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

Document AVC-96
August 7, 1991

SOURCE : Experts Group for ATM Video Coding in CCITT Study Group XV
TITLE : Status Report on ATM Video Coding Standardization
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 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. General B-ISDN terms are defined in Recommendation I.113.

The following items are defined in Annex 1.

backward compatibility
downward compatibility
embedded bit stream
forward compatibility
layered coding
simulcast
standard families
switchable encoder
syntactic extension
upward compatibility

{term, definition, note}

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 overviewed in Table 2, listing opportunities as well as limitations (see AVC-20, AVC-4).

Table 2

A list of questions concerning ATM network characteristics which affect the design of the video coding was formulated as in Annex 2 at the November 1990 meeting for consideration of Study Group XVIII. Two responses from SGXVIII (November 1990, Matsuyama) are also included in Annex 2.

We endorse the IVS baseline document approach which was made in the same SGXVIII meeting and we keep contributing to its upgrading. As a first step, a liaison statement including unsolved questions was drafted and sent to SGXVIII at the second meeting in Paris. The text is included in Annex 3.

{Annex 2 and Annex 3 are to be edited as a Q&A form}

The timing recovery for the AV codecs in B-ISDN was recognized important and requires further study.

4.3 Technical requirements for ATM video coding

4.3.1 Video signals to be handled

Initially we concentrate on video coding of 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.

4.3.4 Average bit rate

It should cover a range from 64 kbit/s 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.

Layering is one of the solutions and the idea of "flexible layering" which exploits B-ISDN characteristics and provides service integration is recognized interesting.

Exact ways to implement compatibilities require further study.

4.5 Requirements for H.26X

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

In addition to this list, the following items are recognized as possible requirements for H.26X;

- Multipoint consideration taking into account B-ISDN capabilities
- Capability to operate in intraframe mode only
- Provision for both backward/downward and downward/upward compatibilities and adaptability to future extension

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).

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?

5.3 Action points

As a bases for answering the above questions, and thereby arrive to a conclusion to the format question, the following action points were listed for the next meeting:

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

- A listing of PROS and CONS for different format solutions.

- 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 model

We feel it would facilitate our making a good progress if we use a common network 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 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 and 61, into consideration.

7. Video coding model

(Reference model for studying video coding appropriate for ATM environments)

8. VBR vs CBR

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.

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.

(This was mentioned several times in the meeting, is this agreed upon?)

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 2)

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

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

Figure 2

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.

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)

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 1.1 and 1.2.

Figure 1.1
Figure 1.2

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 1.3

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.

N.B. For simulations the low cell loss rate can be assumed to be: $P_{loss}=0$

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 $\langle t_i, t_{i+1}]$ 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 2):

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 is one 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 Definition of AAL

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).

Further study is required.

11.3 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) is introduced to clarify the requirements for ATM and AAL layers.

12. 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.

13. Harmonization with other standardization bodies

This group seeks to carry out joint work with TG CMTT/2 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 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.

14. Outstanding questions

14.1 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 ralated 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 trend 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?

14.2 Video coding

14.2.1 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?

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

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

6) 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, ...

7) Cell loss protection and recovery methods?

8) 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?

9) 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?
- ...

14.2.3 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?

14.2.4 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)

See also Figure in Annex 7 to AVC-22R.

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

3) How should it be related to CMTT standards?

14.2.5 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?

14.3 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.

14.3.1 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.)?

14.3.2 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?

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.

14.4 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

14.5 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. 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
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 following;

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 type of Macro	Fixed		–
	Inter+coef		–
	Inter+Coef+MC+Fil		–
	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

