EXPERIMENTAL RESULTS OF MPEG-2 CODED VIDEO TRANSMISSION OVER A NOISY SATELLITE LINK*

Nedo Celandroni#, Erina Ferro#, Francesco Potortì #
Antonio Chimienti^, Maurizio Lucenteforte^

#CNUCE, Institute of National Research Council
Via S. Maria 36 - 56126 Pisa - Italy
Phone: +39-50-593111 - Fax: +39-50-904052 - Telex: 500371
E-mail: {N.Celandroni, E.Ferro, F.Potorti}@cnuce.cnr.it

^CSTV, Institute of National Research Council
Strada delle Cacce 91 -10135 Torino -Italy
Phone: +39-11-3919242 - Fax: +39-11-341882
E-mail: {chimienti, lucente}@cstv.to.cnr.it

ABSTRACT

The results of a transmission experiment of MPEG-2 coded video data over a satellite link affected by noise are presented. The goal was to investigate under which conditions this type of transmission is economically feasible. The scalability feature of the MPEG-2 encoder was used to produce different bitstreams of the same video sequence. We wanted to verify the best combinations among video and channel codings in presence of attenuation on the satellite link, in order to optimize the bandwidth utilization for a requested image quality. The limitations imposed by the satellite equipment at our disposal are outlined. In spite of these limitations, the results obtained give indications about how to deal with different quality of service requirements and highlight the flexibility of the scalable video coding.

I. INTRODUCTION

The transmission of uncompressed, digitized, full-motion, colour video is expensive in terms of bandwidth (about 166 Mbit/s). Video compression technology is particularly attractive for transmission over a satellite link which entails using techniques to optimize the bandwidth occupation and reduce costs. The Motion Picture Expert Group (MPEG) video compression schemes have emerged as standards for multimedia applications from the many video coding schemes proposed. An encoder attempts to keep the quality of the video output constant at the price of changing the bit rate, thus producing a variable bit rate (VBR) bitstream. The burstiness of the resulting traffic depends on the encoding scheme adopted and on the vivacity of the movie scenes. The reduction in the bandwidth needed by VBR video allows data communication networks to agilely support high quality multimedia applications. Scalable video coding is used for a number of applications where it is necessary to display video at different resolutions or quality levels. In a scalable video encoder, two or more layers are generated, coded and transmitted. In the simpler case an enhancement layer encoder utilises information generated by an independent base layer encoder. With regard to error-prone transmission environments, a scalable video coding may perform better than an equivalent non scalable one. In fact an erroneous bit, caused by a channel error, generally has a different effect on the decoded video, if associated with the base or with the enhancement layer. When VBR video traffic is transmitted over a satellite link, in the link budget design the noise level must be carefully calibrated in order to avoid impairness in the reconstruction of the images received. In fact, the target channel signal-to-noise ratio (CSNR), and thus the resulting bit error rate (BER) is chosen as a compromise between quality and cheapness. Moreover, in satellite transmissions over 10 GHz, the signal attenuation due to rain imposes the adoption of fade countermeasures to ensure an acceptable level of link availability with a reasonable channel quality.

This paper reports the results of an experimental transmission that we made over a noisy satellite channel of both a non-scalable version and two different scalable versions of the same strips of "The sheltering sky" movie(1). The goal was to find the most suitable combination of video coding (scalable or non-scalable bitstream) and satellite channel transmission parameters in order to optimise the bandwidth in different link degradation conditions, given a required image quality. The experiment consisted in producing the traces of three strips of film of different vivacity coded according to the MPEG-2 standard with several coding parameters. The data streams obtained were transmitted between two real earth stations connected together via a satellite emulator. The received sequences, relative to different transmission parameters necessary to cope with the channel degradation level, were then analysed to evaluate the resulting video quality.

