New Network and ATM Adaptation Layers for Interactive MPEG-2 Video Communications: A Performance Study Based on Psychophysics

Xavier Garcia Adanez,* Olivier Verscheure

TCOM Laboratory, Swiss Federal Institute of Technology Lausanne

E-mail: {garcia, verscheure}@tcom.epfl.ch

In this paper, we present new Network and ATM Adaptation Layers for interactive MPEG-2 video communications. These layers provide reliable transmission by applying per-cell sequence numbering combined with a selective Forward Error Correction (FEC) mechanism based on Burst Erasure codes. We compare the performance of the proposed scheme with a transmission over AAL5 by simulating the transport of an MPEG-2 sequence over an ATM network. Performance is measured in terms of Cell Loss Ratio (CLR) and user perceived quality. The proposed layers achieve significant improvements on the cell loss figures obtained for AAL5 under the same traffic conditions. To evaluate the impact of cell losses at the user level, we apply a perceptual quality metric to the decoded MPEG-2 sequences. According to the computational metric and subjective rating, the proposed multimedia AAL (MAAL) achieves a graceful quality degradation. The application of a selective FEC achieves an even smoother image quality degradation with a low overhead.

Keywords: ATM Adaptation Layer, Network Adaptation Layer, Forward Error Correction, VBR, MPEG-2, Perceptual Quality, Interactive Multimedia, Statistical Multiplexing.

1. Introduction

For the past several years, a tremendous number of audiovisual services has been emerging (e.g., Video on Demand (VOD), Interactive Distance Learning (IDL), home shopping, etc...). Today, Asynchronous Transfer Mode (ATM) technology, efficient compression techniques and other developments in telecommunications make it possible to offer such services. However, a lot of work remains to be done to optimize these services so that they can be offered at attractive prices (the user expects an adequate perceived quality at the lowest possible cost).

Source coding of audiovisual information (i.e., the MPEG-2 encoding scheme) nat-

* corresponding author

urally generates variable bit rate since the entropy of an audiovisual sequence may vary in time. Such an encoding scheme produces near *constant quality* compressed video streams. However, the majority of the currently available video applications and services deliver constant bit rate encoded video thus achieving a variable quality stream which consequently affects the perceived quality.

ATM, the adopted solution for the Broadband Integrated Services Digital Network (B-ISDN), has the advantage of handling constant as well as variable bit rate communications. Though CBR connections are easier to manage, VBR communications may use the network bandwidth more efficiently through statistical multiplexing [1,2]. Since VBR video has a large peak-to-mean ratio it is *economically* interesting to apply statistical multiplexing to such connections. However, one of the major drawbacks with such statistically multiplexed connections is that the requested QoS can only be guaranteed on a statistical basis and therefore some non-negligible cell and frame loss may occur [3,4]. Since audiovisual applications tolerate some, yet limited, loss it is possible to achieve a better network utilization without noticeable quality degradation.

To overcome this cell loss problem the research community has proposed several techniques.

Layered coding mechanisms have been proposed to take advantage of the scalability of video coders, mainly H.261 [5] and MPEG-2 [6,7]. The syntactic relevance of the different type of pictures has also been used to develop a layered transmission model in [8]. Different QoS target values are applied to each of the flows according to the picture relevance.

However, using layered coding introduces the problem of synchronization between flows which has proven to be difficult to solve for VBR communications.

Error correction techniques have also been developed to reduce the impact of data loss onto video. Traditional closed-loop techniques found in data communications are generally not considered due to the delay generated by retransmission which is incompatible with real-time application requirements.

Instead, open-loop mechanisms such as Forward Error Correction (FEC) are preferred. FEC is able to handle *isochronous* flows, it is well adapted to high bandwidth-delay product networks such as ATM and it generates low delay. The drawbacks of FEC techniques are the overhead required by the redundancy data and the fact that error-free transmission are not guaranteed. Open-loop techniques also avoid the feedback implosion effect observed in multipoint communications, such as the one depicted in Fig. 1, which make use of feedback based correction techniques.

Early studies focusing on the efficiency of FEC over ATM for VBR communications [9,10] showed that FEC efficiency is largely dependent on the cell loss process.

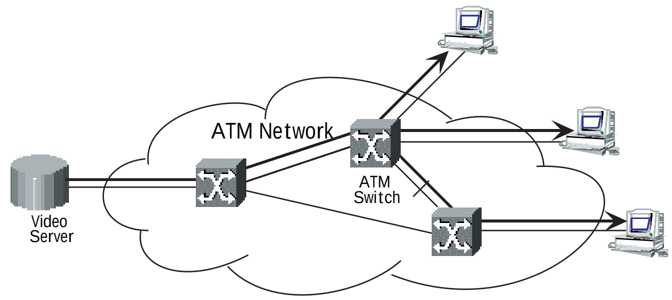

Figure 1. Point-to-Multipoint configuration example

Correlated or bursty losses reduce the efficiency of FEC recovery.

With the increasing interest for VBR video, several contributions proposed FEC mechanisms for ATM networks [11–14]. The proposed AALs where generally enhanced versions of AAL1 which already proposes a FEC mechanism but for CBR services. More recently, within the ATM Forum several contributions proposed the FEC to be performed at the cell-level. Since AAL1 was considered not adapted for the transport of MPEG-2 video, due to the delay and overhead introduced, new AAL mechanisms where proposed [15,16]. The standardization of such a new AAL was however considered as unnecessary by the ATM Forum and AAL5 was chosen instead for the transmission of CBR video services [17]. On this basis, SSCS-FEC mechanisms have been proposed [18,19]. However this approach still relies on a fixed matrix interleaver mechanism which has the problem of introducing variable delay and jitter for VBR services.

Recently a hybrid ARQ-FEC technique has been proposed in [20]. The principle is to offer a reliable service via FEC. Whenever FEC is unable to recover all the lost data, a retransmission mechanism is used, over a high priority ATM virtual channel (VC), to guarantee error free communications. However, this approach still relies on a cell interleaver and on retransmission which in large network configurations will generate too much delay.

Error concealment algorithms have already shown that it is possible to reduce the impact of data loss on the visual information [21–23]. These error concealment algorithms include, for example, spatial interpolation, temporal interpolation and early resynchronization techniques. Early resynchronization decoding techniques limit the spatial propagation of errors by decoding some semantic information that is normally discarded from the damaged MPEG-2 video streams. This method is based on the identification of allowed codewords as proposed in [22] and works only with intra-coded frames.

What stands out from these proposals is that two approaches have been considered up to now:

- The application based approach considers the network as a lossy channel and tries to achieve reliability by including as many error correction and concealment mechanisms at the applications level as possible.
- The network based approach tries to achieve reliability by adding error correction functions at the network level, regardless of the type of data to be transmitted.

This paper addresses the transmission of MPEG-2 VBR video streams on top of new Network Adaptation and Multimedia ATM Adaptation Layers. This *integrated* approach proposes two protocol layers which provide a reliable transmission over ATM under cell loss conditions. We study the performance of the new layers from the network perspective via the cell loss ratio (CLR) and packet loss ratio (PLR) metrics. Since the mapping of data loss onto user perceived quality heavily depends on the type and location of the lost information, we also evaluate performance with a video quality metric. This metric has the advantage of giving results correlated with human perception [24].

We discuss in the next section what the requirements of multimedia communication are and what are the available network services for such applications. Section 3 presents the metrics available to perceptually evaluate video information. The new protocol layers are described in Sec. 4. A performance evaluation of the proposed layers is presented in Sec. 5. Some conclusions are presented in the last section of this paper.

2. Requirements of Multimedia Telecommunication Services

2.1. User Requirements

The user requirements are based on media perception. The most important *perceptual channels* of information for humans with respect to current multimedia applications are the ear and the eye. However, if taken separately, each one has a certain degree of error tolerance. Humans are rather intolerant to audio errors. Conversely, we are quite tolerant to visual errors mainly due to the high spatial and temporal redundancy inherent to video sequences.

The integration of different media is the characteristic of multimedia applications. This integration entails temporal relations between the different media objects. The respect of this timing dependency is called *intra-synchronization*. How the user perceives the loss of synchronization depends on the application and on the media. The lip synchronization or *skew effect* between audio and associated video has been the object of several studies (e.g. [25]). However, there is no exact rule to formulate what the constraints are because the skew effect may depend on the scene contents. Still, as a general rule $\pm$ 80 ms is considered as the threshold beyond which the user notices an out of synchronization effect [26].

Another factor which may influence our perception of the quality is the nature of the application. We distinguish two classes of real-time applications:

- Non-interactive applications, such as VoD, which have strict real-time requirements but can tolerate an initial delay in the playout of the data that will not be noticed by the end-user. These applications require enough bandwidth and a bounded delay but low delay values are not strictly necessary.
- Interactive applications, including video conferencing and live interactive broadcasting, which deliver live audio and video and allow for some level of interaction. These applications have the same bandwidth requirements as non-interactive applications but also require low end-to-end delay to provide the desired level of interaction.

In summary, the typical user requirements are a good synchronization between the media and an acceptable perceptual quality for both audio and video information. All these requirements have a strong influence on both application and transmission requirements aspects that are developed in the next sections.

2.2. Application Requirements

Multimedia applications have to fulfill all the user requirements. Basically, applications have to deliver the different media objects in time, synchronized and with as little perceptive errors as possible. However, their performance depend on the network's ability to fulfill the following requirements:

- Deliver data in time (low delay and low delay variation).
- Deliver multiple data flows with a minimum delay variation between the flows (synchronization).
- Ensure a number of errors below a certain threshold defined according to user perception.

To allow the transmission of audiovisual information, multimedia applications make use of compression techniques. Compression algorithms are among the most important enablers of multimedia communications. By removing redundancy, the applications require relatively little bandwidth but simultaneously become much more sensitive to loss.

