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

Document AVC-109
September 24, 1991

SOURCE : Experts Group for ATM Video Coding in CCITT Study Group XV
TITLE : Status Report on ATM Video Coding Standardization, Issue 2
Purpose: Report

Contents

Note 1 - { } indicates Editor's note, particularly items to be filled.
Note 2 - Editors are listed at the end of this document.

1. Introduction

The Experts Group for ATM Video Coding was established at the July 1990 meeting of Study Group XV to develop video coding standards (Recommendation H.26X) appropriate for B-ISDN providing ATM transport. The group consists of members from FRG, Australia, Belgium, Canada, Denmark, USA, France, Italy, Japan, Norway, The Netherlands, UK, Sweden and Switzerland under chairmanship of Mr. S. Okubo (NTT, Japan). We had three meetings up to now; November 1990 in The Hague, May 1991 in Paris and August 1991 in Santa Clara.

The terms of reference for the group are as follows;

1) to study video coding algorithm appropriate to the ATM environment for conversational services, particularly to study whether modifications are necessary to make the present AV Recommendations applicable to the ATM network;

2) to study the relationships between video coding algorithm and network parameters such as average and peak rates, burstiness, and peak duration to achieve good picture quality and traffic characteristics;

3) to study feasibility of a unified coding standard for various applications in all service classes using the ATM network for which different hardware versions (codecs) can be realized;

4) to investigate potential applications for ATM coding systems (conversational, distributive, retrieval of stored information, etc.) and network-related constraints on potential system performance, and to develop a set of requirements and constraints to guide the work of the Group;

5) to study compatibility of the new algorithms with audiovisual systems covered by Recommendation H.200;

6) to coordinate directly with the experts of other CCITT Groups, CMTT and ISO/IEC on video coding;

7) to draft Recommendation(s) for video coding under the ATM environment.

This document has objectives to form a common ground among members of the group, to clarify the points to be worked out, and thus to promote further progress. It is also intended that this document serves to publicize the activities of the Experts Group. The contents will be updated according to the obtained study results.

It is noted that Study Group XVIII has issued "Integrated Video Services (IVS) Baseline Document" which is intended to provide a basis for harmonizing the work of the wide range of groups involved in video services to ensure consistency with B-ISDN. The group is in a position to contribute to enhancing the parts under its responsibility.

{brief description of each chapter}

2. Terminology

Here is given terminology for terms particular to the ATM video coding. Note that general B-ISDN terms are defined in Recommendation I.113.

The following items are defined in Annex 1.

GENERAL TERMS

CODEC
CIF (Common Intermediate Format)
compression
FEC (Forward Error Correction)
interoperability
lip synchronization
motion compensation coding
multipoint
QCIF
quantization
teleconferencing
videoconferencing
videophone
visual telephone

COMPATIBILITY RELATED TERMS

backward compatibility
downward compatibility
embedded bit stream

forward compatibility
layered coding
simulcast
standard families
switchable encoder
syntactic extension
upward compatibility

3. Applications

Possible applications in B-ISDN and their requirements are listed in Table 1, which are summarized in general term as;

- conversational services,
- distributive services,
- retrieval services,

with stress on their multimedia nature. It is also noted that Draft I.211 gives an extensive list of B-ISDN services.

Our objective is to define a unified coding which can cover the above mentioned services, rather than to confine to a specific service.

Table 1

4. Boundary conditions for ATM video coding

4.1 Target networks

We focus on B-ISDN for networks to which the new video codec is applied but do not preclude such networks as LAN and MAN as far as they are ATM based.

4.2 Network characteristics

ATM network characteristics to which our new video coding should adapt are summarized in Table 2, listing opportunities as well as limitations (see AVC-20, AVC-4).

Table 2

A list of questions have been formulated and sent to SGXVIII concerning ATM network characteristics that affect the design of the video coding. Items of our concern are as follows;

- Cell loss ratio
- Cell loss burst behavior
- CLP bit
- Usage parameters
- Multimedia connections
- Bit Error Rates
- Cell delay jitter
- Network model for hardware experiments
- AAL
- Network interworking

SGXVIII have responded to these questions. These questions and answers are listed in Annex 2.

We endorse the IVS baseline document approach which was made in the SGXVIII meeting in November 1991 and we keep contributing to its upgrading. Comments for this purpose have been sent to SGXVIII in each meeting of the Experts Group. Most recent comments produced at the August 1991 meeting are in Annex 3.

4.3 Technical requirements for ATM video coding

4.3.1 Video signals to be handled

Initially we concentrate on video coding of QCIF/CIF and "CCIR-601 class" standard television signals but at the same time will try to accommodate extension to EDTV and HDTV.

4.3.2 Picture quality target

The target is defined as a range between conversational service quality and distribution service quality, awaiting the quantification in the future.

4.3.3 Processing delay target

Processing delay of the new video coding should be less than that of the current systems, e.g. less than about 150 ms.

> Note - The "less than about 150ms" target comes from the HRD definition (buffer size = 4 times CIF picture period) in H.261.

4.3.4 Average bit rate

It should cover a range up to several tens of Mbit/s.

4.4 Compatibility issues

Further study is necessary on the balance of achieving compatibility and highest coding performance. We agree on a guideline that the compatibility between the new coding system and existing and emerging systems should be highly respected, and that the means for interworking between H.320 terminals connected to the N-ISDN and H.32X terminals connected to the B-ISDN should be developed in this group.

Compatibility is classified into four types, i.e., upward/downward and forward/backward, and the five methods to achieve this property are defined. These are simulcasting, embedded bit stream, syntactic extension, switchable encoder and standard families. See Annex 1.

Layering is one of the solutions and the idea of "flexible layering" which exploits B-ISDN characteristics and provides service integration is recognized interesting. "Flexible layering" realizes interworking among different "service classes" by switching off some constituent layers as necessary. Such layered coding examples are found in Documents AVC-74,94,100,103.

Exact ways to implement compatibilities require further study.

4.5 Requirements for H.26X

In conclusion, a list of H.26X requirements is given in Annex 4.

5. Picture format

5.1 Points of common understanding

The formats QCIF/CIF are already defined as formats for communicational services. Initially QCIF/CIF and the CCIR 601 formats will be used for simulations. It is furthermore agreed to use the CCIR 601 formats for the initial assessment of picture quality (Kurihama test).

We reached the following interim common view at the August 1991 meeting;

The format conversion is not an issue of standardization.

There is a common view that a maximum resolution for a "601" class or services shall be defined. There is furthermore support for the following parameters:

- Progressive format.
- Maximum frame rate of 59.94 frames/sec.
- Maximum 576 vertical lines.
- To support the CIF Pel Aspect Ratio.
- Monitor aspect ratio 4:3.

Agreement on the picture format issue shall be obtained at the earliest occasion in 1992.

5.2 Discussion points

A discussion has been initiated in the Experts Group on whether to define a small number of formats in addition to QCIF/CIF and if the main focus should be on conversational services only. The discussion will be continued during the coming meetings.

The key issues in the discussion may be summarized as:

- Should there be one common(worldwide) format with higher resolution than CIF?
- Should there be a relation between a possible new common format and CIF similar to the relation CIF/QCIF - progressive or interlaced?
- To what extent is it feasible to have different input- and transmission formats?
- When will there be cameras available supporting other formats than CCIR 601?
- Is the maximum format to be defined as SCIF?
- Which subset of the maximum format is allowed?
- Square pels or not?
- How should the "601" class of formats be related to the next class (EDTV or HDTV)?

5.3 Action points

As a basis for answering the above questions, and thereby arrive to a conclusion to the format question, the following action points were listed for the next meetings;

- Simulations to demonstrate the feasibility of conversion between CCIR input formats and a possible common format with acceptable quality loss, delay, loss of coding efficiency, etc.

- A listing of PROS and CONS for different format solutions; single fixed SCIF approach, maximum service class approach, mandatory and optional formats approach, etc.

- Investigations on when cameras (and monitors) for a possible common format other than CCIR 601 could be available.

- Providing source material for demonstration of picture qualities for coding different formats and conversion between different formats.

(At the moment there is no satisfactory solution to conversion between 50 and 60 Hz signals. A common format therefore seems to require the availability of cameras to produce the common format directly. Clarification of this possibility therefore seems to be crucial in the discussion of one or two formats.)

6. Network loading model

We feel it would facilitate our making a good progress of VBR study if we use a common network loading model. The first model has been established and updated as in Annex 5. This is based on the probability that the total of bit rates for multiple calls, each of which has given peak and average rates, becomes greater than the capacity of the transmission pipe assuming on/off model sources.

We feel at the same time this model is not sufficient for more precise study and deeper insight due to the following facts;

- The aggregate model does not take into account correlation between arriving frames.
- Video source cannot be accurately modeled by a memoryless ON/OFF model.
- Document AVC-61 shows that a two-state model which takes into account correlation overestimates cell loss ratio.
- A second order AR (autoregressive) model underestimates cell loss ratio.
- It does not take into account dynamics of statistical multiplexing (i.e. source periodicity effect).

The necessary improvements, however, require further detailed studies taking other models, such as in AVC-43,61,75,97, into consideration.

7. Video coding model

(Reference model, called Test Model (TM) this time, for studying video coding appropriate for ATM environments)

8. VBR vs CBR

8.1 Advantages of VBR

Advantages of VBR over CBR could be expected in following domains:

- statistical multiplexing gains
- reduced coding delay
- picture quality
- others

Therefore, we recognize it as an urgent study item to clarify the advantages of VBR video coding against CBR video coding. It is also a common understanding that applications should be clarified where VBR is most effective.

8.2 CBR as a special case of VBR

It was recognized that CBR is a special case of VBR, and that as such, depending on the application, a VBR codec could operate in CBR or VBR mode.

8.3 Framework for further study

For facilitating the study, a framework for further study has been drafted as in Annex 4 to AVC-65R.

(inputs are expected to elaborate all indicated issues)

The document provides some guidelines to reduce the ambiguity when comparing VBR and CBR:

- to freeze as much as possible of the following variables for the comparison (illustrated in Figure 1)

    * buffer size
    * picture quality
    * bit rate
    * codec complexity

- to define a 'long' video sequence for statistical evaluation

Figure 1

The document classifies the parameters that influence SMG's:

- use of priority bit and layered coding
- whether the B-ISDN can offer different QOS's
- Cell Loss Ratio
- possible correlation between sources (including source periodicity)
- control of VBR operation (including buffer size)
- network model used

The document also indicates clock recovery issues, and the possibility of VBR codecs operating in CBR mode.

More detailed questions and issues that have to be solved are found in the document.

8.4 Comparison between VBR and CBR

A series of study results were presented related with the VBR vs CBR comparison, however a consensus on these results has not yet been reached (except of the fact that the studied issues have an impact on the VBR vs CBR comparison). These issues are covered with the above classification.

The first simplified network model (Annex 4 to AVC-22R) has been updated, but significant insufficiencies have been indicated.

{Contributions are invited to improve the model}

VBR bit rate statistics have been presented. It has been demonstrated that a relatively small buffer makes a significant difference to the shape of the source, and that the number of cells per frame for videoconferencing scene with moderate motion and no scene cuts or changes follows a Gamma distribution when measured with an open loop VBR encoder.

{Also statistics of closed loop VBR encoders are invited. The provision of data files containing bit rate in function of time (both for closed loop and open loop) may help to validate network models}

8.5 Delay of VBR coding

Two contributions are addressing coding and decoding delay for VBR; AVC-56 and 90. With some assumptions for simplification, it is derived as follows;

- The size of the encoder output buffer is given as

$$\text{Sbuf (bits)} = \text{Ave (bit/T)} * \text{Swin (T)}$$

where Ave is average bit rate per frame period T and Swin is window size of the network UPC measured in frame time period T. Average bit rate is policed with moving window of size Swin here. Maximum delay is given as

$$L = (\text{Ave/Peak}) * \text{Swin}$$

where Peak is peak bit rate per frame period T.

