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

Document AVC-401
October 30, 1992

CCITT
STUDY GROUP XV
Geneva, 5-10 November 1992

Temporary Document 23 (XV/1)

Questions: 3/XV, 4/XV

SOURCE: CHAIRMAN OF THE EXPERTS GROUP FOR ATM VIDEO CODING
TITLE : FOURTH PROGRESS REPORT

---

1. General

We have met three times since we presented the third progress report (Annex 1 to COM XV-R97-E) at the previous Working Party XV/1 meeting in May 1992;

7th meeting	1-10 July 1992 in Eatontown (USA) at the kind invitation of AT&T Bell Laboratories and Bellcore and in Rio de Janeiro (Brazil) at the kind invitation of ABNT.
8th meeting	28 September - 1 October 1992 in Tarrytown (USA) at the kind of David Sarnoff Research Center and IBM.
9th meeting	27 October - 6 November 1992 in Ipswich (UK) at the kind invitation of BT Laboratories and in London at the kind invitation of BSI.

We had joint sessions with ISO/IEC JTC1/SC29/WG11 (MPEG) as well as CCITT sole sessions in these three meetings. The list of participants appears in Annex to this report.

This document reports major achievements toward defining Recommendation H.26X for video coding in the ATM environments, and particular items for consideration of the Working Party XV/1.

2. Overall workplan

The current work is proceeding as intended toward freezing technical specifications in March 1993 and completion of the Recommendation in 1994 (see Figure 1/Annex 1 to COM XV-R97-E).

3. Source coding

3.1 Test Model development

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

A preliminary working draft document (PWD) for definition of the first test model was made available as outcome of the Singapore meeting. The first Test Model (TM1) was defined in March 1992 and it evolved into TM2 in July 1992 (revised in September 1992), and TM3 in November 1992. This TM document includes definition of the coding algorithm as well as necessary experiments .

3.2 Elements of coding algorithm

The coding algorithm consists of the following elements;

- Pre/post-processing for source format conversion

- Pre/post-processing for source format conversion
- Layered structure of video data
- Motion estimation and compensations
- Modes and mode selection
- Transformation and quantization
- VLC
- Bitstream syntax
- Rate control and quantization control

3.3 Experiments

Currently the following experiments are described in the TM2 document;

1) Prediction modes including low delay operation

   - efficient prediction methods
   - minimizing steady state delay
   - handling of scene changes

2) Compatibility

   - configuration of flexible layered coding scheme
   - forward/backward compatibility between H.26X/H.261

3) Cell loss and error resilience

   - spatial localization: concealment, structured packing
   - temporal localization: leaky prediction, intra slice/intra picture
   - layered coding

An open question is whether the structure of video is to be related to the structure of cell. Structured packing causes penalty for all users; adjustable mechanism is preferable.

4) Quantization
5) Frequency Scalability

We are particularly concerned with the first three items from communication applications point of view and concentrating our efforts.

4. Materializing the "generic standard" philosophy through the Test Model work

The coding standard shall be "generic". It must therefore fulfill requirements set by different applications. The balance should be carefully chosen between maximizing common elements among applications and optimizing for particular applications. Discussion has been initiated on this topic in parallel with developing effective tools for coding.

5. Network aspects

5.1 VBR vs CBR

Several statistics have been reported on the information generation of VBR mode as a basis to study statistical multiplex gain.

Transmission delay has been 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.

Long term characteristics of cell loss have been 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 is proposed. An implication is that once congestion takes place, very bad condition may continue for tens of frame time due to autocorrelation of the video

5.2 Clock recovery

The following items have been addressed 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 is ongoing.

5.3 AAL

Two example AALs have been provided for video signal support to stimulate the group, while a 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. 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.

5.4 Multimedia multiplexing

The Experts Group considered how we can make progress on the multimedia multiplex study. It is common understanding that we should study several alternatives in parallel until network characteristics (cell loss, cost of separate VCs, differential delay between separate VCs, B-N interworking) become clearer.

Chairman raised a study item on the multimedia multiplex; commonality with MPEG-2 system which is envisaged as extension of MPEG-1 system. This belongs to the user multiplex category; extension of H.221 may be another solution as a user multiplex method. Analysis of the problem is awaited.

6. Picture format

The choice of coding format(s) out of

- single format
- multiple formats
- flexible formats inside an agreed maximum

and a method to guarantee interregional communications have log been studied. In the light of lacking clear application possibilities, a proposal was made to make H.26X a generic coding standard and to commence when appropriate on "application profiles" which will include picture formats, bit rates and other parameters. We agreed to this proposal, recognizing that the issue is transferred to the H.32X terminal specifications.

