ATM Transmission of Wavelet Transformed Video Images: the Impact of Cell Losses and Transmission Errors

Fabrizio Argenti(1), Giuliano Benelli(2), Lorenzo Favalli(2),
Stefano Ferraresi(2), Alessandro Mecocci(2)

(1) Dipartimento di Ingegneria Elettronica
Università di Firenze, via Santa Marta 3,
50139 Firenze, Italy

(2) Dipartimento di Ingegneria Elettronica,
Università di Pavia, via Abbiategrasso 209,
27100 Pavia, Italy

Résumé

Les signaux TV numérique sont, en gèneral, comprimés pour réduir la bande de transmission. Le signal essu de ce processus est plus sensible aux errèurs du canal. Cet article analyse les problèms relatifs à la transmission sur un reseau ATM des signaux TV comprimés. De façon a satisfaire les conditions emposées par le bande de transmission, une quantification adaptative et des mèthodes de compression basées sur la dècomposition en ondelettes de l'image originale ont été developpées. Un nouveau protocole, qui ajoute des informations de controle à la cellule ATM á été developpé pour minimiser les effets de la perte de cellules ATM. Des résultats experimentaux obtenus sur des images conventionelles sont presentés.

Abstract

Digital TV signals are generally compressed to reduce transmission bandwidth requirements. In this process, the signal becomes more sensitive to transmission impairments.The present paper analyzes the major problems related to the transmission of compressed TV signals on an ATM network. In particular, some adaptive quantization and coding strategies based on wavelet decomposition of the original image have been developed, which permit to match the source bit-rate to the current throughput that ATM can provide. A new protocol to add control information to ATM cell, has been developed which minimizes the effects of synchronization loss on the image reconstruction when packets are missed in the ATM network. Some experimental results, using standard images are presented.

I.Introduction

The production, storage, and transmission of TV image sequences are becoming important aspects in many practical applications. Digital TV images require a wide bandwidth, although redundancy provides considerable leeway for reduction in the bandwidth requirement: the most promising compression techniques appear to be the discrete cosine transform (DCT)[1] and subband coding[2].

The broadband integrated services digital network (B-ISDN) which supports voice, data, as well as standard and H-D TV traffic, is expected to become the major transmission vehicle of the next-generation communication systems. The most attractive B-ISDN transfer technique is asynchronous transfer mode (ATM) [3], [4], which can easily accommodate variable bit-rate services within the same network: in this case transmission overloads may produce the loss of one or more packets composed by hundreds of bits, while the channel noise can be considered negligible. The ATM protocol can provide a connection bandwidth to the source which depends on the current traffic load on the network. Therefore, it is quite important to match the source bit-rate to the network bandwidth availability. In this paper some new methods to perform this matching are proposed. These methods use different quantization and coding strategies of the subbands in order to maximize the source quality.

The compressed images are quite sensitive to packet losses, which can occur in the network. A new packetization strategy is proposed, which reduces the degradation of the reconstructed sequence.

Some experimental results relative to the performance of the proposed algorithms, using standard images are presented.

II.Network Transmission

The multilevel decomposition has several advantages and provides useful tools to address different problems that may arise when interfacing to a telecommunication network.

The first is the coexistence of different equipment in the end-user premises or within the distribution network itself. Another important advantage, although less explored, may be found addressing the issue of a **Graceful Degradation** of the service quality in presence of bandwidth constraints. This argumênt is strictly tied to what may happen in B-ISDN/ATM networks. In the usual approach, the station declares its class and the transmission parameters associated during the call set up phase. If the network is capable to meet these requirements, the connection is established and transmission begins.

To ensure the source transmission characteristics don't exceed those established, different mechanisms can be implemented which force the output of the source to a certain maximum or try to smooth traffic peaks[5][6]. In the first case we tâlk of **policing** techniques such as the *Leaky Bucket* and its variations. These techniques ensure that the network never receives traffic exceeding its capacity, but have the drawback that cells are discarded "randomly", without any knowledge of their importance for the source. In the second case (**shaping** functions), cells exceeding the negotiated rate are buffered, thus introducing an additional delay in order to achieve a "constant" bit rate. This delay may be unacceptable for sources with strict timing constraints such as those providing real time video services. The fundamental limitation of these approaches, lays in the fact that the source is considered dumb and only capable to put bits on the link once the connection has been set up. On the contrary, we have many opportunities to operate directly at the source level to tune the output bit rate. Using the source model described in the previous sections, we can reduce the amount of information transmitted by reducing the number of quantization levels and/or suppressing one or more subbands. As a result, the transmitted image may be of a lesser quality than desired, but we may avoid cell dropping within the network under normal operative conditions. In the next section we will describe some *strategies* that may be used to pursue our goal and the results that may be obtained.