For better picture reproduction, we need larger window size of the average policing. This requires larger peak/average values for shorter delay of VBR coding.

9. Simulation guidelines for video coding study

9.1 Introduction

For simulation of video coding for ATM networks many parameters can be used. Some parameter which are relevant to the comparison of the different coding schemes are listed here:

1) Parameters specifying the codec output

- average bitrate
- burstiness or peak bitrate
- variance
- burst duration and frequency
- quality

2) Parameters specifying the network

- link capacity
- cell loss ratio
- queuing technique
- buffer size (delay)
- policing function
- tariffing

3) Parameters important for or resulting from both network and codec

- QOS required
- Priorities required
- Policing function required at the encoder
- Measuring period of all statistical parameters

To be able to compare results in this chapter some guidelines are given for most of these parameters. The following transfer situations may be assumed in an ATM network:

a) One channel with a certain Quality of Service (QOS)

b) Two channels with different QOS, resulting in a layered coding scheme layered coding scheme. This will only be useful if the cell loss rate of the two channels differ significantly.

In the case that one channel is used a simple model can be defined. This model is depicted in Figures 2.1 and 2.2.

Figure 2.1
Figure 2.2

9.2 Reference coding algorithm

For simulations H.261 Encoder Reference Model will be used, as described as a modification to RM8 (according to modified doc COST211ter SIM(90)101).

{A complete reference model for H.261 encoding (based on RM8) will be available in Jan 91: appropriate CCITT reference is required }

In the case of two layered coding the base layer is encoded according to this reference coding algorithm.

For simulations outside the scope of H.261 other coding algorithms can be used.

9.3 Transmitter channel adaptor

The control process adapts the encoder to the ATM channel. It performs packetization of the bitstream at the encoder and controls parameters of the encoder (e.g. stepsize).

9.4 ATM reference channel

9.4.1 Description

The ATM reference channel is a model based on a B-ISDN Class B service (see I.362). In simulations measuring e.g. burstiness and the impact of cell loss the channel parameters "cell length" and "cell los rate" need to be defined.

9.4.2 Cell length

The cell-size for ATM is 48+5 octets. The ATM-payload consists of 48 octets. Each cell has to spend octets for SAR-header and a SAR-trailer.

Figure 2.3

For simulation purposes it is recommended to use:

- 1 octet SAR Header
- 2 octets SAR Trailer
- 45 octets information (SAR-payload)

As the AAL has not yet been fully defined, for simulations the Convergence Sublayer is not taken into account. This results in zero bits for CS-header, CS-pads and CS-trailer; the User PDU length is assumed to have no upper limit.

9.4.3 Cell loss rate

Simulations can include a high and a low cell loss rate.

- $P_{loss}=10^{-3}$ for high cell loss rates.
- $P_{loss}=10^{-8}$ for low cell loss rates.

Note 1 - For simulations the low cell loss rate can be assumed to be: $P_{loss}=0$.

Note 2 - PPD for the Kurihama tests states as follows; "To demonstrate cell loss resilience, it is recommended to simulate 10^{-3} cell loss ratio, e.g. by replacing 384 coded bits every 0.1 sec with '0', decoding the resultant bit stream and reconstructing pictures. Exact method of simulation should be described."

9.5 Receiver channel adaptor

On the receiver side the channel adaptor consists of at least a depacketizer. It can also detect Cell loss and give a "cell loss flag" to the decoder.

9.6 Reference Decoding algorithm

For decoding the H.261 decoder is used. In the case of two layered coding the H.261 is a fall back mode (base channel). For simulations outside the scope of H.261 other decoding algorithms can be used.

9.7 Measurement of statistics

9.7.1 Burstiness

One of the items that seem useful to be measured or limited on a VBR-codec is the burstiness of the bit stream or cell-stream that is produced, but there is no single definition for the burstiness.

For encoder control the maximum number of cells per GOB is defined to limit burstiness.

- measure the cell rate and bit rate for each frame, GOB and MB.
- optional: measure momentary cell rates. The momentary cell rate at cell N is defined as:

  If a cell arrives at time = t_i and the next cell arrives at time = t_{i+1}, the momentary cell rate during the time-interval $<t_i, t_{i+1}]$ is defined as $r_{mom}=1/(t_{i+1}-t_i)$.

  The time base that is used is the average time for a macro block.

9.7.2 Sequences

Three different services can be distinguished:

- Conversational
- Retrieval and messaging
- Distribution

These services all require typical sequences for simulations. Agreed is to use:

- Existing CIF test sequences: SALESMAN, CLAIRE, MISS AMERICA, BLUE JACKET, SWING
- MPEG phase-2 sequences: FLOWER GARDEN, SUSIE, POPPIE, TABLE TENNIS, MOBILE & CALENDAR, TEMPETE
- A long sequence but with average complexity for testing VBR for conversational services

9.7.3 Simulation parameters

Four different bit rates are considered for simulations

mean Bitrate [kbit/s]	Framerate [frames/s]
64	10
320	15[/30]
1024	30
1856	30

For the 25 Hz sequences the frame rates 10,15 and 30 Hz become resp. 8.33, 12.5 and 25 Hz. For 2-layered coding the mean bitrate in the two cannels has to be:

CHANNEL 1: 50 of total (mean) bitrate
CHANNEL 2: 50 of total (mean) bitrate

9.7.4 Network Loading

An important parameter for simulations is the total number of calls. For the loading of networks a number of formulas can be used. See §6 above.

9.7.5 Presentation of results

Depending on the type of simulations that are being performed, the results can be presented with:

- SNR plots
- Plot of Number of bits per frame
- For statistics, see Table 3.1
- Burstiness

1. mean and peak number of cells per macroblock
2. mean and peak number of cells per group of block
3. mean and peak number of cells per frame
4. optional: histograms of
   the momentary cell rates
   the number of cells per GOB

(For items 1,2 and 3 see example Table 3.2)

Coding results have to be presented on U-MATIC TAPE. It is mandatory to show the results full screen (also when a comparison of two sequences is made). In the case of two layered coding the fall back mode (base channel) result has to be shown.

Table 3.1
Table 3.2

{to be discussed and elaborated}

10. Picture quality assessment

10.1 Demonstration of particular topics

For demonstration of hardware processed or computer simulated pictures, the hosting organization is kindly requested to provide the following equipment:

- VCR: multi-standard U-matic, D-1 machine
- monitors: NTSC, PAL, SECAM inputs,
  50 and 60 Hz component inputs (R, G, B)

{monitors to assess progressively-scanned signals, format and hardware for storing progressively-scanned material on D1}

10.2 Formal subjective tests

{Need to have formal subjective tests? When (divergence, convergence)?}

{CCIR method - double stimulus continuous quality-scale method, Rec. 500-3, "Kurihama tests"}

11. Multimedia multiplexing in B-ISDN

11.1 Multiplexing methods

The following methods for multimedia multiplexing are conceivable in the B-ISDN environment (Figure 3);

1) Cell multiplex : each medium is identified by the cell header.
2) Message multiplex: each medium is identified by the IT of SAR.

3) Media multiplex : each medium is identified by the CS header.
4) User multiplex : multimedia signals are multiplexed in the layer above AAL

Figure 3

Cell multiplex has several merits as follows;

- It fits to broadcasting because the selection of media by the receiver is easy
- Usage parameter control for each medium is easy.
- Multiplexing delay is considered to be short.

Therefore we adopt cell multiplex as a reference method for our future work. Further studies are required for the following items;

- How to cope with differential delay between several VCs. To set several VCs in the same VP may be one solution, but desirably network provision of appropriate signalling and control for bounded cross media delay may be a better solution.)
- Whether tariff penalty exists or not to use multiple VCs.
- How to connect with the audiovisual terminals which adopt user multiplex, for example N-ISDN audiovisual terminals (H.320) or MPEG terminals.

The first two items have been sent to SG XVIII as questions.

(It may happen that several multiplex methods are used. For example,

Video and audio : message multiplex to avoid the differential delay.
Video and data : cell multiplex to use different AALs.
Low bit rate data : user multiplex to reduce the cell assemble delay.)

11.2 A method for further study

The following steps are introduced;

1) Clarification of requirements to ATM networks from audiovisual terminal point of view. For example; what kind of clock recovery is required?

2) Concrete multimedia multiplex method to realize the requirements.

Furthermore the idea to assume an intermediate virtual layer (called mux layer, perhaps more appropriately media control layer or media convergence layer) is introduced in AVC-53,92 to clarify the requirements for ATM and AAL layers. Further development of the idea is requested.

12. AAL

12.1 Approaches for AAL suitable for AV systems

The following solutions can be considered for AAL;

1) AAL has minimum functions which are required from all types of audiovisual terminals.
2) AAL has maximum functions which are sufficient for all types of

audiovisual terminals.
3) Middle of 1) and 2).

The Experts Group support the considerations in AVC-77 that AAL Type 2 be commonly used for VBR and CBR audiovisual communications, and recognized the urgency of study on the definition of AAL Type 2 by the Experts Group.

12.2 Bit error and cell loss protections

Since the preventive measures against bit errors and cell losses have crucial impacts to the structure of audiovisual terminals for B-ISDN, the meeting agreed to expand the content in relevant parts of the IVS Baseline Document (p.20 and p.22 of AVC-68) covering the above concerns by notes as a step in the ongoing iterative process.

There was discussion on where necessary error correction and cell loss protection be carried out, at SAR, CS or user layer depending on the network performance. The following factors should be considered in this study:

- Interworking with N-ISDN terminals should also be considered where bit errors are taken care of at the user layer.
- It should be noted that scrambling in ATM layer causes correlated errors in SAR-PDU.
- Order of scrambling, error correction and encryption should be carefully considered.
- If video is layered, different layers may take advantage of different levels of bit error and cell loss performance.

12.3 Interworking between N-ISDN and B-ISDN terminals

There is a question whether audiovisual terminals use Type 1 as well if both VBR and CBR are covered by Type 2 as in §12.1 above. One of the cases identified is interworking between B-ISDN and N-ISDN. We feel ambiguous about whether Type 1 for interworking be required in H.32X terminal, or it may be covered by the interworking unit in the network. It is also pointed out that the existing N-ISDN terminals are quite susceptible to cell loss if it occurs in the circuit emulation mode.

13. Work plan and work method

There is consensus on the following work plan of the group;

- Final Recommendation be made official in 1994, taking into account the completion of the B-ISDN Recommendations in 1992 and subsequent service provision.
- Outline Recommendation be produced at the end of the current study period, which includes scope, list of contents, such parameters as picture formats, framework of coding scheme, etc. to be agreed by that time.

The following methods practiced in the previous Specialists Group for H.261 are supported for the H.26X work;

- Study is phased as "divergence" and "convergence",
- Step by step using Reference Models, and
- Hardware verification at the final stage.

As to the reference model, it was clarified that this time we need two kinds of model; one for network aspects study and the other for video coding aspects study. It was also clarified that the latter includes source coding as well as channel coding.

The group considered when the first Reference Model for video coding should be defined, and over what period various types of candidate algorithm should be tried, concluding that both should coincide with the demarcation between the "divergence phase" and the "convergence phase".

As a summary for the work plan and method, an agreed time table is shown in Annex 6.

14. Harmonization with other standardization bodies

This group seeks to carry out joint work with TG CMTT/2 Special Rapporteur's Group and ISO/IEC JTC1/SC2/WG11 (MPEG) in order to avoid different standards in the same or similar areas and to avoid duplication of standardization work as well.

