ISO-IEC JTC1/SC29/WG11
MPEG92/
July 1992

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

AVC-298
July 1992

SOURCE: Australia
TITLE: Adapting MPEG 1 video for ATM transmission
PURPOSE: Information

Abstract

This paper presents experimental results concerning the transport of MPEG1 coded video over ATM based networks. The important issues are highlighted, and possible approaches which should be considered in the development of coders for ATM networks are raised. This work is important in order to ensure that a suitable quality is achieved when MPEG2 is carried over ATM networks.

1 Introduction

This paper considers several aspects relating to transporting MPEG 1 coded video over ATM networks. The following points are considered:

- Gains and overheads of using the AAL layers.
- Making use of two priority levels for cell loss protection.
- Changing the slice size to vary the error recovery performance.

Simulations are carried out using a SM3 coder at 1.15 Mbits/second (Pattern **A:** N=4, M=2 and pattern **B:** N=9, M=3), under both sporadic and bursty cell loss conditions. 50 frames of the CIF format *BBC Disk* and the *Mobile and Calendar* sequences are used, and cell loss experiments are performed 4 times to obtain a wider sample space. The error resilience of the different strategies are compared using SNR (luminance only) figures calculated on a frame-by-frame basis. Comments are also made regarding the subjective quality.

2 B-ISDN – Coder Interface

In an ATM based network, such as B-ISDN, data is transported in fixed length units called cells. Every ATM cell consists of a 5 byte header and a 48 byte payload. The header includes addressing information for the cell (virtual path and virtual channel identifiers), along with a cell loss priority flag. During times of congestion cells with this flag set (i.e. low priority cells) will be discarded in preference to the cells without this flag set (high priority cells). It is anticipated that low priority cells will be cheaper to transmit than high priority cells, but low priority cells will have a higher cell loss ratio (by, perhaps, 3 to 4 orders of magnitude).

The ATM layer is not concerned with the nature of the payload – voice, video and data are treated identically. The ATM Adaptation Layer (AAL) is used to provide an interface between user applications and the ATM layer, offering functionality in addition to that offered by the ATM layer, which takes into account the varying requirements of different services. The AAL is further divided into two sub-layers, the Convergence Sub-layer (CS) and Segmentation and Reassembly Sub-layer (SAR), which is shown in Figure 1.

Higher layer applications	
ATM Adaptation Layer (AAL)	Convergence (CS)
	Segmentation and Reassembly (SAR)
ATM layer	
Physical layer	

Figure 1: A portion of the B-ISDN protocol reference model.

The segmentation of data in the AAL proceeds as follows: The coder passes variable sized packets of data (called AAL-SDUs) to the AAL. The implementation used to perform the simulation work discussed later places all the bits from coding one slice into one AAL-SDU. The CS sub-layer may add a header and/or a trailer (for timing recovery and forward error correction), and passes this onto the SAR sub-layer. The SAR breaks this into 48 byte segments, each of which can contain a header and/or a trailer (cell numbering, CRC, length indicator). These resulting segments are passed onto the ATM layer for transmission. Figure 2 shows this flow of data. The size of the AAL-SDU is entirely up to the encoder, the larger it is the fewer the number of partly filled cells. However, if a cell is lost, the rest of the cells in a AAL-SDU may not be usable, due to loss of synchronisation in the variable length code (VLC) decoder. See [1] for more details on the use of the AAL.

