ITU Telecommunication Standardization Sector
Study Group 15
Experts Group for ATM Video Coding
(Rapporteur's Group on Part of Q.2/15)

Document AVC-578R
September 10, 1993

Source: RAPPORTEUR (Sakae OKUBO)
Title: REPORT OF THE THIRTEENTH EXPERTS GROUP MEETING IN BRUSSELS (September 2-10, 1993) - Part I and Part II
Purpose: Report

Part I General
Part II Sole sessions in Boston
Part III Joint sessions in New York (see AVC-554R)

Part I General

The twelfth meeting of the Experts Group was held in Brussels, Belgium, as follows;

- ITU-TS sole sessions during 2-3 September at the kind invitation of BELGACOM,
- Joint sessions with ISO/IEC JTC1/SC29/WG11 (MPEG) during 6-10 September at the kind invitation of Commission of the European Communities, DGXIIIB - Advanced Communication.

The list of participants appears at the end of this document.

Part II Sole Sessions

Contents

1. Introduction

The ITU-TS sole sessions were held at BELGACOM during 2-3 September 1993 at the kind invitation of BELGACOM. At the opening session, Mr. O. Poncin made a welcoming address on behalf of the hosting organization.

At the end of the sessions, Chairman thanked the hosting organization for the meeting facilities provided and the excellent secretarial support.

We had also a short closing session on 10 September under chairmanship of Mr. G. Morrison to review both sole and joint sessions as well as to prepare for the Daejeon/Seoul meeting.

Key words headed by • at the start of Section indicate items for discussion in this report.

2. Documentation (TD-2)

For the sole sessions, 26 AVC-numbered documents and 10 Temporary Documents have been made available as listed in Annex 1.

3. Tape demonstration (TD-3)

Two video tape demonstrations in D1 format were given to present experimental results as detailed in Annex 2 .

4. Review of the previous meetings

4.1 Sole sessions in Boston (AVC-553R)

There were no comments to the meeting report.

4.2 Joint sessions in New York (AVC-552,554R,555,556)

Chairman highlighted the following items which had been identified as needing responses of the ITU-T group at the New York joint sessions;

- H.261 compatibility
- Data Partitioning

5. Video source coding

5.1 VBV specifications (Annex C to AVC-555, AVC-563; TD-8)

- handling of 3:2 pulldown
- low delay operational mode (M=1)

The meeting supported the solution in AVC-563 for determination of the data removal timing just after an initial frame.

TD-8 indicated that the current VBV specification is not straightforward for the "low delay operational mode (M=1)" for 3:2 pulldown, but the meeting concluded that this is acceptable because such use would be very unlikely.

Mr. Haskell pointed out that there may be some bitstreams where it is not easy to simultaneously meet VBV in video and STD in systems

5.2 Error resilience by use of Data Partitioning (§3.1/AVC-554R,AVC-565,576,577)

- necessity for ATM environments
- description in the video coding standard (normative? informative?)
- description in the application standard (H.32X?)

The discussion is summarized as follows;

1) Tools currently available in the Video WD can provide cell loss resilience for the single layer coding up to around 10^{-3} CLR. Hence DP as a means of layered coding becomes useful when we use it under lossy environments of CLR higher than that value.

2) To secure the base layer, its bit rate should be constant. Hence, DP should allow fine rate control in the base layer requiring PBP change in the slice header.

3) Price advantage of splitting the bitstream into two channels with different QoS is not certain at this moment.

4) AVC-577 provides specifications for the Data Partitioning as a tool. Though it needs addition of slice layer syntactic elements, this addition does not affect the SP/MP/NP decoder implementation.

The meeting concluded that the Experts Group support the proposal of AVC-577 as a useful error resilience tool for our possible future use.

5.3 H.261 compatibility by use of spatial scalability (AVC-564,571)

- parameter setting in video syntax
- support of external sub-systems
- do we need it?

The H.32X terminal should be able to use the following video coding modes;

	Communication configuration	Video coding mode in H.32X terminal
1)	Point to point communication between two H.32X terminals	H.26X without embedded H.261
2)	Point to point interworking between an H.32X terminal and an H.320 terminal	H.261 (H.32X should operate in an H.320 emulation mode)
3)	Multipoint* among H.32X terminals through video servers such as MCUs	H.26X without embedded H.261
4)	Multipoint* among mixture of H.32X and H.320 terminals through video servers such as MCUs	The following possibilities should be considered; - two layer H.26X with embedded H.261, - H.26X without embedded H.261 (video server makes transcoding to/from H.261), - simulcast of H.26X without embedded H.261 and H.261

*Multipoint includes interactive services (e.g. videoconferencing) and distributive services (e.g. remote video presentation).

At this moment we can not reach a firm conclusion on the choice of the video coding mode for the multipoint operation because of uncertainties both in service and network developments. Hence we conclude that the solution of the "two layer H.26X with embedded H.261" should be covered by the generic video coding standard H.26X|MPEG-2 Video for our possible future use.

5.4 Video mixing bridge for continuous presence multipoint (AVC-501,559,560,575; TD-10)

- technical feasibility of the Bellcore method
- recommendations to WP1/15

AVC-575 is forwarded to SG15 as our reply with the following amendments;

- appreciation of the Bellcore challenge in this new video coding field
- delay implication of increasing HRD buffer size
- reflection of AVC-560
- QCIF only terminals which suffer from receiving lower quality pictures due to limited bit rate in the remote sending side

with conclusion that the current method of video mixing is not yet fully proven to be advantageous compared to the conventional transcoding approach. Annex 3 contains the outcome of this meeting.

Bellcore informed the meeting of the hardware system being experimented which uses PCs as platform, H.261 chipsets operating at up to 384 kbit/s, TCP/IP and Ethernet. A plan of replacing the Ethernet with N-ISDN was also informed.

5.5 Upsampling filters for spatial scalability (§2(3)/AVC-564,AVC-573)

- clarification of specifications in WD

The meeting took note of the experimental results with tape demonstration showing that the simplified upsampling filters does not affect the coding efficiency. It was stated that applicability of this simplification to sampling ratios other than 2:1 is yet to be examined.

Corrections to WD listed in §2(3)/AVC-564 should be conveyed to the joint sessions.

6. Network aspects

6.1 Requirements to H.22X/AAL (AVC-562,567,570)

1) Video clock recovery

AVC-562 provided experimental results showing that asynchronous operation between encoder and decoder are acceptable depending on applications. The meeting appreciated this input based on hardware experiments.

2) Strategy for detection/correction of bit error and cell loss

AVC-567 raised a fundamental question regarding the strategy we should take for bit errors and cell loss. The choice will affect the structure of H.22X/AAL. Since both of network performance and service requirements are uncertain, this is an area where we should apply good engineering sense. Members are requested to respond to this question at the next meeting.

3) PES semantics for **PES_packet_length** field

The meeting supported the proposal in AVC-570 to decouple the semantics of unspecified PES packet length with the nature of elementary stream in the interest of low delay requirement to H.22X/AAL.

At the closing session on 10 September, it was reported that MPEG Systems had not accepted this proposal. It was decided to establish a small correspondence group under Mr. Stuart Dunstan to continue the study. The mandate is to identify a suitable interworking point for MPEG-2 systems on packet oriented networks.

6.2 AALs

6.2.1 Asynchronous clock recovery (AVC-557)

- SRTS for PDH networks
- jitter performance of the adaptive clock method

The meeting became aware of the SG13 activities to elaborate AAL Type 1. There was a comment that asynchronous clock recovery can also be recovered by user methods such as time stamps of MPEG Systems.

6.2.2 Error correction in AAL (AVC-557,569)

- response to SG13's questions (interleaving method, polynomial)
- adaptation to VBR

The meeting agreed on acceptance of the SG 13's method for cell loss correction for delay sensitive signal transport of AAL Type 1; use of RS(94,88) with diagonal interleaving. We also agreed on adoption of the polynomial ($x^8+x^7+x^2+x+1$).

As to the closed form of interleaving in Figure 2/AVC-569, the meeting recognized that this implies 706 Bytes packet transmission. Members are requested to react to this proposal at the next meeting in the light of audiovisual communication applications.

It was clarified that the methods in Figures 1 and 2 in AVC-569 are not compatible because the time order of read out cells at the decoder are different.

6.2.3 Delivery of TS packets over ATM connection (AVC-568,569)

- alternatives
- evaluation items

We will continue the work comparing the alternatives listed in AVC-568 and the one proposed in AVC-569. AVC-569 method can be added to Table 2/AVC-568 as follows;

	Case 2' Use of FEC/CRC and interleave matrix
Cell loss correction	○
Cell loss detection	○
Bit error correction	○
Quick synchronization	○
Transmission efficiency	× 95.87 (47/48 x 188/192)
Generic AAL	△ (new type)
Transmission clock recovery	× (188x47=8836 bytes packet)
Variable bit rate	○

7. Profiles (AVC-572,574)

1) Minimum scalability profile

AVC-572 proposed a new profile (Main + SNR scalability) between Main and Next which is optimized for such applications as graceful degradation, WS/networks with different capabilities. The meeting concluded that we need more time to support this proposal because we should be careful to define new profiles.

It was recognized that proliferation of profiles is against the objective to allow cross application bitstream interchange, but at the same time too big a profile may not be implemented allowing eventual proprietary solutions optimized for particular applications.

Use of DP is another factor to affect this consideration; profile containing DP can not be in the onion ring structure.

