DIGITAL AUDIO AND ANCILLARY DATA SERVICES FOR ATV -- THE WORK OF THE ATSC SPECIALIST GROUP

Graham S. Stubbs
Eidak Corporation

ABSTRACT

This paper describes the advice and suggestions put forth by the Technology Group on Distribution (T3) of the Advanced Television System Committee (ATSC) regarding digital services for Advanced Television (ATV). These recommendations were based on the background work of the Specialist Group (T3/S3) on Digital Services which conducted technical studies, and surveys, and developed the suggestions and recommendations.

Rapid advances in multichannel composite digital audio coding technology now make it possible to plan to provide the consumer (e.g. cable subscriber) with an expanded audio experience to match wide screen high definition TV pictures. This paper summarizes the suggestions adopted by ATSC's T3 Group regarding audio and ancillary data services, including the advice that a standard service for ATV should include capacity for a minimum of five audio channels with composite encoding.

Particular emphasis was placed on the need for flexible allocation of data to audio and ancillary data services on an as-needed basis.

INTRODUCTION

The Advanced Television Systems Committee and other organizations, including the FCC Advisory Committee on Advanced Television Service, EIA, and NCTA, have recognized that much of the effort spent on ATV development and standardization has been put on the development of vision signal encoding schemes and that it is important that appropriate emphasis should be placed on defining the accompanying sound channels, the features of the ancillary data and control services, and the way in which they may be included in the ATV transmission format.

The charter of the ATSC Specialist Group on Digital Services (T3/S3) has been to conduct industry surveys and technical studies and to develop technical information and recommendations on the following subjects. The Specialist Group commenced its studies in December, 1990.

A. Sound & Ancillary Data Services

Identification of the range of ATV sound channel requirements and how they might be satisfied with recent rapid advances in the state-of-the art digital audio coding. Identification of the range of desirable ancillary data services--including those already in use and in some cases mandated--and estimates of the data capacity required for each.

B. Conditional Access

Establishment of system attributes and features required to allow conditional access to the ATV transmission in the various media in particular Cable Television, with emphasis on the structure and capacity of the required control data channel.

C. Multiplex Structure

Determination of requirements for flexibility in the data multiplex struc

ture to permit re-allocation of the available data capacity for the various services.

The goal of this work has been to influence the characteristics of the digital services (other than video) which are to be provided by the ATV emission system selected for use for terrestrial broadcasting in the United States. (However, it is not ATSC's intention to affect the ATV testing program already in progress as part of the FCC selection process.) Specifically, it is suggested that the final ATV system selected for terrestrial broadcasting, cable, and other media follow the guidance detailed in reference 1 (see Appendix for Executive Summary of the ATSC document).

Recognizing that different media (broadcast, cable, DBS, etc.) will use different modulation methods, possibly carry different ancillary signals and have other requirements unique to each medium, T3/S3 has focussed its technical studies to facilitate the adoption of voluntary standards for all media based on the same baseband signal representation for video, audio, data and control to maximize interoperability and to provide a simple common interface to the consumer ATV receiver. Work by T3/S3 on Conditional Access is reported in References 2 and 5.

WORK ON AUDIO AND ANCILLARY DATA SERVICES

Advances in audio bit rate coding technology have occurred extremely rapidly in the past two years, now making it possible to encode multiple audio channels using data rates previously allocated to a single audio channel. It is now technically possible to create multiple channel audio "images" which can truly complement the high definition pictures in a restricted 6 MHZ wide HDTV digital channel. Audio techniques developed for motion picture theaters may become applicable to the consumers' living room.

As a discussion vehicle, the Specialist Group prepared a "strawman" proposal for TV Audio and Data Services for Simulcast ATV systems. The "strawman" proposal suggested desirable program audio attributes and listed possible ancillary digital services, including some of those already in use with NTSC such as closed captioning. It was suggested that advantage should be taken of the opportunity to design a new generation of ancillary data services into the system from the beginning. Starting in April 1992, the "strawman" proposal was circulated for comments and suggestions to a wide range of parties with interest in ATV development and introduction, including the proponents of systems to be tested in the selection process. By an iterative process, the document was refined to reflect as much industry consensus as possible, with particular emphasis on flexible allocation of data capacity to audio and ancillary data services.

The following summarizes some of the principal suggestions as to the characteristics of the audio and digital services which should be provided by an ATV emission system adopted for use in the United States, both for terrestrial broadcast and by alternate media. Certain services are identified which the final distributor of programming may provide in a standardized manner. Standardization of services is necessary so that receiver manufacturers may produce receivers that can receive these services. A number of capabilities are

identified which every receiver should be required to provide.

