PS-WP/4 has identified a number of characteristics that contribute significantly to Interoperability, Scope of Services and Features, and Extensibility. These are based on needs and desires exhibited by alternative media advocates, not only for the delivery of terrestrial broadcast television programming but also for other delivery approaches and applications relating to computing, communications, motion pictures and imaging. In relative order of importance, these characteristics are:

- An all-digital implementation based on a layered architecture model
- The use of universal headers and descriptors (as agreed by industry standards group, for example, SMPTE)
- Transmission of the signal in progressive scan format
- Use of a flexible, packet data transport structure
- Viewer transparent channel re-allocation (limited picture and sound while most of the channel capacity is devoted to data transmission for conditional access addressing or other purposes)
- Ability to implement lower-performance, low-cost ATV receivers (comparable price/performance options to current NTSC receivers)
- Ability to implement low-cost ATV consumer VCR
- System architecture and implementation that will allow improvements and extensions to be incorporated as technology advances while maintaining backward compatibility
- Square pixels or at least the option to select square pixel presentation
- Compatibility with relevant international standards or commitment to this objective
- Easily-implementable and user-accessible "still/motion multi-window transmission"

The PS-WP/4 assessment and evaluation of the proponent systems shows some significant differences under the three criteria. Further, all proponent systems need improvement on one or more of the listed characteristics to achieve a desirable degree of interoperability, extensibility and scope of services and features. It is recommended that the Special Panel, the Advisory Committee and the FCC take these differences into account in the process of selecting an ATV standard. Furthermore, selection of a system that incorporates interoperability features not included in the system as submitted for testing requires verification and/or testing. The system submitted for field testing should also include such features.

BACKGROUND

The objective of Planning Subcommittee Working Party 4 (PS-WP/4) was to study and make recommendations regarding the relationship of terrestrial advanced television systems to alternative media, applications and standards. It was also the objective to investigate approaches for growth paths to the future while, at the same time, supporting timely decisions on an advanced television (ATV) broadcast system with increased performance quality for the end user. Participants of Working Party 4 have addressed issues related to interoperability, scalability and extensibility and more generally, openness. Representatives of the broadcast television, cable television, program production, motion picture, computer, telecommunications, and imaging industries were active in this working party.

In the prior year's effort (1991), PS-WP/4 developed definitions of key terms such as interoperability, scalability and extensibility. Based upon a world becoming more complex and richer in alternatives (media, transmission/distribution, presentations), the working party developed the concept of image data, defined as the digital equivalent of the video information including image, sound and auxiliary data components. As a result, PS-WP/4 recommended the following in its December 1991 Interim Report:

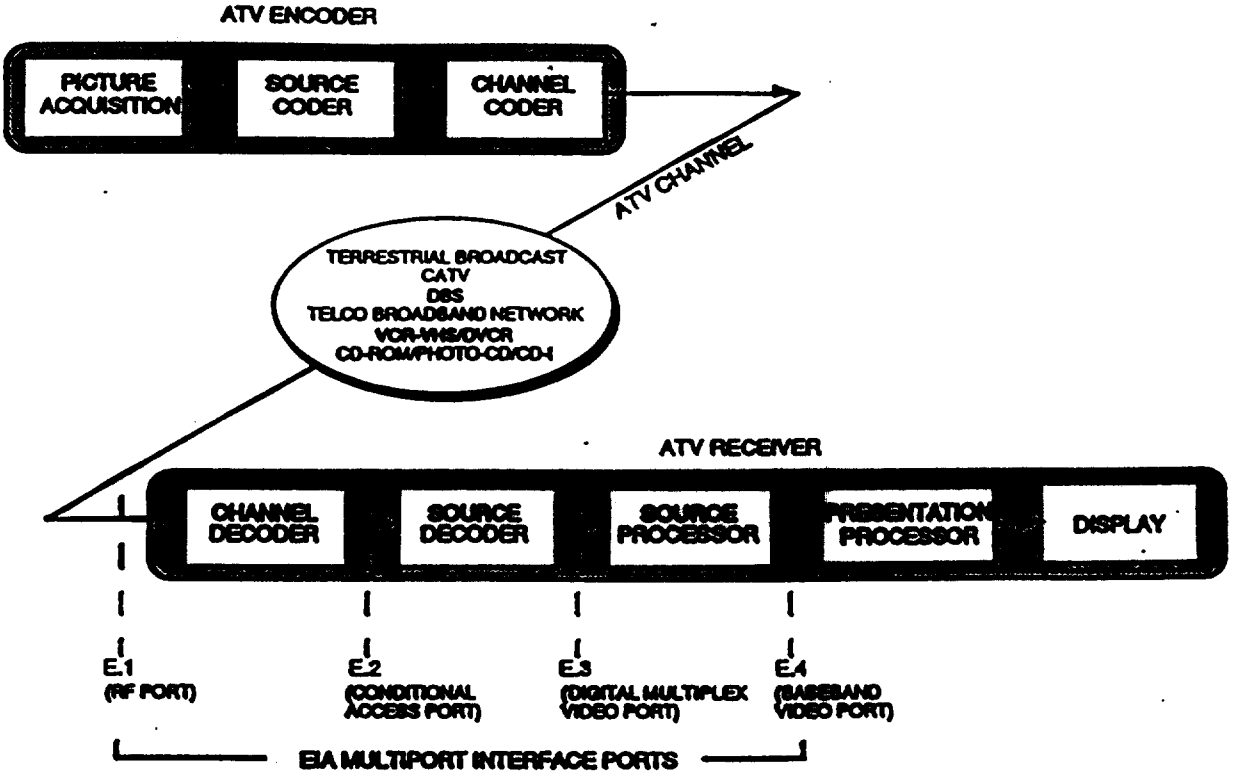
- Maximize utilization of digital video techniques and image data representation.
- Apply HEADERS and DESCRIPTORS (as agreed by industry standards groups) as a method of identifying image data.