2) Level parameters for the Next Profile

The meeting supported the AVC-574 proposal to define the maximum pel rate in terms of the sum of luminance and chrominance pels, since luminance and chrominance pels need the same resource in terms of processing power and memory size. An implication is that NP@ML can handle half SCIF (720x576x30, 4:2:0, progressive).

8. Work plan

8.1 Network standardization (AVC-557,558)

AVC-558 provides for the release time tables of network standardization.

We will continue to contribute to the use of IVS Baseline Document as a coordination tool by providing out study results.

8.2 Hardware trials (AVC-566)

- significance
- time schedule
- system configuration

The AVC-566 proposal to carry out field trails in mid-1995 was generally accepted. Ms. Conte and Mr. Schinkel stated relevant hardware projects in their organizations which might be extended to the proposed field trial.

The meeting concluded to make efforts to materialize the proposed plan, expecting this will assist in encouraging corresponding plans in member organizations.

8.3 H.320 adaptation to LANs (AVC-561)

- timescale
- initial activities for standardization (understanding the state of the arts)

AVC-561 provided a view that H.32X is a long term target but the needs are ambiguous but H.320 over LANs is an urgent short term target, thus the standardization program should reflect the situation. Since we already recognized the necessity to work on audiovisual systems over LANs at the previous meeting, an issue will be weighting on the ongoing ATM work and this new work. We will handle these in parallel for the moment; if it is appropriate, splitting the group may be a solution.

Mr. Morrison requested the participants to ask their commercial people what standards they require.

As a first action for the H.320 over LANs, members are requested to present the state of arts at the next meeting, particularly addressing protocol stack and interworking aspects

The meeting was called attention to that H.32X over ATM LANs should also be in the scope and that network interconnection is going on among ISDN, LAN and ATM.

8.4 Terms of reference of the group (TD-7)

- from pure video coding (H.26X) to total system (H.32X)?
- skeleton of H.32X

Chairman queried the members whether the terms of reference of the group be extended to H.32X from pure video H.26X. Since there was no negative comment expressed during the discussion, this was put forward to the consideration of WP1/15 held in the following week.

TD-7 presented a skeleton of Recommendation H.32X. Comments are welcome by correspondence.

9. Progress report to SG15 (TD-5)

- items needing guidance and/or consideration of SG15

The meeting reviewed the draft progress report and made some editorial amendments. This was presented at the WP1/15 meeting of the following week. Annex 4 contains the outcome of this week with subsequent slight editorial modifications made by WP1/15 and attached to its meeting report.

10. Joint sessions with MPEG

10.1 Documents

- information and discussion documents
- proposal documents

Conclusions obtained in the sole sessions will be put forward to the joint sessions;

Section # in this document	Item	Relevant document
§5.1	VBV	AVC-563
§5.2	Data Partitioning	AVC-577
§5.3	H.261 backward compatibility	none
§5.5	Correction to WD	§2(3)/AVC-564
§6.1 3)	**PES_packet_length** semantics	AVC-570
§7 2)	Definition of pel rate	AVC-574

10.2 Representatives

ITU-T EG	G. Morrison
Requirements	B. Haskell
Video	G. Bjoentegaard
Systems	D. Schinkel
Implementation	G. Morrison

11. Future meetings

Meeting	Date	Sole sessions	Joint sessions with MPEG	Joint sessions with SG13
14th	October - November 1993	October 27-29 in Daejeon (See AVC-535)	November 1-5 in Seoul	-
15th	March 1994	March 16-18 in Paris (?)	March 21-25 in France	March 14 in Geneva (?)

Synchronization of the meetings of this Experts Group and MPEG should be reviewed in the light of new work phase.

END

* * *

Annexes

Annex 1 Documentation
Annex 2 List of tape demonstrations
Annex 3 Video Coding Aspects of the Bellcore 4 QCIF to CIF mixing proposal
Annex 4 Fifth progress report of the Experts Group for ATM Video Coding

Participants of the thirteenth meeting of Experts Group for ATM Video Coding
2-10 September 1993 in Brussels

			S*	J*	
Australia	Mr. M. Biggar	Telecom Australia	X	X	CM
Belgium	Mr. O. Poncin	BELGACOM	X	X	CM
Korea	Mr. J-Y. Nam	ETRI	X	X	CM
USA	Mr. C-T. Chen	Bellcore	X	X	
	Mr. B. Gifrig	Tektronix		X	
	Mr. B.G. Haskell	AT&T Bell Labs	X	X	
	Mr. A. Luthra	Tektronix		X	(CM)
	Mr. P. Rao	CLI	X	X	
	Mr. G. Thom	DIS	X	X	(CM)
	Ms. A. Wong	Bellcore	X	X	
France	Mr. J. Guichard	CNET	X	X	CM
Italy	Ms. L. Conte	CSELT	X	X	CM
Japan	Mr. K. Asai	Mitsubishi	X	X	
	Mr. Y. Kosugi	Tokyo Electric Power Company	X	X	
	Mr. T. Murakami	Mitsubishi	X	X	
	Mr. Y. Nakajima	KDD	X	X	(CM)
	Mr. S. Okubo	NTT	X	X	Chairman
	Mr. K. Sakai	Fujitsu	X	X	
Norway	Mr. G. Bjøntegaard	NTR	X	X	(CM)
Netherlands	Mr. D. Schinkel	PTT Research	X	X	CM
UK	Mr. D. Beaumont	BT	X	-	
	Mr. G. Morrison	BT	X	X	CM
Sweden	Mr. P. Tholin	Telia Research	X	-	
	Ms. C. Verreth	Telia Research	X	X	CM

S: Sole sessions
J: Joint sessions with MPEG
CM: Coordinating Member
(CM): Substitute for CM

Documents for the Brussles meeting (2-10 September 1993)

Normal Documents

AVC number	MPEG 93/???	Pur-pose	Title (Source)
AVC-552	N0491	R	Test Model 6 (Test Model Editing Committee)
AVC-553R	802	R	Report of the twelfth meeting in Boston and New York (July 7-16, 1993) - Part I and Part II (Chairman)
AVC-554R	802	R	Report of the twelfth meeting in Boston and New York (July 7-16, 1993) - Part III (Chairman)
AVC-555	N0502	R	Third Working Draft - Video (WD Editing Committee)
AVC-556	N0501	R	Third Working Draft - Systems (WD Editing Committee)
AVC-557	-	R	Liaison statements to SG15 Experts Group for ATM Video Coding (SG13)
AVC-558	-	R	Report of drafting meeting on the Terms of Reference for JCG on B-ISDN (D. Dorman)
AVC-559	-	D	Comments on Bellcore continuous presence multipoint proposal (BT)
AVC-560	-	I/D	Some comments on video combining from 4 QCIFs to CIF (Bellcore)
AVC-561	-	P	Proposal for work plan of ITU-T WP1/15 (BT)
AVC-562	747	I	Experiments on asynchronous video clock (Japan)
AVC-563	748	P/D	Modification to the current VBV specification (Japan)
AVC-564	-	D	Backward compatibility with H.261 using NEXT Profile (Japan)
AVC-565	746	D	Problems in Data Partitioning (Japan)
AVC-566	-	P	Proposal for field trial of H.32X ATM video codec (Japan)
AVC-567	749	D	Bit error consideration on H.22X/AAL (Japan)
AVC-568	750	D	MPEG2 Transport packet transmission over ATM (Japan)
AVC-569	751	D	Consideration on the diagonal interleaving method for cell loss correction (Japan)
AVC-570	730	P	PES packet length semantics (Siemens)
AVC-571	-	P	H.261 backward compatibility (Telecom Australia)
AVC-572	-	P	Proposal for a profile incorporating minimum complexity scalability (Telecom Australia)
AVC-573	745		The effect of up-sampling on two-layered coding (BT?)
AVC-574	787	I	NEXT Profile Level Parameters for 4:2:0 formats (PTT Research)
AVC-575	-		Video Coding Aspects of the Bellcore 4 QCIF to CIF mixing proposal (PTT Research)
AVC-576	-	D/I	Error resilience based on Data Partitioning (BELGACOM, UCL)
AVC-577	-	P	Data partitioning addition to Working Draft (DSRC)

Abstracts

[???] indicates MPEG Document Number MPEG93/???.

AVC-552 [N0491] Test Model 6 (Test Model Editing Committee)

Three remaining experiments are described, optimization of the upconversion filter for spatial scalability, optimization of macroblock type tables for scalability, and verification of temporal scalability.

AVC-553R [802] Report of the twelfth meeting in Boston and New York (July 7-16, 1993) - Part I and Part II (Chairman)

This document records the outcome of the sole sessions held in Boston July 1993.

AVC-554R [802] Report of the twelfth meeting in Boston and New York (July 7-16, 1993) - Part III (Chairman)

This document records the outcome of the joint sessions with MPEG held in New York in July 1993.

AVC-555 [N0502] Third Working Draft - Video (WD Editing Committee)

This document gives frozen specifications for video coding part which were obtained at the New York meeting.

AVC-556 [N0501] Third Working Draft - Systems (WD Editing Committee)

This document gives frozen specifications for systems coding part which were obtained at the New York meeting.

AVC-557 [-] Liaison statements to SG15 Experts Group for ATM Video Coding (SG13)

This document contains the following liaison statements from SG13 which met during 5-16 July 1993 in Geneva;

- Open issues for AAL Type 1 timing recovery; impacts of Plesiochronous Network Operation, jitter absorbing capability of the Adaptive Clock method
- SG13's choice of the error correction method for delay sensitive signal transport of AAL Type 1 and an open issue - selection of the polynomial
- continuation of IVS Baseline Document activities

AVC-558 [-] Report of drafting meeting on the Terms of Reference for JCG on B-ISDN (D. Dorman)

Mandate, working methods, terms of reference, issues of immediate attention, etc., are described for a new joint coordination group "JCG on B-ISDN". Annex contains the release time table for B-ISDN standardization as a management tool for aligning B-ISDN related Recommendations.