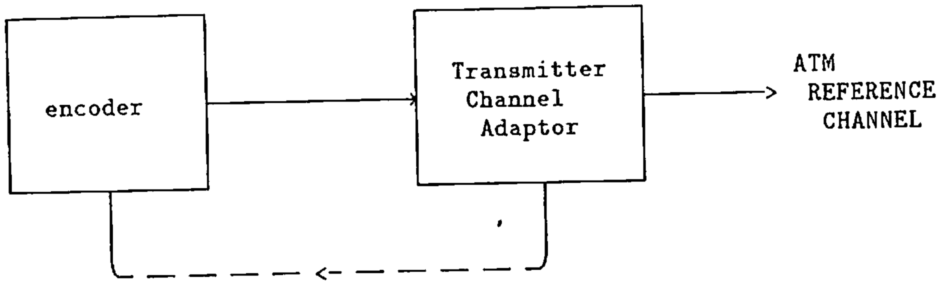


Figure 1.1: ATM Reference Encoder

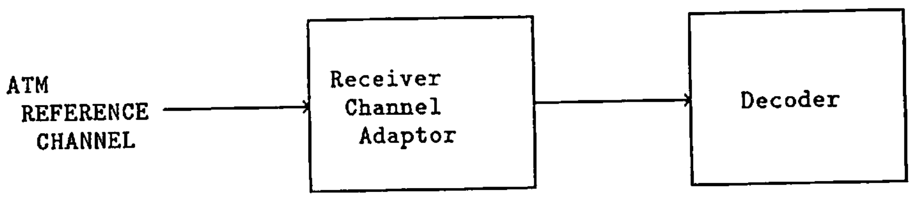


Figure 1.2: ATM Reference Decoder

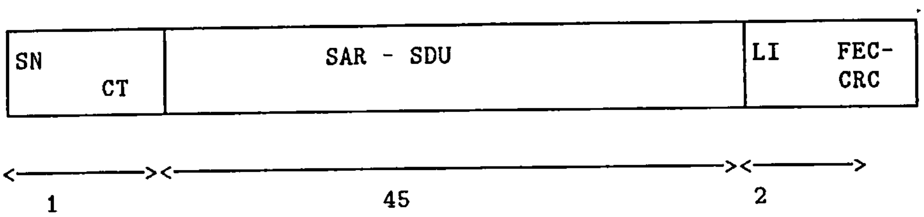


Figure 1.3: ATM cell information field

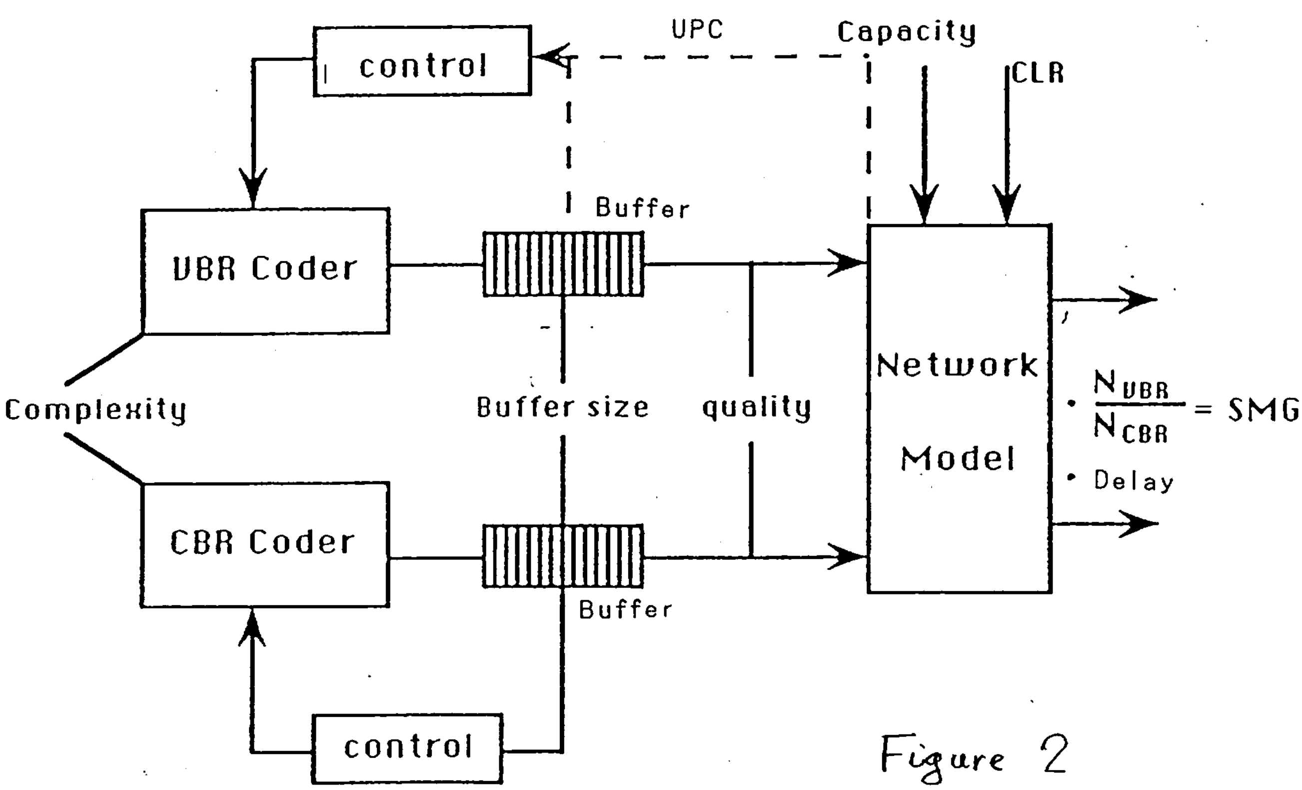


Figure 2

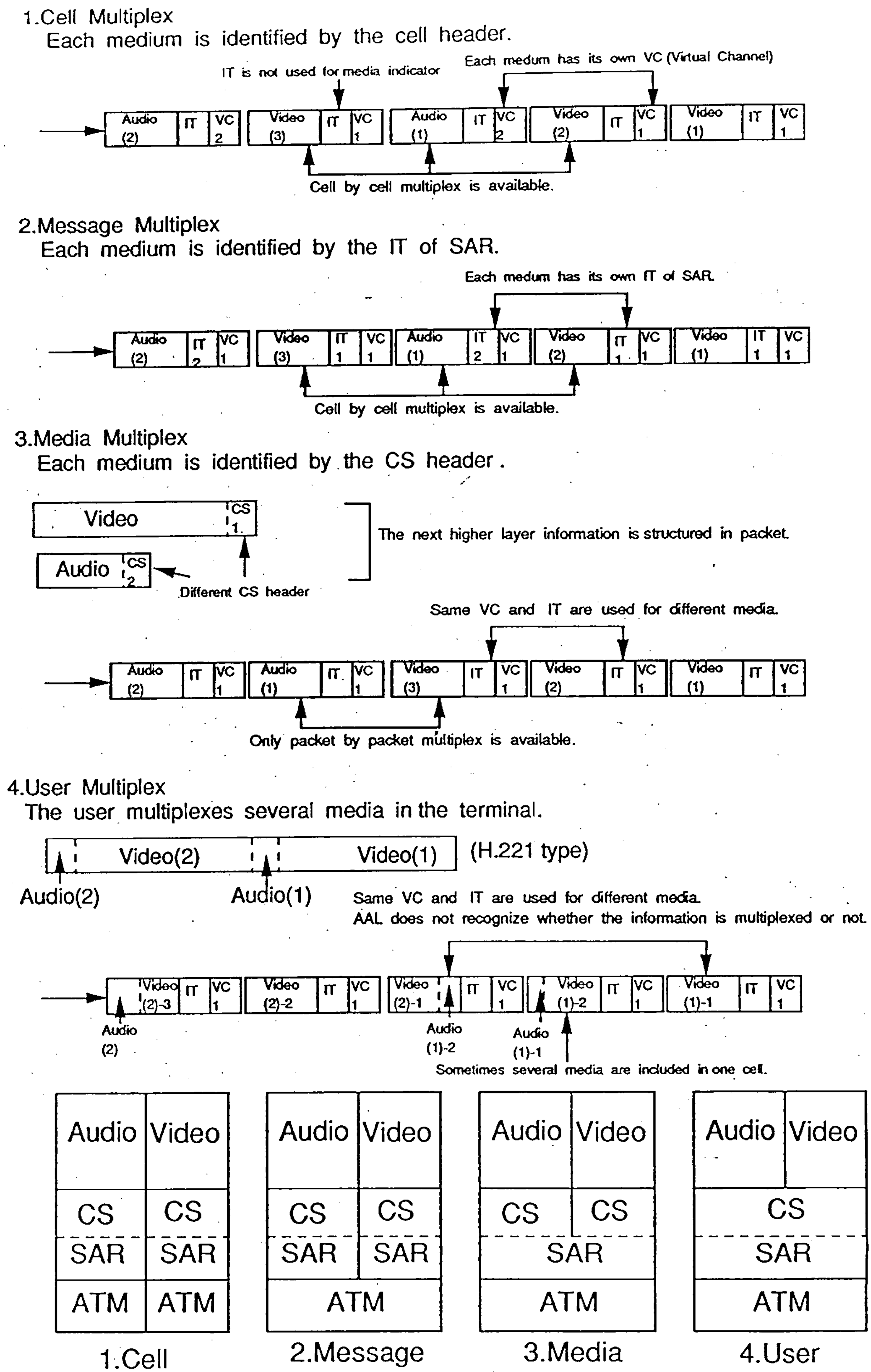


Fig. 3 Multimedia multiplex methods

Terminology

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

Source: Experts Group for ATM Video Coding in SGXV
Title : Liaison statement to SGXVIII on requirements for B-ISDN network model as it impacts on video coding
Date : November 16, 1990

The Experts Group for ATM Video Coding as established at the July 1990 meeting of Study Group XV has initiated its activities by holding the first meeting in The Hague (13 - 16 November 1990).