To this end the CCITT group proposed to MPEG and CMTT/2 SRG that joint meeting sessions be arranged in the areas of overlapping interest and responsibility, namely:

- source video coding algorithm and video multiplexing,
- system issues concerning multimedia multiplexing and synchronization,
- implementation considerations.

The MPEG Berlin meeting of December 4-7, 1990 concluded in response to our liaison statement that the ongoing phase of work on audiovisual coding at bit rates up to about 10 Mbit/s be carried out in collaboration with CCITT, by holding joint meetings on matters of common interest such as video, systems, implementation. It was confirmed there that both groups have a common target date of freezing draft specifications as end of 1992. It was also confirmed that a "Test Model" would be defined after subjective tests of candidate algorithms for further collaborative elaboration.

The TG CMTT/2 Tokyo meeting of March 25-28, 1991 concluded that the aspects of practical collaboration with Study Group XV are addressed through Special Rapporteur and that a delegation of TG CMTT/2 is expected to attend common meetings of Study Group XV and ISO/MPEG.

15. Outstanding questions

15.1 Picture format

For the simulation purpose, we will initially deal with QCIF/CIF and CCIR-601 formats. We recognize the idea of a single coding format for CCIR-601 level pictures, but the details await further study. We should clarify advantages and disadvantages of this idea for making a decision.

1) Single mode using Super-CIF (SCIF) or multiple modes including 625/50 and 525/60?

2) If we adopt the single mode approach, what parameters are appropriate for SCIF? How should SCIF be related to CIF, interlaced or non-interlaced? How should it be related to larger formats such as EDTV and HDTV?

3) How are picture quality, coding efficiency and coding/decoding delay are affected by the use of SCIF compared to the case CCIR 601 signals are directly processed?

4) What trends can we expect for progressive scan cameras and displays?

5) Should we focus mainly on conversational services as a first step, and take initiative to adopt the single picture format approach to other groups as a second step if it is proved viable with evidence?

6) What framework is the best for the service integration on B-ISDN?

7) Shall we take the initiative to adopt the single format approach (if viable) to other groups?

15.2 Video coding

15.2.1 Coding architecture

1) What coding architecture will be appropriate for realizing a universal coding algorithm in terms of service, quality, resolution, application and bit rate each of which is as a range?

2) What specific requirements should be taken into account for the high quality video coding?

3) How are "Flexible Layering" aspects reflected?

4) Is source/video multiplex coding is not separable with transmission coding in B-ISDN environments?

5) Is simple recovery from rare cell loss sufficient or is sophisticated protection against frequent cell loss required?

15.2.2 Variable bit rate coding vs constant bit rate coding

Since VBR (variable bit rate) for constant quality is considered as one of the outstanding features of ATM, we recognize it as an urgent study item to clarify the advantages of VBR video coding against CBR (constant bit rate) video coding. It is also a common understanding that applications should be clarified where VBR is most effective.

For facilitating this study, a network loading model has been defined (see Annex 5 to this document). Comparison in video coding efficiency (e.g. one layer coding vs two layer coding) based on this model as well as proposals for improvement of this model are encouraged. Elaboration of the model may include cluster type of cell loss pattern, one more model appropriate for another QOS and/or higher bit rates. It is noted that the measuring window will be critical when discussing VBR.

1) What does the customer perceive as "constant picture quality"?

2) Does a realistic ATM network model exists so that the performance of various coding schemes can be assessed? How can the network loading model be improved?

- comparison of model with measured data (simulation, HW experiments)
- use of single source model
  * generate synthetic bit rate files
  * use for statistical multiplexing experiments
- refinement of the model
- ...

3) Does VBR offer significant improvements (picture quality, short delay, etc.) over CBR?

4) How coding and decoding delay is estimated for VBR coding?

5) How do codecs specify to the network their call requirements (total mean bit rate, peak bit rate, etc.)?

6) What coding scheme offers the best network loading characteristics?

7) What coding control scheme for VBR?

- averaging windows
  * size
  * jumping vs sliding (or triggered jumping)
  * network & coder desires (requirements, where to meet?)
- network constraints, UPC, ...
- control strategies within the encoder: credit building, ...
- independent control of high and low priority channels

8) Cell loss protection and recovery methods?

9) What clock recovery is required for VBR?

- does VBR requires dedicated clock recovery schemes between encoder and decoder? What is the impact on complexity and performance?

- do layered coding schemes require special clock recovery schemes? Impact on complexity and performance?

10) Can any VBR coder act as a CBR coder?

- relation with buffer length
- symmetry?
- CBR as a special case of VBR:
  * go directly for VBR, CBR as fall back solution?
- ...

15.2.3 Layered coding

The intention of layered coding is either to obtain compatibility among different service classes or to cope with cell loss (equivalently to make use of statistical multiplexing).

1) How can the notion of "layering" be introduced?

2) With what evolution scenario?

3) How can the two priority channels be utilized for efficient video coding?

15.2.4 Compatibilities with other standards

1) To what extent the compatibility with existing AV Recommendations are required?

- Mandatory between terminals
- Forward/backward compatibility
- How about video coding laws? Mandatory / desirable / don't care?

```
 -------   N-ISDN   ---------------   B-ISDN   -------
| H.320 |   UNI    |               |   UNI    | H.32X |
| Termi |----||----|     ISDN      |----||----| Termi |
|    nal|          |               |          |    nal|
 -------            ---------------            -------
```

Audio coding	G.722 etc.	G.72X (G.722 etc.)
Video coding	H.261	H.26X (H.261)
Multimedia mux	H.221	H.22X (H.221)
Communication procedures	H.242	H.24X (H.242)

2) How should H.26X be related with MPEG standards, particularly MPEG-2 one?

3) How should it be related to CMTT standards?

Note - See the first illustration in Annex 4 of this document.

15.2.5 Algorithms for higher quality video coding

1) What algorithm is appropriate for coding video signals with television broadcast or CCIR 601 quality at several Mbit/s to several tens Mbit/s?

2) How is it related to the H.261 algorithm?

3) How can it be extended to such higher resolution pictures as enhanced television or high definition television ones?

4) What new coding elements are promising?

5) How can the impact of cell loss on video coding be simulated?

15.2.4 Multipoint

1) How can B-ISDN properties be utilized for efficient multipoint communications?

2) How can mix of H.320 and H.32X terminals be treated?

15.3 ATM network characteristics

Many of the questions listed here are to be answered by the people who are responsible for the network design. We are in a position to pose better questions to draw out necessary information from them.

15.3.1 Tariffs

1) Is ATM transmission cheaper per information bit than STM?

2) Will VBR bits be cheaper than CBR bits?

3) Is the network's loading efficiency a major contributory factor in determining customer tariffs for the various supported services?

4) How are VBR services going to be charged (by mean bit rate, negotiated bit rate, QOS, etc.)?

15.3.2 Implementation

1) Cell loss ratio: what will be the expected values, for both priority levels? Will VBR and CBR services be subject to the same CLR (in both low and high priority classes)?

2) What is the cell loss burst behavior?

3) How is the CLP bit used?

- Will there be separate negotiations for the two priority levels?
- When will the service provider set this bit?
- What are the restrictions for the use of this bit?
- Is the quality of service selectable?
- The Experts Group wishes to clarify whether the CLP bit could be changed by the network after it has been set to "high priority" by the user. If this change could occur, it would be of serious concern to the Experts Group.

4) Usage parameters:

- What parameters will be used for policing and admission control?
- What policing mechanism will be used?
- What averaging intervals can be used to measure mean, peak, etc.?

5) Multimedia connections:

- How will the admission and monitoring of an ensemble of VC's be handled?
- What is the limit on differential delays between different VC's?
- How should different media signals be synchronized?
- Is there any technique to limit the difference of delay for multiple VCs?
- Will the network offer a mechanism whereby a request that multiple VCs should be supported over the same transmission path can be satisfied?
- Is there any provision to guarantee bounded cross media delay?

6) Bit Error Rates:

- What is the expected rate?
- What is the impact on the AAL?
- Do we need error correction in SAR-PDU?

7) Cell delay jitter: what values are expected?

8) Can ATM networks support differing services with differing QOS or do they have to operate at best QOS for the range of services supported?

9) Will Class 0 be supported by all networks?

10) What is the difference in delivery times at the receiver between 2 channels (i.e. 2 VCIs - this is applicable to 2-layer coding)?

11) Is the proposed sequence number comprising 4 bits of sufficient size to cope with a burst of lost cells?

12) Will the function of AAL Type 2 be determined only from the standpoint of video coding? The required functionality may not necessarily be uniform across the range of services, applications and coding methods.

13) What provision is made for the interworking between N-ISDN and B-ISDN audiovisual communication terminals?

15.4 Multimedia multiplexing

1) Which of the following methods is suitable for multimedia multiplexing in the B-ISDN environment?

- Cell multiplex: each medium is identified by the cell header
- Message multiplex: each medium is identified by the IT of SAR
- Media multiplex: each medium is identified by the CS header
- User multiplex: multimedia signals are multiplexed in the layer above AAL

2) Can the followings be confirmed for the cell multiplex method?

- How to assure cross media synchronization
- Penalty in the use of network resource (Note: Tariff is related to this consideration, which is recognized difficult to handle.)
- Maximum number of VCIs for a multimedia connection
- Existence of user level multiplex to a single bit stream such as N-ISDN audiovisual systems and MPEG systems
- Impact of VBR coding and UPC

15.5 AAL

1) What AAL is most appropriate for audiovisual systems?

2) Detailed specification for AAL Type 2

3) What AAL(s) are used for interworking between N-ISDN and B-ISDN audiovisual terminals?

15.6 Field trials

1) What will be maturity of the B-ISDN network by mid '93?

2) With what system/network can the hardware interconnection tests be carried out?

END

EDITORS FOR "STATUS REPORT"

The editor is in charge of the following tasks;

- to collect materials of common understanding,
- to list items requiring further study indicating different views if any,
- to add any editor's comments in { } to encourage further work.

Title : Status Report on ATM Video Coding Standardization

	Editor
1. Introduction	S. Okubo
2. Terminology	R. Schaphorst
3. Applications	A. Tabatabai
4. Boundary conditions for ATM video coding	M. Wada
5. Picture format	G. Bjoentegaard
6. Network model	D.G Morrison
7. Video coding model	M. Biggar
8. VBR vs CBR	W. Verbiest
9. Simulation guidelines for video coding study	D. Schinkel
10. Picture quality assessment	D. Lemay
11. Multimedia multiplexing in B-ISDN	T. Tanaka
12. AAL	T. Tanaka
13. Work plan and work method	S. Okubo
13. Harmonization with other standardization bodies	S. Okubo
14. Outstanding questions	S. Okubo

Table 1 Matrix of applications and technical issues

Tech. issues / Applications	Network	Storage media	Resolution	Quality objective	Delay
Video conference	N,B-ISDN LAN	-----	CIF ~ CCIR601	~ 3.5	Short
Video conference with wide screen	B-ISDN LAN	-----	EDTV ~ HDTV	~ 3.5	Short
Video conference with multi-screen	B-ISDN LAN	-----	CIF ~ CCIR601	~ 3.5	Short
Videophone	N,B-ISDN LAN	-----	CIF ~ CCIR601	~ 3.5	Short
Video surveillance	N,B-ISDN	----	QCIF/CIF ~CCIR601	3~4.5	Mid
TV broardcasting	DBS CATV-net	-----	Current TV ~ HDTV	~ 4.5	Mid.
Video distribution on storage media	B-ISDN LAN	Disk Tape	Current TV ~ HDTV	~ 4.5	Mid.
Video database	N,B-ISDN LAN	Disk Tape	CIF ~ HDTV	~ 3.5	Long
Videotex	N,B-ISDN LAN	Disk	CIF ~ HDTV	~ 4.5	Long
Video mail	N,B-ISDN LAN	Disk	CIF ~ CCIR601	~ 3.5	Long
Video instruction	B-ISDN LAN	Disk	Current TV ~ CCIR601	~ 4.5	Mid.

note-1: The Quality grade is refered to the follwoing:

Quality	Impairment
5 Excellent	5 Imperceptible
4 Good	4 Perceptible, but not annoying
3 Fair	3 Slightly annoying
2 Poor	2 Annoying
1 Bad	1 Very annoying

note-2: The "Long Delay" means even non-realtime transmission is acceptable.