II. MPEG-2 VIDEO CODING AND SCALABILITY

The MPEG compression algorithms use interframe compression and can achieve compression ratios of 40:1 via a temporal correlation. The MPEG first-phase standard (MPEG-1) is tagged for compression of 320x240 full motion video at rates of 1 to 1.5 Mbit/s. MPEG-2 standard [1] is intended for higher resolutions, similar to the digital video studio standard CCIR 601, EDTV, and further leading to HDTV. An MPEG application determines a sequence of intrapictures (I), predicted pictures (P), and bidirectionally-predicted pictures (B). The first ones are self-contained and provide access points to the coded sequence where decoding can begin; they are coded without reference to other pictures and with only moderate compression. P pictures are generally used for further predictions and are coded more efficiently, using forward predictive coding, where the actual frame is coded with reference to the previous frame (I or P). The compression ratio of P frames is significantly higher than the I frames. B pictures provide the highest degree of compression. They are coded using two reference frames, a past and a future frame (I or P frames) for motion compensation. A sequence is divided into a series of groups of pictures (GOP). A GOP is a flexible set of pictures, being composed by a variable number of I, P and B pictures, according to the intrapicture interval.

The MPEG-2 coding used was performed on a one minute long video sample (1500 frames) extracted from the movie "The sheltering sky". Both non-scalable and scalable traces were generated. To produce scalable video, we used the *SNR* (signal-to-noise ratio) video *scalability*. It provides a *base* bitstream, associated with a lower quality video, and an *enhancement* bitstream, associated with a higher quality video representation. It is impossible to reconstruct the enhancement layer without decoding the base layer in parallel. Without scalability, we produced a VBR single-stream coded with an average rate of 2.37 Mbit/s (hereafter *non-scalable*). When the SNR scalability was used, we produced two base layers coded at a constant bit rate (CBR), refined by two VBR enhancement layers. The two combinations have the following average rates: 1.5 (base) + 1.17 (enhancement) Mbit/s *(scalable1)*, and 1.066 (base) + 1.482 (enhancement) Mbit/s (*scalable2*), respectively. The VBR traces were generated with a quantization step equal to 3 and 6 respectively, but we decided to use only the last one because step 3 produced a bit rate which was too high for our satellite transmission system. As a metric of service quality for video connections we used in our study an objective video quality assessment system [2] which emulates the HVS (Human Visual System). In this method a linear combination is calculated of three complementary video quality measurements, based on spatial and temporal information distortion. The resulting *quality factor* (QF) is an estimation of the subjective mean opinion score: imperceptible (5), perceptible but not annoying (4), slightly annoying (3), annoying (2), very annoying (1). In Table 1 the mean QF and bit-rate for the 3 different video codings used are shown.

	MEAN QF	MEAN BIT RATE [Mbit/s]
non-scalable	4.66	2.37
scal.1 (base only)	4.36	1.5
scal.1 (base+enh)	4.68	1.5+1.17
scal.2 (base only)	4.09	1.066
scal.2 (base+enh)	4.68	1.066+1.482

Table 1. Mean QF and bit rates used in the simulation

The final video quality of all the coding traces is practically the same. The mean bit-rate is higher in the scalable codings, due to the redundancy of the overhead in the enhancement layer. We tested two different bit rates of the base layer, that is to say a different bit rate distribution between the base and the enhancement layer with the same final video quality.

III. THE SATELLITE NETWORK

Figure 1 represents the test environment. The equipment used was previously employed in experiments on the Olympus and Italsat geostationary satellites [3]. This time we used a satellite emulator which introduces the correct round trip time and a calibrated amount of noise to emulate the attenuation of the transmitted signal due to rain fading.

