DRS COMPRESSION APPLIED TO MPEG-2 VIDEO DATA TRANSMITTED OVER A SATELLITE CHANNEL*

Nedo Celandroni, Erina Ferro,
Francesco Potortì
Area della ricerca CNR
CNUCE - via Vittorio Alfieri 1
I-56010 Ghezzano, Pisa
{N.Celandroni,E.Ferro,F.Potorti}
@cnuce.cnr.it

Antonio Chimienti,
Maurizio Lucenteforte
CSTV, Institute of CNR
Strada delle Cacce 91
I-10135 Torino
{chimienti,lucente}@cstv.to.cnr.it

Abstract

This paper shows how the adoption of a Dynamic Rate Shaping (DRS) to variable bit rate (VBR) MPEG-2 (Motion Pictures Expert Group-2) coded video data, in addition to different levels of data protection, adds more flexibility in the way a video data stream can be sent over a faded satellite channel. DRS allows the satellite bandwidth required to be continuously trimmed, within a certain range. Also, some bandwidth can be saved when the video quality is maintained as close as possible to a required level during changing fade conditions. The performance of DRS applied to a VBR MPEG-2 coded video stream is shown. An application of this technique is discussed, and some criteria are given which help the user to choose the most suitable way to send video coded data, according to the service quality requirements.

1: Background

In [1] we analysed the relationships between bandwidth occupancy, satellite link degradation, and the quality of a video stream coded as the Motion Pictures Expert Group-2 (MPEG-2) standard [9], and sent over a satellite channel affected by unfavourable propagation conditions (fades). In satellite transmissions the signal attenuation due to rain imposes that fade countermeasures have to be adopted to ensure an acceptable level of link availability with a reasonable channel quality. The fade countermeasure used in our past experiment was based on the dynamic combination of the channel coding and bit rates variation applied to the data sent via satellite. This technique was made possible by the hardware equipment at our disposal (TDMA controller, burst mode modem, and codec, which together constitute an *earth station*), which had already been used in past experiments [5]. For the present experiment we used an emulated geostationary satellite channel affected by white Gaussian noise. The satellite emulator introduced the correct round trip time and the amount of noise needed to emulate fading situations. The experiment used the traces of the movie "The sheltering sky"[1], coded according to the MPEG-2 standard. The samples obtained were transmitted between two real earth stations connected together via a satellite emulator. The received sequences, relative to the various transmission parameters needed to cope with the channel degradation level, were then analysed to evaluate the resulting video quality. The scenario produced showed the actual possibility of MPEG-2 video transmission over a degraded satellite link with limited resource requirements, by adopting a fade countermeasure technique based on both channel coding and bit rate variations.

The logical continuation of the work in [1] was to study how to trim the data flow more gradually so that a wider variety of channel conditions could be addressed. The goal was to improve the results obtained in [1] by trimming the bitstream so that it could be better protected, while still occupying the same bandwidth, and to find out how to put a variable bit rate (VBR) transmission over a channel with limited bandwidth. Trimming can generally be performed by any intermediate node that knows the real channel conditions. Moreover, for practical reasons,

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

[1] The authors would like to thank Videotime s.p.a. for providing the movie used in the simulations.

the computational load must be very low. With these requirements in mind we designed a rate shaper that is able to trim the MPEG-2 bitstream. This rate shaper should be considered as a parser because it does not perform a decoding and requantization process (this would be too expensive), but simply parses the bitstream and selectively discards an amount of information proportional to the percentage of the flow reduction required. The technique used is called dynamic rate shaping (DRS). Unlike other works in which optimality conditions were imposed on the rate shaper [8] our solution is very simple. Since it is driven only by the flow rate, it can trim the bitstream very finely in order to dynamically adapt it to the bandwidth requirements. The traces on which the DRS and the fade countermeasures technique are applied are the same as the ones produced for the experiment in [1]. This paper only presents the results relevant to the simulations performed over a 60 second video sample produced without the SNR scalability, i.e. a VBR single-stream coded with an average rate of 2.37 Mbit/s, and generated with the quantization step set to 6 in order to maintain the total channel bit rate within the limits of the hardware at our disposal.