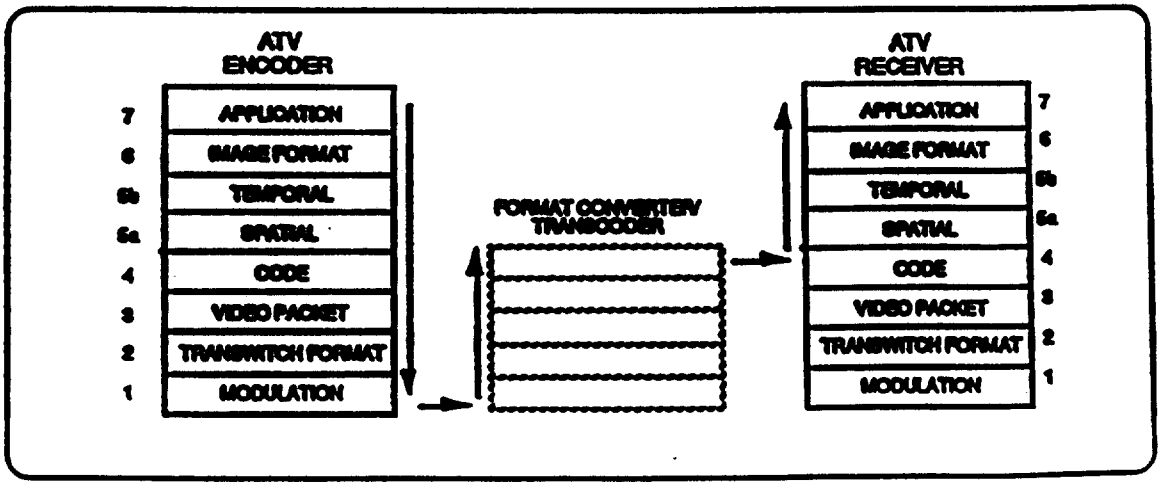
In 1992, PS-WP/4 developed an ISO-like layered architectural model for ATV, consisting of seven layers as shown below:

7	Application
6	Picture Format
5b	Temporal
5a	Spatial
4	Code
3	Video Packet
2	Transwitch Format
1	Modulation

The application of this architectural model in the working party's analyses is shown in the Digital Video Reference Model For Interoperability in the following figure.

ATV REFERENCE MODEL FOR INTEROPERABILITY





Richard Lau
Bellcore
7/9/92
Revised: 11/24/92

PS-WP/4 examined relationships of terrestrial advanced television systems to alternative media, applications and standards in the broad context of recognized and anticipated advances in computing, communications and imaging technology. Weighting the value (e.g., market and social benefit) of interoperability or extensibility other than in established markets (e.g. terrestrial broadcast or cable television programming) was difficult and arguable. Formal positions from affected industries were not available.

Throughout PS-WP/4's investigation, few disagreed that the U.S. FCC ATV decision should endure over some significant, foreseeable technology horizon which should be at least several decades.

Furthermore, few disagreed that over this technology horizon, concepts that are today limited by cost considerations for the consumer market will become reality in the future. These concepts include, at the least, progressive scan (for capture and presentation) and motion image data delivery over packetized wideband networks.

PS-WP/4 believes that the FCC ATV decision must be considered as defining a starting point and assuring a migration path to the future, rather than as defining an end point.

Assessment of Proposed ATV Systems

PS-WP/4 was assigned the task of assessing the interoperability of the proponent systems with media other than terrestrial broadcasting. In addition, the systems were to be evaluated with regard to extensibility and scope of services and features. A preliminary assessment was written in early August by StellaCom, Inc. based on published information available at the time. This report included interoperability questions that were to be answered by the proponents in a planned Interoperability Review in September 1992.

Some 25 generic questions on interoperability, extensibility, and scope of services and features were compiled for all proponents. An additional seven to ten system-specific questions were also addressed to each proponent. Interoperability assessments were made with regard to cable TV, digital technology, headers/descriptors, NTSC, film, computers, satellites, packet networks, interactive systems, format conversion and scalability. The initial use for ancillary data, audio, data, text, captioning, encryption, addressing, capability for a very low-cost receiver with lower performance and VCR capability were assessed as services and features. Extensibility to no visible artifacts regardless of the detail and motion in a scene, to studio quality (editable), to higher resolution such as VHDTV and UHDTV and provision for future compression enhancements were examined.

The proponents answered all questions in writing at the Interoperability Review on September 23-25. They were given additional time to submit any proposed interoperability improvements to their systems. A revised interoperability assessment was then generated including the updated material. The proponents were given the opportunity to critique the

assessment. The assessments were then revised and abbreviated to five to six pages for each proponent to be submitted to the ACATS SS-WP/4 Special Panel. The result is attached as Reference 1.

Interoperability Review

An Interoperability Review was held with the ATV proponents to focus on interoperability, extensibility and scope of services and features. Two-hour presentations were made by each proponent followed by 1/2 hour of questions from the Review Board of multi-industry experts and 1/2 hour by other PS-WP/4 members.

The Review Board members were:

Jules Bellisio	Bellcore
Jules Cohen	Consultant/MSTV
Gary Demos	Demographx
Jack Fuhrer	Hitachi America
Branko Gerovac	DEC
Jukka Hamalainen	Matsushita Applied Research Lab
George Hanover	Electronic Industries Association
Robert Hopkins	ATSC
Ed Horowitz	Viacom
Mike Liebhold	Apple Computer
Craig Tanner	CableLabs
Arpad Toth	Eastman Kodak
Tony Uyttendaele	CapCities/ABC

The proponent presentations in the Interoperability Review gave strong consideration of interoperability, extensibility and scope of services and features. The proponent ATV system descriptions contained evidence of effort toward enabling interoperability with alternative media. The written material distributed at the meeting is found in Reference 3: Minutes of the PS-WP/4 Interoperability Review, September 23-25, 1992.

Additional information was submitted by some proponents on October 1, 1992 and can be found in Reference 4: Minutes of the PS-WP/4 meeting, November 5, 1992. Evaluations from the Interoperability Review were submitted to StellaCom, Inc. by members of the Review Board for compilation and analysis. As much as possible, the evaluations were based on "what the evaluator expects the proponents will deliver commercially". The proponents were also invited to evaluate the proposed systems. The results from these evaluations are attached as Reference 2. It contains summaries of the ratings by the proponents, the Review Board and the ratings by individual experts.

Conclusions

Based on assessments of the five proposed ATV systems and the Interoperability Review conducted on September 23-25, 1992, PS-WP/4 reached the following conclusions with respect to the selection criteria of Interoperability, Scope of Services and Features, and Extensibility. See Reference 2 for numerical ratings and detail.

Overall

During the assessment and review process, all proponents gave strong consideration to meeting the three selection criteria and presented evidence to show how their systems met or could meet the criteria.

The four all-digital systems (AD-HDTV, CCDC, DigiCipher and DSC-HDTV) were found to be superior to the Narrow-MUSE system with respect to Interoperability and Extensibility on all but two or three characteristics where Narrow-MUSE was approximately equal to the all-digital systems. With respect to Scope of Services and Features, the all-digital systems also ranked better than Narrow-MUSE but the difference was smaller.