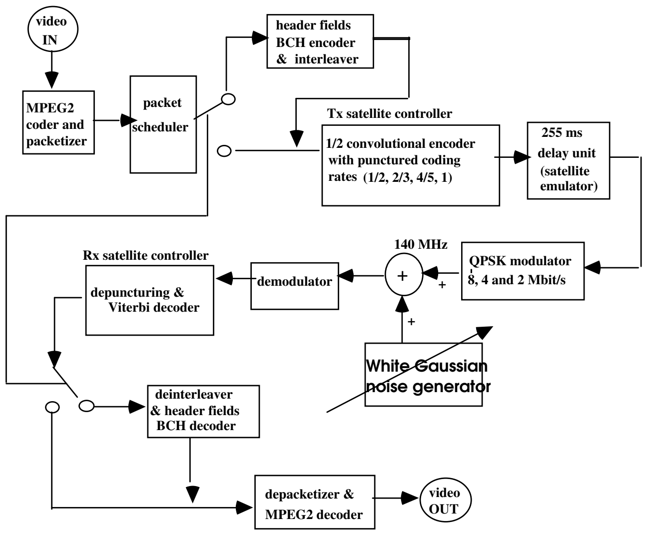

Fig. 1. The satellite network environment

We denote by *link degradation* the difference between the reference and the current value of E_b / N_0 (bit energy to one-sided noise spectral density ratio), which expresses the CSNR. The reference value (12 dB), relative to clear sky conditions, allows a BER of 10^{-8}. Taking the traces from the MPEG-2 encoder, the packet scheduler reproduces packets at exact time instants so as to emulate the real time data in output from a hardware encoder. The transmitting satellite controller adjusts the packet sending times according to the channel access scheme, and applies the channel coding to the base band data stream. The 1/2 convolutional encoder with the puncturing feature allows 1/2, 2/3 and 4/5 coding rates on the satellite channel. Optionally, an outer BCH(2) encoder on the header fields with data interleaving can be applied. No coding at all is possible as well. A base band delay unit introduces the satellite round trip time of 250 ms, then the QPSK modulation of the 140 MHz intermediate frequency carrier allows bit rates of 8, 4, and 2 Mbit/s. Each packet can be sent at individual bit and coding rates. White Gaussian noise is introduced at intermediate frequency level, so as to give the required value of E_b / N_0 at the demodulator input. The receiving controller decodes and packetizes data to be sent to the MPEG-2 decoder which, after data serialisation, reproduces the resulting traces of the video images. At this point we analysed the objective and subjective quality of the corrupted video stream.

We assume the following scenario.

a) The video encoder can operate in one only coding mode, i.e. it can produce only scalable or not-scalable bitstreams.
b) The sending earth station knows the:
-channel degradation of the receiving station [4, 5];
-target quality factor of the video service;
-amount of channel bandwidth available for the video transmission.

All this information is used by the earth station:

- to properly select the suitable channel coding and bit rates, in order to compensate for the different fade conditions;
- in case of scalable transmission, to decide on the convenience to transmit the enhancement layer other than the base layer.

The bandwidth allocation algorithm for video data is beyond the aim of this paper; one of the policies proposed in References [6, 7] is assumed to have been adopted. The residual bandwidth, not actually used by video, can be exploited to send low priority traffic.

After convolutional-encoding/Viterbi-decoding, the residual errors are distributed in bursts of various length, rather than uniformly. In order to compare the quality of the video sequences obtained with bursty and random error distributions, data interleaving was used as well. Moreover, a BCH outer encoder was applied to the header fields to eliminate all the errors, so that we could see the effect of header corruptions on the video quality. This further protection implied the use of the data interleaver.

IV. EXPERIMENTAL RESULTS

Our main goal was to find out which is the best video coding type for use over a noisy satellite link, among the three (non-scalable, scalable1 and scalable2) we have considered. Indeed, the base-only versions of the two scalable codings should be considered as particular cases of the base+enhancement codings. For example, let us assume that one of the two scalable codings is chosen as being the best for this type of transmission. The MPEG-2 encoder at the source thus generates both the base and the enhancement data flows. The transmitting earth station is then responsible for choosing the best channel coding and bit rates for the two flows, and possibly for discarding the enhancement flow. These operations are based on real-time measurements of the attenuation on the satellite link, made by the receiving station. The decisions of the transmitting station influence both the QF of the received MPEG-2 data and the satellite average bandwidth occupancy. In general, a higher channel occupancy leads to a better quality, so the choice is a matter of trading off small bandwidth with high video quality. In order to make such a trade off, we used the criterion of maintaining a given QF while using the smallest possible bandwidth on the satellite channel. Thus, the transmitting station continuously monitors the link attenuation and selects, for both the base and the enhancement flows, channel bit and coding rates such that the QF at the receiver is not smaller than the target one with a minimal average channel occupancy.