AVC-559 [-] Comments on Bellcore continuous presence multipoint proposal (BT)

The method of achieving 2x2 split-screen continuous presence multipoint videoconferencing is commented with respect to use of fill bit, full screen display of pictures coded at 20 kbit/s, picture quality, delay, implementation cost of header manipulation etc. other than pure picture coding, impacts of picture dropping, video clock synchronization among remote terminals, use of existing terminals. It is concluded that the technique is not yet proven and may always suffer from disadvantages compared to the pel domain combining method.

AVC-560 [-] Some comments on video combining from 4 QCIFs to CIF (Bellcore)

The following is presented;

- There is a tradeoff between adding more MCU delay and violating the HRD sooner.
- Frame synchronization can be maintained by mapping input QCIF Temporal Reference on to regular output CIF Temporal Reference with a certain rule.
- Comparison between video combining and transcoding should carefully be studied in terms of picture quality, delay increase, cost and system complexity.

AVC-561 [-] Proposal for work plan of ITU-T WP1/15 (BT)

Taking into account of uncertainty in ATM network performance and timescales, a dearth of real and viable applications, and increasing number of customers asking for H.320 over LANs, this document proposes that H.320 over LANs become a prime focus for the work of SG15 to meet the shorter term needs.

AVC-562 [747] Experiments on asynchronous video clock (Japan)

Measurement and observation results are reported for asynchronous operation of a set of hardware HDTV coder and decoder where the decoder clock frequency is 10/20/30 ppm higher than the encoder one. A field slip appears as one-line downward or upward shift of the reproduced picture. Based on the fact that this one-line shift is seldom visible, it is concluded that asynchronous operation is of some practical use.

AVC-563 [748] Modification to the current VBV specification (Japan)

For 3:2 pulldown VBV operation, a missing specification is provided for the data removal timing at the start of sequence using **vbv_delay** values. A note for attention is also given for the coding of **number_of_field_display_code** at M > 1.

AVC-564 [-] Backward compatibility with H.261 using NEXT Profile (Japan)

Necessary provisions are discussed to implement H.261 backward compatibility by using the spatial scalability in Next Profile; parameter setting in the video syntax, layer indication in the H.22X multimedia multiplex, H.24X capability exchange for the combination of upper layer and lower layer. Some video issues are also listed; VBV, 5 bit TR in H.261, upsampling filter description and indication of progressive/interlace and chroma format for the lower layer. It is concluded that at least he following should be clarifies;

- handling of picture skipping in the lower layer,
- VBV specifications for the upper layer,
- identification of the layer by Systems.

AVC-565 [746] Problems in Data Partitioning (Japan)

Three alternatives for Data Partitioning syntax are compared; 1) DP elements in Sequence, Picture and Slice headers, 2) Extension syntax, 3) Informative description. Necessity of fine rate control in each layer leads to the first alternative. This document also points out possible start code emulation for which inhibition of the layer consisting of priority class 4 is concluded as the most reasonable. Issues of **VBV_delay** and **bit_rate** are also raised.

AVC-566 [-] Proposal for field trial of H.32X ATM video codec (Japan)

A field trial of H.32X terminal interconnection through ATM transmission lines is proposed to verify related draft Recommendations and to encourage early development of such products. Middle of 1995 is suggested for performing the field trial. Specifications looking at the early realistic applications and test methods needing special tools for ATM transmission are also discussed.

AVC-567 [749] Bit error consideration on H.22X/AAL (Japan)

Impacts of bit error are raised to reconsider H.22X/AAL for audiovisual services. Three strategies are discussed; 1) bit error detection and cell loss detection, 2) bit error correction and cell loss detection, 3) bit error correction and cell loss correction. A suggestion is given that the second strategy is preferable to the first one in some situations, though the Boston discussion was based on the first strategy.

AVC-568 [750] MPEG2 Transport packet transmission over ATM (Japan)

The following four methods are compared for transmission of TS packets over an ATM link;

- Byte/bit data transfer (bit pipe approach, AAL Type 1)
- structured data transfer (bit pipe approach, AAL Type 1)
- FEC and interleaving (bit pipe approach, AAL Type 1)
- alignment between a TS packet and ATM cells (packet approach, AAL Type 0 or new Type)

It is concluded that the third method would be a good solution if cell loss correction is necessary but the fourth method if bit error correction is sufficient.

AVC-569 [751] Consideration on the diagonal interleaving method for cell loss correction (Japan)

This document discusses the use of Reed-Solomon code and interleaving for cell loss and/or bit error correction and proposes the following for AAL options;

- RS(94,88) and diagonal interleaving with ring buffer (SG13 proposal) for the lowest delay - bit by bit delivery,
- the same as above but with double buffer for application to the "stop-go" VBR operation - 752 Byte packet delivery,
- RS(192,188) and vertical interleaving for TS packets - 192x47 Byte packet delivery.

AVC-570 [730] PES packet length semantics (Siemens)

In order to achieve low delay, zero value of **PES_packet_lentgh** is proposed to be unspecified even if it is used without support from either the Program Stream or Transport Stream. It is also proposed that this should be true for any elementary streams.

AVC-571 [-] H.261 backward compatibility (Telecom Australia)

Due to the spatial scalable solution constraining the upper layer coding and other reasons, it is proposed that the ATM Video Coding Experts Group adopt a switchable/simulcast strategy to provide interworking between H.32X and H.320 terminals in those situation or at those times when this is required.

AVC-572 [-] Proposal for a profile incorporating minimum complexity scalability (Telecom Australia)

Considering that full support of Next Profile functionalities may be costly, this document proposes definition of a minimum scalability profile, incorporating those features currently in Main Profile, but also allowing 2-layer SNR scalability.

AVC-573 [745] The effect of up-sampling on two-layered coding (BT)

Experimental results are reported for comparison between the TM up-sampling filtering and linear interpolation. It is concluded that the up-sampling filtering process for the lower layer coded picture in a two-layer coder can be simplified to linear interpolation without any significant effect on the upper layer picture quality.

AVC-574 [787] NEXT Profile Level Parameters for 4:2:0 formats (PTT Research)

Taking into account that the processing power and memory size do not differ between luminance and chrominance pels, this document suggests that the level parameter should address total pixel rate (= luminance pixel rate + chrominance pel rate). It is also suggested that the level parameters for Next Profile should make distinction between 4:2:2 and 4:2:0 since progressive pictures are effective for high quality videoconferencing.

AVC-575 [-] Video Coding Aspects of the Bellcore 4 QCIF to CIF mixing proposal (PTT Research)

This document summarizes various comments to the Bellcore 4 QCIF to CIF video mixing proposal in AVC-500 and 501 which were obtained through correspondence. The comments cover questions to be answered, service aspects, approaches to the continuous presence multipoint system, impact on equipment, buffering and delay, picture quality, etc. It is concluded that the proposed method is simple in its basics and does not degrade picture quality, but there are some drawbacks.

AVC-576 [-] Error resilience based on Data Partitioning (BELGACOM, UCL)

Experimental results for cell loss resilience are reported for two different sets of PBP in terms of information splitting ratio and picture quality. It is raised whether we want the PBP information included in the H.26X bitstream at the slice level to allow constant bit rate mode in each of the two layers.

AVC-577 [-] Data partitioning addition to Working Draft (DSRC)

This document proposes syntax and other descriptions for the use of Data Partitioning.

Temporary Documents

TD-1	Chairman	Agenda for the sole sessions
TD-2	Chairman	Available documents for the Brussels meeting
TD-3	Chairman	List of tape demonstrations
TD-4	Chairman	List of open issues - August 1993
TD-5	Chairman	Fifth progress report of the Experts Group (Draft)
TD-6	S. Okubo	ITU-T approval process for Recommendation H.26X|MPEG-2 Standard (MPEG93/752)
TD-7	Chairman	Skeleton of Recommendation H.32X
TD-8	K. Sakai	Supplements to AVC-563
TD-9	Chairman	Agreements
TD-10	Chairman	Video coding aspects of the Bellcore 4 QCIF to CIF mixing proposal
TD-11	Chairman	Agreements and actions (substitute of the meeting report)

END

List of Tape Demonstrations
(2 September 1993, Brussels)

No	Organization	Topics	Tape	Doc.
a	BT Labs	Spatial scalability up-sampling	D-50	AVC-573
b	BELGACOM	Cell loss resilience by Data Partitioning	D-50	AVC-576

INTERNATIONAL TELECOMMUNICATION UNION
TELECOMMUNICATION STANDARDISATION SECTOR
STUDY PERIOD 1993-1996

Temporary Document 43 (15/1)

Geneva 7 - 17 September 1993

Questions: 2/15, 3/15

Source: Experts Group for ATM Video Coding.

Title: Video Coding Aspects of the Bellcore 4 QCIF to CIF mixing proposal

In this paper a response from the Experts Group for ATM Video Coding to the Bellcore 4 QCIF to CIF mixing (Documents AVC 500 and AVC 501) is given as formulated in the meeting of this group on 2 and 3 September 1993 in Brussels.

1. Introduction

This discussion is in response to the request of the Tokyo Rapporteur's meeting. A brief review of the documents at the Boston meeting gave the following comments (see AVC-553R);

- Technical feasibility should be examined in the light of overall system configuration. Existing terminals should be accommodated without modification as far as possible.