Table 2 ATM network characteristics

===

Opportunities

1) Availability of high bandwidths
2) Flexibility in bandwidth usage
3) Varaible bit rate capability
4) Flexibility in multimedia multiplexing or
   multiple connections
5) Service integration

Limitations

1) Cell loss
2) Cell delay jitter
3) Packetization delay
4) Usage Parameter Control (Peak and/or Average)

===

Sequence :
Modification :
Number of tracks :

Institute :
Date :
Temporal resolution :

Item		layer 1	[layer 2]
SNR for luminance SNR for chrominance(U) SNR for chrominance(V)			
Mean value of step size			
Mean value of the number of non-zero coefficients			
Mean value of the number of zero-coefficients			
Block	Fixed		–
type	Inter+coef		–
of	Inter+Coef+MC+Fil		–
Macro	Inter+MC+Fil		–
	Intra		–
	Inter+Coef+MC+Fil+Q		–
	Inter+MC+Fil+Q		–
	Intra+Q		–
	Inter+Coef+MC		–
	Inter+MC		–
	Inter+Coef+MC+Q		–
mean number of bits/frame			
Bits for first frame			
Number of forced to fixed mb's			–

Table 3.1: The form for the assesment of the coding results

		Number of Cells		
		max	mean	max/mean
Dawn 25Hz Fixed Quantizer 8	mb	4	0.28	14.53
	gob	31	9.09	3.41
	frame	285	109.05	2.61
Dawn 12.5Hz Fixed Quantizer 12	mb	3	0.18	16.94
	gob	24	5.85	4.11
	frame	213	70.15	3.04

Table 3.2: Mean and peak cell rate example

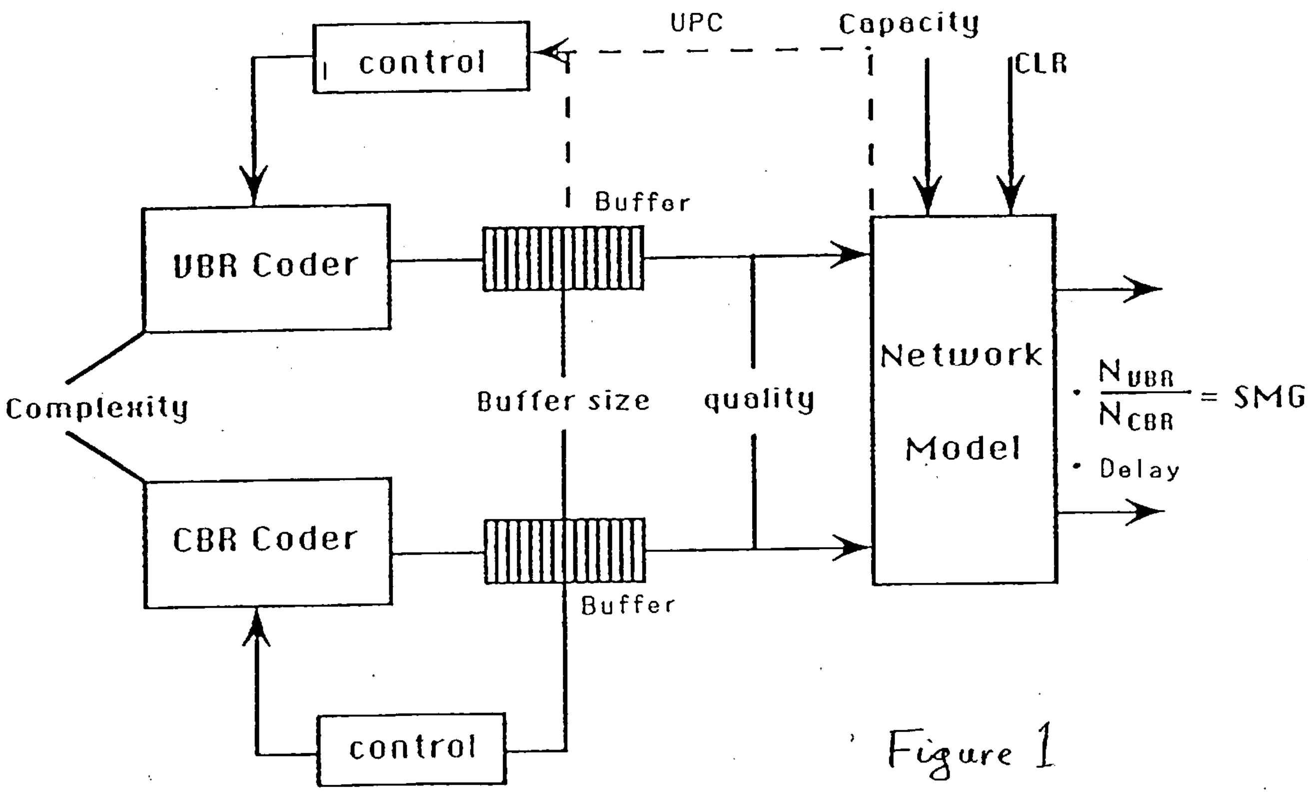


Figure 1

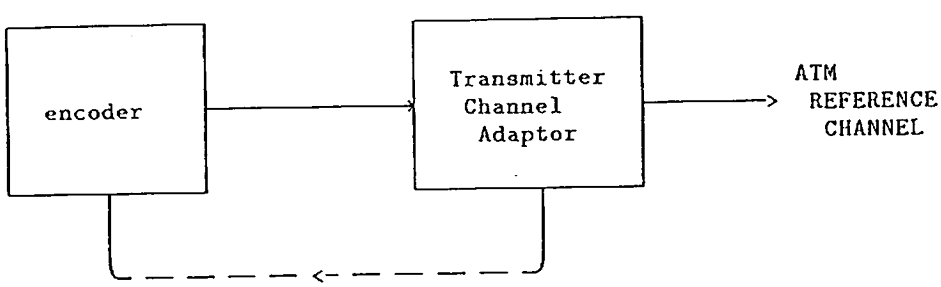


Figure 2.1: ATM Reference Encoder

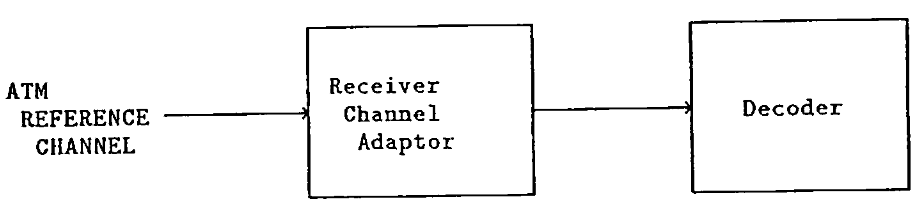


Figure 2.2: ATM Reference Decoder

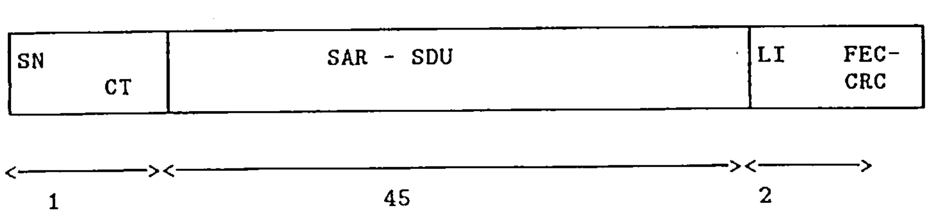


Figure 2.3: ATM cell information field

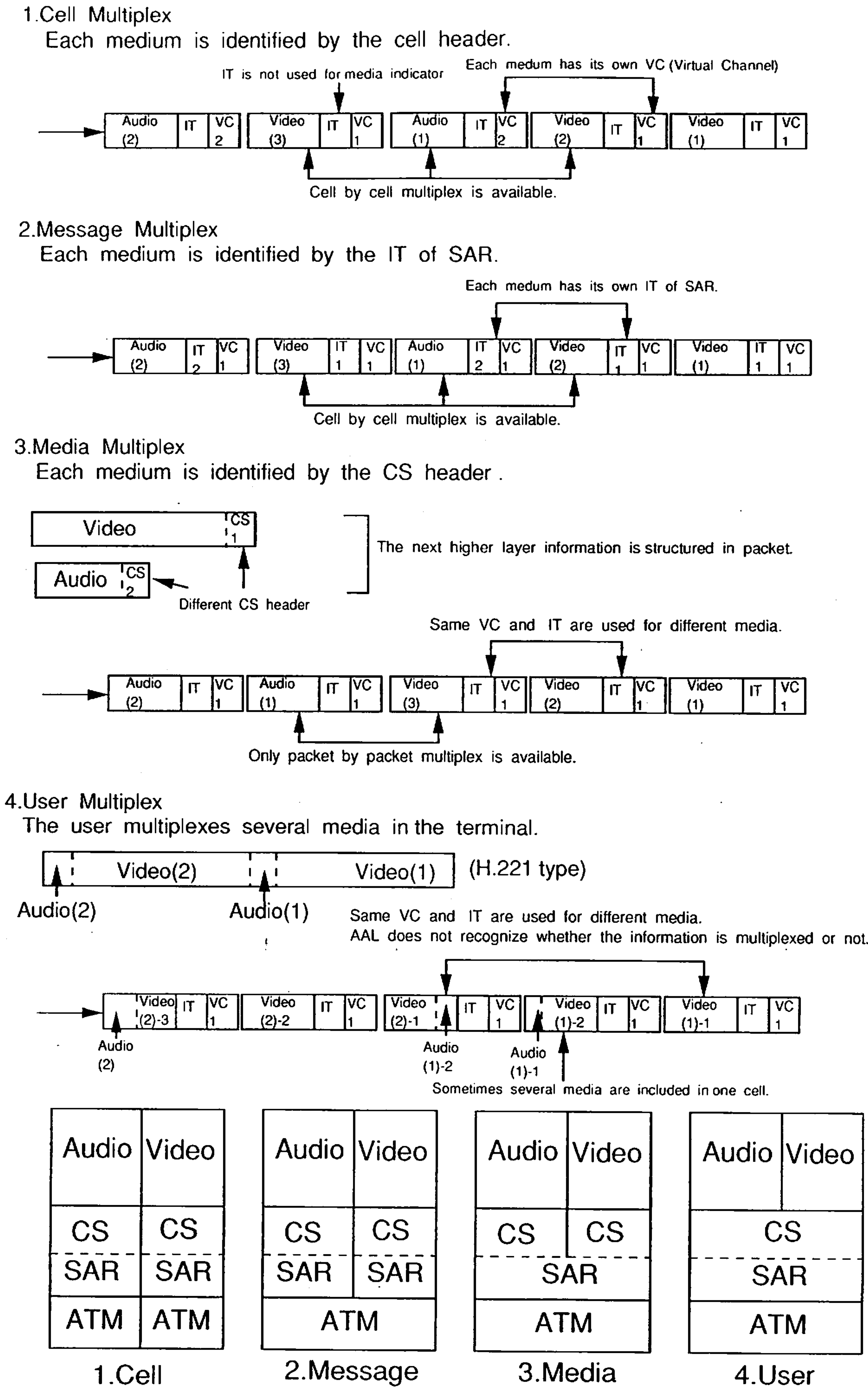


Fig.3 Multimedia multiplex methods

Terminology

GENERAL TERMS

CODEC: Acronym for COder/DECoder. An electronic device that converts analog signals, typically video, voice, and/or data, into digital form and compresses them into a fraction of their original size to save frequency bandwidth on a transmission path or storage media.

compression: The application of any of several techniques that reduce the number of bits required to represent information in data transmission or storage, therefore conserving bandwidth and/or memory.