III.Dynamic Approach

Given a wavelet-decomposed image, the output bit rate will depend on choices such as the quantization law in the different subbands and the actual transmission of all or some of the subbands: the final result may be different according to the order in which we change the quantization levels in the subbands. The basic idea to adapt the actual output bit-rate is to progressively reduce the resolution level of the image starting from the details of lower relevance. We'll see how we can indeed find good compromises between quality and service.

A. Strategies description.

All strategies act at first on the third subband SB_3 corresponding to the edge details, since its energy content is lower than that of the other subbands and the human eye is not very selective about diagonal details. As regards to SB_1 (horizontal details) and SB_2 (vertical details), becomes of great importance the correct evaluation of the energy levels associated with each of them. The best results will be obtained starting to drop bits from the signal associated to the subband with a lesser energy content: this requires the calculation of the energy levels on a per-frame base. Three strategies have been extensively evaluated and are presented in the following.

Strategy 1 (ST1). The number of quantization levels in SB_3 decreases from 64 to 2 halving at every step (64, 32, 16, 8, 4, 2). At this point SB_3 is discarded and the process is repeated for SB_2 and SB_1: in the end, only BB is transmitted. BB itself may be compressed by means of different quantization levels for the DCT values. We have analyzed three cases: BB uncompressed, BB compressed with DCT at 64 or 32 levels.

Strategy 2 (ST2). The quantization levels are the same for all subbands and are uniformly reduced at every step. As a last choice, the three subbands are discarded altogether. The action taken is uniform and the energy levels are not taken into account; again, results will be provided for BB without compression and with DCT at 64 and 32 levels.

Strategy 3 (ST3). It is similar to ST1, but we start reducing the number of levels in SB_i while still acting on SB_{i+1}, not when this has been discarded (64,64,32; 64,32,16; 32,16,8; ...).

B. Experimental Results.

The effectiveness of the three strategies has been evaluated both in terms of signal to noise (S/N) ratio between the original image and the image obtained after the strategies have been applied and in terms of the bit rate generated by the different cases after the coding phase. The purely mathematical approach has been supported by a "visual" approach based on the human evaluation of the actual results. Several images have been used to compare the strategies just described. In this paper we report the results for only two of them that are widely used in the literature. These images are *LENNA* and *VOITURE*: the first is a 256*x*256 pixels image the other is a 512*x*512.

Through the experimental work the quality of the reconstructed image has been evaluated by means of the S/N ratio , defined as the ratio between the variance of the original image vs. the mean square error between the reconstructed and the original images.

This parameter is highly sensitive even to small differences between two images that will be regarded as having a low fidelity, while the human eye roughly notices a negligible degradation. On the contrary, the human eye may be more sensitive on other disturbances which are not revealed by the S/N ratio defined.

Since we are interested in the transmission of our images through a telecommunication network, a further parameter of great importance in our study, is the bit rate corresponding to the image after a given strategy has been applied.

The results for strategy ST1 (Figure 1), show that the S/N ratio between the original and the reconstructed images is roughly constant when the number of quantization levels is reduced from 64 to 4 in a given subband, while a sharp step is noticeable with a further reduction to only 2 levels. The

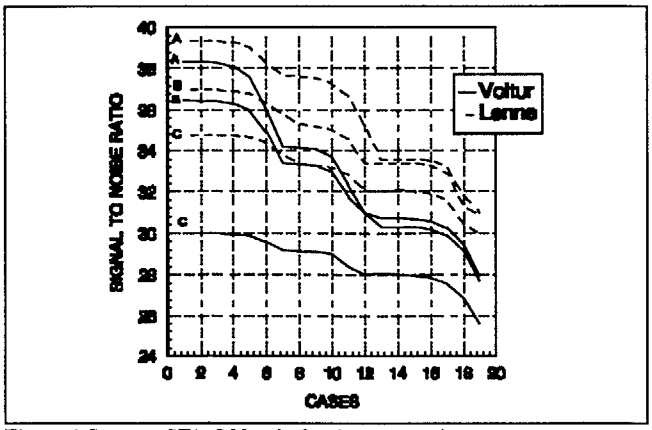

Figure 1 Strategy ST1: S/N ratio for the two test images.
A: BB uncompressed.
B: BB with DCT at 64 levels.
C: BB wit DCT at 32 levels.