Since this Group is intending to make Recommendation(s) for video coding in B-ISDN environments, we need correct understanding of the ATM network characteristics. For this purpose, we have formulated a list of questions at the first meeting as attached which are expected to be answered by SGXVIII at the earliest occasion.

The Experts Group would welcome the participation of a liaison representative from SGXVIII to assist in clarification of the B-ISDN definition and performance. This would be particularly valuable at the second meeting (tentatively scheduled for 12 - 15 March 1991), since multiplexing of variable rate sources under certain cell loss conditions, and refinement of our network model, is expected to be discussed in detail.

Liaison to SGXVIII: Requirements for B-ISDN network model as it impacts on video coding

Abstract

The SGXV Experts Group on Video Coding for ATM held its first meeting in The Hague, 13-16 November 1990. An important outcome was the recognised need for a realistic model of the target B-ISDN to permit progress in the definition of suitable video coding methods. This liaison statement outlines the impact of various network parameters on video coding issues, and requests guidance from SGXVIII concerning parameter values. To permit advancement of video coding work, an initial network model is offered for comment by SGXVIII.

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 itemises the main issues and the addenda will give a more comprehensive explanation.

Issues

1. Cell loss ratio:

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

2. What is the cell loss burst behaviour?

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?

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?

6. Bit Error Rates:

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

7. Cell delay jitter:

   - What values are expected?

8. Network model for hardware experiments:

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

9. 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.

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 characterise 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 I 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 standardised 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 Appendix. 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

Attachment: Annex 4 to AVC-22R

CCITT

Temporary Document 8 (XV/1)

STUDY GROUP XV

Geneva, 18 February - 1 March 1991

Questions: 3/XV; 5/XVIII

SOURCE: XVIII/6 (Matsuyama, November 1990)

TITLE: LS TO SG XV (EXPERT GROUP FOR ATM VIDEO CODING ON SIMULATION RESULTS ON CELL LOSS IN ATM NETWORK

SUBJECT: Information on ATM network performance

Contact Point:
Special Rapporteur for Q.5/XVIII
N. Seitz
US Department of Commerce NTIA

The liaison statement from SG XV Expert Group for ATM Video Coding was considered by Rapporteur Group of Q5/XVIII during Matsuyama meeting(Dec. 1990).

The Q5 Rapporteur Group decided to inform SG XV Expert 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 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.

Response of SGXVIII to Questions 4

CCITT

Temporary Document 10 (XV/1)

STUDY GROUP XV

Geneva, 18 February - 1 March 1991

Questions: 1-4/XV;2,13/XVIII

SOURCE: WP XVIII/8

TITLE: LS TO SG XV (for information)
TRAFFIC CONTROL AND RESOURCE MANAGEMENT ASPECTS

In CCITT SG XVIII, during the Matsuyama meeting (26 Nov, 7 Dec, 1990), a Subworking Party was established to study the Traffic Control and Resource Management aspects of B-ISDN.

First results are relevant to traffic characterizing 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 the already achieved results.

If both low priority and high priority cells are carried within one Virtual Channel (e.g. for layered video coding), CCITT SG XVIII is currently discussing whether both high priority (CLP bit set to zero) and total traffic require 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.

As for the Usage Parameter Control function, SG XVIII 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.

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.

In CCITT SG XVIII, 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 throughout the network.

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

LIAISON STATEMENT TO SGXVIII

CCITT
STUDY GROUP XVIII
Geneva, 11-28 June 1991

Temporary Document (XVIII/8)

Questions: 3,4/XV; 2,13,22/XVIII

SOURCE : EXPERTS GROUP FOR ATM VIDEO CODING IN SGXV
TITLE : LIAISON STATEMENT TO SGXVIII
Purpose: Action

---

1. Introduction

The second meeting of the CCITT SGXV Experts Group for ATM Video Coding was held in Paris, 23-31 May 1991, to progress studies on video coding for services on the B-ISDN.

Many network related issues impact upon the work of the Experts Group, as reflected by the list of questions put to CCITT SGXVIII after the first meeting in The Hague in November 1990. Many of those questions cannot be answered in detail at this stage, but responses from SGXVIII to some of the questions were welcomed. Some additional issues were raised at the second meeting, and these are detailed below in Section 2.

