CCITT SGXV
Working Party XV/1
Experts Group for ATM Video Coding

Document AVC-317R
July 10, 1992

SOURCE : CHAIRMAN OF THE EXPERTS GROUP FOR ATM VIDEO CODING
TITLE : REPORT OF THE SEVENTH MEETING IN NEW JERSEY AND RIO DE JANEIRO (July 1-10, 1992) - PART I
Purpose: Report

Part I (Sole Sessions)

Contents

1. General

The seventh meeting of the Experts Group consisted of two parts; CCITT sole sessions in New Jersey and joint sessions with ISO/IEC JTC1/SC29/WG11 (MPEG) in Rio de Janeiro. The list of participants appears at the end of this report.

The first part was held during 1-3 July 1992 at Sheraton Hotel in Eatontown, New Jersey, USA, at the kind invitation of AT&T Bell Laboratories and Bellcore. At the opening session, Dr. Jules Bellisio, Bellcore, made a welcoming address on behalf of the hosting organizations.

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

The Experts Group also had a short closing session on 10 July in Rio de Janeiro.

The following change of the membership of the group was announced;

- Participation of Korea had been approved at the SGXV meeting in May with its Coordinating Member Mr. J-Y. Nam.
- Mr. H. Carbiere had resigned Liaison Representative to CMTT/2.

The Experts Group regretfully accepted Mr. Carbiere's resignation and solicited his successor.

2. Documentation (TD-2)

For this meeting, 61 AVC-numbered documents and 8 temporary documents were available as listed in Annex 1.

3. Tape demonstration (TD-3)

Several video tape demonstrations were given with D1 or U-matic as detailed in Annex 2 to present experimental results. There were also the following hardware demonstrations;

- AT&T videophone,
- GEC-Marconi videophone.

4. Review of the previous meetings relevant to the Experts Group

4.1 WPXV/1 in May (AVC-261,262,263,264,265)

Chairman presented the outcome of the WPXV/1 meeting held in May, focusing on the following topics;

- H-series new and revised Recommendations
- joint work with MPEG
- PSTN/mobile video telephone

4.2 CCIR, CMTT in May

Mr. Zedler reported that the May CCIR recognized the work of MPEG and decided to prepare requirements from the broadcasting point of view, and that an Special Rapporteur (Mr. D. Nishizawa, NHK - Japan) was appointed with two assisting members for this purpose.

4.3 SGXVIII in June (AVC-311,312)

Chairman introduced briefly the liaisons statements received from Mr. K. Yamazaki, Special Rapporteur SWP XVIII/8-3.

5. Picture format (AVC-274,295)

AVC-274 presented simulation results on 50Hz/60Hz mutual conversion, and AVC-295 reinforced the flexible spatial resolution approach. An answered question is what picture format become a major on B-ISDN. The first document is assuming two versions of CCIR-601, and the second document is assuming window-based multimedia workstations and PCs. There was a comment that US sentiment for HDTV is somewhat in line with the second document. We need application scenarios as recognized in Stockholm.

Since H.26X should be generic and be able to cope with a wide range of formats, the current format issue is rather of H.32X terminal. During the discussion, it was stressed that the problem is how to achieve interworking between various types of terminals with different source formats which are connected through B-ISDN. One way is to define a default format every terminal should have, the other way is to define that every decoder of the same class be able to decode and display any format smaller than a defined maximum.

The meeting confirmed the previous decision in Stockholm to endeavor to reach a common understanding in Ipswich next October.

6. Source coding

6.1 Coding structure

6.1.1 TM1 (AVC-260)

TM1 was defined at the Haifa meeting. The agreed definition of coding schemes and experiments have been edited as AVC-260 by the effort of Editing Committee.

6.1.2 Prediction optimization (AVC-280,281,282,285,301)

It has been recognized that comparison between various predictions (field/frame adaptive, field, pure field, FAMC) need data on the following;

- picture quality,
- inherent delay,
- hardware complexity.

Documents AVC-280,281,282 provided such information, particularly on FAMC. The Experts Group considered this topic from from the low delay mode (see Section 6.2). These and other comparisons will be the major topic in Rio de Janeiro. During the discussion, the following comments were noted;

- The number of memorys is one of the most important factors for implementation study.
- FAMC requires a larger number of memory accesses than other predictions; this will affect hardware complexity.
- Videoconferencing applications will be benefited by a particular prediction.
- Since multimedia applications use film and computer graphics sources, frame based prediction is preferable.
- Decoder complexity does not differ so much even if we allow multiple predictors.

Document AVC-301 proposed an generalization of prediction which consists of a generic prediction vector, up to four pixels in the reference picture which are address by the generic prediction vector, and their weighting values to obtain a prediction value (see Annex 3). The reference pixels and their weighting may be down loaded by the video signal itself or pre-defined according to the application. The current proposal covers all the predictors contained in TM1.

The meeting recognized that this would be a solution for the prediction issue, if we really need a different predictor for each sequence because it is significantly advantageous compared to other predictors. Generally, the following predictions are suited;

- frame coding for still or slowly moving pictures,
- FAMC for moderately moving pictures,
- field coding for rapidly moving pictures.

It was pointed out this solution might ease the standardization process and that it reduces number of macroblock types. However, the following items require further consideration;

- Can one predictor not be practically sufficient?
- Impact of picture distance (value of M) and use of picture dropping
- How B pictures be processed?
- How to cope with adaptive prediction on a macroblock basis, if it is necessary?
- Loading interval for the tables depends on the application.

6.1.3 Adaptation to "tune-in" (AVC-310)

Document AVC-310 proposed inclusion of quantizer matrices in the picture header and use of leaky prediction for channel hopping in broadcasting applications. The meeting was aware of that this tune-in feature also facilitates switching type multipoint communications.

The meeting supported the first proposal. As to the leaky prediction, the meeting decided to consider it in the context of cell loss resilience. During the discussion, it was clarified that the leaky prediction provides graceful build-up compared with the delayed start due to the use of periodic intra pictures, and that the loss of coding efficiency is compensated by non-use of periodic intra pictures.

6.2 Low delay mode (AVC-260)

6.2.1 Performance objectives (AVC-266)

Summary of SGXII study results on the delay in audiovisual services was provided by a liaison statement to SGXV. Recommendation G.114 stipulates that one way delay of 150 ms is "acceptable for most user applications" if echo is adequately controlled. Information on longer delay considerations are also given in this liaison statement.

6.2.2 TM1 experimental results and their implications (AVC-275,276,284,291,302)

Several topics regarding the low delay mode were addressed by respective contributions as follows;

1) coding structure; values of (N,M), use of forced intra slices

AVC-275,276,284,291,302

It was pointed out that pure field coding with M=3 may also provide less than 150 ms coding/decoding delay in addition to the TM1 low delay mode.

2) optimum prediction; field/frame, field, pure field, dual field, FAMC

AVC-275,276,302

Document 302 describes an FAMC like field prediction, which indicates that the idea of FAMC is also applicable to the field coding.

3) first picture in TM experiments

AVC-275,291

There were found two different ways to process the first picture in the TM1 experiment; coding it as a intra picture (field or frame) and then starts forced intra slice from the second picture, and coding the first one/two slices only at the first picture, then the first two/four slices at the second picture and so on.

4) scene change handling

AVC-275

Tape demonstration accompanying Document AVC-275 showed two possible solutions for scene change handling in the low delay mode; picture dropping and wiping. The meeting found the former solution less annoying, though it was noted that the repeating the frame before the scene change may produce motion jittering. It was also commented that if picture dropping takes place in the MPEG1 system multiplex, time stamp for synchronizing audio and video may not work.

There was a question whether the picture dropping be used for VBR, but this is open at the moment.

5) rate control method

AVC-284

Visual moving of forced intra slices were shown in the demonstration accompanying Document AVC-284, but they were not recognized in the demonstration accompanying Documents AVC-302,310. We need further clarification on this aspect.

6.2.3 Core experiments for the low delay mode

After discussion, the meeting asked Mr. Bjoentegaard to coordinate a small group for drafting further core experiments to clarify the issues identified above in Section 6.2.2. The outcome is contained in Annex 4.

The meeting reviewed this outcome and agreed to forward it to the joint sessions in Rio de Janeiro.

6.3 Compatibility and scalability

6.3.1 Experimental results and their implications (AVC-277,286,293,313)

These four experiments provide information on the comparison of embedded vs simulcast, indicating the embedded coding gives better SNR than the simulcast. AVC-277 also provides experimental results on comparison of prediction from the base layer vs prediction of the prediction error, favoring the first scheme in SNR performance.

There was some discussion on necessary hardware for the embedded decoding related to the required number of decoder loops for the full resolution only decoder. If we use high speed processor for two or more loops in a time shared manner, the number of loops does not count. This structure does not increase coding and decoding delay due to the use of low data rate base layer for prediction of the enhanced layer.

It was clarified for the prediction of the prediction error scheme that the enhanced layer decoder receives the base layer motion vector but does not use it for reconstructing the full resolution picture.

6.3.2 H.261 compatibility (AVC-278)

AVC-278 clarified H.261 compatibility requirements for H.26X coder and decoder by analyzing several communication patterns. The meeting noted that all of these situations should be taken into account when we discuss the compatibility issue.

