CCITT SGXV
Working Party XV/1
Expert's Group for ATM Video Coding

AVC-253

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION
ORGANIZATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND ASSOCIATED AUDIO

ISO/IEC JTC1/SC29/WG11
MPEG92/
March 13, 1992

Source : AdHoc Group for Video Requirements Listing
Title : List of requirements for MPEG-2 Video
Purpose: Report

1. Mandate

To produce a neat list of requirements for MPEG-2 video standard with sufficient definition

2. Members

G. Fernando	Sun	USA	gerard.fernando@Sun.Com
A. Gill	Zoran	Israel	aharon@zoran.hellnet.org
B. Haskell	AT&T Bell Labs	USA	bgh%vax135@research.att.com
T. Homma	Toppan	Japan	homma@itl.toppan.co.jp
G. Keesman	Philips	Netherlands	keesman@prl.philips.nl
A. Koster	PTT Research	Netherlands	a.koster@research.ptt.nl
D. LeGall	C-Cube	USA	djl@c3.pla.ca.us
A. Lippman	MIT Media Lab	USA	lip@vertigo.media.mit.edu
D. Mead	Hughes	USA	DMead@NVOSUN.SCG.HAC.COM
J. Morris	Philips	UK	morrisoj@prl.philips.co.uk
G. Morrison	BT Labs	UK	morrison_d_g@bt-web.bt.co.uk
S. Okubo	NTT Labs	Japan (Ch.)	okubo@nttvdt.ntt.jp
O. Poncin	RTT	Belgium	Poncin@tele.ucl.ac.be
J. Princen	Telecom Australia	Australia	j.princen@trl.oz.au
G. Smith	AUSSAT	Australia	FAX: +61 2 238 7839
A. Tabatabai	Bellcore	USA	ajt@nyquist.bellcore.com
T. Tanaka	NTT Labs	Japan	tanaka@nttvdt.ntt.jp
M. Tinker	Intel	USA	mt@provax.intel.com

3. Working method

Through correspondence.

4. Outcome

4.1 Requirements listing

We started from the PPD (MPEG91/100 Rev, August 1992) to collect relevant materials and produced the following two documents for consideration of the Haifa meeting;

- General requirement listing which can be used not only for general guideline inside MPEG but also for consultation with outside communities whether MPEG work is properly oriented (Annex 1). This document provides short heading, definition, examples and comments for each identified requirement.

- Guide for the video work which lists those items giving direct impact on the video coding development and testing thereof (Annex 2). This document is an abbreviated version of the above general requirement listing.

During this correspondence work, it has been pointed out that these two documents place different emphasis on the same material and can cause debate and confusion. A suggestion has been made to make the second one shorter, summarizing guidelines for the development in the other groups by containing only pointers to the appropriate section in the first one.

This and other points for action should be discussed and decided in the Haifa meeting.

4.2 Work method for making the application profile

Another suggestion for consideration for the Haifa meeting is to use the following ranking to each requirement indicating the degree of necessity in the application profile;

0: requirement is not needed in the application
1: requirement is welcomed in the application
2: requirement is strongly recommended for the application
3: requirement is essential for the application

END

Annex 1

List of Requirements for MPEG-2 Video
Issue 3 (13 March 1992)

---

1. Picture Format
2. Picture Quality
3. Bit Rates
4. Flexibility in bitrates
5. Coding Decoding Delay
6. Random access/channel hopping
7. Bitstream scalability
8. Complexity flexibility
9. Compatibility
10. Editing encoded bitstreams
11. Trick mode
12. Repetition of coding and decoding
13. Adaptation to storage and transport methods
14. Real time aspect ratio changes

MPEG phase 2 work is intended to provide a generic coding method of moving picture images and of associated sound for digital storage media having a throughput of up to about 10 Mbit/s. The coding method to be defined is expected to have applications in many other areas such as distribution and communication, so the requirements of other applications will be taken into account during the definition of the MPEG-2 coding system. The coming standard is envisaged to have lifetime of 10 or more years after its establishment {Editor's note: to be confirmed}.

This document lists all the identified requirements for video part of the audiovisual coding, providing with definition, application examples and comments. This work has been carried out in close collaboration with the CCITT Experts Group for ATM Video Coding and in close liaison with the CMTT/2 Special Rapporteur's Group for Secondary Distribution of Digital Television and HDTV.