Interoperability

- All five systems were judged to be easily deliverable by satellite or Cable TV.
- The four all-digital systems were judged to interoperate well with NTSC.
- The four all-digital systems were judged to interoperate well with film with CCDC ranking better than the other three systems.
- Interoperability with computers was judged to be more difficult than with satellites, Cable TV, NTSC and film. DSC-HDTV ranked slightly better than CCDC, with AD-HDTV next, and DigiCipher ranked somewhat worse.
- Progressive scan and square pixels are important for computer and other image applications (rotation, scaling, still-image, LCD displays, ...). Computer industry applications desire greater than 70 Hz display rates although no proponent system offers this feature.
- All of the proponents of all-digital systems now recognize that the use of headers and descriptors is a critical enabling concept for assuring ATV flexibility in the future. Although these four proponents propose to include headers and descriptors in their systems, only AD-HDTV had its final proposal for headers and descriptors fully implemented at the time the system was tested by ATTC, and this system received the highest rating on this characteristic. There is need for definitive work on an industry standard for headers/descriptors. The SMPTE work is a candidate.

System Specific Comments

During the assessment and review conducted by PS-WP/4, the proponents described how they proposed to meet the requirements of the three selection criteria against which they were being evaluated. In some cases, these proposals require changes to the system hardware and software from that which was tested at ATTC. The effects these changes could have on the results of the ATTC and ATEL evaluations of the systems are unknown. In other cases, certain features or characteristics were present in the system when it was tested by ATTC, but were not exercised during the tests. Where changes are required, it must be verified that these changes will not affect the results obtained by ATTC and ATEL. Where the features were not exercised, it must be verified that they work when exercised.

All proponent systems need some improvement in interoperability, scalability and extensibility. For example:

AD-HDTV

- verification of the progressive-scan and square-pixel migration strategy

CCDC

- implementation of universal header/descriptor concept
- implementation of packetized data structure

DigiCipher

- implementation of universal header/descriptor concept
- development & verification of a progressive-scan and square-pixel migration strategy
- implementation of packetized data structure

DSC-HDTV

- implementation of universal header/descriptor concept
- complete the implementation of a fully packetized data structure

Narrow-MUSE

- commitment to digital implementation
- implementation of universal header/descriptor concept
- implementation of packetized data structure
- development & verification of a progressive-scan and square-pixel migration strategy

RECOMMENDATIONS

Based upon the needs and desires exhibited by the alternative media, not only for the delivery of terrestrial broadcast television programming but also for other delivery approaches and applications relating to computing, communications, motion pictures and imaging, PS-WP/4 has identified a number of characteristics that contribute significantly to Interoperability, Scope of Services and Features, and Extensibility. In relative order of Importance, these characteristics are:

- An all-digital implementation based on a layered architecture model
- The use of universal headers and descriptors (reference ACATS Interim Report, March 1992)
- Transmission of the signal in progressive scan format
- Use of a flexible, packet data transport structure
- Viewer-transparent channel re-allocation (limited picture and sound while most of channel capacity devoted to data transmission for conditional access addressing or other purposes)
- Ability to implement lower-performance, low-cost ATV receivers (comparable price/performance options to current NTSC receivers)
- Ability to implement low-cost ATV consumer VCR
- System architecture and implementation that will allow improvements and extensions to be incorporated as technology advances while maintaining backward compatibility
- Square pixels or at least the option to select square pixel presentation
- Compatibility with relevant international standards or commitment to this objective
- Easily-implementable and user-accessible "still/motion multi-window transmission"

The PS-WP/4 assessment and evaluation of the proponent systems shows some significant differences under the three criteria. Further, all proponent systems need improvement on one or more of the listed characteristics to achieve a desirable degree of interoperability, extensibility and scope of services and features. It is recommended that the Special Panel, the Advisory Committee and the FCC take these differences into account in the process of selecting an ATV standard. Furthermore, selection of a system that incorporates interoperability features not included in the system as submitted for testing requires verification and/or testing. The system submitted for field testing should also include such features.

Reference 1

INTEROPERABILITY ASSESSMENTS

December 11, 1992

StellaCom, Inc.
Advanced Television Systems
1901 N. Moore Street - Suite 601
Rosslyn, VA 22209-1706

(703) 243-1695 (p)
(703) 243-1698 (f)

This work was done for ACATS PS/WP-4.

INTRODUCTION

This document was written by StellaCom, Inc. for PS-WP/4, a working party of the Advisory Committee on Advanced Television Systems (ACATS) which operates on behalf of the Federal Communications Commission. The information contained in this document was obtained from documents submitted by the proponents to PS-WP/4 and to other working parties of ACATS, and from oral presentations by the proponents in September, 1992 in which questions from PS-WP/4 were answered.

It was the intention of PS-WP/4 for this document to specify whether each feature discussed was

1) in the system tested at ATTC,

2) promised for the system to be field tested,

3) promised for the system after field testing, or

4) possible/realistic at a future time.

This was done to the extent possible with the information available. However, the proponents are continuing to enhance their systems and revise their time estimates for the implementation of features. Some of the enhancements under way now are in response to the concerns of PS-WP/4 on Interoperability, Scope of Services and Features, and Extensibility.

Financial support was provided by Apple Computer, Inc., Bellcore, Cable Television Laboratories, Inc., Eastman Kodak Company, Digital Equipment Corporation, and Viacom International, Inc.

Advanced Digital HDTV Interoperability Assessment

System Description

The Advanced Digital HDTV (AD-HDTV) System from the Advanced Television Research Consortium is an all-digital system with 1050 scan lines per frame and 29.97 frames per second. Compression and transmission are in a 29.97 hz progressive form. The system tested used a 2:1 interlaced source and a 2:1 interlaced display. The display format of the prototype system has a 16:9 image aspect ratio with 960 lines per frame and 1248 pixels per line, although future versions are expected to use 1440 pixels per line. The proponent also claims that the system can be changed to have square pixels. Features of the system include a prioritized data transport format that separately packages video, audio and data and allows their mix to vary dynamically. It also includes two-tier prioritization where each data segment is assigned to a high-priority (HP) data stream or a standard priority (SP) data stream. Transmission is by Spectrally Shaped QAM (SS-QAM) consisting of two 4- or 5-bit-per-symbol QAM signals fitted between co-channel NTSC carriers. Each data stream modulates a 16- or 32-QAM carrier, with the high-priority carrier transmitted at higher power. This allows a degree of graceful degradation on the edge of the broadcast contour and transmission robustness within the service area, while minimizing interference to and from co-channel NTSC. It also allows a degree of bit-stream scalability.

I. Interoperability

A. with Cable TV

The tests of the proponent systems at ATTC by Cable Labs should determine the critical aspects of interoperability with cable TV. Other factors that may aid cable TV interoperability include channel augmentation to 9 or 12 MHz for improved picture, or better, going to a higher data rate that can be supported in the 6 MHz channel by the relatively benign cable environment. The question remains whether HDTV receivers can be built to support higher numbers of bits per symbol without cost impact.

Encryption and addressing are important service features for Cable TV. See Section II, items F and G.

B. with digital technology

Since this system is all-digital, the advantages of all-digital systems apply.