CIF: Common Intermediate Format. A video format defined in H.261 that is characterized by 360 pels on each of 288 lines, with half as many chrominance pels in each direction, and 29.97 pictures per second.

FEC: Forward Error Correction. A system of error control for data transmission wherein the receiving device has the capability to detect and correct any characters or code block that contains fewer than a predetermined number of symbols in error.

interoperability: The condition achieved among communication-electronics systems or items of communication-electronics equipment when information or services can be exchanged directly and satisfactorily between them and/or to their users. The degree of interoperability should be defined when referring to specific cases.

lip synchronization (lip-sync): The relative timing of audio and video signals so that there is no noticeable lag or lead between audio and video.

motion compensation coding: A type of interframe coding used by picture processors in the compression of video images. The process relies upon an algorithm that examines a sequence of frames to develop a prediction as to the motion that will occur in subsequent frames.

multipoint: A telecommunication system which allows each of three or more sites to both transmit signals to and receive signals from all other sites.

QCIF: Quarter CIF. A video format defined in H.261 that has 1/4 as many pels as CIF per picture.

quantization: A process in which the continuous range of values of a signal is divided into nonoverlapping subranges, a discrete value being uniquely assigned to each subrange.

teleconferencing: Generally, the transmission of audio and/or video communications of a conference such that two or more locations are connected and can function in the live exchange of information.

videoconferencing: Two-way electronic form of communications that permits two or more people in different locations to engage in face-to-face audio and visual communication. Meetings, seminars, and conferences are conducted as if all of the participants are in the same room.

videophone: A symmetrical, bidirectional, real-time, audiovisual communication generally involving two persons in case of point-to-point connection.

visual telephone: A group of audiovisual communications covering videoconferencing and videophone.

COMPATIBILITY RELATED TERMS

Upward and downward compatibility

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

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

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

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

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

Forward and backward compatibility

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

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

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

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

Simulcasting

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

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

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

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

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

Embedded bit stream

In this case the encoder is characterized as follows:

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

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

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

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

Syntactic extension

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

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

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

Switchable encoder

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

The encoder is characterized as follows:

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

Standard families

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

Layered Coding

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

"Flexible layering" was introduced in Doc. AVC-35 (May 1991) and provides for layers which could represent baseband or incremental information as appropriate in a given application. In the terminology above, decoding of data stream M is possible by making reference to data streams B...M, where B is the baseband picture signal and 1 <= B <= M. This system includes single layer coding (B=M) as a special case.

END

Questions/responses to/from Study Group XVIII

Introduction

The Experts Group for ATM Video Coding was established by SGXV to investigate new possibilities for video coding offered by service support on the B-ISDN, and to develop appropriate coding algorithms. There is a significant impact on video coding as a result of ATM transport, and on multimedia system design as a result of virtual connections. The impact of certain network parameters on approaches to video coding and video service provision are outlined below. The intention is to both obtain guidance from SGXVIII regarding suitable parameter values to allow video coding work to progress, and to provide SGXVIII with some input that may influence aspects of network design. The first section briefly itemizes the main issues and the addenda will give a more comprehensive explanation.

Answers from SGXVIII are listed after the corresponding questions.

Issues

Q.1 <November 1990> Cell loss ratio:

- What will be the expected values, for both priority levels?

Q.1a <May 1991> CLR for services with differing rate behavior: Will VBR and CBR services be subject to the same CLR (in both low and high priority classes)?

Q.1b <August 1991> Cell loss ratio

If cell loss ratio is rather low, we believe that the channel coding alone, either at AAL or higher layer, can cope with low cell loss ratio based on acceptable delay requirement. However if the cell loss ratio is extremely high, some technique for cell loss resilience is required also for video source coding.

On the other hand, we intend to select a video source coding algorithm at the beginning of 1992. Therefore after that day, it becomes difficult to implement additional cell loss resilience technique for source coding algorithm.

Therefore we are eager to know the likely value of cell loss ratio as soon as possible. For this purpose we sent the liaison two times. However, the answer from SGXVIII is not clear yet. IVS Baseline Document only says that the requirement from SGXV is that value and does not mention whether this requirement is achievable or not from a network standpoint of view.

Q.2 <November 1990> What is the cell loss burst behavior?

<u>A.1-2 <WPXVIII/6, November 1990></u>

The Q5 Rapporteur Group decided to inform SGXV Experts Group for ATM Video coding of the available performance evaluation results on cell loss in ATm network. The two simulation results show the following:

(i) Cells tend to be lost consecutively.
(ii) The cell loss process is not random process and may be described by the Gilbert Model.

The further information on ATM network performance will be forwarded when it is available.

Reference:

T. Yokoi, N. Kishimoto and H. Fujii: "ATM Network Performance Evaluation using Parallel and Distributed Simulation techniques," to appear in ITC 13, June 1991.

H. Ohta and T. Kitami: "Simulation Study of Cell Discard Process and the Effect of Cell Loss Compensation in ATM Networks," Trans. IEICE, Vol. E73, No. 10, October 1990.

A.1-2 <WPXVIII/8, November 1990>

It is likely that the cell loss will be dominated by network congestion; high priority cells are expected to be discarded with a very low probability (to be defined), except under severe network congestion. Furthermore, bursts of cell losses may happen and the length of the burst is likely to be service rate dependent.

Q.3 <November 1990> How is the CLP bit used?

- Will there be separate negotiations for the two priority levels?
- When will the service provider set this bit?
- What are the restrictions for the use of this bit?
- Is the quality of service selectable?

A.1-3 <WPXVIII/8, November 1990>

First results are relevant to traffic characterization parameters. There is a general consensus to focus the initial activities on a limited set of parameters including peak rate. The future activity will be carried out in such a way not to preclude future refinements and compatibility with already achieved results.

If both low priority and high priority cells are carried within one Virtual Channel (e.g. for layered video coding), CCITT SGXVIII is currently discussing whether both high priority (CLP bit set to zero) and total traffic requires separate characterization. Consequently the expected cell loss rate has not yet been determined. Inputs are required to decide whether the lower priority should provide for assured limits to cell loss rate or not.

Q.3a <May 1991> CLP bit: The Experts Group wishes to clarify whether the CLP bit could be changed by the network after it has been set to "high priority" by the user. If this change could occur, it would be of serious concern to the Experts Group.

A.3 <WPXVIII/8, June 1991> Status of discussions on CLP issues

The followings are questions and answers which have been identified and obtained from services viewpoints by SWP 8-3.

q.1) There are two ways of using CLP capabilities of B-ISDN, i.e. cell by cell basis and per connection basis. In the latter case, will a CLP be defined per VPC basis as well as per VCC basis?

a.1) No constraint from services viewpoints. This should be studied mainly by ATM and Resource Management aspects.

q.2) In the case of CLP bit capability, will a a single level of CLR (Cell Loss Ratio), i.e. high priority level, be defined, or will two levels of CLR be defined respectively?

a.2) CLR for CLP=0 (high priority) should be defined, and assured by the network if a cell traffic does not exceed the negotiated values. For CLP=1 traffic, the following two options exist;

1) defined and assured by the network,
2) not defined and no assurance provided.

In the case of option 2), a given CLR may be maintained by the network engineering.

q.3) Will the CLP bit, set by a user as high priority, be changed by the network, e.g. for violation tagging in case that a traffic exceeds the negotiated values?

a.3) Since CLR of CLP=0 is assured by the network, there may be no impacts on services and for users, where the network will override the bit or not.

q.4) Will the specific value of CLR be explicitly declared by a user, or will the CLR indication be implicitly associated with specific service requests, e.g.

- a standardized service will by definition include the specification of all relevant QOS values, or
- a standardized QOS class will by definition include the specification of all relevant QOS values?

a.4) Under study

q.5) Will the network provide for several CLR, or will the network accommodate CLR requests from users with a very limited number of CLR? And how many CLR will be required?

a.5) Under study.

Q.3b <August 1991> CLP bit

1) Negotiation for two priority flows

The liaison statement from SGXVIII at Geneva meeting, 11-28 June 1991 (annex 2, last paragraph) said as follows;

"There will be negotiation for both priority flows."

It is clear for CBR. However the following two types of negotiation can be considered (Figure 1);

Case 1: Negotiations are done for both flow separately.

In this case low priority flow cannot use the erosion area of high priority flow.

Case 2 : Negotiations are done for high priority flow and sum of both priority flows.

In this case low priority flow is not restricted as Case 1 and can use the erosion area of high priority flow.

The question is which is the likely solution.

2) Merit of using CLP bit

The question is what is the merit of using CLP bit. Layered coding is a suitable technique to use both priority classes. However whether we adopt this technique or not depends on its expected merit. What degree of network resource saving can be obtained by using low priority cells?

This question is also related to the selection of video source coding technique. Therefore quick response is required.

Q.4 <November 1990> Usage parameters:

- What parameters will be used for policing and admission control?
- What policing mechanism will be used?
- What averaging intervals can be used to measure mean, peak, etc.?

A.4 <WPXVIII/8, November 1990>

As for the Usage Parameter Control function, SGXVIII foresees that the monitoring of the peak cell rate is mandatory for any kind of service. A Usage Parameter Control mechanism and the actions to be taken in case of violation are presently under study.

A.4 <WPXVIII/8, June 1991>

SG XVIII-8 agreed on the definition of one traffic parameter: the Peak Cell Rate. The exact definition is as follows:

Location

At the ATM layer SAP for a VCC.

Basic event

Request to send an ATM_SDU (48-octet information field).

Definition

The peak cell rate Rp is the inverse of the minimum inter-arrival time To of the basic event above.

Usage Parameter Control will be based on Peak Cell Rate as defined above. At this moment, it is not expected that a specific UPC mechanism be standardized. A maximum Cell Delay Variation will be allocated to the Customer Equipment (CEQ) between the ATM SAP and the interface at the T Reference Point.

UPC enforcing actions may be dropping cells or optionally tagging high priority cells by changing the Cell Loss Priority bit from 0 to 1. This is agreed because calls subject to tagging are in excess to the negotiated traffic contract. Cells complying with this contract are neither subject to discard nor to tagging. There will be negotiation for both priority flows. This is because SG XVIII-8 came to the agreement that both high priority and low priority flows will be subject to UPC and enforcement. It appears therefore possible, but left to network operators' decision, to provide for a given QOS on each flow.

Q.4a <August 1991> Usage Parameter Control for peak cell rate

The liaison statement from SGXVIII at Geneva meeting, 11-28 June 1991 annex 2 said as follows;

"A maximum Cell Delay Variation will be allocated to the Customer Equipment (CEQ) between the ATM SAP and the interface at the T Reference Point.

Please show us the meaning of "a maximum cell delay variation". When we assume a system configuration as in Figure 2, the minimum inter-arrival time will change because of multiplex of multiple VCs at terminal adapter and NT2 as shown in Fig.3. Therefore, we cannot control peak cell rate at the T reference point. We have two questions;

Question 1: What is the definition for peak cell rate at the S reference point?
Question 2: What technique does SGXVIII recommend for multiplexing in the adapter to keep the peak cell rate at the S reference point?