If embedded scheme is adopted for H.26X, every coder and decoder should have the embedded structure. In a point to point communication between two H.26X terminals, embedded coding may degrade performance, thus a mechanism to switch off the base layer as necessary should be considered.

6.3.3 Scalability (AVC-292,299,306)

Three different scalable structures were presented in these three documents. The meeting simply took note of the information at this moment because the compatibility, H.261 compatibility in particular, is a more urgent issue from the CCITT view point.

6.3.4 Core experiments for H.261 compatibility

After the above discussion, the meeting asked Mr. Parke to coordinate a small group to plan H.261 compatibility experiments. The outcome is contained in Annex 5. The meeting reviewed this report and decided to put forward the experiment proposal to the joint sessions.

6.4 Cell loss resilience

6.4.1 Experimental results and their implications (AVC-269,279,287,298,308; AVC-310)

The following techniques were experimented as cell loss resilience measures;

1) Structured packing: AVC-269,279

Inclusion of video date address information in a cell and localization of damaged area were found very effective for cell loss resilience. There was a question whether this technique is also applicable to multimedia multiplexed audiovisual signals such as H.221 structured signals, and a comment that the number of bits for the overhead indicating address etc. should more conveniently a multiple of 8.

2) Leaky prediction: AVC-279,310

The meeting desired that the contribution of leaky prediction be separated from that of error concealment in AVC-279. Mr. Sakai and Ms. Reibman clarified that the value of leaky factor sensitively affects efficiency and recovery time and that the optimum value depends on the input picture.

3) Concealment: AVC-279,308

Both experiments showed that concealments supported by suitable AAL (including indication of errored cells and slice/MB level re-entry) provide an effective means to make either one- or two-layered MPEG coding cell loss resilient. Dr. Raychaudhuri summarized that a specific AAL/error concealment for MPEG-1 provides reasonable pictures (corresponding approximately to "entertainment quality" in the US) at the following cell loss ratios;

One-layer: CLR ~ 10E-4 to 10E-6 range (depending on decoder concealment capability)

Two-layer: low priority CLR ~ 10E-2, high priority CLR ~ 10E-5

He also informed the meeting that his HDTV experiments (with 120 byte cell payload) indicate that cell size has a relatively small impact on the resilience.

4) Layered coding: AVC-287,298

In response to a question, it was clarified that if CLP is used on a cell by cell basis in a Virtual Channel, the network guarantees cell sequence integrity, thus there is no need to synchronize the two priority bitstreams. The layered coding in AVC-287 uses only the upsampled base layer prediction error with a upscaled motion vector if a macroblock in the enhancement layer is hit by the cell loss, while the one layered coding uses only resynchronization of the next slice.

6.4.2 Error resilience improvement for TM1 syntax (AVC-288)

As to the proposed addition of the two bit after the extension code, the meeting did not understand its effectiveness, thus decided not to support this proposal until further clarification is obtained.

There was a comment on the picture slicing that the real time coding may utilize this technique depending on the transmission media, but the pre-recorded signal can not.

The meeting confirmed that H.261 IDCT specifications stand also at higher bit rates envisaged for H.26X, because refresh frequency is defined in terms of "once every 132 times coded".

6.4.3 Core experiments (AVC-260,300)

After the above discussion, the meeting asked Mr. Biggar to coordinate a small group to design core experiments addressing the following items;

- clarification of the effectiveness of those techniques listed in Section 6.4.1 above.
- Techniques applicable to single layered coding
- Techniques applicable to multiple layered coding

The outcome is contained in Annex 6. After reviewing this report, the meeting decided to put forward this proposal to the joint sessions.

6.5 VBR

6.5.1 Statistics (AVC-267,307,314)

These three documents provided various statistics of the VBR video, which were collected for open loop operation of three different coders (BBE hybrid DCT, MPEG-1, H.261). The meeting appreciated this information for studying multiplex gain of the VBR traffic. There was a question whether these statistics may change if the average rate is regulated according to the preventive policing of the coder.

6.5.2 Impacts of UPC (AVC-271,272,312)

AVC-272 provided a delay analysis for CBR and VBR and indicated a particular case where the sliding window regulation of the average gives benefit to the VBR operation; MPEG-1 like periodical variation of the information generation. AVC-271 reported supporting experimental data.

The liaison statement from SGXVIII indicated that the traffic parameter for the average is likely to be defined in terms of a reference algorithm as currently done for the peak rate in I.371. The Experts Group is greatly concerned with the time constant of the reference algorithm. If it is short in the order of e.g. 1 ms, the coder should regulate the coding rate much more tightly than the current CBR system where a transmission buffer of several tens of ms is used. This should be communicated with SGXVIII.

6.5.3 Cell loss for multiplexed VBR sources (AVC-296)

Long term characteristics of cell loss were presented for multiplexed VBR video sources. An implication is that once congestion takes place, very bad condition with CLR=0.01 may continue for tens of frame time due to autocorrelation of the video source signal. Study of visual effects of CLR=0.01 was proposed. There was a general comment whether the result would change if the source is regulated in average by a leaky bucket mechanism etc.

If the source coding should cope with this high CLR by using e.g. layered coding, redundancy should be included in the coded video, thus the number of channels to be accommodated decreases. Another way to cope with this situation may be to accommodate a slightly less number of channels containing less redundant coded video.

This periodic bunching may happen for non video traffic. The meeting recognized necessity to communicate with SGXVIII on this matter. The coming IVS Technical Session will be a good opportunity to exchange views on this and other network related topics.

Possible feedback of this study result on the current network model was raised, and it was clarified that this result will supplement the network model, but not change. The exact way of reflection is for further study.

7. AAL and other network aspects

7.1 Clock recovery (AVC-273,315)

AVC-273 addressed the following items regarding the video clock recovery;

- supported by AAL or video codec?
- what method be used? what field be prepared for this purpose?
- what method if no common clock is available?

Though we could not reach a firm conclusion, there was general support for the video codec to have this functionality. Study for exact way and other aspects are to be continued.

AVC-315 provided a solution for timing recovery of audiovisual signals transmitted through a variable rate channel where a common clock is not available between the encoder and decoder. This work is based on MPEG system study; members of the Experts Group are encouraged to learn its wisdom. It was also pointed out that this solution is useful for the database access where pre-recorded video is involved and a common clock does not solve the problem even if it is available at both ends.

7.2 AAL for video signal transport (AVC-294,297,312; AVC-298)

Both AVC-294 and AVC-297 provided example AALs for video signal support to stimulate the group, while AVC-312, liaison statement from SGXVIII, listed required functionalities of AAL for video signal support. The key issue is how generic AAL should be, whether full functionality be in AAL or the user data (see also TD-6). One possibility would be to define a null AAL, and contain all necessary fields in the video data.

We should take the following steps;

- To identify required functionalities for audiovisual system, particularly for video,
- To identify where each of those functionality be supported; in AAL or in the user data. The guideline should be that the network related functionalities should be supported by AAL, but video related ones by the user data.

There was expressed some concern about the interface between the AAL and the (video) user. If it is clearly defined, the ATM video codec can also be applied to other networks than B-ISDN.

7.3 Support of H.320 terminals on B-ISDN (AVC-312)

SGXIII sent us a question on the subject matter in the liaison statement. The meeting felt that the content is almost reiteration of our previous question to SGXVIII. Since there is some time to make a reply (SGXV will meet next in January 1993), Mr. Morrison undertook to draft a reply and circulate before the next meeting of the Experts Group.

7.4 LAN (AVC-270)

AVC-270 provided some information on the difference between B-ISDN and LANs. Differing views were expressed whether video signal transport is feasible on existing LANs. Study should be continued on the ground that H.26X codecs may be used either on ATM based LANs or on LANs connected to B-ISDN as NT2.

8. Multimedia multiplex (AVC-268; AVC-297)

AVC-268 provided discussion on the possible interface between the media control layer and the video codec and raised questions on negotiation methods.

The meeting considered how we can make progress on the multimedia multiplex study. Mr. Tanaka commented that we should study several alternatives in parallel until network conditions (cell loss, cost of separate VCs, differential delay between separate VCs, B-N interworking) become clearer. It is a common understanding of the meeting that a particular solution is not wise at this stage.

Chairman raised a study item on the multimedia multiplex; commonality with MPEG-2 system which is envisaged as extension of MPEG-1 system. Analysis of the problem is awaited.

9. Hardware verification of H.26X (AVC-283)

AVC-283 was input as a stimulus to plan hardware trials which is required at the later stage of the Experts Group activities. First of all, at least two volunteers are solicited who independently develop hardware for the testing purpose. The following comments were obtained;

- Definition of the digital interface between Terminal Adaptor and the codec part enables testing in various environments.
- The work plan to test at the end of 1993 is not practical. The hardware may be much more difficult to develop compared to the H.261 Flexible Hardware due to the higher speed and complexity. Developing time of six months is not realistic.
- To transport a simulated coded bitstream through the network and to decode it at the remote end may be a method of verification.
- VBR can be tested at the same time, or CBR is tested first and VBR at a later date?
- Coordination with MPEG on the hardware verification is suggested.