In an MPEG-2 video stream, quality reduction due to data loss strongly depends on the lost information type. Losses in syntactic data, such as headers and system information, affect the quality differently than losses of semantic data such as pure video

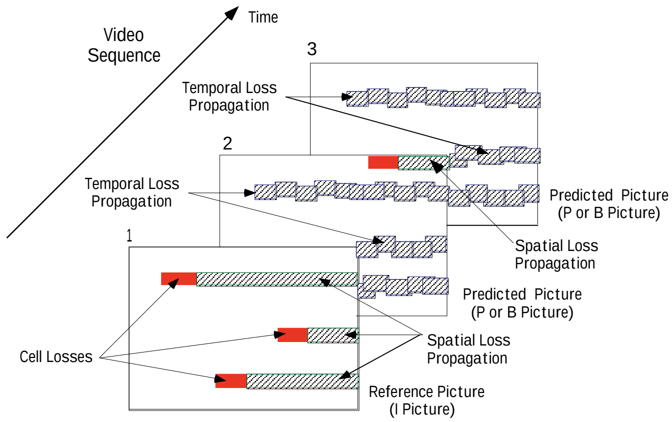

Figure 2. Data loss propagation

(e.g. DCT coefficients, motion vectors, etc. . .) or audio information. Furthermore, the quality degradation depends also on the location of the lost semantic data due to the predictive structure of an MPEG-2 video coded stream.

Let us consider Fig. 2 showing how network losses map onto visual information losses in different types of MPEG-2 pictures. Indeed, data loss spreads within a single picture up to the next resynchronization point (e.g. slice headers) mainly due to the use of variable length coding, run length coding and differential coding. This is referred to as spatial propagation and may damage any type of picture. When loss occurs in a reference picture (intra-coded or predictive frame), it affects the following frames until the next intra-coded picture is received. This causes the errors to propagate across several non intra-coded pictures until the end of the current GOP. This is known as temporal propagation and is due to inter-frame predictions.

The impact the loss of syntactic data may have is in general more important and more difficult to recover than the loss of semantic information. This data loss may induce the whole frame to be lost in the decoded sequence. Indeed, when a frame header (a few syntactic bytes before each frame in the bitstream) is lost, the entire corresponding frame is skipped since the decoder is not able to detect the beginning of the frame. If the skipped frame corresponds to a predictive picture (I or P), it may strongly reduce the perceptual quality due to the predictive structure of the MPEG-2 video stream. Some headers are thus more crucial than others. For instance, sequence headers, predictive (I or P) picture headers, PES headers in some cases, slice headers in intra-coded pictures can be considered as essential in comparison to slice headers in B pictures.

2.3. Network Services

ATM adaptation layers provide specific functions tailored to a set of classes of service. The service classification is based upon three attributes, namely; timing relation between source and destination, bit rate and connections mode. To fit the classes of service early defined in I.362 [27], 5 AALs were defined. Among these AALs, two are available today for the transport of multimedia data: AAL1 and AAL5 [28].

AAL1 was basically designed to cover circuit emulation services. It therefore offers CBR services to the applications. This may be a limitation if constant video quality encoding is used. AAL1 provides a cell level granularity to detect cell losses via a sequence number. It also provides cell loss recovery via Reed-Solomon FEC codes combined with a (124,128) matrix interleaver which is able to correct erasures (cell losses) as well as random errors (impulse noise). However, using interleaving introduces delays proportional to the size of a FEC block at both the sender and the receiver.

AAL5 was developed mainly for data transfer applications but nowadays is considered as a general purpose AAL, partly due to its simplicity and low overhead as illustrated in Fig. 3. This simplicity has, however, some drawbacks for multimedia applications. AAL5 does not provide enough protection against losses due to its original design target. Since such packet-based applications generally rely on robust transport protocols for error detection and correction such as TCP, AAL5 provides a simple *error detection* mechanism able to detect octet errors as well as cell losses. Due to the lack of more sophisticated functionalities, in case of cell losses, AAL5 is unable to know the position of the cells lost inside the PDU and so, no error correction can be applied. Therefore, when cell losses are detected, the packets are discarded. If no retransmission is performed, the packet discard leads to a data loss rate at the application level which is higher than the data loss (cell loss) rate at the network level.

These functions are not adequate for real-time applications for two main reasons. Firstly, a single cell loss causes the loss of several data that could be still used by the decoder with techniques such as early resynchronization [21]. Secondly, since AAL5 does not notify packet discards to the upper layers the decoder may not detect data corruption and may not be able to trigger any error concealment mechanism [29].

Recently, the AAL5 specification has been revised by ITU-T. The updated recommendation [30] now allows the AAL to pass corrupted packets to the upper layers. However, this new feature has its limitations. The most important, is that the only corrupted packets that can be passed are those with an errored CRC check but with the correct length. This automatically avoids passing packets corrupted by cell losses.

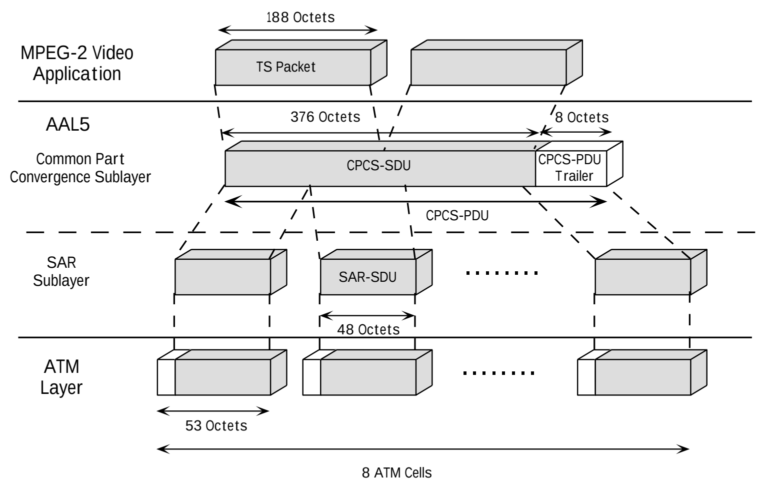

Figure 3. AAL5 segmentation mechanism of MPEG-2 video

Currently, there is no AAL specification to transmit real-time VBR communications over ATM. Originally, AAL2 was defined by the ITU as the adaptation layer for time-constrained VBR communications. However, it has never been specified and recently, this AAL has been devoted to wireless communications. Therefore, due to its ubiquity, to its flexibility to handle both constant as well as variable bit rates and to the lack of a more appropriate AAL, AAL5 has been chosen for the transmission of multimedia applications. The specification describing the transmission of video over ATM is the approved Video on Demand Specification 1.1 [17]. This document describes the encapsulation of MPEG-2 TS packets into AAL5-SDUs. This scheme packetizes two single program transport streams (SPTS) packets regardless of their information contents, being of audio, video or timing nature into a single AAL5-SDU as shown in Fig. 3. The AAL5 adds its 8 byte trailer and the resulting AAL5-PDU is segmented into 8 ATM cells without any padding.

3. User-Oriented Performance Metrics for Video Communication Services

Performance measurement standards generally are addressed to one of two audiences: users or providers. Since telecommunication services exist to fulfill the needs of users, it is important to specify and measure the quality of telecommunication services using performance measurements standards that provide a means for a user to express their satisfaction with the delivered service. Such performance tools are described as "user-oriented".

Traditionally, the quality metric used for audiovisual signals is the Peak Signal to Noise Ratio (PSNR). It is a quantitative measure of the distortion of an image compared to the original defined as:

$$PSNR(dB) = 10 \log_{10} \frac{\sum_{i=1}^{N_p} o_i^2(n)}{\sum_{i=1}^{N_p} (o_i(n) - d_i(n))^2} \tag{1}$$

where $o_i(n)$ and $d_i(n)$ are the luminance values for the i'th pixel of the n'th original and encoded frames, respectively and N_p is the number of pixels in a frame. It has been proved in [24] that the PSNR has no correlation with the user perception of video quality. The linear relationship between the encoding bit rate and the PSNR metric is in contrast with the human inability to perceive a quality improvement beyond a certain encoding bit rate.

An attempt to map network QoS to higher layer QoS is to be found in [31]. The metric called *Glitch* tries to capture the impact of ATM cell losses at a display level for MPEG-2 applications. The problem with this metric is that it does not give any real information of what is perceived by the user. Indeed, a glitch is defined as the interval started by an image partially displayed and ended by an image correctly displayed. The glitch duration and glitch rate are the main metrics. Unfortunately, at a user level, the perception of a glitch depends on the duration but also on the portion of the image that is damaged which is not considered by the glitch metric.

Recent research has addressed the issue of video quality assessment by means of human correlated metrics. One of the first quality metrics, the $\hat{s}$ (SHAT), was developed at the Institute for Telecommunication Science (ITS) in Colorado [32]. The quantitative measure is based upon two quantities, namely, spatial and temporal information, *SI* and *TI* respectively, which tries to map subjective evaluations.

The SHAT metric is a linear combination of three quality impairment measures namely m_1, m_2 and m_3, defined upon the SI and TI metrics as follows:

$$\hat{s} = 4.77 - 0.992 m_1 - 0.272 m_2 - 0.356 m_3. \tag{2}$$

It is however shown in [24] that this metric over estimates the quality in the low bit rate encoding range of MPEG-2.

Several studies have shown that a correct estimation of subjective quality has to incorporate some modeling of the Human Visual System [33]. A spatio-temporal model of human vision has been developed for the assessment of video coding quality [34,24]. The model is based on the following properties of human vision:

- The responses of the neurons in the primary visual cortex are band limited. The human visual system has a collection of mechanisms or detectors (termed "channels") that mediate perception. A channel is characterized by a localization in spatial frequency, spatial orientation and temporal frequency. The responses of the channels are simulated by a three-dimensional filter bank.