For the second phase work of MPEG, the following application areas have been identified;

CTV Cable TV Distribution on optical networks, copper, etc.
ENG Electronic News Gathering (including SNG, Satellite News Gathering)
IPC InterPersonal Communications (videoconferencing, videophone, etc.)
ISM Interactive Storage Media (optical disks, etc.)
NDB Networked Database Services (via ATM, etc.)
RVS Remote Video Surveillance
SSM Serial Storage Media (digital VTR, etc.)
STV Satellite TV Broadcasting
TTV Terrestrial TV Broadcasting

---

1. PICTURE FORMAT (INPUT/OUTPUT)

a) CCIR601 4:2:2
b) 4:4:4 described as tentative specification in Annex I to CCIR601
c) 4:2:0 {Editor's note: are we describing input/output of the codec or that of source coding/decoding?}

/Definition/

	Y	Cr, Cb
4:2:2	720pels x 576/483 lines	360pels x 576/483 lines (co-sited with odd Y samples)
4:4:4	720pels x 576/483 lines	720pels x 576/483 lines
4:2:0	720pels x 576/480 lines	360pels x 288/240 lines
	{Editor's note: the same comment as above, 480/240?}	

/Discussion/

The above are formats of the input and output for the video coding scheme. Inside the coding scheme, other formats such as 4:2:0 (Cr, Cb are vertically 2:1 subsampled) as well as cropping of the horizontal edges may be used. Handling of 3 lines near the blanking period in 525/60 signals need careful consideration in such cases. There may be some applications where significant information is included in these three lines.

/Applications/

Digital satellite broadcasting and secondary distribution of television signals are two demanding areas where 4:2:2 level of quality is required.

High resolution graphics and medical imaging terminals may require 4:4:4 resolution or alternatively three merged bitstreams each of which has "4" level of resolution. {Editor's note: this paragraph is awaiting confirmation.}

Another areas needing the 4:4:4 format are "video theater" using video projection and applications using repeated encoding-decoding.

c) EDTV (16:9 aspect ratio)

/Definition/

e.g. 960 x 240 x 2 x 30; 960 x 288 x 2 x 25
also progressive version?

/Note/

It is likely that there will be 16:9 sources with 720x576 resolution. This may occur for instance if a 16:9 source is recorded on a D1-recorder, which has 720x576 resolution. In terms of source coding, 16:9 pictures will probably need more bits. Hence quality can be a problem depending on the bit rate.

d) Progressive format (film source etc.)

/Definition/

e.g. 960 x 576/480 x 1 x 25/30/50/60

/Applications/

- film source
- EDTV?

- SCIF having twice higher resolutions than CIF in three dimensions, which is now under consideration in CCITT Experts Group for a worldwide unique picture coding format

e) Square pixel

/Definition/

Pel aspect ratio is 1:1.

/Discussion/

It is known that if a CIF picture having 1.07:1 aspect ratio is displayed on a square display monitor, geometrical distortion becomes clearly visible. Square pixel is likely to be adopted in CCD cameras and LCD displays {Editor's note: this sentence should be deleted?}.

/Applications/

Integration of video communication terminals and desktop workstations, multimedia computers etc.

2. PICTURE QUALITY

a) High quality
b) Graceful degradation for critical scenes

/Definition/

The target quality for the second phase video of MPEG is envisaged as not lower than NTSC/PAL/SECAM and up to CCIR-601. Considering prospects for the future technology, target quality and corresponding bit rates for the MPEG Phase-2 work are summarized in the following two categories of signal bandwidth assuming CCIR-601 input;

SIGNAL	LUMINANCE	COLOUR DIFFERENCE (I,Q) or (Cb,Cr)
	Bandwidth in MHz (Approximately)	Bandwidth in MHz (Approximately)
Category 1		
PAL/SECAM	3.5	1.0
PAL/VHS(VCR)	2.5	0.6
NTSC	2.7	1.0,0.6
	quality level within 3-5 Mbit/s	
Category 2		
Close to CCIR Rec.601	5.75	2.75
	quality level within 8-10 Mbit/s	

The users of secondary distribution sources will require high levels of quality and low impairment levels, taking account of the viewing distance (4-6H for CTV/EDTV or 3H for HDTV), the entertainment value of the services and the lengthy viewing times. The following criteria concerning quality are considered desirable in the choice of the coding algorithm:

- For the majority of picture sequences representative of typical programmes, impairments shall be usually imperceptible (i.e. an impairment rating better than CCIR grade 4.5 is to be achieved).

- For a small number of picture sequences that are very critical, there is a probability that some impairment may be visible, dependent on content.

