows the possibility to integrate **local storage** (hard disk/write-able DVD) into the digital TV. Local storage combined with some knowledge of the consumer within the TV, means that advertising can be targeted at the consumer. Local storage also means that

programs can be automatically saved for viewing at the consumers' convenience.

Who will benefit from the new digital age? I believe that the move towards digital gives revenue enhancing opportunities for all levels of the broadcast food chain, as well as giving consumers an enhanced viewing experience.

I am sure that the next few years will see many additions to the above list. For enlightened **broadcasters** who make the right decisions, who are light on their feet, and are not afraid of radical changes, digital broadcasting offers many opportunities. Going digital increases revenue and can achieve a healthy return on investment.

Content owners and creators have, I think, little to fear from digitization. Owners of quality content should find that having more channels to reach the consumer cannot have anything but a positive affect on their revenue earning opportunities. If you own a back catalogue of high quality content or valuable content brand names, then the digital age should be highly profitable.

The digital age also offers many opportunities to **advertisers**. Although many say the advertising dollar is saturated, I believe that technology will have to work a little harder. The advertiser will still be able to reach the required number of viewers and, in fact, will now reach them with far greater precision. After all, no broadcaster wants to kill the goose that provides the cash to make the golden programs. Presently, TV has an absolutely enormous advantage over other media in that for many products and services, TV is simply the only media that can communicate a complex message to millions of consumers. With the addition of interactivity, the effectiveness of advertising on TV and on other media is set to rise yet further.

Driving Technologies

Broadcasters will spend billions of dollars over the next ten years in their move to digital and they will need to make the most of their investment. Some of the key technologies that can both increase revenues and reduce costs are:

Local Storage in the Digital TV/Set Top Box

The capacity-to-cost ratio of local storage is improving so quickly that the digital broadcast of television and data can provide almost all of the benefits of the information superhighway at a much lower cost to the broadcaster and consumer. In effect, local storage combined with knowledge of the consumers' interests, either gained directly from the consumer or derived by learning from a consumer's habits, can give most of the benefits of real-time interactivity. The consumer will surf or watch electronic content pre-served to the local storage device within the digital TV. Smart TVs will decide which program might be of interest to the viewer for storage and later viewing.

Interactive Programming and Advertising

The first adoption of interactive television will likely come from advertisers. For advertisers, interactive TV means a new way to reach their targeted prospects. New opportunities will arise to sell goods and services with special offers, linked information and premiums over the broadcast network using secure electronic-commerce provided by the broadcast systems' Conditional Access. An added benefit, and a very important one for marketers, is the bonus of virtually instant consumers' feedback that is actually quantifiable.

Independent research commissioned by NDS shows that the **types of interactive ads** that are expected to be most popular include *product information ads* that allow viewers to receive detailed information about a product, *incentive ads* that reward viewers for watching, *targeted ads* that target specific market cross-sections, *quiz ads* that increase viewing by running contests during an ad, and *impulse purchasing ads* that enable consumers to make purchases directly from the ad.

Data Broadcasting

Data broadcast systems allow broadcasters to provide additional data services and revenue streams for little extra investment. Broadcasters maximize their precious bandwidth by broadcasting data in off-peak hours and when bandwidth becomes available during peak hours, or from bandwidth gained via the use of statistical multiplexing techniques.

The deployment of data broadcasting to the PC is a commercial reality today. In the next few months and years, data broadcast services will also transmit to consumers' set top boxes and digital TVs. The combination of enhanced services to the subscriber such as low cost access to most of the services of the Information Superhighway, together with the ability to deliver electronic goods such as music, films, and computer games to the consumers' home many times faster and many times cheaper than the Internet and traditional distribution methods, makes data broadcasting attractive to both broadcasters and consumers.

Conclusion

In retrospect, perhaps the title of this talk should have been "Making Digital Television Pay for the Broadcaster, the Content Owner, the Advertiser, and Most Importantly, the Viewer"! The fate of all these groups are so intertwined that Digital TV really does have to work for them all.

In summary, the Digital Age will be a boom time for broadcasters who understand the new opportunities offered by digitization and are not afraid of radically changing the way they make and market TV shows and how they use and sell the advertising slots in the middle. The opportunities offered by digital TV will offer a healthy return on investment for digital broadcasters and great opportunities for the other stakeholders. It is a challenging time, but one full of opportunities for all involved.

ATSC Digital Television Data Broadcast and Interactive Services

Digital TV Application Software Environment and Data Broadcast Protocols

Aninda DasGupta
Philips Electronics
Briarcliff Manor, NY

Abstract

Digital Television allows delivery of data and interactive programming to consumers. In this paper we present ongoing efforts in ATSC to develop technologies and standards for delivery of data over digital TV channels, and for development of a receiver-independent software environment to allow execution of interactive programs. We present a DTV receiver software architecture for decode and display of interactive programming content, and data broadcast protocols.

Introduction

With the advent of digital television broadcast, not only can viewers be provided services consisting of video and audio, but also data and textual information for interactive programming. It has become easier for broadcasters and service providers to distribute program-related and program-independent data along with traditional video programming services over digital media. We refer to these services as data broadcast and interactive services, or data services. Program related data services may be used to enhance audio-visual content. Examples are TV commercials that provide additional product information, or electronic coupons, or allow purchase of merchandise; news programs that can be customized to a viewer's specifications and which provide additional detailed news of the viewer's choice; documentary programs that help educate by letting a viewer peruse Web-based information; musical concerts that allow viewer interaction; sports shows that provide a player's statistics; etc.
Data services may also be independent of audio-visual programming. Examples of such services include customized stock quotes or sports and weather tickers. Program-related data services require that data and interactive content be authored, then combined with audio-visual content during editing or production. A broadcaster or service provider must then transmit this combination of audio, video and data content; such transmission must occur using standardized communication protocols and data formats so that all brands of receivers may receive them unambiguously. Further, content must be decoded and displayed by all brands of receivers, irrespective of what technological choices are made by the receiver manufacturer. Receivers have functions and features that must be made available, in a platform-independent manner, to all interactive and data content. All this calls out for standardization of content formats, communication protocols and access to receiver functions and features. A system architecture for broadcast data and interactive services is shown in Figure 1.

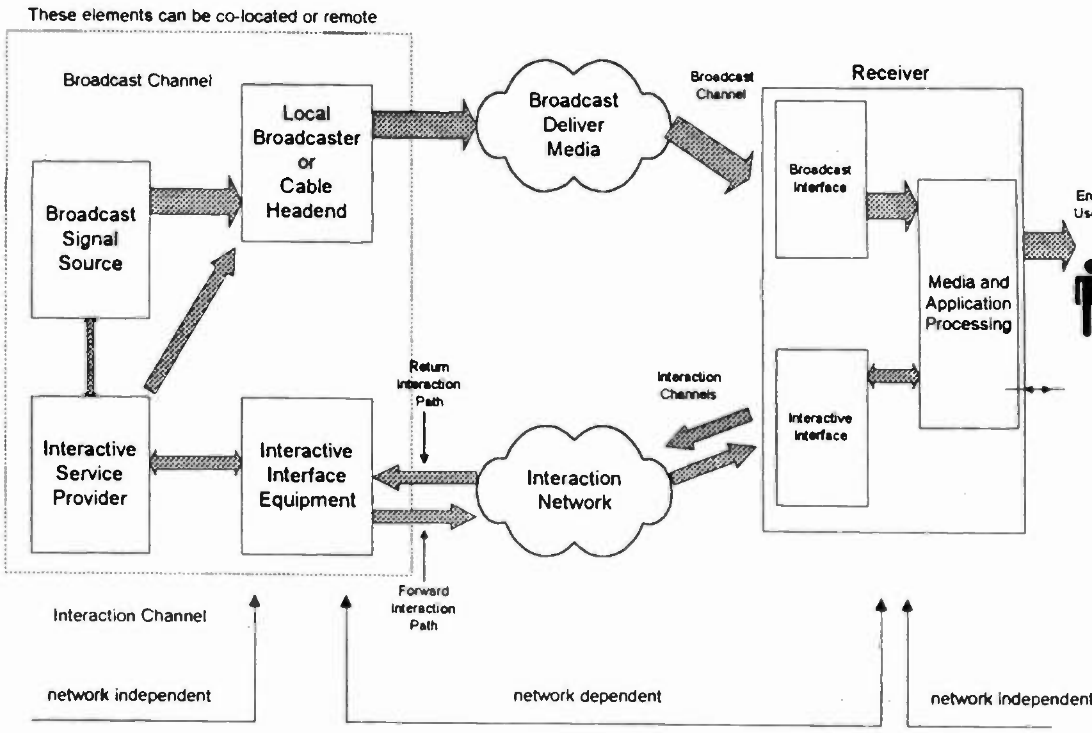


Figure 1.

The Advanced Television Systems Committee (7) is an international standards body that is developing digital TV technical standards. The T3/S13 Specialist Group of ATSC is developing standards for data broadcast protocols. T3/S16, another Specialist Group of ATSC, is developing protocols for interactive services. ATSC's T3/S17 is developing receiver software run-time environments, content formats, broadcast universal resource identifiers, etc.

In this paper, first we briefly present data broadcast protocols being defined by ATSC. Then we present the receiver software architecture. Next we describe ongoing work in defining how HTML and Java content will be integrated for authoring of interactive data services. Finally, we present how receivers may be categorized by profiles of features they provide to interactive content.

Data Broadcast Protocols

Work on developing communication protocols for data broadcast and interactive services continues in ATSC's T3/S13 (5) and T3/S16 (6) Specialist Groups. The basic transport of all packets in a TV channel is via the well-known MPEG transport protocols (1). Using this underlying transport, the following means for delivery of data are proposed.

DSM-CC Data Download Protocol

Carriage of data is possible using the DSM-CC data carousels (2). Data carousels can be thought of as virtual carousels in the air, achieved using repeated broadcast of data, and the receiver may extract data out of a modular unit in the carousel at will. The proposed ATSC use of DSM-CC supports the sending of data modules, asynchronous data streaming, and non-streaming data that is synchronized with associated audio and video that is also transmitted via the MPEG transport.

Addressable Sections

The proposed ATSC data broadcast specification also allows transmission of datagrams in the payload of MPEG transport packets by encapsulating the datagrams in MPEG Addressable Sections compliant with the MPEG-2 private section format. Thus datagrams may be delivered asynchronously.

Synchronous and Synchronized Streaming Data

A third alternative is synchronous and synchronized data streaming using MPEG Packetized Elementary Streams. Synchronous data streaming allows data and clock to be regenerated at the receiver into a data stream with strict timing relationships between data units. Synchronous data streams are characterized by a periodic interval between consecutive packets so that both the maximum and minimum delay jitter between packets is bounded. Synchronized data streaming has the same intra-stream timing requirements as does Synchronous data streaming. Furthermore, synchronized data streaming implies a strong timing association between data, audio or video carried in other elementary streams of the MPEG multiplex.

Data Piping

Data piping allows delivery of arbitrary user-defined data inside an MPEG-2 transport stream multiplex. Data is encapsulated directly into the payload of MPEG-2 transport packets, and none of the MPEG-defined data structures are used. In other words, all interpretation of the data bits application-specific.

DSM-CC Object Carousel

ATSC's T3/S16 Specialist Group is defining protocols required for interactive services. The proposed standard will include parts of the DSM-CC Object Carousel specification. This will allow organization of data in hierarchical virtual file systems so that interactive programs may refer to objects of data, peruse directories that describe organized objects, and retrieve objects from the carousel.

Return Path Protocols

Also included are protocols for the "return path" back to the service provider. TCP/IP protocols are likely to be proposed for this back channel to remote servers and systems. Work in this area proceeds.

Receiver Software Architecture

There are numerous options for the format of the content delivered for data services. Further, the application programs that need to be executed on digital TV receivers to deliver these services must also be able to take advantage of functions available on receivers. Access to these receiver functions must be independent of the hardware or software platform chosen by the manufacturer of the receiver. Receiver manufacturers also must know what content formats are likely to be delivered over broadcast media, and what receiver functions are likely to be required by application programs authored by broadcasters and service providers. Thus, some default content formats and a means to access receiver functions must be agreed upon by authors, providers of data services, and receiver manufacturers.

The ATSC T3/S17 Specialist Group on DTV Application Software Environment (DASE) is working toward standardising the software environment within DTV receivers that deliver data broadcast and interactive services to consumers. We refer to such receivers as DASE-compliant receivers. Below we describe the receiver software architecture for the ATSC Digital Television Application Software Environment (DASE). First, we explain some terms used in this document. Then we describe the key components of the architecture.

Terms

The reader is expected to be familiar with common computing technologies and terms like *operating system, programs, thread, function call, object oriented, methods, object class, software libraries,* etc. The following terms are used in this document:

Service Application (Application): A collection of several HTML files or Java class files, and associated data in files or streams.

Data Service (Service): A collection of Applications intended to be provided together as defined by the content provider. Each service has Service Information (SI) associated with it.

DTV Application Software Environment (DASE): The DTV Application Software Environment includes software modules that allow decoding and execution of applications that deliver interactive and data broadcast services. The standard environment allows service content and applications to be decoded and executed in a manner independent of the receiver's hardware and operating system.

Receiver Platform (Platform): The receiver's hardware, operating system and native software libraries of the manufacturer's choice.

System Services: The receiver platform provide application programs with various common functions that the applications may use to implement data services. We refer to these common services as the System Services.

Application Programming Interface (API): The application programming interface consists of software libraries that provide uniform access to System Services. APIs are often implemented using the receiver platform.

Plugin-type architecture: A software architecture that allows companion modules to be introduced in the receiver to aid in execution of applications and provision of data services.

Interactive and Data Broadcast services will be embodied in application programs (referred to as Applications), and associated data or multimedia content. The DTV Application Software Environment standard will allow data broadcast and interactive applications to be authored once, delivered over broadcast channels and executed over all brands of DASE-compliant receivers. Methods in which data services are selected, and how applications associated with such services are launched, are beyond the scope of this document.

Overview of the DASE System Architecture

The system architecture is illustrated in Figure 2 below. A receiver manufacturer may choose a hardware platform, an operating system and software libraries for implementing the DASE receiver. The DASE adds to this receiver platform several groups of functions, that may in some cases be embodied in implementation specific ways.
A data service consists of several applications and associated data. Applications may be authored in Java, or HTML, or in any of the content formats specified in the DASE standard. Associated data can be in application-specific formats that are understood by the application's code. Applications may use streaming content. ATSC T3/S8, T3/S13 and T3/S16's communication protocols shall be the means for delivery of a

service's applications, data and SI. Several management modules in the DASE-compliant receiver provide management of applications and their lifecycle. Other management modules provide means to allocate, control and revoke system resources. Still other management modules provide functions for management of on-screen elements, their layout on screen as well as temporal synchronization. Event management modules allow applications to register interest in events and to handle events with appropriate handlers. Application Programming Interfaces allow programmatic access to these and other functions in the receiver.

Application Execution Engine

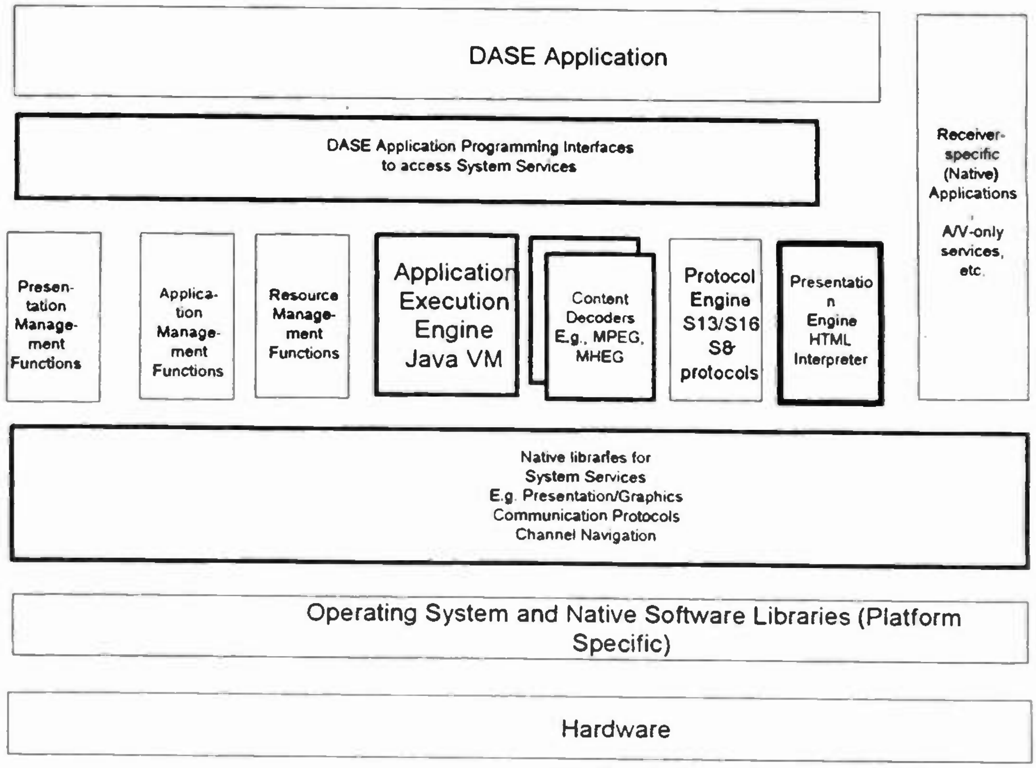
Part of the DTV Application Software Environment (DASE) is an Application Execution Engine. This engine, along with associated software modules, libraries and application programming interfaces, will provide system services to applications in a uniform, platform independent manner. The engine will act as an abstraction of the receiver platform and execute the procedural parts of the application. The Application Execution Engine will provide a uniform platform and support a plugin-type architecture which will allow other software modules to be brought in via a network to enhance the DTV Application Software Environment (DASE). It will act as the integrator of software modules in the DASE, allowing them to interact with each other. The Java Virtual Machine is proposed as the ATSC Application Execution Engine.

Application Management

As applications may be embodied in procedural code (Java class files), declarative code (HTML files), and content in other formats defined in the DASE standard, there is need for management of these applications in the DASE. Application management includes control of an application's lifecycle, and monitoring of an application for the purposes of reporting to privileged status queries. Lifecycle control includes loading and unloading of Java classes, or passing of HTML code to an HTML interpreter, destroying of an application, removal of system resources from an application's control, etc. Application management also includes maintaining a list of running applications; this may be in the form of maintaining a list of object references, or a list of HTML files being displayed, etc. The application management modules in the DASE are to perform all of the above functions. DASE applications may interact with the Application Management modules via Application Programming Interfaces. The Application Management modules are implementation specific.

Content Decoders

Content decoders decode application content of a specific format, e.g., JPEG, PNG , etc.



Presentation Engine

The Presentation Engine is a content decoder for (possibly dynamic) HTML content. It is privileged in that it works very closely with the Application Execution Engine. The Presentation Engine can display the results of other content decoders within the screen real-estate allocated to the Presentation Engine. The Presentation Engine and content decoders will cooperate with the Presentation Management modules (see below) to enable control of screen composition (layout of elements on screen) and temporal control (synchronization) between media elements. DASE compliant applications may use the Presentation Engine and content decoders for any on-screen display, as well as the Presentation APIs (see below) via the application execution engine. (We do not preclude authoring and execution of applications written to the receiver's platform, and which are not DASE compliant.) The DASE standard does not specify means to deliver applications authored for delivery over the World Wide Web. However, implementors who wish to add Internet capabilities to a DASE-compliant receiver may consider adding a JavaScript interpreter to the Presentation Engine. The interaction between the JavaScript interpreter and the HTML interpreter in the Presentation Engine is to happen with a well-defined and restricted Document Object Model such that security issues of JavaScript are resolved satisfactorily.

Presentation Management

Presentation Management requires arbitration between presentation requests from applications, and displaying arbitrated content on screen with precise layout control when necessary. Presentation management modules in the DASE allow presentation functions to be carried out in the receiver via presentation application programming interfaces (using Java code), or via the Presentation Engine (using HTML code), or via content decoders (using content in the format of content decoders listed in the DASE specifications). Presentation Management modules may be implemented in receiver-specific ways.

Presentation Containers

A container is an object that displays screen elements, and maintains a stacking order of elements; containers themselves may be stacked. An application requests container(s) from the Presentation Management modules in which to populate its on-screen objects. It also requests the Presentation Management modules to display its container(s) on the screen. The Presentation Application Programming Interfaces defined in the DASE standard have a container model and application code written in Java make use of these. Application code written in HTML expect the HTML elements to be rendered by the interpreter in cooperation with the Presentation Management modules. An application's container(s) are placed on screen in an arbitrated fashion by the receiver's Presentation Management modules along with containers populated by other applications. In the case of multiple display devices/screens, each display device will have a root container associated with it, and the Presentation Management modules will arbitrate display of elements on these devices' containers; from an application's point of view, this case is analogous to that of Picture-in-Picture displays.

Resource Management

Resource Management functions required in a DASE receiver include allocation, upon request from applications, of resources in the receiver. Examples of resources are local storage space, tuner(s), demultiplexors, etc. These management functions also include arbitration for access to these resources among competing applications. Resource management functions also include means to remove resources from applications and notifying the applications about such removal of resources. Resource Management modules in the receiver provide such functions, programmatically to Java portions of an application, or via the Presentation Engine to application portions authored in HTML, or via the content decoders to application content authored in the format of the decoder.

Event Management

Events may be generated in a DASE receiver from various sources like user action, network data transport mechanisms, timers, applications, etc. Event management functions include means for an application to register with the DASE to listen for events and specify what event handlers need be used. Event management also includes notifying applications of occurrences of events. Event management modules in the DASE receiver provide these functions.

Interaction Between Management Modules

The Application Management modules are to interact with Presentation Management modules to allow association of running applications with presentation elements and containers displayed on screen. The application management modules are also to interact with the Resource Management Modules for allocation and removal of system resources upon request from applications or privileged modules in the receiver. Event Management modules are similarly to interact with the Presentation Management modules and Application Management modules for registering event handlers and passing of events.

Protocol Engine

The Protocol Engine is an embodiment of data broadcast protocols that T3/S13 is defining and interactive service protocols that T3/S16 is defining, plus well-known public communication protocols like

TCP/IP that are in widespread industry use. DASE standards make use of data communication protocols defined by T3/S8, T3/S13 and T3/S16.

Application execution and lifecycle

Application content shall be authored in the languages of the Application Execution Engine or the Presentation Engine, or using the formats of the recommended content decoders. If the part of the application content that must be launched first is authored in the declarative language of the Presentation Engine, the Application Management modules will interact with the Presentation Engine and provide to it this first part for decode and display. The Presentation Engine will interact with the Presentation Management modules in an implementation specific manner to display content on screen. If the first part of the application is authored in the procedural language of the Application Execution Engine, the Application Management modules will pass this part to the Application Execution Engine for decoding and display. Subsequent parts of the application may require interaction between the Application Execution Engine, the Presentation Engine and other Content Decoders, which must all request resources from the Resource Management modules via Resource Management Application Programming Interfaces, request presentation services via the Presentation Application Programming Interfaces or via declarative content interpreted by the Presentation Engine, and other system services via appropriate Application Programming Interfaces.

Further Work in DASE

There is much ongoing work in T3/S17 in defining the DASE. Specifically, the Java Application Programming Interfaces are being defined. Work also continues, in cooperation with the W3C in modularizing HTML and converting HTML into an XML-compliant syntax. Work also continues in defining broadcast-specific extensions to HTML, especially for precise layout control of on-screen elements, for television effects like fades and wipes, for overlay features for tv-quality graphics, for timing and synchronization with audio and video, etc. A Document Object Model (DOM) for DASE will be defined to allow control of the on-screen elements, which are rendered by the HTML interpreter, from procedural code written in Java (and possible ECMAScript). Cooperation continues with other ATSC and MPEG bodies to define means for synchronization of audio and video, delivered via MPEG and IP protocols, with HTML and Java content. We are defining a new broadcast Universal Resource Identifier that will allow reference to broadcast content from within other broadcast content as well as from Web content.

Receiver profiles

Receivers may be implemented in several levels of compliance with the DASE standard. A three-part solution that addresses the issue of receiver profiles is envisioned:

The behavior of receivers in response to executing application programs is classified into three Receiver Profiles.

A minimal set of system services is declared as a DASE Base Level.

Receiver Profile APIs allow an application to query for availability of resources and APIs.

T3/S17 will undertake definitions of profiles and levels in the coming months.

Conformance and testing

Implementation of DASE specifications can be tested for conformance using the conformance tests specified in the DASE standard. Conformance certification is to be carried out by a designated authority.

References

(1) ISO/IEC 13818-1:1996, Information Technology — Generic coding of moving pictures and associated audio — Part 1: Systems.
(2) ISO/IEC 13818-6, MPEG-2 Digital Storage Media—Command & Control
(3) ISO/IEC 8802-2 Logical Link Control (LLC) specification and ISO/IEC 8802-1a SubNetwork Attachment Point (SNAP) specification
(4) http://toocan.philabs.research.philips.com/misc/atsc/dase ATSC T3/S17 (DASE) Web site
(5) http://toocan.philabs.research.philips.com/misc/atsc/t3s13 ATSC T3/S13 Web site
(6) http://toocan.philabs.research.philips.com/misc/atsc/t3s16 ATSC T3/S16 Web site
(7) http://www.atsc.org/ ATSC Web site

Technical Regulatory Hot Topics for Broadcasters

Wednesday, April 21, 1999

2:00 pm - 5:00 pm

Chairperson: Dane Ericksen, P.E.
Hammett & Edison, Inc., San Francisco, CA

***2:00 pm 2 GHz ENG Transition Panel**
David Thomas, Nucomm, Inc., Hackettstown, NJ; Kelly Williams, NAB, Washington, DC; Phil Salas, Alcatel, Richardson, TX; Richard Edwards, SBE Frequency Coordination Committee, Coral Springs, FL; Rick Hollowell, Microwave Radio Communications, Chelmsford, MA

***3:00 pm Ask the FCC**
Keith Larson, Federal Communications Commission, Washington, DC and Bruce Franca, Federal Communications Commission, Washington, DC

***4:00 pm LPFM Technical Issues**
Moderator: Dave Wilson, NAB, Washington, DC
Panelists: Thomas Cornell, Delphi Delco Electronics, Kokomo, IN; Lori Holy; NAB, Washington, DC; Keith Larson, Federal Communications Commission, Washington, DC; Milford Smith, Greater Media, Inc., East Brunswick, NJ

---

*Papers not available at the time of publication

DTV Implementation Workshop

Thursday, April 22, 1999

9:00 am - 12:00 pm

Chairperson: Andy Butler
PBS, Alexandria, VA

9:00 am Principle, Benefits and Applications of Variable Bit Rate Coding for Digital Video Broadcasting, with Statistical Multiplexing Extension
Si Jun Huang
Scientific-Atlanta Inc
Norcross, GA

***9:30 am Scaleable Video: DV - MPEG Transcoding**
Robert Saffari
C-Cube Microsystems
Milpitas, CA

***10:00 am Broadcast Towers - Engineering and Installation Considerations for Digital Television**
Don Doty
Doty Moore Tower Services, Inc.
Cedar Hill, TX
Craig Snyder
Sioux Falls Tower Specialists, Inc.
Sioux Falls, SD

***10:30 am Dealing with System Information**
Matthew Goldman
DiviCom, Inc.
Milpitas, CA

11:00 am Implementing PSIP Solutions
Pierre Clement
Thomcast Communications, Inc.
Alexandria, VA

***11:30 am Progressive Versus Interlace: An Updated Comparison**
Michel Proulx
Miranda Technologies
St. Laurent, Canada

*Papers not available at the time of publication

PRINCIPLE, BENEFITS AND APPLICATIONS OF VARIABLE BIT RATE CODING FOR DIGITAL VIDEO BROADCASTING, WITH STATISTICAL MULTIPLEXING EXTENSION

Si Jun Huang
Scientific-Atlanta, Inc.
Norcross, Georgia

ABSTRACT

This paper studies the principle and benefits of variable bit rate coding of MPEG-2 video compression. The study starts from the fundamental difference of rate control algorithms for constant bit rate (CBR) coding and variable bit rate (VBR) coding. Comparison test results from the real-time video encoder and decoder operation using a PAQ200 picture quality analyzer as the tool for coded video quality measurement are reported and analyzed in association with the rate control algorithms. The combined theoretical analysis and experimental results show that VBR coding has significant bit saving potential over CBR coding for the same coding quality. To effectively utilize the gain from VBR coding, three bit stream transport mechanisms are introduced. This transport discussion is then further extended to the definition of three generations of statistical multiplexers and analysis of their operational features and performance. Finally, potential applications that can benefit from VBR coding have been identified.

1 INTRODUCTION

Currently, most digital video broadcasting systems use the MPEG-2 video compression standard to encode raw video into a constant bit rate (CBR) bit stream. The major advantage of CBR coding is that its bit rate is constant so that it is easy to be transported over conventional communication channels. Since CBR coding rate control aims to balance the desired average bit budget over a specific time period, this implies that the number of bits used for encoding the different segments of a video program is the same. This results in video coding quality variation with respect to video content. For many video broadcasting or transmission applications, higher bit rate is usually used to guarantee the acceptable video quality for the most complex video content segments. At the same time, more bits are unnecessarily wasted for simple video content segments.

To improve the information throughput of digital broadcasting networks, the migration of digital video coding technology from CBR to variable bit rate (VBR) encoding is the next step that could achieve substantial bit savings without sacrificing encoded video quality. The high compression ratio of VBR coding has been evidenced by such applications such as DVD, VCD and video server etc. Depending on the application configuration and encoder and mux design, coding efficiency improvements of up to 40% or more could be reached.

The principle of VBR coding is to encode the video content as much as possible to a constant quality. The number of bits used for coding each frame is a function of spatial and temporal frame content complexity. This results in VBR bit stream.

The above two major features of VBR coding have raised a series of new technical questions for the deployment of VBR coding in digital video broadcasting network. This paper will address some of the issues based on both theoretical analysis and experimental test results.

2 PRINCIPLE OF RATE CONTROL FOR CBR AND VBR CODING

The MPEG-2 video coding algorithm supports two output bit rate operation modes. One is CBR coding that has been studied in detail and used as the default mode of operation for most known applications. The other mode is VBR coding that has been supported syntactically but not studied in detail. This section will explore the fundamental principle difference of CBR and VBR coding methods.

For MPEG video coding, a simplified bit rate reduction model is illustrated in Figure 1. The raw digital video input has a bit rate R_{input} that is typically 216 Mbps for ITU-T 601 video and it is constant over time. The output of encoding process has a bit rate $R_{output}(t)$ that is configurable with an upper limit of 15 Mbps for MPEG-2 video MP@ML. The ratio of R_{input} vs. $R_{output}(t)$ is the compression ratio $\alpha(t)$ of the encoder, which can be represented as Equation 1.

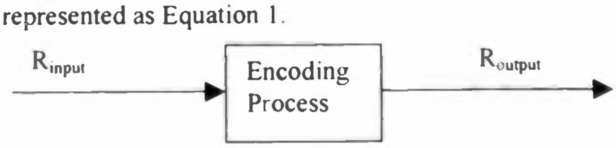


Figure 1 A Simplified MPEG-2 Video Encoding Bit Rate Reduction Model

$$R_{output}(t) = \frac{R_{input}}{\alpha(t)} \quad (1)$$

From this simplified encoding bit rate reduction model, the fundamental difference between CBR coding and VBR coding is that the $R_{output}(t)$ is a constant over time for CBR coding and $R_{output}(t)$ is a variable over time for VBR coding. This fundamental difference results in the following differences between the rate control algorithms.

For CBR coding, the rate control algorithm is designed with the following bandwidth limiting requirements, among many other well-known requirements:

- The decoder's bit receiving buffer is fed with bits from a communication channel or storage media at a constant rate R_{output}, which results in a constant compression ratio α.
- The encoder's bit transmitting buffer shall always have enough bits to sustain the requirement of constant bit rate transmission.
- The encoder should never spend more bits for a picture that could not be transmitted to the decoder's bit receiving buffer before its decoding time at the specified constant output bit rate R_{output}.

Based on the above requirements, the simplest coding rate control algorithm is to encode each input video picture with a fixed number of bits so that the total output bit rate is easily controllable. This is the basic rate control algorithm that could be used for video sequence coding based on a still picture coding algorithm, such as motion JPEG. However it is known that such an intra-picture coding technique is not efficient since it does not deploy the content similarity or redundancy in temporal domain. To overcome this problem, the MPEG video coding algorithm included motion estimation and motion compensation techniques that improved the coding efficiency substantially. One of the impacts of such an inter-picture coding technique is the variation of coded picture size, which is no longer a constant. It is well known that the Intra (I) picture consumes most bits, the forward Predictive (P) picture consumes less bits than the I picture and the Bi-directional predictive (B) picture consumes least bits on average. However the bit allocation differences among different coding picture types are constrained also by the requirement of a constant output bit rate R_{output} for CBR coding. For most MPEG video encoder implementations, the constant bit rate requirement is usually enforced or balanced at the Group of Picture (GOP) layer, which is typically set to 15 for 30 Hz video format. This implies that the rate control algorithm has to spend the budgeted amount of bits within the GOP. This mechanism resulted in variation of coded video quality due to insufficient bits for complicated video content and too many bits for simple and still video content.

For VBR coding, the rate control algorithm has the following differences compared to those of CBR coding algorithm listed above.

- The decoder's bit receiving buffer is fed with bits from a communication channel or storage media at a constant peak rate R_{max}. However this R_{max} is not the sustained transmitting rate rather a burst rate. In other words, the transmitting rate is R_{max} whenever there are any bits and zero when there are no bits in the encoder's transmitting buffer. This mode of bit transmission is called "leaky-mode" in MPEG-2 standard. When the transmitted bits are averaged over time duration, it is a variable. Therefore the compression ratio is a variable over time, $\alpha(t)$.
- The encoder's bit transmitting buffer can be emptied so that a picture can be coded as small as syntactically possible.
- The encoder should never spend more bits for a picture that could not be transmitted to the decoder's bit receiving buffer before its decoding time stamp at the specified constant output bit rate R_{max}.
- The bit allocation is not budgeted for balancing at GOP interval. It is only limited by the R_{max} and the decoder's Video Buffer Verifier (VBV) model.
- The bit spending is controlled by a coding quality parameter, such as the picture or macroblock quant_scale_code or some other parameters, so that only minimal bits are spent for the desired coding quality demand.

3 BENEFITS OF VBR CODING

Comparing the principles of rate control algorithms of CBR and VBR coding, we can find the following features:

- CBR coding has a minimal bit spending constraint due to constant bit rate transmission needs. Even for still sequences, the specified bit rate is still maintained so that most of bits are unnecessarily wasted.
- VBR coding has no minimal bit spending constraint such that minimal bits can be used to code still sequences or other simple sequences.
- VBR coding bit rate control is mainly modulated by the coding quality parameter such that minimal bits are used to achieve the target coding quality.
- The bit saving on the simple video content effectively produces a coded bit stream with a lower average bit rate R_{av} than the peak rate R_{max}.

- Comparing VBR coded average bit rate R_{av} to the constant bit rate R_{output} of CBR coding, VBR coding could achieve higher video coding quality since it can use higher peak rate R_{max} to encode the most complicated content.

To verify the above benefits of VBR coding vs. CBR coding based on the principle of rate control algorithms, a real-time encoding and decoding video quality test has been conducted. The test configuration is shown in Figure 2 where a Tektronix PQA200 picture quality analyzer [1] is used to measure the video content impairment from one generation of MPEG-2 real-time encoding and decoding process. To simplify the discussion, only the 525 lines video content test results are used in this paper. The PQA200 measures the video content impairment using the Picture Quality Rating (PQR) measure based on the JNDmetrix™ human-vision algorithm. The PQA200 comes with a library of original digital video sequences that are used as the video content for the testing. The video sequences are sent to a real-time video encoder in serial D1 signal format. The compressed video bit stream is decoded by a real-time satellite Integrated Receiver and Decoder (IRD). The decoded digital serial D1 signal is acquired by the PQA200 for field by field comparison and PQR measurement.

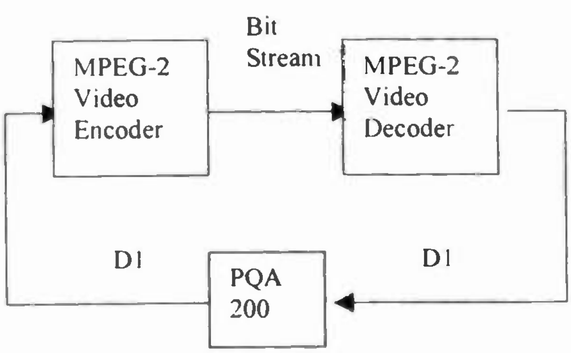


Figure 2 MEPG-2 Video Encoding and Decoding PQA Test Configuration

First CBR coding quality measurements on various video contents were tested over different bit rates. Figure 3 shows the PQR values from the five test sequences, i.e., Flower Garden, Cheer Leader, Mobile and Calendar, Suzie, and Table Tennis. Analyzing the five curves, we obtained the following observations:

- Different video content has a different requirement of bit rate to achieve certain coded video quality. For Suzie sequence, it needs only 2 Mbps to obtain a PQR of about 3, which is considered of low content impairment. For Cheer the coded quality is not high with a PQR of 7 even when the coding bit rate is 7 Mbps.
- Assume the coded video quality with 7 Mbps bit rate is acceptable for broadcast applications, assume further that the quality impairment difference of 1 PQR is not significant, then we can find that many sequences does not need to be coded at 7 Mbps. Table 1 shows the minimal bit rate needs of the 5 sequences under testing that would achieve similar coded video quality in terms of PQR measurement.

Table 1 Minimal bit rate requirement for different sequence

Sequence	PQR @ 7 Mbps	Minimal Bit Rate Requirement
Flower Garden	4	4.7 @ 5 Mbps
Cheer Leader	6	7 @ 5.5 Mbps
Mobile & Calendar	5.7	6.6 @ 5 Mbps
Table Tennis	3.8	4.8 @ 3Mbps
Suzie	2.5	3.4 @ 1.5 Mbps

- From the bit saving perspective, CBR coding will suffer video quality impairment when the bit rate is not enough for high complexity content. On the other hand, there will be significant bit overspending for lower complexity video content that will not improve the coded video quality significantly.
- The variation of coded video quality associated with video content variation will produce so called "pumping" or "breathing" effect.

To illustrate VBR coding benefits with respect to CBR coding, a variable video content sequence is modeled by concatenating five standard test sequences Flower Garden, Cheer Leader, Mobile and Calendar, Suzie and Table Tennis as one single video sequence. The PQA measurements of these five sequences are therefore concatenated in the same order as shown in Figure 4, which compares the PQR of CBR vs. VBR coding for bit rate setting at 3Mbps for both encoders.