3 Simulation of the Network

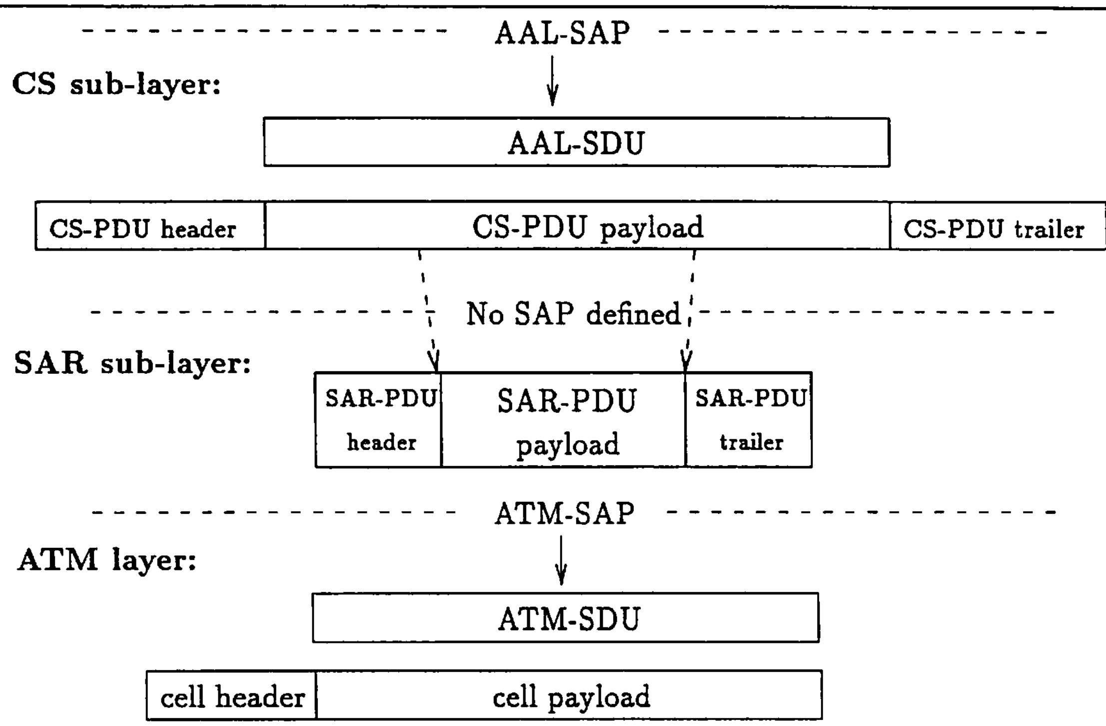
Cell loss in an ATM environment is due to two causes. Firstly, transmission bit errors in the ATM cell header which aren't detected result in mis-delivery of the cell, causing it to either arrive at the wrong address, or to be lost altogether. The second and more dominant source of errors is due to buffer overflow in switching nodes. It is anticipated that cell loss due to buffer overflow will be bursty, as the peak in the data arrival rate at the switch that causes an overflow will not be instantaneous.

The cell loss studies performed for this work use the model proposed in CCITT SGXV [2], whereby cell loss is controlled by two parameters: the average cell loss rate, and the average error burst length (average number of consecutive cells lost). From these parameters, values can be derived which give the probability of cell loss given that the previous cell was either lost or not lost. The model can be applied independently to the high and low priority cells. Whereas the model proposed in [2] is concerned with loss at the cell level, [3] looks at the cell loss problem over much longer periods of time, and can be used to choose appropriate cell loss ratios.

4 The Effect of Packetisation

This section investigates if there are any gains to be made by having the encoder packetise the bit-stream into AAL-SDUs, rather than just filling up ATM cells. The following two arrangements are compared:

1: The bit-stream generated from coding the video is treated as a continuous stream of bits. Bits are placed in cells, without sequence numbers, until the entire video bit-stream has been



AAL-SAP	= AAL service access point	AAL-SDU	= AAL service data unit
CS-PDU	= CS protocol data unit	SAR-PDU	= SAR protocol data unit
ATM-SAP	= ATM service access point	ATM-SDU	= ATM service data unit
ATM-PDU	= ATM protocol data unit		

Figure 2: The flow of data through the AAL showing the header and trailers added by each sub-layer in the AAL.

packetised (i.e. a null, or empty AAL). This represents the case of transporting MPEG coded video over B-ISDN when the coder has no knowledge of the transport medium used. Any cell errors that occur will result in a loss of 384 bits of information, which the decoder will have to detect via loss of synchronisation in the VLC decoder. Once an error has been detected, the decoder will search for the next start code and then resume decoding.