MAJOR PRINCIPLES

Flexible Allocation

The number and type of audio and data services which an ATV system should deliver will vary significantly depending on what services are available to accompany the picture, and the needs of the particular service area for the distribution medium, e.g. broadcast, cable. etc. In order to provide a maximum level of utility, many audio channels and data services might be needed. A fixed provision for many audio channels and data services would require, however, that a significant portion of the available transmission capacity be reserved for these services. This would unnecessarily constrain the video quality, which would suffer if its available data rate were restricted. Since much of the programming may not require many audio channels or data services, a fixed allocation of transmission capacity to these services would be inefficient. Therefore, the ATV transmission system should be capable of a flexible allocation of data to audio and data services on an as-needed basis, with the remaining data available to the video system. At the point of emission or final distribution, a choice may be made based on the available services and the needs of the intended audience as to which services are to be delivered via the final distribution medium.

Extensibility

It is not possible to envision and provide for all potential audio and data services which might be useful components of the ATV service in all media. It is feasible to allow new digital services to be added in a compatible manner by means of the flexible allocation of data capacity. The key is to provide a way to identify new types of allocated data services, so that new or upgraded receivers may make use of the new data types, while older receivers simply ignore them. The new data types could offer new kinds of audio services (perhaps with more audio channels) and new kinds of information services. Some of the additional data types could be intended for private or commercial reception, with the data capacity being sold by the final program distributor to provide a supplemental revenue source.

Multi-Channel Audio

The ATV system should be capable of delivering multi-channel sound appropriate to the wider, higher definition picture. Consistent with demonstrated psychoacoustic principles, the preferred channel assignment is: Left, Center, Right, Left Surround, and Right Surround. A significant advantage of three front channels (rather than two as in present TV stereo) is stabilization of the audio dialogue image in the vicinity of the TV screen.
[Note: This is consistent with the CCIR Task Group 10/1 Draft Recommendation on Multi-Channel Sound (Reference 3), and the SMPTE Film Sound Sub-Committee Report](Reference 4). An optional low bandwidth channel for low-frequency enhancement (subwoofer channel) may also be provided. Monophonic and two-channel stereophonic transmission modes of operation should also be provided, and will offer the most efficient method of delivery for mono or two-channel stereo programming.

Even with the use of advanced low bit-rate audio coding technologies, provision of five high quality independently coded

discrete channels would require significant data capacity. Current estimates of required bit rates per individual audio channel are:

High Quality	128	kbits/sec
Medium Quality	96	kbits/sec
Low Quality	64	kbits/sec

However, recent developments in composite multi-channel audio coding technology show that five-channel audio may be delivered with a data rate only slightly larger than that required by two high quality independently coded channels. It is recommended that a five-channel composite coding mode structure be predefined as part of the ATV system.

Suggested are three composite coding modes, which offer different numbers of high quality audio channels. The numerical designations below (e.g. 3/2) indicate the number of front channels / rear channels. Some audio coding technologies incorporate a low bandwidth channel (<200 Hz) intended to deliver low-frequency enhancement information (subwoofer channel), the use of which would be optional for the program distributor, receiver, and viewer.

Five channels	3/2	300 - 400	kbits/sec
Three channels	3/0	256	kbits/sec
Two channels	2/0	192	kbits/sec

With composite coding, all of the audio channels which have been coded into the composite data stream must be reproduced together (similar to the way the R,G,B colors must be reproduced together to form the viewed picture), although not necessarily out of independent loudspeakers. The channels may be mixed together to reduce the number of loudspeakers required for reproduction.

When surround information is relevant and available, the 3/2 composite coding mode is preferred. Three front channels and two surround channels are provided.

When surround information is not available, the 3/0 mode is sufficient. Three front channels are provided; the 3/2 mode may also be used.

Two-channel stereophonic audio may be most efficiently transmitted with the 2/0 mode. The 3/2 or the 3/0 modes may also be used.

It is preferable that audio programs not be simultaneously provided by a monophonic or two-channel stereophonic service intended for low cost receivers, in addition to a separate multi-channel service intended for high end receivers. That would be wasteful in data capacity, and would inhibit the use and growth of multi-channel audio. It would be necessary, though, that all receivers be capable of decoding a multichannel service into the desired number of reproduction channels. Lower cost receivers would not need to completely decode all five channels, but could decode the five channel service into a conventional monophonic or two-channel stereophonic program with attendant cost savings.

The ATV receiver would be required to produce sound for any of the pre-defined coding modes that the broadcaster chooses to use. This doe not imply that every receiver should be capable of fully decoding a 3/2 service; only that it make sound from it. It would be acceptable for a receiver to decode all modes into monophonic or two-channel stereophonic audio.