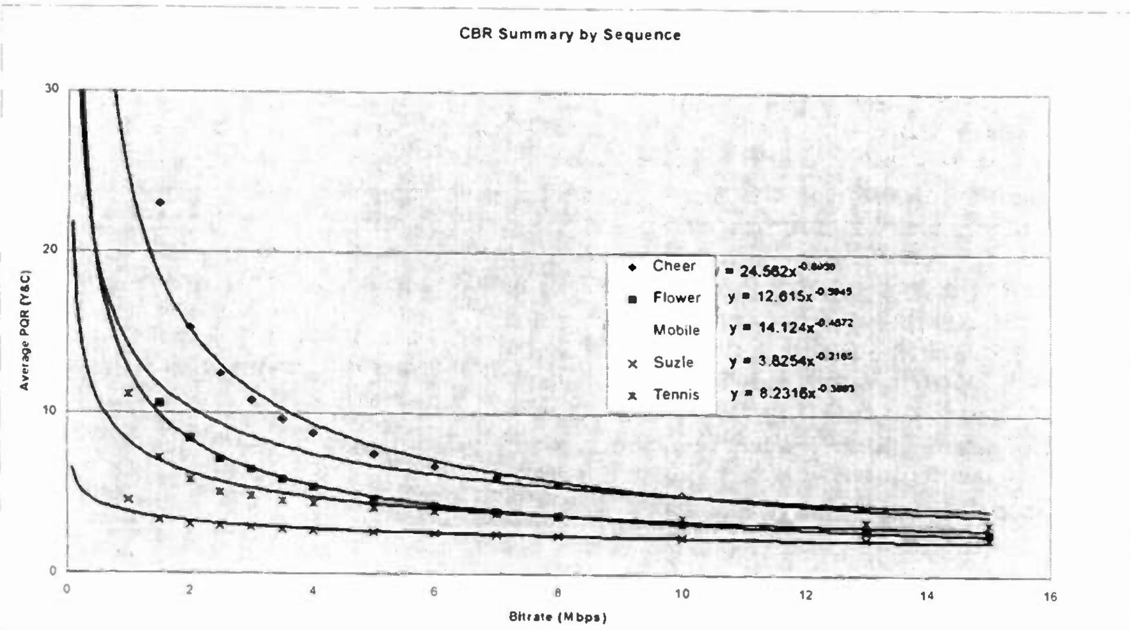


Figure 3. Coded Video Quality Variation w.r.t. Bit Rate and Content

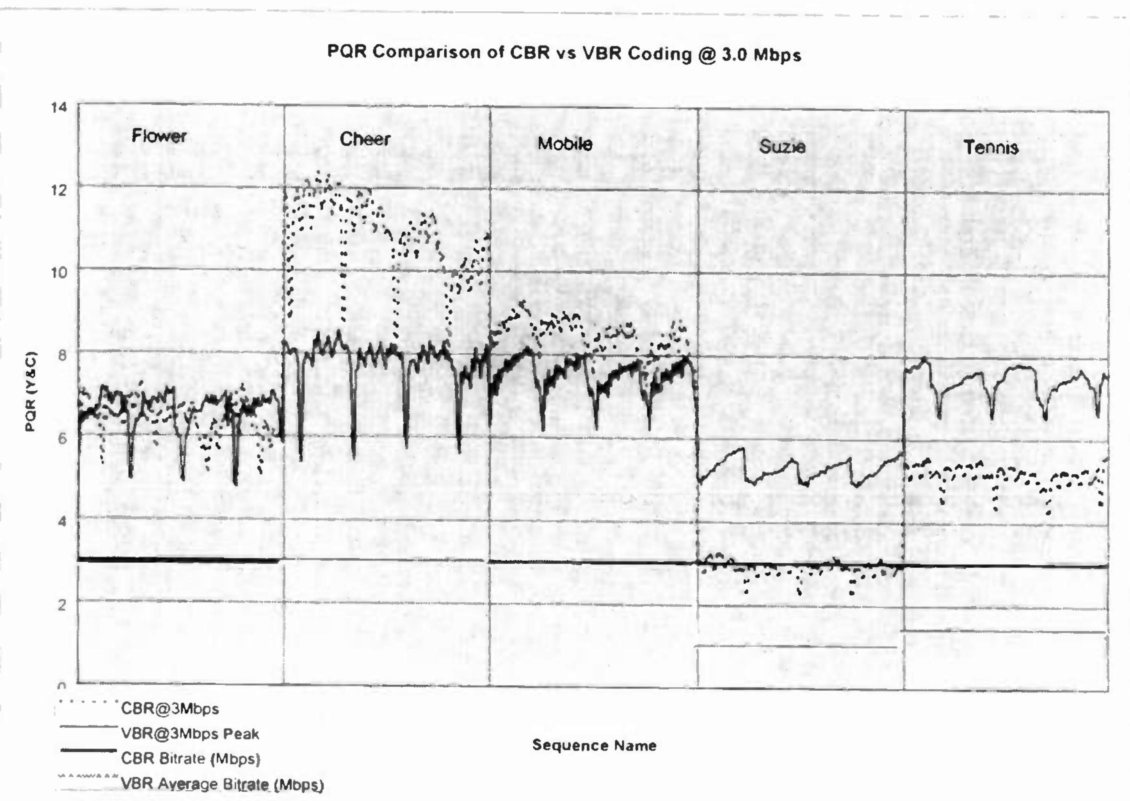


Figure 4 Coded Video Quality Comparison of CBR and VBR Coding

The dotted black line in Figure 4 represents the PQR values of CBR coding at 3 Mbps. The PQR values on average are 7, 11, 9, 3 and 5 in the content order as shown. This represents a wide variation of PQR values and thus the coded video quality. For Suzie sequence, the PQR is very small that represents a low video content impairment, i.e., very high quality. For Cheer Leader sequence, the PQR is very big, which represents a high video content impairment, i.e., very low video quality. This wide range of PQR variation represents the coded video quality variation that has already created a contradiction of high video quality and low bit rate video coding requirement in various digital video compression applications.

The precious bandwidth in any communication applications always demand for low coding bit rate so that more content could be transmitted or delivered in a given bandwidth. However the video content is a random sequence that has a great variation in content. A news channel will have all kind of live content insertion into the scene of talking heads. A movie not only has its own content variation, but also has frequent insertion of commercials. Therefore it is very difficult to set a constant bit rate for almost any video channel. If one sets the bit rate to be high enough, say 9 Mbps, to guarantee certain level of video quality for the most complicated content, a high percentage of bits would be wasted for simple content since the extra video quality difference is not significant for it. This is evidenced by the PQA value of 3 for Suzie sequence at 3 Mbps, which represents moderate video impairment. When CBR bit rate is set to 9 Mbps, the PQA value is only improved to 2.5. On the other hand, the PQA value for Cheer Leader at 9 Mbps is only about 5, which is still worse than the Suzie at 3 Mbps. Therefore potentially more than 66% of bits are wasted for Suzie sequence if the bit rate is set to 9 Mbps. Conversely it is obvious that the video quality of Cheer Leader coded at 3 Mpbs with CBR coding is not acceptable.

The solid red line in Figure 4 represents VBR encoder PQR measurements at the peak bit rate of 3 Mbps. One can read the PQR values are 7, 8, 8, 6 and 7 in the same sequence order. This result is obtained by setting VBR quality control parameters such that VBR PQR value is similar to CBR PQR value for Flower Garden sequence. This setting is based on the philosophy that sets VBR coding quality to the same as CBR coding quality for average video complexity sequence such as the Flower Garden. When the actual video content is more complicated than the Flower Garden, such as Cheer Leader, Mobile and Calendar sequences, a higher coding quality is expected. This is evidenced by the PQR of 8 for the 2nd and 3rd sequences from VBR coding vs. the PQR values of 12 and 9 from CBR coding. When the actual video content is less complicated, a lower coding quality is expected, which is evidenced by the PQR value of 6 for Suzie sequence from VBR coding vs. the PQR value of 3 from CBR coding. This is the case when VBR encoder saves the bits. Figure 4 also plots the average bit rate for both CBR and VBR encoders, the green line represents CBR bit rate of 3 Mbps and the blue lines represents the variation of average bit rate with respect to content. The average bit rates of VBR coding are 2.9, 3, 2.95, 1 and 1.4 Mbps for the five sequences, respectively. When these five bit rates are averaged, it is equal to 2.25 Mbps. In other words, VBR encoder saved about 25% bits vs. CBR coding and at the same time achieved a less variant coded sequence quality.

From video content complexity statistics based on many tests conducted so far, some of them will be reported in our future publications about this topic, the probability of high content complexity is less than 0.6. This implies a potential of up to 40% bit saving capability of VBR encoding vs. CBR encoding for the same video content. However this 40% bit saving potential is a statistical variable that is not guaranteed over any determined time duration. Therefore how to make use of this significant bit saving potential of VBR coding deserves more exploration. Furthermore, the most conventional communication channels are of fixed bandwidth, i.e., fixed bit rate. It is therefore more natural and friendly to CBR bit streams. The next section discusses some transportation mechanisms of VBR bit streams with an overall target of maximizing channel bandwidth utilization.

4 TRANSPORT OF VBR BIT STREAMS

Assuming the transport channel is constant bandwidth, there are the following means of transporting VBR bit stream efficiently.

4.1 VBR bit streams with Available Bit Rate (ABR) data

The Figure 5 is a simple VBR bit stream transporting method where one or multiple VBR bit streams are multiplexed directly without any intelligent processing. This multiplexer has an output bit rate that is equal to the sum of the peak bit rate of all VBR input bit streams. Due to the bit saving nature of VBR coding, there will be more time when the sum of the input bit rate is much lower than the mux output bit rate. These unused bits are therefore used as an ABR data channel for various applications. Figure 6 is an illustration of bit usage distribution over time of this simple mux and ABR data service configuration. Again it has been observed that up to 40% of bits could be used for ABR data channel. The drawback of this multiplexing method is that the bit saving capability of VBR coding is not used to either improve the video quality or increase the number of video programs for the same output bit rate.

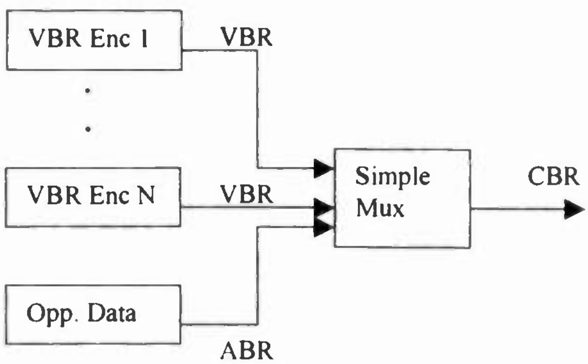


Figure 5 Simple Mux and ABR Data Configuration

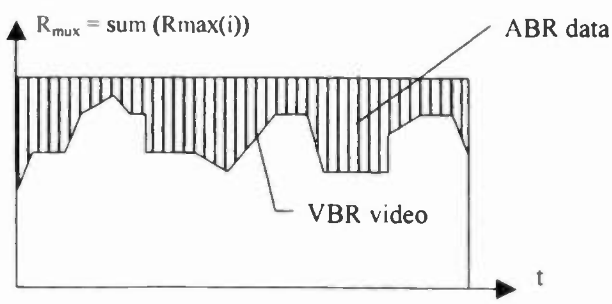


Figure 6 Transport Bandwidth Utilization for Both VBR Video and ABR Data Services

4.2 Multiple VBR streams with a VBR capable, open-loop intelligent multiplexer

An open-loop intelligent multiplexer is defined as a mux that can understand the syntax of the incoming transport packets down to video stream layer. It has a relatively larger packet buffer so that enough time is available for the mux to analyze the input bit stream and make packet output allocation with both inter-channel redundancy and time-shifting mechanisms. Since such an intelligent multiplexer understands the input bit stream, it can conduct efficient statistical multiplexing of input VBR bit streams and has many other capabilities such as bit stream splicing, routing, re-multiplexing etc. This intelligent mux takes open-loop VBR bit stream feeds so that there is no requirement to have a closed-loop to control the video encoders. This feature is specially important since the emerging digital video networks are going to see more and more pre-compressed video feed that is generally not controllable at the transport point of the networks.

Due to the open-loop nature of such an intelligent multiplexer, the application configuration can also be represented by Figure 5, except for the replacement of the simple mux with an intelligent mux. The effect of the intelligent mux can be illustrated in Figure 7 where the total mux output bit rate R_{mux} is a scale factor, $\beta \leq 1$, of the sum of the peak bit rate of each VBR bit stream. For a 10 channel VBR statistical multiplexing operation, a 20% bits saving from VBR coding could be used for video coding, which corresponds to β of 0.8. This is also illustrated in Figure 7 where the ABR percentage is reduced by applying the time-shifting mechanism of the intelligent mux. However it should be pointed out that the 40% total bit saving potential from VBR coding could not be used all for video coding due the limitation of buffer capabilities of both the mux buffer and decoder's bit receiving buffer. For the 10 channel statistical multiplex test example given above, a 20% bit saving for video transport is a reasonable number without a need for feedback control of VBR encoders. The remaining 20% bits are therefore still useful for ABR data services.

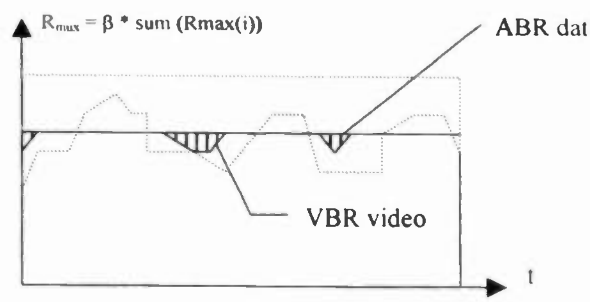


Figure 7 VBR Statistical Multiplexing Using Open-loop Intelligent Mux

4.3 Multiple VBR bit streams with a VBR capable, closed-loop intelligent mux

To further utilize VBR bit saving capabilities for video services, a closed-loop intelligent mux can be used where the peak rate of each VBR encoder is coordinated by a statistical multiplexing rate control agent as shown in Figure 8. The rate control agent will change the peak rate allocation for each VBR encoder frequently such that the sum of the peak rate is always equal to a percentage of the mux output bit rate R_{mux}, i.e., $R_{mux} = \beta * \text{sum (Rmax(i))}$ where $\beta \leq 1$. Obviously one would expect for a lower β value than that of the open-loop intelligent mux operation. With this closed-loop peak rate control mechanism, it is possible then to set the intelligent mux output bit rate R_{mux} to be very close to 40% of the sum of the peak rate, which is the maximum bit saving potential of VBR coding. Due to the time limitations, there has been no testing done for this configuration as of this date.

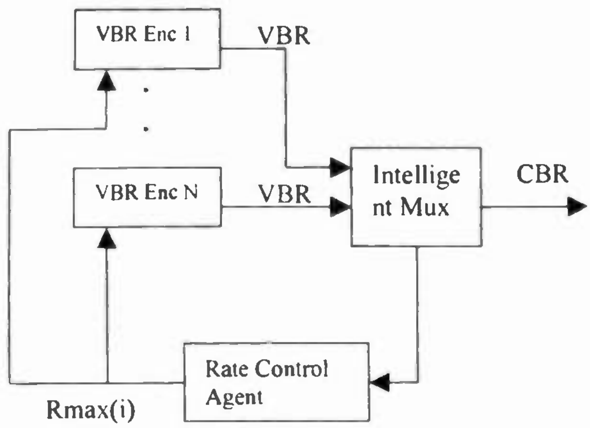


Figure 8 Closed-loop Intelligent Mux for VBR Stream Statistical Multiplexing

As a trade off, one could observe that the ABR data channel may not be available, or have a minimal bit rate, with this configuration since any extra bits are used to improve the video coding quality through the dynamic peak rate allocation mechanism.

5 THREE GENERATIONS OF STATISTICAL MULTIPLEXER AND THEIR OPERATION MECHANISM

The discussion of transportation of VBR bit streams will naturally lead to the topic of statistical multiplexing. Based on the video encoder coding mode, i.e., CBR or VBR, and the statistical multiplexer configuration, three generations of statistical multiplexers are possible.

5.1 First Generation Statistical Multiplexer

It is called G1 Statmux hereafter . It's major features are:

- Video encoder is working in piece-wise CBR mode where the bit rate for each piece of operation is controlled by a statistical multiplexing agent.
- The agent collects the current coding statistics reported by all encoders and make decisions of new bit rates for the next piece of operation of encoders. Therefore it is subject to posterior prediction error.
- The mux is a non-intelligent processor takes constant bit rate inputs and produce a constant bit rate output.
- When video content complexities of multiple encoders line up, all coding qualities of those channels are suppressed.

5.2 Second Generation Statistical Multiplexer

It is called G2 Statmux hereafter. It's major features are:

- The video encoder is assumed with no intelligence of understanding the video content complexity so that an external rate control mechanism is needed to moderate the bit rate allocation in a piece-wise constant bit rate coding fashion.
- Two pass encoding processes are introduced. The first encoding engine will collect the content statistics that is used to control the bit rate allocation of the second encoding engine.
- The mux is still a simple bit stream combiner that takes constant bit rate input and produces a constant bit rate output.
- The early statistical analysis will bring in performance improvement over the G1 Statmux system since the prediction error could be removed.
- It will still suppress the video qualities of those encoders when their content complexity peaks are aligned with respect to time.

5.3 Third Generation Statistical Multiplexer

It is called G3 Statmux hereafter. It has the following major features:

- The video encoders are running in VBR coding mode. It assumes that the video encoding engine knows how to self-adapt to the content changes and allocate bits to each frame intelligently. There is no bit budget balance with a GOP. It is only constrained by the peak rate and the VBV model.
- The mux is an intelligent process as introduced in section 4 of this paper. It takes variable bit rate input and generates constant bit rate output.
- The extra buffer capability of the G3 Statmux introduces the time shifting of peaks so that peak line-up quality suppression could be reduced. The time shifting concept itself is not new and it has been used in telecommunication industries many years. It has also be used for VBR stream transportation in ATM networks and other applications[2,3]
- The open-loop architecture provides many new configuration and application possibilities, which will be further discussed in section 6 of this paper. However it also introduces a need for a recoding process when the sum of the peaks is beyond its buffering capacity.

6 Operation and Performance Comparisons

To further illustrate the operation and performance differences between these three generations of statistical multiplexers, a hypothesis statistical multiplexing operation is modelled where there are three encoders and one mux. Figure 9 a, b, and c show the bit rate requirements of these three encoders with their respective video content for the purpose of achieving a certain constant quality performance. Figure 9d is the mux bit distribution needed.

Figure 10a is a possible operation of G1 Statmux. Figure 11a is the coded quality performance indication. This quality indication is a simple model that uses the ratio of the bit rate that is allocated vs the bit rate requirement from the encoder originally for the certain quality level. If the allocated bit rate is higher than the needed bit rate, a high quality is achieved. If the allocated bit rate is lower than the needed bit rate, a lower quality is achieved. However it should be noted that the absolute value of the quality indicator does not necessarily correspond to the same amount of video quality improvement or degradation since the quality vs bit rate is a nonlinear function as shown in Figure 3. From Figure 11a, one can see that in many instances the video encoder's quality is suppressed, i.e., the quality indicator is less than 1. For many other instances, the video encoder's quality is much higher than what is expected, i.e., the quality indicator is greater than 1.

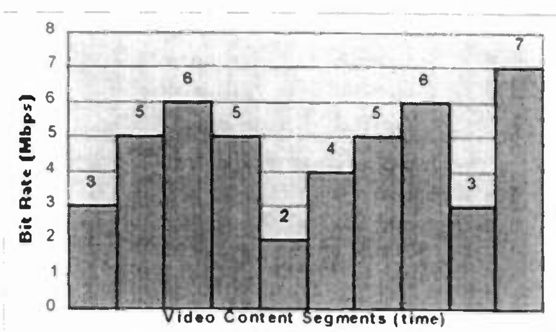


Figure 9a. Encoder 1 Bit Rate Requirement

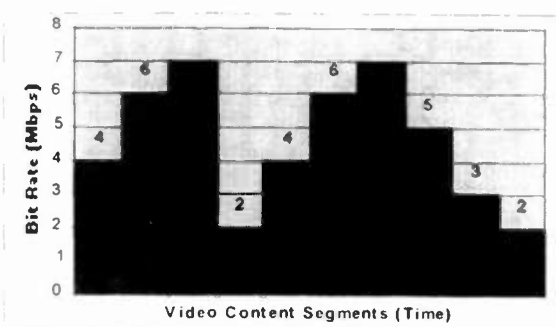


Figure 9b. Encoder 2 Bit Rate Requirement

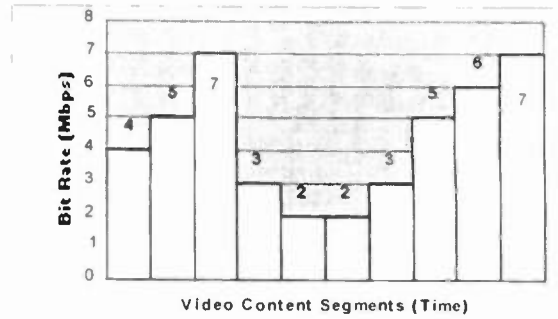


Figure 9c. Encoder 3 Bit Rate Requirement

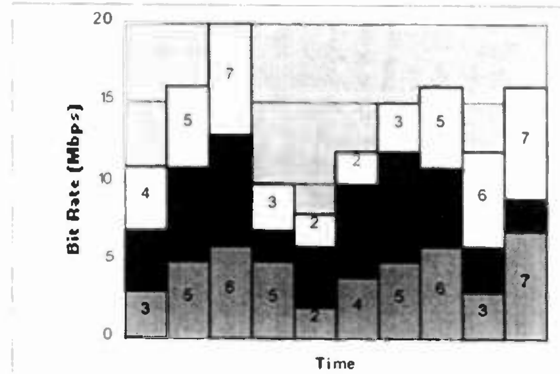


Figure 9d. Mux Bit Distribution

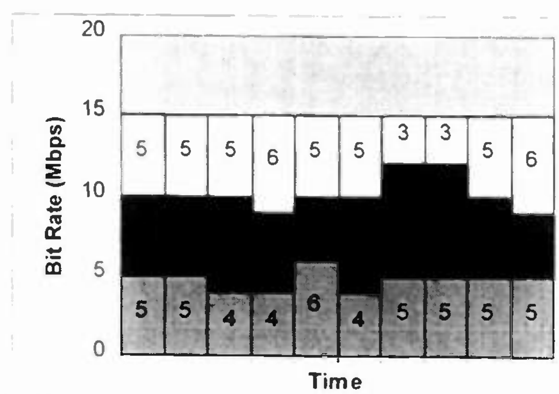


Figure 10a. G1 Statmux Performance

Figure 10b is a possible G2 Statmux operation and Figure 11b is its quality performance indication. One can find that the delayed feedback bit rate control that causes quality penalty in G1 Statmux has been removed. However due to lack of time shifting mechanism, the quality performance indicators have many instances of less than 1 and also many instances of greater than 1. In other words, it over spends bits for simple content time duration that is similar to the G1 Statmux and it also suppresses video quality for complex video content time duration that is similar to the G1 Statmux.

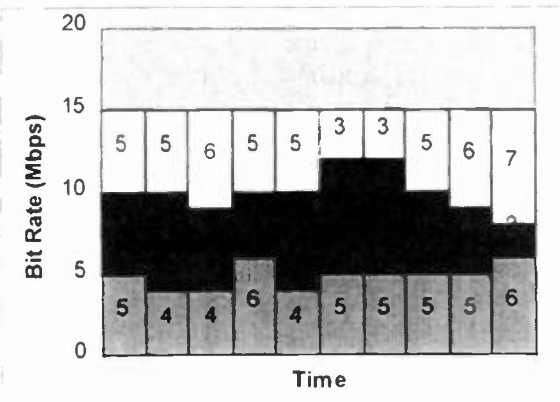


Figure 10b. G2 Statmux Performance

Figure 10c is a possible performance of G3 Statmux and Figure 11c is its quality performance indicator. Figure 10c shows the time shifting effect where the extra bits above 15 Mbps lines are shifted to the time slot of 1, 4 and 9 as indicated in light blue color. Furthermore, there are still unused bits in time slot 1, 5, 6 and 9 that can still be used for ABR data services as discussed in Section 4 earlier. Figure 11c shows that the quality performance of the G3 Statmux is maintained as 1, which indicates that there is no quality suppression and no bit wastes. Of course this is a simple illustration. The actual performance is usually not this ideal, i.e., certain quality variation is unavoidable. However one thing is true that the quality variation will be generally less than the G1 and G2 Statmux systems.

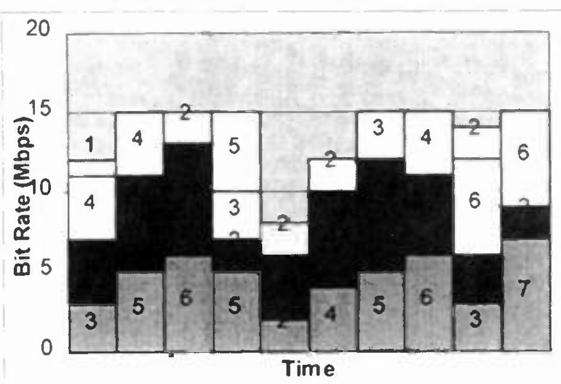


Figure 10c. G3 Statmux Performance

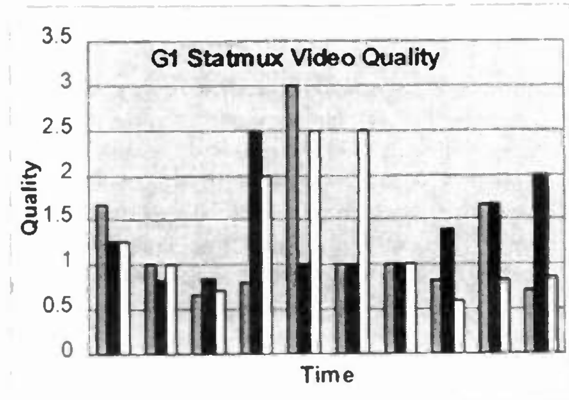


Figure 11a. G1 Statmux Quality Performance

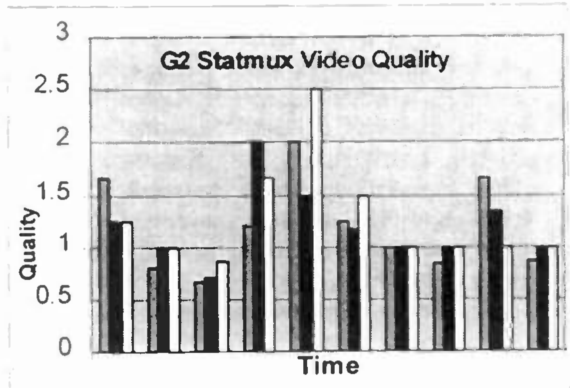


Figure 11b. G2 Statmux Quality Performance

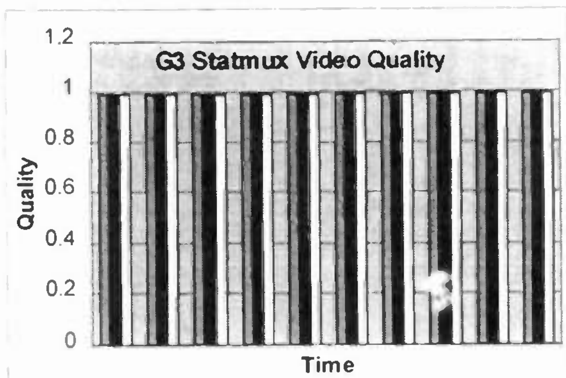


Figure 11c. G3 Statmux Quality Performance

7 APPLICATIONS OF VBR-BASED INTELLIGENT MULTIPLEXOR

The most important feature of VBR coding is its ability to achieve similar or better video coding quality over random complexity video content while using fewer bits compared to CBR coding. This amount of bit saving is a very significant gain for any digital communication system. The open-loop based intelligent mux that can support both VBR and CBR streams without the need of local loop-control of real-time encoders opens the possibility of the following applications:

a) Increased number of video programs per transmission bandwidth through statistical multiplexing.
b) Improved video coding quality through VBR coding using same average bit rate.
c) Use of ABR bandwidth to transmit asynchronous data services such as those defined by ATSC and DVB data broadcasting specifications. These services will be added as if "free" from the original video program configuration, i.e., the same number of video programs with same video quality.
d) Statistically multiplex pre-compressed video such as those from a remote feed or from an MPEG-2 video server.
e) Statistically re-multiplex bit stream from statmuxed video streams.
f) Statistically multiplex VBR coded video directly off a DVD player, which is a professional player that can output coded MPEG streams, without any decoding and recoding processes.
g) Statistically multiplex a VBR based MPEG-2 digital video camera output for news gathering and other live events.

8 CONCLUSION

Due to the utilization of encoding process intelligence of MPEG-2 video encoders, VBR coding algorithm has the potential to save up to 40% of bits needed by CBR coding to achieve similar coded video quality. These benefits are analyzed both from the rate control algorithm principles and also the experimental test results using PQA200 picture quality analyzer. When coupled with an intelligent mux that is VBR statistical multiplexing capable, significant amounts of bit saving could be used to improve the video coding quality, increase video program numbers or add extra data services without new communication channel capacity extension. Applying the combination of true VBR encoding and intelligent multiplexing are of high value to the increasingly busy data communication and broadcasting industries.

9 ACKNOWLEDGEMENT

The author would like to thank his colleague Mark Spittel for his persistent and high quality work in conducting many detailed and time consuming PQA200 tests and data processing.

10 REFERENCE

[1]. Tektronix, "PQA200 Picture Quality Analysis System User Mannual", 071-0091-01

[2]. Amy Reibman and Barry Haskell, "Constraints on Variable Bit-Rate Video for ATM Networks", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 2, No. 4, Dec. 1992

[3]. Barry Haskell and Amy Reibman, "Multiplexing of Variable Rate Encoded Streams", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 4, No. 4, Aug. 1994

IMPLEMENTING PSIP SOLUTIONS

Pierre Clement, James Kutzner, Shannon Skemp
Thomcast Communications, Inc.
Alexandria, Virginia

ABSTRACT

Future digital television station operations will include many of the existing processes and a number of new functions. Included among these new processes will be Program and Specific Information Protocol (PSIP) generation and management. Daily PSIP operations will involve editing, preparation, and transport stream insertion of this new information which describes the DTV transport stream content. This information is required by the DTV receiver for various operations such as locating program services, and building an on-screen Electronic Program Guide (EPG) through which the end user is informed of the programs and other services carried in the DTV transport stream.

PSIP is an ATSC standard that defines and provides the basic set of tables and content information intended for use by DTV receivers. It is also a requirement for the business operations of a DTV station.

Today, PSIP is still at the early age of its implementation and most of the proposed solutions are static, that is, the content of the tables is not changed in real-time as it is carried within the DTV stream. However, in the future implementation solutions for practical station operations will require dynamic and seamless changes where the PSIP stream is fed by external systems such as traffic and automation systems which may also be controlling encoders, MPEG stream servers and multiplexers.

This paper discusses potential future PSIP implementation and the resultant functionality at the consumer level. In order to realize that functionality in a cost-effective operation, methods and internal architecture within a practical and complete station are presented and interface protocols and guidelines are proposed and discussed.

Key Ideas

The goal of PSIP is to supply the decoder with tuning information and to offer to the end user a reliable and complete Electronic Program Guide. The implementation of PSIP does not only take place at the level of broadcasting stations but also at the level of various network head-ends which are responsible for the reliable and timely delivery of accurate EPG data. Both the broadcasters and the network operators must be cognizant and concerned of how their PSIP data are handled before it finally reaches the end consumer.

Introduction

May 30th 1999. Our friend Bernie, fresh from the NAB show, has been convinced by his experiences that he needs to purchase a new digital television system. A friend of his assured him that there were many advantages to do so, most notably the Electronic Program Guide (EPG) which replaces in a considerably more convenient manner his traditional newspaper. Bernie is happy. It's Friday evening, his wife has taken the dog for a walk, and he is now able to enjoy his new purchase. And so Bernie switches on his new digital TV. He presses a remote button to display the Electronic Program Guide. A menu presents him with a list of the different channels, and after having selected an interesting program Bernie happily clicks his remote to access that program's description. He reads the content description, which convinces him to select that program, but to his surprise the program does not correspond to the description. Quite disappointed, Bernie selects a new channel described in the EPG but then again, the set-top box tells him that it is not actually available. Bernie realizes that his friend has emphasized the EPG too much and in dismay throws the remote control on the floor and stomps off to join his wife and the dog.

To prevent this from happening, many issues must be raised and solved. What is the information required in building an EPG? How is this information processed, injected, and then extracted, reprocessed and re-injected into the digital television signal? How can we be sure that information eventually arrives at the end user? These and related issues will be discussed within this paper.

Discussion of the PSIP Standard and its Elements

To provide the decoder with tuning information as well as the user with a complete Electronic Program Guide (EPG), the digital television signal must deliver to the set-top box information which describes system information and program data. Whereas system information allows navigation and access of the channels within the DTV transport stream, the program data gives information for browsing and selection[1]. In MPEG terminology the transport stream, or multiplex, is a digital stream of 19.39 Mbs that in the digital world replaces each of the 6 MHz channels used to carry analog TV signals. The PSIP data are carried via a collection of hierarchically arranged tables described in the ATSC standard (A/65), the Program and System Information Protocol (PSIP). There are in all six types of tables: the System Time Table (STT), the Master Guide Table (MGT), the Virtual Channel table (VCT), the Rating Region Table (RRT), the Event Information Table (EIT) and the Extended Text Table (ETT). A brief description and some typical features of those tables are presented. A discussion will then follow, highlighting issues that may arise with regard to the carriage of PSIP tables in a complex television system.

The STT, or System Time Table, is a small table that carries time information required by any application within the transport stream needing synchronization. This table has a typical length of 20 bytes and must be transmitted at least every second.

The MGT, or Master Guide Table, is like a flag that continuously informs the decoder about the status of all the other tables, except for the STT. It defines table sizes necessary for memory allocation during decoding, the PID that labels each table and a version number to identify tables needing to be updated[1]. The MGT has a few hundred of bytes size directly related to the number of tables within the transport stream, and must be transmitted at least every 150 ms.

The VCT contains a list of all the channels within the transport stream. Brief information is provided for each channel, such as tuning information and the channel name. A VCT is few hundred bytes in size and the table must be transmitted at least every 400ms.

The RRT, or Rating Region table, defines rating rates for different regions or countries. This table of approximately 1 Kbytes must be transmitted at least every minute.

Also a part of PSIP are several Event Information Tables, or EITs, which contain information for events on defined channels. Each EIT covers a time interval of 3 hours. The maximum time span covered by the EITs is 16 days (128 three-hour EITs). The EIT related to the current time span is called EIT-0, and the EITs related to the succeeding time spans are numbered from EIT-1 to EIT-127. Each EIT-k has as many instances as channels within the transport stream. An EIT has a typical size of 2 Kbytes and it is only recommended by the PSIP standard that the instances of EIT-0 be transmitted with a time cycle of less than 500 ms.

An ETT, or Extended Text Table, contains additional information for a program or channel. An ETT can be linked to any of the EITs and to the VCT. The table size is a few Kbytes and there is no requirement for its transmission.

One PID, called the base PID, labels the MGT, the VCT, the RRT and the STT. The EITs and ETTs are labeled by PIDs that can be retrieved from the MGT. Those tables, however, shall not be thought as a unique entity. The MGT and the VCT are used for tuning, and the EITs and the ETTs provide content information for the different programs within the transport stream. If the RRT is unlikely to change frequently, the MGT and the VCT change whenever a change occurs in the multiplex configuration and, by nature, many changes occur in the EITs. Since tuning may be considered more vital than program announcements, the terrestrial and cable network must ensure proper carriage of the SST, the MGT, the VCT, and the RRT in order to be ATSC compliant. EIT-0 through EIT-3 must be carried for a terrestrial network whereas all the EITs are optional for a cable network. The ETT tables are optional for either network.

A broadcaster may only want to transmit the minimum PSIP information required by the standard for proper tuning plus the minimum set of EIT tables. In this case the MGT, the STT, the VCT, and the RRT plus the instances of the first four EITs are transmitted.

For example, four DTV channels are to be transmitted and a single rating region is defined. Typical sizes are then 150 bytes for the MGT, 450 bytes for the VCT, 900 bytes for the RRT and 20 bytes for the STT. Each EIT is supposed to have a typical size of 2 Kbytes and therefore the 16 EITs (there are four instances of an EIT-k, since there are four channels) represents 256 Kbits. The total amount of data is approximatively 268 Kbits. If the maximum time cycles for each of the tables are respected and the broadcaster chooses to send each instance of the EIT-1 to EIT-3 every 2 seconds, then the broadcast of PSIP tables requires a data rate of around 240 Kbps. The major portion of this rate comes from the EITs with 224 Kbps, whereas the MGT, VCT, RRT and STT contribute to the total rate with only 16.1 Kbps. The PSIP data rate has to be linked with the typical throughput of a broadcasting station, which is 19.39 Mbps, and the bandwidth that a broadcaster is ready to provide for PSIP which is a few percent of the total bandwidth. For this example the bandwidth allocated for PSIP represents around 1.2% of total bandwidth. If EITs are not broadcast, then the PSIP bandwidth falls to around 0.1 %.

Broadcasters must also make certain decisions as to the amount of information to be transmitted. As mentioned above, the EPG can also supply the viewers with program information for the next 16 days for each of the channels within the transport stream.

In the case of a broadcaster offering four channels and full EITs but not ETTs for 16 days, the amount of PSIP information transmitted is around 1 Mbytes. In fact, a set of 128 EITs, corresponding to the 16 days divided into 3 hours, is transmitted for each channel within the transport stream. If an EIT size is 2 Kbytes, then the total amount of bytes for the all of EITs is 128x4x2 Kbytes, or 1 Mbytes. Moreover, The more information provided the higher the resultant EPG quality, but also the higher the PSIP data rate. If the required rate is computed with the maximum time cycles of the tables, and with a transmission of the EIT-1 to EIT-127 every 2 seconds, the total rate is around 4 Mbps, that is to say 20% of the total bandwidth. This figure is to be linked with the 1-2% that a broadcaster is ready to invest in an EPG.

A compromise must therefore be found between the acceptable provision of data to a viewer and the use of multiplex bandwidth. Among others, the DVB community provides "guidelines on implementation and usage of Service Information (SI) ", where SI is the PSIP European equivalent, and some repeat rates for the EITs in a terrestrial system. It is recommended in those guidelines to send the EIT-0s and the EIT-1s at least every 2 seconds, whereas the EITs for the first full day may be transmitted every 10 seconds, the EITs for the first 8 days, at least every 30 seconds. After 8 days, the EITs may be transmitted every 300 seconds[4]. Therefore, the total rates for EITS is:

For 4 channels with a guide depth of 16 days:

EIT-0s and EIT-1s :

$$4 \times 2 \times (8 \times 2000) \times ½ = 64 \text{ Kbps}$$

EIT-2s and EIT-7s :

$$+ \ 4 \times 6 \times (8 \times 2000) \times 1/10 = 38.4 \text{ Kbps}$$

EIT-8s and EIT-63s :

$$+ \ 4 \times 56 \times (8 \times 2000) \times 1/30 = 119.6 \text{ Kbps}$$

EIT-64s and EIT-127s :

$$+ \ 4 \times 64 \times (8 \times 2000) \times 1/300 = 13.7 \text{ Kbps}$$

That is $\quad$ 235.7 Kbps

The transmission of such an EPG would therefore require less than 1.5 % of the total bandwidth.

Let's now assume that we have to send the PSIP tables related to the electronic Program Guide related to a complete package of 50 television channels. To reduce bandwidth it is decided to limit the guide depth to 7 days. If the same rules in the preceding example are followed, then the resulted data rate for PSIP is around 2.78 Mbps.

Those previous examples aim at presenting the data rates involved by PSIP tables carriage: from several hundreds of Kbps to several Mbps for a complete EPG, those data rates are in no way negligible in comparison with the data rates involved by video carriage: 4-5 Mbps for a standard definition program to up to 19 Mbps for a high definition program.

Still, the quality of an EPG does not only depend on the guide depth and of the data rate of the PSIP tables. Nothing can work properly if the PSIP tables are not first correctly injected by the broadcasters, and then carried without impairments through the entire telecommunications network to the end user. The following paragraph deals with PSIP implementation at the level of the broadcaster and discusses a possible architecture for a DTV station. PSIP implementation will then be tackled from a network point of view.

Broadcasting Station Implementation Issues

Early implementations of PSIP will be primarily static and manually programmed. In some implementations there will also be a separate and stand-alone programmed sequencer to change the VCTs and EITs on a timed basis. Eventually, stations will want their PSIP information updated automatically. Also, they will want to eliminate and avoid duplicative steps of the workflow within their facilities.

In these early implementations entry to the PSIP tables will often be done via the computer terminal upon which the PSIP generation system resides. However, these operations closely parallel some of the functions of traffic and automation. If the functions of traffic and automation could be made to drive certain functions within PSIP then the station will avoid duplicative work and potentially conflicting and erroneous information within the transmitted ATSC stream. One such potential arrangement is shown in figure 1. In this example the traditional traffic and automation interface is shown, but the additional processes of PSIP generation and encoder and multiplexer management are added.

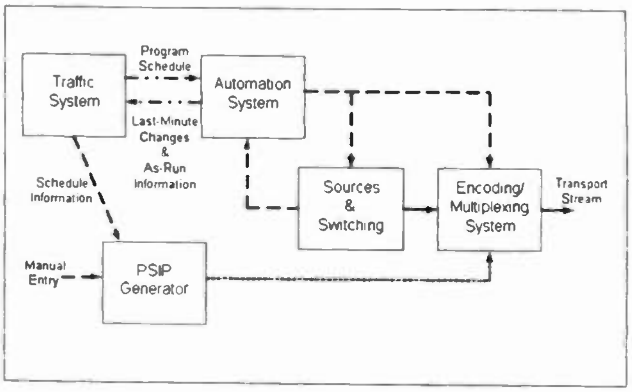


Figure 1

In a typical traffic and automation system future schedules are generated and stored within the traffic system. The current program schedule extending out from a few hours to a day is periodically transferred to the automation system which in turn controls the on-air schedule. Occasionally there are last-minutes changes made to the schedule which are made directly within the automation system. These changes are due to a variety of reasons including live, late-breaking events, damaged or lost media, equipment failure, etc. Traditionally, upon execution or at a later time an 'as-run' schedule is sent from the automation system back to the traffic system for various operations such as billing for paid announcements. Since the EIT-0s are to represent the current schedule extending out three hours, this as-run log will serve no purpose since the as-run log contains a record of events that have already occurred. What is needed is a new process of an immediate feedback of program log changes from the automation system to the traffic system so that the current EIT-0s can be correctly updated.

At the time of broadcast in current traffic and automation systems there is information on current and future broadcasts in two locations, and immediately before and at the instant of a change this program information in those two locations will likely be conflicting. If the PSIP generator were to be accepting information from both then some decision process would be required in cases of these conflicts. It would be preferred to have all decisions made by the traffic or automation systems and feed the PSIP generator with a single data stream from one or the other, leaving the generator as a background task. Figure 1 shows one potential process where the traffic system accepts immediate updates back from the automation system and then creates a data stream consisting of programming and channel information for the EITs and VCTs, respectively. Other arrangements are possible as long as the EIT-0s information is correctly updated. The PSIP generator generally sits in the background and provides its output stream to the multiplexer where the data is multiplexed into the transport stream. Manual input to the PSIP generator may be done as required.

Head-end Implementation Issues

It is important to note that PSIP tables correctly injected by the broadcasters can be highly impaired during their carriage through a larger telecommunications network. A broadcaster may in fact be connected to a network head-end that may receive PSIP information for several other broadcasters. The head-end may also play the role of a pass-through and be cascaded to other head-ends. Furthermore, at the opposite of the broadcasting range, local actions may extract PSIP tables in order to make local program insertions. All those operations can in one way or an other deeply degrade the integrity of the PSIP information originally sent by the broadcasters, such that when arriving to the set-top boxes the PSIP tables are of no use.