C. with headers/descriptors

The AD-HDTV system formats all its data in a layered manner following the MPEG syntax. After synchronization, the data link layer identifies the service type. Within the adaptation level, the adaptation header contains information governing the packing of variable-length code words for video, and information used in recovery after channel changes or errors. The video service level includes the actual encoded video information.

All data sent by the AD-HDTV system is grouped into fixed-length cells that contain data of a single particular type. The cells are 148 bytes long, starting with a single byte for synchronization, an 8-bit Barker sequence {00011101}. This is followed by a single byte for a service header that defines cell

types at the link layer. Next is a 4-byte adaptation level header that contains pointers to the start of a slice within the picture frame, as well as the identity of the slice within the frame. These are followed by a 120-byte information payload containing actual data for the cell's particular service type. At the end of the cell is a 2-byte frame check sequence (FCS) trailer, and a 20-byte forward-error-correction (FEC) trailer.

The adaptation layer pointers and slice identification information provide re-entry points within the codec video data, making it possible to begin decoding at a known point after an error event or channel change that requires a restart for some or all of the video decoding.

D. with NTSC

The proponent selected the field rate of 59.94 Hz for compatibility with NTSC. The number of active video lines, 960, was selected to be double the number of active NTSC lines, 480. (NTSC normally has 483 active video lines, but 483 is an awkward number for conversions.) Since the aspect ratio of this system, 16:9, is different from the aspect ratio of NTSC, 4:3, a choice must be made for the conversion. Two choices for down-conversion are 1) "Edge Crop", in which the HDTV picture fills 480 lines on NTSC with loss of the sides of the HDTV picture, and 2) "Letterbox", in which the full width of the HDTV picture is displayed in the full width of the NTSC picture, but leaving unused areas at the top and the bottom of the NTSC screen. Pixel values from HDTV lines are stored in memory and read out at reduced speed to make NTSC lines. The range of pixels read and the clocking rate depend upon the choice of edge-crop or letterbox. The down-conversion involves interpolation between HDTV pixels in a line and between HDTV lines.

E. with film

It is claimed that AD-HDTV will support an "electronic film" format that makes use of the redundant field to improve quality. Since film has a slower (24 frames per second) temporal rate, AD-HDTV will scan its 1050-line raster progressively for film at 24 frames per second, with the same 1440 x 960 pixel format that it uses with video sources. However, in film productions where computer graphics are used extensively, square pixels may be more desirable in the image representation. It is claimed that AD-HDTV will also provide a progressively scanned 1440 x 810 square pixel format to accommodate this source. Most receivers will perform 3:2 pull-down to convert to their 59.92 Hz field rate, but higher-cost receivers could use 3:1 frame repeat to display at 72 Hz.

F. with computers

Progressive scanning and square pixels are important factors for interoperability of an HDTV system with computers -- nearly all existing bit-mapped computer graphics displays have these features. Progressive scanning and square pixels are most critical for real-time applications such as display, scan-conversion, frame capture, and video effects. Computers are expected to play an increasing role in video image generation and production and it is desirable to have an HDTV format which allows direct display and manipulation of HDTV video on the computer. Progressive scanning is preferable for computer applications to avoid artifacts that are common with interlaced display of computer generated imagery. Square pixels are inherent to graphics display hardware for all popular computers employing bit-mapped displays. The main reason is that it facilitates processing of 2D transformations, especially rotations.

Encoding and transmission in AD-HDTV are done in progressive form, favoring interoperability with computers, although testing of this system has been done with interlaced sources and displays requiring format conversions. ATRC has suggested that the system will eventually use progressive sources and displays. With 1440 pixels/line, the pixels in AD-HDTV are 18.5% wider than high, and

with 1248 pixels/line they are 37% wider than high. ATRC has suggested that their system can provide square pixels by reduction of the number of active scan lines to 810.

This system also offers the possibility of interoperability with MPEG-based computer applications. See Section J.3 below.

G. with satellites

Digital signals with bit rates used by the proponents for terrestrial transmission can generally be accommodated in the available bandwidth of a satellite transponder. The total data rate for the AD-HDTV system in the mode tested by ATTC is 24.0 Mbps. The proponent has suggested removing the 0.9-rate trellis code used with SS-QAM, making the net data rate 21.6 Mbps. The proponent does not anticipate the need for any additional error correction for satellite transmission, although convolutional coding is normally used. The proponent stated that three AD-HDTV channels may be carried in a transponder. However, it is unlikely that more than two AD-HDTV channels will be carried in a typical 36-MHz transponder. See the link analysis done for the DSC-HDTV signal. The proponents also point out that it is possible to carry AD-HDTV and NTSC signals on the same transponder.

H. with packet networks

The data link sublayer format is based on a "cell relay" asynchronous time-division multiplexing concept similar to the asynchronous transfer mode (ATM) standard that will be used for broadband integrated services digital network (B-ISDN). The data-link header contains information such as priority indicator, service ID and cell sequence number. This provides service-independent transport services such as priority support, service multiplexing, and cell-error detection and correction. For the received bit-stream, the transport decoder performs Reed-Solomon decoding and a cyclic redundancy check (CRC) for error detection. Cells received in error after correction are discarded by the demultiplexer. Each cell contains 148 bytes.

The proponent claims that minor picture artifacts may be noticed at a packet error rate of 10^{-3}. However, AD-HDTV will make viewable pictures and perfect sound even if only its High Priority cells are received. This is effectively an error rate of 8 out of 10 packets lost. Out-of-order packets can be put back in order at the receiver.

I. with interactive systems

The total rate control buffer delay is nominally 250 msec for both the high priority and standard priority buffers. In addition, there may be image build-up after a switch or a scene cut depending on the place in the 9-frame Group-of-Pictures (GOP) cycle. Total acquisition time has been designed to be limited to a worst case value of about 1 sec. This number is based on 0.5 sec for the modem and equalizer and 0.5 sec for the video/transport acquisition.

Latency is the time delay between a video frame going into the encoder and the corresponding frame coming out of the decoder in the back-to-back mode. It can be important in interactive applications. According to the proponent, a full-quality AD-HDTV encoder requires 4 frames of latency, due to its use of bi-directional motion compensation with an MPEG GOP structure of M=3, N=9. An additional frame is needed for interlaced-to-progressive conversion. Similar delays are present at the receiver. Rate buffers at the transmitter and at the receiver cause additional delays. Furthermore, the system implemented for test may not be representative of the minimum delay possible. Comparison of latency between the proponent systems is difficult at this time.

This proponent claims that for interactive applications where latency is a concern, an encoder can produce an MPEG bit stream using only forward motion compensation to reduce the coding part of the latency.