- In a first approximation, the channels can be considered to be independent. Perception can thus be predicted channel by channel without interaction.
- Human sensitivity to contrast is a function of both frequency and orientation. The *contrast sensitivity function* (CSF) quantizes this phenomenon by specifying the detection threshold for a stimulus as a function of frequency.
- Visual masking accounts for inter-stimuli interferences. The presence of a background stimulus modifies the perception of a foreground stimulus. Masking corresponds to a modification of the detection threshold of the foreground according to the local contrast of the background.

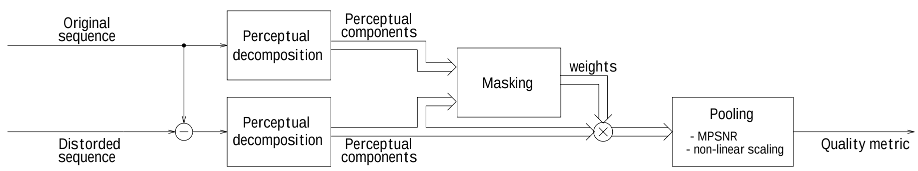

Figure 4. Moving Pictures Quality Metric (MPQM) block diagram

The vision model described in [34] has been used to build a computational quality metric for moving pictures depicted in Fig. 4 [24] which proved to behave consistently with human judgments.The metric, termed Moving Pictures Quality Metric (MPQM), computes the quality according to the rating described in Tbl. 1 [35].

Table 1
Quality scale that is often used for subjective testing in the engineering community

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

Figure 5 presents the MPQM quality assessment of MPEG-2 video for the Mobile & Calendar sequence as a function of the bit rate. An important result that can be extracted

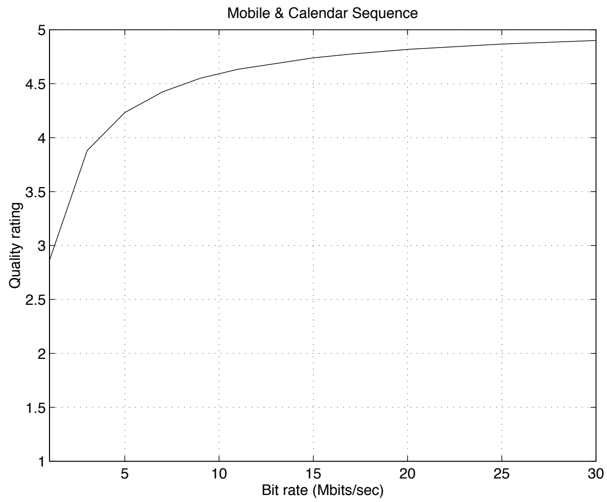

Figure 5. MPQM quality assessment for the Mobile & Calendar sequence as a function of the bit rate

from the graph is that the perceptual quality saturates at high bit rates. Increasing the bit rate may thus result, at some point, in a waste of bandwidth since the end user does not perceive an improvement in quality anymore. Additional compressed video sequences show results with the same behavior in [24].

4. Proposed Layers for Interactive MPEG-2 Video over ATM

4.1. Introduction

The foundations to build the new protocol layers are twofold: the requirements of interactive multimedia applications and the nature of compressed data.

An AAL, by definition, has to be generic, so, the way to take into account the nature of the data without developing an AAL per compression algorithm is to split the required functions into two sets: the network services to be provided by the AAL and application specific services to be provided by a *Network Adaptation Layer* (NAL). Fig. 6 depicts the proposed protocol scheme.

Such a two-layered scheme provides application specific functions while keeping the network part, in this case the AAL, generic. This comes with the advantage that NALs could be used over other network or transport protocols if the underlying layers provide the required functionalities.

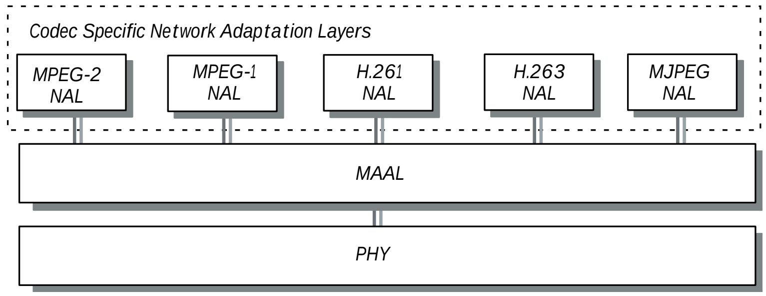

Figure 6. Two layer reference scheme

4.2. A Multimedia ATM Adaptation Layer

The AAL enhances the services provided by the ATM layer to support the requirements of a specific service, for instance the transmission of real-time variable bit rate applications. Multimedia data sources are of diverse nature. Several compression algorithms exist with their own characteristics and syntax. Unlike multimedia, the data in file transfer applications does not have any particular syntax relevant to the network. All information is of equal importance and has to be transmitted without any error to the receiver. This is also the case in today's schemes for multimedia transmission. All data is considered with the same level of importance even if solutions based on a combination of priorities and layered coding tend to prove that this is not the case [5].

However, common characteristics between very different data syntaxes and semantics exist. All the currently available encoding algorithms use a set of headers to structure the information. In addition, every compression algorithm uses prediction which leads to two main types of pictures; reference and predictive.

This drives us to formulate a set of design principles we will use to specify the functions to be performed by the new multimedia AAL, we will refer to as MAAL, which are the following:

- Genericity: the AAL has to support any type of codec.
- Reliability: errors or cell losses have to be minimized.
- Low delay: due to the nature of the data and the requirements of both the user and the application.

The next sections propose a rationale for each of the AAL functions. To derive a first approximation required to evaluate the performance of the chosen functions, we make use of the *low traffic source* assumption [4,36,37]. The utilization of such a model is justified by the fact that high quality multimedia applications using compressed video should not make use of average bandwidths beyond 10 to 15 Mbits/s, even if peak rates could go beyond these values. We therefore assume that the cell loss process seen by this type of connection is independent and identically distributed which allows to develop some analytical models. The simulation experiments show the accuracy of such models.

4.2.1. Cell Loss Detection

There are actually two ways of detecting missing cells: the first one consists of using a packet length information used to monitor the length of the received packet. The second one consists of using sequence numbers used to monitor the continuity of the cell sequence. An advantage of the former packet-oriented mechanism is that it generates low overhead. On the other hand, in case of cell loss this mechanism is unable to pass enough information to allow the higher layers to perform error correction or concealment. It is therefore difficult to take advantage of corrupted information. If, instead, a packet discard scheme is used, the probability exists that the dropped packet contained syntax information that was correctly received as depicted in Fig. 7.

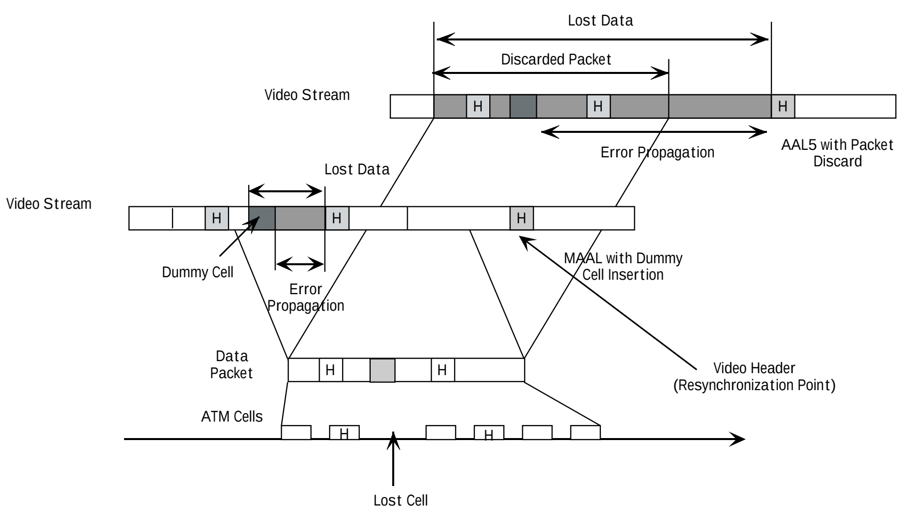

Figure 7. Cell loss propagation in a video stream

Alternatively, a cell based loss detection has several advantages. Among them the fact that a cell loss within a packet will not directly affect the remaining information in the packet because the corrupted packet could be passed to the upper layer with an error indication containing the number as well as the position of the lost cells in the PDU. This could easily be exploited by higher layers to take corrective actions.

On this basis we propose the MAAL to provide a cell sequence number for cell loss and misinsertion detection.

Having selected a cell sequence number for the MAAL, requires that at least an octet is used from the SAR-SDU to carry the information. We proposed here a preliminary study to fix the size of the *SN field* based on the low traffic source assumption.

A field size of n bits is able to number modulo 2^n consecutive cells. Once the maximum value is reached the counter is reset and numbering starts again from zero. This means that the AAL is able to detect up to $2^n - 1$ consecutive losses. To dimension the SN field one has to evaluate the probability of observing n consecutive cell losses.

Assuming an independent and identically distributed (iid) cell loss process, we define the random variable X as the number of consecutive cell losses observed. Then if the cell loss ratio is equal to p the probability of observing n consecutive losses is given by:

$$P(X = n) = p^n. \tag{3}$$

Figure 8 shows the evolution of the probability for n between 1 and 8 bits. We have to consider that the assumption is the best case possible and that the real behavior would lead to higher probabilities of consecutive losses. We therefore chose $n = 5$ leaving three extra free bits that will be used for FEC information.