- Comparison should be made against transcoding MCU in terms of picture quality, delay and other factors as listed in AVC-501.

The meeting decided to continue the study toward the next meeting and the SG15 meeting (both in September) through correspondence.

The task is to draft a reply to WP1/15 with respect to video coding techniques which are necessary for terminals and MCU to constitute the video mixing multipoint system.

2. Questions to be answered

The question we should answer is whether the video coding solution proposed by Bellcore as in AVC-501 (or modified if necessary) can be practical so that SG15 can recommend this method for the continuous presence multipoint system. More specifically,

1) What does this method require for the video coding/decoding in the terminal?

2) What does this method require for the video signal processing in the MCU?

3) What system performance can we obtain in terms of picture quality and delay? Actually unless requantization process is involved, the picture quality of each constituent QCIF picture is maintained even after video mixing (but motion rendition may be affected?) Our prime concern will be

buffering delay at the output of the video mixer.

4) What can be gained with this method? This is a more service related question, but can not be separated from the video related questions.

3 Service Aspects

When participating in a conference it is very helpful to be able to see all the participating sites, but in many situations conversation is carried on between two parties with other parties listening in. During the conference the dominant parties may change and at times several parties may contribute, but not all the time. Ideally, each site would be able to see all other sites to have a feeling of their presence, whilst also seeing the current speaker in more detail. So for example the current speaker would appear as a CIF image on one screen and up to 4 locations would appear combined in 1 CIF picture on a second screen.

Most future terminal equipment in multipoint conferences are likely to be workstation or PC based systems. These will provide dialogue functions via a "common screen" or a "shared screen" and just seeing the other partners will be only one aspect of the conference.

Demonstration systems for multipoint video conferences between 4 to 6 partners as shown by different companies (e.g. DEC spin, SUN spin) already show that participants can be presented with sufficient visual quality in relatively small formats as 'background arrangements', leaving room for a common object of interest in the form of:- a document, graphic, dialogue cursor/pointer or for the actual speaker.

Future MCU's are required to support:
- the dialogue control
- dialogue (foreground image switching on demand on speaker control)
- common screen encoding.

4. Approaches to the multipoint Systems

There may be two approaches;

1) Existing terminals are kept intact and all the necessary measures are provided by the MCU. The current H.231/243 system is based on this approach. Though there have been defined some enhanced functionalities such as chair control on an optional basis, existing H.320 terminals can participate in multipoint communications without any change.

2) We introduce a new terminal design for this purpose. Hence, only new generation terminals can take part in the continuous presence multipoint system.

There is a significant difference between the two approaches from a service point of view. The Bellcore approach seems to be close to the first one, but the terminal is required to operate in R/4R asymmetric video rates. This functionality is not supported by the existing H.320 terminals.

5. Impact on equipment

Although the use of the fill bit requires no changes to H.261 as a standard, it does constitute a significant change in the way that existing H.261 terminals use the fill bit. Normally the fill bit is only asserted when buffer under flow is imminent. For this reason and for others discussed later in this document it is not possible to implement this proposal with existing terminals without modification to those terminals. It is likely that some existing terminals and H.261 chip sets will not be capable of using the fill bit to deliberately reduce the active video rate.

The proposal requires the introduction of new capability and command codes that must be transmitted and understood by the video terminals. This precludes the use of existing terminals without an update, if possible, to their H.242/H.230 software. As mentioned above the terminals also have to be modified to use the error corrector fill bit for quarter rate video.

6. Some observations

6.1 Buffering and hypothetical reference decoder

There is an extreme case where the video mixing as in AVC-501 can work properly;

- Picture frequencies are locked to each other among the terminals.
- Coded picture rates of different input QCIFs are exactly the same (30Hz, 15Hz, 10Hz, or 7.5Hz) and constant.
- Number of bits per QCIF frame are the same throughout each sequence and among the sequences.

In this case it is obvious that the video bridge can multiplex 4 QCIF coded data into a CIF coded data with minimum buffering delay (one picture time delay may be incurred to accommodate different phases of the input pictures). This case is illustrated in the following diagram with an assumption that all the input QCIFs are synchronised in picture phase.

Unit 1 indicates number of bits per QCIF picture. The number of bits of each output CIF picture is exactly 4 units. In this case the video mixing buffer will work with no problems and no delay is involved. HRD specification is also met, thus the existing CIF decoder can decode this split-screen picture.

```
     ---------------------------------------------------------
In-1 | 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1|
     ---------------------------------------------------------
In-2 | 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1|
     ---------------------------------------------------------
In-3 | 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1|
     ---------------------------------------------------------
In-4 | 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1|
     ---------------------------------------------------------

     ---------------------------------------------------------
Out  | 4| 4| 4| 4| 4| 4| 4| 4| 4| 4| 4| 4| 4| 4| 4| 4| 4| 4| 4|
     ---------------------------------------------------------
```

A next extreme, but not the most extreme, case is as before but In-4 includes a scene cut, thus intra picture with 12 unit of coded data bits.

```
     ------------------------------------------------------------
In-1 | 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1|
     ------------------------------------------------------------
In-2 | 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1|
     ------------------------------------------------------------
In-3 | 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1| 1|
     ------------------------------------------------------------
In-4 | 1|              12                | 1| 1| 1| 1| 1| 1|
     ------------------------------------------------------------
     ------------------------------------------------------------
Out  | 4| 3| 3| 3| 3| 3| 3| 3| 3| 3| 3| 3|15| 4| 4| 4| 4| 4| 4|
     ------------------------------------------------------------
Out  | 4|3*|3*|3*|3*|3*|3*|3*|3*|3*|3*|3*|15        | 4| 4| 4| 4| 4| 4|

      * stuffing bits
                                                <------->
                                                  delay
```

In this case, the output picture with 15 unit of coded data causes stationary delay of 15/4 picture time. Furthermore, the 4 picture time HRD specification will be violated if the intra picture becomes larger than 16.

There are complicated situations in between as combination of asynchronous nature of input pictures and different and time varying coded picture rates (within a limitation of defined minimum coded picture interval) among input pictures. The stationary video processing delay may vary from almost zero to the worst case mentioned above. Unfortunately there is no time to analyse these intermediate situations.

An essential problem is that the video mixer is passive, hence it has very limited power to control the output buffer occupancy or bit generation there. Once a large picture arrives, it causes corresponding stationary delay at the output of the video mixer. In case of point-to-point communications, we can isolate the large picture delay by using picture skipping, thus keeping stationary state delay small.

6.2 Synchronisation of the data streams

The video data streams have to be synchronised before they can be mixed. In order to synchronise the video data streams the mixer has to wait for new picture start codes in each of the involved data streams and then has to wait until it has received the complete pictures before the mixed pictures can be redistributed. Or it can send a new picture when it has received a new full picture in one of the channels. In both cases you need large buffers in the video mixer because:

1. The pictures in the different video data streams do not start simultaneously but have offsets to each other.

2. The length of the pictures in the different video data streams may vary due to the fact that

a) the contents of the pictures are different, this results in different compression ratios,

b) one picture may be coded in INTRA mode, whereas others are coded in INTER mode,

3) All video pictures which contain INTER-coded blocks can not be left away because they are necessary for the reconstruction at the receiving side.

A sophisticated management mechanism is needed, and requires large buffers.

This leads to a large delay of the video data.

In any case the rate of the CIF output bit rate has to be higher than the sum of the four incoming bit rates to compensate for the stuffing.

6.3. More delay considerations

The simulation studies paper demonstrates end to end delays for a 256 kbit/s video bit-rate (64 kbit/s per QCIF source) of between 266.67 and 644.44 ms. The mean delay appears to be approximately 440 ms. At 256 kbit/s the end to end delay for a single CIF encode and decode is approximately 260 ms and does not vary much with picture content. This gives a back to back delay of 520 ms without any optimisation. The proposed method of combination therefore appears to give only a marginal improvement in end to end delay and in some cases causes a degradation.

6.4 Effect of the distribution of FUR's

It is not clear whether the simulations performed (by Bellcore) have taken into account the dropping of frames by the source encoders. At low bit rates it is quite normal for an H.261 encoder to drop frames rather than quantise a picture more heavily in order to limit the coded bit-rate. The dropping of frames results in those frames that are transmitted being allocated more bits than average. The effect of this will be similar to the intra coded frame in response to a fast-update request (FUR). Unlike the FUR the timing of the dropped and transmitted frames can not be controlled. It is quite likely that all 4 QCIF sources will drop frames at the same time and then start to code pictures with a large number of bits in phase. This has been shown to result in longer end to end delays. It is questionable whether there would be much gain in distributing the FURs in this case.

The dropping of frames also has an implication on the way the 4 QCIF pictures are combined. The Bellcore proposal suggests that the Temporal References from all but 1 source are ignored and that pictures are combined as they arrive to achieve minimal delay. With dropped frames this could result in temporal distortion.

6.5 Synchronisation of terminal Clocks

Document AVC-501 Section 4.2 states:

"The videoconferencing terminal clocks can be locked to the network clock so that data transmission is synchronous between the terminals and the network. These synchronous terminal clocks also ensure the frame rates of different QCIFs are exactly the same but their frame phases may be different."

Such locking is not common-place in terminal designs. We are not aware of any which incorporate it. While it is technically feasible, it does place restrictions on the selection or design of cameras. For integrated terminals these may be under the control of the designer, but for additional plug in cameras and other sources such as video tape playback machines, the locking requirement is a drawback.