- No picture sequence shall create a catastrophic failure of the coding algorithm, but a gradual, graceful and progressive raising of the level of the impairments may be visible as extremely critical sequences are introduced, that are above the activity level of typical, critical programme sequences.

c) Minimum coder-decoder combination should provide targeted performance

/Discussion/

There is an unanswered question whether a video coding standard, generic one in particular, should or could guarantee quality in addition to bitstream interchange capability. H.261 or MPEG-1 video guarantees the latter only, but audio coding standard usually guarantees quality as well.

3. BIT RATE

a) NTSC/PAL/SECAM level at 3-5 Mbit/s
b) Close to CCIR 601 level at 8-10 Mbit/s

/Definition/

The target bit rates for video are defined not at specific fixed values but as a range so that the standard cover a range of picture quality corresponding to a range of bit rates. The upper bound is 10 Mbit/s. The lower bound, however, is open at the moment. Depending on the applications, the mean bit rate for video is envisaged as follows;

- 5 to 10 Mbit/s (CTV, ENG, SSM?, STV, TTV)
- 2 to 10 Mbit/s (IPC, ISM, NDB, RVS)

4. FLEXIBILITY IN BIT RATE

a) CBR and VBR
b) Rate control for VBR under bit generation constriction

/Definition/

CBR (constant bit rate) coding

Number of transmitted bits per unit time is constant on the channel.

Since the encoder output rate generally varies depending on the picture content, it should regulate the rate constant by buffering etc. In CBR, picture quality may vary depending on its content.

VBR (variable bit rate) coding

Number of transmitted bits per unit time may vary on the channel under some constriction. VBR is expected to provide constant quality coding. In case of ATM network, this network constriction is called Usage Parameter Control (UPC). If the coded bitstream does not meet UPC requirements, some part of the bitstream may be discarded at the input of the network. In case of multichannel transmission in the CBR environment, each channel rate should be controlled so as to maintain the aggregate bit rate constant.

/Applications/

- Video transmission in ATM environments
- Multichannel transmission in DBS etc.

5. CODING/DECODING DELAY

/Definition/

End to end delay of the system

Time elapsed between the camera output of a particular part of the picture and the monitor input of the corresponding part.

Encoding and decoding delay

Time elapsed between the encoder input and and the decoder output when the encoder and the decoder are connected back-to-back. Major contributing elements are frame reordering, format conversion, source coding + buffer delay, transmission coding and decoding.

Representation of the delay

Both in number of frames (fields) and in ms are recommended.

/Note/

See MPEG92/010 (AVC-179) for detailed analysis of delay.

/Target values/

For conversational services (IPC), encoding and decoding delay of less than 150 ms is targeted at bit rates higher than 2 Mbit/s. Other applications are not so sensitive to delay except in distribution applications (CTV, STV, TTV) where a conversation takes place between the studio and the end user or where betting for horse races etc. is involved. Less than 500ms delay may be required for this purpose.

6. RANDOM ACCESS / CHANNEL HOPPING

/Definition/

Random access

The process of beginning to read and decode the coded bitstream stored in DSM at an arbitrary point. Access time is defined as the time lap between 1) selecting a specific program and 2) displaying the first picture from the corresponding bitstream.

Channel hopping

The capability of selecting a television channel and having the service fully operating. This response time is not yet specified but CMTT/2 considers a range of 0.3 to 1 second for the complete system.

/Applications/

- Interactive storage media access
- TV distribution
- Switched video multipoint where a picture of site containing the current speaker is selected for distribution. This functionality is also required when pictures should appear quickly after the channel is connected.

7. BITSTREAM SCALABILITY

a) Spatial resolution scalability
b) Temporal resolution scalability
c) Coding noise scalability
d) Complexity scalability

/Definition/

A scalable bitstream is one where you can neglect some of the bits in the bitstream, and still decode a useful picture. Demonstration of bitstream scalability could be a 9 megabit bitstream subsampled successively to form approximately 6 and 4 megabit bitstreams. Syntactically, this implies that the bitstream is formatted so that bits within it can be ignored. That is, the six and four megabit subsampled streams defined here are not created by re-formatting the data but by simply neglecting parts of the stream.

Resolution scalability is when the video can be decoded at different resolutions or sizes directly. This is distinct from decoding a full size image and then decimating it.

Temporal scalability is when the decoded bitstream results in a sequence of different frame rate.

Complexity scaling is when a bitstream can be decoded by systems of varying complexities.