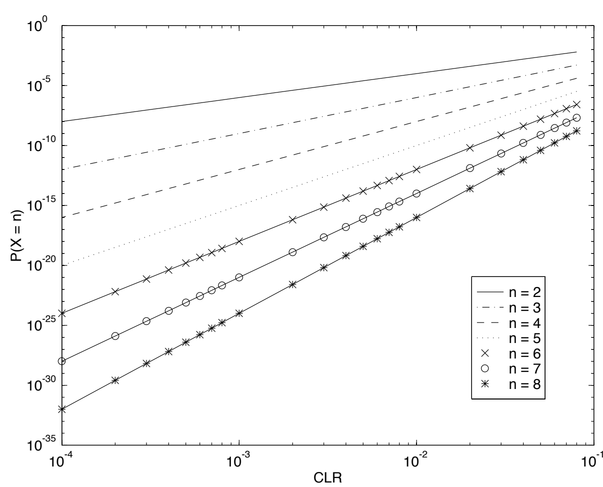

Figure 8. Probability of losing n consecutive cells

4.2.2. The Dummy Cell Insertion Mechanism

Dummy cell insertion consists of inserting *dummy* data to replace the lost data within a packet. This mechanism is used by AAL1 to keep the data aligned in the interleaver to allow recovery of the lost cells.

If interleaving is not used then the only reason to use dummy cell insertion is to keep packet size integrity. It could be useful in the sense that if there is a packet length check at higher layers it could pass through avoiding discard of the information due to a length mismatch.

We see a third possible utilization of dummy cell insertion related to video. Since all compression algorithms use run length coding, it is impossible that long sequences of the same value appear. Therefore the existence of an unallowed codeword could be interpreted as an error by the decoder which could then drive the error concealment mechanisms.

Taking into account that different types of applications could make use of the MAAL, we propose the dummy cell to be selected by the user at connection setup.

4.2.3. Packet Delineation

If variable packets are used, then it is necessary to provide a packet delineation mechanism. The ATM cells have a 2 bit field in the header, the Payload Type Identifier (PTI), whose purpose is to delineate the packets (see Fig. 9). If only the PTI field is used to delineate the packets then the probability exists that the delineation cells are lost leading to extra packet loss.

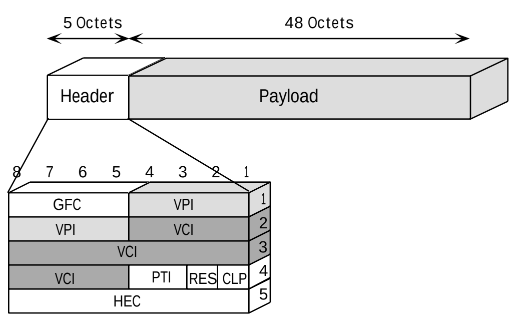

Figure 9. ATM cell structure

A first evaluation could be done by calculating the probability that this happens assuming an iid cell loss process. Given a cell loss ratio of p then the probability of losing the first or the last cell in a packet is the same and equal to the probability of losing any

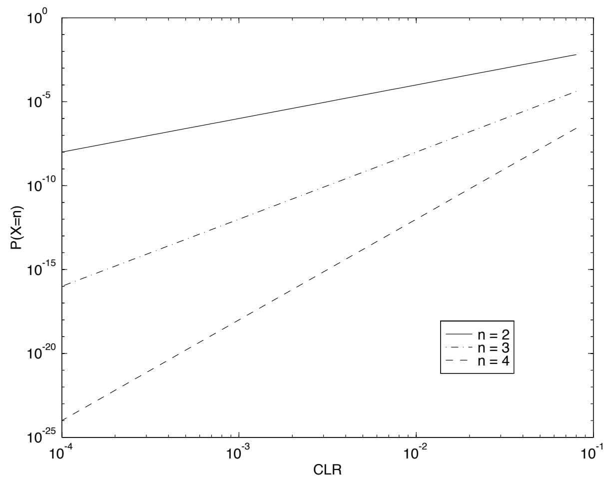

Figure 10. Probability of consecutive packet losses due to cell delineation loss

other cell. Thus, if we define $P(A)$ and $P(B)$ as being the probabilities of losing the first and last cells of any packet respectively, the probability of losing the delineation cells for packets of size m is equal to:

$$P(A \cap B) = P(A) \times P(B)$$
$$P(A \cap B) = (p)^2$$

and does not depend on m. The probability X of observing n consecutive packets, and therefore $n \times m$ cells, lost due to loss of delineation cells is given by:

$$Prob(X = n) = \left(p^2\right)^{(n-1)} . \tag{4}$$

As Fig 10 shows, the probability of multiple consecutive packet losses due to loss of delineation packets is small. This means that under the low traffic sources assumption, the utilization of the PTI fields in the ATM cells should be enough to reliably delineate and transmit variable size packets. Note that AAL5 uses only the last cell of the packet to delineate the PDUs. In this case the probability will be:

$$Prob(X = n) = p^{(n-1)} \tag{5}$$

which leads to a higher probability of packet loss due to delineation cell losses.

The disadvantage of transmitting variable size packets is that they do not necessary segment into an integer number of SAR-SDU payloads. Therefore, the CPCS has to provide a boundary alignment function which requires a CPCS-SDU header or trailer to indicate the number of extra bytes transmitted, thus adding overhead. We therefore propose that in addition to using the PTI field the size of the PDUs is kept constant for the duration of the connection. The size will be selected by the user at connection setup.

4.2.4. Packet Level Forward Error Correction

Working on top of AAL5 gives a packet visibility to the error correction mechanism. If we consider the protection scheme of Fig. 11, we are protecting k packets with h overhead packets. To guarantee no loss it is necessary to receive k out of $k+h$ packets.

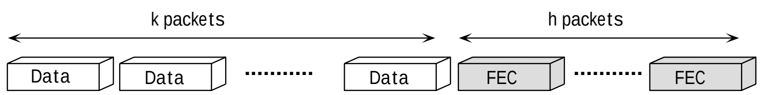

Figure 11. RSE based forward error correction mechanism

To prove the error recovery efficiency reduction for the packet level scheme lets assume an iid cell loss pattern. Each packet segments into m cells. We want to calculate the probability of receiving at least k among $k+h$ correct packets.

The probability p of losing a packet is given by the probability of losing at least one and at most all the cells in a packet if we consider that AAL5 is used. We want to calculate the number of successful events (cell losses) of a Bernoulli process. Thus the probability of losing a packet is given by a binomial distribution:

$$p = \sum_{n=1}^{m} \binom{m}{n} CLR^n \times (1 - CLR)^{m-n}. \tag{6}$$

As Fig. 12 illustrates, AAL5 amplifies the CLR by a factor equal to m mainly due to the *packet discard* mechanism.

It is worth to note that a packet based recovery mechanism will not benefit from I.363.5 recent revision [30]. The new recommendation allows AAL5 to pass corrupted packets that are correct in length (CRC mismatch). This is incompatible with error recovery techniques because using corrupted packets to recover missing packets will lead to wrong results.

If the cell loss process is Bernoulli, then the packet loss process is binomial with a packet loss probability p given by Eq. 6. If we define the random variable X as being the number of received packets, the probability of receiving k among $k+h$ packets is again given by a binomial distribution:

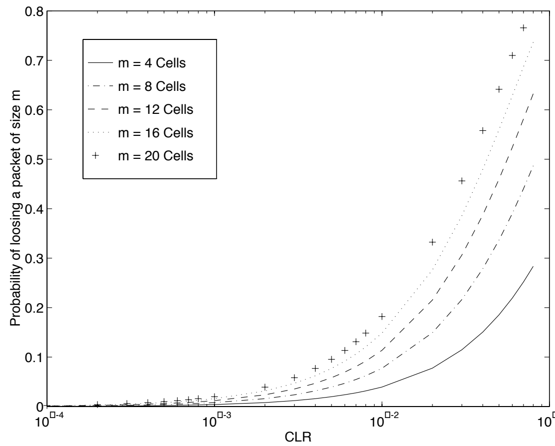

Figure 12. Probability of losing a packet of size m cells

$$Prob\{X \geq k\} = \sum_{n=k}^{k+h} \binom{(k+h)}{n} (1-p)^n \times p^{k+h-n}. \tag{7}$$

Figure 13 shows the probability of receiving k among $k+h$ packets correctly based on Eq. 7. We have fixed $k = 8$ packets and $h = 1$ packet. The packet size has been varied between 4 and 20 cells. This protection scheme tolerates a single packet loss regardless of the number of cells lost *within the packet*. This scheme is therefore sensitive to the packet size as well as to the distribution of the cell losses. For a given CLR the probability of losing a packet increases according to Eq 6. Thus, the recovery efficiency is reduced by a factor m. With a bursty loss process, which increases the probability of consecutive losses, such a mechanism would perform better than under the iid assumption.

4.2.5. Cell Level Forward Error Correction

If we now consider that the FEC data is generated at the cell level, in this case the SAR sublayer, to achieve the same overhead as in the packet level protection case we have to send $k \times m$ cells protected by $h \times m$ FEC cells. Then, using the same random variable as in Eq. 7, the probability of receiving $k \times m$ cells among $(k+h) \times m$ is given by:

$$Prob\{X \geq (k \times m)\} = \sum_{n=mk}^{m(k+h)} \binom{m(k+h)}{n} (1-CLR)^n \times CLR^{m(k+h)-n}. \tag{8}$$

Figure 14 shows the probability of receiving k among $k+h$ packets correctly for the cell level recovery mechanism according to Eq. 8. We keep the same k, h and m

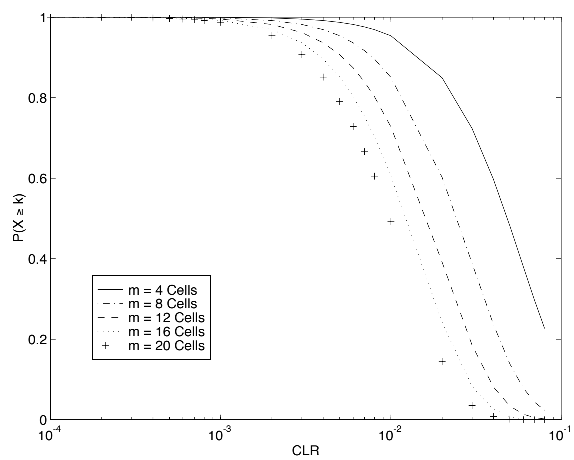

Figure 13. Probability of receiving k among $k + h$ packets with packet level protection

parameters of the packet case. In the cell level scheme, all the redundancy cells protect all the data cells and therefore recovery depends only on the number of lost cells and not on the distribution within the *FEC block* which explains the improved efficiency compared to Fig. 13. Note that increasing the packet size *increases* the recovery efficiency of the cell level mechanism. This is due to the fact that for a given CLR, the probability of losing m cells is a function of the CLR while the probability of receiving $m \times k$ cells is a function of Eq. 8.