Q.5 <November 1990> Multimedia connections:

- How will the admission and monitoring of an ensemble of VC's be handled?
- What is the limit on differential delays between different VC's?

A.5 <WPXVIII/8, November 1990>

In CCITT SGXVIII, the issue on how to provide multimedia services with different virtual channels is still under discussion. At this point in time, the assessment of the limit on differential delay between different virtual channels requires further study; in any case, this assessment has to be based on a reference path through the network.

A virtual path connection may be used to limit the differential delay between virtual channels.

Q.5a <May 1991> Differential delay between virtual channels: The Experts Group recognizes the advantages offered by the multiplexing of multimedia connections on the basis of virtual channels (VCs). However, it will be necessary to limit any differential delay between VCs. Is there any technique to limit the difference of delay for multiple VCs? Will the network offer a mechanism whereby a request that multiple VCs should be supported over the same transmission path can be satisfied?

Q.5b <August 1991> Signaling for multimedia synchronization

One obvious consequence of mixing media on a single virtual path is that the Quality of Service (QOS) required must correspond to that of the most sensitive service, and for many applications this may not be cost effective solution (i.e. mixing loss sensitive data and delay sensitive video traffic) and multiple virtual paths may be required. Multiplexing all services onto a single VP results in zero cross media delay at the cost of a potential mismatch of QOS. Separate VPs ensure a QOS matched exactly to the media.

We are considering to use one VC for each medium for each service, in other words, each service component (multiplex in ATM layer). For multimedia communications, a cross media maximum differential delay should be guaranteed. The typical case is the lip-sync between video and audio. For this requirement it seems that multimedia call should be marked/indicated. Signalling for call establishment and for the addition and deletion of media components must therefore be capable of indicating that particular services are associated for the purpose of synchronization. Further study is requested for signaling and control to handle the cross media delay to be minimum.

Q.5c <August 1991> Technique to support low bitrate information

We intend to use multiplex in ATM layer as mentioned above. However there exists very low bitrate information such as pointing or telewriting. What technique does SGXVIII recommend to transmit that kind of low bitrate information (e.g. 300 bit/s) in multimedia connections?

Q.6 <November 1990> Bit Error Rates:

- What is the expected rate?
- What is the impact on the AAL?

Q.7 <November 1990> Cell delay jitter:

- What values are expected?

Q.8 <November 1990> Network model for hardware experiments:

- What will be maturity of the B-ISDN network by mid '93?

Q.9 <November 1990> AAL

The Experts Group for ATM Video Coding is willing to collaborate with SGXVIII in defining the AAL for video services, and will forward appropriate input to SGXVIII.

Q.9a <May 1991> AAL Type 2: The Experts Group is considering the functionality that it may require from the AAL for video services support. Will the function of AAL Type 2 be determined only from the standpoint of video coding? The required functionality may not necessarily be uniform across the range of services, applications and coding methods.

Q.9b <August 1991> Requirements for AAL Type 1

We are concerned about the circuit emulation mode for existing standard audiovisual terminals. Some tentative requirements have been identified and are listed below. Further requirements may be identified as the work of the Experts Group progress.

1) Interleaving

CMTT suggested to SG XVIII that the CS layer should be capable of interleaving. Considering that the delay produced by the interleaving processing depends on transmission rate, the interleaving function should be optional, not mandatory.

2) Cell loss notification

Since not all erroneous information can be corrected by AAL, cell loss notification is indispensable for the decoder to lessen the damage to the reconstructed picture.

3) 8 kHz timing

When conventional terminals are used, 8 kHz timing is necessary to synchronize the first bit of each octet between the sender and the receiver. CMTT also requires 8 kHz timing on behalf of conventional codecs, so SG XV also requires 8 kHz timing for the same reason.

Q.9c <August 1991> AAL Type 2 SAR

The major user of AAL type 2 will be video services. Therefore we agree that we have responsibility for providing the major inputs to SG XVIII leading to the definition of AAL type 2.

Given the necessity to support a wide range of video services with different requirements, rates, etc., it appears that the fields given as examples in I.363 AAL Type 2 example may be restrictive. For example, an AAL Type 2 with minimum functionality may offer the flexibility to accommodate a wide and diverse range of video services, but the minimum functionality is now under discussion.

<u>Commonality of the SAR-Sublayer for CBR & VBR video</u>

It is the common view held by the Experts Group that issues in relation to SAR functionality to support VBR video using AAL Type 2 apply equally to the support of CBR video. Therefore, in line with the Experts Group's desire to contribute to the development of requirements of the AAL Type 2 for the support of VBR video, AAL Type 2 will be developed within the Experts Group to support both VBR and CBR video services.

Q.10 <August 1991> Network interworking

1) When network provides interworking function between N-ISDN and B-ISDN by network gateway as shown in Fig. 4, what types of AAL are required for B-ISDN side terminal?

2) How does B-ISDN intend to provide transparent N-ISDN circuit emulation especially those for cell loss sensitive and time delay critical services such as visual telephone using H.261, H.221 etc.?

3) What slip rate is expected when providing N-ISDN circuit emulation between N-ISDNs via B-ISDN?

4) The Experts Group is considering the conversion of N-ISDN user multiplexed signals to B-ISDN VC multiplexed signals either in Terminal Adapters (TA) or B-ISDN/N-ISDN Inter-Working Units (IWU). Current user

multiplex structures (e.g. H.221) can reconfigure their internal rate allocation in the order of 20msec. The Experts Group requires information on the possibility of;

a. Associating a group of virtual channels with different QOS requirements with a a single resource allocation, and

b. Resource allocation renegotiation in the order of 20msec.

5) What N-ISDN bit rates will be supported by the B-ISDN circuit emulation mode?

Addenda

Ad.1 Cell los ratio

Various cell-based video coding systems have been developed or simulated, collectively capable of satisfactory performance in the face of a variety of cell loss ratios. However, the actual figure to be expected from the network for a particular video service application and bit rate will determine both the need for cell loss protection or recovery and the method to be used. The Experts Group expects to identify appropriate cell loss ratios for video services on the B-ISDN, and will input this information to SGXVIII when available.

The cell loss ratio has fundamental implications for the video coding strategy and its efficiency. If, for example, layered video coding systems are to be used, exploiting the availability of the cell loss priority indicator, an indication of the cell loss rate for each priority level is necessary. Figures for expected cell insertion rates are also required.

Ad.2 Cell loss burst characteristics

The question of whether cells are lost in isolation or in bursts is fundamental for the video coding approach. Some coding schemes are proposed which provide a means of protecting against bursts of cell loss, but they may not be necessary if cells are lost in isolation (i.e. if cell losses are uncorrelated).

- Will cell loss be dominated by network congestion?
- Will bursts of cell loss result from network congestion?
- Will the cell loss burst length be service rate dependent?
- Will high priority cells be affected by network congestion?

Clarification of these points is sought from SGXVIII, and additional guidance to an appropriate statistical model to characterize bursts of cell loss would be welcome.

Ad.3 Use of CLP indicator

The use of CLP is useful for some coding schemes to provide tolerance to cell loss.

- Under what circumstances would the service provider set the CLP indicator?
- Could the CLP be changed by the service provider after a user has set it?

- Will the usage monitoring structure encourage the use of both high and low priority cells?
- Will the rate of high and low priority cells be negotiated independently with the network?

Other information concerning call admission control and usage monitoring that would impact on the user's choice of a combination of high and low priority cells would be welcome.

Ad.4 Usage parameters

Ad.5 Multimedia connections

The ability of the B-ISDN to perform the multiplexing task provided by the terminal on circuit-switched networks makes it attractive to consider cell-by-cell multiplexing (by use of different Virtual Channels, or possibly on a single VPI/multiple VCI's) for the provision of multimedia connections. However two issues arise here, as discussed below:

- Will the network be capable of providing connection admission and monitoring based on the group of VC's constituting a multimedia connection? If not, would users see a penalty in the use of multiple VC's, and be encouraged to perform multimedia service multiplexing at a higher layer? SGXVIII should be aware of this possibility and consider whether this capability can be accommodated. Previous experience with multimedia services suggest that a group of at least seven VC's may be necessary, but we would like to know if there is an upper limit.

- Differential VC delay. If multimedia connections (video and associated audio in particular) are supported over multiple VCs, there exists the possibility of differential delay. If excessive, this may require end-to-end signalling overheads to add time-stamps and permit resynchronisation. What is the expected limit on differential delay between VCs?

Ad. 6 Bit error rates

We assume that cell payloads are subject to a small probability of transmission bit errors. The statistics of such errors will determine the need for, and type of, error correction mechanism and the overhead necessary to achieve this. It could also influence approaches to, and efficiency of, video coding and choice of code word assignment schemes.

What is the expected probability of transmission bit errors, and are these errors likely to be uncorrelated or bursty? Draft Rec. I.363 notes (section 2.3 and 3.3) that one of the functions of type 1 and 2 AALs is the '...monitoring of user information field or bit errors and possible corrective action...'. The Experts Group wishes to work with SGXVIII to further clarify the functionality of the AAL in this respect.

Ad.7 Cell delay and jitter

The fixed component of end-to-end delay is an important factor for conversational video services. It will impact on the choice of coding method and allowable buffering within the encoder and decoder.

What is the expected maximum B-ISDN delay, including processing and queuing within the B-ISDN switching equipment?

The variation in delay, or jitter, determines the size of receive buffers necessary for its removal, and therefore again influences the total end-to-end delay. What are the expected statistics of cell delay jitter? Is a hypothetical reference connection available or planned, that would assist in these matters?

Ad.8 Network model for hardware experiments

It is the initial intention of this Experts Group to target hardware trial of ATM video codecs for the second half of '93. Success of such trials will depend on the availability of network equipment or simulators. Would these be expected in this timeframe?

Ad.9 AAL

Draft Rec. I.363 describes AAL type 1 & 2 structures which could be used for real-time video services. To make progress, the experts group intends to distinguish between:

- an AAL suitable for existing video services (e.g. H.261), that could be standardized in the relatively short term,
- for future ATM video coding standards, an AAL matched to the specific coding algorithms will be necessary. It is premature to define an AAL for these applications at this time.

Conclusion

Close liaison between the Experts Group for ATM Video Coding and SGXVIII will be necessary to harmonize and optimize B-ISDN network design and video service provision. It is the Experts Group's intention to provide input to SGXVIII on requirements for network performance and to assist SGXVIII in the definition of the AAL for type 1 & 2 service categories.

In the first instance, however, the Experts Group requires guidance to provide bounds for certain network parameters crucial to the development of appropriate video coding methods. The main issues have been highlighted in this document. A model of the network is necessary to permit commencement of video coding studies. The parameters of such a model, along with some estimates of possible parameter values, is provided in the Annex 4 of this document. This is a first attempt, to initiate studies. Refinements will be made at future meetings. SGXVIII is invited to comment on, or correct this model.