J. Format conversion

1. with 1125/60

Up-converting to the Common Image Format (1920 x 1080) requires a 8:9 vertical interpolation and a 3:4 horizontal interpolation from 1440 x 960. SMPTE 240M currently uses 1035 active lines. Colorimetry used by AD-HDTV is intended to be consistent with SMPTE 240M.

2. with 1250/50

Conversion between systems with different frame rates is the most difficult type of conversion presently being done. Digital conversion between 59.94 fields per second and 50 fields per second requires a number of frame stores and very large processing capability. Simple methods that involve frame dropping lead to jerky motion, but other techniques produce acceptable images under most conditions. This difficult conversion is not simplified by the fact that both the source system and the target system are interlaced 2:1.

3. with MPEG

AD-HDTV's use of MPEG-based video and audio compression provides the possibility of interoperability with MPEG-based computer multimedia applications directly in the compressed bit-stream format. The underlying video compression algorithm adheres to the MPEG1 standard from ISO, in that parameters allowable within the MPEG1 definition are used although they are not the MPEG1 default parameters. Prior to entering the prioritization and transport processors, the compressed video representation produced at the encoder fully conforms to the MPEG1 specification. An MPEG bit-stream can be obtained from the output of the compression encoder at the interface to the priority processor. Note that the AD-HDTV encoder prototype uses an internal fixed-length representation for MPEG code words at the interface between its compression and prioritization stages, so that a standard MPEG1 bit stream is not currently available as an output from the prototype hardware.

4. with still image

AD-HDTV's MPEG1 compression, based on the DCT, is generally compatible with JPEG, the ISO standard for still-image compression.

Photo CD uses a proprietary subband coding approach for compression. Picture interchange with AD-HDTV requires full decompression and recompression.

CDI uses the MPEG1 compression syntax for motion video. Therefore, AD-HDTV will be directly compatible with CDI.

K. Scalability

Easy scalability is claimed by the proponent. The MPEG1 bit-stream produced by the encoder is organized into several layers. The ISO-MPEG standard not only specifies the form of the compressed video bit-stream, but it incorporates several higher layers for specifying video parameters. In particular, the "sequence header" provides a mechanism for dynamic specification of the picture size (in terms of pixels x lines), the pixel aspect ratio and the frame rate. With this mechanism, AD-HDTV supports dynamic reconfiguration among its nominal and optional spatial resolutions and raster formats (e.g., interlaced HDTV and progressively scanned film modes). These capabilities are not implemented in the AD-HDTV prototype hardware tested at the ATTC.

The picture produced by AD-HDTV's HP signal alone is a reduced-quality image. The proponent has announced his intention to reprioritize the data partition to make the HP-only picture useful as a

digital hierarchy. The present form is considered by the proponent as sufficient to carry the viewer through short SP channel outages, and as a base for error concealment for high SP error rates. The decoded artifacts observed in an HP-only reconstruction will depend on the exact priority processing algorithm. Typical priority processor operation results in a lower resolution and slightly jerky motion rendition since the B-frame data is carried primarily in the SP channel.

For picture-in-picture and picture-out-of-picture, only the reprioritized HP signal with its 4.8 Mbps data steam needs to be processed, with savings in hardware over full AD-HDTV processing.

For multiple programs in a single channel, AD-HDTV's Prioritized Data Transport layer provides for asynchronous delivery of multiple service types. Multiple video streams can be assigned individual service types.

II. Scope of services and features

A. Initial Use of Ancillary Data

Because of AD-HDTV's asynchronous data multiplexing, there is no hard partitioning of ancillary data needed. Unassigned service types provide for the delivery of many types of ancillary data.

B. Audio

AD-HDTV offers flexibility in the mix of audio, video and data and therefore the number of audio channels. The audio system chosen for AD-HDTV uses a coding algorithm called MUSICAM, which is related to standards endorsed by the ISO in conjunction with MPEG. The audio coding accepts 16-bit samples of audio at a rate of 48 k-samples/second, for each channel of the stereo pair. The MUSICAM algorithm compresses the 1.54 Mbps to 256 kbps. AD-HDTV provides an additional 5 dB of robustness for audio, as it transmits audio as High Priority data. Thus, audio remains when only the HP picture can be decoded.

C. Data

An allocation of 256 kbps has been made for auxiliary data, and that data has been set aside within the prototype equipment as SP data. A standard communications interface port for this auxiliary data was provided for the prototype.

Auxiliary data is normally sent in standard priority (SP) mode. The actual priority used in an operational system is a broadcaster option and will depend on the type of data service. If <100 kbps of HP capacity were used for data, the proponent claims that the impact on HP-only picture quality would be small; for higher auxiliary data rates, use of SP is recommended, since additional reliability in data services can be provided via appropriate retransmission schedules, etc. Also, see G. Addressing.

D. Text

Text delivery is discussed below in Section I on Overlay.

E. Captioning

It was suggested that 9.6 kbps of the ancillary data capacity be used for closed captioning.

F. Encryption

The prototype hardware did not include encryption, but AD-HDTV's transport structure has been designed to readily accommodate encryption. The proponent expects to detail the encryption method with industry participation.

G. Addressing

The system provides opportunities for high-data-rate burst-mode delivery of auxiliary data. As a limit, the entire channel capacity, 18.5 Mbps, could be dedicated to addressing a large number of receivers with decryption keys during the few seconds preceding a popular pay-per-view program. A still picture with audio using 4 Mbps could be transmitted while 14.5 Mbps is devoted to subscriber addressing. Conditional access data can be treated as a special service type and packaged into its own transport cells, or included within the video and audio data. It can be decoded without decoding video and audio.

H. Low-cost receiver

AD-HDTV is partitioned into High Priority and Standard Priority data streams that modulate separate carriers. A low-cost receiver can be built to use only the HP stream with savings on the tuner and processing components. The HP-only mode has been demonstrated, showing high resolution but poor response to motion. The proponent anticipates improvements to the prioritization algorithms to make HP-only pictures that are much better suited for small, low-cost receivers.

I. VCR capability

No hardware development of VCR's has been reported. However, it is thought that digital recording is within current technology for consumer use.

MPEG has periodically occurring frames that are entirely spatially coded. This is said to provide the ability to reconstruct pictures in fast-forward and reverse scanning modes from digital storage media.

MPEG compression has periodic spatially-coded frames, allowing splices and inserts to be made on GOP boundaries in the compressed transmission format. Limited picture cropping can be handled in compressed form if it aligns with macroblock boundaries.

AD-HDTV has provision for carrying text and graphics overlay data. This data can be sent as a separate service type, to be superimposed on the display at the receiver. The proponent recommends that receivers have, as a minimum, overlay capability for simple text. This would permit the local broadcaster to insert text of local importance into the data stream containing the compressed signal that he passes through from the network source to the viewers, without decompressing and recompressing the video.