This represents the worse case situation, to which all other schemes should be compared, to see if they actually make a difference.

2: The bits generated from coding a slice are placed in an AAL-SDU. This requires the MPEG coder to have some knowledge of the transport medium (in that information is passed in packets, rather than as a stream of bits), but it does not impact on the coding algorithm.

The AAL routines use sequence numbering to detect missing and inserted cells. If an erroneous cell is detected, the AAL-SDU currently being assembled at the decoder is discarded, an error message is passed up to the decoder and assembly of the next AAL-SDU commences.

In the decoder an AAL-SDU is read in, and the slice data it contains is then decoded. If an AAL-SDU is discarded, the decoder knows immediately that data has been lost, and can then continue decoding the next slice, without the need to search for the next start code (as it will be at the start of the next successfully received slice).

Both of these schemes conceal the effect of lost data using the following strategy, which depends on the type of picture being decoded: For I and P frames, replace the lost slice data with that from the same spatial location in the previous I or P frame. For B frames, use an interpolation of slice data from adjacent I and P frames.

The simulations performed indicated that in scheme **1** not all errors are detected, resulting in corrupted data being processed. This manifests itself as incorrect motion vectors, splashes of colour due to corrupted DCT coefficients, and shifted data due to extra decoded macroblocks.

The advantage of using scheme **2** is that notification of cell errors is given directly to the decoder before it tries to decode the errored data, making recovery easier and eliminating incorrectly decoded data. The drawback of this scheme is that not every cell transmitted will be full (as the length of AAL-SDUs doesn't need to be a multiple of the cell payload length). In the simulations performed, 12% of the output bit-stream was due to unused cell payloads and AAL sub-layer headers and trailers.

The AAL routines used at the decoder for this simulation discard an entire AAL-SDU when a missing cell is detected, even though the missing cell may not have been the first cell in the packet. An extension to this study will be to consider AAL routines which can work around this problem. For the experiments performed in this paper, plain packetisation (scheme **1**) will be able to decode (possibly incorrectly) some of an errored slice, whereas scheme **2** will discard the entire slice.

Table 1 shows the results of this experiment. Only the average sequence SNR is given as an approximate guide to relative quality. The average cell loss rate was 0.002 which, although high, gives between 4 and 20 errors over the 50 frames. The analysis in [3] suggests that during times of congestion the cell loss rate may be as high as 0.01.

The SNR for error burst lengths of 3 ($BL = 3$) tends to be higher than for $BL = 1$ because there are longer error free periods. When the decoded sequences are viewed, the quality of the two schemes are similar: scheme 1 is characterised by incorrectly decoded data, and scheme 2 is characterised by entire slices being concealed. On some sequences (such as *BBC Disk*) the effect of error concealment is noticeable, but on the majority of sequences, incorrectly decoded data is a far more noticeable artifact. The reliability of error detection obtainable by using the AAL interface outweighs any drawbacks due to unused capacity in ATM cells.

			SNR of decoded errored sequences			
		SNR with	Scheme 1		Scheme 2	
Sequence	Pattern	no errors	BL=1	BL=3	BL=1	BL=3
BBC disk	**A:**	29.57	25.43	26.98	24.26	26.62
	B:	31.54	21.96	25.48	21.39	25.61
Calendar	**A:**	24.44	21.74	20.05	23.77	24.24
	B:	25.92	24.56	25.25	24.07	25.29

Table 1: The average sequence SNR for the two different packetisation schemes discussed: without error, and with error burst lengths of 1 and 3. The two test sequences are used, along with the two MPEG coding patterns (**A**: N=4, M=2 & **B**: N=9, M=3).