Members are requested to study various aspects of the testing process to materialize the action plan.

Related to this discussion, a view was also expressed that the current plan of "freezing specifications" in March 1993 is difficult to achieve if the amount of work to do is considered. Reactions to this view are awaited. At this moment, we should take our efforts to make the "frozen specifications" appropriate for our communication systems.

10. Very low bit rate video coding

10.1 Background (AVC-263)

At its May 1993 meeting, Working Party XV/1 asked Chairman of the Experts Group to make an initial consideration in his group on the very low bit rate coding for PSTN and mobile/cordless networks. The meeting had one hour free discussion after having reviewed relevant documents and tape demonstrations.

10.2 Experimental results (AVC-289,303,305,309)

Several simulation and experimental results were presented on video coding around 8-16 kbit/s. The following comments were noted in particular;

- Modem initialization took very long time (e.g. 20 second) at the call set up (AVC-289), though it might be a problem of the particular equipment.
- Significant improvements have been found by using half-pel accuracy motion compensation instead of integer motion compensation plus loop filter employed in H.261 (AVC-303).
- Improvements by modifying VLCs and video multiplex are rather minor (5%), the used picture format is 1/9 CIF, the monitor size must be less

than 3-4 inches. Susie is a stress material in these low bit rates (AVC-305).

10.3 Policy (AVC-290,304)

The following is a summary of views expressed during the free discussion. Chairman of WPXV/1 will be consulted with for suggested further actions of the Experts Group to prepare for a reply to the November meeting of WPXV/1.

1) Video coding possibility

Algorithms of presented simulation or hardware results were in the category of H.261 +/- delta, providing more or less similar performance. There was no indication for possibility of quite new algorithms at this meeting.

2) System and terminal

PSTN videophone standardization should cover all of the following aspects for the global terminal;

- video source coding
- audio source coding
- multimedia multiplex
- channel coding
- modem, including fall back capability
- call set up, delay

PSTN videophone system may be differentiated with cordless system which is used as a leg to the ISDN system.

It was noted that V.FAST modem (28.8 kbit/s) is being studied and requirements from conversational services should be reflected.

3) Interworking between PSTN and ISDN videophones

When a PSTN videophone interworks with an ISDN videophone, the ISDN side should prepare for interworking capability. Since existing ISDN videophones do not have this capability, we should be careful of making retrospectively mandatory specifications. Existing ISDN videophones are required to be able to interwork PSTN telephones, thus one way to cope with the interworking requirement may be to deal with it in the context of ISDN videophone and PSTN telephone interworking.

This consideration should be a part of the general PSTN/N-ISDN/B-ISDN interworking issue.

4) Standardization program

We discussed necessary standardization in terms of "short term" and "long term". The former means a very quick work and the latter may require 5 or more years time. There was no claim for "middle term" standardization work which lasts 2 or 3 years. Reasoning is;

- N-ISDN will prevail in this time frame.
- Current proprietary equipment may prevail.

It was a common understanding that longer term standardization of low bit rate video coding is required. Very low bit rate coding significantly better than H.261, which may or may not be model based coding, is applicable to mobile applications and it will also provide very good pictures for the ISDN videophones.

There were expressed opposing views for the short term standardization. Supporting views are;

- De facto situation should be avoided.
- Technical solutions are feasible; video coding close to H.261 but optimized at low bit rates which can provide reasonably good pictures with reasonable complexity, 5" LCD without extra cost, around 6 kbit/s audio coding, V.FAST modem, etc.

Opposing or reluctant views are;

- The essential factor in PSTN videophone is rather pricing than technical specifications, price related technology is not appropriate for standardization. If short term standardization is necessary, CCITT should adopt one of the existing ones, possibly through licensing (?). Defining the third one does not help the industry.
- It is questionable that the current performance be accepted by customers. Answers may be available in six months or one year.

5) Harmonization with other standardization bodies

It is a unanimous view that CCITT should carry out the global system standardization as mentioned 2) above.

11. Preparation for the joint sessions

11.1 Documents

TD-5,6,7 are are forwarded as proposal from the Experts Group. The following documents are forwarded for information or discussion with each original source; AVC-275,276,278,279,296,297,298,300,307,308,315.

11.2 Representatives

The following is appointed as representative to the joint session;

Requirements/Test	S. Okubo
Video	G. Bjoentegaard
System	B.G. Haskell
Implementation	D.G. Morrison

12. Others

12.1 IVS Technical Session (AVC-311)

1) Speakers

Messrs S. Okubo, D.G. Morrison and A. Tabatabai will speak on behalf of the SGXV Experts Group.

2) Participants

Those who intend to attend are requested to advise Chairman by August 10.

12.2 Update of the "status report" AVC-109

1) Time schedule

The meeting agreed to complete revision by the end of September so that we can distribute it at the IVS Technical Session. Cooperation of the editors are requested.

2) Editors

As listed in AVC-109.

12.3 Future meeting plan

- 8th meeting: September 28 - October 1, 1992 in Tarrytown, USA.

  Sole sessions: September 28 (half day)
  Joint sessions: September 29 - October 1

- 9th meeting: October 28 - 30, 1992 in UK

  Sole sessions: October 28 - 30 in Ipswich
  Joint sessions: November 2 - 6 in London

- 10th meeting: January 1993?

- 11th meeting: March ?? - April ?, 1993 in Australia

  Sole sessions: after the joint sessions?
  Joint sessions: March 29 - April 2

Note: IVS Technical Session is planned on October 26-27, 1992 in Ipswich.

END

* * *

Annexes

Annex 1 Documentation
Annex 2 List of tape demonstration
Annex 3 Generic predictor in AVC-301
Annex 4 Report form the small group on low delay mode
Annex 5 Report of the small group on H.261 compatibility
Annex 6 Cell loss resilience: Issues and core experiments

Participants of the seventh meeting of
Experts Group for ATM Video Coding
(1-10 July 1992, New Jersey and Rio de Janeiro)

			N	R	
FRG	Mr. M. Kuehn	DBP Telekom		X	
	Mr. F. May	Daimler-Benz	X	X	CM
	Mr. G. Zedler	DBP Telekom	X	X	CM
Australia	Mr. M. Biggar	Telecom Australia	X	X	CM
Belgium	Mr. O. Poncin	RTT Belgacom	X	X	CM
	Mr. B. Voeten	Bell Telephone	X		
Korea	Ms. Sang-Mi Lee	ETRI	X	X	(CM)
USA	Mr. B.G. Haskell	AT&T Bell Labs	X	X	
	Mr. S. Kumar	Wiltel	X		
	Mr. A. Luthra	Tektronix	X	X	
	Mr. D. Klenke	CLI	X		
	Mr. D. Hein	VideoTelecom	X		
	Mr. N. Randall	DIS	X		(CM)
	Mr. D. Raychaudhuri	David Sarnoff	X		
	Ms. A. Reibman	AT&T Bell Labs	X		
	Mr. R. Schaphorst	DIS		X	CM
	Mr. A. Tabatabai	Bellcore	X		CM
	Mr. F. Tobagi	Starlight Networks	X		
	Mr. J. Zdepski	David Sarnoff		X	
	Mr. Y. Xiancheng	PictureTel	X		
France	Mr. G. Eude	CNET	X	X	
	Mr. J. Guichard	CNET	X		CM
Italy	Ms. L. Conte	CSELT	X	X	CM
Japan	Mr. S. Okubo	NTT	X	X	Chairman
	Mr. T. Murakami	Mitsubishi	X	X	
	Mr. K. Sakai	Fujitsu	X	X	
	Mr. Y. Takishima	KDD	X		(CM)
	Mr. T. Tanaka	NTT	X		CM
	Mr. H. Ueno	Toshiba		X	
	Mr. T Yukitake	Matsushita Communication	X	X	
Norway	Mr. G. Bjoentegaard	NTA	X	X	(CM)
Netherlands	Mr. D.A. Schinkel	PTT Research	X		CM
	Mr. A. Koster	PTT Research		X	
UK	Mr. I. Parke	BT	X	X	
	Mr. D.G. Morrison	BT	X	X	CM
Sweden	Ms. C. Verreth	Telia Research	X	X	

CM: Coordinating Member
(CM): Substitute for CM

Documents for the seventh meeting of the Experts Group
1-10 July 1992, New Jersey and Rio de Janeiro

Normal Documents

Note: Contributions with "*" have also been sent to MPEG for consideration at the joint sessions. Some documents are also registered directly through the MPEG channel, which are not marked with * but have been considered at the joint sessions. See AVC-318R.

AVC-256R REPORT OF THE SIXTH MEETING OF THE EXPERTS GROUP FOR ATM VIDEO CODING IN STOCKHOLM/HAIFA - PART I (CHAIRMAN)
AVC-257R REPORT OF THE SIXTH MEETING OF THE EXPERTS GROUP FOR ATM VIDEO CODING IN STOCKHOLM/HAIFA - PART II (CHAIRMAN)

Achievements and action points obtained at the sole sessions in Stockholm and the joint sessions in Haifa are recorded.