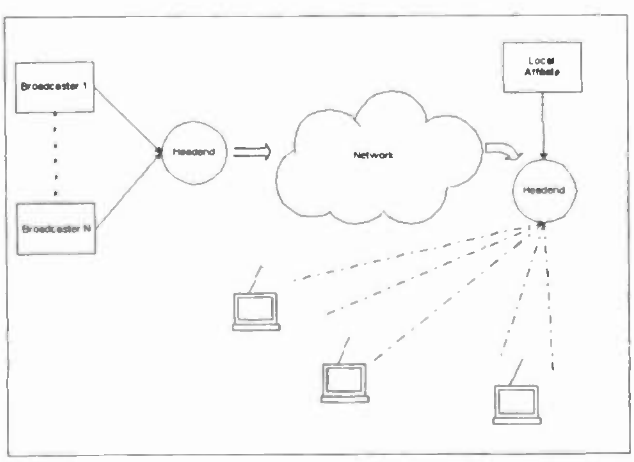


figure 2

It has been previously shown how a broadcaster could efficiently feed a program multiplexer with PSIP information. Assume now that the broadcaster supplies at its station output a digital multiplex containing its programs and related PSIP tables, and the broadcaster's multiplex has to be sent to a head-end equipment to be shipped to the final user. The head-end is an entry point to a medium that can be a terrestrial, a cable or a satellite network. Each of those media is characterized by a typical bandwidth than can be seen as a big cupboard with some drawers that are exactly shaped to carry a digital television multiplex. The head-end must put in these dedicated drawers the television multiplexes coming from the different broadcasters (figure 3).

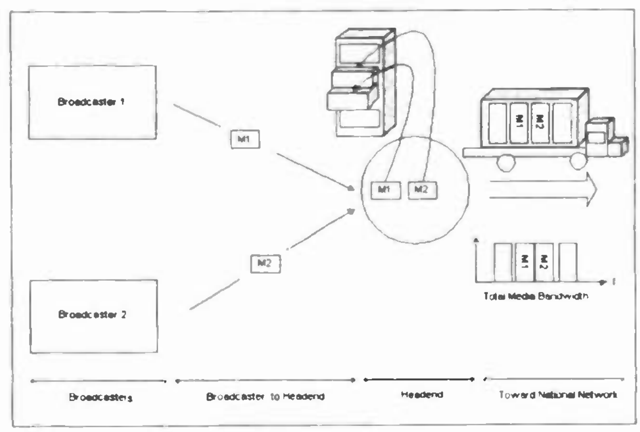


figure 3

The bandwidth allocated to the different DTV multiplexes is then transmitted to local network head-ends which pick up the contents of some drawers and then refill or re-feed the regional networks with the selected multiplexes.

It is said in the technical corrigendum to ATSC A/53 that each ATSC compliant 'Transport Stream shall include system information and program guide data'[3]. In others words, each drawer shall have its dedicated PSIP information and in the above scheme, both multiplexes M1 and M2 have the proper tuning information that a set-top box can use for channel hoping.

However, this scheme may prove itself to be inefficient with regards to the amount of bandwidth consumed for EPG-related information that is rarely consulted. In fact, if a broadcaster carried a full EPG, that is all the information related to 16 days of programs, it may be preferable for him to keep in his multiplex the minimum information required for basic tuning information (MGT and VCT), and to put content information (EITs and ETTs) in another multiplex partly dedicated to the EPG. In other words, the broadcasters may choose a common and dedicated drawer for the EPG to put together their program information but keep in their own drawer the basic PSIP tables required for tuning operation. In this case, the head-end has to fill in correctly the EPG drawer and then to send the resulting cupboard. Below, figure 4 represents the inside process of a head-end which retrieves from the incoming multiplexes M1 and M2 their related PSIP information, EPG1 and EPG2. The head-end then reconstructs new multiplexes, M'1 and M'2, which contain the minimum required PSIP information, and contructs a new EPG from EPG1 and EPG2. This new EPG is carried within a dedicated multiplex:

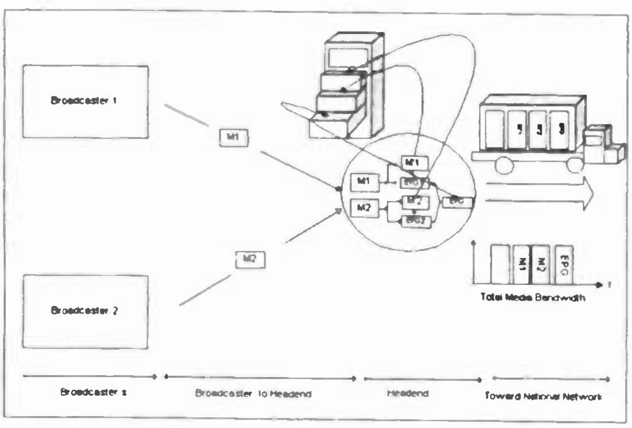


figure 4

In the above scheme, the end user can either select a multiplex and handle on it basic tuning operations or open an EPG by tuning to the dedicated multiplex.

The rest of this section discusses a possible technical solution of this idea, and reveals two real cases in which Thomcast has been involved: England and Sweden.

Let's now consider the head-end in the above diagram and let's assume that its fundamental tasks are to collect PSIP information and program data from the incoming streams, to compile them within a same transport stream, and to redistribute the subsequent stream to the proper networks. Since Broadcaster 1 is not informed of the doings of Broadcaster 2, possible conflict may appear in PID assignment. Broadcaster 1 may in fact have chosen to carry its video stream within transport packets labeled with a PID that, unfortunately, Broadcaster 2 has also selected. In the same way, conflicts may appear in the PID assignment for EITs and ETTs of our two broadcasters. Moreover, a conflict systematically occurs for the MGT, VCT, RRT and STT, since those PSIP tables are carried within the same PID. The ATSC standard solves the uniqueness problem, but this means that some proper tables are to be reconstructed from the incoming PSIP tables. Those conflicts appear typically at the level of the head-end shown in figure 4, and it is the responsibility of those head-ends to solve these issues.

Let's look at a full head end of this type and find a possible implementation for the PSIP part of the problem. It has been shown above that the basic functionality of that head-end is:

- The retrieval of PSIP tables
- The compilation task
- The redistribution task
- The Editing and Supervision task

These tasks can be thought as functional blocks, which are described in detail hereafter:

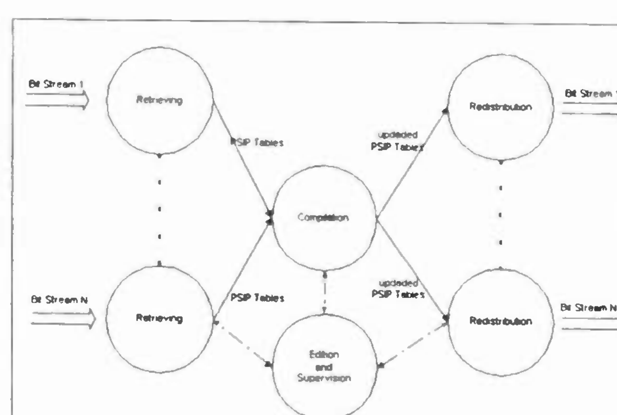


figure 5

Retrieving of PSIP information

The first operation to perform is to retrieve from each of the head-end incoming streams the set of PSIP tables originated by the broadcaster. Since the information arrives at the input of this functional block as a raw bit stream, there must be some kind of process able to reconstruct from the bit stream the different transport packets. Then, the transport packets may be parsed PID-by-PID to pick up all the packets labeled with the base PID. A new filtering operation on the type of table may be used to reconstruct the MGT, the VCT, the STT, and the RRT coming from the different inputs. The EITs and ETTs may at this stage be retrieved thanks to the reconstructed MGTs. The retrieving task is duplicated into as many copies as there are incoming streams.

Only the tables with a new version number, which are the tables needed to be updated, are sent to the compilation task. The other tables may be directly send to the redistribution task.

Compilation task

The input of this functional block is parsed to retrieve the new incoming PSIP tables.
The conflicts in PID assignment may then be solved, at the exception of the base PID.
Once the other PIDs are correctly reassigned, a new VCT may be generated which takes into account the changes that have occurred during PID reassignments. A new RRT can also be created in which all the rating rates are added. At that time, the MGT can be build, for the other PSIP tables except for the STT are informed. Some tools to check the compliance with ATSC standard are to be implemented in these blocks.

At the output of the compilation block, a set of new PSIP tables is supplied. If no change occurs, that is, the input of the compilation block does not provide any updated tables, the compilation task does nothing.

Redistribution task

In the VCT as well as in the MGT parameters related to transportation may need to be changed. This may be required in such cases where a terrestrial network feed a cable network. Then the modulation may change from VSB to QAM ; the frequencies as well as TSID may also be affected. The redistribution task is duplicated into as many copies as there are outgoing streams.

Editing and Supervision task

Whereas the editing task allows to feed the compilation task with user's edited PSIP tables, the supervision task gives access to different system parameters, such as the PSIP tables to be filtered and the type of modulation to be used at the output.

Those tasks are to be performed in real time and the head-end has to ensure that the time cycles of each PSIP table within the output transport stream are respected.

Let's now consider a head-end of a local network. This head-end is typically located at the end of the broadcasting range. This may be for example, a station responsible to collect DTV signal from a national station and to redistribute it to a regional network. This station may also filter some channels and interact with local players who may do local program insertion. At the PSIP level, a local insertion implies not only some retrieving and presentation operations but also some editing operation on the PSIP tables. Those issues of editing and filtering are revisited hereafter.

Editing task

The PSIP standard A/65 stipulates to carry PSIP tables within transport packets whose tranport_scrambling_control is set to zero, that is to say that those packets are not scrambled. A local head-end can therefore potentially retrieve all the PSIP tables of a particular transport stream, manipulate them and then re-inject them into a regional DTV stream. Some mechanisms must thus be implemented to avoid illegal operation. A head-end may handle the task to check legality of changes before sending the subsequent information to the local network.

Filtering task

A local affiliate may also filter a national program and replace it by its own. This task implies a corresponding filtering on the PSIP information in order to send only the relevant information to the regional network.

These different issues of retrieving, compilation, editing, filtering and redistribution have been extensively discussed within Thomcast and different systems have been developped and implemented both in Europe and in the US. Each system has been customized to take the specificity of the countries into account. The cases of England and Sweden, where Thomcast has been heavily involved, are described hereafter.

In 1996, United Kingdom launched its Digital Terrestrial Television Project and Thomcast was selected to implement an EPG system. The issues to be resolved were those previously discussed in this article. The EPG data needed to be collected by a central system prior to redistribution to local networks.

To target more closely the audience, the United Kingdom was divided in 29 broadcasting regions. A Central Service Information system was agreed upon to collect the EPG data for all of the UK. All this information is stored in a database that each operator can update in real time through a remote and secure gateway. The EPG information is then filtered by Thomcast equipment and sent to the 29 regions. The operation of filtering is done in real time on an 8 Mbps EPG data flow coming from the Central Service Information System and produces a 500 Kbps output EPG flow containing the relevant information for a particular broadcasting region. This output flow is then carried to each of the regions via the national network. An operation of local insertion can be made both at the level of programs and related EPG data.

For its EPG project, Sweden has agreed upon a slightly different approach than the United Kingdom. The country has also chosen to centralize all the EPG data in a common database, but Sweden has opted for a Thomcast system able to sustain in parallel fifteen compilation tasks at the level of EPG data to obtain specific EPG data flows. Those flows generated and sent to the different Swedish broadcasting regions.

Both of the countries have now a complete and secure solution, which ensures the proper carriage of EPG data.

Conclusion

PSIP is a required component of a complete DTV system. It is the underpinning of the Electronic Program Guide. At the minimum it is needed to ensure proper configuration of set-top boxes as well as reliable channel hoping and must be properly carried through the network. An erroneous set of PSIP tables can deprive many viewers of scores of television channels, including Bernie's set-top box.

PSIP implementations involve many processes even for basic configurations. Since PSIP will continue to evolve throughout the years, it is of the utmost importance that the actual implementations should be flexible and modular. This will allow the incorporation of new PSIP figures as they evolve.

PSIP therefore requires from the different partners in the television industry not only a strong know-how but also a good anticipation of what the PSIP future may be.

Further developments

The EPG quality suffers today from the small amount of bandwidth that the broadcasters are ready to invest and will also suffer from the small bandwidth that the cable network will negotiate with the broadcasters. It is nevertheless understandable that a service such as the EPG, which is consulted very rarely in comparison with the video programs, be assigned a bandwidth in direct relationship with the profits that it can generate. Still, at the edge of the multimedia revolution, the EPG may be considered as one of the great commercial arguments for DTV. With good software downloaded in the set-top box, the end user can be offered a powerful browser to navigate through the scores of channels within the DTV transport stream. Moreover, it may be useful to recall that the more information there is, the more crucial the navigation tools. If it is true that a good EPG leads to less time spent in navigation, it is also true that the aftermath of a bad EPG may well be a smaller audience who is disappointed by the time spent in browsing instead of watching. With the current PSIP standard and the bandwidth constraints, such a good browser, if purely textual, can be achieved.

Still, much greater bandwidths than nowadays can be achieved: Terabits are now obtained thanks to wavelength multiplexing technology and fiber optical networks. Better compression rates will also come along thanks to new mathematical tools such as wavelets. In a near future, the EPG can thus not only be textual but also visual, showing pictures, diagrams, and even clips. Those new types of data can be added in the PSIP tables without breaking the whole system. As an open standard, PSIP allows the creation of new types of tables referenced thanks to the Master Guide Table. An Extended Video Table may then be thought of, containing a pointer to a clip related to a particular event, and a decoder aware of this new extension may display the clips along with the textual information.

Because the underpinning of the future EPG is the EPG of today, PSIP implementation issues are as many challenges to solve as new steps to take towards a complete digital world. The extra costs involved by PSIP implementation are therefore also a long-term investment, which may in turn become a guarantee for the investments related to digital video encoding and multiplexing equipment.

About the Main Author

Pierre Clement joined Thomcast in 1997 and worked in the development of a System Information Management System. He is now Product Development Engineer for CDS of Thomcast Communications, Inc., where he is responsible for the adaptation to the ATSC market of MPEG based products which were originally developed for the DVB market.

Mr. Clement holds a Diploma of Engineer from the *Institut National des Telecommunications*, a major French telecommunications school (equivalent to a Master of Science in Telecommunications).

Contact:

Phone: (703) 519 63 05
Email: pclement@comarkcom.com

References:

[1]'ATSC Standard A/65 *Program and System Information Protocol for Terrestrial Broadcast and Cable*'

[2] 'Draft Corrigendum for Program and System Information Protocol for Terrestrial Broadcast and Cable (A/65)'

[3]'Proposed Technical Corrigendum No. 1 to ATSC Document A/53, 16 Sep 95, ATSC Digital Television Standard'

[4] 'DVB Standard *Digital Broadcasting systems for television; Guidelines on implementation and usage of service information (SI)*'

Remote Broadcasting Techniques and Technologies

Thursday, April 22, 1999

9:00 am - 12:00 pm

Chairperson: Jeff Andrew
Gannett Broadcasting, Arlington, VA

***9:00 am DSNG: Getting the News Feed to the Studio**
Robert Pape
NDS Limited
Southampton, United Kingdom

9:30 am FQPSK Use for Electronic News Gathering (ENG), Telemetry and Broadcasting
Hans Emmenegger
Broadcast Microwave Services, Inc.
San Diego, CA
Kamilo Feher
FQPSK Consortium and Digicom
El Macero, CA

***10:00 am 64 QAM Digital Radio Propagation Testing**
Jerry Brown
Alcatel
Richardson, TX

10:30 am Multi-channel HDTV Transmission System Using Dense WDM for Full-scale On-site HDTV Broadcasting
Shuichi Fujisawa
NHK Engineering Services, Inc.
Tokyo, Japan

11:00 am New Internet and Broadband Video Technology for Remote Television Production
Etienne Mirlesse
Obvious Technology, Inc.
San Francisco, CA

11:30 am Live Transmission System for the Tokyo Marathon Using 800 MHz COFDM Links from Moving Vehicles
Jiro Hirono
Fuji Television Network, Inc.
Tokyo, Japan

---

*Papers not available at the time of publication

FQPSK Use for Electronic News Gathering (ENG), Telemetry and Broadcasting

Kamilo Feher, Ph.D.
University of California, Davis and
FQPSK* Consortium – Digcom, Inc.
44685 Country Club ,El Macero,CA 95618
Tel 530-753-0738; Fax 530-753-1788

Hans Emmenegger
Broadcast Microwave Services, Inc. (BMS)
A Cohu Company, 5636 Ruffin Road
San Diego, CA 92123-6388
Tel 619-560-8601; Fax 619-560-1637

ABSTRACT

A new generation of spectral efficient, robust performance Electronic News Gathering (ENG) and broadcasting systems is described. These systems, as well as other wireless and telemetry "dual-use" commercial and U.S. Government DoD and NASA systems, use Feher patented Quadrature Phase Shift Keying (FQPSK) transceivers [1]. A typical ENG transmission at 7.5Mb/s requires 7.5 MHz of RF bandwidth. High spectral efficiency, combined with C-class RF power efficient amplifiers and robust performance, is the reason why the military test ranges are mandating the use of FQPSK for high data rate transmission from airborne test objects to the ground receive site. Test results of commercial ENG transmission from HELICOPTERS to central receive site using FQPSK driving a BMS transmitter will be shown.

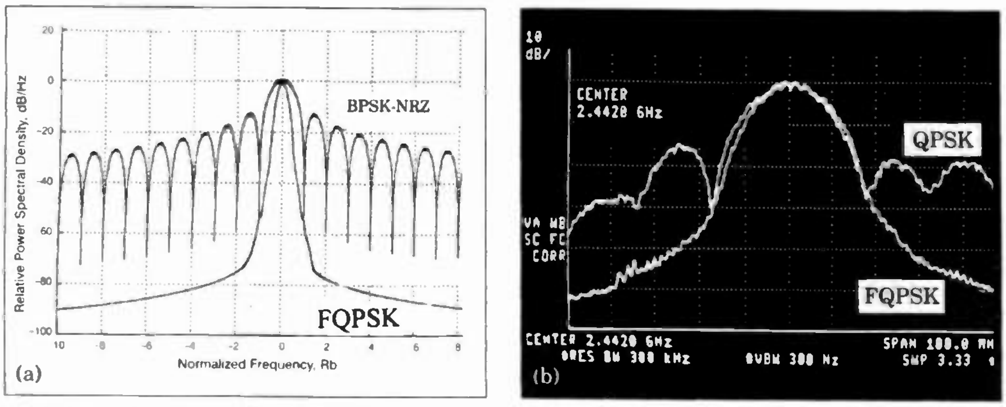
1. INTRODUCTION

Transmission of increased bit rate Electronic News Gathering (ENG) signals (audio, video, data) requires RF spectral efficient solutions. Related synergism with the broadcast industry, telemetry applications and with wireless systems lead to significant new discoveries, technology and product developments for spectral (bandwidth) efficient broadcast solutions.

Feher patented Quadrature Phase Shift Keying (FQPSK) technologies and products demonstrated significant spectral saving and RF power efficient robust BER performance advantages. These Digcom, Inc. licensed bit rate agile modems and Non Linearly Amplified (NLA)-or C-class amplified transceivers, DSP and hardware implementations, and in some instances "software-radios" (20kb/s to more than100Mb/s) and RF frequency agile (from 150MHz to more than 40GHz) developments and systems have recently been demonstrated and deployed. The spectral efficiency, i.e., data throughput capability of the 1st generation of FQPSK, as demonstrated in US Department of Defense (DoD) - Advanced Range Telemetry (ARTM) flight tests and other NASA and commercial tests approximately <u>doubles</u> while 2nd generation "FQPSK-2" systems have the potential to <u>quadruple the spectral efficiency</u> of operational systems. It is also demonstrated that the spectral efficiency advantage of FQPSK over that of NLA power efficient GMSK, OQPSK and QPSK modulated transceivers is in the 50% to 300% range and that the potential spectral efficiency advantage of FQPSK-2 over GMSK [1] is in the 200% to 500% range.

Based on extensive multi-year studies of alternative solutions for spectral and RF power efficient, robust BER performance systems, several commercial U.S. and international organizations, AIAA, CCSDS, NASA, ESA, CCSDS and various programs of the US Department of Defense (DoD) concluded that FQPSK offers the most spectrally efficient high performance-high speed proven technology solutions and recommended FQPSK standardization for several data links. Initial DoD-ARTM Program Office Air-to-Ground L-band and S-band jet airborne telemetry Test and Evaluation (T&E) data, obtained during 1998 are briefly highlighted. The American Institute of Aeronautics and Astronautics (AIAA) draft modulation standard recommended to the DoD, NASA and CCSDS that FQPSK modulation be immediately adopted as the interim *increment–1* standard." HELICOPTER Air-to Ground ENG test results will be included in the conference presentation.

**Significant parts of the material in this publication are based on publications, inventions and patents of K.Feher et al. and the rights for these parts remain with K.Feher-Digcom,Inc.-Ref.[1] (This File:NAB.99.paper.1.29.99b)*



Figures 1a and 1b. Spectral results of NLA (Non Linearly Amplified) FQPSK [1], and of filtered QPSK and BPSK after the output of the NLA RF amplifier. (a) From a NASA/JPL report , note that FQPSK (lower trace) is considerably more spectrally efficient than BPSK-NRZ and that FQPSK is the most spectrally efficient NLA alternative. (b) Hardware measurements on this Lockheed Martin (L-3 Communications, Inc.) designed FQPSK-QPSK modem over a 2.44 GHz, 1 watt commercial 34 Mb/s system demonstrate the significant spectral advantages of FQPSK over that of filtered QPSK.

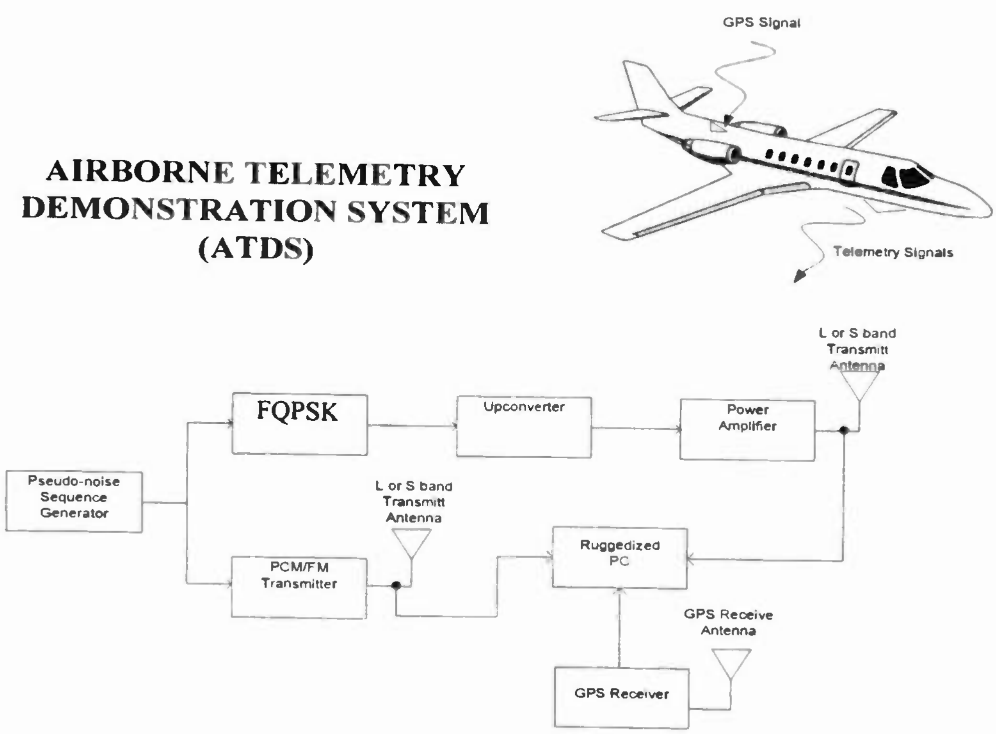


Fig. 2 US DoD's ARTM Program- Airbome Telemetry Equipment Demonstration System and Ground Station Telemetry used for FQPSK and PCM/FM flight tests in the L-band and S-band, with transmit power of about 10 Watt & flight distances of about 300 km [3-5]

2. FQPSK Offers a Proven Technology Solution for Robust Performance Increased Data Rate in Decreased Available Spectrum

The *most important - efficient high performance communications requirements* include:

- spectral efficiency (e.g., out-of-band Integrated Adjacent Channel (ACI) spectrum @ –70dB)
- robust BER = $f(E_b/N_0)$
- Non-Linearly Amplified (NLA), i.e., fully saturated or C-class transceivers

In this paper NLA spectrally efficient FQPSK (Feher patented QPSK), See Reference [1], transceiver developments having significant advantages over linearly amplified modulated and over other NLA systems are highlighted. References, including [1-5], present technical information and other relevant data related to FQPSK. Performance charts demonstrate that RF power efficient FQPSK systems double (200%) the spectral efficiency over that of compatible OQPSK, GMSK systems (having a comparable simple hardware and software - DSP implementations) and that FQPSK is more than 400% spectrally efficient than filtered NRZ-BPSK, MSK or NLA conventional QPSK [1-5]. The spectral efficiency (data throughput capability in an authorized RF spectral band of FQPSK) is double that of the currently-operational PCM/FM telemetry systems. It is also demonstrated that FQPSK operates over the PCM/FM installed base infrastructure, including entire receivers and down-converter IF stages.

In an AIAA January, 1998 announcement it is stated that approaches for new standards should consider only proven techniques and should meet the following performance guidelines:
(1) capable of operation at bit rates of 1Mb/s and above while achieving; (2) high RF spectral efficiency using; (3) non-linearly amplified (e.g., fully saturated) RF devices without additional IF or RF filters, and
(4) displaying robust bit error rate performance without coding.

The FQPSK technologies meet and exceed the aforementioned AIAA-stipulated requirements and, as of the Summer of 1998, FQPSK [1] has been recommended by the AIAA for standardization by the DoD, by NASA and by the international CCSDS[22] for DoD-NASA–CCSDS applications and other "dual-use" spectrally and RF power efficient standards [15-23].

For **dual-use commercial and defense** technologies and products for U.S. and international applications, FQPSK spectral saving and bit rate & RF frequency agile digital radio transceivers demonstrated better BER performance than compatible GMSK, MSK, OQPSK and QPSK, PCM/FM.

3. Increased Data Rate Requirements in Reduced Spectral Environments:

A CHALLENGE for Electronic News Gathering(ENG) ,Telemetry, Wireless Services and the BROADCASTING industry

An Illustrative Example: A 34Mb/s rate filtered QPSK [2], fully saturated or C-class RF power efficient 0.5Watt to 100 Watt RF transmitter operated between 150 MHz and 5.7GHz or other RF bands exhibits a significant spectral restoration; See Fig.1. Linearly operated high power RF amplifiers are too expensive or too large, have unacceptably large gain and power variations and/or are not available for low dc voltage/low power, e.g., 3V dc battery operation. For these reasons C-class fully saturated cost/power efficient smaller NLA transceivers have to be implemented. The NLA spectral efficiency improvement attained by **FQPSK over filtered QPSK** in the critical –40dBr to –70dBr range is **more than 300% and over filtered OQPSK more than 200%** The Adjacent Channel Interference (ACI) results, –65dBr, demonstrate the approximately 2:1 FQPSK (and 4:1 of FQPSK-2) data packing advantage over GMSK (matched 4^{th} order Gaussian receive filtered GMSK .

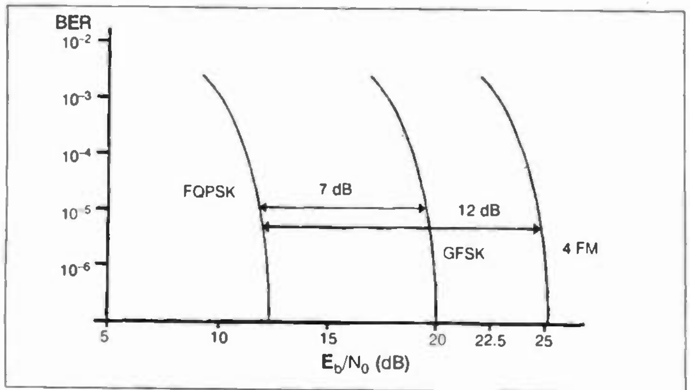


Fig. 3 **Measured BER = $f(E_b/N_0)$ curves of several class C RF IC/NLA modulated systems at f_b = 1 Mb/s and 2 Mb/s rate. FQPSK, GFSK with 160 kHz deviation and digital 4FM shown. Illustrative experimental data was submitted to WLAN and PCS standardization committees such as IEEE 802.11 and TIA/JTC. The experimental data show that at the specified BER = 10^{-5}, FQPSK is 7 dB and 12 dB more robust than GFSK and 4FM, respectively. Such dramatic performance improvement in an interference controlled environment, e.g., FCC Part 15, can increase the throughput rate about 100 to 1000 times. Relatively small 160 kHz deviation is specified in order to meet the FCC Part 15 and IEEE 802.11 spectral efficiency and out-of-band attenuation requirements [2].**

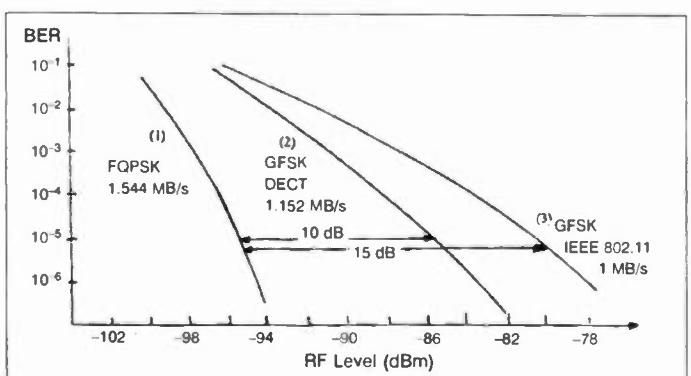


Fig. 4 **Performance comparison of measured, typical BER curves of NLA constant envelope standardized DECT (GFSK, 1.152 Mb/s at 1.9 GHz in 1.75 MHz); IEEE 802.11 WLAN (GFSK, 1 Mb/s at 2.45 GHz in 1 MHz with FCC-mandated −20dBr) and superior performance Lockheed Martin-manufactured EB200KF with Celeritek IF/RF (FQPSK 1.544 Mb/s at 2.45 GHz). For FCC Part 15 required −20dBr, the spectral efficiency of FQPSK is double that of the IEEE 802.11 and DECT-standardized and other MSK-based systems. The above FQPSK measurements have also been confirmed with RF Networks, Inc. 1 Mb/s "clear mode" and also** FQPSK **DS-SS (spread spectrum 17 Mchip/s) systems**

4. A COMPARISON of FQPSK and GMSK

COTS products: Coherent QPSK based systems, such as OQPSK and inter-operable and compatible FQPSK have been used and manufactured in large volumes in the USA and globally at higher than 1Mb/s rates – to several 100Mb/s. GMSK has only lower speed coherent/high performance (e.g., 270.833kb/s GSM) COTS products.

FQPSK performance advantages over that of GMSK include:

BEP robustness (approx. 1-2dB) FQPSK advantage over GMSK @ BEP=10^{-2} to 10^{-4} range

FQPSK spectral efficiency advantage 25% to 100% range if Nonlinearly Amplified (NLA) Transceivers

2nd generation FQPSK-2 spectral efficiency advantage over GMSK 300% to 500%

Simpler Product, e.g. FQPSK has 4th order Tx-Rx filters versus 100-plus taps GMSK

Smaller size - less dc power and simpler DSP requirements – e.g. 4 sample/ symbol versus 8 sample/symbol

Considerably lower cost implementation for above 1Mb/s rate

FQPSK high speed (1Mb/s, 3Mb/s and 17Mb/s to 40Mb/s) hardware products COTS have been demonstrated - while high speed GMSK is much more complex; still in R&D.

5. Recent Background Information Highlights

- DoD's RDT&E Spectrum Requirements Working Groups (WG), based on previous CBD solicitations, considered several spectral efficient modulation proposals.
- AIAA NASA/JPL, DoD, industry and university extensive multiyear studies found that FQPSK is the most spectrally efficient robust BEP performance RF power efficient COTS modulation.
- The availability of Feher patented FQPSK and GMSK licensing and technology transfer on equal-opportunity, non-discriminatory fair market value basis for dual-use commercial and military applications has been announced [1].
- AIAA-approved recommendation to DoD, NASA and CCSDS to standardize on FQPSK

6. FQPSK Used in AIAA-NASA-DoD Specifications

Following the CBD announced critical community review of FQPSK and GMSK and of other proposed alternative technologies, FQPSK, the most spectral efficient solution and robust BER performance high bit rate hardware and software proven technology with COTS (Commercially Off-The-Shelf) available products has been specified for the AIAA-NASA-DoD standardization project

Electronic News Gathering (ENG) and Broadcasting HELICOPTER based Air- Ground Transmission Tests

During 1999 several ENG and Broadcasting Helicopter Air-to-Ground transmission tests are planed. Test results will be presented during the conference presentation.

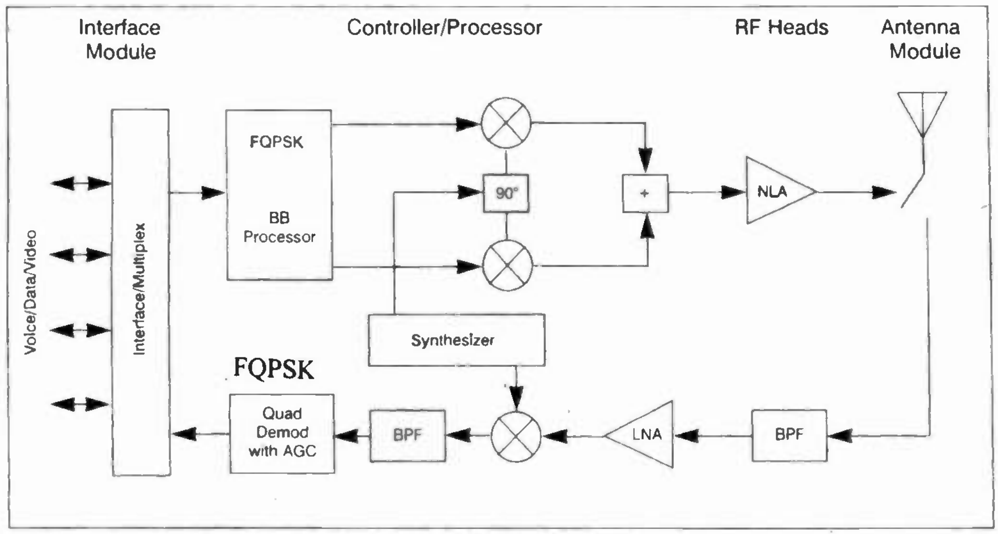


Fig. 5 **Transceiver block diagram of FQPSK**

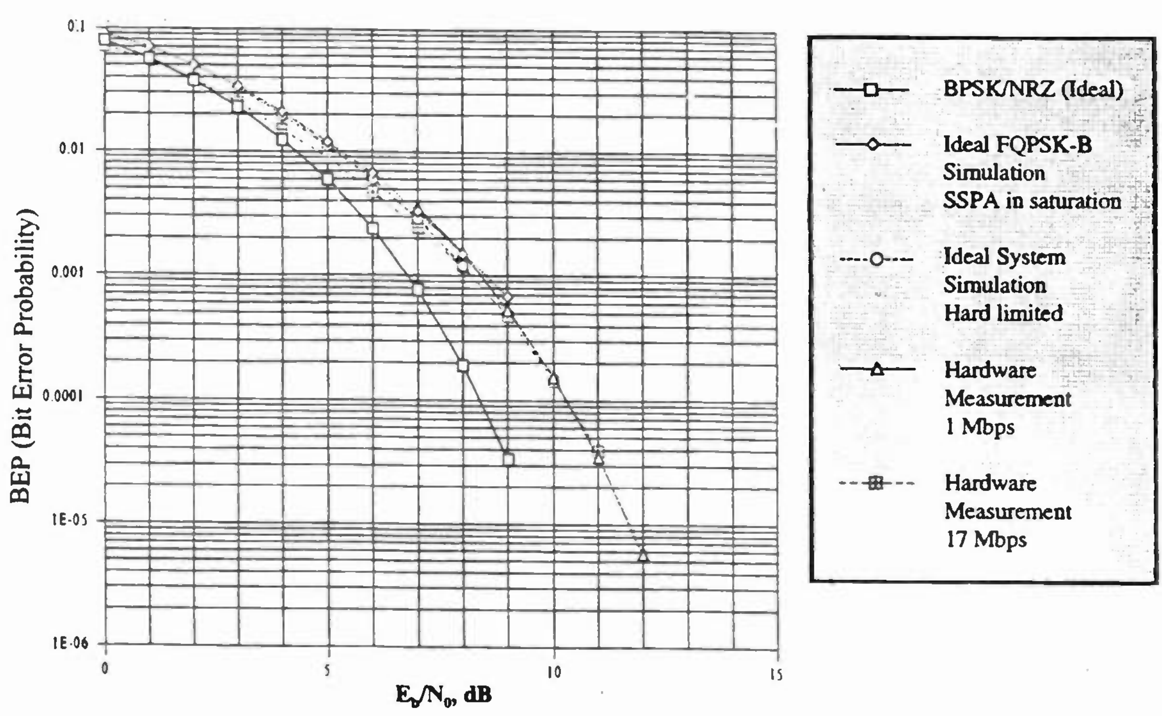


Fig. 6 NASA/JPL test results of FQPSK Transceivers operated at 8.4GHz, in NLA(Class C) mode at 17Mb/s over Lockheed Martin (now L-3 Communications, Inc.) manufactured modems and at 1Mb/s over RF Networks,Inc. manufactured FQPSK systems

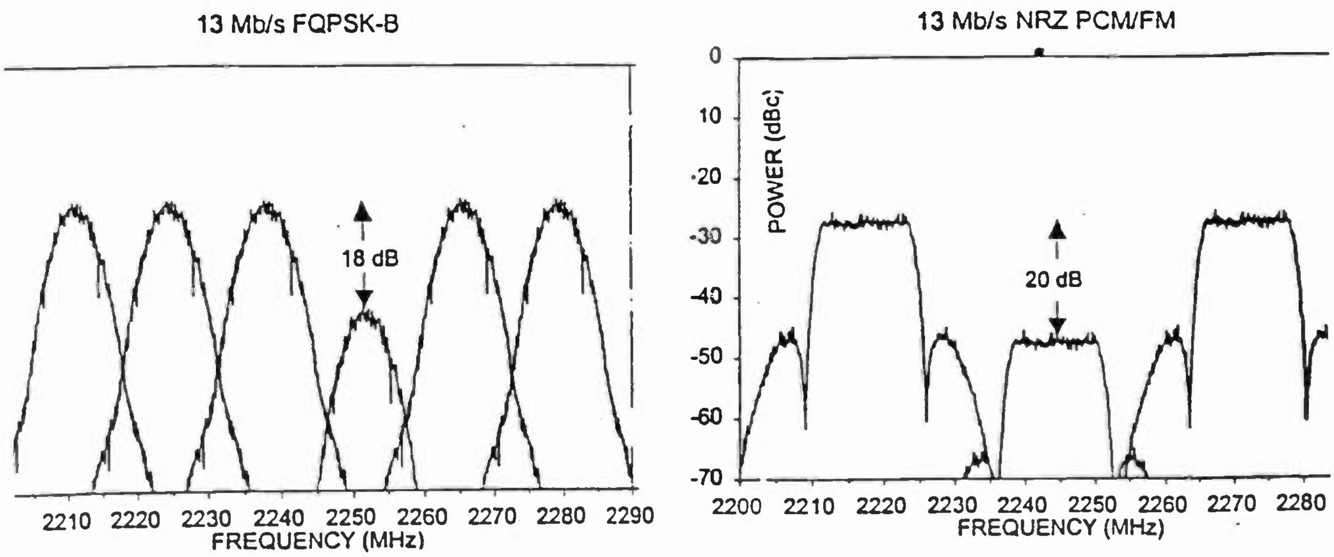


Fig. 7 Experimental hardware results indicate that FQPSK has the potential of doubling (200%) the data rate throughput capacity of operational PCM systems Based on preliminary measurements with FQPSK (6*13Mb/s) instead of PCM/FM (3*13Mb/s) would be attainable in the authorized 90MHz S-band.