7. Picture quality

The simulation study paper (AVC-501) concedes that although the pel-domain combining results in slightly degraded picture quality, the degradation is insignificant and almost undetectable subjectively. This conclusion is based on comparing the picture quality before and after a decoded picture has been re-coded and decoded a second time. A more relevant comparison might be between the following two cases. The first is a double encode and decode entailing an initial encoding and decoding using CIF at the full bit rate available from a terminal followed by coding and decoding with QCIF at one quarter of the bit rate. The second case is the single encode and decode with QCIF and quarter bit rate which corresponds to the Bellcore proposal. It is expected that the objective and subjective degradations introduced by the double encoding will be significantly less than those observed in the Bellcore experiments.

One of the features of the video mixer is that it does not involve picture quality degradation at MCU. Instead of using rate R toward the MCU, we can use rate 4R for the transcoding MCU. This may reduce picture quality degradation due to transcoding. It is quite obvious that if the rate of the channel toward MCU is very high, there is no degradation involved at all.

The comparison should be between the following;

```
                                  rate R
Video mixer    Coder =================================> Decoder

                            rate 4R                 rate R
Transcoder     Coder =======> Decoder -> Coder ======> Decoder
```

The chosen rate should range from 19.6 kbit/s up to 480 kbit/s (ISDN rates 2B, 6B and primary access).

The transcoding MCU presumably places no limitations to the operating mode of the terminal. If the picture quality and delay are comparable, this can be a practical solution as well for continuous presence multipoint systems.

With the Bellcore proposal only a low bit rate QCIF picture is available from each terminal. For a 2B conference the video rate, assuming 48 kbit/s G.722 audio, no data and no encryption, is 78.4 kbit/s. Each individual QCIF rate will be limited to 19.6 kbit/s. The display of this picture as a full size image on anything other than a small screen would not be acceptable. (This can be confirmed by pictures seen at the meetings of the Rapporteur for Very Low Bit Rate Coding).

Experiments carried out SEL Alcatel using this method on 1B (R = 17.2 kbit/s), lead to unacceptable results, because the four QCIF images gave an image quality which was too poor for the multipoint service, because the frame rate dropped below 1 Hz.

8 ALTERNATIVE METHODS

The hardware to support the combination of pictures at the GOB level may not be as complex as that required to decode the picture and re-code it but all of the complexity saving lies in the compression process for which chip sets and complete codecs are readily available. (Line terminating, clock extraction, data extraction etc. are necessary for both approaches.) The cost of an H.261 encoder and decoder is decreasing all the time, and soon its marginal cost over finding and manipulating the picture and GOB headers and controlling the input buffers will be insignificant in the overall consideration. It is arguably a more complex task to build a 4 QCIF to CIF combiner than to put a decoders back to back with an encoder through a 4:1

spatial decimator/combiner.

9. Impact of the method

1) What does this method require for the video coding/decoding in the terminal?

The video part of the terminal (H.261) has to be changed to asymmetric mode. When delay needs to be reduced definition of Hypothetical Reference Decoder (HRD) needs to be changed. The method will not give an acceptable solution for QCIF only terminals.

2) What does this method require for the video signal processing in the MCU?

The proposed method avoids decoding, recoding. The mixing can be carried out in the H.261 domain and is relatively simple. However today's technology provides single chip solutions for H.261 encoders and decoders.

3) What system performance can we obtain in terms of picture quality and delay?

Delay:

If the definition of hypothetical reference decoder is (HRD) not changed, the method proposed by Bellcore gives a delay which is the same range as transcoding (or might in some cases even be higher). Changes to the HRD move the delay from the MCU to the decoder.

Picture quality:

Although intrinsically the proposed method gives no degradation, the limitation to QCIF and a bit rate of R in stead of 4R. A fair comparison has not been shown yet. Experiments on 1B (total bit rate 64 kbit/s) carried out by SEL Alcatel gave an unacceptable picture quality (frame rate dropped below 1 frame/s). A 2B operation (128 kbit/s using G.722 for audio) would give an unacceptable quality (extrapolation from results shown in the meetings of the Rapporteur for Very Low Bit Rate Coding)

4) What can be gained with this method?

Obviously continues presence is a highly desirable feature in multipoint conferencing. For the service, however, other advanced display methods such as e.g.

- a common screen (in high quality), or
- a high quality picture of the person who is speaking combined with a lower quality picture of the other participants

may be preferable. The proposed method does not give the flexibility required for these display method. A transcoding method can provide these features with no limitations.

10. Conclusion

The method proposed by Bellcore is simple in it's basics. It requires small modification to the definition of the terminal. The MCU can do a relatively simple mixing in the H.261 domain. The method gives no degradation due to transcoding. Therefore it was very valuable to deeply look into this unused hook of H.261.

Drawbacks of the method are:

- There is no proof that the method is better than a pel domain combining.
- Changes to the H.261 terminal equipment that already exists is a problem.
- Technology development makes transcoding and pel domain mixing cheap.
- The method still gives substantial delay.
- The method gives no flexibility in bit rates and to the users.
- The method gives no solution for QCIF only terminals.

END

INTERNATIONAL TELECOMMUNICATION UNION
TELECOMMUNICATION STANDARDIZATION SECTOR
STUDY PERIOD 1993-1996

Temporary Document 42 + 77(15/1)
Revised

Geneva, 7-17 September 1993
Questions: 2/15, 3/15

SOURCE: RAPPORTEUR FOR Q.2/15
TITLE : FIFTH PROGRESS REPORT OF THE EXPERTS GROUP FOR ATM VIDEO CODING

----

1. General

We met four times since we had presented the fourth progress report (Annex 4 to COM XV-R 105-E) at the previous Working Party XV/1 meeting in November 1992;

Meeting	Sole Sessions	Joint sessions with MPEG
10th meeting in Italy	21-22 January 1993 in Torino at the kind invitation of CSELT	25-29 January 1993 in Rome at the kind invitation of the Ministry of Posts and Telecommunications
11th meeting in Australia	5-7 April in Melbourne at the kind invitation of Telecom Australia	29 March - 2 April 1993 in Sydney at the kind invitation of Standards Australia
12th meeting in USA	7-9 July 1993 in Boston at the kind invitation of PictureTel	12-16 July 1993 in New York at the kind invitation of Columbia University
13th meeting in Belgium	2-3 September 1993 in Brussels at the kind invitation of BELGACOM	6-10 September 1993 in Brussels at the kind invitation of Commission of the European Communities, DGXIIIB - Advanced Communication

Our meetings have been synchronized with those of ISO/IEC JTC1/SC29/WG11 (MPEG) since May 1991. We have held joint session with MPEG as well as ITU-T sole sessions. Lists of participants of the four meetings appear in Annex to this report *(omitted here).*

This document reports major achievements toward defining Recommendation H.26X for video coding in the ATM environments and its surrounding network aspects, and particular items for the consideration of Study Group 15.

2. Overall work plan

Recommendation H.26X is intended to provide a generic video coding standard which serves various services such as communicative, distributive, retrieval and messaging services in ATM environments. SG15 decided that this Recommendation should be developed as common text standard H.26X|MPEG-2 between ITU-T and ISO/IEC.

The current work is proceeding as planned; technical specifications were frozen in July 1993 toward final approval of the Recommendation early 1995. See Figure 1 below for more details.

3. Source coding

3.1 Test Model development

Test Model is a reference video coding scheme against which proposed improvements are tested. If test results are promising, those improvements are incorporated into the next generation Test Model for further comparison.

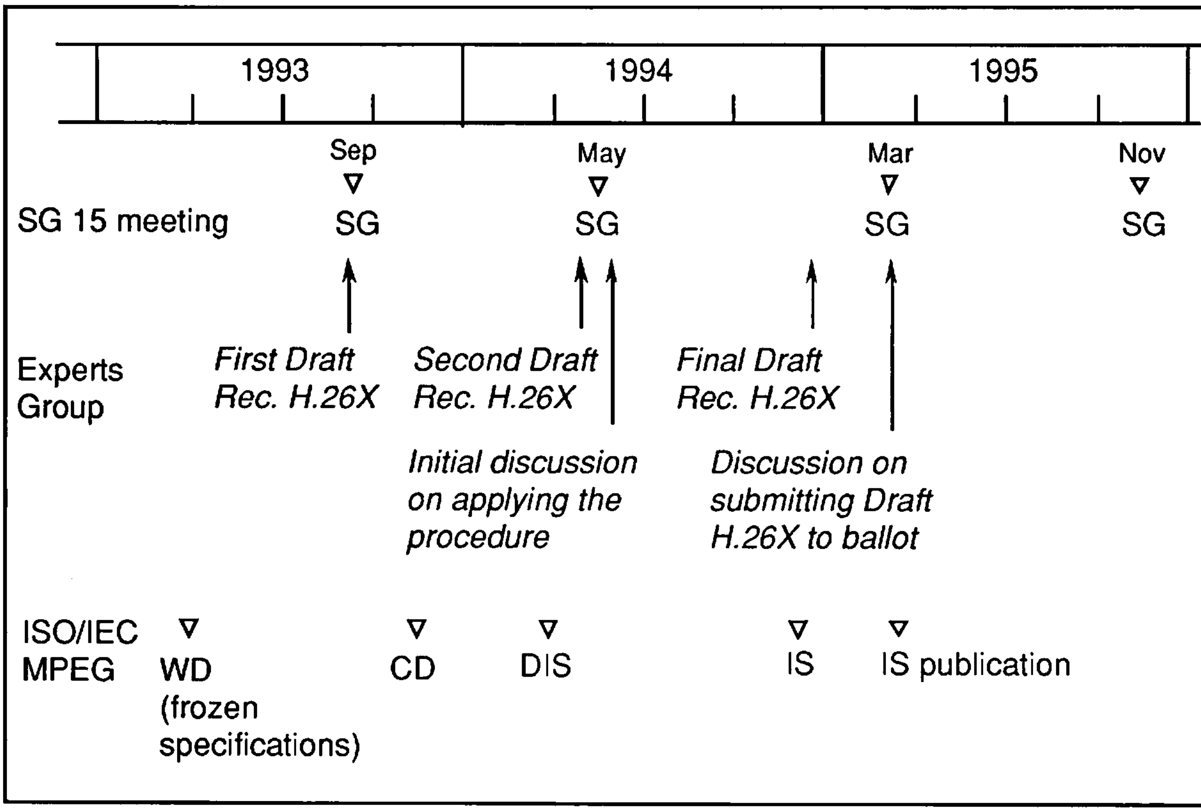


Figure 1 Approval process of Recommendation H.26X|MPEG-2 standard

Since the first Test Model (TM1) had been defined in March 1992, it evolved into TM5 in April 1993. This TM document includes definition of the coding algorithm as well as necessary experiments .