AVC-258 LIST OF REQUIREMENTS FOR MPEG-2 VIDEO (REQUIREMENTS GROUP)

This document lists all the identified requirements for video part of the audiovisual coding, providing definition, application examples and comments, which were updated in Haifa.

AVC-259 GUIDE FOR THE VIDEO WORK (REQUIREMENTS GROUP)

This document extracts from the general requirements listing only those requirements that directly impact the development of video coding standard and the testing thereof, which were updated in Haifa.

AVC-260 TEST MODEL 1 (TEST MODEL EDITING COMMITTEE)

A comprehensive description of Test Model 1 (TM1) is given, incorporating coding schemes and various experiments. This model is used in the course of research for comparison purposes.

AVC-261 THIRD PROGRESS REPORT (CHAIRMAN)

This is a progress report to Working Party XV/1, covering major achievements obtained in the three meetings (Yokosuka, Singapore, Stockholm/Haifa) toward defining Recommendation H.26X "video coding in the ATM environments", and particular items for consideration of WPXV/1.

AVC-262 LIAISON STATEMENTS SUBMITTED BY THE EXPERTS GROUP (CHAIRMAN)

This is a collection of three liaison statements that the Experts Group sent; two to SGXVIII and one to CCIR Ad-Hoc Group on Digital Coding.

AVC-263 MEETING REPORT (WPXV/1)

This is a report of the Working Party XV/1 meeting held in Geneva during 7-13 May 1992. The following items particularly concerns the Experts Group;

- Very low bit rate video coding for PSTN and mobile telecommunications,
- SGXV endorsement of collaborative work between CCITT EG and MPEG,
- Request for inputs on likely applications of B-ISDN,
- Handling of MPEG "system" in the framework of AV Recommendations (H.200).

AVC-264 RULES FOR PRESENTATION OF CCITT:ISO/IEC COMMON TEXT (CCITT)

AVC-265 SOME HELPFUL INFORMATION TO EDITORS OF CCITT/ISO COMMON TEXT (CCITT SECRETARIAT)

These two documents provide "presentation rules" for documents which are intended to be both CCITT Recommendations and ISO or ISO/IEC International Standards. These rules are called fictitiously Recommendation A.1000 and ISO/IEC 0001.

AVC-266 NETWORK PLANNING FOR VIDEOTELEPHONY & VIDEOCONFERENCE SERVICES (CCITT SGXII)

Current status of study is given for the effect of end-to-end delay (transmission and processing delay) on the communication quality. Revised G.114 recommends 400ms as the upper limit of acceptable one way end-to-end delay, recognizing that this limit will be exceeded in some exceptional cases (including videotelephony over satellite circuits). A summary of various evaluation results is also given which addresses the subject in audio only as well as audiovisual communications, indicating that highly interactive tasks require less than 400 ms pure delay (without echo) but that the effects of pure delay on speech quality appears to modestly increase as the delay is increased, hence that further study is necessary.

AVC-267 BIT RATE STATISTICS OF A TV DISTRIBUTION CODEC (RTT BELGACOM)

Using a hardware hybrid coder (without motion compensation), several VBR statistics are collected for TV programmes in ATM environments; cell inter-arrival times, cell rate for unbuffered packetization, cell rate for slice buffered packetization, and bit rate for image buffered packetization. It is suggested that data may be used for estimating statistical multiplex gain and determining the policing function for TV codecs.

AVC-268 MULTIMEDIA MULTIPLEX AND NEGOTIATION METHODS (JAPAN)

Three design examples of the interface between the video codec and the media control layer are discussed considering cell loss resilience and H.221 multiplex requirements. Negotiations for the communication mode in B-ISDN, including cases of uni-directional and non-symmetrical transmission, are also discussed from the view point of possible negotiation channels; outband, dedicated channel, inband.

AVC-269 SEGMENTATION TRICK TO IMPROVE STRUCTURED PACKING EFFICIENCY OF CODED SIGNALS (JAPAN)

An improvement for structured packing method is presented with experimental results where truncation of transmission unit takes place if more than a certain number of cells are already filled and if the remaining capacity of the current cell is less than a certain value. This method provides a better trade-off between transmission efficiency and the size of area damaged by a cell loss. It is noted that the cell packing structure requiring cell boundary information must be standardized, while the segmentation algorithm is not a matter of standardization.

AVC-270 CONSIDERATION OF LAN (JAPAN)

Difference between B-ISDN and LAN are reviewed with respect to network performance (delay, transmission error, clock). Impact of LAN on H.32X terminals is also discussed.

AVC-271 MERIT OF VBR TRANSMISSION (JAPAN)

Simulation results are provided which indicate that VBR with sliding window policing of average rate can benefit a coding structure with periodical variation of the coded bit generation like MPEG1. It is shown that the best benefit is obtained if the window size is exactly an integer multiple of coding structure parameter N.

AVC-272 IMPROVEMENT OF DELAY UNDER AVERAGE BITRATE CONSTRAINT (JAPAN)

Transmission delay is analyzed for CBR and VBR using an information generation model. Sliding window and leaky bucket are compared as average rate UPC methods for VBR assuming the same mechanism is applied to the coder for preventive policing. It is concluded that leaky bucket VBR can provide shorter delay than CBR, but sliding window VBR can not except for a particular case of periodical variation of information generation.

AVC-273 CLOCK RECOVERY FOR VIDEO (JAPAN)

It is first discussed whether video clock recovery be carried out in AAL or video codec, suggesting that the latter method can be more flexible for multimedia multiplexing. Then, three alternative methods of clock recovery in the video codec are compared; buffer control, stuffing and frequency counting methods. The frequency counting method is suggested from adaptability to VBR services. It is also pointed out that the clock recovery should be studied for the case where a common network clock is not available at both the encoder and the decoder.

AVC-274 PICTURE QUALITY COMPARISON OF FORMAT CONVERSION THROUGH / NOT THROUGH SCIF (JAPAN)

Conversion between CCIR 625/50 and 525/60 signals is experimented for three different schemes; two schemes use SCIF in the intermediate of conversion and one scheme uses direct conversion. It is concluded that conversions with small delay give similar picture quality regardless of being through or not through SCIF.

AVC-275* SCENE CHANGE HANDLING IN LOW DELAY MODE (JAPAN)

Scene change which requires a large number of bits per picture and consequently additional buffering delay is discussed with TM experimental results, comparing a strategy which assigns almost constant number of bits per picture and a strategy which allows occasional picture dropping. The second strategy needs definition of skipped pictures in the coder and the decoder.

AVC-276* COMPARISON OF PREDICTION METHODS OF THE LOW DELAY MODE OF TM1 (JAPAN)

Three experimental results are reported for the TM1 low delay mode; the first experiment compares four frame-based prediction methods indicating FAMC gives the best SNR for M=1, the second experiment compares pure field predictions with frame-based ones indicating that "M=3 pure field" gives better SNR than "M=1 frame/field adaptive", the third experiment shows that half-pel MC referring to the reconstructed pictures gives slightly better SNR than the one referring to original pictures.

AVC-277 SIMULATION RESULT ON COMPATIBILITY (JAPAN)

Simulation results for the two compatibility experiments are given; "prediction from the base layer" and "prediction for the prediction error". It is concluded that compatible type switching defined in the TM1 syntax is effective for the fist scheme and that the prediction from the base layer provides better performance than the prediction for the prediction error even after having been improved by remaking the prediction from the lower layer depending on the upper layer prediction mode.

AVC-278* H.261 COMPATIBILITY REQUIREMENT (JAPAN)

After analyzing several H.26X/H.261 interworking scenarios (point-to-point, multipoint, multicast and database access), requirements for the H.26X decoder and coder are concluded as follows;

- The decoder in B-ISDN should be able to receive both of H.26X bitstream and H.261 stream, but one at a time, in all the situations including point-to-point, multipoint, multicast and database access.
- H.26X encoder should be able to transmit either of H.26X bitstream or H.261 stream at a time for point-to-point communications or multipoint communications where a common operating mode is used or where MCU carries out transcoding. It should also be able to transmit two bit streams simultaneously for multipoint communications where MCU does not provide transcoding or where a mesh connection is employed, multicast communications and in the database.

AVC-279* CELL-LOSS COMPENSATION SCHEME (JAPAN)

Structured packing (use of macroblock start pointer and absolute block address in the cell header), leaky prediction and concealment at the decoder are simulated for TM1 IP mode (M=1, N=150) and low delay mode, concluding that the scheme is effective, particularly for cell loss sensitive sequences.

AVC-280 SIMPLIFICATION OF FAMC (JAPAN)

Simplification of FAMC in motion compensation is experimented for address generation and interpolation to show that no or only slight degradation in SNR is incurred. Simplification of motion estimation is also experimented for a two-step search comprising of frame-based ME and FAMC ME. As a conclusion, simplified FAMC43 is proposed for prediction.

AVC-281 COMPARISON OF REQUIRED PROCESSING BETWEEN FRAME/FIELD AND FAMC (JAPAN)

The number of operations per second is estimated for adaptive field/frame and simplified FAMC as in AVC-280 with respect to addition and multiplication. FAMC requires 60% operations of the adaptive field/frame, but multiplication is involved and maximum memory bus load is doubled.