We started by analysing the QF of the received flow as a function of the satellite channel bit and coding rates, and of the channel attenuation. The channel bit and coding rates used for the experiment (summarised in Table 2) are selected among the ones allowed by the burst modem available to us, which dictated most of the parameters chosen for the experiment. In particular, the raw throughput of the video stream after channel coding does not exceed 7.2 Mb/s, which is the maximum we could afford, given our modem and the overhead needed by the available satellite access scheme.

Bit rate	Coding rate	Redundancy
8 Mb/s	1:1	1
8 Mb/s	5:4	1.25
8 Mb/s	3:2	1.5
8 Mb/s	2:1	2
4 Mb/s	5:4	2.5
4 Mb/s	2:1	4

Table 2. Bit and coding rates used for the experiment.

Each combination of channel bit and coding rates for both the base and the enhancement flows corresponds to an average bandwidth occupancy, which thus assumes discrete values. Nevertheless, representing the QF on a bandwidth-attenuation plane gives a thorough insight into the behaviour of the system as the attenuation changes. Figures 2, 3, and 4 show regions of the bandwidth-attenuation plane where the QF is comprised between some given threshold values. In these figures, M indicate the non-scalable coding (monoflow), B stands for base flow and E for enhancement flow.
Figure 2 shows the simplest case.

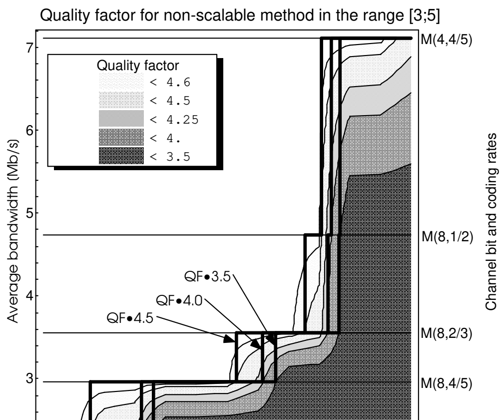

Figure 2. Isoquality regions in the attenuation /bandwidth plane, and paths of constant QF for non-scalable MPEG-2 coding.

The usable average bandwidth values are those that correspond to the horizontal lines, which are labelled with the relative bit and coding rates. The other horizontal lines represent other channel bit and coding rates, with increasing redundancy and, consequently, increasing data protection from corruption. The white region is the set of points (attenuation, bandwidth) whose corresponding QF is greater than 4.6, out of an attainable QF of 4.66 for the used MPEG-2 coding. This region is practically error-free, and the quality of the received data is the same as the one of an uncorrupted video stream. The darkest region is where

the quality of the received data is between 3.5 and 3, which we considered as the floor QF value. The three thick lines are the paths that the transmitting station follows in order to maintain a quasi-constant QF, jumping from one channel coding rate to the next, as the channel attenuation changes. The line for QF not less than 4.5 is the nearest one to the top-left corner of the graph, with respect to the lines for 4.0 and 3.5. This visual indication shows the increasing cost, in terms of average bandwidth, needed in order to maintain a QF not smaller than 4.5, with respect to maintaining a QF not smaller than 4.0 or 3.5.
Figures 3 and 4 are more complex to read, because of the non monotonic behaviour of scalable coding coupled with channel coding, and the possibility of eliminating the enhancement data flow.

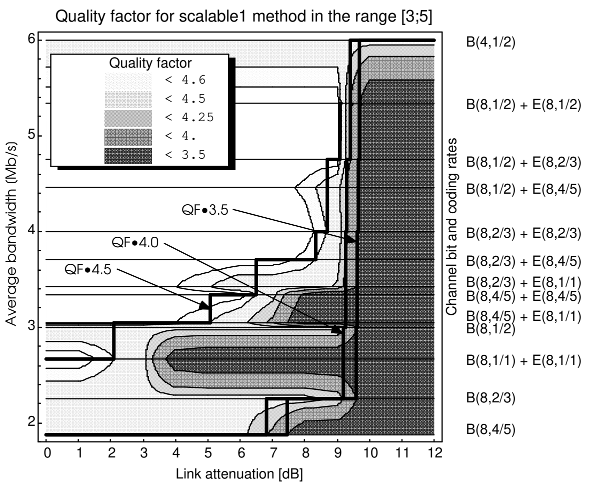

Figure 3. Isoquality regions in the attenuation/ bandwidth plane, and paths of constant QF for scalable1 MPEG-2 coding (see Table 1).