5 Using Two Cell Loss Priority Levels

This section investigates the gain in quality possible when some of the bit-stream travels in high priority cells. B frames will be transmitted in low priority cells, as errors in these pictures do not propagate, whereas errors in the other pictures do. Segregating the sequence into cells in this fashion is a form of temporal layering. Both this section and the next use the scheme 2 AAL.

A temporally layered sequence is to be compared against a sequence which only uses low priority cells. Two scenarios will be considered, firstly one in which no errors occur in the high priority channel, and then one in which some errors occur, but at a rate $\frac{1}{5}$ of that used for the low priority channel. (The factor of $\frac{1}{5}$ is used so that several errors occur during the 50 frame test sequence.) Cell loss rates for the high and low priority cells are chosen so that approximately the same number of errors occur in the single layer case as in the two temporal layered cases, with an average overall cell loss rate of approximately 0.003.

Table 2 show the results of this test. For the case with a burst length of 1, the temporally layered decoder performs better than the single layer version, even when high priority cells are lost. The quality of the decoded errored sequence when no high priority cells are lost is quite good, at times within 1 dB of the unerrored sequence. It is interesting to note that when the error burst length is 3, the layered decoder is better than the single layer coder without high priority loss, but is worse with high priority loss.

When the sequences are viewed, the best results are obtained with the layered high priority lossless decoder, in which errors are very short lived, being quickly refreshed. In the other cases, errors persist for a greater number of frames, and unerrored cells obtain prediction information from errored pictures, spreading the effect of the error. Many studies have been performed which consider spatial layering for cell loss resilience, and have concluded that there are advantages, but some small loss of efficiency may occur. The temporal layering tested here imposes no overheads, and can easily be used with MPEG 1 video. The visible effect of errors in temporal layering is a brief flash due to error concealment, which is a bit more noticeable than the temporary loss of high frequency detail which occurs in spatial layering.

The results in Table 2 also show the percentage of the cells that are high priority. While this is currently quite high, it can be reduced down to 25–30% by either increasing the number of B frames, or moving some or all of the P frames into low priority cells.

6 Varying the Size of Slices

The area of a picture that is contained within an AAL-SDU also effects the level of visual degradation that results when cells are lost. The experiments performed in the previous two sections use a slice containing 22 macroblocks. The size of a slice determines how much data is lost when an AAL-SDU is discarded in the AAL due to cell loss. The smaller the slice the less the area of picture affected, but the greater the bit-rate overhead introduced by an increase in the number of slice headers, and the larger the number of partly filled ATM cells. When coding at 1.15 Mbits/second, one ATM cell contains on average $3 \rightarrow 4$ macroblocks, so it requires $5 \rightarrow 7$ ATM cells to transport a 22 macroblock slice.

Simulations were performed which looked at the effect of different slice sizes, containing 11, 22, 44 and 66 macroblocks, which equates to $\frac{1}{2}$, 1, 2 and 3 stripes of the picture. The results in Table 3

Sequence	Pattern	% CLP = 0 cells	SNR with no errors	CLP = 0 cells lost	SNR of decoded errored sequences: Priority used BL=1	Priority used BL=3	Priority not used BL=1	Priority not used BL=3
BBC disk	A:	81	29.57	yes	22.54	23.02	18.82	24.12
				no	27.08	28.45	21.30	25.06
	B:	68	31.54	yes	19.54	19.74	15.62	21.15
				no	26.02	27.78	18.16	24.62
Calendar	A:	78	24.44	yes	23.48	23.59	21.80	23.93
				no	24.20	24.33	23.01	24.10
	B:	62	25.92	yes	23.64	23.82	20.13	24.76
				no	25.28	25.58	22.96	25.20

Table 2: The average sequence SNR for the bit-streams transported with and without making use of high priority (CLP = 0) cells. Two cases are considered, that in which high priority cells are not lost, and that in which they are. Experiments are performed for both test sequences, using two coding patterns and error burst lengths (BL) of 1 and 3.

show the average sequence SNR for sequences decoded: without error, with error burst length of 1 and 3 (cell loss ratio = 0.002). The '% excess' column shows the total number of cells transmitted, as a percentage increase over the case in which no AAL is used at all. The increase is due to CS and SAR headers, and unused payload capacity.