AVC-282 CODING/DECODING DELAY AND REQUIRED MEMORYS FOR THE PREDICTION IN TM1 (JAPAN)

Necessary number of field memorys and coding/decoding delay are estimated for the pure field, the frame-based field structure and the frame-based frame structure.

AVC-283 HARDWARE VERIFICATION OF H.26X SPECIFICATIONS (JAPAN)

Study items to prepare for the hardware verification of H.26X are listed for discussion. It is pointed out that the Experts Group need a policy to cope with those items surrounding the coding algorithm, which are not subjects of verification but required to test H.26X.

AVC-284 EXPERIMENT ON LOW DELAY MODE (JAPAN)

TM1 low delay mode is first compared with frame-based {N=15, M=1} in coding efficiency, concluding that there is no significant loss of coding efficiency by use of the low delay mode. Then, it is pointed out that the current coding control method for the low delay mode causes visible moving of INTRA slices, proposing a modified coding control method to mitigate the effect.

AVC-285 TM1 PURE FIELD CODING SIMULATION RESULTS (PTT RESEARCH - NETHERLANDS)

Pure field coding {N=12, M=3} is compared with the adaptive frame/field coding. It is concluded that about equal performance is obtained in picture quality.

AVC-286 TM1 PYRAMID CODING AND COMPATIBILITY VERSUS SIMULCAST (PTT RESEARCH - NETHERLANDS)

A two layer compatible coding (1.5 Mbit/s for base layer and 2.5 Mbit/s for enhanced layer) as defined in Appendix G of TM1 is compared with simulcast (2.5 Mbit/s). Prediction for the prediction error is used for odd numbered fields of CCIR-601 signals. It is concluded that an obtained gain of 0.5 to 0.7 dB in SNR and other factors (full compatibility, scalability and error resilience) favors this hierarchical structure.

AVC-287 ATM CELL LOSS EXPERIMENTS WITH TM1 (PTT RESEARCH - NETHERLANDS)

A two layered coding using pure field based coding with upsampled MPEG1 prediction error for odd numbered field prediction error is compared with one layered coding in cell loss resilience. It is concluded that layered coding can be very useful for transmission over ATM-networks when the base layer bitstream can be sent over a guaranteed channel, pointing out that more study and information is necessary for multi-layered coding.

AVC-288 ERROR SENSITIVITY OF THE TM1 SYNTAX (PTT RESEARCH - NETHERLANDS)

Error sensitivity of TM1 syntax is discussed in resynchronization of bitstream, FEC, forced updating, layered coding, and picture slicing. A specific proposal is made to add two bit identification fields after each extension start code.

AVC-289 PSTN VIDEO CODING ON 8 AND 16 KBIT/S USING A DOWN SCALED H.261 (PTT RESEARCH - NETHERLANDS)

Simulated and hardware processed pictures are presented which have been obtained using H.261 or its subset (without MC) operating at QCIF, 8-16 kbit/s. It is concluded that a video bitstream conforming to H.261 can be a step towards interworking of PSTN videophone service with ISDN one. It is pointed out that call set up duration, end-to-end delay, spatial and temporal resolution, and interworking between PSTN and ISDN are items of further study.

AVC-290 RACE 2072 MOBILE AUDIO VISUAL TERMINAL; PROJECT INFORMATION (PTT RESEARCH - NETHERLANDS)

The objectives are to find a powerful audio and video coding algorithm for transmission of moving and still video in a mobile environment and to implement these algorithms on a demonstrator. Technical approach, key issues, expected impact and participants are also presented.

AVC-291 LOW-DELAY CODING EXPERIMENT (BELLCORE - USA)

Three coding structures are compared in terms of SNR and buffering delay; A(N=15,M=3), B(N=15,M=1), C(N=150,M=1 with forced intra-slices). Simulation results show that low delay mode defined in TM1 reduces the buffering delay by 2 to 3 frames, but at the same time it also degrades the picture quality by as much as 1.84 dB. It is concluded that B and C performs very similarly, but C gives slightly larger buffering delay than B due to that it encodes 28 more intra-slices at the first 15 frames.

AVC-292 FREQUENCY SCANNING AND ENTROPY CODING USING MUVLC (HHI - FRG, RTT BELGACOM, SIEMENS GERMANY)

In this contribution, UVLC technique is mentioned, an improved version MUVLC concerning coding efficiency is described, a proposed syntax addition to TM1, C-programs for bitstream encoding and decoding a slice of macroblocks, and results on coding gain are given.

AVC-293 TM1 COMPATIBILITY EXPERIMENTS (UK)

Use of compatible prediction from the base layer is compared with simulcast in TM1 to conclude that it can improve the SNR picture quality.

AVC-294 A PROPOSAL FOR AAL TYPE 2 (BELGIUM, GERMANY, ITALY, NETHERLANDS, NORWAY, SWEDEN, UK)

Requirements for AAL to support VBR video transport are listed; efficient usage of bits, low delay, extendibility, cell loss detection, bit error protection, compatibility issues, multipoint and multimedia considerations, end-to-end timing and multi-layer synchronization, and interleaving. A generic AAL Type2 and three extensions thereof are proposed to stimulate discussion. It is also pointed out that VBR will most benefit low bit rate video coding in picture quality and delay.

AVC-295 ISSUES CONCERNING THE SUPPORT OF FLEXIBLE SPATIAL RESOLUTIONS (AUSTRALIA)

Necessary side information and encoder/decoder implementation are discussed for two possible solutions of the flexible spatial resolution approach; to encode the signal as being input and to pad the maximum picture size. It is concluded that the first alternative is more appropriate from the view points of transmission efficiency and effective use of codec processing power. It is also pointed out that a flexible width stripe be adopted to support the flexible spatial resolution approach.

AVC-296* CELL LOSS CHARACTERISTICS FOR STATISTICALLY MULTIPLEXED VIDEO SOURCES (AUSTRALIA)

Longer term characteristics of cell loss are analyzed for multiplexed VBR sources assuming a simplified time correlation model of of the video sources and a multiplexing process. Average CLR, average congestion time, congestion time distribution and average CLR during congestion are estimated. The results show that bursts of cell loss can last several frames up to tens of frames and that CLR can be of the order of 0.01 during the congestion.

AVC-297* THE ATM ADAPTATION LAYER FOR VIDEO SERVICES IN THE B-ISDN (AUSTRALIA)

This document gives an overview for the AAL for transport of video services. Methods of multimedia multiplexing are described, presenting features of ATM layer multiplex. Required functions of AAL Type 2 are listed with an example of specific fields. It is also pointed out that cell loss experiments should incorporate some assumptions regarding AAL such as shown in this document.

AVC-298* ADAPTING MPEG1 VIDEO FOR ATM TRANSMISSION (AUSTRALIA)

Experimental results are provided for transporting MPEG1 coded video over B-ISDN. Use of AAL error detecting functionality, dividing picture types into high and low cell loss priority, and varying the size of slice are addressed.

AVC-299 SCALABLE CODING ARCHITECTURE (AUSTRALIA)

A scalable encoder architecture is proposed which uses 4x4 DCT for the second layer, 2x2 DCT for the third layer and prediction for the prediction error. This scheme incorporates multiple independent loops to give quality control of all layers and allowing the rate control to be applied to each layer. In order to prevent loss of coding efficiency, independent motion vectors and coding modes for each layer are introduced. It is proposed that the features discussed in this document be included in the scalability core experiments.

AVC-300* PROPOSAL FOR CELL LOSS CORE EXPERIMENTS ON LAYERED AND NON-LAYERED CODERS (AUSTRALIA)

A proposal is made to define a cell loss core experiment by incorporating short term and long term cell loss characteristics, separation of high and low priority information, reference AAL, use of ATM cell loss priority, use of scalability syntax extension. The aim is to compare non-layered and layered coding structures.

AVC-301 FLEXIBLE ENCODER DEFINED PREDICTIONS (NTR - NORWAY)

A methodology to unify many prediction modes is proposed whereby the encoder is able to specify and down load predictors. It is characterized as follows;

1) The predictors are not fixed in the standard but may be optimized for each application.
2) The method gives room for almost all the known predictors like frame based, field based and FAMC.
3) It gives room for defining new predictions also after the standard is fixed.
4) The predictors may be down loaded on the sequence level (or picture level).
5) Only one vector is used for each macroblock for P-frames/fields.
6) The implementation complexity is comparable with traditional half-pixel motion prediction.

AVC-302 SIMULATION WITH FIELD CODING AND M=1 FOR LOW DELAY (NTR - NORWAY)

Simulation results are given for the TM1 low delay mode (field coding) in terms of SNR and buffering delay. The definition of prediction follows the "encoder defined prediction" method proposed in AVC-301.

AVC-303 SIMULATIONS WITH VERY LOW BITRATES - (8-16) KB/S (NORWAY)

Simulation results are presented for a hybrid DCT scheme operating at 8-16 kbit/s and 32 kbit/s with the picture format QCIF. The average picture frequency of 8-10 Hz is reported.