7. Hardware verification

We initiated to plan hardware trials which are 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. Various components and their interfaces of the trial system should be clarified. Another difficulty is that there may not be appropriate networks available. late 1993 or early 1994.

8. Very low bit rate video coding

At the July meeting, the following discussion took place;

8.1 Background

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.

8.2 Experimental results

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 , 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.
- 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.

8.3 Policy

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.

At the September meeting, the following discussion took place;

European countries expressed that it is appropriate to develop standards for the PSTN videotelephony complementing ISDN ones and suggested guidelines. Japan stated that they will submit a paper to the WPXV/1 meeting in November, addressing interworking between

PSTN and ISDN videophones. Mr. Schaphorst informed the meeting that T1A1 is establishing a related project subject to the last approval.

The meeting considered what should we recommend to Working Party XV/1 as a continuation of the New Jersey meeting in July. Chairman proposed to recommend organizing an experts group for this task, covering both short term and long term study items.

At the November meeting, the Experts Group agreed to the above conclusion and drafted the terms of reference of the new group as attached in Annex.

9. Future activities (CCITT sole sessions / Joint sessions with MPEG)

The 10th meeting	21-29 January 1993	Italy
The 11th meeting	29 March - 7 April 1993	Sydney, Melbourne
The 12th meeting	July 1993	New York

10. Specific items requiring the consideration of Working Party XV/1

1) Work plan for broadband audiovisual Recommendations, liaison to SGXVIII
2) Common text approach for H.26X|MPEG-2

We have been concentrating so far on achieving a generic video coding standard H.26X aligning with MPEG-2 according to the decision at the first meeting of the group in November 1990. Most of our efforts have been to incorporate functionalities of our concern (such as low delay, cell loss resilience, H.261 compatibility) from the audiovisual communication system point of view. A difficulty is in the discrepancy between the freezing time of specifications (March 1993) and the time when B-ISDN network services will be made available for our applications (1995 or beyond?), but this way of working is generally considered right and profitable to CCITT.

Chairman sought the support of the group to take the common text approach for H.26X|MPEG-2. Due to the time discrepancy mentioned above, there may happen a situation that new unknown requirements are found after the specification freezing time. This should be taken into account and necessary procedures for adding new specifications in the future should be checked against CCITT and ISO/IEC rules. With this proviso, the meeting supported the common text approach, subject to confirmation of Working Party XV/1. Desire of MPEG to seek the common text approach wit h CCITT should also be confirmed.

We recognize that the generic coding approach leaves many system issues to the terminal Recommendation H.32X. To make a complete interworkable system, we need standards for multimedia multiplex (H.22X), communication procedures (H.24X), call control (AV42X) and whole terminal (H.32X). Though strictly speaking these are not under responsibility of the Experts Group, this group is an only body in WPXV/1 working for B-ISDN audiovisual systems, thus wishes contribute to the progress. We should draw attention of Working Party XV/1 to this situation.

3) Very low bit rate videophone activities

As in Section 8.

END

Annex

Proposed terms of reference
of
Experts Group for Very Low Bitrate Visual Telephony

The terms of reference for the short term standardization are as follows;

1) to study video coding algorithm derived from existing standards such as H.261 applicable to several kbit/s and above,

2) to study speech coding at low bit rates around several kbit/s and above,

3) to study multiplexing methods (bit multiplexing and packet technique) for audio, video and other auxiliary signals,

4) to study methods for error correction and/or other error resilience,

5) to study modulation and demodulation methods to transmit the multiplexed audiovisual signal on PSTN, low bandwidth telecommunication networks, and/or mobile environments

6) and then to draft a Recommendation for the visual telephone terminal by the frist meeting of Study Group XV (September 1993) if required.

It is suggested to proceed as follows;

1) not to develop a new specific ISDN teleservice because of the severe impact on the network,

2) interworking between new very low bit rate video telephones and ISDN terminals is important, but to achieve it without any modifications of existing standard H.320,

3) not to go below 9.6 kbit/s for the total bit rate of audio and video from service quality consideration,

4) to use existing or emerging CCITT standardized modems.

Since it is recognized that the long term study for very low bit rate audiovisual communications is essential for CCITT, the group is also charged with recommending WPXV/1 of the following;

1) potential audiovisual applications at very low bit rates,

2) long term study plan (for the CCITT study period 1993-1996 or beyond) for audiovisual applications at very low bit rates (this work plan is expected by the end of 1993),

3) work method to collaborate with other standardization groups, particularly ISO/IEC JTC1/SC29/WG11.

END