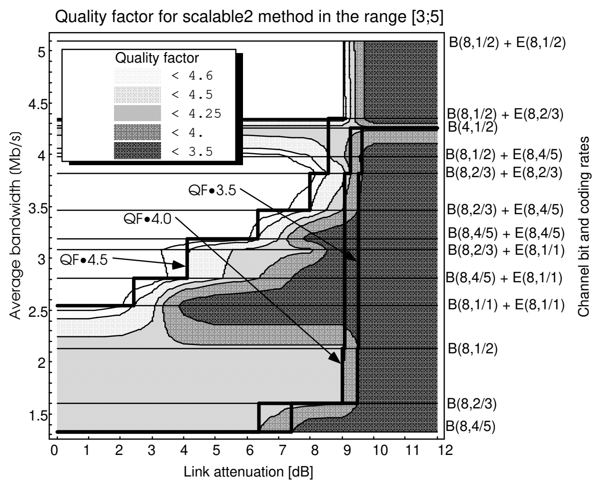

Figure 4. Isoquality regions in the attenuation/ bandwidth plane, and paths of constant QF for scalable2 MPEG-2 coding (see Table 1).

The number of possible combinations of channel bit and coding rates for the scalable video codings is higher than in the case of the non-scalable one (13 versus 5). Having at its disposal a greater number of possible choices for bandwidth and channel codings, the transmitting station can make a better job of choosing the channel bit and coding rates, while keeping the average bandwidth occupation as small as possible, and maintaining the requested video quality.

This is particularly true when low qualities are requested, that is when the target QF is 4.0 or 3.5. Indeed, in this case, the base flow alone is able to deliver a quality better than 4 with both the scalable codings considered, so the average bandwidth occupancy is generally smaller when using scalable codings than the non-scalable one where all the video information is sent to the receiver, no matter what the minimum tolerated video quality is.
Figures 5, 6, and 7 compare the performance of the three MPEG-2 coding methods used. For each video coding, given a target QF, the upper boundary of the bandwidth occupied by the transmitting earth station is depicted for a given attenuation. An MPEG-2 coding method thus performs better than another one if it uses less bandwidth for a given attenuation value. Even if it depends on the attenuation value, the graphs nonetheless give a clear idea of what is the best video coding method in most conditions.

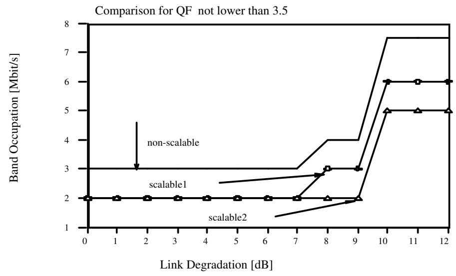

Figure 5. Bandwidth occupation by the transmitting earth station when maintaining a QF not less than 3.5 for the three video codings considered.

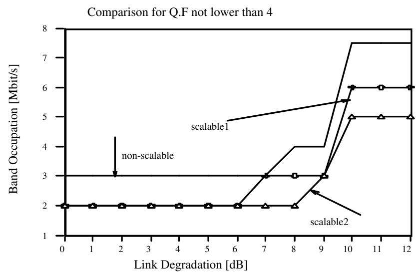

Figure 6. Bandwidth occupation by the transmitting earth station when maintaining a QF not less than 4 for the three video codings considered.

In Figures 5 and 6, in fact, the scalable2 method is clearly the winner, because it uses significantly less average bandwidth than the other methods to obtain the same QF. The main reason why scalable coding performs better than the non-scalable one — for target QF of 3.5 and 4.0 — is that the former can drop the enhancement flow. This advantage over non-scalable coding is more significant when the ratio between the base and the total flow rates is lower. The situation is less definite when we consider high video quality. Figure 7 illustrates the performance of the three MPEG-2 codings when the target quality is not lower than 4.5. In this case, none of the two scalable video codings can drop the enhancement flow. The non-scalable coding method takes advantage of a smaller intrinsic overhead (hence a smaller average bandwidth occupancy), but on the other hand the transmitting station has greater flexibility in handling the scalable codings, because it can attribute different channel bit/coding rates to the base and the enhancement flows. The only video coding usable for attenuations greater than 10 dB is the non-scalable one, but this result is not particularly significant as it is a consequence of our experiment constraints, which limit the average bandwidth to 7.2 Mbit/s.