Uniform Loudness

The ATV audio system should provide the means to control loudness in a uniform manner among various programs and delivery channels. The viewer should perceive the same subjective dialogue loudness when a program ends and a new program begins, or when channels are changed. The ATV audio data should inform the receiver of a dialogue reference level so that the receiver can reproduce the normal spoken dialogue at an acoustic level chosen by the viewer.

Dynamic Range Control

There is a conflict between the needs of many viewers for program audio to have a narrow dynamic range, and the desires of some viewers to reproduce the audio in the full dynamic range intended by the program producer. The ATV audio coding system should incorporate an integrated dynamic range compression method which delivers data to the receiver representing the compression characteristic employed.

Receivers may include circuitry which gives the viewer the ability to control the reproduced dynamic range. Using the data representing the compression characteristic employed, the receiver may partially or completely reverse the dynamic range compression intentionally introduced by the program provider.

Error Correction and Concealment for Audio Services

It is recognized that error correction and concealment is a more critical issue for an audio service than for the video service, as reproduced audio errors may be more annoying than reproduced video /errors. In addition, some types of delivery (terrestrial, DBS, etc.) will be used by some viewers at the threshold of the service area. Therefore, the ATV audio system should include effective concealment methods to minimize audible disturbances caused by uncorrectable errors.

Audio Services to the Visually and Hearing Impaired

The ATV audio system should allow programmers to deliver special audio services to the visually impaired (VI) and hearing impaired (HI) using the flexibly allocable channels. The VI service would typically contain a narration describing the picture content, and would be reproduced along with the main audio program in the receiver. The HI service would typically contain only dialogue, and would be processed for improved intelligibility.

SERVICE IDENTIFICATION DATA (SVID)

Service Identification Data (SVID) should be incorporated into the ATV data stream. The SVID supplies descriptors for all digital services delivered by the ATV signal. The descriptors identify the digital services and indicate their locations within the overall data multiplex. Several methods may be used to deliver this information (such as packets with headers) providing they meet the requirements for this function. This issue has not been studied in depth and no recommendation is made as to the technique to be employed.

The SVID must be very reliable, since errors could cause total loss of ATV

service. The SVID should be recognizable by the receiver as soon as possible after a channel change. The information carried by the SVID should be repeated frequently, so that a receiver can quickly recognize the available services after a channel change, and so that the redundancy can be used to improve reliability of the data. The format that is adopted for the SVID must be very flexible and allow new digital services to be defined and delivered in a manner that is compatible with all ATV receivers.

ANCILLARY DATA SERVICES

Several types of data services should be pre-defined. Others may be added using the flexible allocation capability. Some types of data services, such as conditional access, may only be used by the alternative media.

Captioning

The ATV emission system should provide a captioning system capable of delivering multiple versions of captions. Different versions of captions may be used to provide service in multiple languages, for the hearing impaired as well as the non-impaired. A minimum requirement on the captioning system should be that it allows three versions of captions: primary language, secondary language, and primary language for slow readers. A data rate sufficient to support this level of service should be allocable to the captioning service when captioning information is available. It should be possible to allocate additional data to the captioning service to support delivery of additional versions of captions (or subtitles). All ATV receivers should be required to decode and display captions.

Program Guides

The program guide is an optional service. The program guide is intended to inform the viewer about the programming available on the particular ATV channel being viewed. The program guide service should be capable of delivering text accompanied by graphics. For the benefit of the viewer who is quickly scanning channels, lines of text identifying the current and upcoming programs should be provided and the program guide data should the found quickly in order to minimize the delay in displaying the current program information.

If an additional frequency agile tuner is available, either in the ATV receiver or at a cable headend, all available ATV channels may be scanned and the program guide information for all available channels combined to form a multichannel program guide.

Conditional Access

Some applications of the ATV system will require data to be allocated to conditional access control. The required data capacity will depend on the size of the audience being served and the type of service being offered (i.e. pay-per-view or monthly subscription). There should be no fixed allocation of data to conditional access. Data capacity should be allocable to conditional access control at the option of the particular service provider. The T3/S3 Specialist Group made no recommendation as to the minimum data rate required to support a conditional access system.

(NOTE: Meetings of the ATSC Specialist Group on Digital Services T3/S3 were held concurrently with meetings of the ATSC Specialist Group on Interoperability and

Consumer Product Interface T3/S2. The two Specialist Groups worked closely on Conditional Access issues. This work on desirable attributes and features of encryption systems is reported in a companion paper, Reference 5.)

CONCLUSIONS

The "strawman" approach to soliciting industry opinion and advice regarding digital audio and ancillary data services for ATV resulted in responses from a broad range of parties with ATV interests. The results of the T3/S3 studies indicate that an acceptable set of audio and ancillary services can only be provided by a flexible, extensible system which allocates data capacity to these services only on an as-needed basis. All remaining data capacity should be used by the video system for improved picture quality.