III. Extensibility

A. to no visible artifacts

AD-HDTV has, as its core, the MPEG1 compression standard that is extensible to virtually any data rate. The proponent points out that AD-HDTV at 17.7 Mbps is already an extension of the baseline MPEG1 parameters which encode low-resolution video at 1.5 Mbps.

B. to studio-quality data rate

AD-HDTV was designed with the anticipation of several levels of related MPEG compression. The proponent suggests that a studio standard could be set at 216 Mbps, the data rate of existing studio D-1 recorders. He believes that this standard would be of extremely high quality and would permit studio processing without degradation.

C. to higher resolution

The proponent claims that AD-HDTV potentially supports the delivery of other video and image formats over appropriate bandwidth channels to special receivers with increased memory. The MPEG1 core allows resolutions up to 4095 x 4095. Examples given by the proponent follow. One is a 1280 x 1024 progressive scan, logically square pixel format that is appropriate for computer workstations. Another is a 65,520 x 65,520 very-high-resolution still-image format (the limit of JPEG addressing capability). Even higher resolutions can be delivered by extending the JPEG addressing capability. Such formats would likely prove useful in medical imaging, computer graphics, and space and defense applications.

The proponent has discussed the possibility of introducing ultra-high-definition television by sending augmentation data packets assigned to a unique service type that will be disregarded by older receivers but processed by new receivers.

D. Provision for future compression enhancement

The proponent suggests improved calculation of motion vectors and improvements in bit allocation and prioritization as likely means for picture quality improvement without changing the receivers or the data rate.

Channel Compatible DigiCipher HDTV Interoperability Assessment

System Description

The Channel Compatible DigiCipher (CCDC) HDTV System, also known as the Advanced Television Alliance (ATVA) Progressive System, is an all-digital system that is progressively scanned with 787.5 lines per frame and 59.94 frames per second. The display format has square pixels with 720 active lines by 1280 pixels per line in a 16:9 image aspect ratio. Several types of digital coding are used for compression and signal robustness. Transmission is by quadrature amplitude modulation with 4 or 5 bits per symbol, offering the broadcaster the choice of two data rates involving a tradeoff between picture quality and coverage area. The system accommodates either 16- or 32-QAM modulation. The proponent emphasizes that the system was designed to be highly modular, so that the video processing, the audio processing, and the transmission system can be used independently.

I. Interoperability

A. with Cable TV

The tests of the proponent systems at ATTC by Cable Labs should determine the critical aspects of interoperability with cable TV. Other factors that may aid cable TV interoperability include channel augmentation to 9 or 12 MHz for improved picture, or better, going to a higher data rate that can be supported in the 6 MHz channel by the relatively benign cable environment. The question remains whether HDTV receivers can be built to support higher numbers of bits per symbol without cost impact.

Encryption and addressing are important service features for Cable TV. See Section II, items F and G.

B. with digital technology

Since this system is all-digital, the advantages of all-digital systems apply.

C. Headers/descriptors

A frame header identifies the video source material, the frame rate, resolution, aspect ratio, and other system data.

D. with NTSC

As the CCDC system is directly related to NTSC, transcoding to NTSC is straightforward. Edge-crop conversion to CCIR 601 involves discarding 160 HD samples from each end of the horizontal line and 4:3 interpolation of the remaining samples. It can involve 3:1 and 2:1 vertical interpolations in alternate fields. For letter-box, a 16:9 horizontal interpolation is required, as well as a 3:1 vertical interpolation of the active lines. For edge-crop display, the 480 lines can be placed in a specific set of the 483 active lines. The first (odd) field can be displayed on lines 21, 23, ..., 499, and the second (even) field is displayed on lines 22, 24, ..., 500. The active lines 501 and 503 of the odd fields and 502 of the even fields can be displayed as black. For letter-box display, the vertical interpolation is more

complicated. Up-conversion from NTSC requires line tripling, horizontal line-rate conversion and interpolation.

E. with film

The prototype has the following film mode. Film is displayed with the three-two pulldown process for 24 frame/sec film and with simple frame repetition for 30 frame/sec film. The proponent claims actual frame rates of 59.94, 30 and 24 frames/second, implying that no film slowdown is needed to accommodate the standard video frame rate. The encoder automatically detects the presence of 24 frame/sec or 30 frames/sec scene material from film sources. When a film source is detected, an alternate buffer control algorithm is used which takes advantage of repeated frames in the source. With the scanning method used in CCDC, only two out of each five TV frames need to be sent for 24 frame/sec film. The proponent claims that the video quality in the film mode is essentially equivalent to the uncoded source. He also claims that full-resolution color is possible in the film mode, though that capability is not in the prototype.

F. with computers

Progressive scanning and square pixels, both used in this system, are important factors for interoperability of an HDTV system with computers -- nearly all existing bit-mapped computer graphics displays have these features. Progressive scanning and square pixels are most critical for real-time applications such as display, scan-conversion, frame capture, and video effects. Computers are expected to play an increasing role in video image generation and production, and it is desirable to have an HDTV format which allows direct display and manipulation of HDTV video on the computer. Progressive scanning is preferable for computer applications to avoid artifacts that are common with interlaced display of computer generated imagery. Also, scan conversion between interlaced and progressive systems can produce undesirable artifacts. Square pixels are inherent to graphics display hardware for all popular computers employing bit-mapped displays. The main reason is that it facilitates processing of 2D transformations, especially rotations.

G. with satellites

Digital signals with bit rates used by the proponents for terrestrial transmission can generally be accommodated in the available bandwidth of a satellite transponder. The total data rate for the CCDC system in the mode tested by ATTC is 26.4 Mbps. The proponent suggests that 8-PSK modulation would permit two CCDC signals per 36 MHz transponder. However, normal transmission by satellite is QPSK (4-phase). CCDC, which has the highest data rate of the all-digital proponents, would be the most difficult to transmit with two channels per transponder. Nevertheless, using the 19.9- Mbps information rate of CCDC, Reed-Solomon coding, and rate 7/8 convolutional coding, two channels can probably be transmitted in a 36-MHz transponder.

H. with packet networks

CCDC data is organized into 525 data lines per frame. These data lines could be used as packets if augmented with packet assembly information. Error concealment, already implemented, would ensure some resistance to packet loss. The proponent has suggested implementation of a packetized structure after field testing.

I. with interactive systems

A form of interactivity important to cable operators is the speed with which a customer can scan through many channels. This system is reported to acquire a channel in 0.4 seconds, providing the sound and beginning to build the picture. The picture is at full resolution in 0.73 seconds.

Latency is the time delay between a video frame going into the encoder and the corresponding frame coming out of the decoder in the back-to-back mode. It can be important in interactive applications. Frame delays are required at the transmitter and at the receivers for coding and decoding. Rate buffers at the transmitter and at the receiver cause additional delays. Furthermore, the system implemented for test may not be representative of the minimum delay possible. Comparison of latency between the proponent systems is difficult at this time.