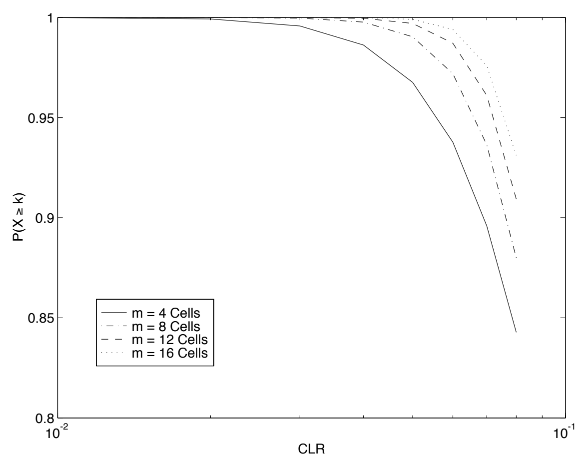

Figure 14. Probability of receiving k among $k + h$ packets with cell level protection

4.2.6. Comparative Performance

Figures 13 and 14 show that the cell level protection mechanism performs more than an order of magnitude better than the packet level scheme for the same *protection overhead*. The reasons to this are that the recovery efficiency of the packet level protection scheme directly depends on the cell loss process and on the packet size while the recovery efficiency of the cell level protection scheme depends only on the number of redundancy cells.

Figure 15 (a) shows a comparative plot of the recovery efficiency achieved for both packet and cell level schemes. We have fixed for the packet level scheme $k = 8$ and $h = 1$ packets and for the cell level scheme, $k = m \times 8$ and $h = 1$ cells.

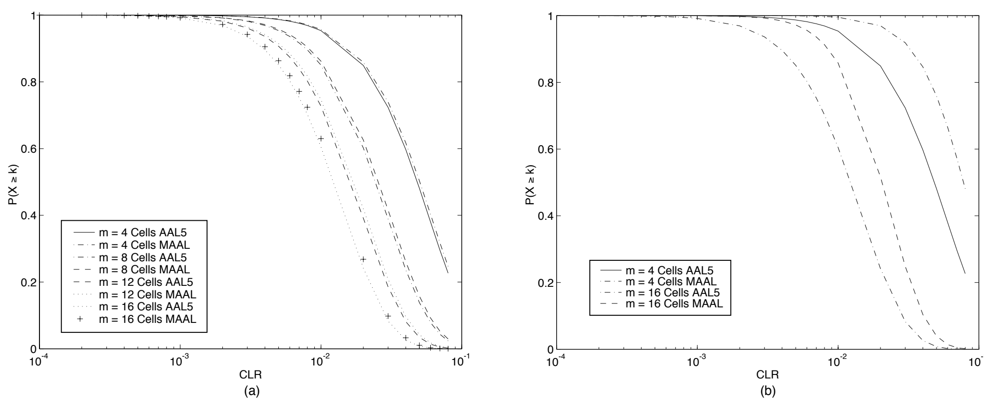

Figure 15. Comparative packet and cell based probability of receiving k among $k + h$ packets

The cell level mechanism achieves the *same recovery efficiency* as the packet level with an overhead of $\frac{1}{m}$ *less* than for the packet based case for all the packet sizes considered.

Figure 15 (b) shows again a comparative plot for both protection mechanisms. In this case the cell level overhead is $\frac{2}{m}$.

In summary, and under the uniformly distributed cell loss process assumption derived from the low traffic source hypothesis, a cell level protection mechanism achieves a much better performance than a packet level one. We therefore propose for the MAAL a mechanism based on burst erasure codes (RSE) [38]. The advantage of this method is that it takes into account the specifics of ATM as depicted in Fig. 11. It relies on the fact that erasures are limited to fixed boundaries (cells) to correct cell losses only. Conversely, it cannot correct octet or bit errors (impulse noise). This is not a fundamental problem since we assume that the ratio of errors to erasures in fiber based networks is very small and can be considered as negligible. Also, since RSE does not need interleaving, the delay introduced is small (see Fig 16).

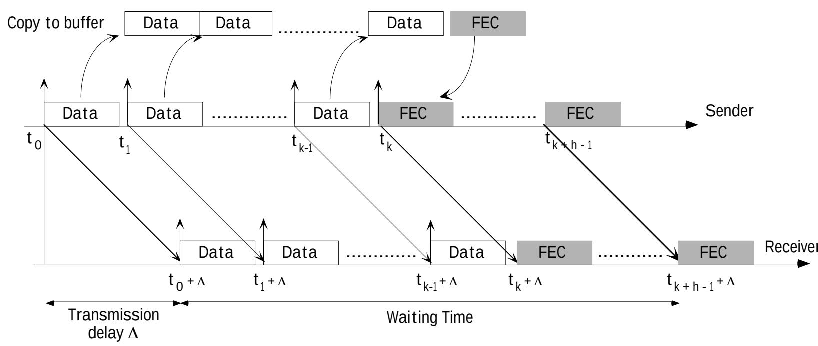

Figure 16. RSE based forward error correction end-to-end delay

4.3. MAAL Proposed Functions

To meet the requirements of real-time multimedia applications, we propose an AAL that includes the following mechanisms:

- Transmission of constant size packets.
- Cell sequence numbering.
- Passing of corrupted PDUs to the upper layers with an error notification.
- User selected dummy cell insertion mechanism depending on the type of application or coder.
- Selective RSE based Forward Error Correction.

The structure of the MAAL SAR-PDU is shown in Fig. 17. The sequence number provides a cell level granularity to improve the error detection capabilities and combined with dummy cell insertion allows to pass incomplete packets to the upper layer. The RSE based FEC selectively provides protection of essential information thus achieving a low overhead. This functions require a SAR-SDU header of one octet which adds a per cell overhead of 2%. However, since the PTI field is used for packet delineation no overhead is added at the convergence sublayer.

Concerning the packetization of MPEG-2 TS packets, the proposed packetization has the advantage of always giving an integer number of cells, since the length of the TS packets is 188 bytes which gives exactly 4×47 byte payloads.

In summary, the structure of the proposed MAAL, depicted in Fig. 17 will be as follows:

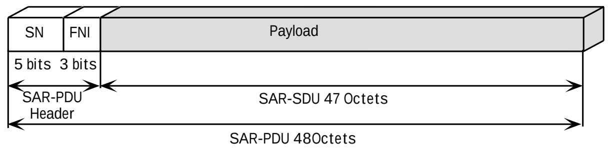

Figure 17. MAAL cell structure

- SAR-SDU size of 47 octets.
- SAR-SDU header of one octet divided into two fields: a sequence number field and a FEC number field.
- Selective SAR-level FEC.
- Null SSCS and CPCS.

Such an AAL fulfills the proposed design principles. The MAAL is generic since there is no application specific function included, even if some functions are user configurable. It is reliable because it provides an efficient cell loss detection mechanism in addition to a cell level FEC scheme. Finally, by deliberately limiting the AAL functions, we keep processing delays to small values thus leading to low delays.

4.4. An MPEG-2 Specific Network Adaptation Layer

One of the major arguments against FEC is that it adds overhead that may contribute to the network congestion it tries to solve. We have seen that FEC still presents several advantages for multimedia communications such as low delay and sender independence for multipoint. As we have discussed in Sec. 2.2 compressed audiovisual information can be classified according to the perceptual importance that its loss has onto video. It could therefore be possible to take advantage of such perceptual relevance to reduce the overhead of the protection mechanism by providing a selective FEC.

To achieve such protection scheme it is necessary to have an a priori knowledge of the type of data to be transmitted to identify the sensitive information. This requires a layer, in our case we propose a *Network Adaptation Layer*, to be codec specific.

The *Perceptual Syntactic Information Protection* (PSIP) takes into account both the hierarchical structure of compressed audiovisual information as well as its perceptual relevance.

The PSIP algorithm takes into account both the picture and the syntax levels of audiovisual bitstreams. Firstly, a picture classification is done according to their perceptual relevance (reference or predictive frames). Secondly, a header classification is established according to the hierarchy, and therefore the perceptual relevance, within the bitstream. Headers being included into pictures, a complete loss impact classification

Table 2
Example of loss impact classification of MPEG-2 data

	Frame Type		
	I (1)	P (2)	B (3)
Picture Header (1)	1	2	3
Slice Header (2)	2	4	6
TS Header (3)	3	6	9

could be done using a two-entry table as shown in Tbl. 2. If we set priority values for the 3 picture types as well as for the different headers, a hierarchy could be obtained by multiplying the values of the joint picture-header combination. The lowest value being the highest priority.

This classification does not include the GOP header, actually optional in MPEG-2, and the Sequence Header which both encompass the pictures. Clearly these have an even higher priority, say priority 0.