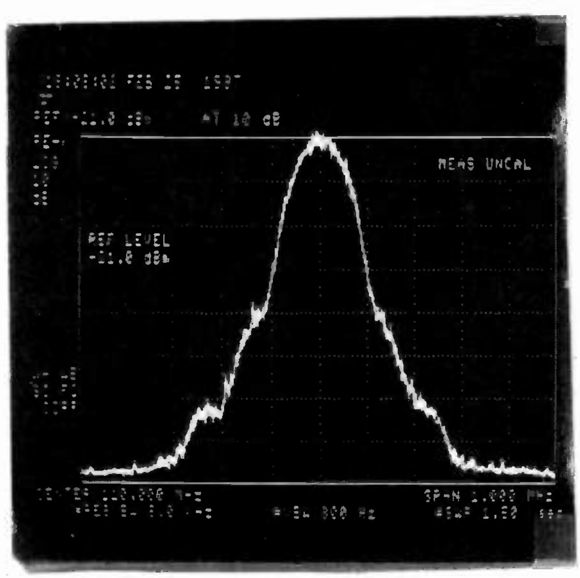


Fig. 8 Spectrum of a C-class-Non-Linearly Amplified(NLA) FQPSK transmitter operated at a relatively low bit rate of 192kb/s is illustrated in this 2.250GHz 10 Watt RF power measurement. Note the steep almost "water-fall" FQPSK spectrum to -76dB

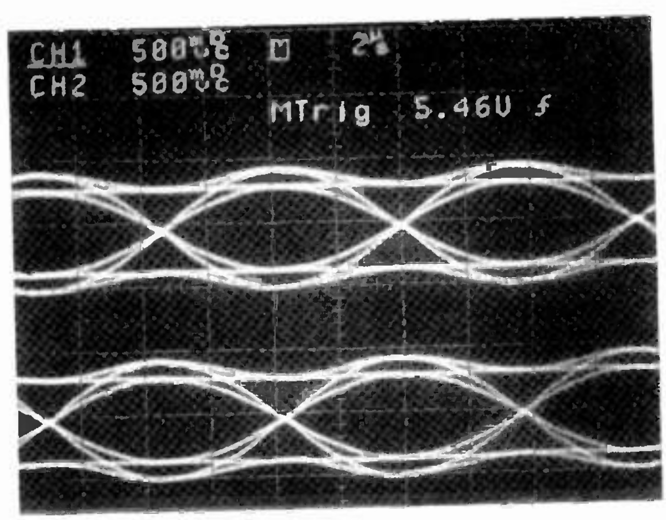


Fig. 9 The demodulated I and Q eye diagrams of this FQPSK system illustrate the robust performance attainable with C-class amplified FQPSK commercial products

7. ARTM Air-to-Ground Flight Tests of FQPSK at Edwards AFB, CA

The Advanced Range Telemetry (ARTM) tri-service Program Office, an Office of Undersecretary of Defense CTEIP funded program (of the U.S. Air Force, U.S. Army and U.S. Navy) developed an extensive Test and Evaluation (T&E) facility for new generations of telemetry systems. During the Summer of 1998, initial aeronautical simultaneous jet aircraft air to ground flight tests of FQPSK and PCM/FM have been undertaken at the US Air Force, Edwards AFB [3-5]. The predominant objective of these flight tests has been to compare the performance of FQPSK and PCM/FM modulation methods and transceivers using the aeronautical telemetry environment with PCM/FM as the baseline. Bit Error Probability (BEP) performance and total link availability assessment has been undertaken by the ARTM team.

During the Advanced Range Telemetry (ARTM) Conference Session at the International Telemetering Conference, ITC-'98 in San Diego, October 98, further details and test results were presented by members of the ARTM-DoD team and by NASA/JPL. Illustrative sample BEP and datalink performance results measured in the L-Band (approx. 10 Watt RF power at 1480.5MHz & 1485.5MHz and in the S-band indicate that the data link and BEP performance of the FQPSK is very similar to that of the baseline PCM/FM system. Similar conclusions have been reached for a large class of "Flight Decks" that is, different flight trajectories. T&E flights have been performed for the initial set of FQPSK and PCM/FM comparisons for distances of up to 300km at low altitudes of only 150 meters above ground level, medium altitudes of about 4,000 meters above ground, and of 10km and higher altitudes (for "supersonic flight corridors"). In the initial set of measurements the airborne transmit antennas were omnidirectional "blade" and "button" antennas located on the belly of the jet aircraft.

The primary objective of the flight tests for TIER-1 of the ARTM Program has been to Test and Evaluate comparable datalink availability and BEP performance of the FQPSK and of the PCM/FM systems. The spectral efficiency advantage of FQPSK (of approximately 2:1 of FQPSK) has also been demonstrated to the respective Spectral Managers. It is significant to note that in Line-of-Sight (LOS) environments, even at distances of about 300km, error-free intervals of 20 minutes or longer have been recorded on both systems, i.e., for LOS environments there was no noticeable "error-floor." For example, in one of the 20 minute measurement intervals at a 1Mb/s rate, 1,200 seconds x 1Mb/s = 1.2×10^9 bits were transmitted with a 20-minute average BEP of less than BEP=10^{-9}. While the aircraft was maneuvering, selective fade – NLOS (Non-Line-of-Sight) induced burst errors have been observed on both systems.
During the 3rd quarter of 1998 –ARTM flight tests with 5Mb/s rate FQPSK-B have been undertaken at Edwards Air Force Base, CA.

References

1. K. Feher et al.: US Patents 4,567,602; 4,339,724; 4,644,565; 5,491,457; 5,784,402 and Canadian Patents: 1,211,517; 1,130,871; 1,265,851 and post-patent improvements, including Digcom's proprietary 'know-how' and other US and international patents pending/in applications-in-process. These have been licensed on an equal-opportunity, non-discriminatory fair market value basis for "dual-use" Feher-patented Quadrature Phase Shift Keying (FQPSK) and Feher patented GMSK commercial and military use, make, sale, developments and other use and applications to large, medium size and small corporations and organizations in the USA and internationally by Digcom, Inc. For FQPSK Consortium, FQPSK–GMSK Feher patented licensing and technology transfer information and for FQPSK-2 (Feher Quadrature Phase Shift Keying 2nd generation) write to:
   FQPSK Consortium-Digcom, Inc., c/o Dr.K. Feher 44685 Country Club Dr.; El Macero, CA 95618; Tel.530-753-0738; Fax 530-753-1788.
2. K. Feher: *"Wireless Digital Communications: Modulation and Spread Spectrum Applications."* Book published by Prentice Hall, 1995,
3. E.L. Law: "Robust, bandwidth efficient modulation methods for aeronautical telemetry applications; spectral occupancy, adjacent channel interference and bit error probability comparisons," *Proceedings of the of European Telmetry Conference, ETC'98*, Germany, May 5-8,1998.
4. K. Feher: "FQPSK Doubles the Spectral Efficiency of Operational Telemetry Systems," *Proc. of European Tel. Conf., ETC'98*, Garmish-Patternk., May1998.
5. K. Feher :"FQPSK Transceivers Double Spectral Efficiency of Operational Systems: Advances, Applications, Laboratory and Initial Air to Ground Flight Tests" Proceedings of the International Tememetering Conference ,ITC/USA'98, San Diego, California , Oct.26-29,1998

Multi-channel HDTV transmission system using dense WDM for full-scale on-site HDTV broadcasting

Shuichi Fujisawa, Yasuaki Nishida, Shun Morisawa, Yosuke Endo
NHK (Japan Broadcasting Corporation)
Tokyo, Japan

1. Abstract

When sporting events, such as golf tournaments, are broadcast using high-definition television (HDTV), many metal cables have to be laid between the cameras, the broadcasting vans, and the switching center. To simplify the cable laying, we have developed an HDTV transmission system capable of transmitting 30 HDTV serial-digital-interface (SDI) signals maximally through a single-mode optical fiber. The system uses such advanced technologies as dense wavelength-division multiplexing (DWDM) and optical signal amplification using an erbium-doped fiber amplifier. The input/output SDI signal format of the system is SMPTE 292M (BTA S-004).

The system consists of three transmitter/receiver devices, optical-fiber cables, and an optical multiplexer. The optical cables are the ones commonly used for HDTV cameras. We have also developed a prototype camera and an adapter for connecting the camera to the camera-control unit (CCU) by using DWDM. The CCU and camera adapters are cascaded with the camera cables, making it easy to lay the cables. These adapters handle the camera-control signals and intercommunication signals in addition to the HDTV signal.

2. Introduction

NHK is preparing to begin satellite digital broadcasting in the year 2000, featuring digital high-definition television (HDTV) and multimedia data broadcasting. To encourage the widespread use of satellite digital broadcast receivers, it is essential to offer attractive services, such as HDTV news, drama, and large-scale live events. For that purpose, NHK is developing an HDTV news center and studio, as well developing HDTV equipment for broadcasting live programs.

This report describes the multi channel HDTV transmission system we have developed for broadcasting large-scale live events, such as golf tournaments.

3. Dense wavelength division multiplex (DWDM) technology

3.1 DWDM transmission

The DWDM transmission mechanism is shown in Fig. 1. Optical signals of different wavelengths are output by optical transmitters A through N. These signals are wavelength multiplexed by an optical coupler and input to an optical receiver via a single-mode fiber. A tunable optical band-pass filter (BPF) in the receiver is used to select the required signal from among the multiplexed signals. In addition, the system uses an erbium-doped fiber amplifier.

3.2 Erbium-doped fiber amplifier (EDFA)

The EDFA is used to compensate for the optical losses, such as transmission loss in the optical fiber and the insertion loss at connectors and couplers. The EDFA can amplify an optical signal with any wavelength of around 1550

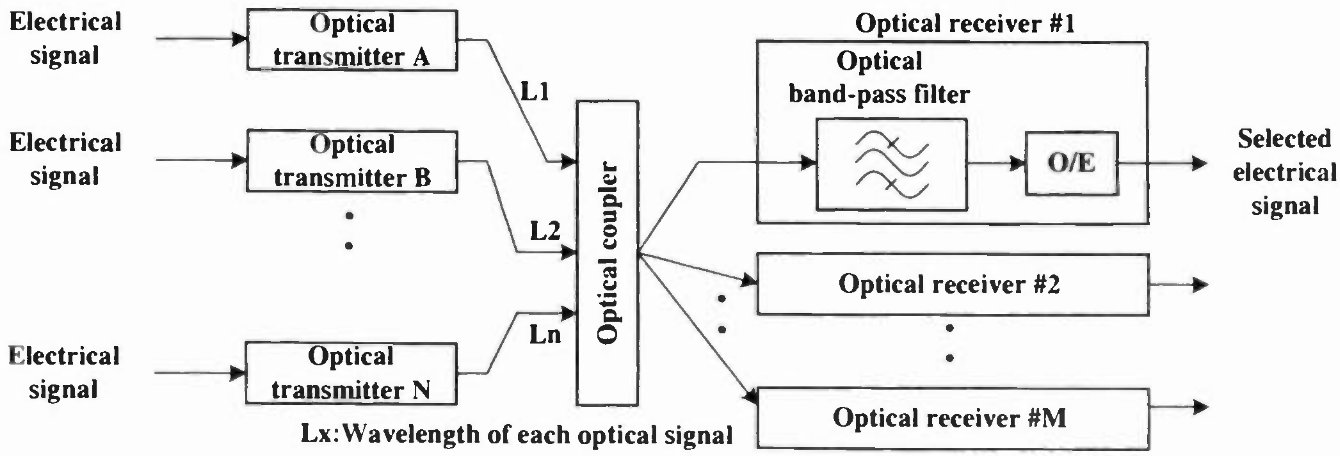


Figure 1. Fundamental configuration

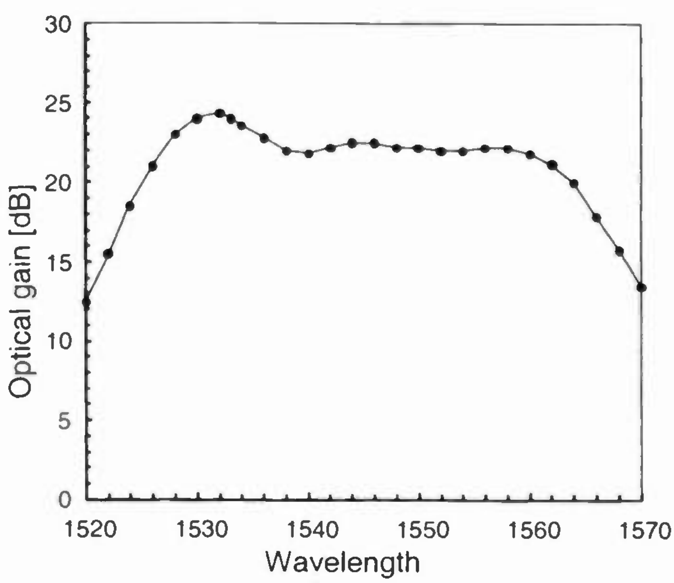


Figure 2. Band pass characteristics of EDFA

nanometer with a gain of 20-25 dB. An example of an EDFA band-pass characteristics is shown in Fig.2.

3.3 Tunable optical band-pass filter

The number of optical signals transmitted by the DWDM system depends on the frequency band-pass characteristics of the EDFA and on the interval between the wavelengths of adjacent signals. The wavelength interval should thus be made as small as possible to maximize the number of signals that can be carried. If the wavelength interval is too small, however, the optical BPF of the receiver will not be able to properly select the desired optical signal. Therefore, the characteristics of the optical BPF are the key to determine the number of optical signals that can be multiplexed in a DWDM transmission system. Furthermore, the optical BPF characteristics must not be affected by changes in temperature or by vibration. For these reasons, we chose to use a multi-layer dielectric interface filter as the optical BPF in the transmission system. As shown in Fig.3, this filter is mechanically inclined relative to the inserted light so as to change the pass wavelength.

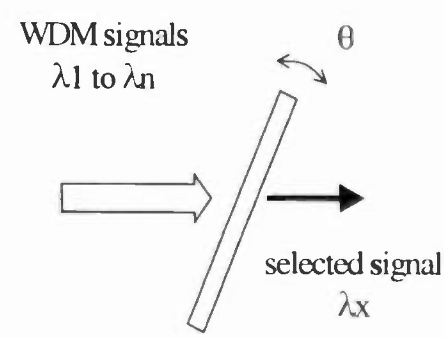


Figure 3. Multi-layer dielectric interface optical filter

The characteristics of the bit error rate vs. the received optical power are shown in Fig. 4. Because a 0.5 dB power penalty is required, an optical filter must be used to suppress the interference signal by at least 10 dB. In an actual system, adjacent interference signals exist at the upper and lower side, so the interference signals must be suppressed by at least 13 dB. In addition that, a margin of about 2 dB must be guaranteed. Totally, the interference signals must be suppressed by at least 15 dB.

Figure 5 shows the band-pass characteristics of the multi-layer dielectric interface filter. For a signal wavelength interval of 1 nm, the level of the adjacent optical signals can be suppressed by 20 dB or more.

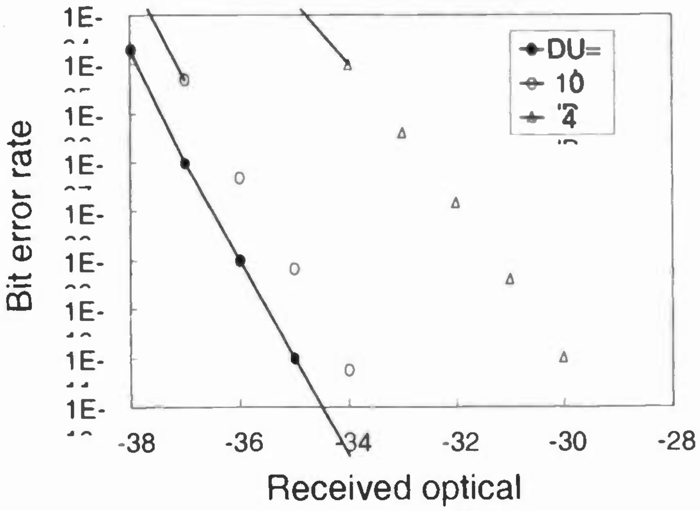


Figure 4. Characteristics of bit-error-rate vs. received optical power

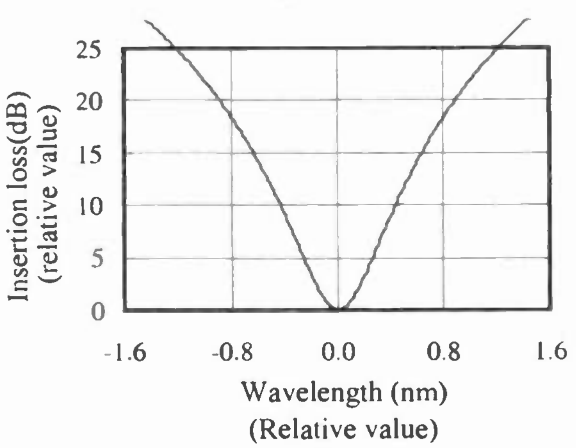


Figure 5. Band-pass characteristics of multi-layer dielectric interface filter

3.4 Wavelength interval

Experimental results show that a wavelength interval of about 0.6 nm between multiplexed optical signals is sufficient to suppress interference by 15 dB or more. On the other hand, the ITU has specified a wavelength interval for DWDM transmission of 0.8 nm (ITU-T Recommendation G.692: Optical interfaces for multi-channel systems with optical amplifiers). Therefore, we use a wavelength interval of 0.8 nm to comply with the ITU standard with the margin of 5 dB.

4. Basic system configuration

The basic system configuration of HDTV transmission using DWDM technology is described here.

4.1 Star-connection system configuration

We use the star configuration illustrated in Fig. 6. The outside broadcast vans (OB vans) are connected to a switching center by a star network. In the switching center and in each van, the HDTV serial digital signals are converted into optical signals of different wavelengths and then multiplexed by the optical couplers. The multiplexed signals output from the switching center and the vans are fed to an optical multiplexer to be multiplexed once again. In this way, all the HDTV signals from the switching center and the vans are multiplexed. The signal output from the optical multiplexer is returned to the receivers in each van by another optical fiber, making it possible for each receiver to select the required signal by means of the tunable optical BPF. The transmitting and receiving device (Tx/Rx device) and the optical multiplexer can thus be connected by HDTV camera cables consisting of two single-mode fibers. By using the camera cables for connecting the vans to the switching center, the signals can be transmitted up to a distance of three kilometers.

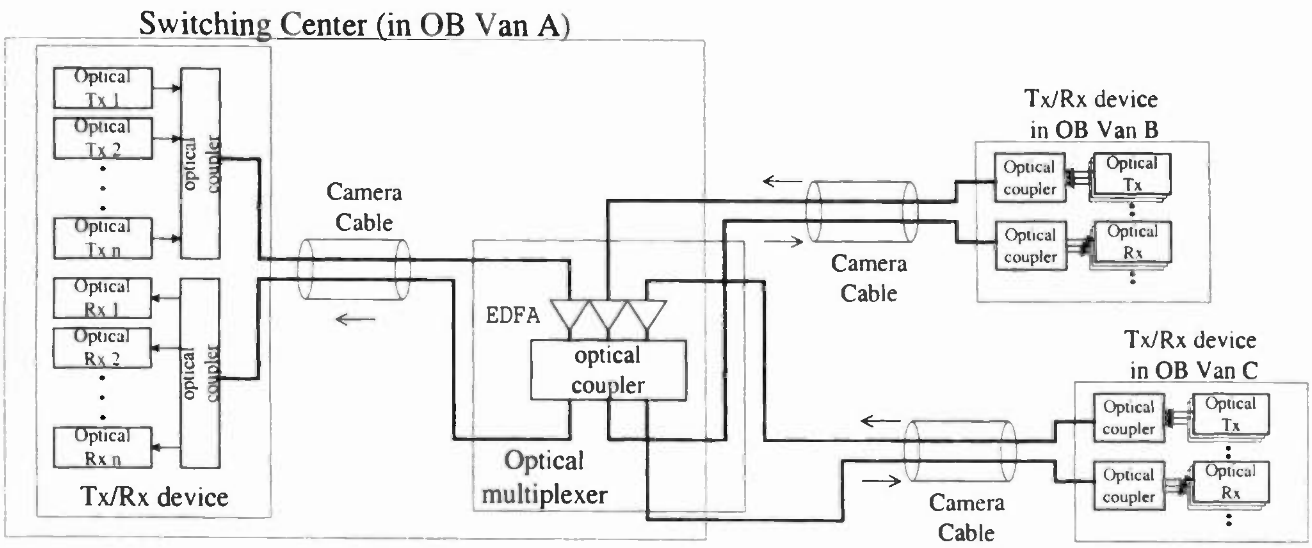


Figure 6. Star connection system configuration

4.2 Cascade connection system configuration

The system was designed to allow the use of the cascade-connection configuration shown in Fig. 7, as well as the star configuration. This configuration provides more flexibility in the placement of the switching center and vans; however, the maximum distance from the switching center to the furthest van is 3 km. In this configuration, camera cables can also be used.

4.3 Camera/CCU adapter-system configuration

The DWDM technology enables two-way transmission between an HDTV camera and a camera-control-unit (CCU). The connection between the camera and the CCU can be transparent with this configuration. The signals transmitted in the configuration are the video signal from the camera to the CCU, the control signals from the CCU to the camera, and the intercommunication signals. The camera-adapter is placed outdoors, so it must work well under harsh conditions, such as

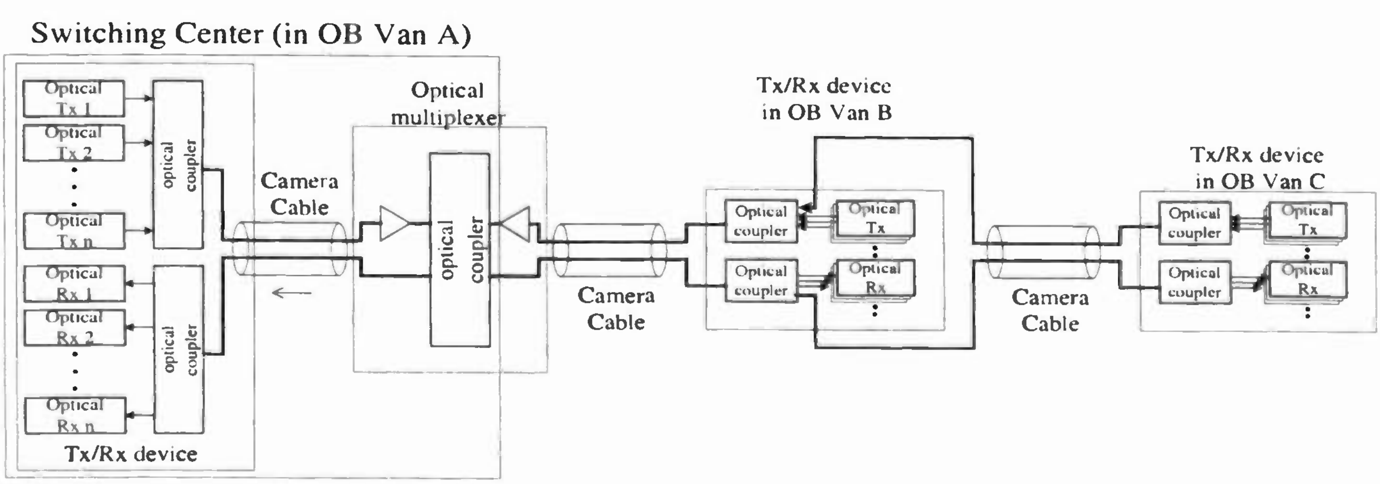


Figure 7. Cascade-connection system configuration

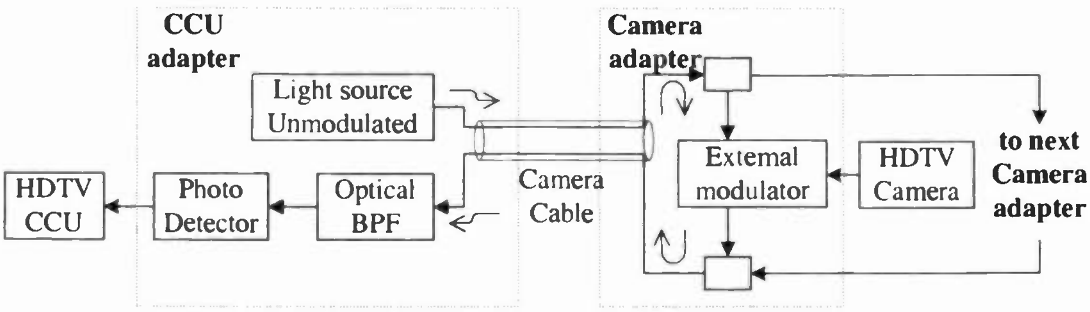


Figure 8. Camera and CCU adapter system configuration

during a hot summer or cold winter. For this reason, temperature-sensitive laser diode, used to transmit the HDTV signal from a camera to a CCU, is installed not in the camera adapter but in the CCU adapter, which is placed in the van. This configuration is shown in Fig. 8. This configuration enables 3-km transmission with three camera adapters.

5. Developed multi-channel HDTV transmission equipment

5.1 Configuration

The system consists of three transmitter-/receiver devices (Tx/Rx devices) and one optical multiplexer shown in Fig.9. The Tx/Rx devices and the optical multiplexer are connected by HDTV camera cables in a star network configuration or in a cascaded network configuration. The Tx/Rx devices have eight circuit-board slots which can be used to install either a transmitter board or a receiver board. For example, if Tx boards are inserted into all eight slots, eight HDTV serial-digital-interface (SDI) signals can be output. It is also possible to insert eight Rx boards or four of each type of board into the eight slots. Each Tx/Rx device has a control panel on the front side for controlling whichever type board is inserted. The control panel also can display the channel numbers of signals that are being transmitted within the network. Each Tx/Rx device can also be connected to a remote controller that has the same functions as the control panel. The appearance of the remote controller is shown in Fig. 10.

The camera adapter is shown in Fig. 11. The Tx/Rx device operates as the CCU adapter in an indoor location, such as in an OB van. The camera adapter is a compact device designed for outdoor operation.

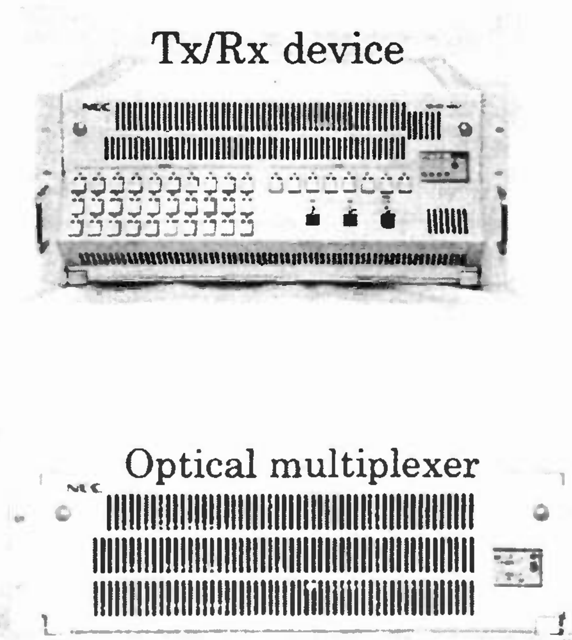


Figure 9. Tx/Rx device and optical multiplexer

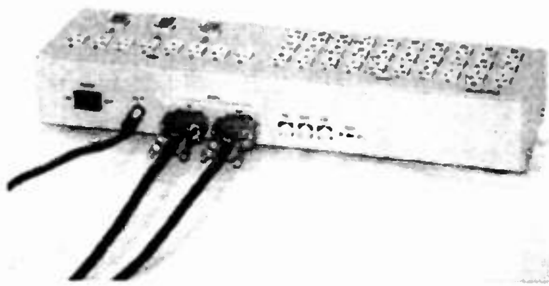


Figure 10. Remote controller

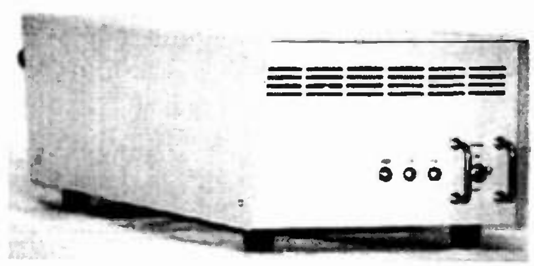


Figure 11. Camera adapter

5.2 Specifications

The specifications of the HDTV transmission system are listed in Table 1. The maximum transmission distance is 3 km when 250 m camera cable is used for the connections. The optical loss at each connection is approximately 1dB. With a longer cable, transmission of over 10 km is possible. Moreover, although the input signal format for this system is SMPTE 292M (BTAS-004), a format-free network can be constructed by changing the electrical signal receiver in the optical transmitter and the electrical signal driver in the optical receiver. It is also possible to increase the number of optical signals multiplexed by this system by using time-domain-multiplexing of the electrical signals.

Wavelength	1535 - 1565 nm
Number of signal	30 max.
Wavelength interval	0.8 nm
Input/Output signal format	SMPTE 292M (BTAS-004)
Bit rate	1.485 Gbps (1000/1001)
Type of optical filter	multi-layer dielectric interface filter

	Tx/Rx device, Optical Mux.
Dimensions	200(H) * 480(W) * 600(D)
Weight	31 kg
Consumption power	455 VA

	Optical Multiplexer
Dimensions	140(H) * 480(W) * 600(D)
Weight	15 kg
Consumption power	140 VA

	Camera adapter
Dimensions	240(H) * 320(W) * 730(D)
Weight	6 kg
Consumption power	45 VA

Table 1. Specifications of the HDTV transmission system

6. Conclusion

The multi-channel HDTV transmission system we have developed using DWDM techniques has two basic functions: route switching and video distribution. Accordingly, in addition to serving as a video-material switching system at the site of a live broadcast, it can be used as a video-signal distribution system for delivering video signals to the monitors of announcers and program directors. Because all of the camera video signals are multiplexed within the network, the announcers or directors can select the desired video signal.

Testing of this system for live broadcasting of golf tournaments and other such events demonstrated that it operates satisfactorily.

New Internet and Broadband Video Technology for Remote Television Production

By Etienne Mirlesse
President, C.E.O, Obvious Technology Inc.

Introduction:

The convergence Paradigm

In March 1997, the FCC granted US TV broadcasters the right to transmit Digital Video on terrestrial bands and via satellite in the United States.(The same revolution having taken place in Europe as early as 1994 with the "call for proposals" by the European Commission to produce and broadcast HDTV & 16/9 DVT contents.) The immediate effect of these initiatives was a global conversion of television and video production from analogue to digital formats. Broadcasters and equipment manufacturers are now switching to the new digital standard, which will be fully operational and world-wide within a few years.

By the same token, this revolutionary change from analogue to digital Video and multimedia standards such as Mpeg 1,3,4 & 7 has now made video broadcast fully compatible with computer networks – and thus "broadcastable" by conventional non-hertzian /network telecommunication means – such as telephone lines, fiber optics, LAN & ATM Networks, and even Cable TV.

This merge of Video, Telecommunication & computer Network technologies has brought about the birth of a multi-billion dollar industry –commonly referred to as "The Convergence Market". This market, in turn, generated new digital tools and media transaction based market, and with it, the need for a universal and affordable approach to the creation, indexing, query, management, and delivery of Broadcast content over computer networks.

TV network affiliates and advertisers to home viewers with set-top boxes, the entire broadcast industry value-chain is now seeking ways to exploit rich media content in revolutionary ways.

Hence it is safe to assume that the birth of the "convergence market" may in fact be the single largest post industrial revolution brought about since the invention of Television, the telephone and the PC.

However, while this unprecedented revolution is taking place, pre-existing telecommunication, and computer networks have proved totally incapable of handling the massive arrival of digital video efficiently on the Web .

Therefore, both new and conventional systems (digital video and computer management systems) are now forced to co-exist painfully and no integrated solution currently exists to efficiently combine both digital TV & new computer media formats.

The ability to easily manage video content in a clear, unambiguous – and most importantly - visual manner and sharing it over digital networks has thus emerged as the single biggest challenge to information based industries.

The timed media revolution

The early answer to Digital Video distribution on computer networks simply consisted of a breathless race for new compression formats which gave birth to such technologies as the Vxtreme , Quicktime and Real Media / Netshow PC video players.

Later came the birth of a "timed media" concept associating events contained in a Video Stream – with actual HTML documents. This concept translated in programming languages: The Real Media SMIL format or the Microsoft SAMI formats – both allow a web viewer to trigger the play of a ultra-compressed Video content simply by selected an HTML title corresponding to that content.

Simultaneously, new pioneering Technologies such as Virage – to cite but one - began utilizing close caption acquisition and its translation into text along with its association to video contents through SMPTE Video Time code as a means to index – and thus catalog or query those contents over the networks.

While these technologies constitute a clear progress in the management and distribution of Video content, and thus a step towards "remote Video production" capability, the computer industry at large still faces an enormous challenge to meet the growing broadcast-grade requirements for:

- Remote video content identification, and editing on the network
- Remote indexing, query, selection and download, of broadcast-grade videos in any format, from any network station, any where
- The annotation, and simultaneous collaboration of any video sequence over any existing network
- The remote selection , assembly and editing of Video sequences of interest, and their automated conform edit on Broadcast masters
- The embedding of any interactive data types in Video content and the simultaneous transmission of that Interactive data to computer or TV network clients
- Protection of Intellectual Property, Patents & Trademarks associated with the broadcast of Video content on computer Networks

And last but not least:

- the association of Broadcast Video Content with any other media-rich interactive content such as scripted annotations, closed captions, HTML pages, electronic transactions forms, watermarks and real time financial data.

The successful combination of all the above requirements would indeed allow for remote Television Production, editing & distribution on computer networks.

Redefining the priorities

Let us make here a parenthesis to better translate & redefine those requirements. The computer and broadcast industries are still separated by a communication and semantic vocabulary issue. Regardless of technology, the fundamentals of film, Video or Broadcast TV production and distribution have not changed since the birth of motion picture:
The translation of a few terms could help better identify the process:

- What was then called "filming" or "shooting" is now called " content acquisition"
- "Viewing dailies" is now translated into "Media content identification"
- Assembly & fine cut Editing is now called "non-linear editing"
- And "going to the movie" or watching TV" is gradually being replaced by "media streaming" or " chapter selections" on DVD menus.

But the fundamentals have not changed: In order to distribute and transact on broadcast content – it must first be scripted, shot, edited, and advertised for distribution.

So while the computer industry races to develop automated algorithms, new compression formats, wider band transmissions on computer networks it is in fact by-passing this fundamental process, which simply consists of allowing the collaboration of a wide number of individuals – ranging from producers, to directors, to editors, not to forget the viewers – to share the same content at one time or another – as to be able to collaborate and transact on it.

But more importantly, it is to allow individual to collaborate on the interpretation and associated use of that content. For instance, the decisions made by a TV producer on the set or TV viewer at home do not really differ in essence in so far as they are based on content identification & evaluation. It is then and only then, than content can be transacted on – whether the transaction consists of the creation of an EDL, the purchase of a product, or the association of events to the Video content.
The big difference in this case being that the Broadcast producer has access to the totality of the information (…or wished he did!) whereas the viewer only get a very small amount of information – I.e: from a TV program – and even worse – in our case - from a small compressed video segment streaming on the Web.

We're now at the core of this issue. The fact is, that in the current state of the technology, computer networks in general & the web in particular do not offer the producer and much less the viewer – the final client – the possibility to make an informed decision – and to thus transact on any content. The same comment is valid at every level of the Broadcast production, post production and distribution chains.

So remote content identification and transactions – in the generic sense of the term - constitute the core issues for Remote Television Production

This paper being entitled "New Internet and Broadband Video Technology for Remote Television Production" might make you wonder what a course on the "semantics of Broadcast production" might bring to the table.

The answer is simple: By clearly identifying the core issue at the heart of "Remote Broadcast TV production" which is the necessity to identify and transact on Video content at any stage of the work flow process - production, post production – conform editing, distribution or Viewing we can now address the corresponding technology paradigms and start looking for real answers.

Resolution of the technology paradox:

On one hand, the TV industry broadcast digital formats are not really manageable – that is "indexable" and "tranportable" on global computer networks – and offer no option of interactivity – except on broadcast grade machines (Avid) & installations.

On the other hand computer networks, including the web – do offer interactivity but are unable to accommodate the transport of broadcast-grade digital formats.

This paradox can be further developed by considering that even if broadcast DTV were to meet computer network standards, one day, two fundamental problems would remain:

1)DTV files would still have to be associated with Interactive Media-rich content to allow multiple sharing collaboration and transaction.

2) They would have to fit point to point protocol in order to allow full one on one collaboration between individuals at every stage of the work flow process

So how do we resolve the paradox?

One solution is to cross our fingers and hope that bandwidth and transport technologies can one-day resolve the problem. The other is to look at the current state of the technology and search for a combination of elements that would in fact allow this process to take place.

As stated earlier, the single most important issue at the heart of "Remote Broadcast TV production" – the necessity to identify and transact on Video content at any stage of the work flow process.

But does this in fact mean that such Video content would have to be viewed every time, by every single individual in its original broadcast grade format (I.e: DVB) – in order to transact on it?

The answer is: Not necessarily – for as long as a direct synchronicity can be established between the different file formats used for remote content identification and the original broadcast grade formats generated by the DTV industry.

If complete – frame-accurate synchronicity between both realms can be achieved, then this means that any transaction or decision executed on a proxy or "mirrored" network format can be implemented directly on the source format.

Media Synchronicity: The key to remote Television production

Synchronicity: How is it achieved and implemented in the context of remote TV production?

I was visiting with a large database company some time ago to discuss the use of their video servers in the context of a video network collaboration project. That company –as it was their policy – had sent the SDK of their Video Server software to our lab for testing – but our engineers soon called me back: "We have a problem" they said: " we cannot access the Video time-code on those Video servers." In a subsequent meeting I relayed the information to a prominent sales executive in that company who seemed surprised "Why do you need to access a timeline?" he wondered "how would you use that in your environment?". "The reason is simple" I answered: "100% of the broadcast industry uses time code". ...That particular executive – otherwise a brilliant fellow – seemed puzzled. The fact was and still is that the computer industry measures Video time in frame numbers as opposed to Smpte time code. And it had apparently not occurred to this company's R&D staff that Time code could be utilized to synchronize and otherwise associate selected "events" in a Video sequence – with interactive media content.

The fact is that if we can achieve the frame-accurate synchronization of lightweight Video proxy file formats – compatible with Web bandwidth - with Broadcast Grade Video content, then everything becomes possible:

One can then annotate, edit, transact collaborate or otherwise interact with Broadcast content on any network including the web. Simply because the very association of conventional computer software with lightweight Video proxies – themselves synchronized by time-code to the original Video broadcast formats generates end-to-end synchronicity throughout the workflow process. Ranging from remote content identification ("viewing the dailies") to the automated conform editing of broadcast dailies – the entire production can be addressed through this architecture.

So again, the principle is simple: Rather to force broadcast formats onto narrow bandwidth networks – through compression for instance – and rather than trying to "embed" multiple data types in broadcast transmission protocols, which is feasible but does not allow interactivity - an affordable solution is to simply synchronize both realms via at third connecting paradigm.

Buddhists would perhaps call this "the way of the middle": I call it the "obvious" way.

Again professional broadcasters might argue that this would constitute a "cope-out" answer to real broadcast issues – which among other things – require a certain quality of remote viewing .

But that argument does not stand in any configuration because in this model – the video proxy used for network collaboration – only needs to have sufficient definition to allow the users to ID the basic content. Besides, this model does not imply the use of any particular proxy or bandwidth. In fact it would authorize the use of several proxies of the same content – ranging from high-quality Mpeg-1 – good enough for a DP to view dailies - to Java Video players – all linked together through the same Smpte Time code:

The use of a particular proxy could be determined through the use of a "band width" analysis system – which will determine the actual bandwidth used by the remote viewer and thus stream the corresponding format. For instance if someone wants to access video content on a LAN or ATM network, the bandwidth analyzer can detect a 10 Mbit/S bandwidth and stream an Mpeg1 format. If the same user – for instance a TV show producer – is travelling away from the set – and using a laptop in a hotel room, the same bandwidth analyzer can stream a Java Video – roughly "weighting" 100K/minute. But in both cases, it is the same Video content that is being streamed – and along with it any associated file format.

Again you might ask: How can you apply Smpte Time Code to Mpeg – well that problem was resolved a long time ago and one can now access Mpeg content at a frame accurate level – and thus link it both to network Media rich contents and – say – a Digital Beta Broadcast tape via the use of EDLs.

The list of possibilities is endless. But before closing let's address three basic issues:

Remote broadcast content Acquisition, Editing, and Distribution.

Acquisition:

This is the premise: We have a DTV broadcast Video content that needs to be used for remote content Identification and transaction – so it then needs to be converted to a variety of format – and furthermore remain linked to those various network format through time-code. How is that done? And how is it done if we're talking about a finished edited program – such as a show or a commercial – where we want to keep the totality of sequence, but be able for instance to view them in any order we choose?

One way to do this is to automatically segment various portions of that content – ie: Scene cuts, or particular sequences – into a set of "Video blocks" –, which display, for instance, as storyboard frames. But in this case those blocks are automatically created by the initial import of the program's EDL – or editing decision list – which will be used by the acquisition software to "read" the Video and extract the corresponding blocks. (Nothing by the way – prevents us to create Video blocks manually – down to a single frame if necessary for editing or annotation purposes).

At the same time, the Video itself is converted to an initial proxy format – such as Mpeg 1. So we now have the proxy video – itself associated with a set of extracted blocks – or frames - which are in fact "pointers" to the Video content, and can trigger its playback. At that point the Time code can be reset to "0" at the beginning of the file – for as long as a translation table is recorded equating any proxy time code number with the Original broadcast time code.

Step two consist of associating every segment of the video – down to a single frame if necessary – with a variety of files – these can range from simple text annotations to HTML pages, Database tables – Watermarking - additional graphics or sound files – and even electronic transaction pages.

One this proxy is in place it can then be saved as a file for distribution on the server. But in effect the file itself does not need to contain the Video proxy. It is enough for this file to contain pointers to the location of the Video – for instance on a computer network Video server – and can furthermore point to a variety of different formats of the same video – which can be duplicated in various sizes. I.e: A real Media format, a Netshow format , and a Java Video format).

Once the content is acquired, annotated, and saved on a server it is now ready for editing and distribution.

Editing

Let me clarify a terminology point here: In the broadcast realm – "editing" means the actual "cutting" of a tape in various segments – or scenes belonging to various sources – that are then rearranged and re-ordered to create a program. In the computer realm – "editing" means in fact "publishing" a content. What we are talking about here is editing in the broadcast sense of the term.