The Experts Group also considered the Integrated Video Services (IVS) Baseline Document initiated by SGXVIII in Matsuyama in Nov./Dec. 1990. The SGXV Experts Group welcomes this initiative taken by SGXVIII under its coordinating role, and wishes to advise that it intends to offer substantial contributions during the evolution of the document. As an initial step, some text is offered in Section 3 below.

2. Network Related Questions

The SGXV Experts Group would welcome a response from SGXVIII on the following issues to assist in the progress of video coding and video system architecture developments.

<u>CLR for services with differing rate behavior</u>

Will VBR and CBR services be subject to the same CLR (in both low and high priority classes)?

<u>Differential delay between virtual channels</u>

The Experts Group recognises 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?

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.

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.

3. IVS Baseline Document

The SGXV Experts Group welcomes the creation, and recognises the value, of the IVS Baseline Document as a means of coordinating video coding studies for the B-ISDN. Continued support of the document, and contributions to it, will be provided.

3.1 Video Service Interworking

The CCITT SGXV Experts Group for ATM Video Coding proposes that the text of the IVS Baseline Document of Dec. 1990 (Matsuyama) be modified as follows.

The text contained in Annex 4 (Video Service Interworking) should be deleted, and replaced with the following:

Annex 4. Video Service Interworking

Integration of video services is recognised as a key objective for ATM Video Coding. It is an agreed target for the video coding system under study by the SGXV Experts Group. Several options exist for interworking between services:

Negotiation Approach:

At the commencement of a connection, terminals negotiate a set of parameters with which both can cope. A set of standards of increasing quality would be defined and a basic capability assumed for all terminals.

Simulcast Approach:

Transmitting terminals contain multiple encoders, operating at a variety of resolutions and quality levels so that broad interconnectivity can be achieved. Receiving terminals could be simple devices able to receive one of the bit streams, or could contain multiple decoders allowing a selection.

Layered Signal Approach:

A hierarchical representation of the video signal is defined. Coders transmit a baseband signal which provides a basic quality service. Incremental signals, which can be used along with the baseband to recover a high quality signal, are also transmitted. Receiving terminals utilise the baseband and an appropriate number of incremental signals to recover the video signal to the quality which they are capable of displaying. Transmitting terminals provide the number of signals which is commensurate with their input signal quality. Note that "embedded bitstream" and "syntactic extension" techniques are also versions of layered coding (see TG CMTT/3 liaison statement to SGXVIII dated 17 April 1991 for the terminology).

A range of issues needs to be considered in comparing these different approaches, including complexity, coding rate penalties and performance. Negotiation would seem inappropriate for multipoint and distribution services, whereas simulcast seems inappropriate for storage applications (e.g. store and forward video mail). Layered coding seems suited to the widest application range. "Flexible layering" in which any number of layers can be used in any particular application, appears to provide broad interworking capability with few restrictions, and is currently one of the options being studied.

It is recognised that to provide easy interworking or conversions between services, and to use common display components on a terminal device intended to access multiple video services, the definition of a family of picture formats would be beneficial. Picture formats represent an important area that will influence video coding and it is being studied actively in the SGXV Experts Group.

- END -

3.2 Network Issues

The CCITT SGXV Experts Group for ATM Video Coding proposes that the text of the IVS Baseline Document of Dec. 1990 (Matsuyama) be modified as follows.

The text contained in Annex 2 (Network Aspects) should note the needs of the Video Coding Experts to be advised of certain parameters having important impact on the coding:

Annex 2. Network Aspects

Add a Section with title "Network Parameters Impacting on Video Coding Definition" as follows:

A number of parameters and operational procedures concerning the B-ISDN network will have significant impact on the definition of appropriate coding schemes for the support of video services. The areas requiring definition are listed below:

Cell loss ratio

This is an important determinant of the quality of service achievable for a video application. It determines the means, and even necessity, for providing cell loss protection for different services. It is recognised that there is a degree of flexibility in this figure, since the network operators have some flexibility to dimension the network to provide certain cell loss ratios if they are considered essential for some video services, while the codec design can also be changed to accommodate different figures. Progress needs to be made, though, perhaps by considering the impact of a range of cell loss ratios on both network and codec. The cell loss ratios for both priority levels need to be defined. The SGXV Experts Group believes that guaranteed overall cell loss ratios, for both priority levels, will be essential to satisfy video quality of service requirements. Guaranteed performance, at least within certain time intervals, will also be required.