From this table it is easy to achieve a protection scheme on the basis of priority values. This mechanism does not allow for all combinations of protection but gives a very simple priority-based protection algorithm. This type of classification could be further improved by adding the MPEG audio part. Several experiments done with interactive audiovisual transmission [39] have shown that the users are much more sensitive to audio degradation than video. It is therefore natural to consider the audio information as a high priority flow. Since the audio flows require relatively low bandwidth compared to video the protection of the audio flow generates a low overhead and can therefore be set to priority 0.

The PSIP mechanism is simple to implement does not require much processing time and is easy to apply to any kind of hierarchically structured audiovisual information.

Due to its flexibility, the PSIP could be improved in many ways. It could be possible to achieve an adaptive PSIP (given that a feedback channel exists) that would adapt the level of protection according to network conditions. For example the PSIP could switch from a level 3 to two protection in case of low CLR.

Also a multi-level PSIP which would generate a given number of FEC cells according to the priority level could be developed. As an example, all picture headers could be protected with a larger number of FEC cells than, say the slice or PES headers, regardless of the type of frame. This will reduce the probability of losing complete frames due to a frame header lost.

If layered coding is used it is also possible to apply priority levels by layer or by type of picture regardless of the layer. It is actually up to the implementer to develop its own protection algorithm.

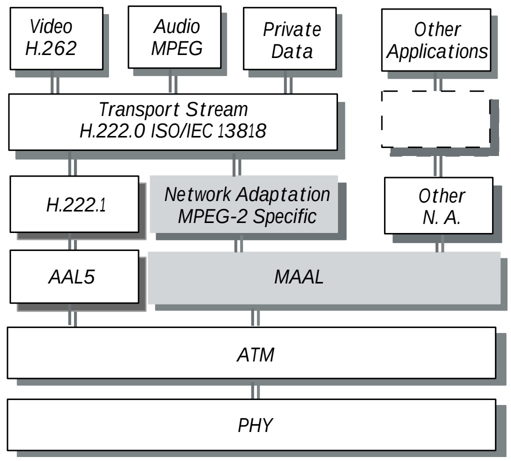

Figure 18. ATM Forum and proposed frameworks for transmission of MPEG-2 video over ATM

In summary, we propose a two-layered framework that splits the functions required for multimedia into two sets: network services provided by the MAAL and application specific services to be provided by the NAL. Fig 18 depicts the current framework in white. The grey layers show how the proposed protocol layers seamlessly integrate into the existing protocol stack. Such scheme provides codec specific functions while keeping the network part, the AAL, generic.

5. Performance Evaluation

5.1. Experimental Framework

5.1.1. Network Setup

To evaluate the performance of the chosen mechanisms for the MAAL, we have performed a set of simulations based on the ATM network simulator setup depicted in Fig. 19. The simulator is composed of four multimedia workstations and two ATM switches. The workstations are connected as two point-to-point communications. Both switching stages, implemented as multiplexers with limited buffer size, are loaded with several On-Off background traffic sources widely used to simulate a multiplex of traffic.

To guarantee the same CLRs to both cell streams, the background traffic is replicated and sent simultaneously to both multiplexing stages. To avoid correlations between different sources, random phases are used at simulation startup. Also, to avoid transient phenomena, the traffic under test (TUT) is not sent until the network reaches steady state. The TUT is then sent to the switch where it is multiplexed with the background traffic. Since the switch buffer is limited in size, some cells may be lost. The TUT is then routed to the receiver end system where the data is reassembled. The background traffic is assumed not to interfere further with the TUT and thus is directly routed to a traffic sink after leaving the switching stage.

In all our experiments we compare the proposed MAAL against an equivalent transmission over AAL5.

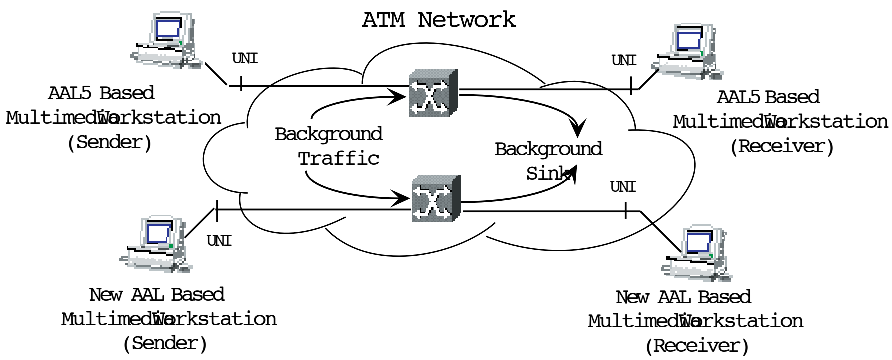

Figure 19. Simulation scenario

5.1.2. Video Setup

The foreground traffic used for our experiments has been obtained by encoding real video sequences. By doing this we avoid the accuracy problems due to mathematical models of VBR video sources. In addition, to be able to use a video quality metric, we needed real bit streams.

The main video sequence used to generate the TUT consists of a ski sequence of 1000 frames of ITU-R 601 format ($720 \times 576, 25fps$). We have generated VBR video streams with a TM5 MPEG-2 software encoder as interlaced video, with a structure of 12 images per GOP and 2 B-pictures between every reference picture, and a single slice per line (i.e. 45 macroblocs per slice). The sequence provides several scene changes as well as fast motion. Figure 20 shows a frame of the ski sequence.

In our experiments, we have used the encoder in open-loop with a constant quantizer scale. Two different quantizer scales have been considered, 26 and 28 which both produce a CATV quality, in such a way that 10% more bandwidth in average is required by the former. The saved bandwidth will be used for data protection (PSIP). These two quantizers imply different perceptual video encoding qualities as well as traffic characteristics (see Fig. 25).

Before being transmitted, the MPEG-2 bitstream is encapsulated into 18800-bytes

Figure 20. Sample of the original video sequence used for transmission

length Packetized Elementary Stream (PES) packets and divided into fixed length Transport Stream (TS) packets by the MPEG-2 system encoder.

The decoder provides some error concealment techniques. These techniques have been used for different reasons. The first is to be consistent with real implementations and the second one to be able to perform the perceptual measurements. Indeed, the human visual system models currently developed and the derived metrics have been tested for errors below what is called the *suprathreshold*. The problem is that in general, the errors due to cell losses generate highly visible artifacts, holes, in the sequence and these errors are all above this suprathreshold. By using concealment techniques, most of the artifacts may be considered as being below the suprathreshold of vision, making the perceptual measure accurate [24].

5.2. Experimental Results

5.2.1. Provider-Oriented Performance Evaluation

In the precedent section we have used the low traffic source assumption [4,36,37] to dimension the AAL and the FEC mechanism. This assumption has been studied under specific traffic profiles. There is no actual proof of its applicability to VBR video traffic. Indeed, the utilization of an open-loop VBR encoder leads to a bursty traffic profile. It is therefore to be expected that the cell loss process will also show some burstiness. This should entail a reduced FEC recovery efficiency as shown in [9].

We first study the behavior of the cell loss process for the global cell stream as well as for the single VBR video stream. We perform this by means of signal processing techniques. We sample in our simulations simultaneously both the global and the TUT cell loss processes. We obtain a random on-off process of the cell loss occurrences. We derive from these traces the Power Spectral Density (PSD) functions $\Phi(f)$. Since the inverse Fourier Transform of a stationary random process is the autocorrelation function we can derive from the spectra the main characteristics of the loss process.

Figure 21 depicts the Power Spectral Density (PSD) of the global process as a function of the network load. The peak at $\varphi(0)$ depicts the total power of the process. The amplitude of $\varphi(0)$ indicates large fluctuations around the mean value and suggests a correlated process. This is to be expected because during congestion periods, consecutive cells are dropped by the switches. Note also that when the traffic intensity decreases the peak also decreases and some noise appears in the curves. This is due to the relative low frequency of the occurrences. Still with increasing background loads, the spectra tend to precise figures. This same behavior is observed in the other PSD figures.

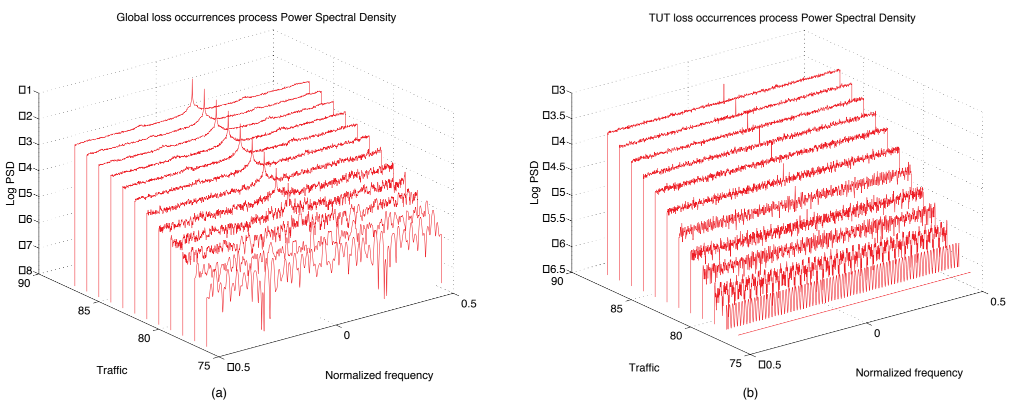

Figure 21. Power spectral density of the global and video cell loss processes

In contrast, the TUT PSD shown in Fig. 21 (b) suggests a non correlated process. The PSD of an ideal uniformly distributed process is flat. The spectra obtained from the simulations shows a peak at the null frequency $\varphi(0)$ which actually depicts the total power of the signal. In fact the spectrum could not be totally flat because the signal's energy is finite which explains the amplitude at $\varphi(0)$. We can therefore, deduce that the cell loss process of a *single connection* tends to be non-correlated. This could be explained by the fact that as suggested in [36], if the fraction of the link bandwidth used by a connection is small enough then the cell loss process can be accurately approximated by a uniformly distributed process.