How can we then achieve remote Network editing of a Video content – which in fact is not in the original format?

After the acquisition of a Video content, and its saving on a network server – in association with a file containing frame-accurate pointers to any designated portion of the Video file – it is then possible to download or transmit this "pointer-container" to any remote network station in order to re-access playback of any portion of the video.

Now , remember that this network user is looking both at a set of extracted frames – or "video blocks" – each connected to a portion of the video playback through time code.
Let us now imagine that the blocks can then be rearranged in any order to reorganize shots and/or sequences in the video. And let us further assume that this new arrangement of the blocks can be translated into an EDL (Editing Decision List) and then re-exported back to the server. What we have done here is in fact created a new EDL – that can easily be translated back to the original EDL of the Broadcast tape. It is then simple to deduct that this new EDL could be use to conform-edit the original tape in a conventional manner. Of course, the problem remains to implement such complex editing procedures such as special effects additions, and or complex sound editing. But remember that the initial purpose of this architecture is not to replace conventional editing, but to allow remote intervention &/or collaboration on broadcast content by the various parties associated to the workflow.

The beauty of this approach is that – at no time – the original master is touched. Theoretically you could have an infinite number of users creating an infinite number of annotations or EDLs on the same source – for collaboration purposes, without affecting the source tape. That should make directors and producers quite happy- shouldn't it?

Distribution

The key problem of future broadcast distribution will be:

1) the ability for broadcasters to distribute broadcast quality content using point to point protocols –as opposed to multi point protocols. In other words – it is the difference between being able to stream content on demand directly to a single user as opposed to a group of users who all get the broadcast simultaneously.

But furthermore – the real future broadcast challenge is to permit interactivity on the user/viewer end – in order to enable:

- The remote query of broadcast contents
- The remote transaction on such contents (such as ordering the playback of a film or buying goods associated with a commercial content – or even ordering a cruise to the Bahamas after viewing a documentary on the subject)
- The subscription to a variety of similar content through pre-organized distribution channels
- (I.e: Sport, News, Documentaries)

The architecture described here does not address the point to point protocols – which is a different subject all together – but to enhance the service offers at the user interactivity level. Here again the "obvious way" is the simplest possible solution to this problem:

By simply using a 56k "return channel" on any set top box – we can allow the user to interact not with the broadcast source itself – but with all associated data mirrored from that source onto network servers. By virtue of the fact that the broadcast program remains permanently associated and connected to the "pointer containers" – connected to interactive content – via the web for instance - the viewer/user can now interact with the broadcast content in real time.

Conclusion

This proposed architecture is designed to link broadcast video content to any type of existing media on a frame-accurate basis (text, sound, HTML, still images, ODBC database tables, referred to by IP addresses)- not to forget close captions.

It is also designed to allow complete remote control of this video content and its seamless integration in conventional media network management systems.

This system further enables the user to quickly identify key video content using conventional office management software and collaborate on that media content with a global networked production team.

Last but not least this architecture also enables the complete, automated integration of video with all types of media for archiving and network distribution. It is further enhanced with plug-ins for e-commerce, legacy database connectivity and multi-channel streaming of video content and related metadata.

Also included in this design is an extensive set of interactive client tools enabling remote video content identification & editing, video-on-demand, and on-line transactions.

For example, if a viewer/client receives an "pointer-container" made from a commercial advertising a product, the user can view associated graphics while the video is playing, read through product spec sheets, even complete related electronic web-transaction forms and payments. All of which can take place within the Application's window on their PC, laptop or TV screens.

For professional broadcasters, this means unprecedented versatility for creating rich multimedia content for broadcast programming. The resulting integrated media adds a whole new dimension to traditional broadcast content distribution.

Producers, broadcasters, and audiences alike benefit from multimedia content that may be easily associated with those "synchronized media" files. Businesses can interact with real-time financial market data/analysis embedded in financial news reports. Home users are able to browse sports statistics associated with the event they are watching. Advertisers can use "synchronized media files" to preview "dailies" during production of a new commercial, link their comments and other multimedia information to specific key frames, then email the lightweight media to colleagues.

While no current solution is perfect to resolve the painful coexistence of broadcast DTV and computer network Video formats – we advocate "the way of the middle" the "obvious way" as we call it - as a practical an effective technique for the remote management, editing & distribution of Broadcast contents

... And do you have any questions?...

Etienne Mirlesse
Phoenix, AZ, February 12 1999

Live Transmission System for The "Tokyo Marathon" Using 800MHz COFDM Links from Moving Vehicles

Jiro HIRONO Toshiyuki KAWASHIMA
Fuji Television Network, Inc.
Tokyo, Japan

ABSTRACT

By introducing 800MHz COFDM (Coded Orthogonal Frequency Division Multiplexing) FPUs (Field Pick-up Units), the overall transmission system for the "Tokyo Marathon" has been significantly simplified. The number of necessary relay transmission sites has been reduced by about 70% and a link helicopter is no longer required. This is because the COFDM modulation scheme has excellent characteristics in a multi-path fading channel such as mobile links even within a large metropolitan environment like Tokyo and the transmission links with it are not disrupted when passing under bridges and approaching obstacles (buildings, trees, etc.). This paper describes practical operation examples of a live-relay transmission by 800MHz COFDM FPUs.

1.0 INTRODUCTION

Broadcasters must simultaneously reduce production costs of live sports events and improve production ability in order to survive in the coming digital era. Fuji Television Network, Inc. which is actively working on methods for applying digital transmission links for broadcasting live sports events such as marathon races, has confirmed that COFDM FPUs might be one of the solutions for making an efficient live broadcasting system.

We broadcast the "Tokyo Half Marathon" and the "Tokyo International Marathon" as live-programs mainly consisting of pictures by relay transmissions from vehicles moving along the marathon courses. 800MHz COFDM FPUs were introduced this year making it possible to feed stable high quality video from moving vehicles and to simplify the overall transmission systems significantly by reducing the number of necessary relay-sites by about 70% compared with existing systems with analog FM (Frequency Modulation) FPUs.

In the case of mobile transmission links within a large city like Tokyo, interference of radio transmissions often occurred when passing under bridges and approaching obstacles (buildings, trees, etc.) due to multi-path fading. Therefore, many relay transmission sites, strategically placed throughout the course, and a helicopter linked with van-vehicles were required. Moreover, redundancy systems were usually prepared for foul weather days when the link-helicopters were unable to fly further increasing the number of terrestrial relay-sites necessary. The traditional systems for broadcasting both marathon races, thereby, became large scale and the operations were complicated.

The COFDM modulation scheme has excellent characteristics in the multi-path-fading environment and transmission links with COFDM FPU are not disrupted even when not having a line of sight path. By introducing 800MHz COFDM FPU for mobile links, many relay transmission sites consequently became unnecessary and in the case of the "Tokyo Half Marathon", no link-helicopter was used at all. Thereby not only were the production costs and frequencies saved but also complicated switching operations were greatly simplified. Moreover, transmission by omni directivity antenna became possible, which had been avoided in the case of analog due to quality deterioration by interference.

The following is a comparison of transmission systems for outside broadcasting with typical analog FPUs and the COFDM-FPUs for two separate marathons. The method for designing transmission links with COFDM and propagation characteristics in an urban region is also mentioned.

2.0 THE TOKYO HALF MARATHON

2.1 Planning criteria with analog FM FPUs

In the case of analog FM FPUs, the quality of video and audio directly depends on the received power and the existence of a reflected or diffracted ray. Considering such a property, we determined the allocation of terrestrial relay-sites to be set up by the following criteria.

(1). It is to be within the line of sight between the transmission antenna and the receiving antenna.
(2). The reception power is to be over –60 dBm.
(3). A directional antenna of curricular polarization (typically eight element cross Yagi) is recommended for both transmission and reception in order to suppress a reflection by the cross polarization discrimination.
(4). The influence of reflection and diffraction is to be minimized at the reception point.

2.2 The transmission system design with traditional analog FM FPUs

2.2.1 The overview of the course

The Tokyo half marathon was held in a seaside area, which was newly developed in front of Tokyo Bay. The course map is shown in Fig.1. Most of the course is laid along an overhead railway of a monorail called "Yurikamome" and there are many high buildings along the course as well. Helicopter flight is also very restricted because this area is located near the Haneda Airport making this location possibly one of the worst environments for broadcasting a live marathon.

2.2.1 The site allocation

It was estimated by transmission tests executed in advance that 12 relay transmission sites, as shown in Fig.1, were required if mobile links for the whole course were to be established by only terrestrial links with analog FM FPUs. For example, a relay-site had to be allocated by the bridge shown in Fig.2(a) in order to get a line of sight.

2.2.2 The set-up of vehicles and sites and the operation

For this live marathon, two moving vehicles locally linked with a motorcycle camera crew by microwave (Fig.2 (c)) had to be prepared, which is one of the typical styles for live marathon broadcasting in our company. The sources gathered in the vehicles are switched and then transmitted to the terrestrial relay sites by mobile links above. Those sources received at each site are retransmitted to the main Switching Center by stationary microwave links (Fig.2(b)), where they are combined or switched with other sources (VCRs, landscape cameras etc.) and super-imposed and finally emitted to the air.

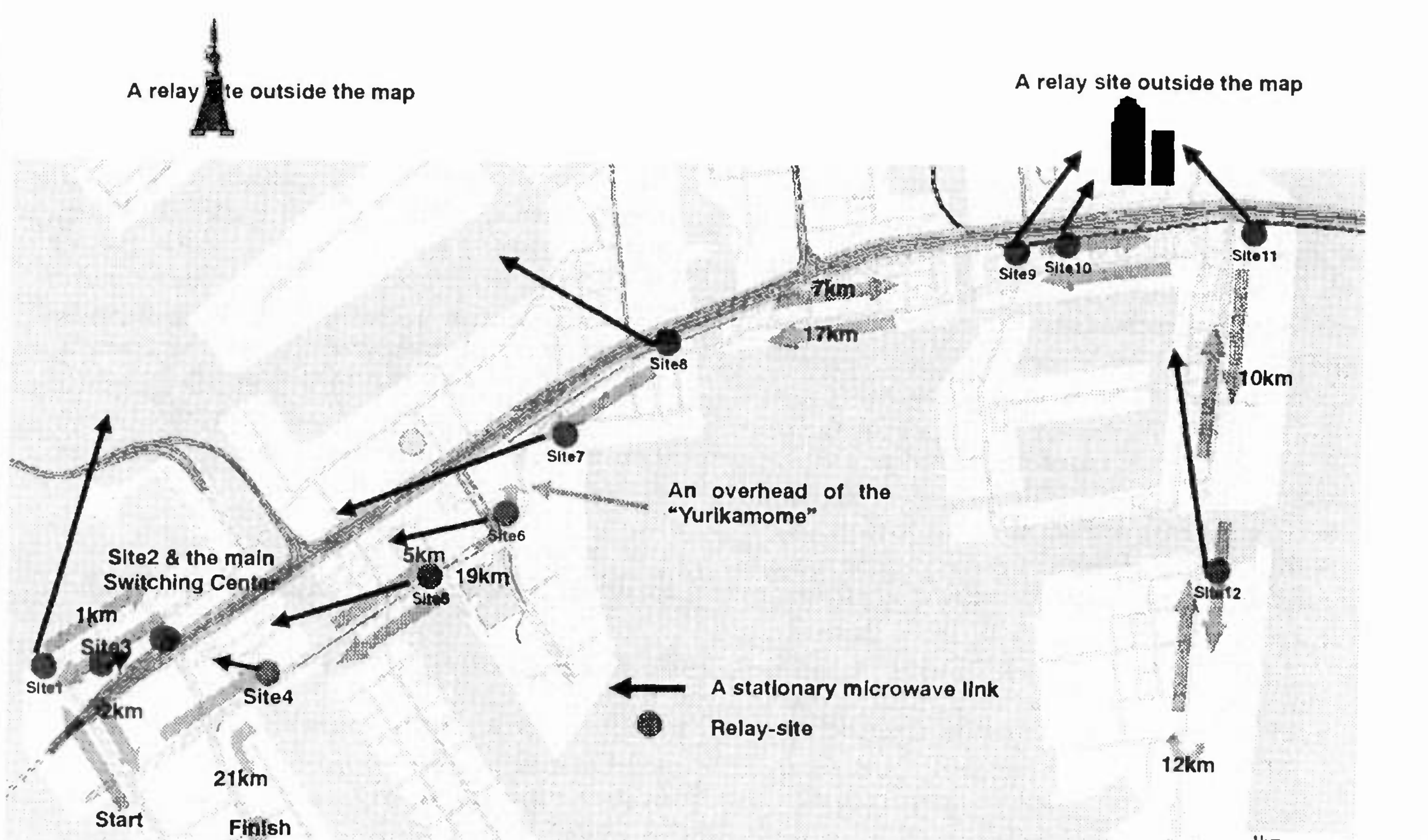
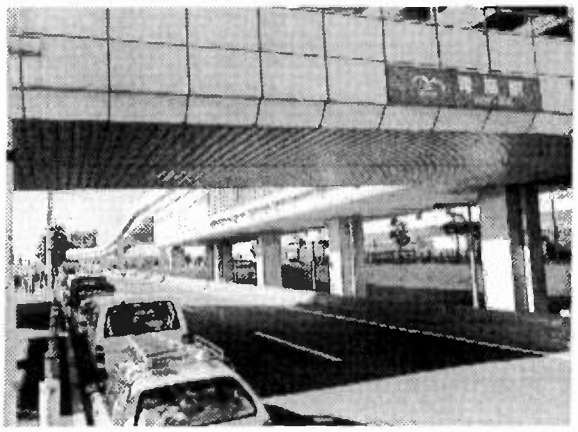
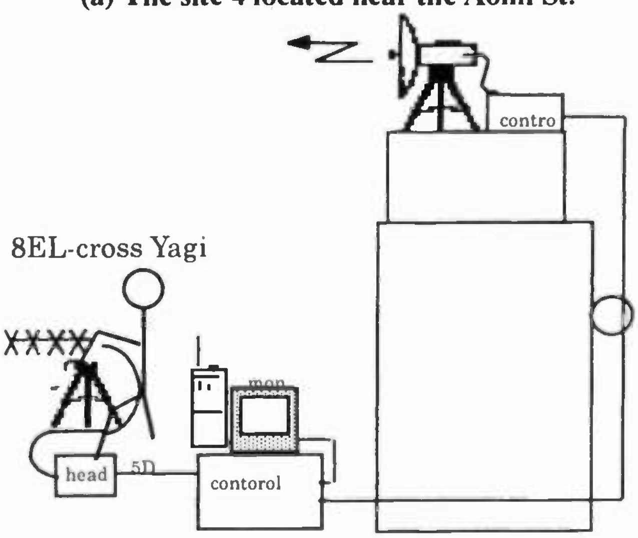


Fig 1. The overview of the Tokyo Half Marathon course and the sites allocation with analog

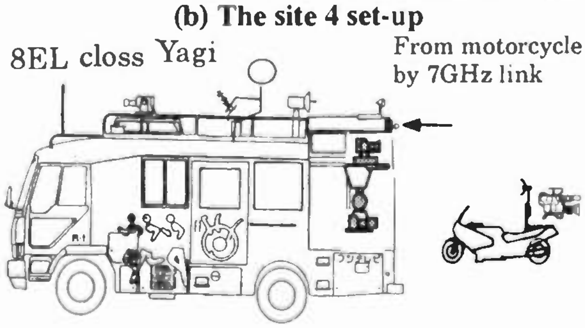
Because there are not enough frequencies to assign a separate channel to each link, some have to be shared. So as more terrestrial links become necessary, switching operations of the frequencies and sources become even more complicated resulting in the need for more equipment and personnel to run the equipment.



(a) The site 4 located near the Aomi St.



(b) The site 4 set-up



(c) The moving vehicle set-up

Fig 2. The moving vehicle and a typical site set-up for analog

2.4 The antenna type selection

An eight-elements-cross-Yagi is usually adopted for the antenna of both vehicles and sites in order to suppress a reflection by the cross polarization discrimination. And antenna experts aim their antennas toward each other making the proper adjustments for the vehicles moving along the course with the marathon runners. Because the quality of the transmission source directly depends on the skill of the antenna experts, sufficient transmission tests must be executed in advance.

Live marathons are very popular in Japan and more than ten such programs are produced and broadcast each year. Therefore it is necessary to discover ways to control costs in production as soon as possible.

2.3 Transmission system design with COFDM FPUs

2.3.1 The site allocation

Many test transmissions with COFDM FPUs were conducted to find a way for reducing the number of necessary terrestrial sites. Appropriate antenna types and their specific placements were also examined. Consequently it was found that only three terrestrial relay-sites, as shown in Fig.3, enabled stable pictures to be transmitted from a moving vehicle to the Switching Center for the entire marathon course. If it had been permitted this year, landscape or VTR pictures could have been substituted during the time when images were not able to be sent from the vehicles traveling in the small NG areas (see Fig. 3) thereby reducing the relay sites from three to just one site (d1).

The location coverage for a mobile link of each site allocation is indicated in Table.1.

Because the site d1, which conveniently includes a Switching Center, is located 400 feet above at the top of a building, the other two sites may not be set up for next year's race in order to save on production costs if the quality of the program is judged to be not significantly lowered.

It goes without saying that the transmission quality was improved in comparison of analog.

Table 1: Location coverage versus site allocation (Tokyo Half)

	the location coverage
Site d1	96%(20.2km/21.1km)
Site d1, d2 and d3	quasi 100%

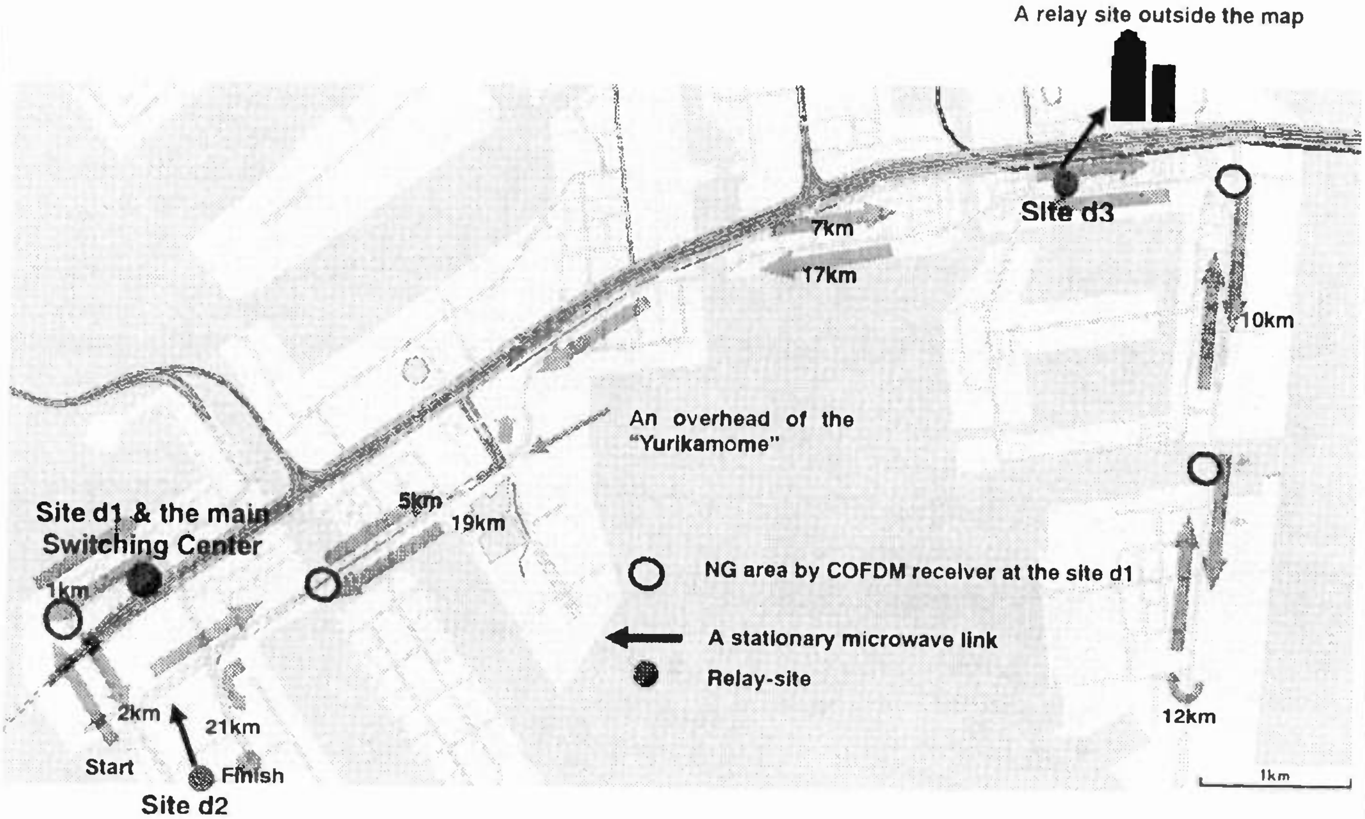


Fig 3 the overview of the Tokyo Half Marathon course and the sites allocation with COFDM

In comparison, this year the total production costs for the COFDM plan is approximately 30 % lower than that of the analog one used and production of next year's race could be even less expensive.

2.3.2 The set-up of vehicles and sites and the antenna selection

Concerning the set-up of both the moving vehicles and the sites, it was found by test transmissions that the use of a non-directional antenna such as a dipole or a ground-plain became possible and training and deployment of antenna-men was no longer necessary.

Table 2: The antenna selected (Tokyo Half)

	antenna type
transmission from vehicles	Type: A omni directional grand-Plain Gain: 8dBi Polarization: Vertical
reception at the site	Type: 26 EL Yagi Gain: 20dBi Polarization: Vertical

Although it is true with respect to both transmission and reception, this year grand-plain antennas were applied only on the transmission side and Yagi were on the reception side. A lot of comparison tests of the antenna type (directivity and polarization) were conducted at the site d1 and the type shown in Table 2 was selected as the best combination for covering this specific course.

The reasons are described below:

(1) The vertical polarization gets best coverage.

The reason is that reflections by buildings are main component of the total reception power in the case of an out of sight transmission and that the circular polarization, which is popular with analog, is not appropriate for out of sight transmissions in which case the cross polarization discrimination does not act properly.

(2) The combination of omni direction transmission and high gain reception gets best coverage and reception is most stable.

The case that both are omni directional antennas is best for operations, but it is unfavorable for a long distance transmission and consequently is not cost-efficient for this marathon race. On the other hand, the case that both are high gain antennas is best for a stationary long distance transmission, however in the case of mobile links out of sight the reception power doesn't behave stably, because the antenna-men cannot find the best direction.

(3) The high gain antenna is easy to operate in the case of COFDM links.

The antenna direction is less sensitive to the reception quality than that of an analog system and it is easy to follow a moving vehicle even if the target cannot be seen from behind a building, etc. This is because the transmission quality is independent of the reception power even with the existence of reflected or diffracted rays, as long as it is more than the threshold power. So even a high gain antenna like 26EL-Yagi can be easily used.

2.4 The COFDM equipment and the transmission characteristics

The COFDM FPU introduced this year employs the specification in Table 3, which is compliant to the ARIB standard STD-B13 (VER 1.0) titled “800MHz-BAND OFDM Transmission System for Television Program Contribution”. Because this was the first practical application of COFDM mobile links, the strongest mode (convolution code 1/2)

Table 3: COFDM FPU specification

parameter	specification				
modulation	DQPSK				
center frequency	774.5, 783.5, 792.5 801.5MHz				
occupied BW	8.5MHz				
number of carrier	544				
FFT size	1024				
carrier spacing	15.625kHz				
effective symbol length	64μs				
guard interval length	3μs				
Frame period	60.3ms, 900symbol including 6 sync. symbols				
outer error correction	RS(204.188)				
inner error correction	1/2	2/3	3/4	7/8	1
bitrate after RS Mbps	8.0	10.66	12.0	14.0	16.0
transmission power	5[W]				
encoder	MPEG 2 MP@ML				

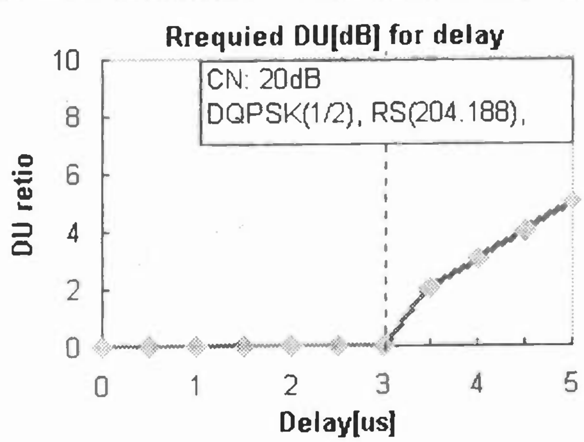


Fig.4 the required DU ratio for delay

was selected, considering enough allowance.

The graph in Fig.4 indicates the required DU ratio for quasi error free, which indicates COFDM characteristics in the best way. The delay of echoes within the guard-interval 3 μs gives no influence to the reception quality and the required DU ratio for a delay over it is gradually increased according to its delay time. Considering this property of COFDM scheme, how to design COFDM links are mentioned next.

2.5 The planning for COFDM links and the propagation characteristics in this area

2.5.1 the planning criteria for COFDM mobile links

In a sense the planning of analog links is easy, because it is used for a line of sight transmission and their propagation characteristics are easily predicted. And the prediction often proves right on the on-air day.

On the other hand COFDM links are mainly used for an out of sight transmission and it is difficult to predict their propagation. Successful transmissions may have not been consistent during the tests. Considering the characteristics described above, however, it was found that a satisfactory planning of COFDM links was possible by checking the following points.

The planning criteria for COFDM mobile links are:

(1) The reception power is over -85dBm, which means CN ratio approximately 20dB.

(2) The delay of echo is within the guard-interval or the DU ratio of outside it is enough (5-10dB).

2.5.2 The measurement results

To check the points above, the measurement system shown in Fig.5 was prepared. The delay profile is obtained by the correlation of the charp symbol which is inserted into every OFDM frame as a sync symbol. And the thermal power meter installed in the receiver roughly measures the reception power and the spectrum analyzer checks the CN ratio after AGC.

Fig.6 gives the measurement results of the relay site d1. Fig.6(a) and Fig.6(b) the distributions of the reception power and the strong echoes (over DU ratio 5) respectively.

The areas where the COFDM links were disrupted are marked in Fig.3 and Fig.6. It is found by the above results that the existence of a strong echo outside the guard-interval caused the disruption of transmission in the almost cases. Although location probability of outside guard-interval is small as shown in Fig 6(b), some countermeasures overcoming it is necessary in the case of contribution, because it often gives disruption of transmission.

Some methods described below might be effective to overcome delays outside the guard-interval length and to magnify the coverage of the site d1 to the whole course.

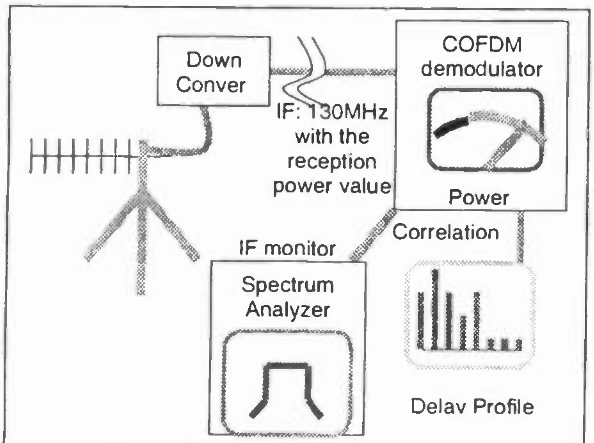


Fig.5 The measurement system

(1) To adopt longer guard-interval according to measured delay spread at the cost of the room for the transmission source.

(2) To predict the disruption area and alter the direction of the reception antenna directivity to weaken the interfering delays.

(3) To predict the disruption area and altar the directivity to weaken the interfering delays.

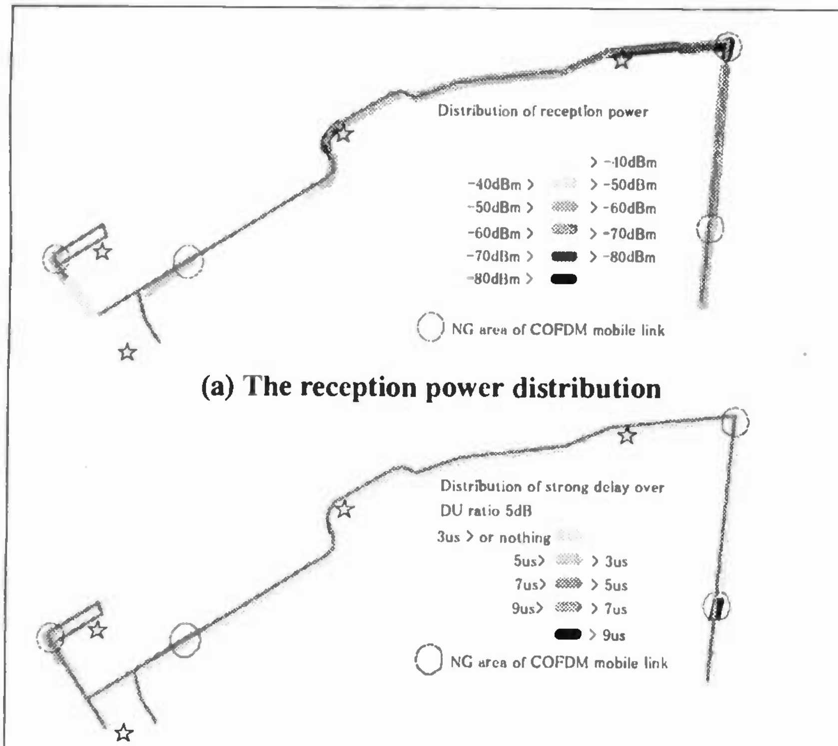
Regarding the solutions mentioned above, test transmissions are planed and the results will be reflected to the plan for next year.

3.0 THE TOKYO INTERNATIONAL MARATHON

3.1 The comparison the transmission system design with analog FPUs of that with COFDM FPUs

3.1.1 The overview of the course

The Tokyo international marathon course is laid from the center of Tokyo to the south. High buildings surround almost all the course and some part indicated in Fig.7 is along the overhead of highway. And a flight of helicopters is sometimes restricted in the south part, because it is near the Haneda airport.



(a) The reception power distribution

(b) The distribution of strong delays (DU ratio over 5dB)

Fig.6 The measurement results at the site d1

3.1.2 The overall transmission system design with analog

Mobile links with analog FPUs were set up mainly by link-helicopters but some part had to be covered with terrestrial relay site. And some additional terrestrial sites were prepared as back up for foul weather days when the link-helicopters were unable to fly. Consequently 17 terrestrial relay sites were set up throughout the course as shown in Fig.7.

3.1.3 The overall transmission system design with COFDM

On the other hand, by applying COFDM FPUs for the mobile links, the necessary relay sites of this year's race were reduced to only five shown in Fig.7. Three of those were reception sites with COFDM, which covered over 90% of this course and the rest were analog ones which was allocated in order to patch the rest of this course. And the COFDM link from the moving vehicle to the helicopter was very much

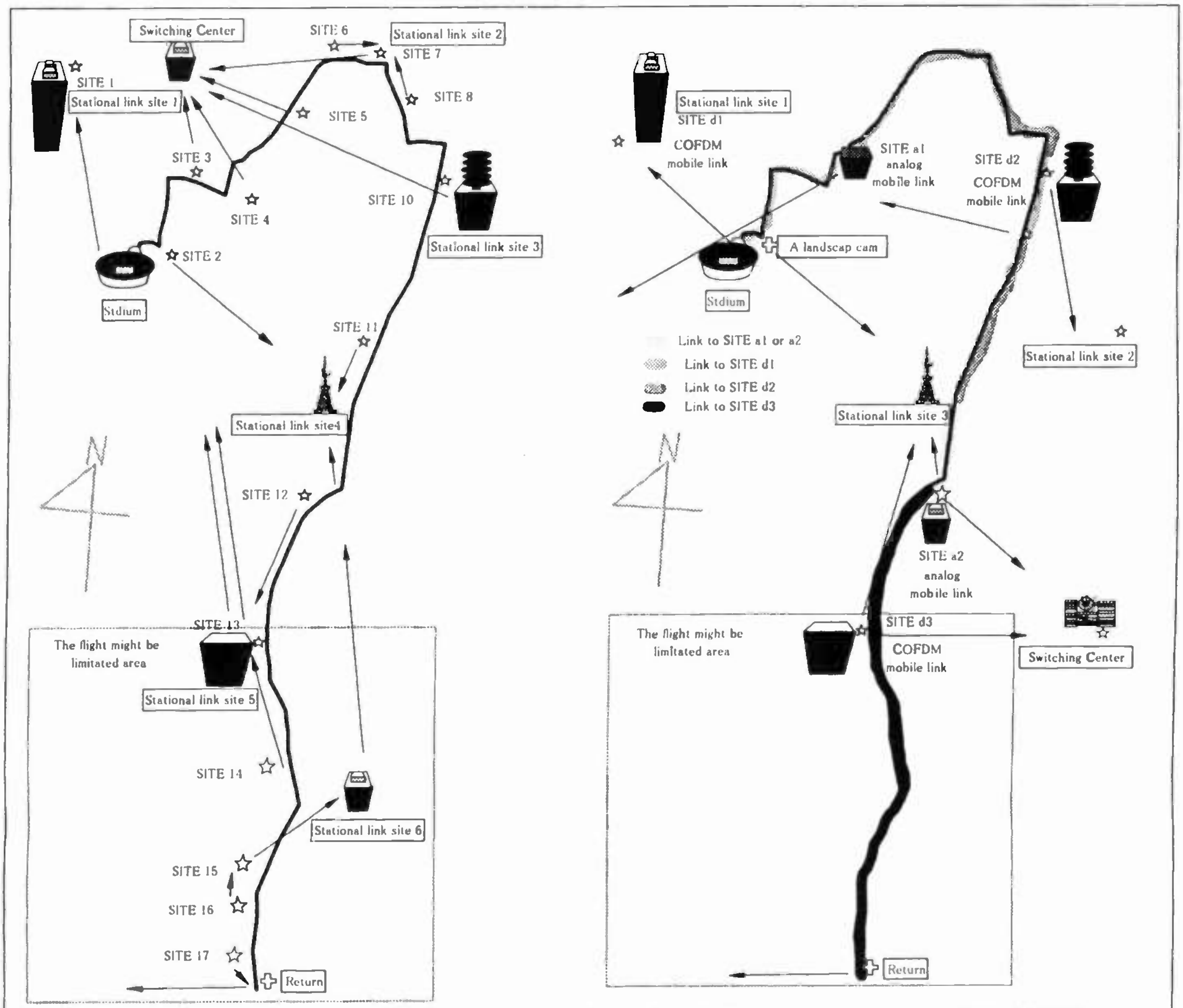


Fig.7 Comparison the overall link plan with analog FPUs of that with COFDM FPUs

effective as well. Because the transmission quality did not so much depend on the position of the helicopter, it could fly relatively freely, which enabled one helicopter to link more than two moving vehicles.

Talble.4 Comparison of the site number

	The number of sites	
	The existing system with analog	This year's with COFDM
For the case link-helicopter is unavailable	17	5 (3 of COFDM, 2 of analog)
For the case link-helicopter is available	9 (5 are for the case of the flight limitation in south area)	1 (all are for the case of the flight limitation in south area)

Table 4 is comparing the site number of sites between this year's plan with COFDM and the existing plan with analog. It was found that COFDM mobile links are efficient for the overall transmission systems to be remarkably simplified.

4.0 CONCLUSIONS

By success of Fuji TV's practical use in the two marathon races, it was confirmed that COFDM FPUs might be one of the solutions to make an efficient live broadcasting system and be also applied in another application such as ENG. More effective use can be expected in the future by overcoming the subjects previously described.

Corresponding to the digital satellite HDTV broadcast that will start in the year 2000; the establishment of mobile links of HDTV is expected early. In order to realize this, it is necessary to develop the standard employed such as 16QAM or 64QAM modulation and that of microwave band where the assigned bandwidth is broad and is favorable for high bit rate transmission such as HDTV contribution.

MORE INFORMATION

We are welcome to visit our booth, North Halls N3 & N4 area, L10334, Las Vegas Convention Center, where the COFDM FPUs are demonstrated, and if you need more information, please contact via E-mail.

E-mail address; jiro.hirono@fujitv.co.jp

WCBS-DT DTV FIELD TEST REPORT

Prepared by:__________________
Walter Sidas, P.E.

Approved by: __________________
Robert P. Seidel

March 22, 1999

TABLE OF CONTENTS

WCBS-DT DTV FIELD TEST

1. INTRODUCTION

CBS Corp. owns 14 full power and three satellite broadcast stations throughout the United States. Its flagship station, WCBS-TV Channel 2 in New York City, operates from a transmitter and antenna located on the North Tower of the World Trade Center, in downtown New York City. WCBS-TV also has an auxiliary transmitter and antenna at the Empire State Building, located at the corner of Fifth Avenue and 34th street in Midtown Manhattan, 2.9 miles (4.7 km) at a bearing of 29.8 degrees from the World Trade Center. **Figure 1 is a map showing the location of the Empire State Building and the World Trade Center.**

CBS is participating in a joint regional study, conducted by the Television All Industry Committee, to find a common location for the DTV antennas for all the television stations in the New York Metropolitan area. In the interim, WCBS-TV decided to replace its five bay VHF auxiliary antenna (Channel 2) with a two bay auxiliary antenna, and use the aperture of the top three bays for a DTV antenna. On November 26, 1996, CBS was granted an experimental license to broadcast DTV on Channel 33, from the Empire State Building site, and commenced operation on April 4, 1997. When the Federal Communications Commission released the DTV allotment Table in 1997, CBS was assigned Channel 56. On December 12, 1997 CBS was granted a Construction Permit for WCBS-DT. WCBS-DT went on air as New York City's first and only DTV station on November 1, 1998, at which time CBS discontinued its experimental operation on Channel 33.

On July 2, 1997 the Federal Communications Commission issued OET Bulletin No. 69 entitled "Longley-Rice Methodology for Evaluating TV Coverage and Interference".
The Longley-Rice radio propagation model is used to make predictions of radio field strength at specific geographic points based on the elevation profile of terrain between the transmitter and each specific reception point. A computer program of the Longley-Rice model, developed by Techware, uses the coordinate location and height of the center of radiation of both the transmitter and the receiver to determine a path profile, and identifies obstructions in the transmission path. A digital terrain database is used to calculate elevations.

The effective radiated power and the antenna pattern of the transmitter determines the radiated power in the direction of the receive site. For digital television stations, service is evaluated inside contours determined by the FCC planning factors in combination with field strength curves derived for 50% of locations and 90% of the time. Within this area, the computer program uses the terrain dependent prediction model to calculate the value of the field strength. Interference data from adjacent NTSC and DTV stations may also be included.

2. GOAL OF THE TESTING PROGRAM

The goal of the field test program was to verify the quality of reception of an ATSC DTV signal transmitted from a major metropolitan center, and compare actual measurements versus predicted coverage using the Longley-Rice calculation. Measurements to be performed included: range of reception, effect of ghosting due to reflections from tall buildings, in particular the World Trade Center twin towers, and reception in densely populated metropolitan areas with tall buildings. Another objective of the test program was to evaluate the interference caused by a short-spaced adjacent NTSC Channel 55.

3. WCBS-DT FIELD TESTING OVERVIEW

CBS used the testing methodology procedures of the Advisory Committee on Advanced Television Service (ACATS), and the previous experience of field tests at WRAL-TV and WGN-TV. The methodology developed during these tests, as well as the instrumentation guidelines were used to develop the DTV test program for WCBS-DT. However, the unique circumstances in New York, consisting of many tall buildings, a large urban sprawl area and wide variety of terrain, from relatively flat areas to mountainous suburbs, provided an opportunity to evaluate the data on a few levels. Areas that were at relatively large distances from the transmitter utilized the Longley-Rice methodology to predict DTV reception. Areas in New York City close to the transmitter are subject to blocking by tall buildings. These sites were evaluated using visual observation of the antenna paths to the transmitter to evaluate blocking, rather than terrain database used with the Longley-Rice analysis.

3.1 TRANSMITTER SITE

The transmitter is a two-tube IOT Harris Sigma CD 190 DTV Transmitter installed on the 83rd floor of the Empire State Building. The transmitter operates at a TPO of 23.8 kW. The antenna system is a Harris broadband multi panel array antenna, side mounted on the northwest corner of the mast on top of the building. The maximum effective radiated power (AVG) is 349 kW (25.4 dBK), height of radiation center above ground is 395 meters (1296 feet), and height of radiation above average terrain is 397 meters (1303 feet). **Figure 2 is a simplified block diagram of the transmitter installation.**

The Empire State Building has a tall tower on which are mounted numerous VHF, UHF, FM radio and microwave antennas. **Figure 3 is a photograph of Empire State Building mast with the WCBS-DT panel antenna**. Since some distortion of the antenna's azimuth pattern may be expected due to the side mounting of the antenna system on the mast structure, the antenna was mounted on the northwest corner of the mast. Thus the blocking effect of the far field, due to the mast structure, is localized to an area southeast of the building, generally over water. The data collected by WRAL-HD in Raleigh, NC suggests that side mounting of DTV antennas is not as critical as would be expected. The Harris broadband antenna system is capable of transmitting DTV Channels 14 through 69. **Figure 4 shows the calculated elevation and range-tested azimuth patterns for Channels 28, 45, 56 and 61.** These are channels that have been allotted by the Commission for DTV service in New York City. The azimuth pattern of the antenna system was positioned such that true north is –28° from the panels face.