END

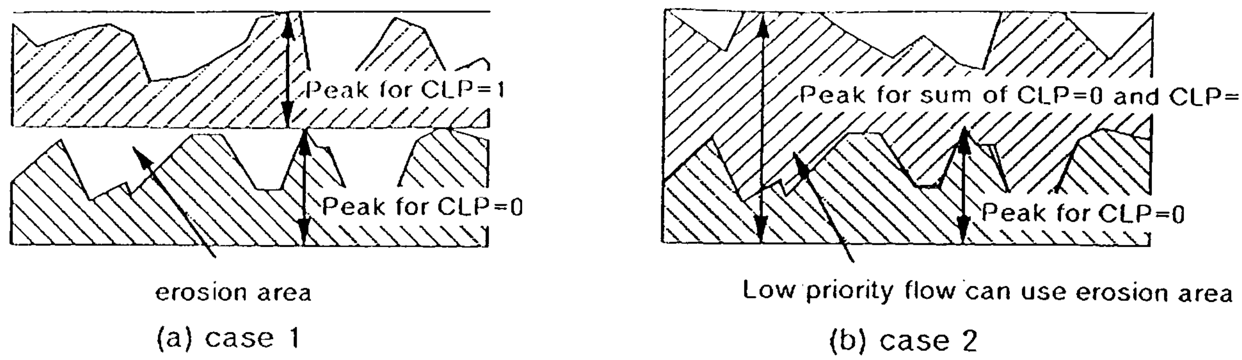


Fig.1 Peak bit rate negotiation

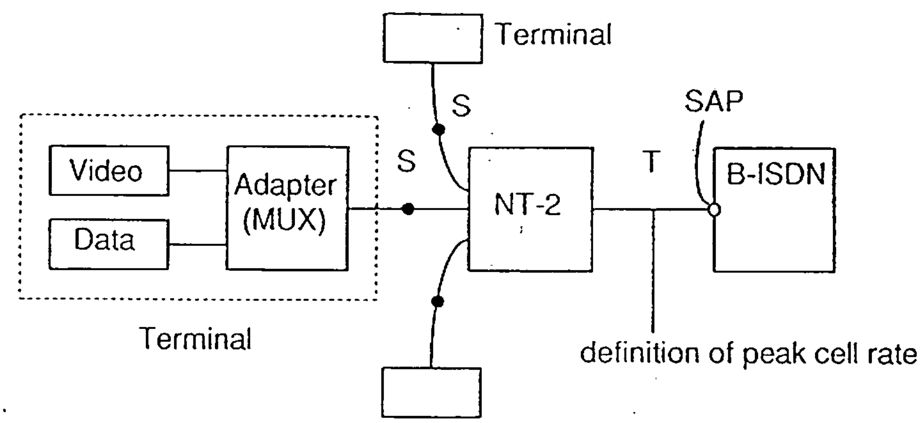


Fig2. An example of system configuration

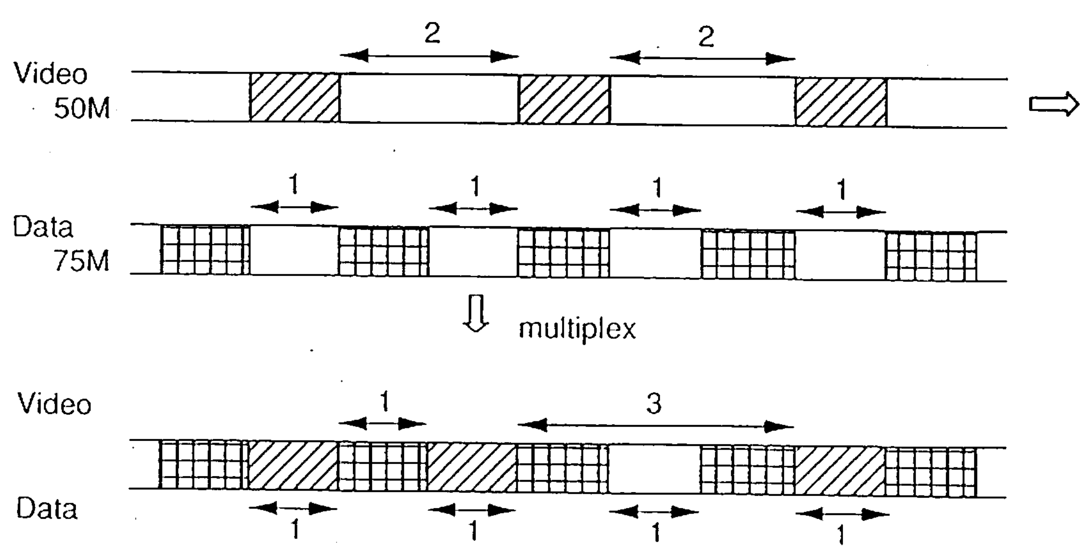


Fig3. The variation of peak cell rate caused by multplex

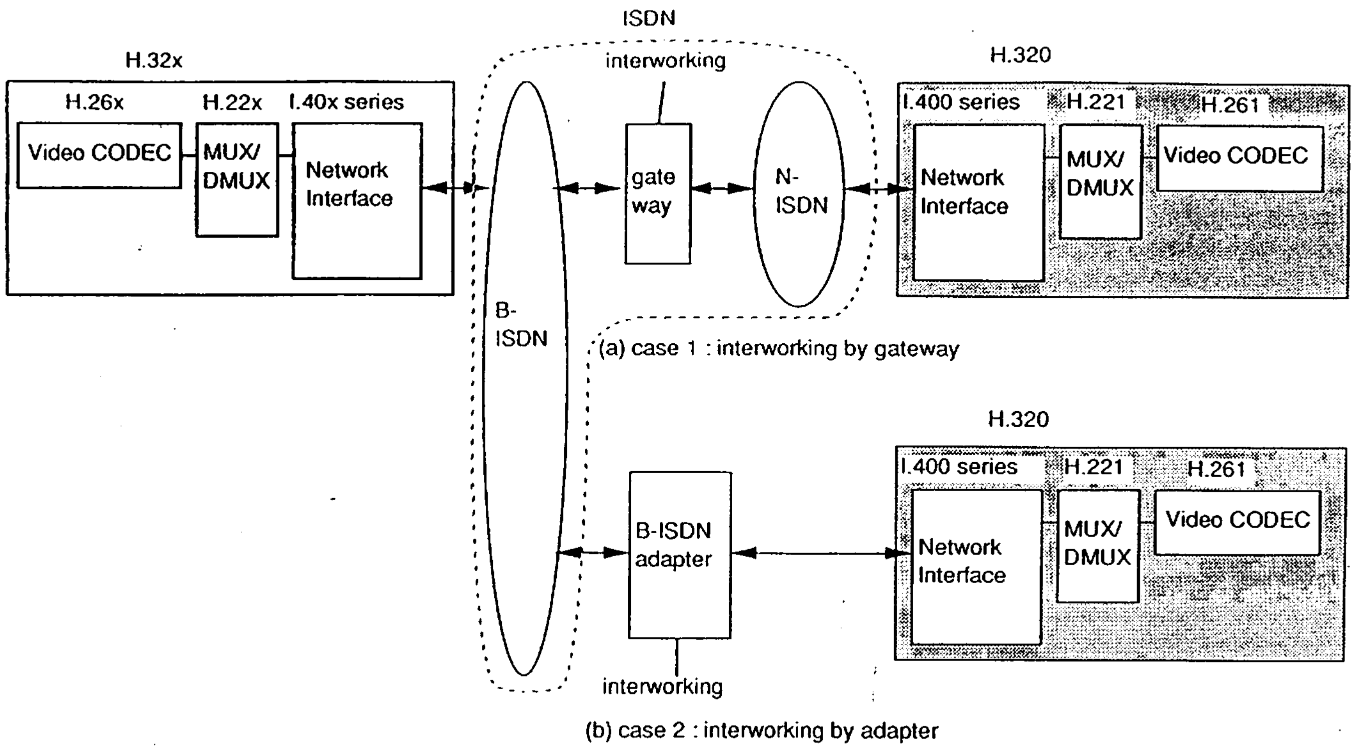


Fig. 4 Two cases of interworking between H.32x and H.320 terminals

COMMENTS ON IBS BASELINE DOCUMENT (JUNE 1991 EDITION)

The following comments of the Experts Group were obtained at its meeting in August 1991.

---

1. Cell loss ratio (page 16, 2nd paragraph, 2nd and 3rd sentences)

Please revise the sentences as follows:

Current sentences;

"For example, a high quality videoconference connection... This may be..."

New sentences;

"Table 1 (as attached at the end of this Annex) provides some network performance requirements obtained from some example service quality figures. The table concentrates on bit error and cell loss error correction techniques. Layered coding concealment techniques are however under consideration and lead to different figures."

2. A2.6 Service bitrate (page 11, 2nd paragraph)

Usage parameter control and network parameter control, CLP bit.

"When the cell tagging option is exercised, non compliant CLP=0 cells may be overwritten to CLP=1".

We are concerned that the network can modify the CLP bit. Some layered coding techniques intend to use CLP bit for layer indicator. For such a case, changing the CLP bit may cause more problems than discarding the cell.

This fact has already been mentioned in the last liaison statement from SGXV;

Use of CLP bit (page 16)

"The CLP bit is seen as a useful mechanism to provide protection against cell loss by controlling that information which might be lost. It is crucial that, after a cell is labeled 'high priority' by a terminal device, this is not changed by the network."

3. QOS related to Cell Loss Priority (bottom of page 12)

QOS related to Cell Loss Priority

"The CLP indicator in the cell header may be set by the user or the service provider. In the case of video services, the CLP bit is set by the layered coding provider..."

The wording should be clarified; e.g. who is "the layered coding provider"? Is he a user?

4. List of H.26X requirements

We have identified such functional requirements as in Annex 4 of this document for the high quality video coding standard H.26X, which we propose to be included in the IVS Baseline Document.

END

Table 1 SERVICE AND NETWORK REQUIREMENTS

Service	Bit rate	QOS requirements (***)	Required BER/CLR without error handling in AAL	AAL type	Required BER/CLR after single bit error correction on cell basis in AAL (*)	Required BER/CLR after single bit EC on cell basis and addit. cell loss correction in AAL (**)
Communication						
videophone	64kbps/2Mbps FBR (H261)	30 min error free	BER<1.e-6 CLR<1.e-7 (BCH(511,493) FEC in user layer)	type 1	in user layer	BER<... CLR<8.e-5
videophone	2Mbps VBR	30 min error free	BER<3e-10 CLR<1e-7	type 2	BER<1.2e-6 CLR<1e-7	BER<2.3e-5 (CLR=1e-6) CLR<8e-5
videoconference	5Mbps VBR	30 min error free	BER<1e-10 CLR<4e-8	type 2	BER<8e-7 CLR<4e-8	BER<1.8e-5 (CLR=1e-6) CLR<5e-5
videodistribution						
TV distribution	20-50Mbps VBR	2 hours error free	BER<3e-12 CLR<1e-9	type 2	BER<1.2e-7 CLR<1e-9	BER<6e-6 (CLR=1e-6) CLR<8e-6
MPEG1 core	1.5Mbps VBR	30 min error free	BER<4e-10 CLR<1e-7	type 2	BER<1.4e-6 CLR<1e-7	BER<2.5e-5 (CLR=1.e-6) CLR<9.5e-5
MPEG2 core	10Mbps VBR	30 min error free	BER<6e-11 CLR<2e-8	type 2	BER<5.4e-7 CLR<2e-8	BER<1.5e-5 (CLR=1.e-6) CLR<4.e-5

(*) Payload scrambling polynomial 1+x**43 produces double, correlated bit errors.
(**) Based on parity cell built from 31 consecutive data cells (see further in annex 1). The cell losses are assumed to be isolated. With this simple correction scheme, single cell losses can be corrected if combined with cell loss detection by cell numbering. Also non-corrected but detected bit errors in a cell are handled by replacing this faulty cell by a dummy cell followed by correction of this cell by the cell parity mechanism. The BER calculations are done in the assumption that all double ATM link errors (2 times 2 correlated errors due to payload scrambling) can be detected.
(***) QOS requirements, as visualized by viewers; not directly related to channel errors.

Notes

- These values are calculated under the assumption that cell losses are isolated. If cell losses tend to occur successively, another cell loss ratio and another cell loss correction technique may be required.

- We assumed that one cell loss always causes a picture degradation. The visual perception of the picture, however, may be acceptable even if cell loss concealment technique is not used. Therefore there is a possibility that these requirements will be relaxed.