/Discussion/

Each of the elements described above is independent although they are usually linked in a given implementation. For example, a reduced rate bitstream can result in an image that has more coding noise, is smaller, or has a lower frame rate. Each case is useful for a different application.

We distinguish between scalable coding and layered coding. The latter is a subset of the former. A scalable system implies a broader range. {Editor's note: "layered coding" needs definition?, see the note in the compatibility section as well}

/Applications/

The following applications are presented to help clarify the meaning and intent of a scalable system. Scalable systems inevitably carry a cost in performance (higher bit-rate or lower quality and greater complexity) compared with non-scalable systems. The features listed below can be provided either by a scalable system or by other methods (e.g. multiple decoders or simulcast). The relative advantage of a scalable system compared to non-scalable solutions should be studied. This list is not exhaustive.

- Picture in Picture, Picture out of Picture: In broadcast television, it is desirable to show more than one program on the screen (e.g. a 16:9 screen with three images alongside a large 4:3 image, or a reduced size image overlaid onto the larger image). The smaller image might be between 1/4 and 1/16 of the full screen area. It is desirable to implement this feature without requiring multiple full scale decoders and full screen intermediate image stores.

- Retransmission: Occasionally a digital video signal is distributed via a satellite link then re-transmitted through a cable system at reduced bandwidth because the cable system has less capacity than the satellite system and the operator does not wish to sacrifice program diversity. One solution to this is to accept a 10 Mbit/s stream at the headend, and retransmit a 4 Mbit/s version of it along the cable. In this case, the reduced rate signal is to be decoded at full size but with some sacrifice in image quality.

- Desktop videoconferencing: In a network environment, one may wish to view more than one participant in a visually accompanied telephone conversation (continuous presence multipoint scheme). To solve this, it is desirable to request network bandwidth dynamically and allocate local reconstruction processing resources. Since many viewers may be viewing the same image at different bandwidths, it will be encoded once and decoded in a variety of different ways. This is an example of the more general case of heterogeneous networks and decoders.

- Cheap TV receiver : very small screens (as for TV Walkman) do not require the full definition. Partial scale decoders being cheaper, the global cost of such small TV receivers might be reduced.

/Ranges of values for scalable parameters/

In the case of television, all reconstructions of the image must be of high quality (for entertainment viewing). The desired range of bitrates varies from full rate to 1/100 that rate or less, with a limited number of intermediate values that are logarithmically distributed (e.g. 1/2, 1/4, 1/8, 1/16,...).

Resolution scalability should be at least by successive factors of two.

Multirate reconstruction similar to MPEG-1 by omitting interpolated frames is probably suitable for most applications.

Complexity scaling is not easy to quantify. A software decoder (using a general purpose processor such as *386 or *486 or *040) is mandatory for computer applications. Therefore there should be no elements of the decoder that preclude such decoding. Depending on the extent of the definition of the core bitstream, it may be desirable to implement variable complexity decoding so that decoders can interpret a larger set of bitstreams and produce reasonable results without incurring the expense of complete decoding of all possible bitstreams.

/Note/

Scalability is distinct from the ability to encode at a wide range of rates.

8. COMPLEXITY FLEXIBILITY

a) Flexibility to allow tradeoff between performance and complexity
b) Intraframe mode only codec

/Definition/

There are three cases of different symmetry in terms of allowable complexity:

- high performance encoder with low complexity decoder
  e.g. CTV, ISM, NDB, STV, TTV
- balance of performance / complexity between encoder and decoder
  e.g. IPC, SSM
- low complexity encoder with high performance decoder
  e.g. ENG/SNG

The standard should be flexible enough to allow both high performance / high complexity and low performance / low complexity coders and decoders. Capability to operate in intraframe mode only may offer advantages in some cases (e.g. for low complexity).

/Applications/

Intra mode only coding may be applied in recording applications where editing on picture basis is required, without introducing too many new errors.

9. COMPATIBILITY

a) forward compatibility with MPEG-1/H.261
b) backward compatibility with MPEG-1/H.261
c) upward/downward compatibility with EDTV, HDTV, SDTV

/Definition/

Upward and downward compatibility

Compatibility here refers to a transmission system, where different picture formats are used for the video encoder and video decoder. Different picture formats do not imply different standards. The system is:

- upward compatible if a higher resolution receiver is able to decode pictures from the signal transmitted by a lower resolution encoder.

- downward compatible if a lower resolution receiver is able to decode pictures from the signal or part of the signal transmitted by a higher resolution encoder. Two ways of downward compatibility can be discerned:

    * The decoder reconstructs the entire picture at lower resolution.
    * The decoder reconstructs a window of the input picture.