3.2 Basic coding algorithm

The basic coding algorithm is a hybrid one consisting of motion compensated interframe prediction to remove temporal redundancy and DCT to remove spatial redundancy. Motion compensated interframe prediction includes both of forward and backward predictions. This algorithm has been proven effective for wide range of picture formats and bit rates.

3.3 Improvements of coding efficiency

Particular efforts have been made to improve coding efficiency for interlaced signals which are common in the current television broadcasting. The following is a list of such techniques adopted in the standard;

- field/frame adaptive motion compensated prediction,
- field/frame adaptive DCT,
- special prediction (dual-prime).

In the quantization process, a new non-linear quantizing step size assignment has been defined to allow finer quantization control for high quality coding. A new VLC table has been adopted for intra pictures, and a new DCT coefficient scanning method has also been adopted to cope with field pictures.

3.4 Functionalities

In order to make the generic coding standard applicable to as many applications as possible, the following functionalities have been incorporated;

1) Multiple formats

The coding algorithm can handle interlaced or progressive formats of various sizes.

2) Low delay operational mode

A cyclic refreshing method called "intra-slice" has been devised to reduce buffering delay. This technique refreshes one or more strips of macroblocks at a regular interval. Picture skipping is also introduced to cope with large pictures containing a number of bits due to e.g. a scene change. Special prediction (dual-prime) is effective for the low delay operation.

3) Random access/channel hopping

Fast access can be implemented by selecting refresh parameters.

4) Scalability

Resolution, SNR and temporal scales are supported by spatio-temporal prediction and re-quantization. The resolution scalability technique uses the enlarged low resolution picture as a possible prediction mode. Hence if the decoder uses only the base layer bitstream, small size pictures (quarter size in the Example of Figure 2) are reproduced while the two bitstreams are used together, full size pictures are obtained. If the base layer uses existing coding standards and formats, then this spatial scalability techniques provides backward capability if multiple bitstreams are supported by the system.

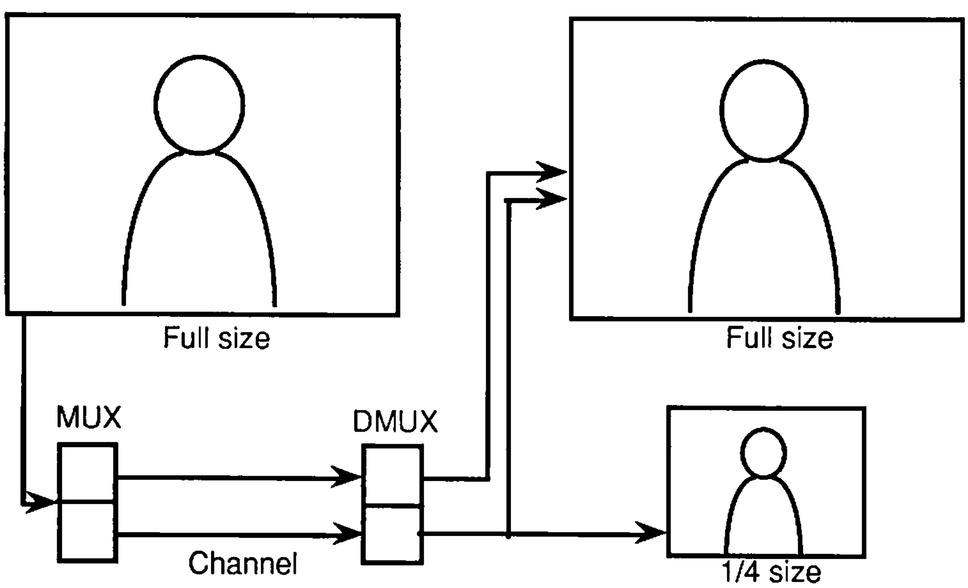


Figure 2 Spatial resolution scalability

SNR scalability provides two different versions of the same spatial resolution picture with different level of quantizing noises. This may facilitate graceful degradation when transmission channels become vulnerable to errors. The temporal scalability is useful for migration from interlaced television to progressive television. Prediction errors from the interlaced signal to the progressive signals are transmitted as an enhancement signal to the encoded interlace signal.

5) Cell loss resilience

The developed coding standard can support modest level of cell loss resilience by selection of appropriate cyclic refresh parameters and slice sizes in addition to concealment techniques which replace damages areas with motion compensated surrounding blocks. Motion vectors for intra macroblocks can optionally be sent for this purpose.

If we need stronger cell loss resilience for CLR higher than 10^{-3} (Note), we need some layered coding. Above mentioned SNR scalability is a tool to achieve this property. Another tool is data partitioning which split DCT coefficients into high and low priority data. Low priority data are much less sensitive to cell loss than high priority data.

Note - It should be noted that this figure is nothing more than a very general guide. Layered coding may be the most appropriate choice in some circumstances at lower CLRs, and the error resilience performance of any system depends on the tools used, application, choice of parameters, detailed error conditions and the picture materials.

3.5 Profiles and Levels

H.26X is intended to be generic in the sense that it can serve a wide range of applications, bit rates, resolutions, qualities and services. Applications should cover, among other things, digital storage media, television broadcasting and communications. In the course of creating this standard, various requirements from typical applications have been considered, necessary algorithmic elements have been developed, and they have been integrated into a single syntax so that this standard will facilitate the bitstream interchange among different applications.

Considering practicality of implementing the full specifications of this standard at early stages, however, a limited number of subsets are also stipulated by means of "profile" and "level".

Table 1 Profile/Level structure of H.26X|MPEG-2 Video
(Upper bounds of parameter values)

High (up to 60 Mbit/s) Note 4	Pels/line Lines/frame Frames/sec Pels/sec	1920 1152 60 62.7 million	1920 1152 60 62.7 million	1920 1152 60 62.7 million
High-1440 (up to 60 Mbit/s) Note 4	Pels/line Lines/frame Frames/sec Pels/sec	1440 1152 60 47.0 million	1440 1152 60 47.0 million	1440 1152 60 47.0 million
Main (up to 15 Mbit/s) Note 4	Pels/line Lines/frame Frames/sec Pels/sec	720 576 30 10.4 million	720 576 30 10.4 million	720 576 30 11.06 million (Note 2)
Low (up to 4 Mbit/s?)	Pels/line Lines/frame Frames/sec Pels/sec	352 288 30 2.53 Million	352 288 30 2.53 Million	Not decided
Level (Note 1)		Simple 4:2:0 single layer (Note 5)	Main 4:2:0 single layer	Next 4:2:2 scalable
			Profile	

Note 1 - Level for the Next profile indicates the upper layer of the two resolution scales.
Note 2 - 720x512x30 has been considered to accommodate 483 active lines of 525/60 TV. The extension of the upper bound of the pel rate to cover half SCIF is an open issue.
Note 3 - Multiples of 16 up to the upper bound are supported in number of pels per line and number of lines per frame.
Note 4 - Data rates for other than MP@ML are to be determined .
Note 5 - Main minus B-pictures

A "profile" is a defined sub-set of the entire bit stream syntax that is defined by this standard. Within the bounds imposed by the syntax of a given profile it is still possible to require a very large variation in the performance of encoders and decoders depending upon the values taken by parameters in the bit stream.

In order to deal with this "levels" are defined within each profile. A level is a defined set of constraints imposed on parameters in the bit stream. These constraints may be simple limits

on numbers. Alternatively they may also take the form of constraints on arithmetic combinations of the parameters.

For each defined subset (a profile at a certain level), conformance testing will also be defined.

The current structure of Profiles and Levels are shown in Table 1 with upper bounds of parameter values.

3.6 VBR

Study on the VBR coding has been continued to investigate its performance under Usage Parameter Control. One is to experiment on the relationship between leaky bucket counter size and encoder buffer delay and its relevance to cell interleaving method for cell loss correction. Another is to experiment on the utilization of periodic coding structure (I/B/P pictures) to obtain low delay in the sliding window environments.

SG13 indicated in its liaison response that UPC time constant in the range of seconds or more is not likely implemented in the network. This UPC time constant affects the VBR effectiveness for video coding; if it is not sufficiently long compared to the picture period, we may not be able to fully exploit the advantages of VBR.

3.7 Working Draft for H.26X|MPEG-2

As the outcome of the New York meeting, the third Working Draft for H.26X|MPEG-2 was produced as in a separate document TD 39 (15/1).