3.2 TEST VEHICLE

CBS adopted the Advanced Television Systems Committee (ATSC) Field Test Vehicle Design Information authored by Gary Sgrignoli of Zenith Electronics (Reference 5) as the basis of the test vehicle construction. In 1997 CBS prepared and issued a specification for the design of the test vehicle (Reference 6). The 13 foot 3 inch high truck is a 1998 Ford E350 van with an 18-foot long body. It is equipped with a 10 kW AC generator and a compressed air extendable mast capable of raising the antenna to the 30-foot AGL level. An electronic rotor points a Wade Model WL30-83 antenna used to receive the Channel 56 DTV and Channel 41 NTSC signal used for comparison of reception. **Figure 5 contains the electrical specifications and azimuth pattern of the Wade antenna used in the DTV test truck**. Its output is transformed from 75 ohms to 50 ohms to match the 50-ohm coaxial downlead cable.

Figure 6 is a simplified block diagram of the test vehicle electronics. An RF preamplifier system built by Zenith (referred to as the "works-in-a-drawer") and described in Reference 5 is used to: (1) amplify the RF signals, (2) level control via an attenuator, and (3) add white noise from an external generator.

The minimum field strength for DTV threshold operation of the test vehicle, at Channel 56, is 40.9 dBuV/m. Using the FCC's planning factors, the minimum field strength required to receive the Channel 56 DTV signal is 42.4. dBuV/m. This means that the DTV test truck is 1.5 dB more sensitive than the FCC planning factor criteria for reception, which was taken into account in the margin calculations.

The test equipment monitoring the 8VSB signal is an HP89441A Vector Signal Analyzer (VSA), an HP8560E analog spectrum analyzer, and a Grand Alliance (GA)/Zenith "blue rack" 8VSB demodulator. The HDTV monitor is a 17" Ikegami 2003D high definition monitor capable of scanning at the 1080I rate, and audio is monitored by a stereo amplifier/speaker system. A Tektronix DS1000 television demodulator with a Panasonic S901 color video monitor is available for monitoring the received NTSC signals. A Tektronix 1740A vectorsope and waveform monitor provides means to measure the analog video.

A Pentium-based computer runs DeLorme Street Atlas map software and custom-designed EXCEL spreadsheet software developed by Zenith for recording all the measured data. The Techware P-Pro_1 software is used to calculate the terrain profiles, the Longley-Rice parameters, and the FCC 50,90 Propagation Model field strength values. The map software is used to accurately calculate the distance and bearing from the transmitter to each test site. A Garmin GPS II Plus Personal Navigator™, was programmed with the coordinates of the Empire State Building transmitter site, and showed the distance and bearing from a test site to the transmitter. An HP Deskjet 670C color printer is used for printing from the VSA, the spreadsheet, mapping software, and terrain profiling software.

3.3 FIELD TEST PLAN

The field test plan used during the WCBS DTV field test was essentially identical to that used in Charlotte, NC for the ACATS testing in 1994 and 1995 (Reference 1), in Raleigh NC at WRAL-HD (Reference 3) and WGN, Chicago, IL (Reference 4), KICU, San Jose, CA, and KING and KOMO in Seattle, WA.

A detailed transmitter calibration was performed at the transmitter site in the Empire State Building at the start of the measurement program to establish a performance benchmark. Average DTV power in the 6 MHz channel, inband spectrum tilt and Signal- to-Noise Ratio (S/N) were measured. Automatic monitoring equipment was used to constantly monitor transmitter performance throughout the test. Testing was discontinued if transmitter parameters were at variance with the initial transmitter calibration.

A field truck receiver calibration was performed before the start of each test day. The van was parked at the same location, on West 56th Street, with a line of sight to the Empire State Building. The test consisted of using a test signal generator in the truck, routed to the uplead / downlead coaxial cable system, the RF preamplifier system, the VSA and other receivers. This allowed the truck's receiving system to be accurately calibrated for gain from the antenna output to the VSA/VSB receiver inputs. At the same time, the truck's noise floor was verified prior to each day's testing.

The test procedures at every test site, as called out in the generic field-test plan, called for the truck's antenna to be raised 30 feet above ground level (AGL). The antenna was aimed, using the electronic mast rotor, for maximum signal strength, and a compass used to verify that it was indeed pointed towards the transmitter (within +/- 20 degrees). In some cases, reception was peaked with the antenna pointed away from the transmitter, indicating that reception was via a reflected signal. At each location the following data was recorded in the spreadsheet: site name, site's latitude and longitude, distance and bearing from the Empire State Building. The data between the map program and the handheld GPS receivers was compared to verify position accuracy.

Using the attenuator in front of the RF preamplifier (Refer to Figure 6), the received signal's average power was adjusted to be about –30 dBm. This is done to prevent RF preamplifier overload at locations near the transmitter that had large field strength levels. The attenuator setting, the exact value of the 8-VSB average power in 6 MHz, noise floor, and S/N (as received) were recorded. The DTV field strength was calculated in the spreadsheet using the value of the truck's system gain, the log periodic antenna gain, attenuator setting, and DTV average power. White noise was added to determine the Signal-to-Noise ratio at threshold of errors. The VSB demodulator's equalizer parameters (S/N in, S/N out, tap energy) and segement (packet) error rate were also measured and recorded. If video and audio were present, they were also monitored and recorded.

The test plan also included a measurement of an NTSC UHF television signal, at each site. While an on-channel NTSC signal, at the same location as the DTV signal would have been ideal for comparison, in New York the closest UHF signal available is Channel 41, 90 MHz from the 725 MHZ center frequency of Channel 56. **Table 1, below, shows a comparison of station parameters for the four television stations considered in this field test**.

Table 1 Summary of Television Station Parameters

	WCBS-TV	**WXTV**	**WLNY**	**WCBS-DT**
	NTSC	NTSC	NTSC	DTV
Channel	2	41	55	56
Transmitter Location	WTC	ESB	Riverhead, NY	ESB
ERP (kW)	21	2,340	5,000	349
HAAT(meters)	482	421	194	397
HAAT(feet)	1,581	1,381	637	1,303

The NTSC peak sync power was set to –30 dBm, and a subjective CCIR 5-point impairment rating (1.0 very annoying, 5.0 imperceptible) was recorded. The delays of visible ghosts were observed on the NTSC monitor, measured with a ruler and recorded.

The test site locations were chosen to accomplish the goals of the test program: (1) measure range of coverage, (2) compare measurement to prediction using the Longley-Rice coverage model, (3) determine the effects of ghosting to reception, (4) determine the effect of inner city "canyon" effects, and (5) measure the effect of lower adjacent NTSC interference. The radial sites were used to measure the range of transmission. They were selected to approximate the field test measurement radials used in 1988, to measure the coverage of WCBS-TV, from the World Trade Center. **Figure 7 is a map from a report prepared by Jules Cohen & Associates, on November 23, 1988**. **Figure 8 is a map of the sites selected for the DTV Channel 56 measurements.**

Eight radials were chosen, at 003, 45, 80, 88, 188, 220, 268, and 312 degrees of true north. On each radial, sites were spaced every 5 miles starting at the 10-mile point and proceeding out to 55 miles. If reception was received at 55 miles, in some cases an additional special site was measured to find the "cliff effect".

Radial R080 was specifically chosen to measure the DTV reception in the presence of adjacent channel NTSC interference. WLNY-TV, Channel 55, licensed to Riverhead, NY on Long Island is located 56.9 miles (91.6 km) at a bearing of 79.2 degrees from the Empire State Building.

The arc sites were selected to measure ghosting to the Northeast of the Empire State Building resulting from reflections by the World Trade Center. The four arcs consisted of points spaced approximately 0.5 miles apart, at 20, 25, 30, and 35-mile distances from the Empire State Building. The main comparison at each site was the NTSC Channel 41 signal. While this was used to determine presence of ghosts, it was recognized that the 15 channel, 90 MHz separation resulted in differing ghosting characteristics.

Grids were measured to obtain data on the inner city coverage and ghosting due to the "concrete canyons".

Finally three grids were measured. One grid in Brooklyn, NY consists of 28 points at increments of 1 mile, and another grid in Manhattan, north of the Empire State Building, and a third grid in Long Island approximately 30-40 miles from the transmitter. Other special test sites were measured to verify reception in Queens, NY and other areas of special interest. **Table 2 contains a summary of the DTV field test sites.**

4.0. TEST RESULTS

The outdoor DTV field test conducted in October 1998 through February 1999 included 158 sites, covering a 55-mile radius.

4.1 DATA SUMMARY

Table 3 is a summary of the DTV field test results.

The first column in Table 3 contains the site type (i.e. radial, arc, grid or special). Adjacent to this, is the site name such as: R003 for the radial on a bearing of 003 degrees true North from the Empire State Building, or A1 for the first arc, or G2 for the second grid. The third column (A) details the total number of sites on the radial, arc, or grid. Column B is a number of sites, as predicted by the Longley-Rice methodology that will be able to receive DTV. Column C represents the number of measured sites that could actually receive DTV signals as predicted by the Longley-Rice methodology. Column D is a number of Longley-Rice sites that were predicted to have no DTV reception. Column E represents the number of measured sites that could not receive DTV signals as predicted by Longley-Rice due to terrain obstructions or interference. Column F, on the right, is a ratio (expressed in percent) of the number of predicted Longley-Rice sites divided by the actual measured Longley-Rice success rate (C/B*100=F).

Table 3 has been grouped into two sections. The first section contains the sites that utilized the Longley-Rice model to predict the reception of DTV or the lack of reception due to terrain obstacle interference. The second half of the table contains grids and special points that do not lend themselves to Longley-Rice analysis due to man-made obstructions that can not accurately be predicted in the Longley-Rice model.

A site was deemed to have successful DTV reception if, during a 20-minute visit, less than 2.5 segment errors/second were observed. Longley-Rice analysis was used to predict where the DTV signal could and could not be received. **Figure 9 shows the calculated Longley-Rice Field Strength vs. Accumulated Sites.**

The test data for the radials shows that the coverage of WCBS-DT extends out to a radius of approximately 55 miles unless limited by terrain obstructions or adjacent channel NTSC interference. Terrain blocking limited coverage to the North (R003) and to the Northwest (R312). Coverage on radial R080 was limited to 35 miles, due to terrain blocking. Many of the remaining radials had a 100% success rate.

Four arcs were measured to the Northeast at a range of azimuths 35 to 52 degrees and distances of 20 to 35 miles from the Empire State Building. Analysis in Reference 8 shows that the ratio of the total reflected signal from the both World Trade Center towers to the direct signal from the transmitting antenna located at the Empire State Building at Channel 56 is a nonlinear function. The ratio of the reflected signal to the direct signal provides reflection at numerous discrete bearings in degrees relative to true north. The range of bearings of these reflected signals is less than three degrees. The goal of the WCBS-DT field tests was not to confirm the actual calculated areas of reflection, but rather to determine that the equalizer correction in the receiver was able to overcome any ghosting that may result from the World Trade Center towers at the Channel 56 frequency. With the exception of one site, a success rate of 96 percent was achieved for this test condition.

The grid measurements showed that reception within the "concrete canyons" of a large city is variable. In general, sites that do not have a line of sight to the transmitter do not receive DTV unless an alternate path signal via a "bounce" can be established. It should be stressed that measurement of the received signal at a height of 30 feet AGL at these sites is not an actual test of the method of over the air reception. For example in G1, a grid in Brooklyn, measurements showed the effect of many tall multistory apartment buildings on reception. However, in reality an actual DTV viewer is able to receive an over the air signal from a rooftop antenna, which in many cases would be high enough to achieve a line of sight to the Empire State Building. This is also true for grid G3, in Manhattan. Measurements were also performed at sites termed Special, or S sites. These were sites that did not fit the other categories or to determine localized coverage as for example areas in Queens, to the east of the Empire State Building.

4.2 SITE MARGIN

The site margin that was measured at each site can be plotted versus the measured average DTV field strength (in 6 MHz), as shown in **Figure 10.** As expected, the site margin decreases with decreasing field strength. The predicted margins should decrease linearly with field strength. **Figure 11 is a plot of the Tap Energy vs. Field Strength, and Figure 12 is the Field Strength as a function of Sites.**

4.3 ANALYSIS OF SITES WITH NO DTV RECEPTION

Table 4 contains a summary of sites with no DTV reception.
Similar to the grouping in Table 3 the sites are grouped into two sections. The first section contains relatively distant sites that used the Longley-Rice and terrain database to predict coverage. The second section contains sites in New York City, close to the transmitter.

The first column in Table 4 contains the site type. The second and third columns show the distance and bearing of the site from the empire State Building. The fourth column shows the number of obstructions in the path from the Empire State building to the site.
Next are the measured field strength, the tap energy, and finally the Channel 41 NTSC CCIR rating. Some remarks associated with the site complete the table.

Table 4 shows that most of the sites that were 14.8 to 56.6 miles from the Empire State Building did not receive DTV due to terrain blocking. The sites close to the transmitter suffered from either blocking by multistory buildings, or severe multipath.

4.4 INDOOR TESTS

The scope of these field tests did not include indoor measurements. For a comprehensive test of indoor reception many variables need to be quantified, i.e.: type of building construction, location of the site within a building, and other factors that are difficult to determine. However, CBS Engineering has recently installed DTV receivers in executive offices using "double bow-tie" or rabbit ear antennas as sole means of reception. Many of these offices do not have line of sight to the Empire State Building.

4.5 COMPARISON OF WCBS-DT TEST DATA TO PREVIOUS TESTS

The WCBS-DT test data is consistent with test data obtained at WRAL-TV and WGN-TV, and increases the size of the overall test database.

In the WRAL-TV field test (Reference 3) a summary of the tests shows that 146 of the 163 points tested received DTV. This was termed a service availability of 89.6 percent. In the WGN-TV outdoor field test it was reported that 104 of the 112 sites tested received DTV, or a 93 percent success rate. The WCBS-DT success rate was computed by evaluating the predicted sites, for signal level and terrain obstructions, against actual measured sites. As shown in Table 3, sites at distances where the Longley-Rice methodology was used to predict coverage, 96 % of the predicted sites received DTV. In the "concrete canyons" 89% of the sites that were not directly blocked received DTV.

SECTION 4.6 NTSC/DTV LONGLEY-RICE AND TASO COVERAGE ANALYSIS

Figure 13 contains a map of the Longley-Rice calculated coverage of WCBS-TV, Channel 2 operating from the World Trade Center. On the color version of this map interference from other NTSC stations is shown in green and interference from other DTV stations is shown in blue. Terrain obstruction is shown in red, and occurs mainly to the north and northwest of the Empire State Building. A portion of the WCBS-TV coverage area receives interference from other NTSC channels to the east and southwest of the station.

Figure 14, contains a map of the calculated ideal coverage of WCBS-HD, Channel 56, operating from the Empire State Building. This is an ideal case coverage map, as it assumes a omnidirectional antenna pattern, capable of radiating the maximum authorized effective radiated power of 349 kW in all directions. On the color version of this map interference from other DTV stations is shown in green and interference from other NTSC stations is shown in blue. Terrain obstruction is shown in red, and occurs mainly to the north and northwest of the Empire State Building.

Figure 15 contains a map of the calculated coverage of WCBS-HD, using the measured antenna azimuth pattern. This map shows the coverage of Channel 56, operating from the Empire State Building, with the azimuth pattern of the Harris Model TAD7342A01H antenna, as measured during testing on the antenna test range. On the color version of this map interference from other DTV stations is shown in green and interference from other NTSC stations is shown in blue. Terrain obstruction is shown in red, and occurs mainly to the north and northwest of the Empire State Building.

Figure 16 contains a terrain profile to a CBS test site, R312-08, where no DTV was received. This profile map was generated using a computer program developed by Techware, Inc. which uses the Longley-Rice propagation model with a digital terrain database. The figure shows: the predicted Longley-Rice field, the FCC Propagation Model field, the transmitter and receive site parameters, bearing, distance and obstructions. As predicted by the Longley-Rice model no DTV reception was measured at this site. The terrain profile shows that Radial 312 has mountains that obstruct DTV reception.

Table 5 contains a comparison of the WCBS-HD DTV coverage to WCBS-TV NTSC coverage at sites where data was available for both. The source of the NTSC picture quality data is the measurement report of Jules Cohen prepared in 1988. The report
Contains data gathered at distances of 10 to 50 miles from the World Trade Center, at increments of approximately 2 miles. The TASO rating of points that most closely correspond to the DTV test sites were divided into two categories: those sites with a TASO rating of 1,2 or3, that is a picture considered excellent to fine, and TASO 4 or 5, that is a marginal to inferior picture.

A total of 50 sites had NTSC measurement data and corresponded closely to DTV measurement sites that predicted DTV service, according to the Longley-Rice methodology. As shown in Figure 5, the 47 sites that had a TASO 1,2,or 3 rating all received DTV, or a 100% success rate. Of the remaining three sites, with a TASO rating of 4 or 5, only one received DTV as predicted, or a 33% success rate.

5. SUMMARY

The WCBS-DT field test in New York demonstrated that the digital transmission from a tall structure in the center of a large metropolitan area is successful over a large area.
Many of the radials measured achieved a 100% success rate. Overall, the radials, arcs, and grids, at least 10 miles from the Empire State Building achieved a success rate of 96%. Sites that were in New York City achieved an 89% success rate, unless tall buildings blocked them.

The DTV coverage, unless it was significantly blocked by terrain, was found to extend out to approximately 55 miles. On radials that were terrain limited (for example on R312), blocked by hills and mountains resulted in loss of signal. Testing also showed that, at Channel 56 the reflections caused by tall buildings, in particular the World Trade Center, did not significantly impair DTV reception.

It was demonstrated that in the inner city, "the concrete canyons of New York", reception cannot be realistically evaluated by a test methodology of evaluating service at a 30 foot antenna height above ground, and pointing the antenna at the transmitter. Service in these areas is a complex mechanism consisting of reflected signals, blockage by tall building and dynamic ghosting due traffic flow. This resulted in variable data accumulated in the two grid areas that were measured in Brooklyn and Manhattan. In fact at some sites DTV service was obtained by lowering or raising the antenna height above ground to a level other than 30 feet, or changing the direction at which the receive antenna was pointed. These conditions were termed as appropriate for this environment, and data when obtained under these circumstances was considered valid. The inner city "concrete canyon site ", G3, achieved an 83% success rate. Realistically, over-the-air service in these areas would be obtained using antennas mounted on top of tall buildings that generally would have access to a MATV system and line of sight to the transmitter at the Empire State Building.

The test also showed that a combination of terrain blocking and the presence of a lower adjacent NTSC Channel, resulted in predictable loss of service to an area on the North Shore of Long Island past approximately 40 miles. This is a unique circumstance that may benefit from a translator. It was generally found that DTV service was available even in areas that suffered poor UHF reception at Channel 41.

6.0 REFERENCES

(1) "FIELD TEST RESULTS OF THE GRAND ALLIANCE HDTV TRANSMISSION SUBSYSTEM", September 1994

(2) "ATSC Digital Television Standard", Document A53, by Advanced Television System Committee

(3) "WRAL-HD DTV FIELD TESTING", Paper by Luther Ritchie, Capitol Broadcasting Co., Presented at the 1998 National Association of Broadcasters Convention.

(4) "TRIBUNE/WGN DTV FIELD TEST", Report by. M.McKinnon, M. Drazin, and G. Sgrignoli.

(5) "ATSC FIELD TEST VEHICLE DESIGN INFORMATION". G. Sgrignoli, April 14, 1997.

(6) "SPECIFICATION FOR A HDTV FIELD STRENGTH MEASUREMENT VEHICLE", G. Coppa and R.P.Seidel, CBS Document, June 9, 1997.

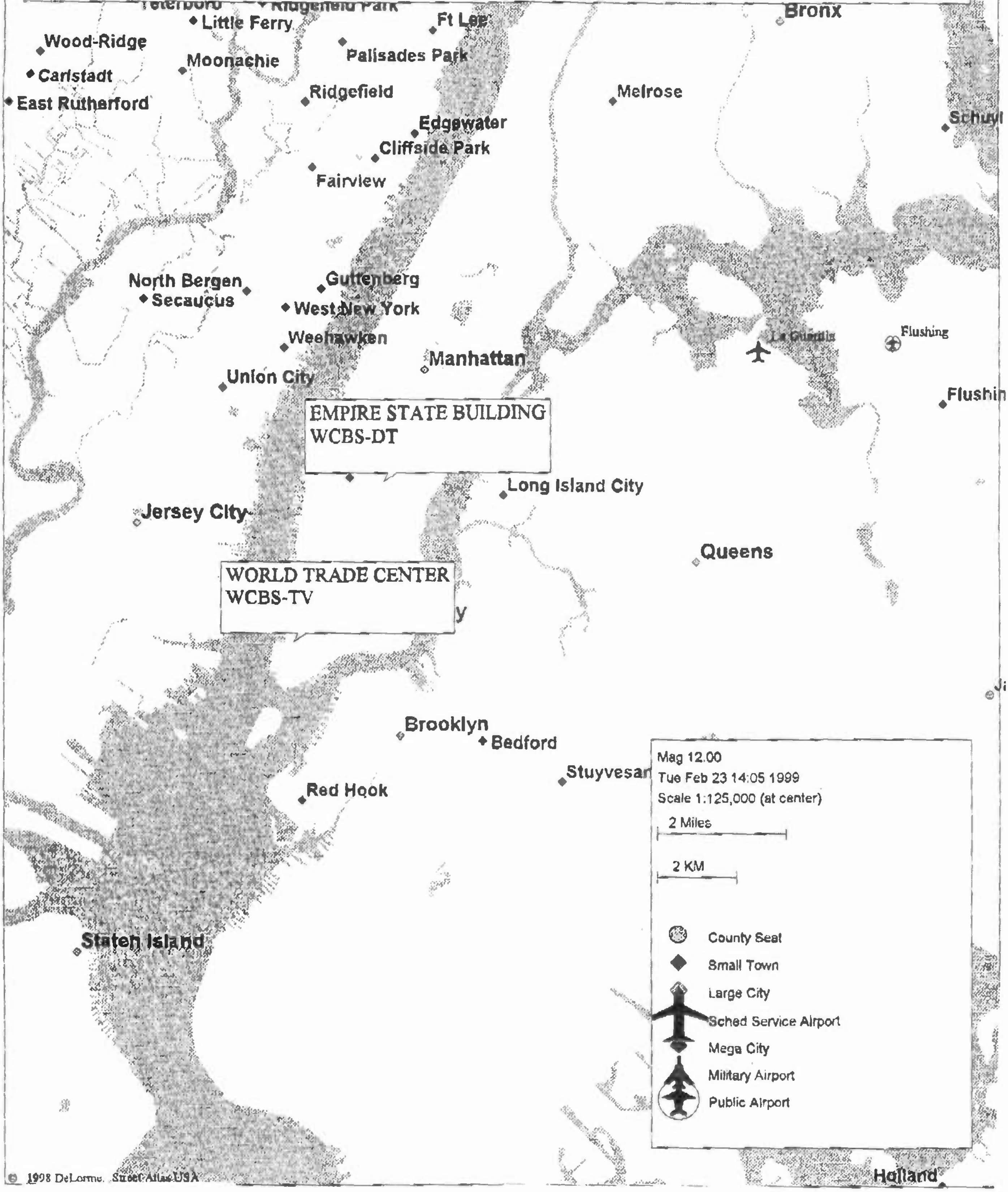
(7)"TELEVISION RECEPTION STUDY PREPARED FOR CBS INC., STATION WCBS-TV", Jules Cohen & Associates, P.C., November 23, 1988.

(8) "TECHNICAL STUDY OF TELEVISION BROADCASTING FROM THE WORLD TRADE CENTER", D. Fink, date N/A.

7. ACKNOWLEDGEMENTS

The following individuals participated in the New York DTV Field tests: Gary Sgrignoli and Steven Heinz of Zenith Electronics Corporation, and Joseph Palucci, Kevin Coleman, Adrianna Higuera and Walter Sidas of CBS Corporation. Greg Coppa and Robert Seidel prepared the specification for the DTV test vehicle.

Figure 1- Location of Transmitters



WCBS-DT TRANSMITTER BLOCK DIAGRAM

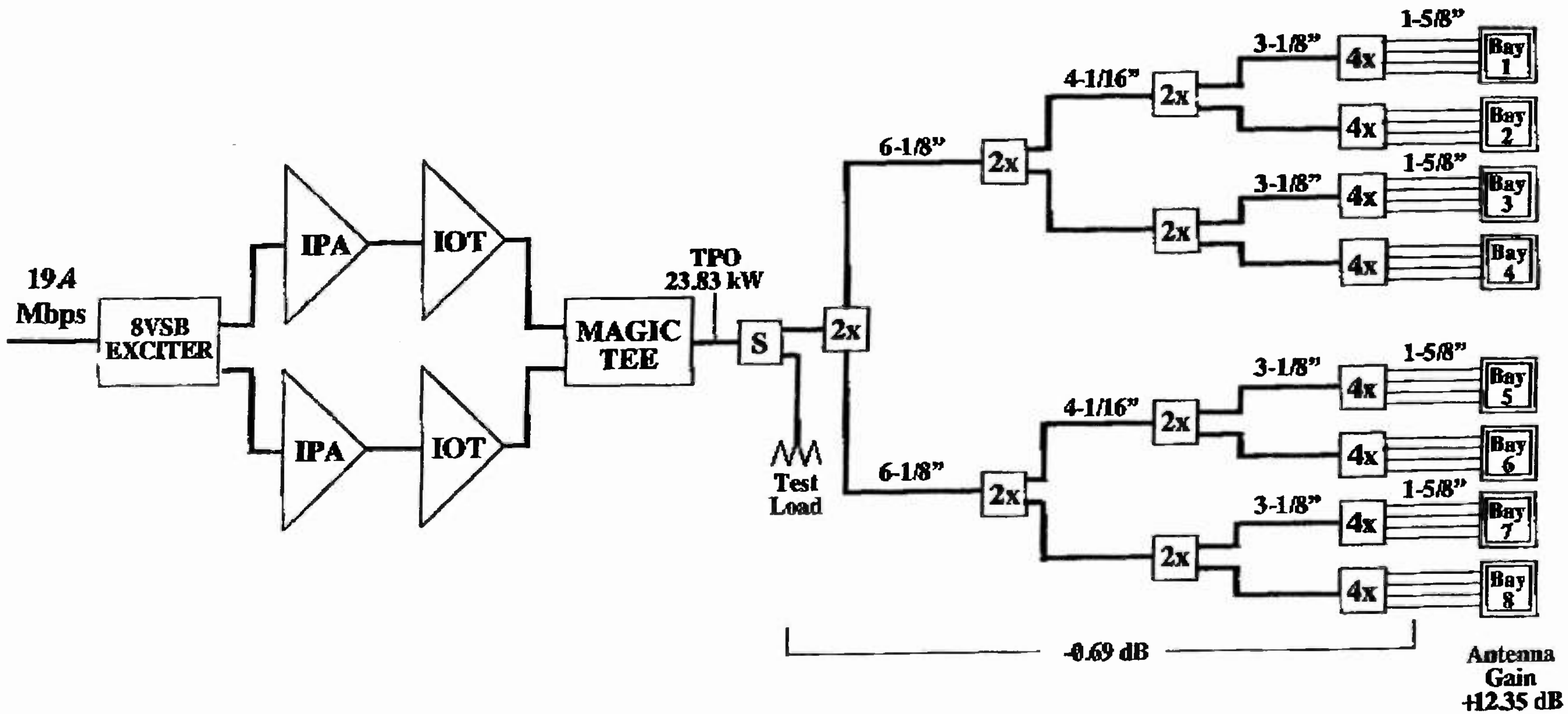
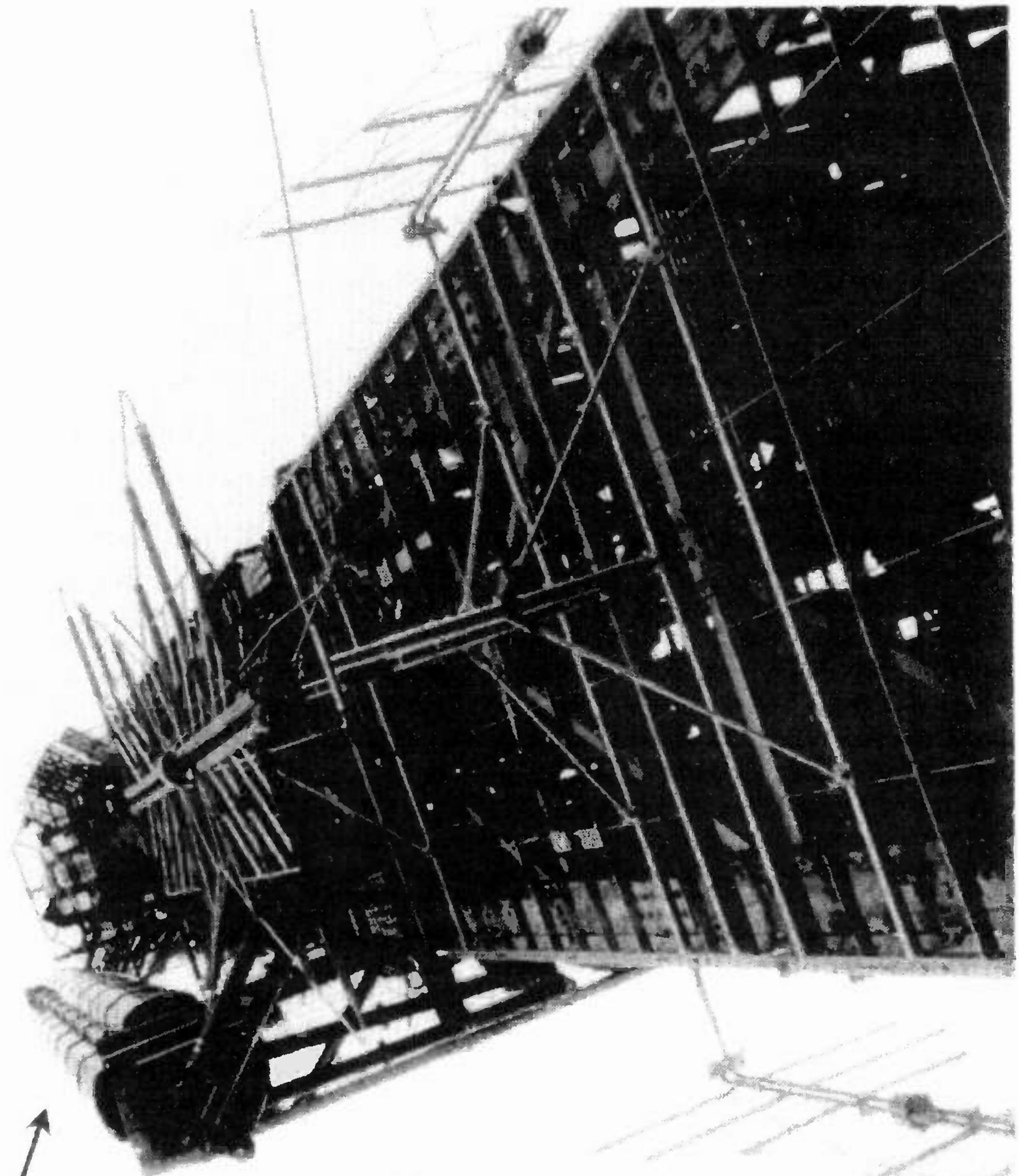