H.26X REQUIREMENTS

Status notation

(A) Agreed
(P) Preferable
(M) Mandatory
(T) Target
(FS) Implementation method is for further study

1. BIT RATE

up to several 10s Mbit/s	(A)

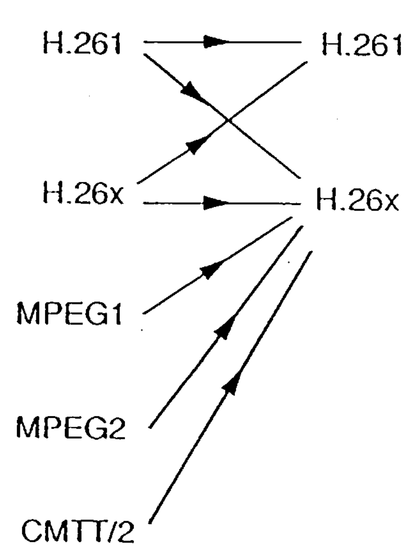
2. CODEC SOURCE FORMAT

QCIF/CIF	(A)
"601" class	(FS)
EDTV	(?)
HDTV	(?)

3. COMPATIBILITY

encoder		decoder		
H.320	--->	H.32X	(terminal)	(A,M)
H.32X	--->	H.320	(terminal)	(A,M)
H.261	--->	H.26X		(P,FS)
H.26X	--->	H.261		(P,FS)
MPEG1	--->	H.26X		(P,FS)
MPEG2	--->	H.26X		(P,FS)
*"CMTT/2"	--->	H.26X		(P,FS)

* Secondary distribution, which may include classes above "601"

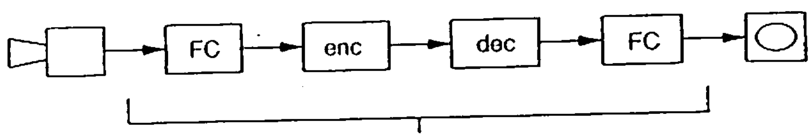


4. PICTURE QUALITY

"PAL/NTSC" at 3-5 Mbit/s and delay=?	(T,FS)
"Rec. 601" at 8-10 Mbit/s and delay=?	(T,FS)

5. DELAY

less than about 150 ms at bit rate > 2 Mbit/s	(FS)



6. CODEC COMPLEXITY

complex/high performance
vs
simple/low performance
ex. pure intra-codec

7. APPLICATIONS

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

8. ATM

VBR and CBR	(A,M)
Cell loss resilience	(M,FS)
Bit error resilience	(M,FS)
High/low priority cell utilization	(P,FS)
High/low priority cell independent rate control	(P,FS)
Usage Parameter Control	(M,FS)

9. MULTIPOINT

Continuous presence possible	(P,FS)

- Time-sliced decoding
- Editing without decoding-recoding

Mix of H.320 and H.32X	(M,FS)

10. H.32X TERMINAL

Interwork with

H.320 terminal	(A,FS)
Network database	(P,FS)
Distributive service	(P,FS)
Multipoint	(A,FS)
Stored bitstream	(P,FS)

Multimedia multiplexing	(M,FS)
Audio quality > ?	(FS)
Relative audio/video delay < ?	(FS)
Video clock recovery	(FS)
Encryption/scrambling	(FS)

END

Update of the First Simplified Network Model

The first simplified network model, Annex 4 to AVC-22R (Hague), has been updated to calculate the cell loss ratio. See AVC-97. It is noted that the cell loss ratio CLR is calculated in general according to the following equation:

$$CLR = \int_0^{\infty} L(R)*p(R)dR \Big/ \int_0^{\infty} R*p(R)dR$$

where R is instantaneous rate of multiplexed signals, p(R) is probability density for instantaneous rate being R, L(R) is loss function. L(R) can be approximated as

$$\begin{aligned} L(R) &= 0 && \text{for } R < CAP \text{ (capacity of the multiplex)} \\ &= R - CAP && \text{for } R \geqq CAP \end{aligned}$$

1) A single stage multiplex is assumed.

2) The network is assumed to exhibit a cell loss/network load characteristics as shown in Appendix 1. An example of these characteristics are shown in Appendix 2.

3) The multiplex is assumed to have a maximum available bandwidth (CAP) of 100 Mbit/s.

4) Cell loss is assumed to be random.

5) The model is independent of the shape of the source.

6) The model provides the lower limit for the network loading, ie it is a conservative model because of point 5 above.

<u>Note</u> - We need more accurate model(s) for the precise study of VBR vs CBR issue and to obtain better understanding and deeper insight to the statistical multiplexing problem. Suitable models for this purpose need continued study.

Appendix 1

Psat = exp(-n * k) where Psat is the probability of saturation

k = (a * ln(a/p)) + ((1-a) * ln((1-a)/(1-p)))

a = CAP / (n * Peak)   Note: 0 < a < 1

CAP = Maximum capacity of the multiplexer output in Mbit/s

n = percentage of network loading

Note: n is the number of sources, 0 < n <= 100. The illustration corresponds to the case of Mean = 1 Mbit/s.

Peak = The maximum bit rate over simulation interval (measured on a per frame basis)

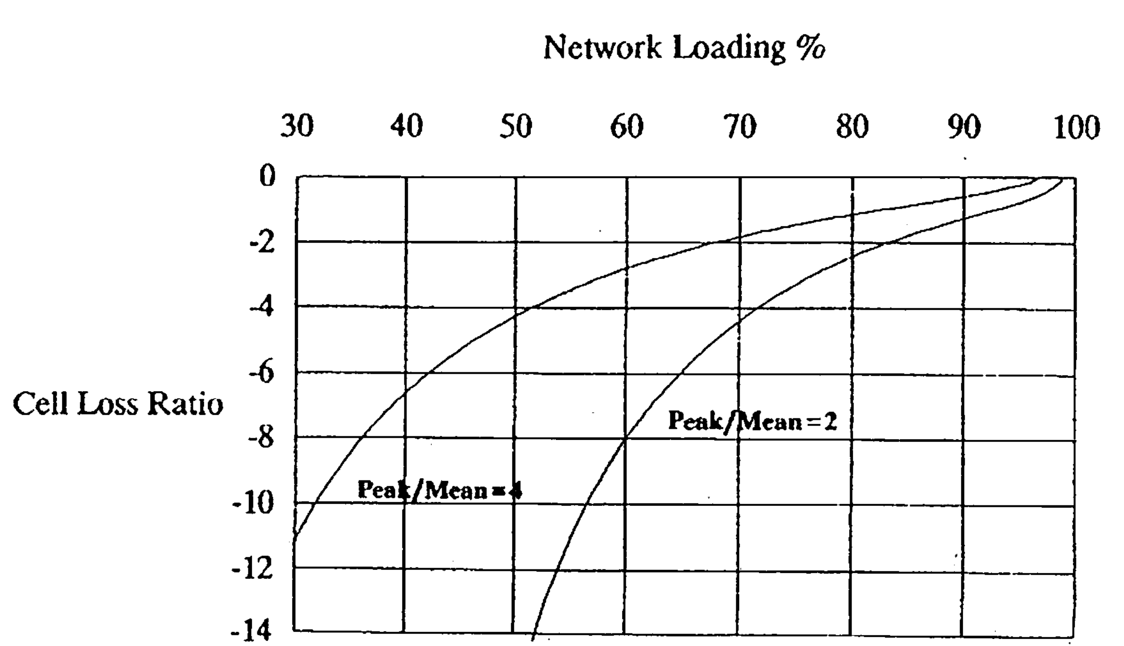
p = Mean/Peak   Note: 0 < p <= 1

Mean = The average bit rate over the simulation interval

CLR = Psat / (n * Mean * ln ((a * (1-p))/(p * (1-a))))
where CLR is the cell loss ratio.

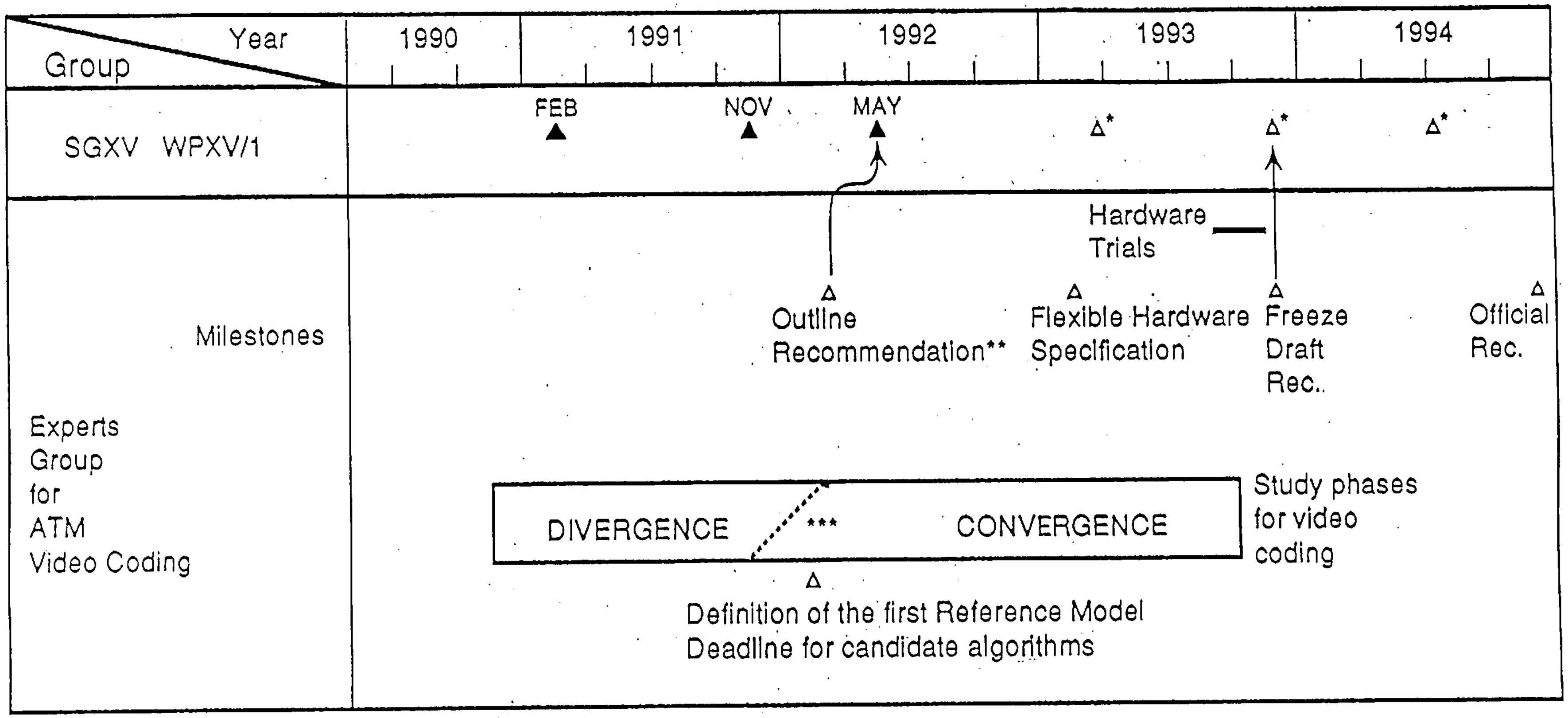
ln = log to the base 'e'

Appendix 2



Cell Loss Ratio vs. Network Loading (Example)

Work Plan for the Experts Group



* Meeting schedules for the next sutdy period (1993-1996) are not yet decided. These are copied from those of the current study period.

** This outline Recommendation includes scope, list of contents, such parameters as picture formats, framework of coding scheme, etc. which are agreed by that time.

*** This demarcation may vary according to the progress of the coming year.