When no further notice is made, it is assumed the decoder reconstructs the entire picture at lower spatial resolution. The frame rate is not necessarily equal.

Forward and backward compatibility

Here, compatibility refers to a transmission system where different standards are used for video encoder and video decoder, i.e. an existing standard and a new standard. The picture formats of these standards can, but need not differ. The system is:

- forward compatible if the new standard decoder is able to decode pictures from the signal or part of the signal of an existing standard encoder.

- backward compatible if an existing standard decoder is able to decode pictures from the signal or part of the signal of a new standard encoder.

It is assumed the entire input picture is reconstructed by the decoder, possibly at different spatial or temporal resolutions.

/Implementation methods/

There are several possible implementation methods for these compatibilities; simulcasting, embedded bit stream, syntactic extension, switchable encoder, standard families.

Simulcasting

In this case the encoder system is characterized as follows. Typically, two encoders operate in parallel, one according to an existing standard and picture format, the other according to a new standard and/or picture format.

a - It transmits N (with $N > 1$) multiplexed streams of data, which may be separated at the decoder.
b - Data streams 1..K-1 (with $K <= N$) are decodable by an existing standard decoder after demultiplexing.
c - In a new standard decoder pictures are decoded from a set of one or more data streams K..N without making reference to data streams 1..K-1.

Backward compatibility is achieved by feature b, whereas forward compatibility is not guaranteed. A new standard decoder will discard the existing standard data streams 1..K-1. Decoding of the existing standard may or may not be included as a special option.

Considering upward and downward compatibility, downward compatibility is achieved by feature b, while upward compatibility is not guaranteed, as the information for the two picture formats is transferred and processed independently.

In principle, this compatibility method under certain circumstances could be wasteful of bandwidth as the same picture information is transferred several times in different multiplexed data streams.

Embedded bit stream

In this case the encoder is characterized as follows:

- a - It transmits N (with N > 1) multiplexed streams of data, which may be separated at the decoder.
- b - Data streams 1..K-1 (with K <= N) are decodable by an existing standard decoder (backward) or a decoder with smaller picture format (downward) after demultiplexing.
- c - From data streams 1..K-1 pictures may be decoded without reference to the other data streams, but decoding pictures from a data stream M (with K <= M <= N) is not possible without making reference to one or more of the data streams 1..K-1. Data streams K..N carry information additional to data streams 1..K-1.

Backward or downward compatibility are achieved by feature b. Forward compatibility is achieved as the new standard decoder can decode pictures of existing standard quality from data streams 1..K-1 only.

This also implies that upward compatibility is achieved, as data streams 1..K-1 carry lower resolution pictures, while data streams K..N carry the additional information for full resolution pictures.

In principle there is no waste of bandwidth since the N multiplexed data streams carry complementary information only. In practice however, the constraint of an existing standard or for the data streams 1..K limits the achievable coding efficiency when comparing with an equivalent stand-alone system.

Syntactic extension

In this case only one data stream is transmitted. The data stream produced by the new standard encoder has a syntax which is an extension of the existing standard. This allows for forward compatibility, as the new standard decoder is equipped for the syntax of the existing standard and may decode the existing standard when little adaptations in the decoding process are made.

A similar description is possible for upward compatibility: the data stream for the full resolution pictures is an extension of the data stream for the lower resolution pictures, such that the full resolution decoder can decode the signal of the lower resolution encoder.

Backward or downward compatibility is not achieved by this method, as the signal for the existing standard or lower resolution decoder as such is not embedded or simulcasted in the data stream. A transcoder with more than a demultiplexer and multiplexer would be needed to obtain the existing standard or lower resolution signal.

Switchable encoder

This method of compatibility is mainly intended for services where the type of receiver(s) can be identified by the transmitter, e.g. for point to point conversational services.

The encoder is characterized as follows:

a - It transmits one stream of data only.
b - To achieve forward and backward compatibility, the encoder is capable to operate in new standard or existing standard mode. For upward and downward compatibility, the encoder must be capable to produce the signal for full resolution or for lower resolution decoders.
c - Encoder and decoder(s) negotiate to determine which standard and/or picture format will be used for the connection.

Standard families

This is not a compatibility method, but allows for joint developments for several standards. A new standard having many commonalities with an existing standard, or a family of standards for several picture formats, may reduce efforts for development and optionally facilitate development of dual(or multiple) standard equipment. This may be beneficial for introduction of a new service.