The paper is organised as follows. Section 2 presents the DRS implementation and its performance. Section 3 shows the results of the simulation of the transmission over a faded satellite channel of the VBR video stream, with and without DRS. Section 4 depicts possible applications of this novel technique, and indicates interesting directions for future work.

2: Dynamic Rate Shaping: implementation and performance

Rate shaping is a technique to adapt the rate of compressed video bitstreams to variable rate constraints. It provides an interface between the encoder and the network, which allows a perfect match of the encoder's output with the network's quality of service characteristics.

The approach does not require any interaction with the encoder, so it can be applied in all precoded and stored video services (i.e. video-on-demand video systems) where a wide variety of video programs may be available. In these applications a multiresolution coding with too many layers would be undesirable, due to the loss in channel efficiency. In general, a rate shaper works in such a way that an input MPEG-2 stream with instant bit rate of R(t) Mbps is squeezed in order to achieve a constrained instant bit rate of $R_T(t)$. The instant output bit rate, R'(t), is defined as $R'(t) \cong R_T(t)$ if $R_T(t) < R(t)$, and as $R'(t) = R(t)$ if $R_T(t) \geq R(t)$. In a general situation, R(t) and $R_T(t)$ can either be constant or variable, and are associated with the characteristics of the source and channel.

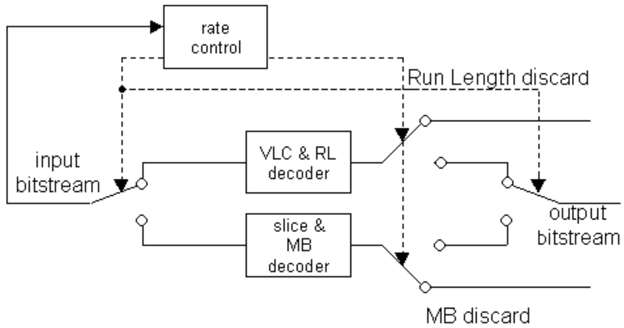

Figure 1. Rate shaping module scheme. VLC stands for variable length coding.

Assuming a general hybrid motion-compensated block-based transform coding technique, there are two fundamental ways to reduce the rate: i) to adopt a transcoder, composed of a cascaded decoder and encoder; ii) to selectively discard coded information: transform coefficients, macroblocks, slices, etc. Although different levels of complexity are allowed in a transcoder [4], the first approach is significantly more complex than the second one because it needs a coefficient requantization that includes Huffman decoding, run-length decoding, dequantization, requantization, run-length (RL) encoding, and Huffman encoding. Despite this, an additional reconstruction error will occur, caused by a drift between the decoder's prediction signal and the original (not requantized) prediction signal, which will introduce signal degradation and coding delay. Since the scheme adopted must be as simple as possible in order to install several cheap modules in suitable points of the distribution network, in this paper we consider a method based on the second approach, whose implementation is depicted in Fig. 1. To further justify our choice, note that we refer to an unbalanced coding scheme (the MPEG-2 encoding process is more complex than the decoding one) and thus we generally assume pre-recorded video traces. We use a rate shaping algorithm based on a selective information discard approach, applied to MPEG-2 VBR sources, which tests several output constrained bit rates in order to match the channel bandwidth requirements. A very important feature of this method is that the resulting bitstream syntax is faithful to the MPEG-2 standard. Rate reduction is obtained by eliminating a contiguous string of DCT coefficients at the end of each block. When, with a small target bit rate, the rate-control is not able to track the bit rate reduction, selective slice discarding is performed, for bidirectionally-predicted pictures only.

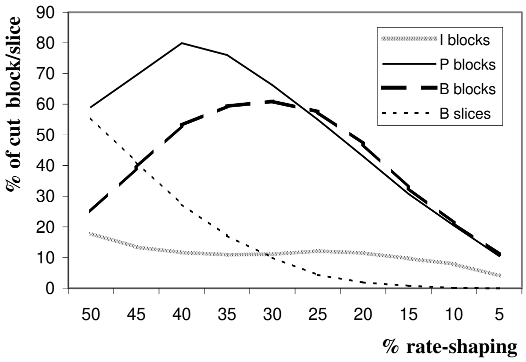

Figure 2. Cut macroblocks/slices versus DRS compression rate.