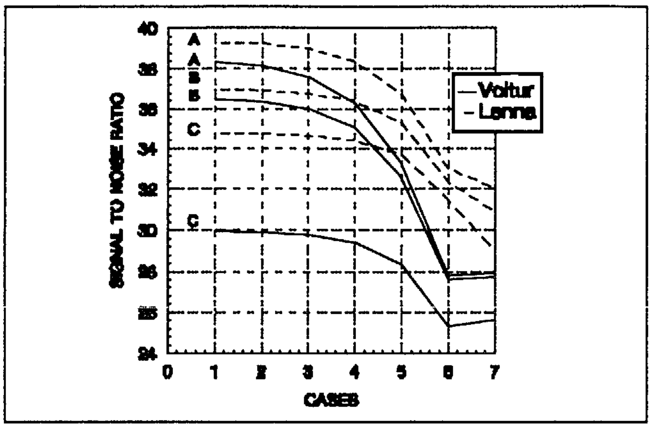

Figure 2 Strategy ST2: S/N ratio for the two test images.
A: BB uncompressed.
B: BB with DCT at 64 levels.
C: BB with DCT at 32 levels.

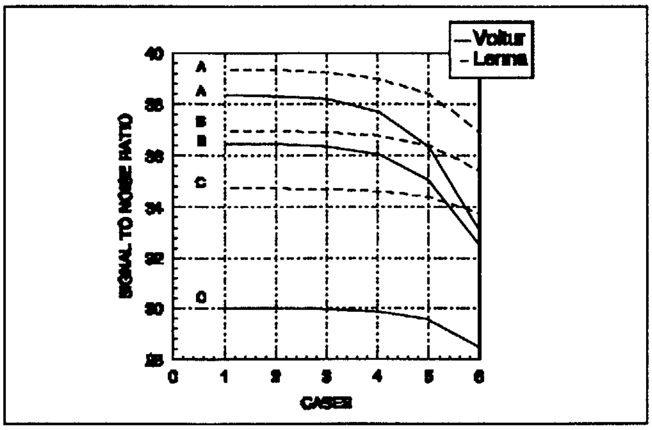

Figure 3 Strategy ST2: S/N ratio for the two test images.
A: BB uncompressed.
B: BB with DCT at 64 levels.
C: BB with DCT at 32 levels.

application of DCT at 64 or 32 levels on BB simply has the effect of shifting the graph to lower values without changing the qualitative behaviour of the strategy. Strategies ST2 (Figure 2) and ST3 (Figure 3), show a stronger reduction of S/N compared to ST1.

The output bit rate of the source (considering images generated every 50 Hz and Huffman coded) changes almost linearly for all three strategies and show that it is possible to keep an almost constant image quality, reducing the number of quantization levels from 64 to 4, while obtaining an evident reduction in the bit rate. In Figure 4 the difference in bit rate values between the two images is due to their different resolution. A comparison of the different strategies on different images, reveals that the reduction in the S/N ratio is mainly influenced by the energy content of the subbands (the more the energy, the more the details contained in that particular subband).

These graphs, provide an easy way to match the required output bit rate (as it may be imposed by the ATM admission control function) with the best S/N ratio with a given strategy and case.

IV.Transmission and Image Recovery

Once the output stream is ready, we need to split it into cells and provide some recovery mechanism to allow an easy reconstruction of damaged images. In this work, we make the hypothesis that bit errors due to disturbances or to the transmission media are negligible since transmission over a fiber optic link is assumed: the principal cause of errors at the receiver is then the loss of cells due to network congestion.

In the approach we have studied, before being compressed, eight adjacent screen lines are grouped to form *stripes* each of which is delimited by appropriate flags. The information is then splitted into small ATM cells. We can insert additional control information in the payload to help the receiver in its work. The scheme we have devised reserves the first two payload bytes for the additional control information thus leaving 368 bits for effective data. The 16 extra-control bits are organized as follows: *a)* an n bits "SEQ" field to indicate the progressive number of the cell transmitted according to a windowing mechanism. The counter is reset at the beginning of every stripe; *b)* m bits in field "START" for the offset of the first complete codeword contained in the cell; *c)* t bits for the location of last useful bit in case of the last cell of the stripe ("LAST" field).

The exact values for n, m, t may be tuned depending on the information that need to be sent: in our experiments we have always used n=4, m=4, t=8 and these values have always proved satisfactory.

The last cell of the stripe is indicated by the all 1s configuration of the SEQ field so that the receiver may resynchronize on the new beginning of stripe, since the beginning of a stripe is coincident with the end of the previous stripe with no gaps in between. If a cell is not the first of the stripe (or conversely the last of the previous one) the LAST field may be neglected and its bits may be used to increase the capacity of the

START field thus reducing the possibility of an overflow of the value of the offset.