If the cell loss ratio is sufficiently small, no cell loss protection may be necessary. For example, a high quality videoconference connection operating at 10 Mbit/s would suffer only one cell loss every 10 hours with a CLR of 10^{-9}. This may be acceptable even if the cell loss caused visible degradation.

Studies are required to determine the quality of service parameters available to the user, and to relate these to cell loss ratio.

Cell loss burst behavior

It is understood that cell losses may occur in bursts. This impacts on the means of cell loss protection; the use of forward error correction may be too expensive and delay may be excessive for conversational services if multiple consecutive lost cells must be detected and corrected. Cell loss burst behavior may be modeled by the Gilbert model (a two-state Markov model requiring four transition probabilities, with one state representing no cell loss and the other constant cell loss).

Open questions remaining are:

- How will the cell loss burst behavior depend upon the service rate?

- Will the burst behavior of high priority cells differ from that of low priority cells and, if so, how?

- How can we estimate the average interval time, T, in which no cell loss occurs? If T>>1/(bitrate x CLR), the requirement for CLR might be relieved.

Use of CLP bit

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.

Open questions:

- Will there be separate negotiations for the two priority levels?

- Will the usage monitoring structure encourage use of both high and low priority cells?

- What options are available in selecting the quality of service?

Usage parameters

The rate statistics required of a video encoder have a significant impact on its performance (in terms of picture quality and delay). For circuit switched networks, the target was straightforward: minimise the rate and keep it constant. For the B-ISDN (with the possible advantages of variable rate over constant rate operation), entirely different rate control strategies may be appropriate, and these could have a significant impact on codec performance. At this stage, the only clear decision is that peak rate will be an important parameter that is monitored.

In our group the term "window" means the policing time for the average bit rate. The following methods are considered for policing in the network:

<u>Jumping window</u>:

There is no time interval between two successive windows.

<u>Moving window (sliding window)</u>:

The window is sliding at a time step smaller than the window size.

<u>Stepping window</u>:

There is a time interval between two successive windows, which always start at a valid cell.

<u>Leaky bucket</u>:

Cells are put into a buffer and taken from the buffer at an average bit rate. If the buffer overflows, cells are discarded.

If a codec does not know when the network measuring window starts, it should control the bit rate by sliding window (the most severe method). Is there any way in which the starting time of the network measuring window can be known?

Open questions:

- 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.? Longer intervals (significantly greater than a video frame period which is typically 33-40 ms) are preferred for video services.

- When the network capacity is very large, the bit rate requirements of a single user will be relatively small. In this situation it seems there will be very little difference in the required network resources for low and high priority cell loss classes. Will the high priority cell loss class continue to exist in the future?

<u>Multimedia connections</u>

Multiplexing of multiple media has been carried out within the terminal device for circuit switched networks. The B-ISDN already offers the flexibility to use cell-based or virtual channel based multiplexing instead. An important factor in the choice between terminal-based or cell-based multiplexing is whether there will be a penalty caused by the use of an ensemble of virtual channels instead of one composite one, although the overall rate characteristic, for example, would be the same. Most importantly, would the two options have the same transmission costs?

Some multimedia connections (most obviously associated audio, stereo in particular, and video channels) require synchronism. A concern arises, therefore, if the differential delay between virtual channels became noticeable in some service applications. This is unlikely to be a problem unless the cumulative differential delay exceeds some tens of milliseconds from end to end.

Open questions:

- How will multimedia services be handled in the B-ISDN?

- What signalling methods are being proposed?

- What kind of multimedia multiplexing method is preferred from the standpoint of network resource management?

Bit error rates

Cell payloads will be subject to a small probability of transmission error on the B-ISDN. 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 codeword assignment scheme. Estimates of the likely bit error rates are required by those working on video coding schemes for the B-ISDN.

For interworking between video codecs on N-ISDN and B-ISDN networks, the B-ISDN bit error rate must be no greater than that for the N-ISDN. It should also be noted that the H.261 coding scheme for N-ISDN provides bit error correction, so this would not be a necessary function of the AAL in this case.

SGXVIII should work in close collaboration with the video coding experts to define any capability within the AAL concerning bit error detection or correction.

Cell delay and jitter

The fixed component of end-to-end network delay contributes to the total service end-to-end delay and therefore is a determining factor in the overall quality of service. Estimates of the limits of B-ISDN delay are required to quantify such performance and determine its impact on video encoders and decoders.