The correlated characteristic of the global cell loss process is corroborated by the plots of Fig. 22 (a) which shows a detailed view of the PSD and the autocovariance functions for a fixed load. The lag plot shows that beyond 50 cell slots the process is still correlated which suggests relatively long bursts of cell losses.

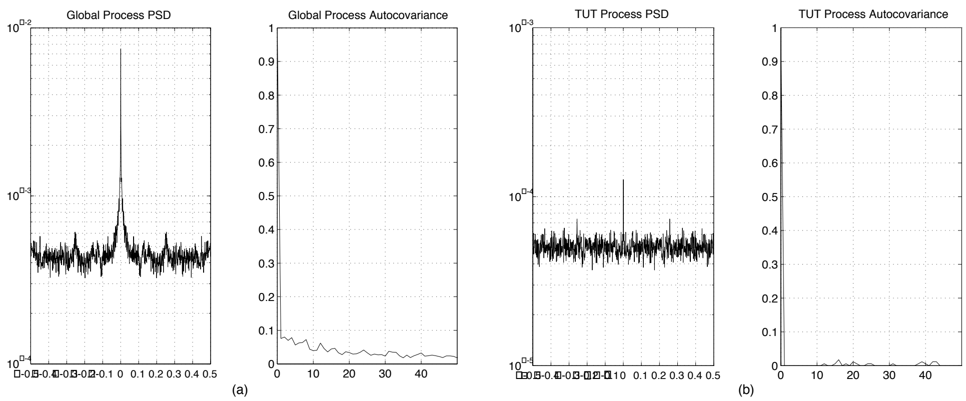

Figure 22. Power spectral density and normalized autocovariance function of the global and video cell loss processes for a background load of 80%

As depicted in Fig. 22 (b), the lag plot of the TUT loss process shows a very light correlation in a relative long range of up to 40 cell slots. This is due to the burstiness of VBR video. When high activity periods occur, the cells are sent at cell rates high enough to increase the probability of observing consecutive losses. Actually, the cells are close enough in time to observe such cell loss clustering effect. However, the relative low frequency of these peaks results in an almost negligible cell loss correlation.

In addition, if we now compare the Packet Loss Ratio (PLR) observed for AAL5 and we plot it as a function of the CLR (Fig 23), it turns out that the simulated figure follows very closely the theory given by Eq. 6. We can therefore conclude that at the packet scale the loss process behaves close to a uniformly distributed process. This is actually the worst case for a packet-oriented AAL that makes use of a packet discard mechanism.

The presented figures corroborate the low traffic source assumption for a VBR video connection using, in average, a small fraction of the link rate under an On-Off background traffic profile. Certainly increasing the burstiness of the multimedia stream or the average bandwidth used will increase the correlation factor. However, given the very small values observed, the cell loss process of a CATV quality video stream can accurately be approximated by a non-correlated process. This leads to optimal conditions to use a cell-level FEC recovery mechanism.

Figure 24 shows a comparative plot of the CLRs measured at the receiver, for both AAL5 and MAAL with PSIP protection. The AAL 5 curve includes all the data lost due to the packet discard mechanism and the MAAL the FEC data lost.

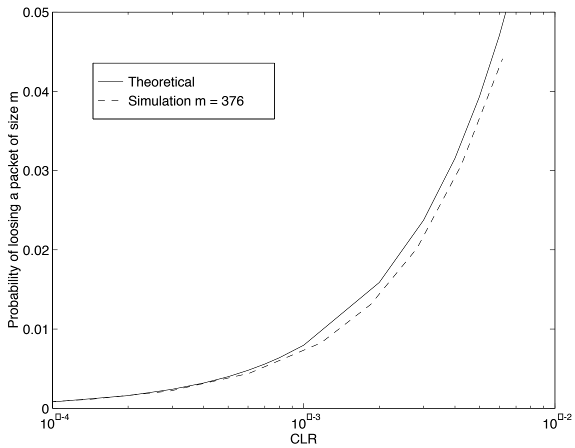

Figure 23. Packet Loss Ratio vs. Cell Loss Ratio for open loop VBR video traffic

The CLR measured for the MAAL is actually equal to the network CLR since there is no extra discard of information. Comparatively, the CLRs achieved by AAL5 are 4 times higher mainly due to the packet discard mechanism which clearly amplifies the CLR seen by the application. The size of the packets used in the simulation is of 8 ATM cells. The fact that the CLR difference between both AALs is not closer to the packet size is due to the clustering effect observed in the cell loss process which reduces the negative impact of the packet discard mechanism. In addition the PSIP overhead has been calculated to achieve *in average* the same bitrate as the unprotected bitstream. This means that in high activity periods FEC data may have contributed to extra cell loss in the MAAL cased. It is interesting to note that changing the size of the packets to be transmitted does not affect the performance of the MAAL. This is not the case for AAL5 whose performance depends on the packet size all the more that the cell loss process is almost not correlated which is the worst case for a protocol using a packet discard mechanism.

5.2.2. User-Oriented Performance Evaluation

Figures 25 (a) and (b) show the perceptual quality assessment as a function of the CLR for the transmission of MPEG-2 VBR streams over AAL5 and MAAL. Figure 25 (a) compares the quality of the same video stream transmitted with both AALs *without any*

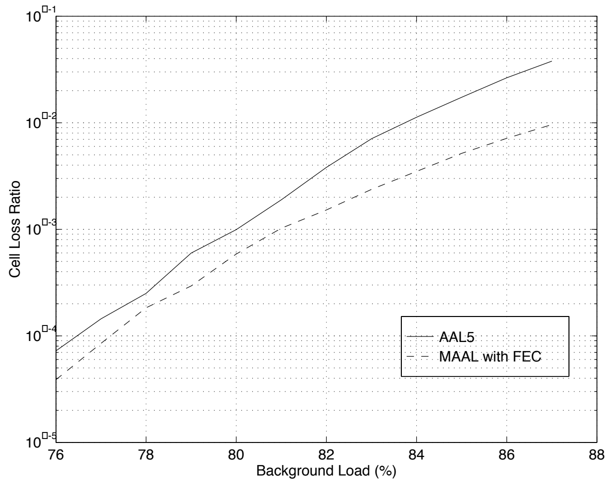

Figure 24. Comparative AAL5 and MAAL Cell Loss Ratios observed for VBR video transmission

data protection. The MAAL shows a better behavior under medium to high CLRs. Note that both the AAL5 and MAAL figures begin to diverge for a CLR value close to 10^{-5}. However, under very high CLRs both figures converge again. This behavior is explained by the higher granularity of the MAAL which allows to pass incomplete data packets to the upper layers. The lack of such possibility in AAL5 penalizes its performance from a perceptual point of view.

The reason why both curves converge for very high cell losses is due to two reasons. Firstly, with increased network load, the cell loss process becomes more bursty and groups the losses into small clusters reducing the negative effect of AAL5's packet discard mechanism. Secondly when large cell loss ratios occur, the pictures are so severely damaged (including lost frames) that the perceived quality drops very fast for both AALs. However, the MAAL proves to be more reliable without any extra overhead in a large range of cell loss ratios.

Figure 25 (b) compares the results obtained for the transmission of both a higher encoded quality unprotected video stream over AAL5 against a lower encoded quality PSIP protected one over MAAL. The figure which gives the best quality is the transmission of the lowest video encoding quality stream with MAAL and PSIP. It is interesting to note that even if the quality of the encoded flow is lower, the quality at the receiver is better for CLR values beyond 10^{-4}. For higher CLR values both figures diverge increasing the difference in quality achieved by the MAAL. Even if we have shown that the burstiness of the cell loss process is limited it is true that the FEC mechanism is penalized under such conditions. Bearing in mind that the protection mechanism adds a single redundancy cell, configuration which is not optimized for such bursty loss pro-

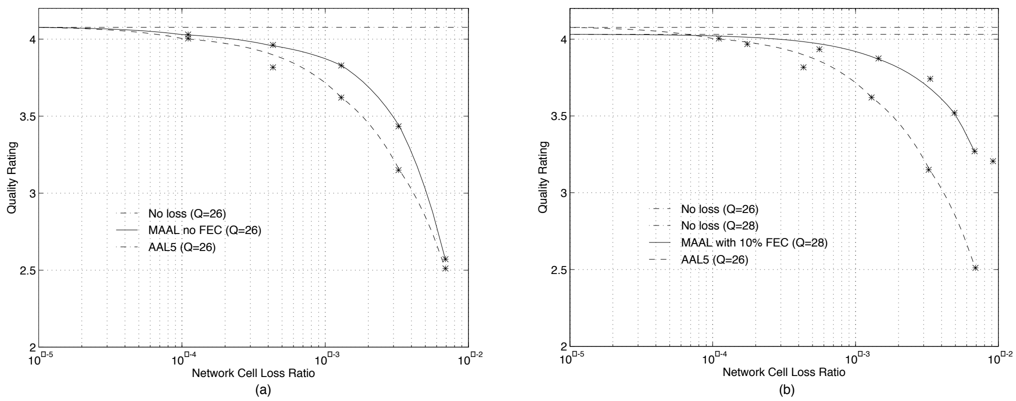

Figure 25. Quality rating vs. network CLR for VBR video transmission over AAL5 and MAAL

cesses, we can conclude that further improvements on the PSIP mechanism could lead to better results. We can then conclude that a tradeoff may be found which optimizes the perceived quality (the bitrate) and the protection efficiency (the overhead).