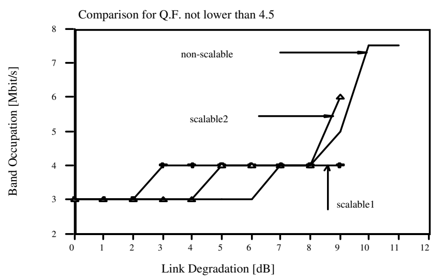

Figure 7. Bandwidth occupation by the transmitting earth station when maintaining a QF not less than 4.5, for the three video codings considered.

V. CONCLUSIONS

The experiment gave some interesting results. First of all, when the transmission policy is to keep a medium quality (quality factor not lower than 3.5 or 4) at the MPEG-2 receiver, a scalable coding is advantageous. The advantage increases with the ratio of the average total bit rate over the base flow bit rate. A higher bandwidth occupation allows a higher data redundancy, so it gives a better video quality for a given value of the satellite channel attenuation. Scalable methods take advantage of dropping the enhancement flow, when necessary, and obtain a quality factor not lower than 4 even at deep fade levels. When higher quality factor values are required, the intrinsic overhead of scalable methods is balanced by the possibility to protect suitably the two flows, and results are comparable with non-scalable method.

Due to problems of space, this paper does not present the results about a further protection of the headers' field using a BCH code. We can summarise by saying that the experiment showed that the headers' protection did not produce any appreciable difference in the resulting average QF, because the better protection of the headers is compensated for by the spread of erroneous bits produced by the data interleaver, which is needed by the BCH decoder to recover all the errors. We are aware that the results of this study do not have a general validity, because they are related exclusively to the traces of the video sequence we examined, and, moreover, the choice of the video encoder parameters to produce the MPEG-2 traces was obliged by the limits of the satellite channel we had at our disposal. Nevertheless, the results obtained encourage us to continue the investigation of the transmission of MPEG-2 video codings on a noisy satellite link, and we think that the scenario presented may be helpful for the design of the earth stations and the payload required for VBR video transmissions.

REFERENCES

[1] ISO/IEC 13818-2, ITU-T Rec. H.262, "Generic coding of moving pictures and associated audio information: video", 1995.

[2] A. Webster, C. Jones, M. Pinson, S. Voran, S. Wolf, "An objective video quality assessment system based on human perception", Proceedings of SPIE Human Vision, Visual Processing and Digital Display IV, 1-4 February 1993, San Jose, California, Vol. 1913.

[3] N. Celandroni, E. Ferro, F. Potortì , A. Bellini, F. Pirri: "Practical experiences in Interconnecting LANs via Satellite", ACM SIGCOMM Computer Communication Review", Vol. 25, No. 5, October 1995, pp. 56-68.

[4] N. Celandroni, F. Potortì, S. T. Rizzo: "An inexpensive rain fade countermeasure technique for DA-TDMA satellite systems", proceedings of IEEE GLOBECOM' 96, London, Nov. 18-22 1996, pp. 1000-1004.

[5] N. Celandroni, E. Ferro, F. Potortì: "Quality estimation of PSK modulated signals", IEEE Communications Magazine, July 1997, pp. 50-55.

[6] N. Celandroni, E. Ferro, F. Potortì: "Satellite bandwidth allocation schemes for VBR applications", CNUCE Report C94-24, December 1994.

[7] N. Celandroni, M. Conti, E. Ferro, E. Gregori, F. Potortì: "A bandwidth assignment algorithm on satellite channel for VBR traffic", to appear in Intern. Journal of Satellite Communications.

---

* Work carried out in the framework of the Italian co-ordinated project "*Advanced applications for next generation packet-switching networks*" funded by the National Research Council (C.N.R.)

(1) The authors would like to thank Videotime for providing the movie.

(2) Bose-Chaudhuri-Hocquenghem