AVC-304 VERY LOW BITRATE VIDEOTELEPHONY STANDARDIZATION (BELGIUM, FRANCE, GERMANY, ITALY, THE NETHERLANDS, NORWAY, SWEDEN, UK)

It is claimed that videotelephony service on PSTN and mobile network be initiated to cover all areas of service description, infrastructure, systems, terminals and call control. It is also claimed that video coding algorithm should interwork over various networks. Information on relevant European activities is given.

AVC-305 VERY LOW BIT RATE "H.261 LIKE" VIDEO CODING SIMULATION (CNET - FRANCE)

Using a picture format of 1/9CIF and H.261 coding scheme with modifications in GOB structure (removal of GOB) and VLC (3-D VLC), simulations were carried out at 8.4 and 16.8 kbit/s. It is concluded that only 5% improvement in coding efficiency is obtained by those modifications and that the picture quality is not acceptable if the size of display is greater than 3-4 inches.

AVC-306 TM1 PYRAMID CODING FOR THE SCALABILITY REQUIREMENT (PTT RESEARCH - NETHERLANDS)

A proposal is made for a compatible scalable codec structure which uses prediction for the prediction error, pointing out that the structure can be extended for CTV/HDTV compatibility and that it can form a starting point to reach a common solution for the compatibility and scalability requirements.

AVC-307* VBR MPEG BIT-RATE CHARACTERISTICS (DAVID SARNOFF - USA)

This document provides detailed statistics for VBR operation of MPEG1 coder, which have been collected through simulation of more than 3000 frames of video;

1) frame-by-frame bit rate, probability density function and autocorrelation plots for each test sequence,
2) SNR versus average and peak rates,
3) effect of different peak rate control levels on the image quality,
4) comparison of CBR versus VBR image quality for the same average bit rate.

Cell loss resilience and its impact on VBR mode are also discussed.

AVC-308* ERROR CONCEALMENT FOR MPEG VIDEO OVER ATM (DAVID SARNOFF -USA)

Simulation results for cell loss concealment of MPEG1 coded video are reported where CRC in the transport-level detects cell errors and identify the spatio-temporal position of MPEG slices, replacing errored DCT coefficients with EOB for low priority errors, replacing each errored macroblock with an estimate from the spatio-temporally surrounding macroblocks for high-priority errors (or for one-layer transmission). It is concluded that a reasonable quality of restructured pictures can be obtained on one-tier transmission media with packet/cell loss rate as high as 0.1%.

AVC-309 VIDEO TRANSMISSION OVER A RADIO LINK USING H.261 AND DECT (BT - UK)

This document describes an adaptation method of H.261 video codec to a radio link. The improved error resilience comprises of FEC, interleaving of the data stream and repeating corrupted blocks. For both of video and radio parts, experimental hardware can be demonstrated.

AVC-310 MODIFICATIONS TO TM1 TO SUIT BROADCAST APPLICATIONS (AT&T -USA)

Two proposals are made to adapt TM1 to the "tune-in" situation in broadcast applications; inclusion of quantizer matrices in the picture header, and use of leaky prediction where leak factor is transmitted for every P picture in the form of $(1-2^{-n})$.

AVC-311 IVS ACTIVITIES (SWP XVIII/8-3)

This document provides a draft agenda for the IVS Technical Session to be held during 26-27 October before the Experts Group meeting. The objectives are to promote further mutual understanding and to achieve an optimized association between networks and coding techniques.

AVC-312 LIAISON STATEMENTS (SGXVIII)

This collection of liaison statements contain the following items of the Experts Group's concern; possible functions for video signal support of AAL Types 1 and 2, support of H.320 terminals in B-ISDN, traffic control and usage parameter control issues. The second item requires specific answers from the Experts Group.

AVC-313 NON-COMPATIBLE VS SIMULCAST VS COMPATIBLE - EXPERIMENT G.3 (BELLCORE - USA)

Simulation results are compared between non-compatible at 4 Mbit/s (A), simulcast at 2 Mbit/s (B), compatible at 4 Mbit/s with upscaled prediction from HHR base layer at 2 Mbit/s (C), and HHR at 2 Mbit/s (D). It is concluded that A is preferred over B, perceptual quality of C is not far off from A, some prediction improvements are necessary for P and B pictures, and perceptual difference between B and D is greater than it appears on the SNR values.

AVC-314 STATISTICAL ANALYSIS OF VIDEO TELECONFERENCE TRAFFIC - III (BELLCORE - USA)

Analyzing a 32,000 frame traffic data produced by a H.261 codec with motion compensation, it is confirmed that the number of cells per frame for video conferences follows a gamma distribution. It is also shown that the density function fits with that of a Markov chain.

AVC-315* TIMING RECOVERY FOR VARIABLE BIT-RATE VIDEO ON ATM NETWORKS (AT&T - USA)

For the MPEG system (multimedia multiplex), several methods are described to recover timing in video systems where the transmission delay jitter may be substantial as in packetized networks. It is also pointed out that for high quality real time audio and video display, the stability specification of MPEG may not be stringent enough, requiring additional time base correction.

AVC-316 RECOMMENDATIONS FOR CELL LOSS EXPERIMENTS (DAVID SARNOFF - USA)

Some core experiments are suggested, taking into account codec mode, ATM transport mode, ATM AAL features, ATM network scenarios, and decoder concealment features.

Temporary Documents

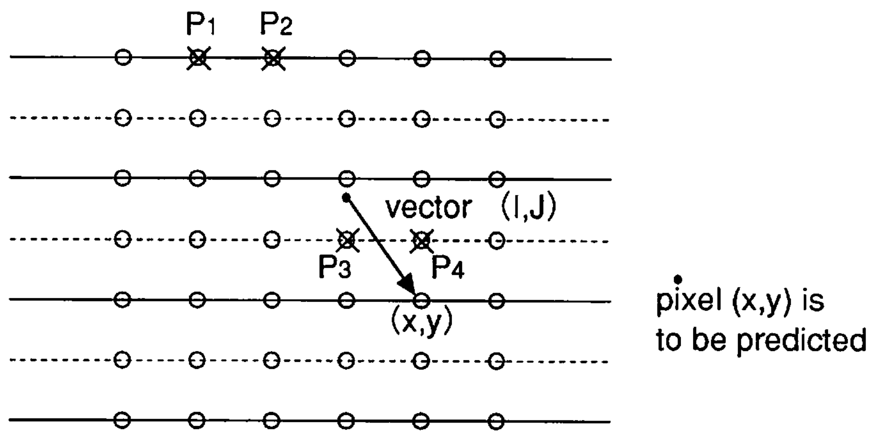
TD-1 Agenda (Chairman)
TD-2 Available documents (Chairman)
TD-3 List of tape demonstrations (Chairman)
TD-4 Generic prediction in AVC-301 (Chairman)
TD-5* Report of small group meeting on low delay mode (Mr. Bjoentegaard)
TD-6* Report of small group meeting on cell loss resilience (Mr. Biggar)
TD-7* Report of small group meeting on H.261 compatibility (Mr. Parke)
TD-8 Draft meeting report for the sole sessions (Chairman)

END

List of Tape Demonstrations
(1 July 1992, Holmdel)

No Organization	Topics	Tape	Doc.
a. NTR	- Down loadable prediction - Low delay mode	D-50	AVC-301 AVC-302
b. NTR	- 8/16 kbit/s coding	U-50	AVC-303
c. BT	- TM1 compatible experiments	D-50	AVC-293
d. BT	- H.261 on DECT	D-50	AVC-309
e. PTT Research	- TM1 pure field coding	D-50	AVC-285
f. PTT Research	- Compatibility - TM1	D-50	AVC-286 AVC-306
g. PTT Research	- Cell loss	D-50	AVC-287 AVC-288
g. PTT Research	- PSTN videotelephony	D-50	AVC-289
h. CNET	- Very low bit rate video	U-50	AVC-305
i. Fujitsu	- Cell loss compensation scheme	D-60	AVC-280
j. Mitsubishi	- Scene change handling in low delay mode	D-60	AVC-275
k. NEC	- Experiments on low delay mode	D-60	AVC-284
l. Matsushita, JVC	- Comparison of prediction methods for the low delay mode	D-60	AVC-276
m. Bellcore	- Low delay mode simulation	D-60	AVC-291
n. Bellcore	- Simulcast vs compatible	D-60	AVC-313
o. AT&T	- Leaky prediction	D-60	AVC-310
p. NTT	- H.261 operation at low bit rates	U-60	
q. KDD	- Picture quality comparison of format	D-60	AVC-274
r. Toshiba	- Simulation results on compatibility	D-60	AVC-277

Generic prediction in AVC-301



Decoder

(I,J) → generic prediction vector

$$\tilde{f}(x,y) = \sum_{i=1}^{4} w_i \cdot f(x+\Delta x_i,\ y+\Delta y_i)$$

→ $\tilde{f}(x,y)$ prediciton value for pixel (x,y)

Loaded table

entry	Value
I,J	$\Delta x_1, \Delta y_1$; w_1 $\Delta x_2, \Delta y_2$; w_2 $\Delta x_3, \Delta y_3$; w_3 $\Delta x_4, \Delta y_4$; w_4

$\Delta x_i, \Delta y_i$ deflection
w_i weighting

ISO/IEC JTC1/SC29/WD11
MPEG92/ 397
July 1992

Title: Report from the small group on low delay mode.
Source: CCITT Experts group for ATM Coding.
Purpose: Proposal.