Sequence	slice size	Coding pattern A: %excess cells	Coding pattern A: SNR of decoded sequences No errs	Coding pattern A: BL=1	Coding pattern A: BL=3	Coding pattern B: %excess cells	Coding pattern B: SNR of decoded sequences No errs	Coding pattern B: BL=1	Coding pattern B: BL=3
BBC disk	11	31.7	28.91	24.35	26.07	19.9	31.39	22.95	25.85
	22	12.7	29.57	23.17	25.73	12.4	31.54	20.17	24.77
	44	9.3	29.53	19.53	22.89	8.9	31.54	18.70	21.95
	66	9.2	29.45	18.25	23.96	7.7	31.56	16.19	20.84
Calendar	11	20.0	24.36	22.82	23.64	19.6	25.80	24.77	25.55
	22	12.5	24.45	23.64	24.16	12.4	25.92	23.97	25.26
	44	9.0	24.46	22.91	23.68	9.2	25.92	22.96	24.69
	66	7.8	24.46	21.83	23.54	8.3	25.89	21.95	24.25

Table 3: The average sequence SNR for different slice sizes under error conditions. Uses the two MPEG coding patterns and two different average error burst lengths (BL=1 and BL=3).

When the decoded sequences are viewed, the visual effect of having large areas of the picture (up to a sixth) replaced by a previous picture, to cover the effect of cell loss is quite noticeable, especially in the *BBC Disc* sequence. For a burst length of 1 and a slice size of 11, the visual effect of concealment is quite small, and only slightly more noticeable for slice sizes of 22. When the burst length is increased to 3, the small slice size sometimes results in two errored slices, increasing the area of concealment to the same as that for slice sizes of 22. On the whole though, the results obtained for slice sizes of 11 and 22 are quite comparable, and far better than the larger slice sizes. Due to the large overhead with the smallest slice size, 22 appears to be a reasonable compromise.

7 Conclusions

This paper has looked at how to transport MPEG 1 video data in ATM cells, investigating how the parameters controlling MPEG can be tailored to best suit the conditions encountered under ATM.

- Using the AAL layer guaranteed that every error is detected, eliminating the chance of incorrectly decoded data, and thus improving the quality.
- Dividing picture types into high and low cell loss priority streams greatly improved the overall quality, without any additional overhead.
- Controlling the amount of picture data transmitted in a slice also affected the error resilience. The smaller slices offered better quality under errors, but at the cost of ATM cell packing efficiency, especially for slice sizes less than the picture width.

It is worth noting that video traffic will be one of the major users of broadband networks, and so work needs to continue to investigate the best coding schemes to use. All of the features investigated in this paper were performed within the framework of the MPEG 1 standard, indicating that existing coders could be adapted readily to delivery over broadband networks. It is clear, however, that the error handling capabilities could be improved by extensions beyond the framework of MPEG 1. In particular, the use of spatial scalability for cell loss protection should be investigated.

References

[1] Australian contribution to CCITT SGXV Experts group for ATM video coding. *ATM Adaption Layer Type 2 Functionality*, March 1992. Document AVC-222.

[2] CCITT SGXV Experts Group for ATM Video Coding. *Cell loss experiment specifications*, January 1992. ISO/IEC JTC1/SC29/WG11 MPEG 92/027.

[3] Australian contribution to CCITT SGXV Experts group for ATM video coding. *Cell Loss Characteristics for Statistically Multiplexed Video Sources*, July 1992. Document AVC-296.