The variation in delay, or jitter determines the size of receiver buffers necessary for its removal, and therefore again influences total end-to-end delay. The expected statistics of cell delay jitter need to be known to determine the impact on the video coding system and overall quality of service.

- END -

4. Conclusion

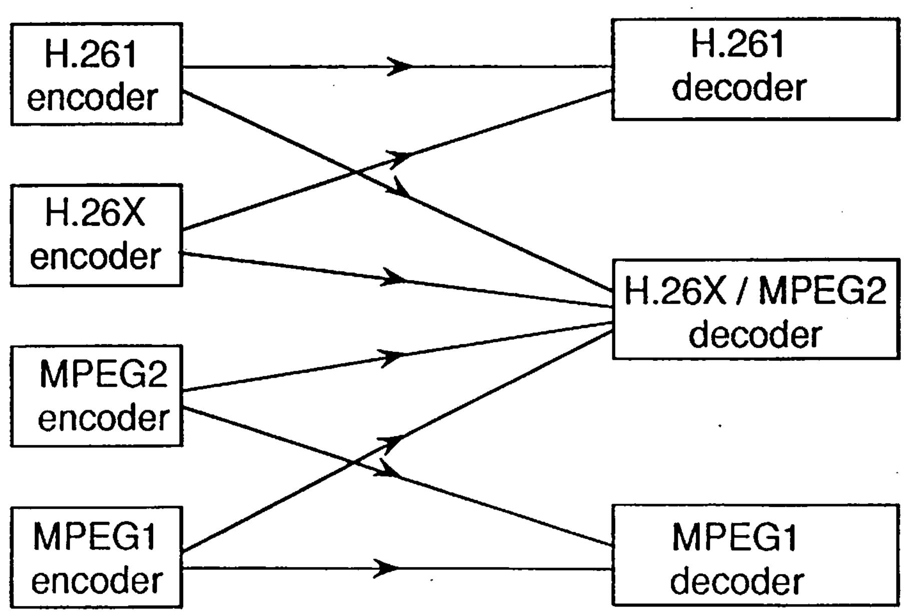
This document has raised some important network related questions and offered text for the evolving IVS Baseline Document, as part of continuing collaboration between the SGXV Video Coding Experts Group and SGXVIII on the development of video services for the B-ISDN.

END

Provisional List of H.26X Requirements

- Interconnectability on equipment level
- Bit rate range: 64 kbit/s - several tens Mbit/s
- CCIR 601 capability
- Various picture materials
- ATM network capability
    * cell loss resilience
    * variable/constant bit rate
- Consideration of conversational services
    * end-to-end delay ≦ ~150 ms
    * Multipoint
- Hardware verification

END



Annex 5 to AVC-96

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. Changes are highlighted by marker lines in the margin.

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.

(The old point 5 has been removed as it is not relevant to the network model)

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

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.

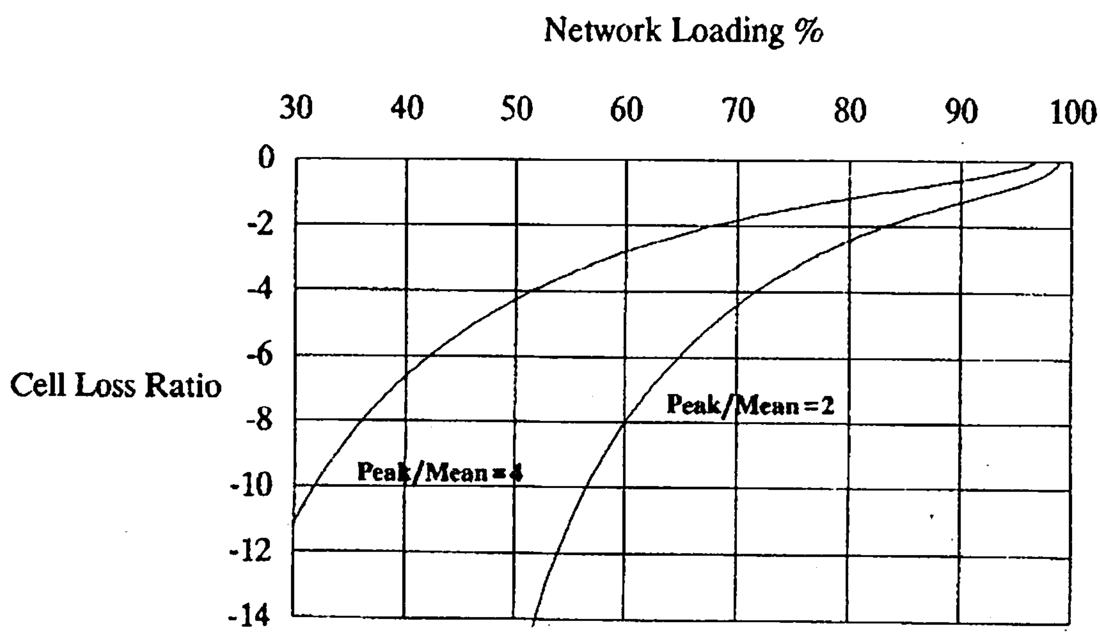
Peak = The maximum bit rate over simulation interval (measured on a per frame basis)