The following items are dealt with :

- Coding structure - field/frame, value of M and preferred prediction.
- Handling of scene change to maintain low delay.
- How to handle the first picture in test model simulations.
- Definition of intra slice coding.
- Buffer control.
- Influence of leaky prediction on low delay coding.

Coding structure.

In the low delay profile it is assumed that the total coding/decoding delay shall be kept below 150 ms. However, it is realized that coding/decoding delay considerably below that limit could be desirable for many applications (e.g. two way communication over satellite links).

In the following the delay will be measured in "field periods" - fp. For 50 and 60 Hz pictures the 150 ms correspond to:

- 50 Hz -> 150 ms = 7.5 fp.
- 60 Hz -> 150 ms = 9.0 fp.

The coding/decoding delay is considered to consist mainly of two parts:
- Buffer delay. A typical value is 5 fp. This is mainly due to the regular INTRA picture update. With a different update mechanism it is assumed that the buffer delay may be reduced to 2 fp.
- Delay resulting from frame/field reordering. This will be called basic delay.

The sum of these two contribution should be below 150 ms - or as low as possible. The possible coding modes fulfilling this criteria are listed in the table below.

Predictor performance:
Coding performance is important also for low delay coding. Performance depends to a large extent on the quality of the predictor. The preferred predictors for the different modes are:

- Field coding and M>1: The predictors in TM1 performs well.
- Frame coding and M=1: FAMC - or similar - are the preferred modes.
- Field coding and M=1: A prediction mode similar to FAMC have shown very promising performance and would be preferable. The syntax is the same as for field coding with one vector pr. MB as defined in TM1.

"FAMC - or similar" is mentioned. This takes into account the proposal of flexible encoder defined predictors which may define "FAMC like" predictors for field predictions as well as for frame prediction. In addition this method leaves room for future defined predictors.

The table below list:
- Possible coding modes that fulfills the delay criteria.
- Preferred prediction modes (for M=1).
- Different delays.
- Preliminary performance ranking. This is far from complete due to lacking results from core experiments. (For "field, M=1" there is one simulation but no direct comparison).

Coding mode	M	Buffer delay	Basic delay	Total delay	Performance ranking (low="good")
Frame FAMC - or similar	1	5 fp	2 fp	7 fp	2
Field FAMC - or similar	1	5 fp	0 fp	5 fp	-
Frame FAMC - or similar	1	2 fp	2 fp	4 fp	2
Field FAMC - or similar	1	2 fp	0 fp	2 fp	-
Field structure	2	2 fp	4 fp	6 fp	1
Pure field 2-B	3	2 fp	3 fp	5 fp	2
Pure field 3-B	3	2 fp	5 fp	7 fp	1

Handling of scene change to maintain low delay.

A major contribution to the buffer delay is the complete INTRA frames/fields coding. To get around this delay it is necessary for the low delay profile to have the possibility of picture skipping. The encoder may decide to skip frames after a "large" picture due to scene cut. The frame dropping must be signaled to the decoder which will display the last decoded field until the first picture after the scene cut is reconstructed. The decoder will skip displaying the pictures skipped by the encoder.

It should be considered whether the picture skipping will have any influence on the syntax or operation of the VBV.

How to handle the first picture.

In the low delay profile it is of interest to study the "steady state" performance of buffer delay. It is therefore desirable to pretend that we are in the steady state situation from the beginning of the sequence. The proposed way to obtain this is the following:

- Code the first picture INTRA with QP=16 (for 4 Mb/s).
- After the first picture the buffer is set to a value corresponding to 1/60th of a second (66667 bits for 4 Mb/s).
- The number of bits for the first picture - used for buffer regulation - is set to the average number of bits for the sequence.
- In evaluating the delay only buffer filling after the first picture is considered.

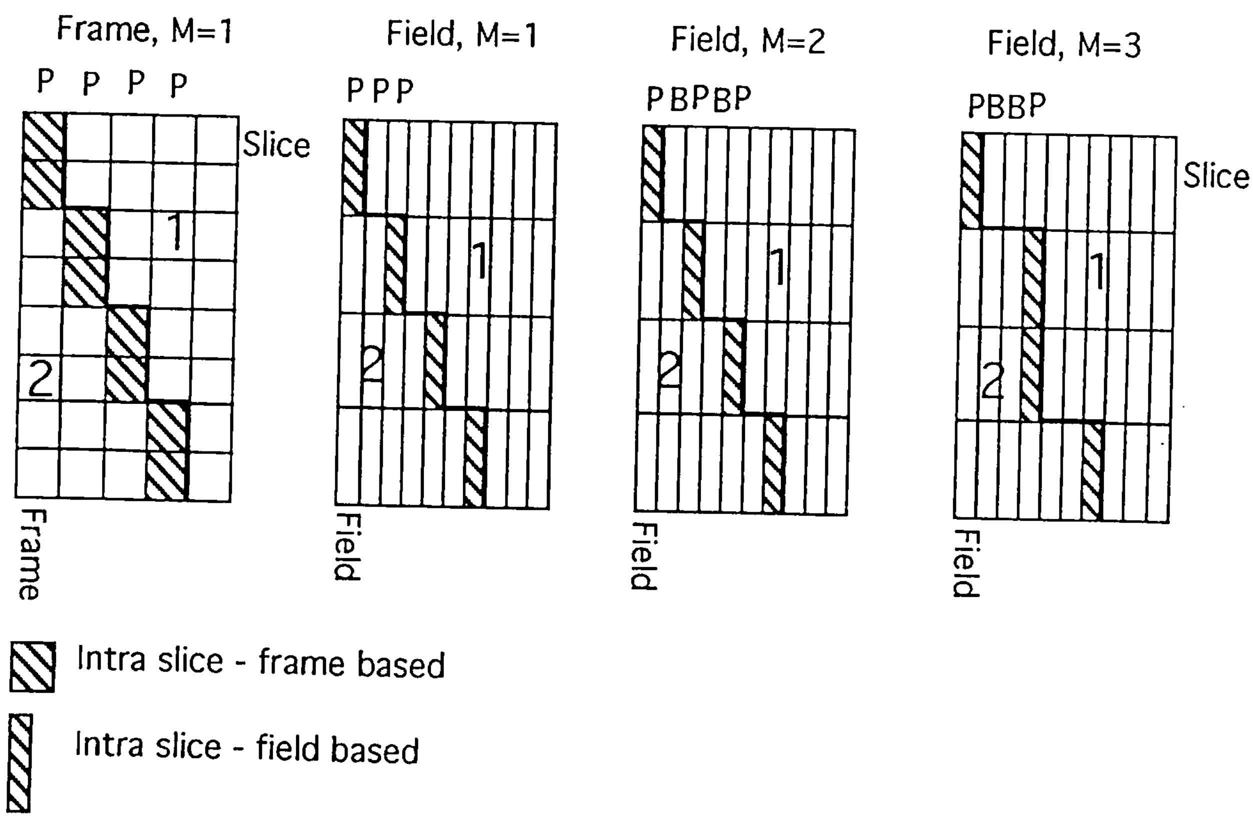
Definition of intra slice coding.

To reduce buffer delay the updating is made with forced INTRA slices rather than forced INTRA pictures. The procedure for doing the forced INTRA slice coding is shown in the figure below. All the modes given in the table above are covered. The updating for the different modes are arranged so that the time for total update is the same for all modes.

- Frame mode, M=1: Two slices pr. frame.
- Field mode, M=1 and 2: One slice every other field.
- Field mode, M=3: One slice for even P-fields, two slices for odd P-fields.

To guarantee that errors do not propagate there is restriction on predictions in region 1. Motion vectors in region 1 may not refer to areas in region 2 (refer to the figure).

With this method the time for total refresh is 500 ms for 60 Hz sequences and 720 ms for 50 Hz sequences. It should be noted that the maximum entry times for channel hopping are twice as large as the mentioned values.



TD7

Report of the small group on H.261 compatibility

The discussion focused on two main areas:

- the status of H.261 compatibility, and
- definition of a core experiment to help determine the most suitable compatible prediction method.

H.261 Compatibility status

At present there have been no experiments performed on H.261 compatibility.

For interworking between video services of B-ISDN and N-ISDN (point-to-point, multipoint, multicast and database access) a decoder must be able to accept both H.26x and H.261 bitstreams. In some configurations the encoder must be able to produce both H.26x and H.261 bitstreams (AVC-278).

There have been several investigations into MPEG-1 compatible coding. These have shown that compatible coding with embedded bit streams (by means of layered coding) performs better than a simulcast approach (AVC-293, AVC 286) see table below.

Several embedded coding schemes have been proposed. We need to select one. Some direct comparison has been performed (AVC-277). To help make the selection a core experiment has been defined to compare the layered coding schemes. See ANNEX 1.