Advantage should be taken of the opportunity to design into a new ATV standard the ability to distribute multi-channel audio with qualities commensurate with wide screen high definition pictures. At the same time, a flexible approach to data allocation will provide new ancillary data services and should satisfy the conditional access needs of Cable and other media.

Acknowledgements

The author wishes to acknowledge that portions of this paper are extracted from "Recommendation on Digital Audio and Ancillary Data Services for an Advanced Television Service", distributed by ATSC in February 1992. The author is grateful to Craig Todd (of Dolby Laboratories) and Tom Keller (Consultant to Cable Labs) for their considerable contributions to the generation and authoring of the ATSC document. Appreciation is also expressed to the many individuals who participated in the Specialist Group's work, and who responded to the numerous requests for advice and comment.

References

1. "Digital Audio and Ancillary Data Services for an Advanced Television Service" February 3, 1992 (Doc. T3/186) Advanced Television System Committee, 1776 K Street, NW #300, Washington, DC 20006.

2. "ATV Encryption System Characteristics" May 16, 1991 (Doc. T3/180) Advanced Television System Committee, 1776 K Street, NW #300, Washington, DC 20006.

3. CCIR Task Group 10/1 Draft Recommendation on Multi-Channel Sound, November 7, 1992.

4. SMPTE Film Sound Sub-Committee Report.

5. "A Progress Report on the Work of the ATSC Specialist Group on Interoperability and Consumer Product Interface". Bernard Lechner, NCTA Convention, May, 1992.

Appendix

Executive Summary of "Digital Audio and Ancillary Data Services for an Advanced Television Network", ATSC, February 3, 1992. (T3/186)

The Specialist Group on Digital Services (T3/S3) of the Technology Group on Distribution (T3) of the Advanced Television Systems Committee (ATSC) has been working for two years to identify the digital audio and data services that should accompany the Advanced Television (ATV) picture. T3/S3 has conducted technical studies and industry surveys, and has circulated a "strawman" proposal to stimulate industry discussion and feedback. Based on this work, T3 offers the guidance contained in this document. On February 3, 1992 the ATSC Executive Committee approved release of this material. It is hoped that the ATV system selected for use in the United States for terrestrial broadcasting, as well as ATV systems for the alternative media, will follow this guidance.

A major finding is that it is not desirable to select a fixed set (number and type) of digital audio and data services for inclusion into the 6 MHz ATV channel. This is because the data rate required for all potential services would negatively impact the picture data rate and affect the picture quality. It is recommended that the ATV system allow data to be allocated to digital audio and data services only on an as-needed basis. A flexible system of data allocation will require the use of only the minimum data capacity necessary for digital audio and data services at any time. Flexible allocation also allows the addition of new types of digital services in the future, with older receivers ignoring the new data types.

The ATV service will offer widescreen pictures, and this feature is expected to increase consumer interest and enjoyment of this new format. The audio corollary to the wider aspect ratio is multi-channel audio. Recognizing the significant limitations of two-channel stereophony for sound accompanied by pictures, it is recommended that the ATV service be capable of delivering five channel audio (left, center, right, left surround, and right surround). This is generally consistent with recent trends in the application of digital audio in the motion picture industry, and the draft recommendation from CCIR Task Group 10/1. Recent advances in multi-channel audio coding technology have reduced the data rate required for five channel audio nearly to that required for two independent audio channels. All ATV receivers would need to decode the provided service (which could vary from one to five channels) into the number of loudspeaker channels to be used (e.g. the five channel audio service may be decoded into mono for the low-cost mono receiver).

The ATV audio system should provide a solution to the problem of loudness uniformity among programs, channels, and delivery media. The average perceived loudness of dialogue should be uniform. The coded audio data should indicate the average dialogue loudness within the dynamic range of the coding system. Different programs may have varying amounts of headroom above this level which is available for dramatic effect.

The ATV audio system should allow audio service to be provided with a wide dynamic range. Since experience indicates much of the audience will prefer a restricted dynamic range, the audio coding system should incorporate an integrated dynamic

range compression system. Information about the compression introduced may be incorporated into the audio data stream so that the receiver may optionally reverse the compression to restore the original dynamic range of the program.

The ATV system should allow service to be delivered to the visually and hearing impaired. The visually impaired may be served by an optionally allocated narrative audio channel. The hearing impaired may be served by the captioning system, or by an optionally allocated audio channel containing only dialogue which has been processed for improved intelligibility.

Flexible allocation of data capacity to audio allows audio service to be optionally provided to specialized audiences, including those with other languages. Audio services should be tagged to indicate language and type, so that the receiver may assume the burden of choosing the correct language audio service for the viewer (assuming the receiver has knowledge of the viewers preferred language), and so that the viewer may determine the types of audio available.

* * * * * * * * * * * * * * * * *