/Note to "Layered Coding"/

Both "Embedded Bitstream" and "Syntactic Extension" compatibility methods are versions of layered coding. The difference between the methods is that all the data passes through the decoder (though some may be ignored) in the case of "syntactic extension". With the "embedded bitstream" method, the unused bitstreams need not be presented to the decoder.

/Guideline/

MPEG will seek to provide compatibility with CD-11172 to the maximum extent.

/Discussion/

Generally new standards are linked with higher resolutions, thus they may require higher bit rates. Existing or lower resolution decoders may only be able to accept lower bit rates. When we discuss downward or backward compatibilities of the coded bitstream, we assume that multiplexing and demultiplexing of two or more bitstreams are supported at the system level, one of which conforms to the low resolution or existing standard.

/Applications/

- MPEG-1 disks are played by an MPEG-2 deck.
- MPEG-2 disks are played by an MPEG-1 deck.
- H.32X terminal containing H.26X video coding interworks with H.320 terminal containing H.261 video coding.
- Digital HDTV program signal is decoded by CTV receiver.
- Digital CTV program signal is decoded by HDTV receiver.

d) Provision for future extension (e.g. extension to larger formats as HDTV)

/Definition/

Hooks for future extension are prepared in the standard, such as "extension_start_code" in MPEG-1 and "PEI/GEI" in H.261. Equipment conforming to the standard can ignore the data following these hooks so that future backward compatible extension is allowed.

10. EDITING ENCODED BITSTREAMS

/Definition/

A coded bitstream for a new picture is produced from multiple coded bitstreams representing different pictures without decoding and recoding.

/Discussion/

The capability in defined circumstances of cut-and paste editing at GOP boundaries provided by the MPEG-1 Video algorithm is adequate for this purpose.

/Applications/

- Multipoint communication is one of the most desired features in this category, where multiple terminals participate in a communication. In order to present multiple sites in one picture (called continuous presence scheme), mixing of several pictures is necessary. Better picture quality and less processing delay would be obtained if coded signals could be multiplexed without decoding and recoding.
- Picture database generation
- Multimedia computers

11. TRICK MODE

a) Fast playback (forward, backward)

/Definition/

The process of displaying a sequence, or parts of a sequence, of pictures in display/reverse-order faster than real time.

b) Normal reverse playback
c) Slow motion
d) Pause

12. REPETITION OF CODING AND DECODING

/Definition/

Coded picture is decoded, and then decoded picture is coded again. This may be repeated several times.

/Applications/

- Editing of multiple coded materials to obtain a new picture.
- Continuous presence multipoint system where MCU (Multipoint Control Unit) generates a new picture containing multiple sites participating in a videoconferencing and transmits it to a particular site.

13. ADAPTATION TO STORAGE AND TRANSPORT METHODS

a) Bit error resilience (random and burst)
b) Cell/packet loss resilience (random and burst)
c) Graceful degradation for severe cases

/Required resilience/

Robustness to data corruption can be stated as follows from the implications of the transport media;

- High quality service at BER of 10E-4 including error bursts of up to 30 bits (necessary for ENG, STV, TTV)
- Service continuity at BER of 10E-3 including error bursts of up to 30 bits (necessary for ENG, STV, TTV)
- High quality service at Cell Loss Ratio of 10E-n for high priority channel (necessary for IPC and NDB via ATM networks)
- Service continuity after loss of multiple successive ATM cells, i.e. up to e.g. about 1500 missing bits (necessary for IPC and NDB via ATM networks)

{Specific required values need reference to well established papers}

/Note/

Error protection will also be supported at the system level.

d) Recovery of synchronization after an arbitrary point

/Definition/

Recovery after the channel is lost for a while.

e) Particular requirements of the media or transport methods

/Definition/

DSM

{to be provided}

Secondary distribution

While the channels used for secondary distribution of television are generally near error-free, unusual circumstances may introduce significant levels of errors. In that situation, it is believed desirable that;

- in the presence of uncorrectable errors, the decoder should fail gracefully (progressive and gradual increase of impairment levels with increasing error-rate and may resort to error mitigation techniques (e.g.. error concealment) to maintain acceptable output.

- a priority must be placed on the maintenance on the recovery of synchronisation during error overloads.

ATM networks

"ATM networks" are not a type of media, but a transport method which is independent of the medium. It could be used as the transport for the satellite/terrestrial broadcasting applications mentioned below, for example.