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

Mean = The average bit rate over the simulation interval

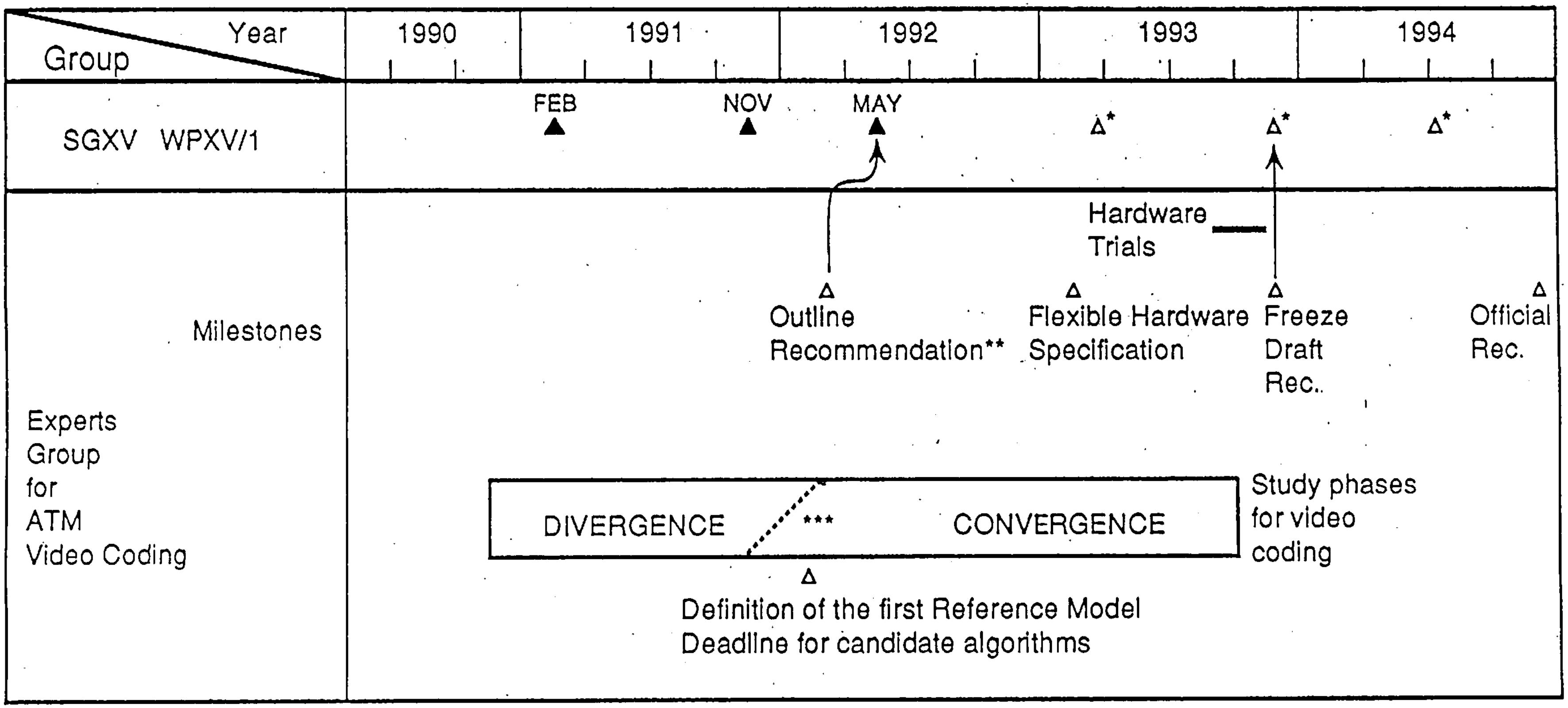
CLR = Psat / (n * Mean * ln ((a * (1-p))/(p * (1-a))))
where CLR is the cell loss ratio.

ln = log to the base 'e'



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.