Figure 2



WCBS-DT
CH-56

HARRIS

Calculated Elevation Pattern

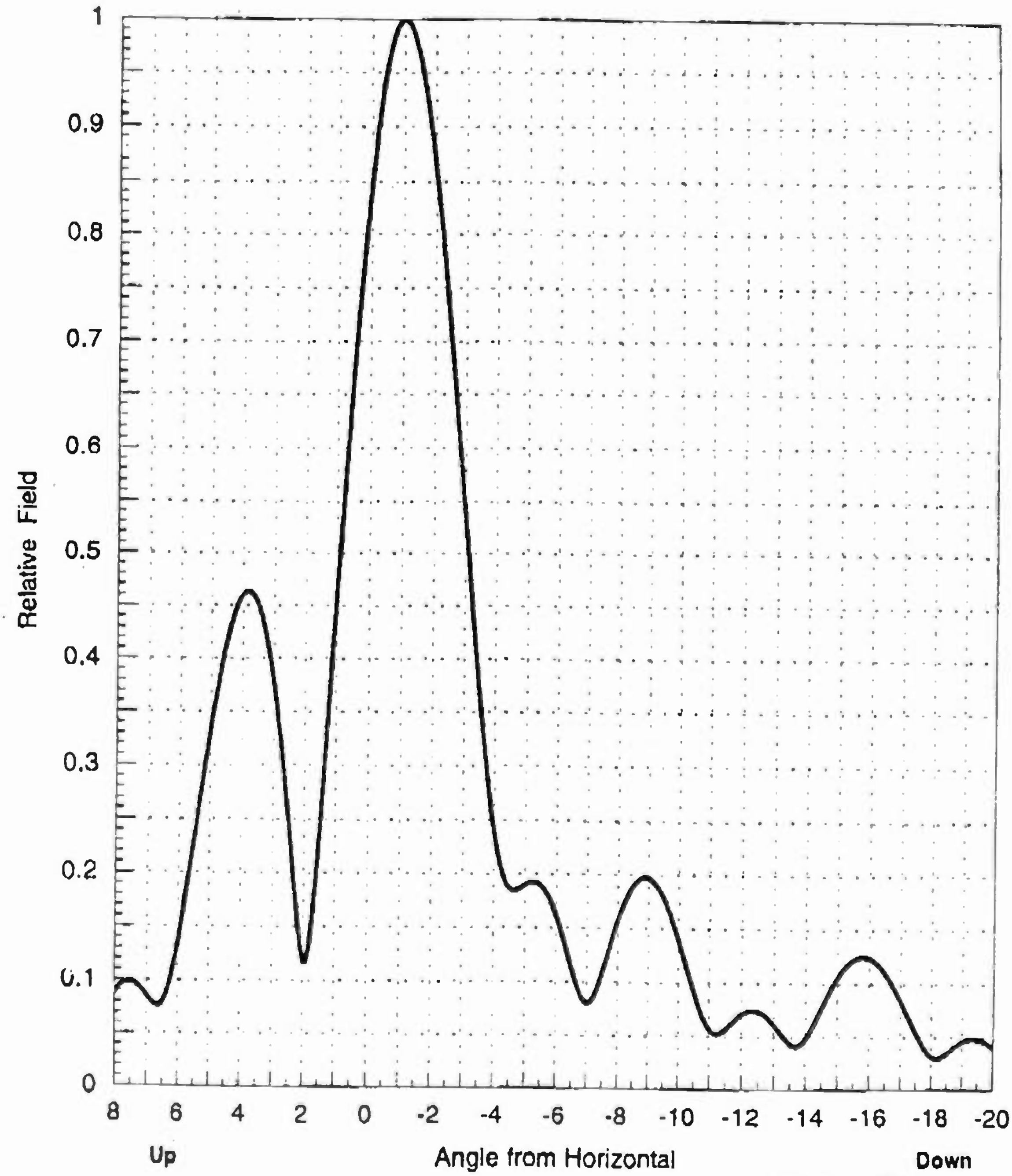


Figure 4.1
Channel 28 Elevation

Series: TAD
Harris Pattern No.: 7342E02H

HARRIS CORPORATION BROADCAST DIVISION PO BOX 4290 3200 WISMANN LANE QUINCY, IL 62305 U.S.A. 217/222-8200

HARRIS

Measured Relative Field Pattern

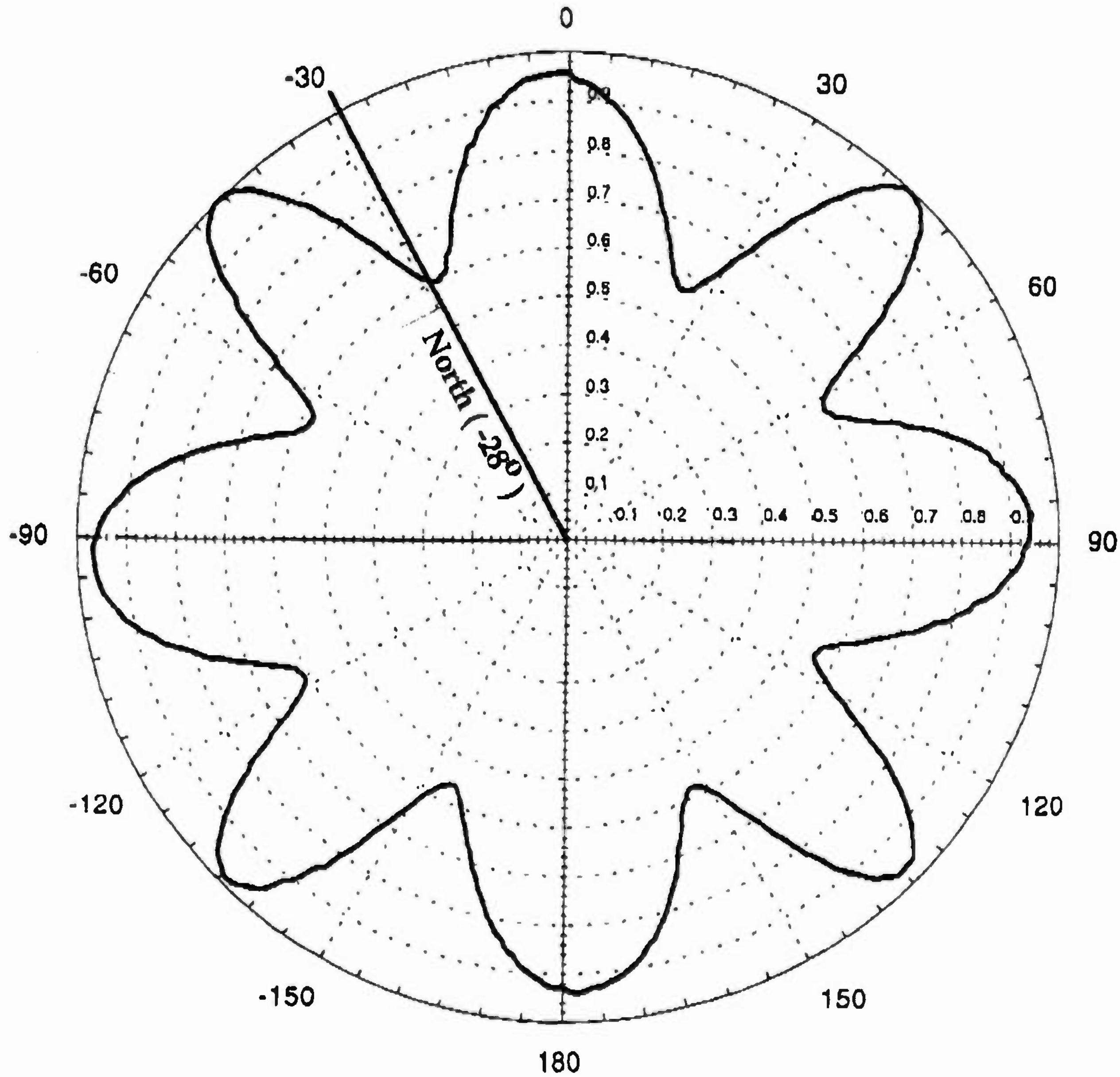


Figure 4.2
Channel 28 Azimuth

Series: TAD
Harris Pattern No.: 7342A02H

HARRIS CORPORATION BROADCAST DIVISION PO BOX 4290 3200 WISMANN LANE QUINCY, IL. 62305 U.S.A. 217/222-8200

Calculated Elevation Pattern

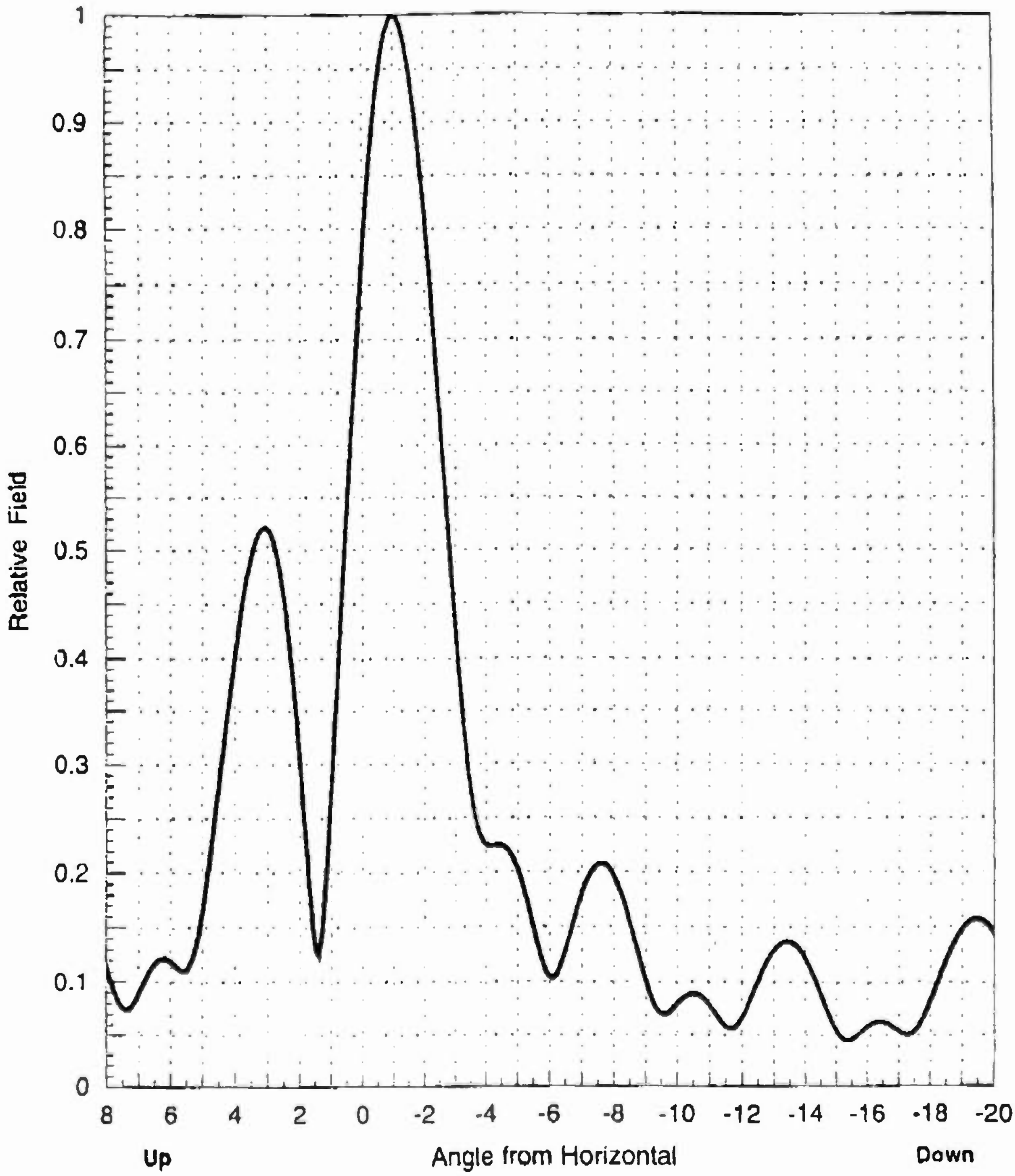


Figure 4.3
Channel 45 Elevation

Series: TAD
Harris Pattern No.: 7342E03H

HARRIS CORPORATION BROADCAST DIVISION PO BOX 4290 3200 WISMANN LANE QUINCY, IL 62305 U.S.A. 217/222-8200

HARRIS

Measured Relative Field Pattern

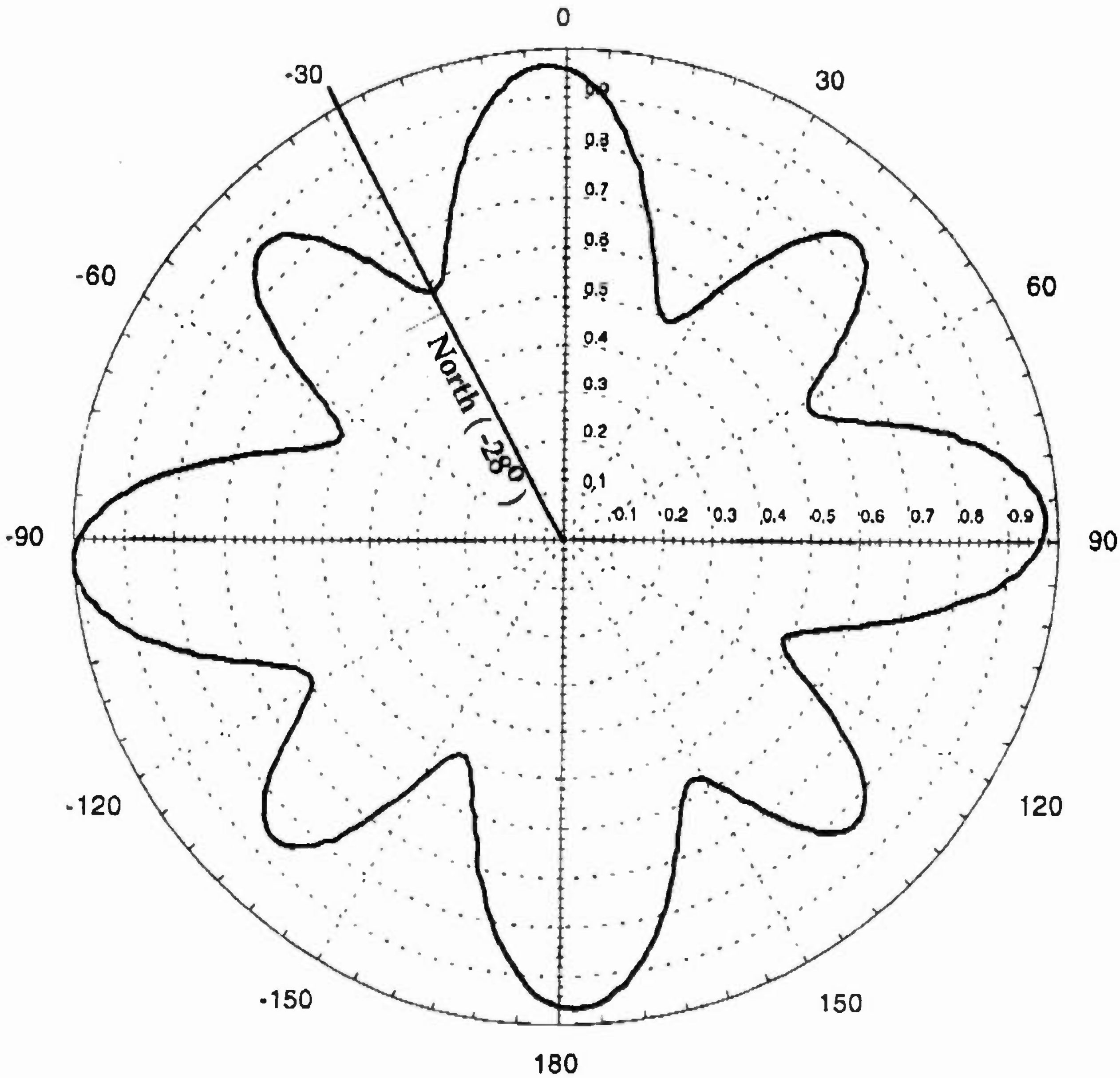


Figure 4.4
Channel 45 Azimuth

Series: TAD
Harris Pattern No.: 7342A03H

HARRIS CORPORATION BROADCAST DIVISION PO BOX 4290 3200 WISMANN LANE QUINCY, IL. 62305 U.S.A. 217/222-8200

HARRIS

Calculated Elevation Pattern

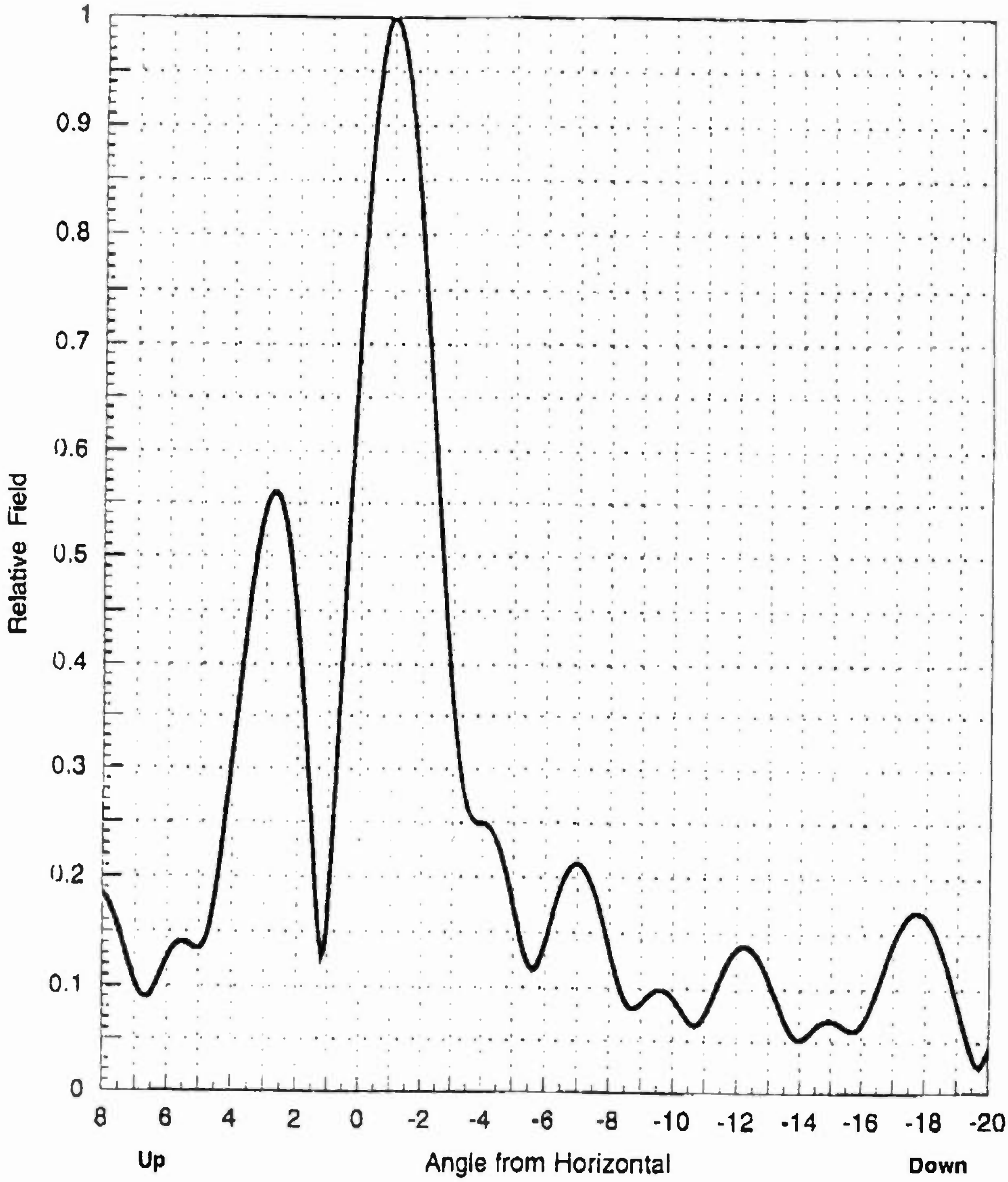


Figure 4.5
Channel 56 Elevation

Series: TAD
Harris Pattern No.: 7342E01H

HARRIS CORPORATION BROADCAST DIVISION PO BOX 4290 3200 WISMANN LANE QUINCY, IL. 62305 U.S.A. 217/222-8200

HARRIS

Measured Relative Field Pattern

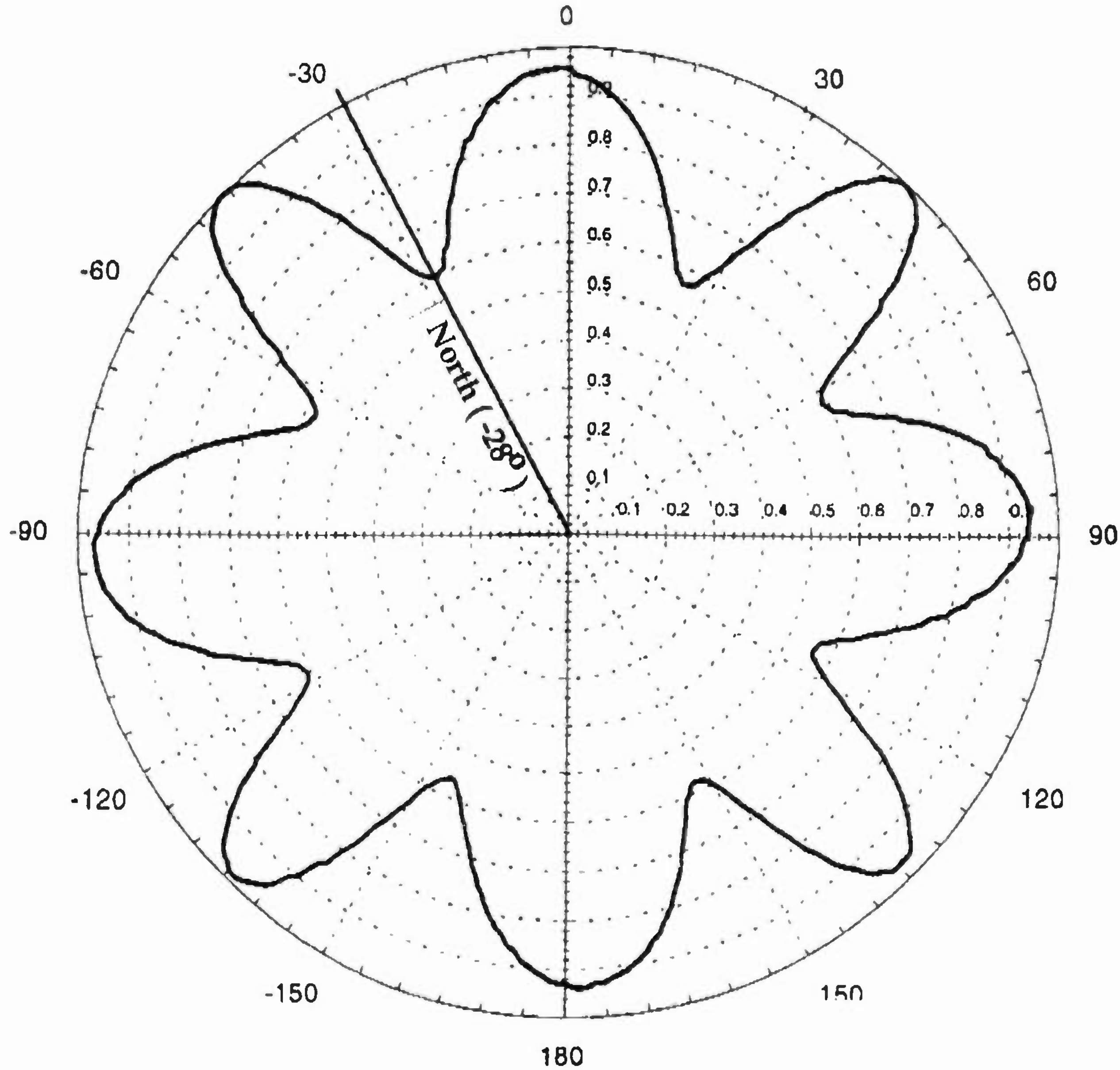


Figure 4.6
Channel 56 Azimuth

Series: TAD
Harris Pattern No.: 7342A01H

HARRIS CORPORATION BROADCAST DIVISION PO BOX 4290 3200 WISMANN LANE QUINCY, IL. 62305 U.S.A. 217/222-8200

HARRIS

Calculated Elevation Pattern

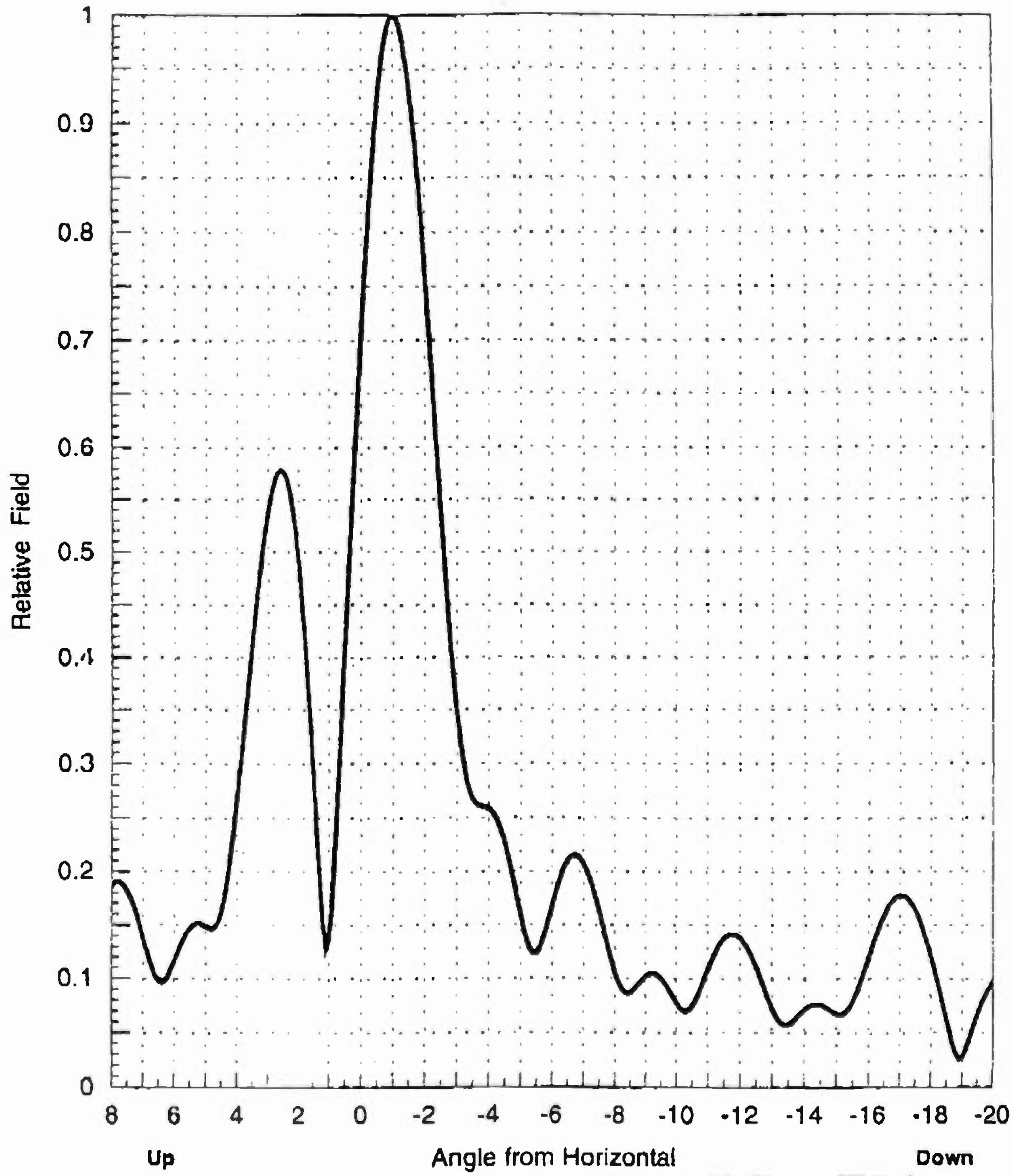


Figure 4.7
Channel 61 Elevation

Series: TAD
Harris Pattern No.: 7342E04H

HARRIS CORPORATION BROADCAST DIVISION PO BOX 4290 3200 WISMANN LANE QUINCY, IL. 62305 U.S.A. 217/222-8200

HARRIS

Measured Relative Field Pattern

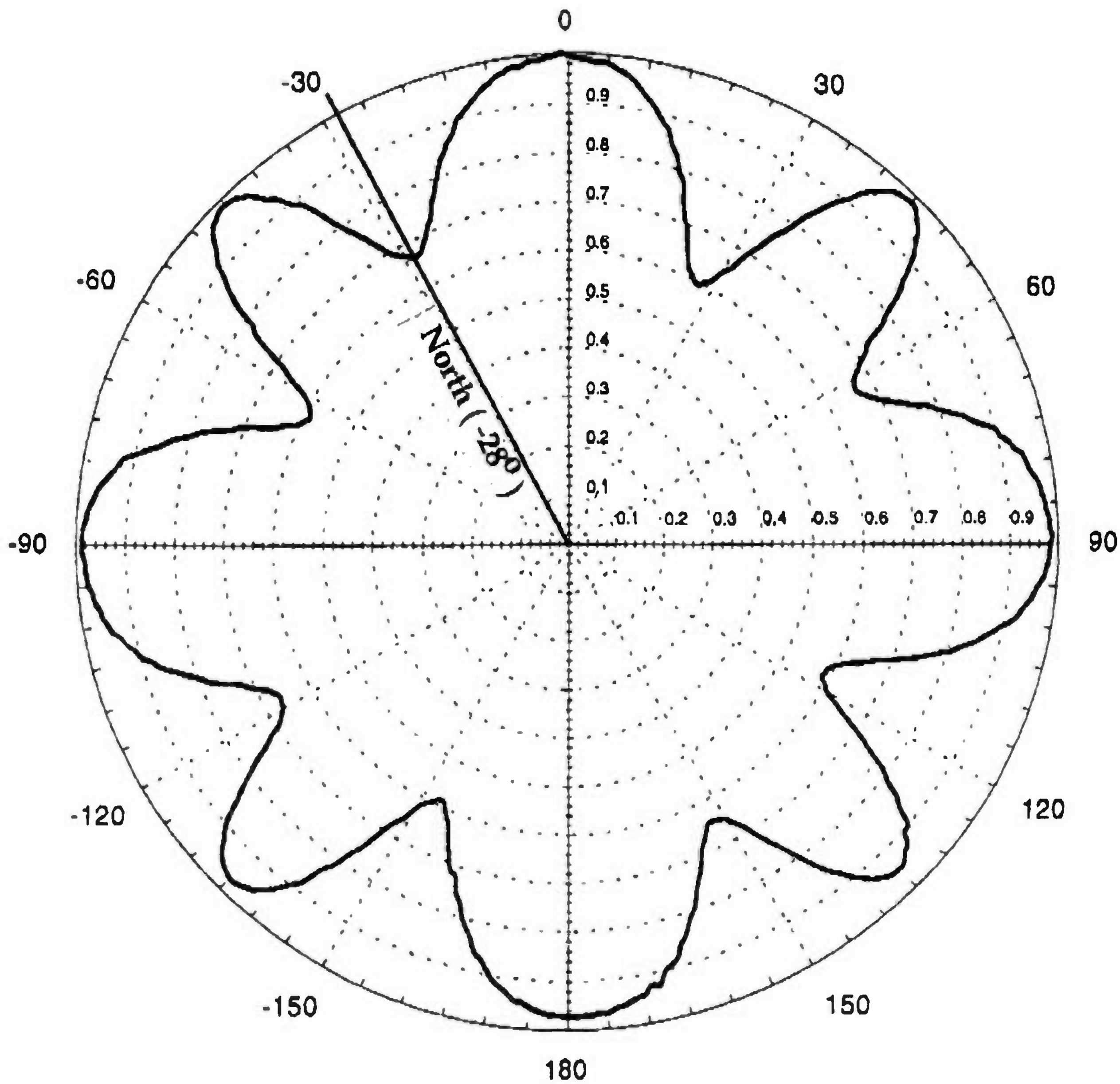
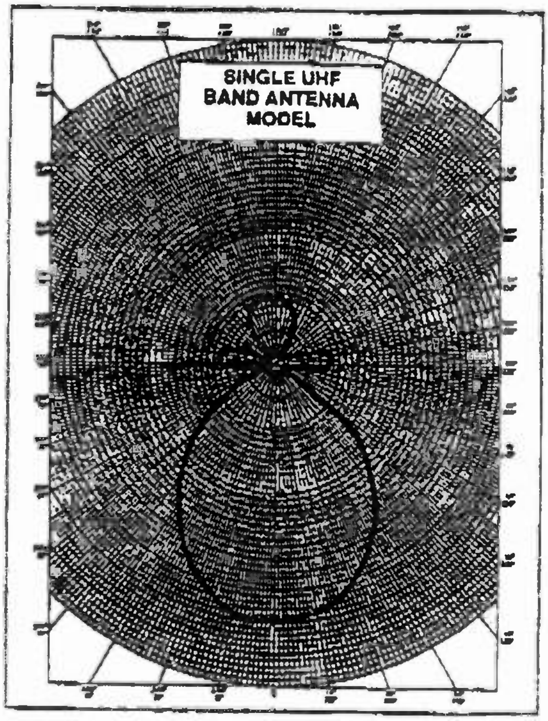


Figure 4.8
Channel 61 Azimuth

Series: TAD
Harris Pattern No.: 7342A04H

HARRIS CORPORATION BROADCAST DIVISION PO BOX 4290 3200 WISMANN LANE QUINCY, IL 62305 U.S.A. 217/222-8200



WADE ANTENNA

SINGLE UHF ANTENNA MODELS:

- WL 14-35/S
- WL 30-83/S

UHF ANTENNA

There are two rugged UHF models. Each antenna is designed for optimum performance over the desired band. The 75 ohm feed point is sealed within the boom. A short length of cable is fitted with a standard "F" connector for connection to the down lead. These light weight, high quality antennas are small in size and big on performance.

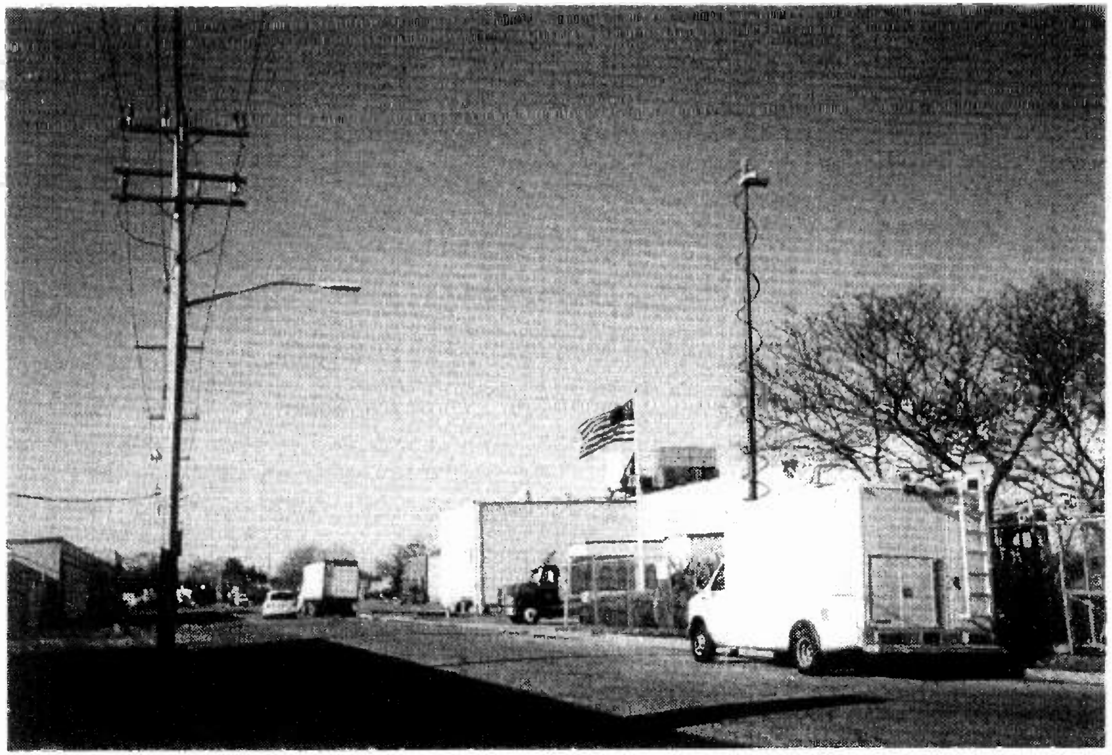
WADE LOG PERIODIC ANTENNA

SINGLE ANTENNA — ELECTRICAL SPECIFICATIONS

SPECIFICATION	WL 14-35/S	WL 30-83/S
FREQUENCY RANGE	470-602 MHZ	566-890 MHZ
CHANNELS	14 To 35	30 To 83
GAIN	11 dBi	11 dBi
IMPEDANCE	75 Ohms	75 Ohms
VSWR	<1.25:1	<1.25:1
FR:BK RATIO	>25 dB	>25 dB
POLARIZATION	H or V	H or V
H. BEAM WIDTH	45 deg.	48 deg.
V. BEAM WIDTH	65 deg.	65 deg.
SIDE LOBE SUPPRESSION*	>30 dB	>30 dB
CONNECTORS	Weather Proof "F"	
STD. MOUNT	3/8" U-Bolts to Fit 2" O.D. Pipe	

*WHERE INTERFERING SIGNALS SUCH AS CO-CHANNEL, ADJACENT CHANNEL AND GHOSTING ARE PRESENT, SPECIAL ARRAYS CAN BE DESIGNED TO REDUCED THE LEVEL OF INTERFERENCE BY AS MUCH AS 40 dB IN MOST CASES.

Figure 5



DTV FIELD TEST TRUCK BLOCK DIAGRAM

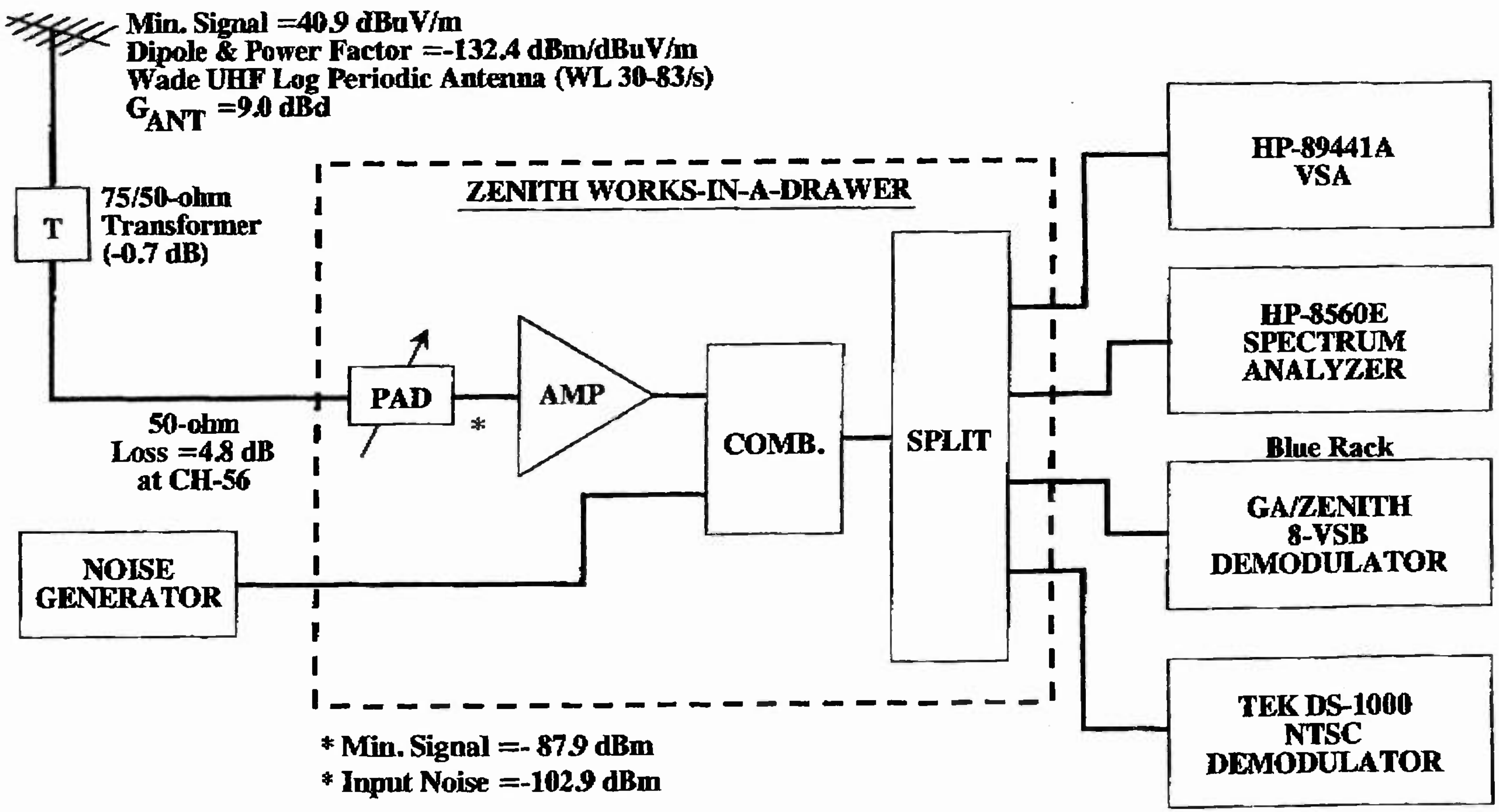


Figure 6.1

ZENITH WORKS-IN-A-DRAWER BLOCK DIAGRAM

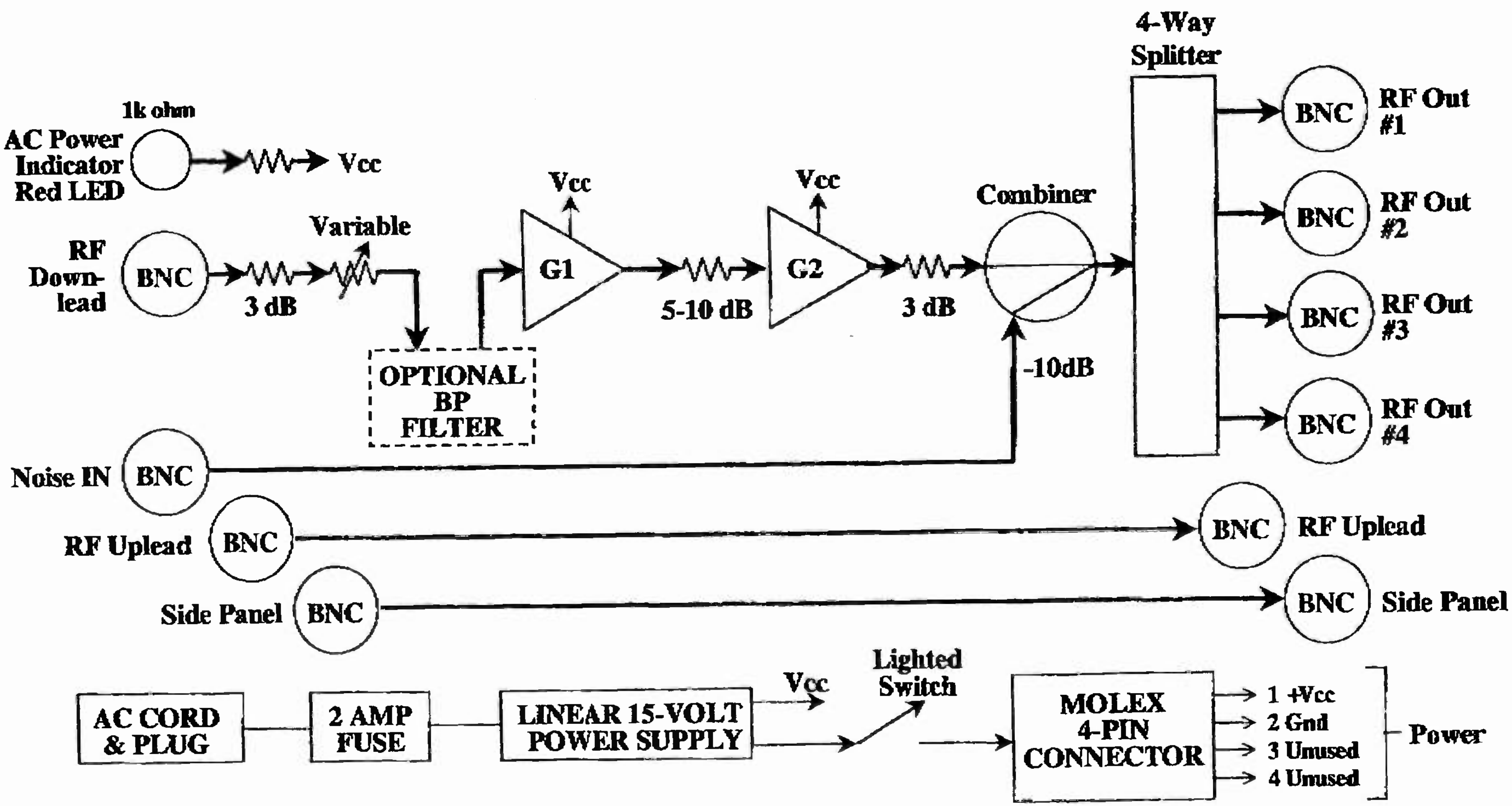


Figure 6.2

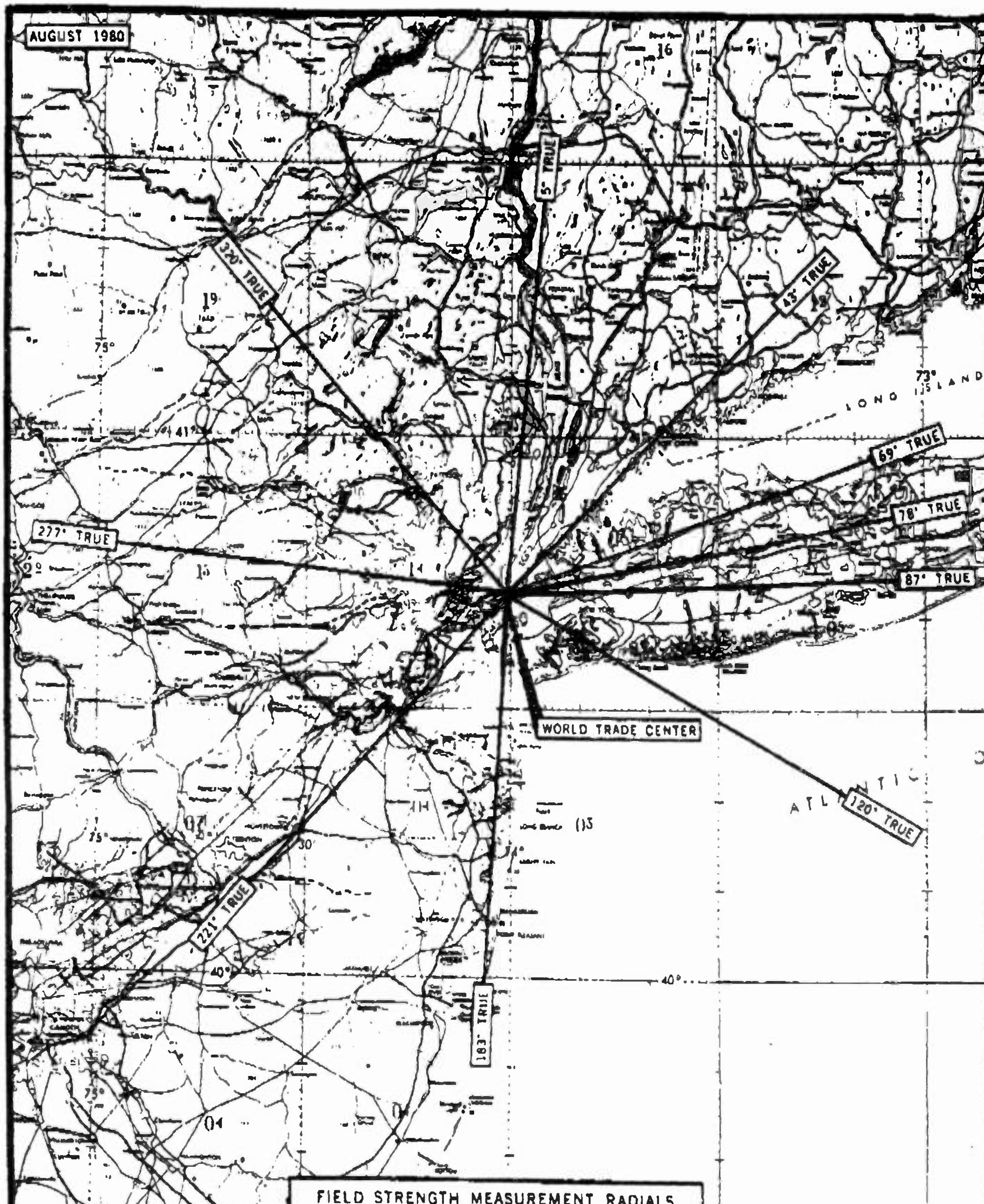
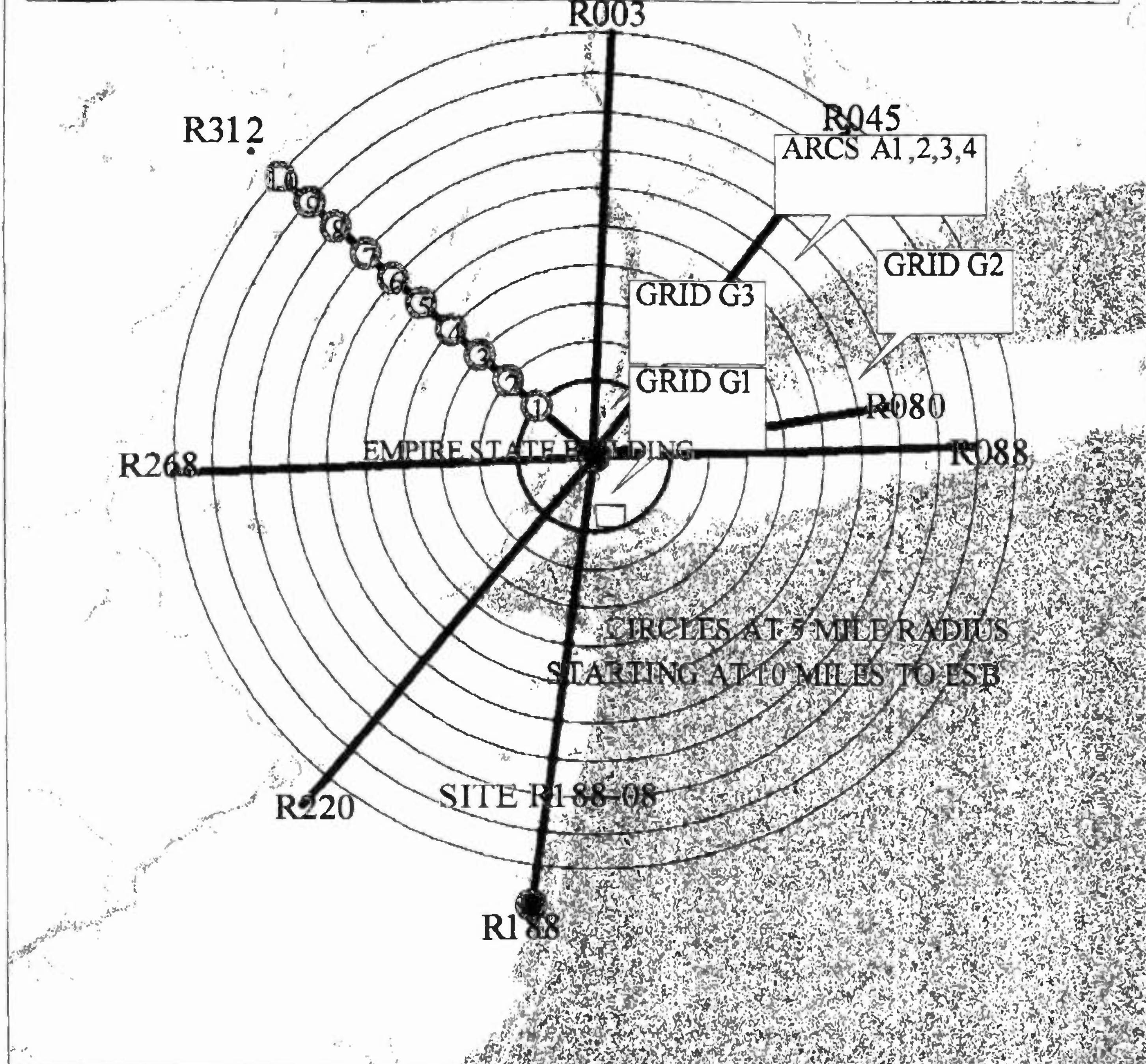

AUGUST 1980
5° TRUE
320° TRUE
43° TRUE
LONG ISLAND
69° TRUE
78° TRUE
277° TRUE
87° TRUE
WORLD TRADE CENTER
ATLANTIC
120° TRUE
221° TRUE
183° TRUE
FIELD STRENGTH MEASUREMENT RADIALS
CBS INC.
STATION WCBS-TV NEW YORK, NEW YORK
CH 2 21.4 KW 1580 FT
Jules Cohen & Associates Consulting Electronics Engineers

Figure 7

Figure 8 - DTV Test Sites



Mag 8.00
Mon Mar 08 07:41 1999
Scale 1:1,200,000 (at center)

Land
Water

20 Miles

20 KM

Table 2 WCBS-DT Field Test Site Summary

Site Type	Site Name	Total Number of Sites	Distance from TX Minimum (miles)	Median (miles)	Maximum (miles)	Bearing Minimum (degrees)	Maximum (degrees)
Radial	R003	10	9.9	32.5	55.1	2	4
Radial	R045	10	10.1	32.5	54.9	45	45
Radial	R080 (Note 1)	7	10.0	25.1	40.0	79	80
Radial	R088	10	10.1	32.6	54.9	88	89
Radial	R188	8	10.0	37.4	55.1	188	188
Radial	R220	10	9.9	32.9	55.0	220	220
Radial	R268	10	9.8	32.8	55.0	266	266
Radial	R312 (Note 2)	9	10.0	35.2	55.1	312	312
Arc	A1	3	20.1	20.2	20.5	45	48
Arc	A2	6	24.9	25.0	25.1	35	44
Arc	A3	6	29.9	30.0	30.1	44	49
Arc	A4	10	34.8	35.0	35.1	45	52
Grid	G2	3	35.5	37.6	38.8	71	72
Special SA	S	9	24.7	50.4	57.9	51	313
Totals		**111**					
	Following Test Sites Were in New York City						
Grid	G1	28	6.2	8.5	11.0	143	185
Grid	G3	10	1.6	4.2	6.2	5	350
Special SB	S	9	1.0	3.4	9.9	27	352
Totals		**47**					

Notes: 1 - Coverage on this radial is limited by terrain and interference from NTSC Channel 55
2- This radial is terrain limited by blocking from high elevation mountains
3 - Considers the blocking of signal by tall buildings.