The probability of observing such high CLRs in ATM networks due only to congestion is relatively reduced. Nevertheless, our results show that under high CLRs the same or even better perceived quality could be achieved with the combined NAL-MAAL layers. This leads to two possible utilizations: the first is the possibility of applying the proposed mechanisms to a wireless environment where such high CLRs are expected to occur. In this case further investigations are necessary concerning the cell loss profiles in such environments. The second one is related to cost. If a network operator applies statistical multiplexing the way the resources are allocated depends, among others, on the cell loss ratio an application tolerates. Since the MAAL with the PSIP mechanism achieves the same perceptual quality for higher CLRs than AAL5 it is possible to multiplex more connections per link thus obtaining a better network resource utilization.

6. Conclusions

We describe in this paper an AAL targeting VBR interactive multimedia services. It provides a cell level granularity view for loss detection and correction. Considering the tight timing constraints that such applications have as well as the characteristics of video which tolerates some loss, a FEC-based correction mechanism is also provided. The AAL functions are kept simple to avoid extra delays due to processing. To overcome the overhead problem due to FEC data, and taking into account the hierarchical structure that characterizes compressed audiovisual information, we also describe the principles of a video-oriented Network Adaptation Layer, NAL. We applied these principles to the design of an MPEG-2 specific NAL. The NAL provides means to selectively protect the most relevant information of the video streams. The utilization of the Perceptual Syntactic Information Protection algorithm, PSIP, keeps the FEC overhead to a minimum value while still achieving efficient data protection.

We have studied the performance of this approach in a simulated environment. The study of the cell loss process shows that if we consider a single cell stream, for instance the video stream, the cell loss process can be accurately approximated by a uniformly distributed process, under certain background traffic conditions. This validates the choice of a cell-level FEC mechanism without interleaving. Comparisons made with the current framework for the transmission of MPEG-2 video over ATM have shown from both the network provider and the user perspectives that under medium to high error rates the proposed MAAL achieves better results with no supplementary overhead. When the MAAL is combined with the PSIP provided by the NAL, our proposal achieves a graceful quality degradation under heavy loss conditions.

Acknowledgements

The authors are grateful to Pascal Frossard. We have appreciated his invaluable assistance in the simulation work and his relevant comments and remarks on the paper.

References

[1] B. Maglaris, Dimitris Anastassiou, Prodip Sen, Gunnar Karlsson and John D. Robbins, Performance Models of Statistical Multiplexing in Packet Video Communications, IEEE Transactions on Communications, 1998, vol. 36, pp. 834-844

[2] George Kesidis, Jean Walrand and Cheng-Shang Chang, Effective Bandwidths for Multiclass Markov Fluids and Other ATM Sources, IEEE/ACM Transactions on Networking, 1993, vol. 4, pp. 424-428

[3] Thomas Stock and Reto Gruenenfelder, Frame Loss vs. Cell Loss in ATM Concentrators and Multiplexing Units, EFOC&N'94, 1994, pp.174-177

[4] Israel Cidon, Roch Guerin, Asad Khamisy and Kumar N. Sivarajan, Cell vs. Message Level Performance in ATM Networks, First International ATM Traffic Expert Symposium, Basel Switzerland, April 1995

[5] Mohammad Ghanbari, An Adapted H.261 Two-Layer Video Codec for ATM Networks, IEEE Transactions on Communications, September 1992, vol. 40

[6] Pramod Pancha and Magda El Zarki, MPEG Coding for Variable Bit Rate Video Transmission, IEEE Communications Magazine, May 1994, vol. 32, pp. 54-66

[7] Wenjun Luo and Magda El Zarki, Adaptive Data Partitioning for MPEG-2 Video Transmission over ATM Based Networks, IEEE Multimedia, 1995

[8] Iain E. G. Richardson and Martyn J. Riley, Intelligent Packetising of MPEG Video Data, Proc. SPIE Visual Communications and Image Processing, Taipei, 1995

[9] Ernst W. Biersack, Performance Evaluation of Forward Error Correction in an ATM Environment, IEEE Journal on Selected Areas in Communications, 1994, vol. 11

[10] Masahisa Kawashima, Analysis on the Effectiveness of Forward Error Correction Code in VBR Video Transmission, Sixth International Workshop on Packet Video, Portland USA, 1994

[11] Hiroshi Ohta and Tokuhiro Kitami, A Cell Loss Recovery Method Using FEC in ATM Networks, IEEE Journal on Selected Areas in Communications, December 1991, vol. 9, pp 1471-1483

[12] Yoshiaki Kato, Kazuhiro Matsuzaki, Tokumichi Murakami and Hideo Ohira, A Study of FEC Methods for the ATM Transmission of MPEG-2 Transport Stream Packets, Sixth International Workshop on Packet Video, Portland USA, 1994

[13] Hidenobu Harasaki and Mitsuharu Yano, Interleaved Forward Error Correction for Variable Bit Rate Video Codec, GLOBECOM, 1994

[14] Thomas Stock and Xavier Garcia Adanez, On the Potentials of Forward Error Correction Mechanisms Applied to Real-Time Services Carried over B-ISDN, International Zurich Seminar, Springer, 1996, pp. 107-118

[15] Mike Nilsson, David Beaumont and Geoff Morrison, An AAL for Class B Services, British Telecom ATM Forum Contribution af94-0780, 1994

[16] Kei-ji Tsunoda, Kumiko Kanai and Hiroshi Esaki, A Reed-Solomon Erasure Code and its Applications to AAL, Toshiba Corporation ATM Forum Contribution af95-1162, 1995

[17] The ATM Forum Technical Committee, Video on Demand Specification 1.1, October 1996

[18] Georg Carle, Aloke Guha, Tim Dwight, Kei-ji Tsunoda, Kumiko Kanai and Hiroshi Esaki, A Cell-Level Forward Error Correction Scheme (FEC-SSCS) for ATM Networks, IEEE/ACM Transactions on Networking, August 1995

[19] Martyn J. Riley and Iain E. G. Richardson, An FEC Protocol for MPEG-2 Video over ATM, Technical Report, The Robert Gordon University, 1996

[20] Georg Carle, Framework with Scaleable Error Control for Reliable Multipoint Services in ATM Networks, International Conference on Multimedia Networking, Aizu-Wakamatsu, Fukushima, Japan, September 1995

[21] Carlos Lopez Fernandez, Andrea Basso and Jean-Pierre Hubaux, Error Concealment and Early Resynchronization Techniques for MPEG-2 Video Streams Damaged by Transmission over ATM Networks, SPIE, San Jose CA, 1996

[22] S. Lee, J. Youn and S. Jang, Transmission Error Detection, Resynchronization, and Error Concealment for MPEG-2 Video Decoder, SPIE, 1993, pp. 195-204

[23] S. Aign and K. Fazel, Temporal and Spatial Error Concealment Techniques for Hierarchical MPEG-2 Video Coded, ICC'95, 1995

[24] Christian J. van den Branden Lambrecht and Olivier Verscheure, Perceptual Quality Measure using a Spatio-Temporal Model of the Human Visual System, in Proceedings of SPIE, San Jose CA, 1996, pp. 450-461

[25] M. Goldstein, N.Björkmann and A. Latour, Cell Loss Effects on Perceived Audio/Video Quality and Quality of Service (QoS), 6th Workshop in Packet Video, Portland USA, 1994

[26] Ralf Steinmetz and Klara Nahrstedt, Multimedia: Computing, Communications & Applications, Prentice-Hall, 1995

[27] Recommendation I.362 B-ISDN ATM Adaptation Layer Functional Description, ITU-T, Geneva, 1993

[28] Recommendation I.363 B-ISDN ATM Adaptation Layer (AAL) Specification, ITU-T, Geneva, 1993

[29] Mohammad Ghanbari and Vassilis Seferidis, Cell Loss Concealment in ATM Video Codecs, IEEE Transactions on Circuits and Systems for Video Technology, 1993, vol. 3

[30] Recommendation I.363.5 B-ISDN ATM Adaptation Layer Specification Type 5, ITU-T Draft Ed. Geneva, 1995

[31] Ismail Dalgic, Performance of Ethernet and ATM Networks Carrying Video Traffic Based on Accurate Characterization of Video Sources, PhD Dissertation, Stanford University, 1996

[32] Arthur A. Webster, Coleen T. Jones, Margaret H. Pinson, Stephen D. Voran and Stephen Wolf, An Objective Video Quality Assessment System Based on Human Perception, SPIE Human Vision, Visual Processing, and Digital Display IV, San Jose CA, 1993, pp. 15-26

[33] Serge Comes, Les traitements perceptifs d'images numerisées, PhD Dissertation Université Catholique de Louvain, 1995

[34] Christian J. van den Branden Lambrecht, A Working Spatio-Temporal Model of the Human Visual System for Image Restoration and Quality Assessment Applications, in Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, Atlanta GA, 1996, pp. 2293-2296

[35] M. Ardito, M. Barbero, M. Stroppiana and M. Visca, Compression and Quality, Proceedings of the International Workshop on HDTV 94, Springer Verlag, Brisbane Australia, 1994

[36] Henning Schulzrinne, James F. Kurose and Donald Towsley, Loss Correlation for Queues with Bursty Input Streams, ICC'92, Chicago, Illinois, 1992, pp.219-224

[37] I. Norros and J.T. Virtamo, Who Loses Cells in the Case of Burst Scale Congestion?, ITC-13, North-Holland, Copenhagen, 1991, pp 829-833

[38] A.J. McAuley, Reliable Broadband Communications Using a Burst Erasure Correcting Code, ACM SIGCOMM'90, Academic Press, Philadelphia, PA, 1990, pp. 297-306

[39] Simon Znaty, Thomas Walter, Marcus Brunner, Jean-Pierre Hubaux and Bernhard Plattner, Multimedia Multipoint Teleteaching over the European ATM Pilot, International Zurich Seminar, Springer, 1996, pp. 225-238