The underlying idea is to use a rate control in a similar way to the one used in the Test Model 5 [10] of MPEG-2, using the parser in order to select the information to discard. A picture target bit allocation is thus calculated, based on the picture type and the target bit rate. The rate control mechanism sets the bits available to each macroblock (MB), by comparing the overall spent bits with the target bits by means of a virtual buffer associated with each of the three picture types defined in MPEG-2 (intra-pictures I, predicted pictures P, and bidirectionally-predicted pictures B). Within each picture type, the macroblock bits are distributed to the various blocks (luminance and chrominance), by assigning them the number of bits that can be re-transmitted.

In normal conditions the rate shaper parses and counts the input bitstream, by copying it in output. The discarding of the DCT coefficients starts when, at a block level, the number of input bits exceeds the target one determined by the rate control. In this situation, the last run-length code encountered is followed by an end-of-block code by discarding the remaining coefficients associated with the higher spectral frequencies. We require that at least one DCT coefficient remains in each block in order to avoid re-calculating and re-coding some syntax fields, such as the coded block patterns (which indicate which blocks in each macroblock are included in the bitstream), DC prediction, where DC is the DCT coefficient associated with a frequency of zero (mean value), and the macroblock address within the current slice. In this way, breakpoint values range from 1 to 64. The fundamental elements of this method are a bitstream parser and a rate control module. The DCT coefficients' discard is applied to I, P and B pictures.

Sometimes discarding DCT coefficients is not sufficient to track the shaping requirement. This is especially true when the source input bitstream has a high compression ratio, as in our case. In this situation the source coder applies a coarse quantization and only a small number of coefficients are coded in the original bitstream. The basic idea is that the slice is completely discarded when the picture target rate is lower than a threshold that depends on the local sequence activity, and the number of bits used for all slices up to the current one exceeds the target value by a given percentage. In this situation only the first and the last macroblocks are maintained.

The selective discarding of coded information has a predictable impact on the resulting video quality. Discarding the higher frequencies corresponds to a low-pass filtering process. Since it is applied to I and P pictures as well, a drift error associated with the wrong temporal prediction is added to the error caused by the coefficient discard. At high compression rates, the shaping reduces the coded signal to the side information fields and the lower frequency DCT coefficients, thus causing an annoying block effect artefact. The slice discard in B pictures, on the other hand, has a more evident effect on the decoded pictures, since the decoder retrieves the information of the missed macroblocks as if they had been skipped (i.e. the motion vector associated with the first macroblock in the slice is used to perform the motion-compensation of the missed macroblocks).

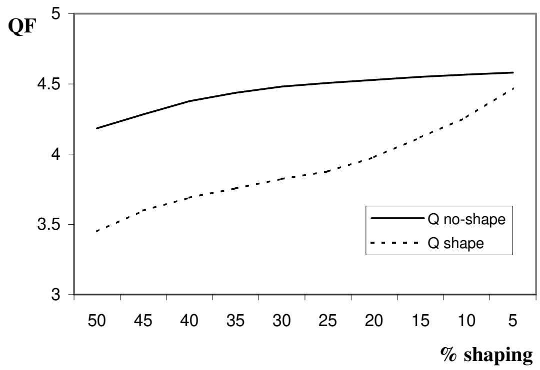

Figure 3. QF comparison between the sequence decoded at different levels of shape (5% to 50%) and the corresponding standard simulation coded at the same bit rate.

To measure the loss of video quality of the proposed method, the decoded video can be compared with the original sequence at different levels of compression. The quality of the video may be computed by using an objective video quality assessment method [7] that emulates the HVS (Human Visual System). With this method, the resulting quality factor (QF) is strictly correlated with the subjective mean opinion score: imperceptible (5), perceptible but not annoying (4), slightly annoying (3), annoying (2), and very annoying (1). All our results will be expressed by using this quality factor assessment method.

To understand how the rate control performs the rate reduction, Fig. 2 plots the rate of block and slice cutting for the above specified rate-shaping range. As can be seen, for high percentages of rate shaping, a high B-slice cutting is the only way to track the rate control target. For this reason the B-blocks and the P-blocks cutting rate is rather low. This rate increases for medium shaping rates (20%-40%), when the rate-reduction initially leans on block cutting. Another significant fact is the matching of video shaped at a rate of R'(t) with an unshaped trace coded at the same medium rate. Figure 3 compares the video quality between the sequence decoded from bitstreams obtained from the original at different levels of shape (5% to 50%) and the corresponding standard simulation coded at the same bit rate.