This scheme introduces an additional load approximately equal to 4%, but allows a very easy and effective way to confine errors and to recover the correct synchronism: in case a single cell is lost, all pixels are correctly decoded until the one corresponding to last codeword completely contained in the cell preceding the lost one. Correct decoding may restart form the START bit (corresponding say to pixel x) of the successive cell: information from all pixels after this is buffered and reconstruction is started from the last pixel of the stripe backward to pixel x. This is possible as we know the exact position on the screen corresponding to the end point of each stripe. The strategy for the substitution of the lost pixels differs from detail subbands SB_i to BB: in the first case, the value 0 is inserted in all corresponding positions since this is by far the most probable value in these subbands. If the lost cell belongs to BB then more sophisticated substitution strategies are

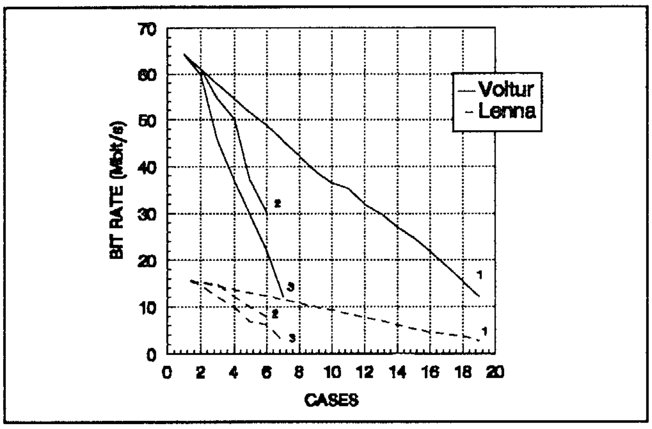

Figure 4 Bit rate corresponding to the two test images.
1) ST1; 2) ST2; 3) ST3.

required. A detailed study of these strategy is out of the scope of the present paper, since we only want to introduce what we think is a promising approach. In our experiments, the simplest recover technique has been used: if part of the information in BB is lost, it is substituted with information from the previous frame.

If more than one cell is lost, only the bit stream between the first and the last is lost. The receiver hooks to the cells preceding and to those following the lost data. Correct cells belonging to the current stripe eventually received between these two extremes are also lost, since there is no way to recover the exact position of the corresponding pixels. The effect of the loss of a whole stripe in the BB subband is depicted in Figure 5: it particularly evident on the gate that is the part which is moving faster along the visual plane. This degradation may be greatly reduced by using motion compensation techniques.

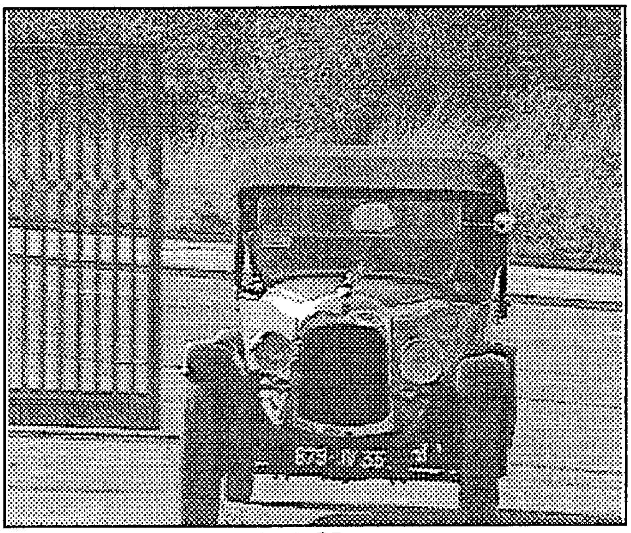

Figure 5 Loss of an entire stripe in BB.

V.Conclusions

This paper has presented a new approach to the transmission of video images in an ATM environment which combines a dynamic adaptation of the coding parameters and a control protocol to protect the information sent from cell dropping within the network. The results provided show how this technique may be both simple and effective. Further work is being currently carried out to obtain a better knowledge of its behaviour in a "real" environment in comparison with other techniques. Also, more sophisticated control mechanisms and compression-recovery mechanisms based on motion compensation schemes are being implemented.

References

[1] W.Chen, C.H.Smith, "Adaptive Coding of Monochrome and Color Images", IEEE Trans. Comm. vol. COM-25, pp.1285-1292, Nov. 1977.

[2] J.Woods, S.O'Neil, "Subband Coding of Images", IEEE Trans. Acoust. Speech, Signal Process., Vol. ASSP-34, pp.1278-1288, Oct. 1986.

[3] F.Argenti, G.Benelli, A.Mecocci, "Source Coding and Transmission of HDTV Images Compressed with the Wavelet Transform", Journal on Selected Areas in Communications, in press.

[4] S.E.Minzer, "Broadband ISDN and Asynchronous Transfer Mode (ATM)", IEEE Communications Magazine, pp. 17-24, Sept. 1989.

[5] J.J.Bae, T.Suda, "Survey of Traffic Control Schemes and Protocols in ATM Networks", Proc. of IEEE, vol. 79, pp. 170-189, Feb. 1991.

[6] AA.VV. Journal on Selected Areas in Communications, Special Issue on "Congestion Control in High Speed Packet Switched Networks", Vol. 9, n. 7.