CCDC reports a video delay of 5 or 6 frames, corresponding to 83 to 100 ms. The exact time is said to depend on how the frame buffer is used, with the video/film selection a factor.

J. Format conversion

1. 1125/60

Upconverting to the Common Image Format (1920 x 1080) is easily done by 2:3 interpolation horizontally and vertically. SMPTE 240M currently uses 1035 active lines. Colorimetry is SMPTE 240M.

2. 1250/50

Conversion between systems with different frame rates is the most difficult type of conversion presently being done. Digital conversion between 59.94 fields per second and 50 fields per second requires a number of frame stores and very large processing capability. Present methods that involve frame dropping lead to jerky motion, but other techniques can produce acceptable images under most conditions. This difficult conversion may be easier from a progressive source than from an interlaced source.

3. MPEG

There is no direct compatibility in terms of bit stream. However, see the MPEG discussion for DigiCipher.

4. with still image

The capture of still images from video is favored by progressive scan.

K. Scalability

The proponent claims that the system is scalable in the sense that the decoded signal may be interpolated or decimated to suit various displays. Scalability by picture interpolation can be implemented in any proposed system. It is simplified by the progressive scanning in this system. The system may also be scalable in the sense that a subset of the bits, such as those that correspond to the lower-order discrete cosine coefficients, could be used by lower-cost VLSI to obtain a lower-resolution signal. However, the entire bit stream must be available for the subset to be extracted, and the other digital systems have similar capabilities. Picture-in-picture and picture-out-of-picture are handled by standard methods in the receiver.

II. Scope of services and features

A. Initial Use of Ancillary Data

Assignments have not yet been made. See Data.

B. Audio

The CCDC system is designed to have four channels of audio using 503 kbps or six channels of audio using 755 kbps. Unused audio capacity is used to carry additional video data. The audio system uses a coding algorithm called MIT Audio Coder, presently implemented with Motorola DSP-96002 floating-point digital signal processors. It accepts 16-bit audio samples at a rate of 48 kHz for each channel. All audio channels are coded independently. Capacity has been reserved for dynamic bit allocation for different parameter sets. The proponent claims that this will permit the system to incorporate future improvements in audio compression without changing the receivers.

C. Data

Auxiliary and control data has been allocated 251.8 kbps. This may be implemented with an RS-422 interface. In the current hardware, the only access to the auxiliary data channel is via four asynchronous 9600-bps RS-232 interfaces. However, the design can be modified to permit data capacity, not used by audio, to be reassigned to data services.

D. Text
E. Captioning

Teletext and captioning is sent in the ancillary data channel. Additional ancillary data capacity will be provided if needed.

F. Encryption
G. Addressing

The first byte in each data line is reserved for control information, described as including decryption keys and subscriber data. There are 525 data lines per frame and 59.94 frames per second. Thus, there are about 252 kbps of capacity for this kind of data. Encryption has not yet been implemented.

H. Low-cost receiver

This system will permit lower receiver cost using lower-resolution displays. The proponent claims that progressive scan used in this system favors such scalability. The low-cost receiver would have to extract the information it can use from the complete data stream, perhaps just the DC DCT coefficients, but would save on processing capability as well as display cost. The proponent claims that the standard HDTV receiver can be built with 12 custom VLSI's and 2 Mbytes of memory.

I. VCR capability

The proponent reports no hardware development of VCR's specific to CCDC, but refers to the DigiCipher/Toshiba VCR that has been demonstrated within ATVA. The CCDC data stream, about 20 Mbps, is within the capability of current technology for consumer use.

It is claimed that a rapid search mode can be implemented by reconstructing the images from those blocks coded with no temporal predictor. This gives at least three displayable frames for every

60. Additional intra-coded blocks may also be used as they occur. The resultant picture would have full resolution, but may include artifacts. The reverse playback cannot be done with full quality because predicted frames cannot be generated.

Splice and insert could be handled by forcing the receiver to reacquire. Crop and overlay would require that the data stream be decompressed first.

Square pixels and progressive scanning simplify the implementation of special effects such as zooming and panning.

III. Extensibility

A. to no visible artifacts

Based on simulation tests, the proponent believes that the compression algorithm will produce no visible artifacts at a data rate of 50 Mbps, regardless of the difficulty of the camera-generated source material.

B. to studio-quality data rate

According to the proponent, the intraframe encoding mode for the whole frame can be used for a production standard. Here, every frame is encoded without motion prediction. Production-quality video with a resolution of 1280 x 720 can be stored with 3 Mb/frame using the intraframe compression method included in this system. At 60 frames per second, the bit rate is 180 Mbps, an acceptable rate for studio use. The proponent claims that the frame can be decoded and re-encoded many times with little degradation.

C. to higher resolution

Currently the system is designed to display 1280 x 720 image sequences, but larger sizes can be specified as part of the frame header.

D. Provision for future compression enhancement

The compression algorithm can be improved by performing better motion estimation, and including better perceptual criteria at the transmitter. These involve no changes at the receiver.

DigiCipher HDTV Interoperability Assessment

System Description

Also known as the Advanced Television Alliance (ATVA) interlaced system, the DigiCipher HDTV system is an all-digital system with 1050 scan lines per frame, 59.94 fields per second, and 2:1 interlace. The display format has a 16:9 image aspect ratio with 960 lines per frame and 1408 pixels/line. Several types of digital coding are used for compression and signal robustness. The primary picture coding is discrete cosine transform with motion compensation. Transmission is by quadrature amplitude modulation with 4 or 5 bits per symbol, offering the broadcaster the choice of two data rates involving a tradeoff between picture quality and coverage area. The system accommodates either 16- or 32-QAM modulation.

I. Interoperability

A. with Cable TV

The tests of the proponent systems at ATTC by Cable Labs should determine the critical aspects of interoperability with cable TV. Other factors that may aid cable TV interoperability include channel augmentation to 9 or 12 MHz for improved picture, or better, going to a higher data rate that can be supported in the 6 MHz channel by the relatively benign cable environment. The proponent states that it is likely that a rate higher than 32-QAM will be used for multichannel NTSC, and has expressed the intent to maintain compatibility between DigiCipher Multichannel NTSC and HDTV. The question remains whether HDTV receivers can be built to support higher numbers of bits per symbol without cost impact. The proponent has conducted demonstrations of one-way and two-way operation over existing cable TV plants.

Encryption and addressing are also important service features for Cable TV. See Section II, items F and G.

B. with digital technology

Since this system is all-digital, the advantages of all-digital systems apply. The proponent has agreed to determine the feasibility of a common standard between the DigiCipher digital bus and the EIA CE bus.