The layered coding schemes have been extended to H.261 compatible coding in the experiment. In defining the experiment simplifications have had to be made. These are:

- The picture format chosen for the H.261 base layer is SIF.
- The H.261 base layer (RM8) does not allow picture dropping.

These simplifications do not invalidate the comparison being made in this experiment. However it does raise the following questions:

- What picture format/s is H.261 to be compatible with?
- What target bit rates are to be used in H.26x?
- How do we handle picture dropping?

The picture format/s will have impact on the up-conversion method needed for the compatible prediction.
The target bit rates will determine when simulcast or embedded coding is used.

Further work

The compatible coding performance is one aspect to how H.261/H.26x compatibility is achieved. Other factors have been highlighted above. Further study is needed in these areas.

Suitable test sequences are required for the H.26x/H.261 compatibility investigations.

Comparison of SNR figures.

AVC Doc		Flower	Table	Mobile	Football
285 & 286	Error Pred	29.63/29.92	33.89/33.05		
	TM1 2.5M	28.39/28.67	33.15/32.97		
	TM1 4M	30.90/31.15	34.99/34.82		
293	Base Pred	28.44		28.06	
	TM1 2.5M	27.09		27.33	
	TM1 4M	29.44		29.65	
277 & 234	Error Pred	28.56			35.60
	Base Pred	28.91			36.25
	TM0 2.5M	27.86			33.85
	TM0 4M	29.86			35.71

ISO/IEC JTC1/SC29/WG11
MPEG92/...... 399
July 1992

Title : An experiment to investigate H.261 compatibility and to help determine the most suitable compatible prediction method.

Purpose : Core experiment

Source : CCITT Experts group for ATM Video Coding

At the present moment there are two basic methods to achieve compatibility; simulcast and embedded bit streams. Many experiments have been performed which compared simulcast and embedded bit streams.
These have shown embedded bit streams to give better performance than simulcast. However there has been no direct comparison of the embedded bit stream approaches. There have been several embedded bit stream approaches proposed. These have used two layer coding technique with some means for a compatible prediction in the top layer. The compatible prediction modes proposed have been :

1. Prediction from the locally decoded base layer picture
2. Prediction of the prediction error in the second layer using the prediction error of the base layer.

In this experiment we propose to compare these two methods and to also compare against simulcast.

EXPERIMENT DETAILS

Picture format

The base layer is SIF.
The conversion to SIF is as in TM1, section 3.3.2 of TM1.

Coding of SIF

The coding is performed using RM8.
The motion estimation is full search with a motion vector range of +/- 15.
For SIF 240 the GOB's 11 and 12 are coded as black.
Frame dropping is not allowed.

Compatible prediction method

This is done on a macro block basis.
Depending on the compatible prediction method the prediction is generated from the corresponded 8x8 prediction error block of the lower layer or the corresponding coded 8x8 block of the lower layer.
This 8x8 block is than up sampled to a 16x16 block using a [1/2,1,1/2] filter.
A similar operation is performed to the 4x4 chrominance subblocks to give a 8x4 prediction block.

Compatible mode selection

Mode selection is made on 16x8 blocks of the macro block separately. The selection is based on the minimum power of the prediction errors.
A two bit code is used to indicate which fields in the macro block are coded compatibly; page 56 of TM1.

Top layer coding

This is TM1 using frame/field adaptive prediction only : no FAMC, no dual field.
The coding structure is I,P,P,P,P,P ie. M=1 , N= length of sequence

Bitrate :

H.261 bitrate are :	112 kbit/s	2B
	320 kbit/s	6B
	1500 kbit/s	24B

Total bitrate are 2 and 4 Mbit/s

H261 bitrates picture rate / H.26x bitrates	112 kbit/s 10/8.33Hz	320 kbit/s 30/25Hz	1500 kbit/s 30/25Hz
2 Mbit/s	x	x	
4 Mbit/s		x	x

These are the simulcast bitrates :

H261 bitrates / H.26x bitrates	112 kbit/s	320 kbit/s	1500 kbit/s
2 Mbit/s	1888kbits/s	1680kbits/s	
4 Mbit/s		3680kbits/s	2500kbits/s

Test sequences :

Susie
Table Tennis

Results

The following statistics are needed:
SNR figures for the H.261 base layer,
SNR figures for the H.26x top layer,
SNR figures for simulcast,
SNR figures for non-compatible coding at 2 and 4 Mbits/s,
Percentage of macroblocks predicted from H.261 base layer.

end

CCITT SGXV Experts Group on ATM Video Coding
July 1992

TD-6

MPEG 92/398

Cell Loss Resilience: Issues and Core Experiments

This document is intended to summarise the main issues that should be considered in the definition of core experiments to determine suitable means of protecting against cell loss.

Approaches to Cell Loss Resilience.

Cell loss resilience techniques can be divided into two groups; reactive and proactive.

Reactive techniques

In this document, the term *cell loss concealment* will be taken to represent methods applied only at the receiver to conceal the effects of cell loss. These techniques need not be standardised because they apply only at one end of the communication link, but they have a significant impact on the visual effect of cell loss and could be sufficient either alone or in conjunction with other methods to provide adequate picture quality in the event of cell loss.

Standardisation is not necessary or desirable, but reports of cell loss experiments must include complete descriptions of any optional cell loss concealment methods employed, and all experiments that use concealment must also be conducted without.

Proactive techniques

Proactive techniques are applied at both ends of the link. Four approaches are identified here:

Structured Packing

Video data may be transmitted in a continuous "streaming mode", that might involve addition of pointers within cells to picture structure boundaries (e.g. macroblocks) as a means of allowing resynchronisation in the event of a cell loss. Alternatively, a more structured approach may be used, in which the AAL functionality defines discrete picture units (e.g. slices) that are segmented and transmitted as discrete entities. The AAL approach may involve a lower cell packing efficiency, but this loss should be quantified and weighed against the benefits of AAL functionality and the complexity of indexing for extraction of the required macroblock data in the streaming mode.

In discussing AAL functionality in this document, we are concerned with the functionality itself rather than its actual location in a field of the AAL layer or in the higher layers. It is recognised that there remain many open issues concerning the use of the AAL for video service support, including:

- What functionality should the AAL provide?
- How generic should the AAL be?
- Should the AAL be tailored for the specific application and specified at call establishment?

Whether the functions reside in the AAL or above, the following functions (at least) are expected to be required to support the video service over ATM:

- Sequence Number, to assist with cell loss detection;
- Error protection of the larger data unit (e.g. CRC);
- Segment Type (i.e. Beginning of Message, Continuation of Message, etc.).

Protection of the AAL header fields will also be important.

Studies are required to determine the appropriate fields (again, whether they reside in the AAL layer or above), as well as their length.

It is noted that these functions may play a major role in several cell loss resilience schemes and so definition of a suitable reference system is a high priority.

Leaky Prediction

Leaky prediction allows controlled recovery of the picture quality in the event of an error, at the expense of some coding efficiency. It is also recognised as having other advantages, in particular that of permitting rapid establishment of a picture when decoding from an arbitrary point in the bit stream (channel changing).

A core experiment is recommended to determine the costs and benefits of leaky prediction in its role in combatting cell loss. The aim is to determine the relative efficiency between a Group of Pictures approach and leaky prediction, and to compare cell loss resilience for equivalent efficiencies. The core experiment conditions are defined as follows:

An AAL is required. Its fields and functions are open to the investigator, but it must be completely described.

The Leak Factor is to be transmitted in 3 bits in each P picture header. Calculation of the Leak Factor is as defined in Document AVC-310. Leakage is towards mid-grey (value 128).

The precision of multiplication used in the experiment must be specified; the reference method is floating point calculations with truncation towards 128 in 8 bits.

Cell loss is defined as in TM1.

Spatial Layered Coding

Layered coding offers the possibility of using lower resolution information to complete a picture even when higher resolution information is lost due to cell loss. The current work on scalability provides a layered coding description that is suitable for use in cell loss resilience core experiments utilising layered coding. The aim of the experiment is to determine the advantage of layered over non-layered coding under cell loss conditions, and whether this outways any loss of coding efficiency that layered coding imposes.

Core experiment conditions are as follows:

An AAL is required. Its fields and functions are open to the investigator, but it must be completely described.

The layered coder must conform to the scalability or compatibility profile syntax.

The non-layered coder has exactly the same form, but with all lower layers turned off.

Total rate of both coders should be constant and the same.

The same cell loss model (from TM1) is applied to both coders. The layered coder utilises both cell loss priorities; the nonlayered coder utilises only the low priority.

It is recognised that there are limitations to this experiment, since it ignores such issues as the impact of variable rate coding and possible different network utilisation that may occur when both cell loss priorities are used as opposed to when only one is used. The above definition is a reasonable first step, that may be further refined as more knowledge is gained.

Use of CLP with non-scalable coders

The CLP capability of ATM transmission may be utilised even when spatial layering is not provided, by separating the most important from the least important information in a coder stream, and transmitting these separately with approporiate priorities.

The choice of parameters and quantities (for example, I, P and B frames, addressing, motion vectors, low and high frequency coefficients, motion vectors) to place in each priority stream is unknown at this stage, and inputs are requested.