4. System aspects

4.1 Audiovisual system configuration

According to the practice of N-ISDN audiovisual standardization, Recommendations in Figure 3 are assumed. Dotted partitions between multimedia multiplex and AAL etc. are intended to represent the open nature of the current discussion.

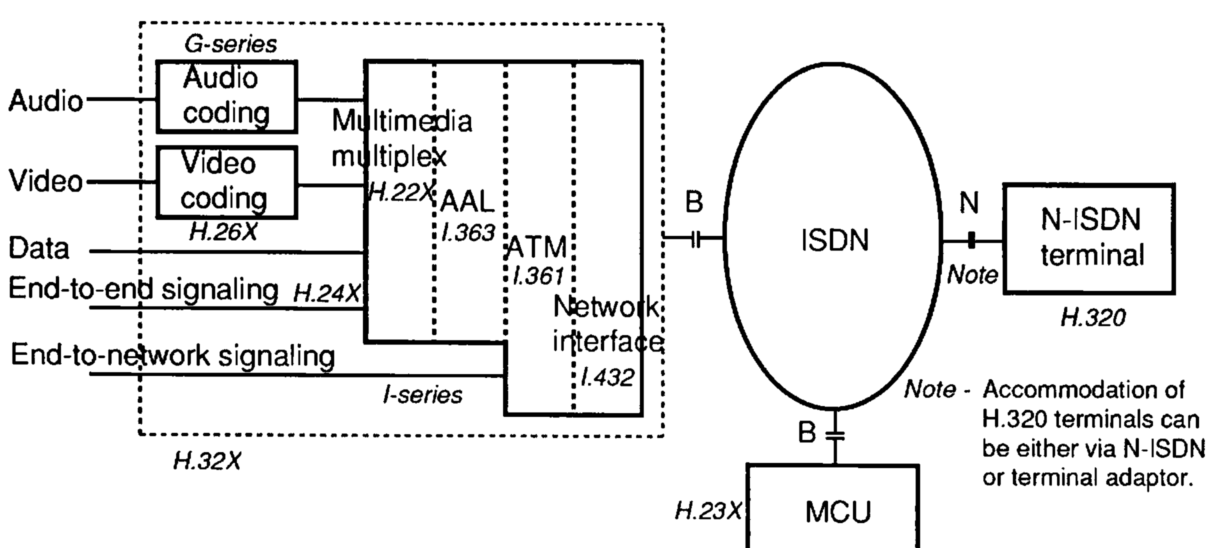


Figure 3 Audiovisual communication terminal and system configuration

4.2 Cell loss correction in AAL Type 1

SG13's liaison response on support of H.320 terminals (which are cell loss critical) in B-ISDN indicated necessity of cell loss correction at bit rates not lower than 384 kbit/s and that requirements from CMTT/3 on high quality sound transmission indicate the same solution for cell loss. Another factor is that we may need some bit error correction mechanism as well in video signal transport through ATM networks.

An AAL Type 1 solution which uses Forward Error Correction and cell interleaving has been studied both in this group and SG13. Low delay characteristics have been sought for conversational applications. As an example, use of Reed Solomon (94,88) and diagonal scanning will give processing delay of 15 ms at 384 kbit/s.

4.3 Multimedia multiplex and video support AAL

In the system model of Figure 3, we are now going to have video coding standard H.26X while ATM layer specifications have already been fixed. To connect these two layers, we need multimedia multiplex and AAL which should have the following functions;

- transfer of variable length data units. Interval between data units may be fixed or variable.
- multiplexing and synchronization of elementary streams
- capability to synchronize source and receiver clocks
- error detection and/or correction capability

In the H.32X terminal these functions are to be distributed between;

- H.22X - multimedia multiplexing and synchronization (possible common use of MPEG-2 Systems)
- I.363 - the ATM Adaptation Layer

As to the multimedia multiplex and synchronization, MPEG-2 "system coding" provides a packet and time-stamp based method. Each elementary bitstream is segmented into Packetized Elementary Stream (PES), and then respective packets are multiplexed. MPEG-2 Systems consist of the following syntax;

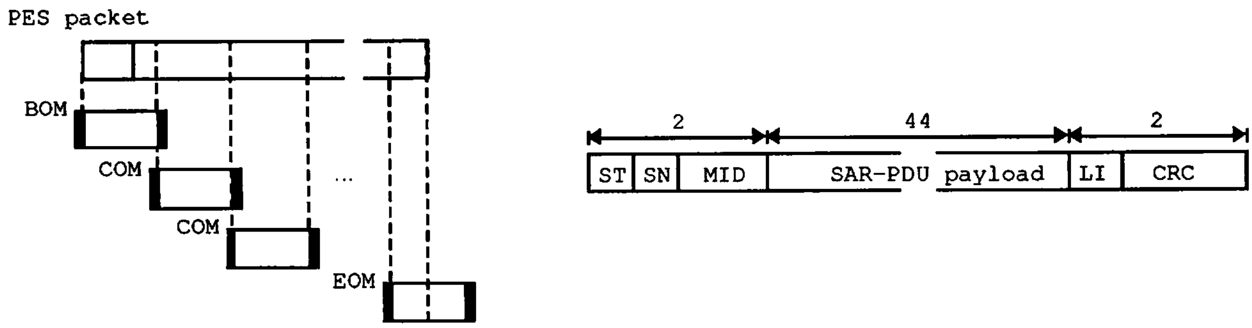
- "Program Stream" which is a multiplex of variable length PES packets and is similar to MPEG-1 Systems in functionality. This is designed for use in error free environments.

- "Transport Stream" which consists of 188 byte fixed length packets and has functionality of multiple program multiplexing (typically for broadcasting applications, each program having independent clock reference). This is designed for use in error prone environments.

In the interest of service integration on B-ISDN, it is desirable that multimedia multiplexing is also common among various applications. WP1/15 already expressed this view at its meeting in November 1992 and MPEG welcomed this policy.

It should be noted, however, that the above mentioned Transport Stream has close relevance to AAL/ATM specifications in functionality. We are now finding an appropriate combination of part of MPEG-2 Systems, H.22X, and audiovisual support AAL (AAL Type 2 in particular). The following is considered as alternatives;

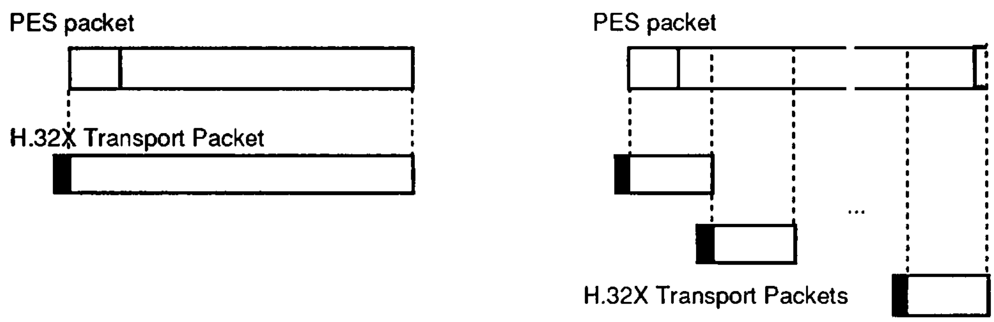
1) PES packet + modified AAL Type 4 SAR



a) PES and SAR relationship b) SAR-PDU structure

Figure 4 PES packet + modified AAL Type 4 SAR sublayer.

2) PES packet + modified Transport Stream packet (H.32X transport packet)+ AAL Type 5 CPCS/SAR



a) Error free b) Error prone

Figure 5 PES packet + modified Transport Stream packet + AAL type 5 CPCS/SAR

3) PES packet + AAL Type 5 CPCS/SAR

4) PES packet + AAL Type 1

4.4 Delivery of MPEG-2 Transport Stream packets

Given that it is required to deliver a Transport Stream across a number of systems, one part of which may be ATM/B-ISDN. An issue is what is the best way to adapt Transport Stream (TS) packets to the ATM/B-ISDN part of the link.

For the B-ISDN connection,

- some error detection/correction capability is required,
- it is believed that SAR/CS mechanisms common to other AALs should be used if possible.

AAL	ATM Adaptation Layer
BOM	Beginning of Message
COM	Continuation of Message
CPCS	Common Part of Convergence Sublayer
CRC	Cyclic Redundancy Check
EOM	End of Message
LI	Length Indicator
MID	Message Identification
PDU	Protocol Data Unit
PES	Packetized Elementary Stream
SAR	Segmentation And Reassembly sublayer
SN	Sequence Number
ST	Segment Type

Consideration is given to the use of AAL Type 1 and AAL new Type mechanisms. The latter example is illustrated in Figure 6.