Table 3 Summary of DTV Test Results

Site Type	Site Name	A Total Number of Sites	B L-R Predicted As Successful	C Measured As Successful Receive	D L-R Predicted No Reception	E Measured No Reception	F=C/B Percent of Predicted Sites Successful
Radial	R003	10	6	5	4	5	83%
Radial	R045	10	10	10	0	0	100%
Radial	R080 (Note 1)	7	7	6	0	1	86%
Radial	R088	10	10	10	0	0	100%
Radial	R188	8	8	8	0	0	100%
Radial	R220	10	10	10	0	0	100%
Radial	R268	10	10	10	0	0	100%
Radial	R312 (Note 2)	9	3	3	6	6	100%
Arc	A1	3	3	3	0	0	100%
Arc	A2	6	6	6	0	0	100%
Arc	A3	6	6	6	0	0	100%
Arc	A4	10	10	9	0	1	90%
Grid	G2	3	3	3	0	0	100%
Special SA	S	9	7	6	2	3	86%
Totals		**111**	**99**	**95**	**12**	**16**	**96%**
	Following Test Sites Were in New York City						
					Note 3		
Grid	G1	28	26	23	2	4	**88%**
Grid	G3	10	6	5	4	5	**83%**
Special SB	S	9	5	5	4	4	**100%**
Totals		**47**	**37**	**33**	**10**	**13**	**89%**

Notes: 1 - Coverage on this radial is limited by interference from NTSC Channel 55
2- This radial is terrain limited by blocking from high elevation mountains
3 - Considers visual observation of tall buildings in evaluation of predicted reception

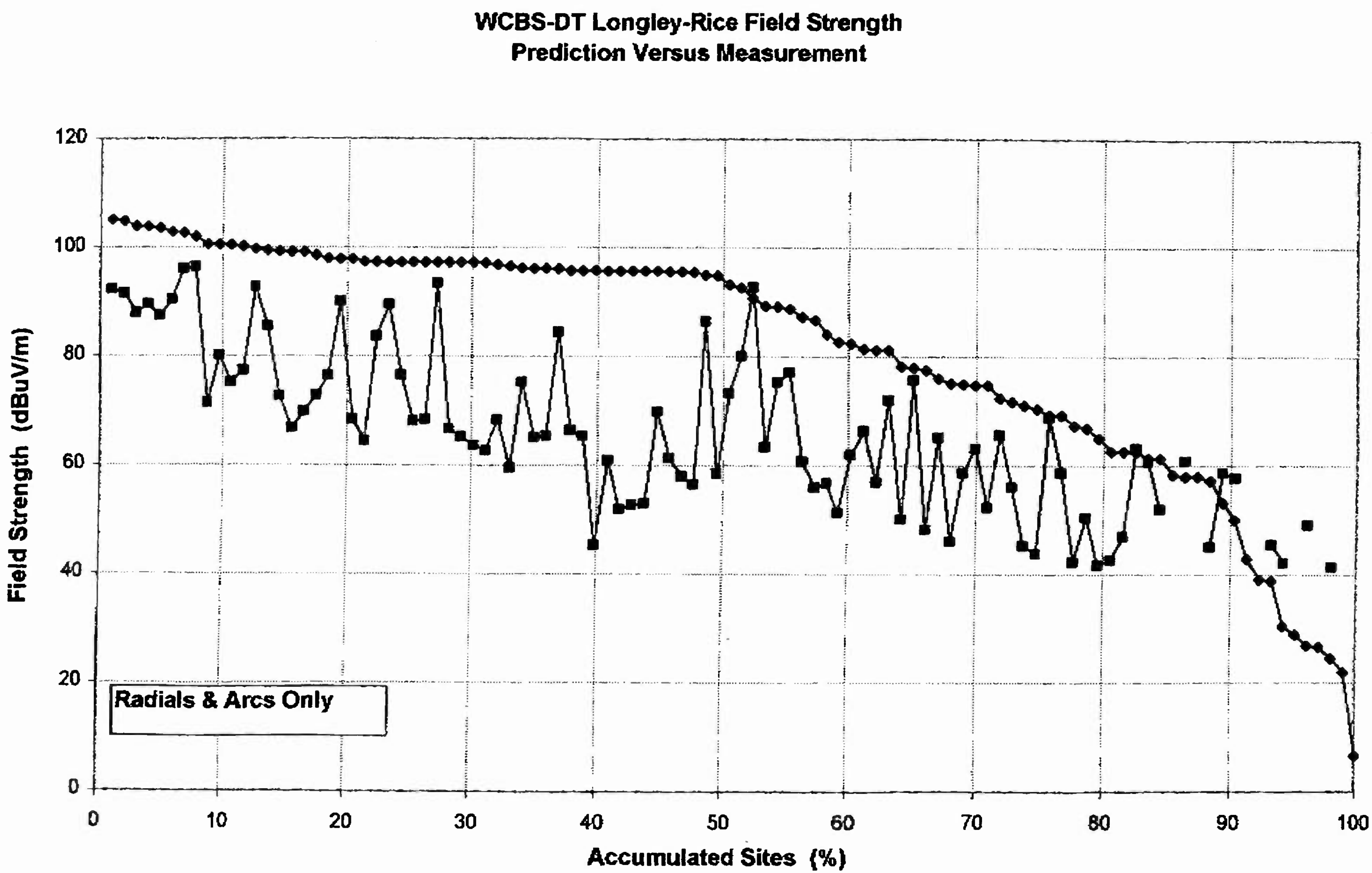

WCBS-DT Longley-Rice Field Strength
Prediction Versus Measurement
Field Strength (dBuV/m)
Accumulated Sites (%)
Radials & Arcs Only
0
20
40
60
80
100
120
0
10
20
30
40
50
60
70
80
90
100

FIGURE 9

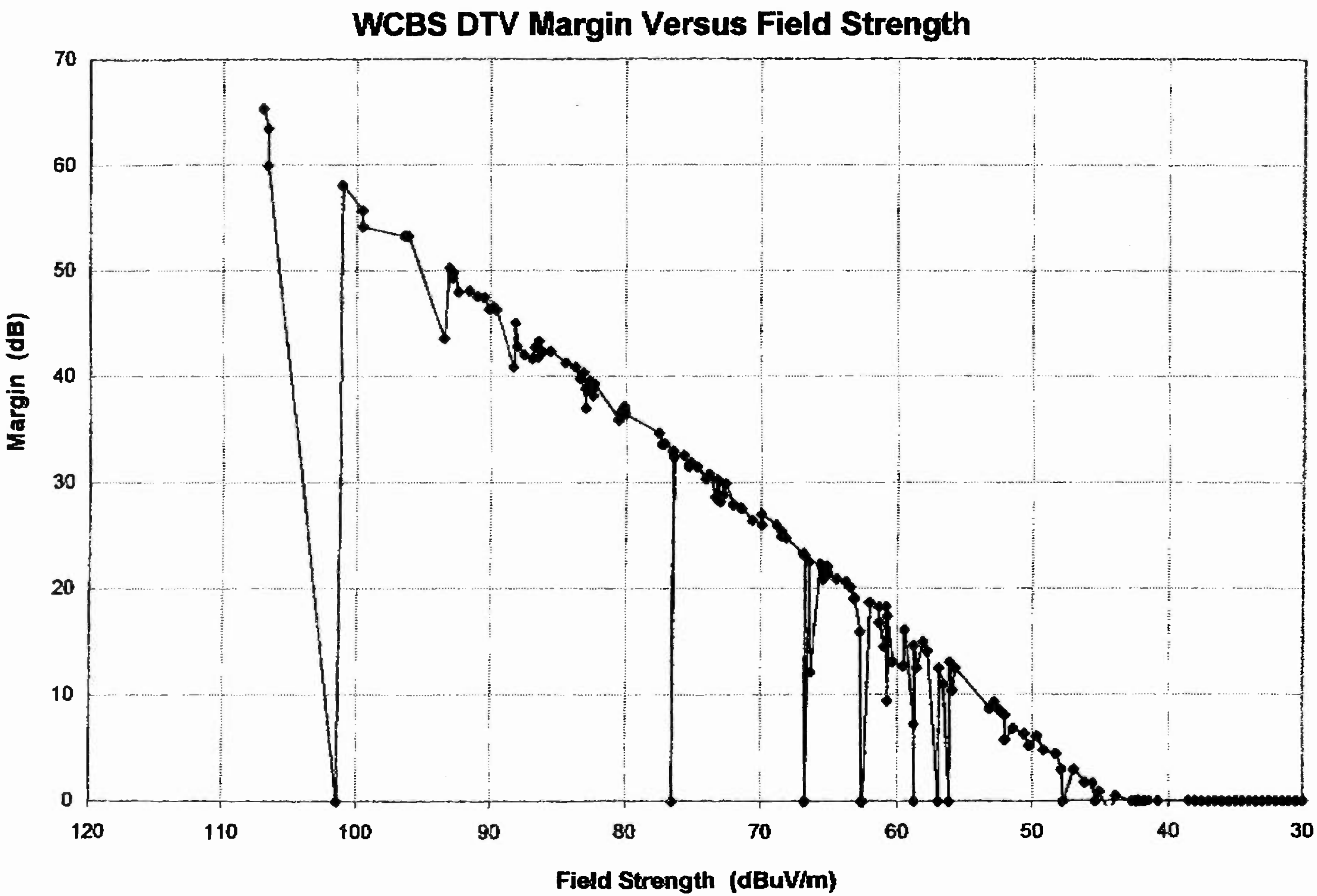

WCBS DTV Margin Versus Field Strength
Margin (dB)
0
10
20
30
40
50
60
70
120
110
100
90
80
70
60
50
40
30
Field Strength (dBuV/m)

FIGURE 10

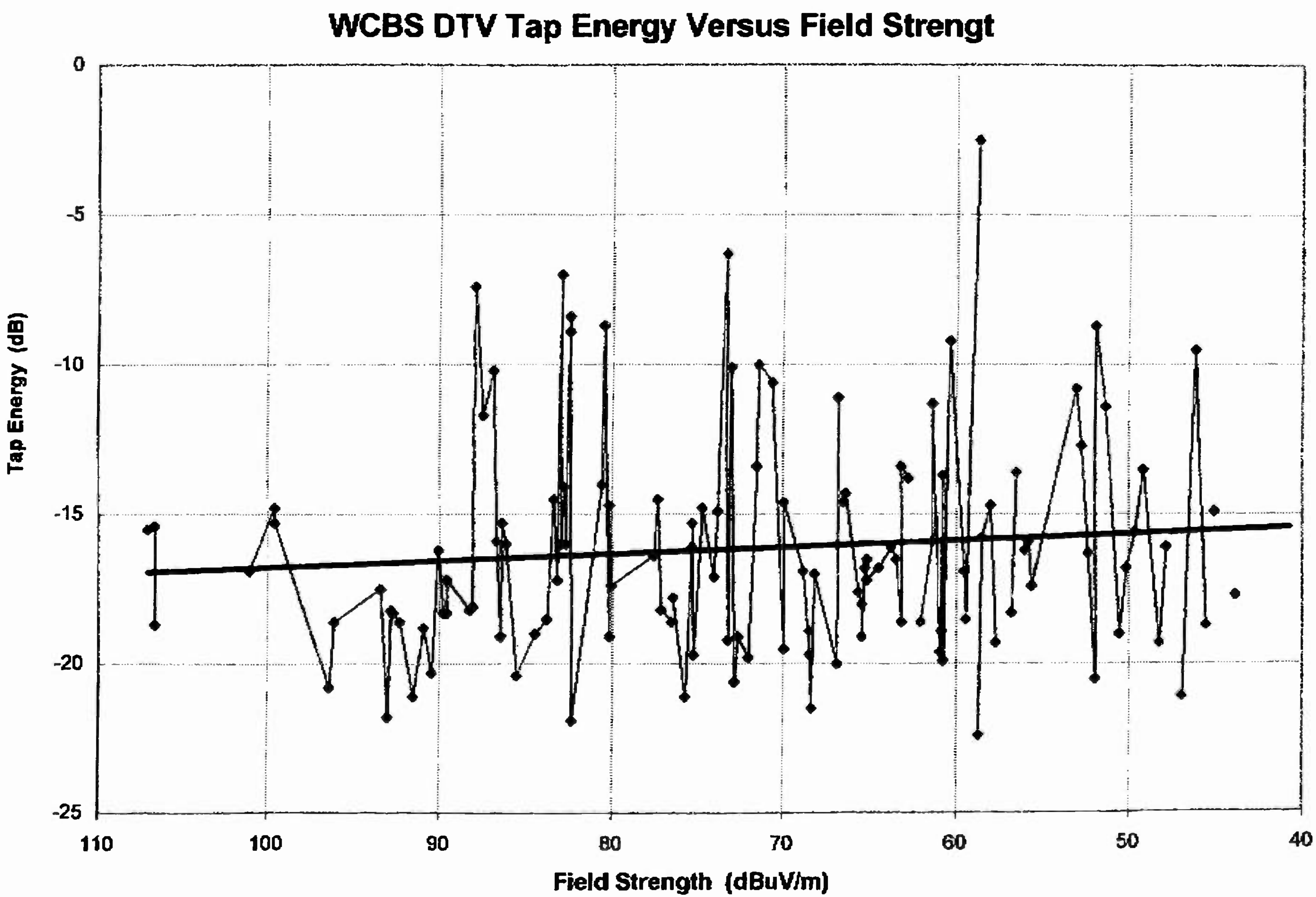

WCBS DTV Tap Energy Versus Field Strengt
Tap Energy (dB)
0
-5
-10
-15
-20
-25
110
100
90
80
70
60
50
40
Field Strength (dBuV/m)

FIGURE 11

WCBS-DT Field Strength Versus Accumulated Sites

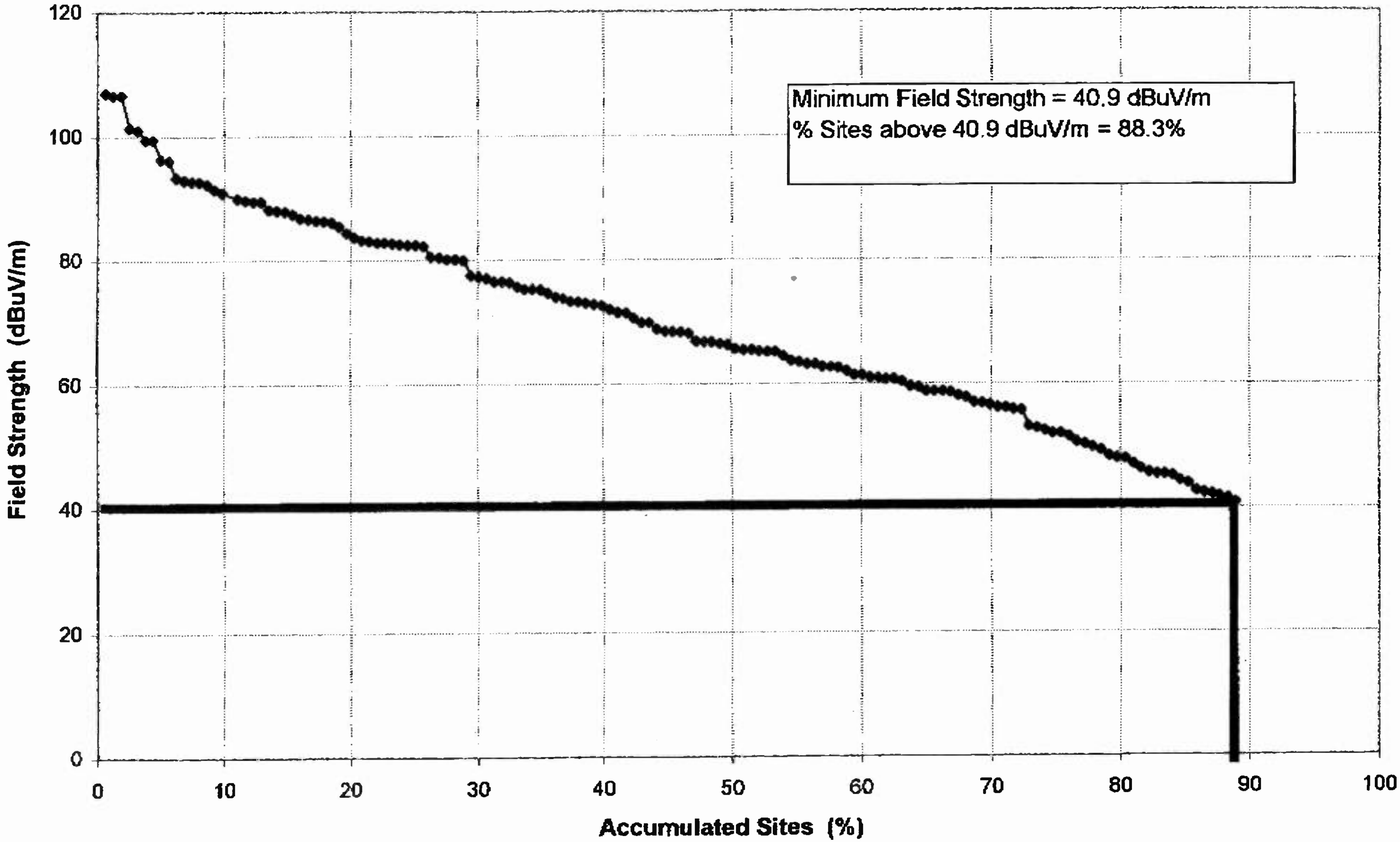
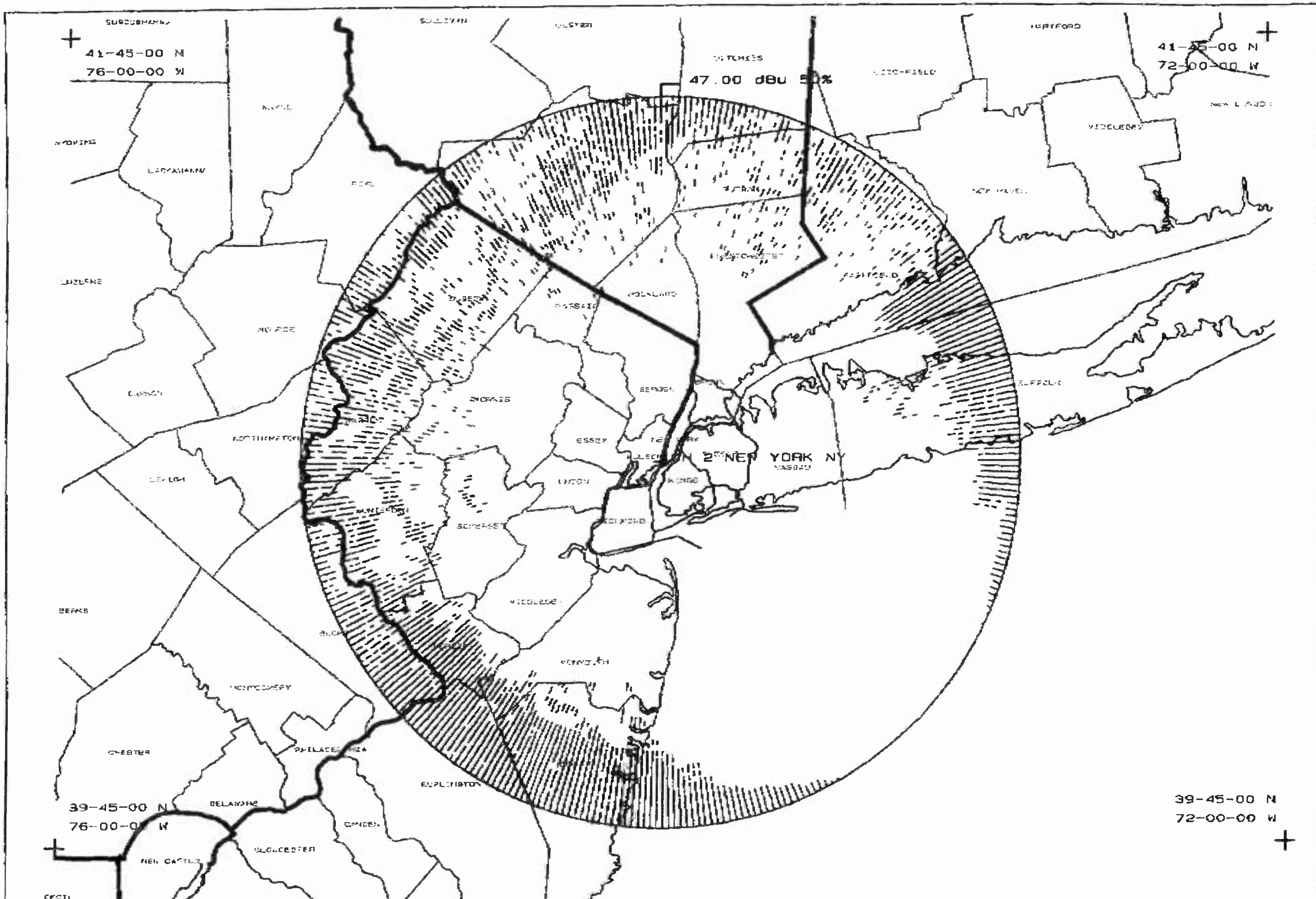


FIGURE 12

Table 4 - SITES WITH NO DTV RECEPTION

Site Name	Distance (mi)	Bearing (degrees)	Number of Obstructions (Note 1)	Measured Field Strength (dBuv/m)	Tap Energy	CH 41 NTSC CCIR	REMARKS
A4-06	34.9	47	0	45.4	-2.2	3	8 usec ghost on Channel 41
R003-06	35.1	3	1	58.8	n/a	5	
R003-07	40.2	2	1	n/a	n/a	5	
R003-08	44.2	3	2	42.2	-7.5	1	
R003-09	50.0	4	1	n/a	n/a	n/a	
R003-10	55.1	3	3	69.7	n/a	n/a	
R080-07	40.0	80	2	41.8	-16.4	3	
R312-02	14.8	312	1	56.2	n/a	n/a	
R312-04	25.0	312	2	45.4	n/a	n/a	
R312-06	35.2	312	1	42.4	-14.8	4	Terrain Block
R312-07	40.3	312	3	n/a	n/a	n/a	Terrain Block
R312-08	45.1	312	5	n/a	n/a	n/a	Terrain Block
R312-10	55.1	312	5	n/a	n/a	n/a	Terrain Block
S5	45.1	52	0	62.7	-12.6	5	
S6	40.9	51	1	61.4	-15.9	5	
S17	56.6	87	2	40.8	N/A		
Following Test Sites Were in New York City							
			(Note 2)				
G1-03	6.4	166	B				Raising antenna to maximum obtained some DTV reception
G1-06	7.7	143		66.8	n/a	3	Notch at 725 MHz
G1-12	8.7	147	B	95.8	n/a	2	Severe Multipath
G1-14	8.3	177		n/a	n/a	n/a	30 usec ghost on Channel 41
G3-02	2.4	5	B	n/a	n/a	n/a	No Channel 41 reception
G3-03	3.4	14	B	95.6	n/a	1	Severe ghosting on Channel 41
G3-05	5.3	18	B	123.7	n/a	1	Ghosting on Channel 41
G3-09	3.0	30	B	n/a	n/a	n/a	Lowered antenna to minimum to get some DTV
G3-12	6.2	31		n/a	n/a	n/a	Lowered antenna to minimum to get some DTV
S09	1.0	300	B	101.5	n/a	n/a	Severe Multipath
S1	3.4	30	B	88.3	-18.9	1	Antenna path blcocked by large building
S14	2.0	43	B	n/a	n/a	n/a	Antenna path blcocked by large building
S9	8.0	168	B	62.6	n/a	n/a	Burst errors on GA rack, dynamic ghosting

Notes: 1. Obstructions analyzed using terrain database and software
2. Obstructions observed, buildings, etc.
3. N/a - measurements was not made



Terrain loss - Red, IX NTSC - Green, ATV - Blue
WCBS-NTSC Ch. 2 ERP 21.4 HAAT 482.6
NEW YORK NY Dir. Ant.
Date: 03 12 99

MAJOR DIVISIONS ARE: 10.0 KM
SCALE IS: 38.38 KM/INCH
23.84 MILES/INCH

Figure 13

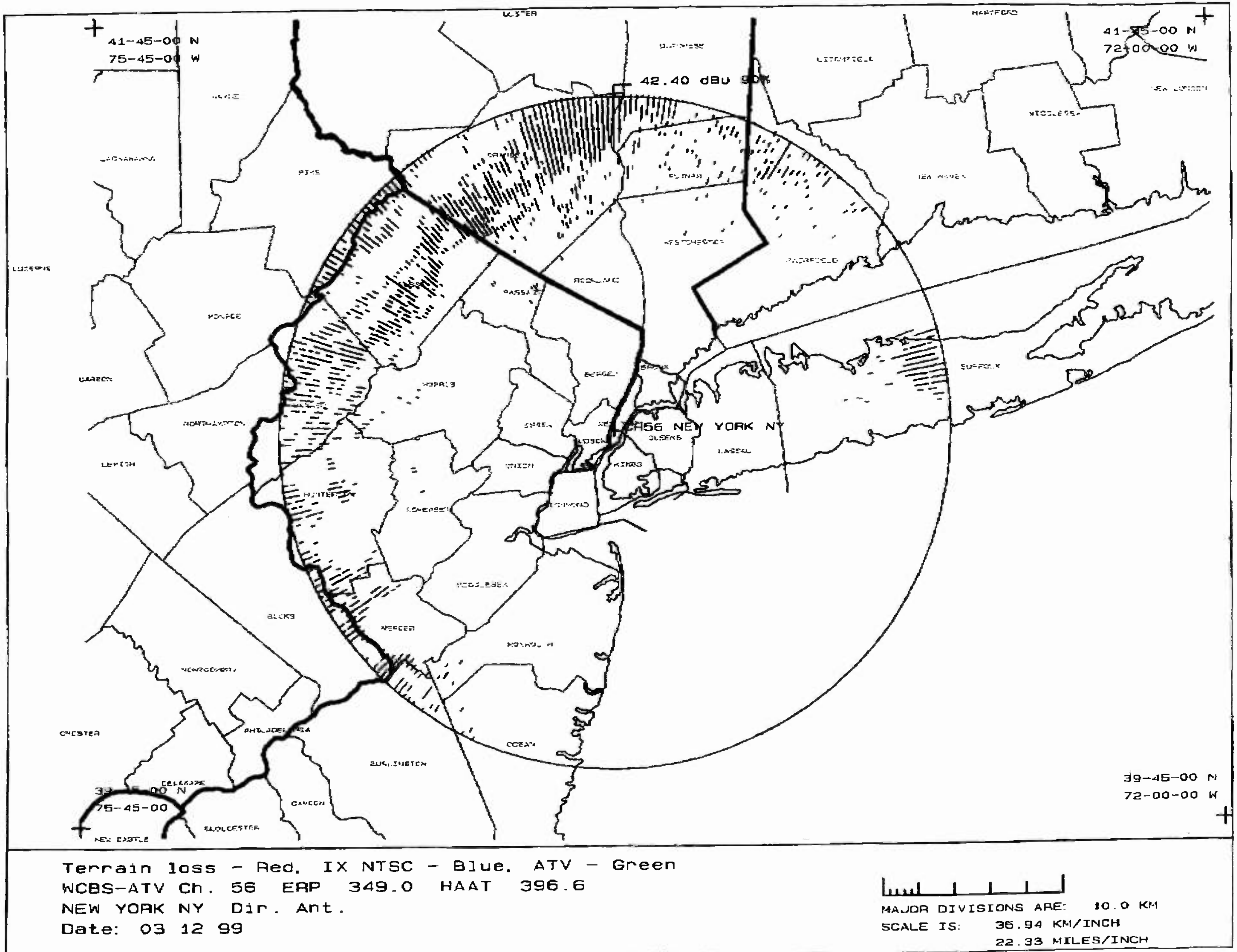
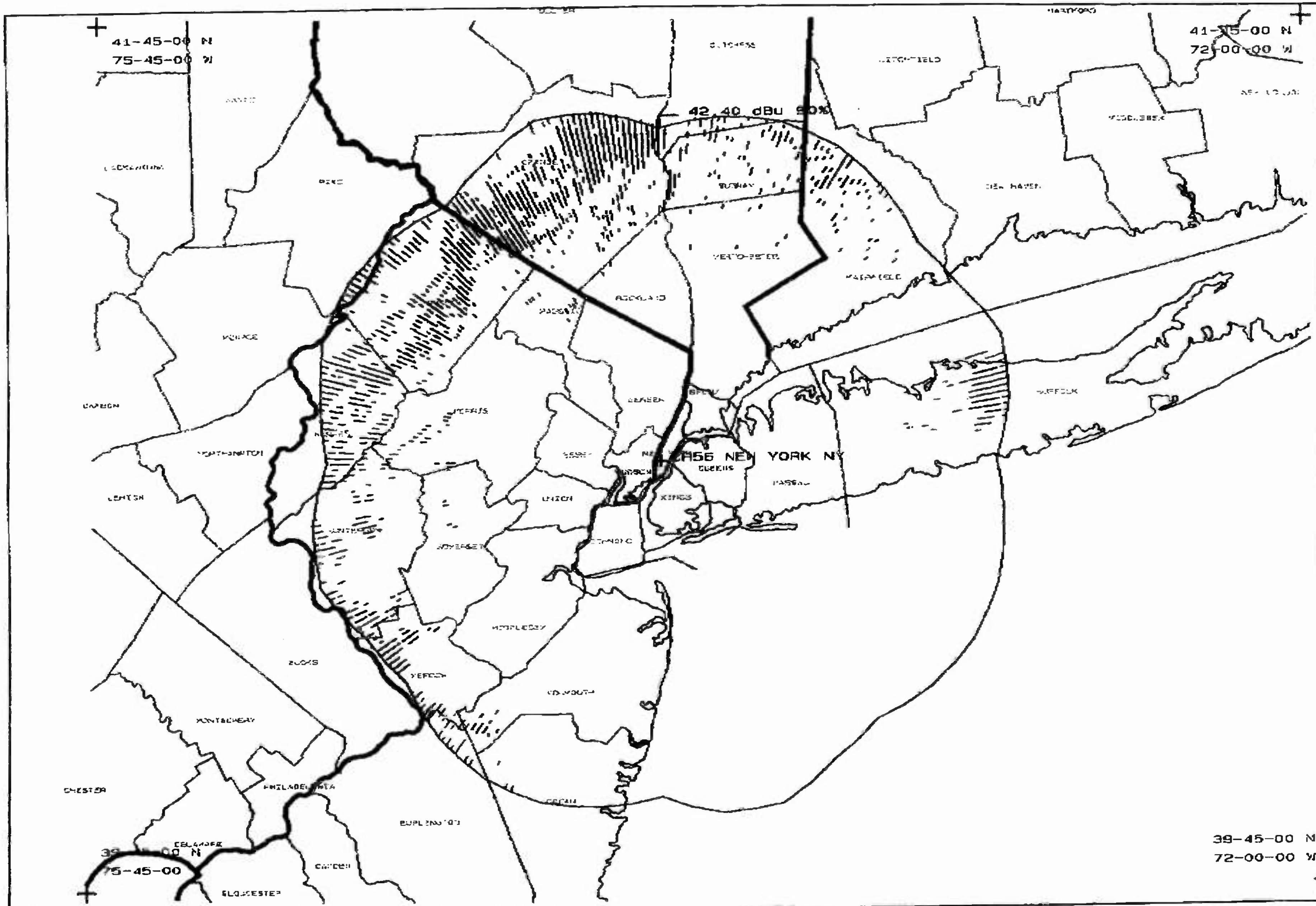

41-45-00 N
75-45-00 W
41-45-00 N
72-00-00 W
42.40 dBu 90%
CH56 NEW YORK NY
39-45-00 N
72-00-00 W
75-45-00
Terrain loss - Red, IX NTSC - Blue, ATV - Green
WCBS-ATV Ch. 56 ERP 349.0 HAAT 396.6
NEW YORK NY Dir. Ant.
Date: 03 12 99
MAJOR DIVISIONS ARE: 10.0 KM
SCALE IS: 35.94 KM/INCH
22.33 MILES/INCH

Figure 14



Terrain loss - Red, IX NTSC - Blue, ATV - Green
WCBS-ATV Ch. 56 ERP 349.0 HAAT 396.6
NEW YORK NY Dir. Ant. cbs 1
Date: 03 12 99

MAJOR DIVISIONS ARE: 10.0 KM
SCALE IS: 35.76 KM/INCH
22.21 MILES/INCH

Figure 15

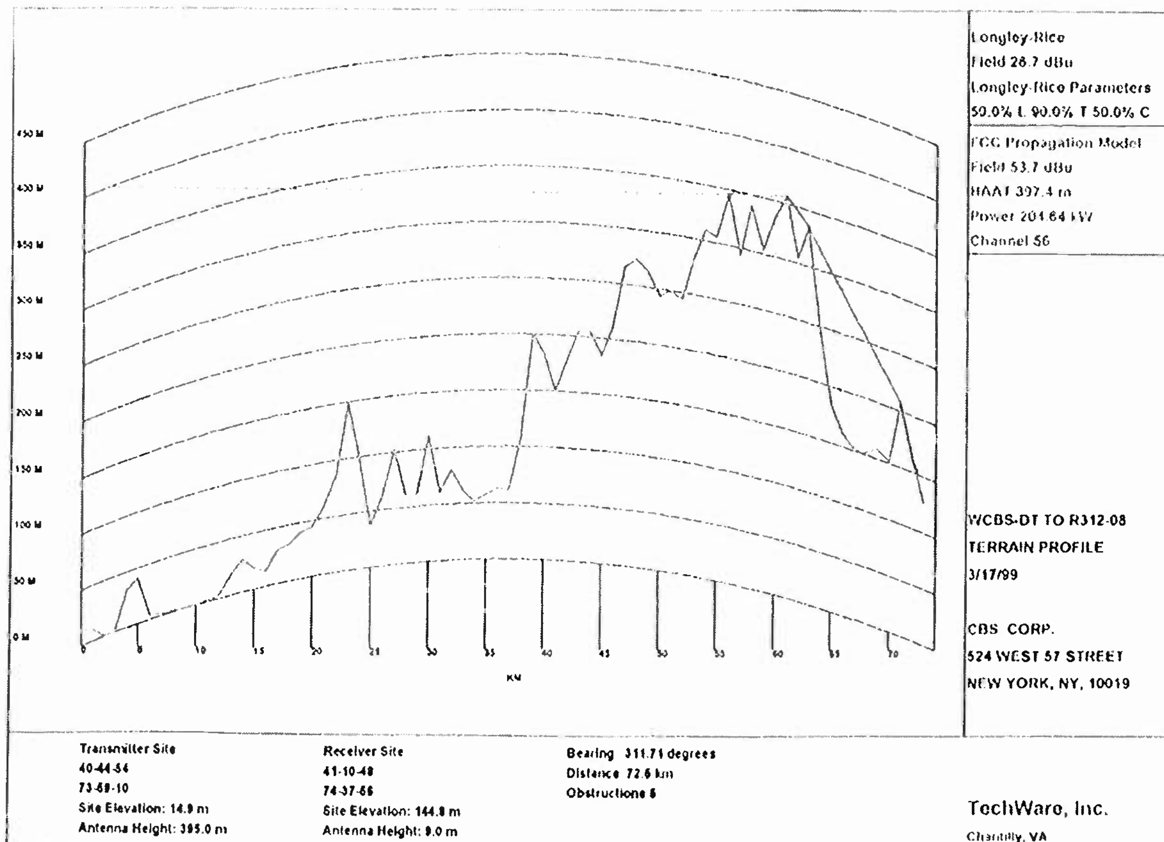

Longley-Rice
Field 28.7 dBu
Longley-Rico Parameters
50.0% L 90.0% T 50.0% C
FCC Propagation Model
Field 53.7 dBu
HAAT 397.4 m
Power 204.64 kW
Channel 56
450 M
400 M
350 M
300 M
250 M
200 M
150 M
100 M
50 M
0 M
KM
WCBS-DT TO R312-08
TERRAIN PROFILE
3/17/99
CBS CORP.
524 WEST 57 STREET
NEW YORK, NY, 10019
Transmitter Site
40-44-54
73-59-10
Site Elevation: 14.9 m
Antenna Height: 395.0 m
Receiver Site
41-10-48
74-37-56
Site Elevation: 144.8 m
Antenna Height: 9.0 m
Bearing 311.71 degrees
Distance 72.6 km
Obstructions 6
TechWare, Inc.
Chantilly, VA

Figure 16

Table 5 Comparison of WCBS-DT Channel 56 Coverage to NTSC WCBS-TV								
	WCBS-TV WITH NTSC TASO 1,2,3				WCBS-TV WITH NTSC TASO 4,5			
	A	B	C	D	E	F	G	H
Number of Sites With Predicted DTV Service Where NTSC Was Measured	Number of Channel 2 Sites Measured	Channel 56 Sites With Successful DTV as Predicted With L-R	Channel 56 Sites With No DTV	DTV Success Rate	Number of Channel 2 Sites Measured	Channel 56 Sites With Successful DTV as Predicted With L-R	Channel 56 Sites With No DTV	DTV Success Rate
50	47	47	0	100%	0	0	2	#DIV/0!

Note: TASO Grading Scale is: 1- Excellent, 2-Fine, 3-Passable, 4-Marginal, 5-Inferior,6-Unusable

AUTHOR AND PRESENTER INDEX

COMPANY INDEX

3891
Price: $80.00 EACH
DESC: BROADCAST ENG. CONF. PROC

NAB™
BROADCASTERS

National Association of Broadcasters
1771 N Street, NW
Washington, DC 20036-2891

Item #3891