C. Headers/descriptors

The proponent discussed the use of the ancillary data space for transmitting the program name, remaining times and program rating. In the prototype, there is a 7-byte header at the beginning of each data frame. Four bytes are used, and three bytes are available. There is a one-byte header at the beginning of each video frame, of which one bit is available. There is a fully defined two-byte header at the beginning of each macroblock. In his certification disclosure, the proponent discussed his intention to implement adaptive allocation between video, audio and ancillary data. The proponent has now announced an improvement involving expansion of the video frame and macroblock headers, and packetization at the transwitch layer. Transwitch layer packetization makes the data allocations adaptive, allowing flexible allocation of capacity among video, audio, and data applications. The proponent states that this modified data format will be implemented in the prototype prior to field testing.

D. with NTSC

The proponent selected the field rate of 59.94 Hz for compatibility with NTSC. The number of active video lines, 960, was selected to be double the number of active NTSC lines, 480. (NTSC normally has 483 active video lines, but 483 is an awkward number for conversions.) Since the aspect ratio of this system, 16:9, is different from the aspect ratio of NTSC, 4:3, a choice must be made for the conversion. The limiting choices for down-conversion are 1) "Edge Crop", in which the HDTV picture fills 480 lines on NTSC with loss of the sides of the HDTV picture, and 2) "Letterbox", in which the full width of the HDTV picture is displayed in the full width of the NTSC picture, but leaving unused areas at the top and the bottom of the NTSC screen. Pixel values from HDTV lines are stored in memory and read out at reduced speed to make NTSC lines. The range of pixels read and the clocking rate depend upon the choice of edge-crop or letterbox. The down-conversion involves interpolation between HDTV pixels in a line and between HDTV lines.

E. with film

The prototype system accepts film shot at 24 frames per second as 59.94-Hz video, 2:1 interlaced, having been converted by the three-two pulldown technique. The DigiCipher encoder recognizes the redundancy in each five-field sequence as having originated in 24-frame film, and converts the 59.94-field video back to 23.98 frames/second. The image is processed and transmitted as 23.98 frame progressive, and converted back to 59.94-field interlace in the decoder, using 3:2 pull-down. Future receivers could alternatively use 3:1 frame repeat to display progressive at 72 Hz. Similarly, 30-frame film source which is delivered to the encoder as 59.94-field video is processed and transmitted as 29.97-frame progressive. The benefit is more efficient coding, and thus higher quality. The proponent has proposed a system improvement which would allow receiving and processing images directly in 24- and 30-frame progressive. He also suggests that in the future, spatial resolution could be increased as temporal resolution is decreased.

F. with computers

Progressive scanning and square pixels, not included in the DigiCipher system tested, are important factors for interoperability of an HDTV system with computers -- nearly all existing bit-mapped computer graphics displays have these features. Progressive scanning and square pixels are most critical for real-time applications such as display, scan-conversion, frame capture, and video effects. Computers are expected to play an increasing role in video image generation and production and it is desirable to have an HDTV format which allows direct display and manipulation of HDTV video on the computer. Progressive scanning is preferable for computer applications to avoid artifacts that are common with interlaced display of computer generated imagery. Also, scan conversion between interlaced and progressive systems can produce undesirable artifacts. Square pixels are inherent to graphics display hardware for all popular computers employing bit-mapped displays. The main reason is that it facilitates processing of 2D transformations, especially rotations. Non-square pixels do not present a problem for the display of RGB signals from a computer's video card on HDTV receivers, but would complicate more sophisticated attempts at interoperability, such as the display by an HDTV receiver of a picture from a digital data stream generated by a computer. This system has pixels that are 21% wider than high.

The prototype hardware was built to select between field processing and frame processing for each superblock, depending upon its motion, in order to provide optimum motion handling. However, computer interoperability would be enhanced if the encoder were forced to do frame processing on all superblocks. With this feature, coding and transmission would be in progressive form. The proponent has proposed adding this feature as an option at the encoder.

G. with satellites

Satellite transmission of the DigiCipher HDTV signal was recently demonstrated using QPSK in a 24-MHz bandwidth, achieving a raw data rate of 39 Mbps. Instead of the trellis coding used in the terrestrial system, convolutional coding with a Viterbi decoder was used. The coding was rate-1/2, so that the data rate after Viterbi decoding was 19.51 Mbps. Reed-Solomon coding was also used, with the information rate being 18.2 Mbps, identical to the terrestrial signal. A 5.5 dB C/N threshold was achieved, an improvement over the 8+ dB threshold typically achieved in NTSC satellite transmission. The proponent recommends using rate-3/4 coding to yield a 50% increase in the information rate. This would support a higher-level compressed HDTV signal or an NTSC signal sharing the channel with the HDTV signal. In a 36-MHz transponder, two transmission-quality HDTV signals, or alternatively, one distribution-quality, 40-45 Mbps signal can be transmitted.

H. with packet networks

In the prototype tested at ATTC, the video data was packetized, with each packet, about 2 kb long, containing one macroblock of video data. For lost packets, the decoder would have used error concealment which was already implemented to handle transmission errors.

The proponent has since offered a redesign of the data structure with headers/descriptors. The benefits of this approach include improved packet transmission. Packets will be organized by data type. Packet length will be 155 bytes, with the payload being 141 bytes. Out-of-order packets can be reordered in a suitably designed receiver. Implementation is expected after field testing.

I. with interactive systems

A form of interactivity important to cable operators is the speed with which a customer can scan through many channels. This system is reported to acquire a channel in 0.4 seconds, providing the sound and beginning to build the picture. The picture is at full resolution in 0.77 seconds.

Latency is the time delay between a video frame going into the encoder and the corresponding frame coming out of the decoder in the back-to-back mode. It can be important in interactive applications. Frame delays are required at the transmitter and at the receiver for coding and decoding. Additional delay may be needed to facilitate frame coding in an interlaced system. Rate buffers at the transmitter and at the receiver also cause delays. Furthermore, the system implemented for test may not be representative of the minimum delay possible. Comparison of latency between the proponent systems is difficult at this time.

According to the proponent, the latency of DigiCipher is 83 msec.

J. Format conversion

1. with 1125/60

Up-converting to the Common Image Format (1920 x 1080) requires a 8:9 vertical interpolation from 1408 x 960. SMPTE 240M currently uses 1035 active lines. Conversion from 1408 pixels/line to 1920 pixels/line requires an 11-to-15 interpolation. The proponent says that 15 interpolation filters are required, each having 5-6 taps. Colorimetry used by DigiCipher is the same as SMPTE 240M.

2. with 1250/50

Conversion between systems with different frame rates is the most difficult type of conversion presently being done. Digital conversion between 59.94 fields per second and 50 fields per second requires a number of frame stores and very large processing capability. Present methods that involve frame dropping lead to jerky motion, but other techniques can produce acceptable images under most