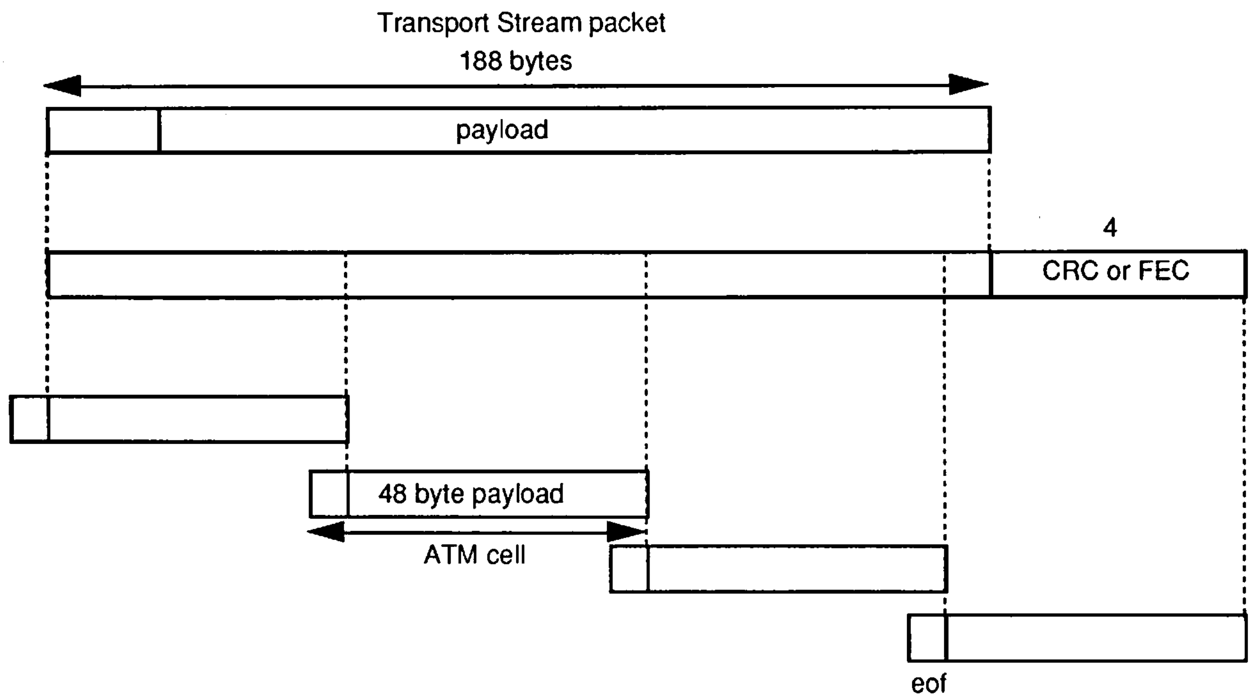


Figure 6 Transport Stream packet and new AAL.

4.5 Independent clock sources and their recovery at the decoding end

There can be several independent clock sources in the audiovisual terminal. They should be identified and the recovery method for each source should be clarified. This is yet to be worked out for the H.32X terminal.

4. 6 Joint meeting with SG13 experts on AAL

A joint meeting with SG13 experts on AAL Type 1 and Type 2 for audiovisual communications has been suggested by Mr. Yamazaki, Rapporteur for AAL Type 1/2 in SG13. We supported this idea, awaiting detailed arrangements to be settled between the two groups. The topics will be cell loss correcting AAL Type 1, video support AAL, and TS packet delivery AAL.

5. Video mixing for continuous presence multipoint systems

This discussion is in response to the request of the Tokyo Rapporteur's meeting that the Experts Group be referred to the video aspects of the video mixing bridge proposal. Our discussion results and some suggestions are summarized in a separate document TD 43 (15/1).

6. Audiovisual communication terminals as extension of H.320

In addition to audiovisual terminal Recommendation H.32X fully exploiting B-ISDN features, we have recognized the following two type of terminals which need early standardization;

- H.32Y: Adaptation of H.320 terminals to B-ISDN environments
- H.32Z: Adaptation of H.320 terminals to LAN environments

It is noted that functions of H.32Y should also be supported by H.32X, since H.32X terminals are required to interwork with H.320 terminals which are accommodated in N-ISDN. This is similar to that H.320 terminals are required to interwork with analog telephones.

Necessity of H.32Z standardization was first discussed at the Rapporteur's meeting in Tokyo last May. There are now several commercial products or announcements for desktop videoconferencing and videophone which are designed for LANs and WS/PC environments. A stack of protocols for this purpose is an area of ITU-T SG15 responsibility. SG15 should consider to provide Recommendations necessary to allow interworking between independently designed desktop videoconferencing and videophone systems over LANs.

At the subsequent meeting of the Experts Group in July 1993, there was a relevant input document discussing necessary functionalities of multimedia multiplexing, and several organizations expressed prompt interests in standardization of this area. Furthermore a UK contribution at the September meeting proposed that this should be the prime focus of SG15. We concluded that this Experts Group should draft necessary Recommendations and that time frame should be set for this work, seeking advice of WP1/15.

7. Interaction with other groups

In addition to the joint meeting sessions with MPEG in the area of common interests, the Experts Group had interaction with SG13 through correspondence.

from the Experts Group to Rapporteur for AAL Type 1/2 of SG13

July 1993 meeting sent the following;

- Cell loss correction method for AAL Type 1 to support H.320 terminals in B-ISDN

from Rapporteurs of SG13 to the Experts Group

January 1993 meeting of SGXVIII (now ITU-T SG13) sent the following;

- Update of the IVS Baseline Document
- Request of information on maximum allowable AAL delay for supporting H.320 terminals in B-ISDN by introducing a forward error correction in AAL Type 1.
- Traffic parameters for average cell rates, particularly measuring time constants of seconds or more
- Interworking issues when using CLP bit and resource allocation options

June 1993 meeting of SG13 sent the following;

- AAL Type 1 timing recovery issue
- An error correction method for delay sensitive signal transport of AAL Type 1
- IVS Baseline document
- Updated release time table for B-ISDN standardization

8. Intellectual property

A contribution from FRG reminded the participants of the ITU-T "code of practice" in TSB Circular 6, particularly the disclosure of established or pending patents related to technical proposals. The same contribution also proposed that the Experts Group should adopt the waiving option for Recommendation H.26X, but we felt that it should first be advised to and consulted with patent experts in each participating organization.

9. Work plan

We confirmed that the Tokyo Rapporteur's meeting had supported to set a target of early 1995 for "minimum system" Rec. to deter from prior divergence; also to show "committee draft" on the plan (May 1994?) to focus attention on preparation of first solid draft. Here "minimum system" is meant to be a minimum total audiovisual communication system (H.32X) consisting of video/audio coding, multimedia multiplex, communication procedures, etc.

Chairman reminded the members of hardware verification trials which should take place in the final stage of standardization work to secure our Recommendations. Since video coding/decoding specifications will be verified mostly through the bitstream exchange currently ongoing, our task is to test the above mentioned minimum system between two or more independently designed equipment. The date could be sometime during 1995. Interested organizations are requested to express willingness of participation and to contribute to materializing the plan.

10. Future activities (ITU-T sole sessions / Joint sessions with MPEG)

Meeting	Date	Sole sessions	Joint sessions with MPEG
14th	October - November 1993	October 27-29 in Daejeon (See AVC-535)	November 1-5 in Seoul
15th	March 1994	in France (?)	in France

11. Specific items requiring the consideration of Study Group 15

11.1 Intellectual property processing for H.26X

We owe to MPEG Convenor (Dr. L. Chiariglione) for his efforts to collect patent statements related to H.26X|MPEG-2 at this stage. More than 30 organizations submitted statements indicating "non-discriminatory licensing on reasonable and fair terms". These are addressed to MPEG Convenor.

The ITU-T should collect patent statements addressed to Director of ITU-TSB from those organizations by the time when SG15 initiates the approval procedure. Though Chairman of the Experts Group advised the participants of the necessity at the September joint sessions with MPEG, Study Group 15 is requested to take appropriate actions to secure this statements collection.

11.2 Alignment between ITU-T Recommendations and ISO/IEC standards

The MPEG-2 standards are published in four Parts;

Systems - specifies the system coding layer of the Specification. It defines a multiplexed structure for combining audio and video data and means of representing the timing information needed to replay synchronized sequences in real-time.

Video - specifies the coded representation of video data and the decoding process required to reconstruct pictures.

Audio - specifies the coded representation of audio data for multichannel (five plus one)

Conformance - specifies the procedures for determining the characteristics of coded bit streams and for testing compliance with the requirements stated in Systems, Video and Audio.

WP1/15 decided at its November 1992 meeting that the Video Part should be of common text with Recommendation H.26X. How to handle the Systems, Audio and Conformance Parts should be decided. MPEG is going to publish all of these in common text format even if ITU-T decides to take another approach. For the consideration of Study Group 15, Working Drafts for Systems and Audio are also input as separate documents TD 37 (15/1) and TD 38 (15/1).

It is noted that the current Systems include DSM CC (Digital Storage Media Control Commands). This is a set of commands between DSMs and the user to perform such operations as set up connection, selection of non-multiplexed bitstreams, playback, storage, edit, information request. In our terms, DSM CC is a set of C&Is.

11.3 Time schedule

It is quite obvious that the video coding cannot stand alone; we need audio coding (G.XXX), multimedia multiplexing (H.22X), communication procedures (H.24X), call control (AV42X), whole terminal (H.32X) etc. as illustrated in Figure 1 to make interoperable communication systems. WP1/15 made a small step at the November 1992 meeting by declaring that they will be "1994+" Recommendations in the "IVS Baseline Document" edited by SG13.

One serious difficulty is that B-ISDN is not so visible as N-ISDN, but the other way to look at the situation is that we are in a good position to do the work before incompatible equipment appears in the market and to reflect audiovisual requirements to the network standardization.

As the SG15 meeting schedule is given for 1993-1995, we may be able to produce a WP work plan more precisely; what Recommendations we intend to make and at what time. We need good target dates for Recommendations including H.32Y and H.32Z described in §6 above.

11.4 Terms of reference for the Experts Group

Currently the Experts Group is charged with mostly pure video coding in its terms of reference. Considering

- that the generic coding Recommendation H.26X is going to become complete on schedule,
- that the application Recommendation should specify how to use the generic standard,
- that interworking is not possible without the total system and terminal specifications,

the role of the Experts Group should be reviewed in the light of the WP work plan for this study period and onward.

END