- The video coding system must optimize performance, in terms of picture quality and end to end delay, on the B-ISDN. This must include cell loss tolerance and could take advantage of VBR coding.
- The video coding system must provide for maximum integration of all video services on the B-ISDN.
- The video coding system must permit advantage to be taken of the flexibility of transport offered by the B-ISDN; high/low priority channels, VC (Virtual Channel) based multimedia multiplex, multipoint distribution in the network, etc.

Satellite broadcasting

- The video coding system must have adequate performance at low C/N levels.

  * The video coding system must provide a graceful failure mode during heavy rain attenuation periods,
  * provide robustness against interference,
  * able to operate in saturation region of a satellite TWT transponder.

Terrestrial broadcasting

- Provide robustness against ghosting.
- Provide robustness against interference
- Optimized performance achievable possibly with CO FDM techniques

14. REAL TIME ASPECT RATIO CHANGES

a) Pan/scan
b) Letter box changes

/Definition/

Pan/scan

Indicating sender's desire concerning what part of the 16:9 picture be displayed on 4:3 monitors. Desired portion may be panned or scanned.

Letter box

16:9 signal is reduced in vertical resolution and displayed on a 4:3 monitor maintaining original aspect ratio.

END

Annex 2

Guide for the video work

This document intends to extract from the general requirement listing only those requirements that directly impact the development of video coding standard and the testing thereof {Editor's note: section order should be aligned with Annex 1}.

1. Picture Quality

The target quality for MPEG-2 is commensurate with the data rate. At 4 Mbit/s, we envision images better than broadcast television with most entertainment television CCIR-601 source material. At 9 Mbit/s, we envision substantially unimpaired images. It must be taken into account that 16:9 pictures require a substantially higher bit rate.

2. Range of Use

MPEG-2 video can be used from bit rates of approximately 1 Mbit/s to over 20 Mbit/s. {Editor's note: "20 Mbit/s" conflicts with the "10 Mbit/s" MPEG-2 terms of reference?} At lower encoding rates one may choose to process a lower resolution raster than, but the standard has been optimized for CCIR 601 interlaced image sources.

3. Picture Format

MPEG-2 should support interlaced and sequentially scanned source material (such as film). We anticipate use of 4:2:2, 4:2:0 and 4:4:4 color ratios for applications from entertainment video distribution to computer networks and small scale production. The syntax should support these options, however the tradeoffs between picture quality and color resolution are a matter for applications to determine in use. We see no relevance to testing and development of the standard {Editor's note: this sentence should be deleted?}.

Similarly, aspect ratios will be determined through use. The syntax should support a variety ranging from square to at least cinemascope (2.55:1).

4. Data Rate Control

MPEG-2 should support constant bit rate and variable bit rate applications. The model for the latter is constant-quality coding over B-ISDN channels that apply Usage Parameter Control (UPC). In this case, some part of the data stream may be discarded in the case of unanticipated overload in the network. Therefore testing should provide for graceful recovery in the face of missing video data packets.

It is important that core experiments to develop the standard demonstrate the impact of proposed improvements on decoding video with missing packets.

5. Scalability

Bitstream scalability is defined in another document (MPEG92/078) In short, scalability is defined as a bitstream that need not be completely transmitted or decoded, a subset of the data should produce useful images. When a lower rate than is encoded is transmitted or processed at the decoder, the image will be either (1) noisier, (2) smaller, (3) at a lower frame rate, or requiring fewer processing operations to decode. Not all of these options are jointly supported; some of them may be included in the design of the bitstream at the option of the encoder.

For example, a bitstream can be designed so that decoding a part of it will result in reduced scale images in a stepwise fashion (quarter scale, 16th scale, etc.). Alternatively, another bitstream can be designed to reproduce images at the same scale as the input but with more coding noise as the decoding rate is reduced. This is the option of the encoding, and the same bit stream need not explicitly support both options simultaneously.

MPEG document {to be defined, Annex D of PWD?} explicitly describes options and applications for scalability. Digital VCRs and other applications require SNR scalability so that a full size image can be recorded without recoding even though the source rates may vary. Heterogeneous computer systems and networks require broadcast of a single bitstream that can be decoded using varying complexity hardware. Similarly, multipoint videoconferencing that accommodates different class of terminals requires this scalability. Decoding using software on modern PCs (I486, 68040) is essential. Television distribution requires a high quality (e.g. HDTV) bitstream that can be viewed on existing TV receivers with full entertainment value. Picture-in-Picture requires simple decoding of a reduced size image sequence. Thus all options in the definition of scalability are required in at least some applications. Therefore, all should be supported in the syntax and the core standard.

It is important that core experiments demonstrate the impact of any proposed alteration of the coding model on reduced scale decoding. For the purpose of testing, we strongly urge all core experiments that do not directly address scalability to demonstrate decoding a reduced rate bitstream at an agreed upon reduced rate, thus showing picture quality at both full rate and with partial decoding.

Since there are multiple dimensions of scalability, it is sufficient to show at least two reduced scale images, one at reduced rate and the other produced by a core standard software decoder.

[Argument: This entails more work for core experimenters however, it is the price for a generic standard. The additional work is small since the low rate image is quarter scale and thus computationally simpler to produce. Such a practice insures that no improvement to the coding model is made without knowledge of the effect of such an improvement on scalability.]

6. Codec Delay

An envisioned requirement is low encoding and decoding delay for real-time video communications such as telephony, conferencing, monitoring and synchronization with real-time sound. A test case is videotelephony with no delay in the sound path. A sound-to-picture asynchrony of approximately 16 milliseconds is a reasonable limit. [Note: CCIR Report 1081-1 indicates that 20ms advance or 40ms delay of audio relative to video is detectable.]

For model development, any core experiment that impacts the quality of the low-delay image should be demonstrated in low-delay mode; those that do not should provide argument to support such an assertion.

7. Error Resilience

In any real system, there are always transmission or storage errors. Shannon showed that optimal communication results when all redundancy is removed from the source through coding and transmission error resilience is introduced by a channel coder. However, MPEG-2 will often be used with no channel error correction overlay or with channel coding which has residual errors. Therefore resilience is deemed to be a feature (requirement) of the source coding itself. In particular, cell loss resilience in ATM networks and resilience to channel errors in storage and transmission media of up to one part in 10E-4 (ENG, STV, TTV) are required.

For model development, any core experiment that impacts the quality of an image with such an error rate should be demonstrated; those that do not should provide argument to support such an assertion.

8. Re-coding loss

It is desirable that as many operation on the undecoded bitstream as are practical be supported directly through such requirements such as scalability and error resilience but there are always occasions where the image sequence will have to be decoded, processed and later re-coded. Examples include video effects such as keys, mattes and windowing, when the images are encoded in one standard and re-interpreted for another, or when multiple images are composed into an image to support continuous presence multipoint videoconferencing. Therefore, it is desirable to minimize recoding loss.

The requirements subcommittee has no criteria for development of the standard that exemplifies this criterion.

9. Compatibility

We identify three aspects of compatibility:

1) forward compatibility with MPEG-1 and H.261
2) backward compatibility with MPEG-1 and H.261
3) upward/downward compatibility with EDTV, HDTV and SDTV.

In addition, we define the notion of "compatible enough" as being compatible when a programmable decoder can be trivially altered to decode another standard even though the bitstream is not identical. All aspects of compatibility are defined below:

A bitstream is forward compatible if the new decoder can produce their full quality images when presented with data from an existing standard such as H.261 or MPEG-1.

A bitstream is backward compatible if there exists a subset of the data stream that can be processed directly by a standard decoder, for example, MPEG-1, H.261, coming US-FCC HDTV decoders, and others.

Upward compatibility is similar to forward compatibility. The goal is a bitstream that can be processed by a decoder designed for a higher data rate or higher resolution version of the same standard. An example, is an MPEG-2 HDTV receiver displaying MPEG-2 CCIR 601 transmissions.

Downward compatibility is the converse of the above and is covered in the definition of scalability.

The implementation subcommittee recognizes that there are many possible solutions to the various compatibility requirements including simulcasting, scalable bitstreams and programmable decoders. We also recognize that in some applications simulcasting may be appropriate.

At present, there is no consensus about what systems MPEG-2 must be compatible with or the optimal manner to do so. We therefore reserve discussion until a later date.

10. MPEG-1 Modes

There are a set of requirements used to define MPEG-1 that remain requirements of MPEG-2. Included are the following:

1) Bitstream editing. This is defined as concatenation of segments of various bitstreams to produce a new one;

2) Scan modes and slow motion. These are defined by reference to common VCR features. It is a requirement that the digital standard reproduce these features.

We also recognize that MPEG-1 does not support these modes perfectly in the core standard but provides for full support through the description in the annex to the standard of a set of recommended practices. Any tests that impact these features generally is obvious by argument and they are well understood by most workers. Therefore, there is no recommendation for test and development procedures in this area.

11. Implementation complexity

{Flexibility to allow both high performance/high complexity implementation and reasonable performance/low complexity implementation}

END