3: The transmission over the satellite channel: results

On satellite channels the signal *attenuation* is defined as the dB difference between the power received under clear sky conditions and the power currently being received. Unfavourable propagation conditions (*fades*) are a major problem in transmissions above 10 GHz in that they attenuate the signal and increase the noise level thus worsening the channel signal-to-noise ratio (CSNR), which is the indicator of the channel quality. We express CSNR in terms of E_b/N_o (bit energy to one-sided noise spectral density ratio). Hereafter *link degradation* will denote the difference between the reference (12 dB) and the current value of E_b/N_o. The reference value, relative to clear sky conditions, allows a BER of 10^{-8}. All the tests in our experiment were made in steady state conditions, as far as the link characteristics are concerned, that is, the QF of the movie was estimated while maintaining a constant satellite link quality. This procedure prevents biased results due to video scenes having specific characteristics, because the quality factor measured is averaged over a video sequence of a significant length. The mean QF and mean bit rate we used are 4.66 and 2.37 Mbps, respectively.

We measured the video quality using the QF parameter, and we analysed its relationship with the bandwidth needed by the video bit stream, averaged over the video sample duration, while varying the data coding and bit rates as well as the rate shaping. Hereafter *bandwidth of the video stream* will be used to mean the average share of the link bandwidth which would be used by an uncoded sequence sent at the maximum bit rate on the satellite link, without any DRS. Figure 4 reports the QF versus DRS for various data sequences in different channel degradation and coding conditions.

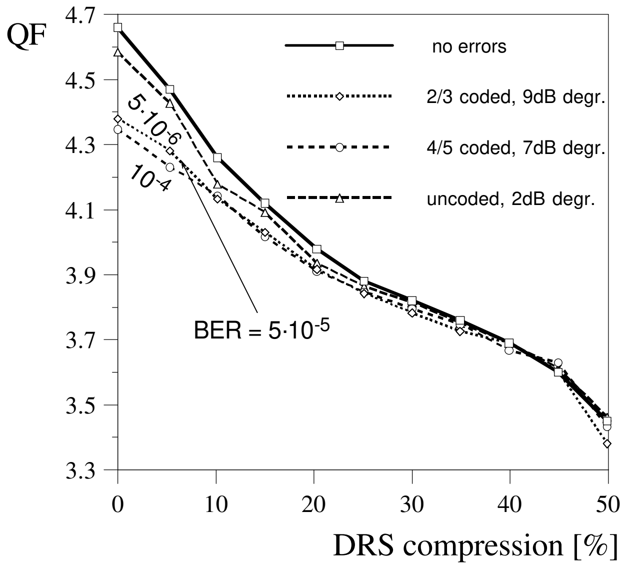

Figure 4. QF versus DRS compression levels for various coding rates and channel degradation conditions.

The relevant BERs, ranging from 10^{-4} to zero (no errors case) are shown as well. The figure highlights that all the curves, whose QF is spread over the interval 4.35-4.66 when data are not compressed (DRS = 0), exhibit practically the same QF for DRS from 35% and above. Using high data protection, which wastes bandwidth, is thus not very convenient, when we need a high rate shaping on the video data stream. This result is confirmed if we consider the effect of DRS from an energy point of view, i.e. considering the QF versus link degradation with and without DRS. Figure 5 gives the QF versus the link degradation at different data coding rates, without any DRS. In all cases, we can observe a quasi-constant QF up to a certain degradation, then a knee follows, after which QF degrades steeply. A DRS compression applied in the leftmost part of each curve causes a small loss in signal power, while if it is applied after the knee some gain is made. As an example we consider the uncoded curve when no link degradation is present (E_b/N_o = 12 dB, 8 Mbps). A 50% flow compression with DRS is energetically equivalent to 3dB of channel degradation, because by halving the bandwidth a reduction of 50% of the signal power does not lead to any change in E_b/N_o. Such a compression by means of DRS makes QF fall to 3.45 (see Fig. 4), while a channel degradation of 3.65 dB is needed to fall to the same value (Fig. 5), with however a loss of 0.65 dB.

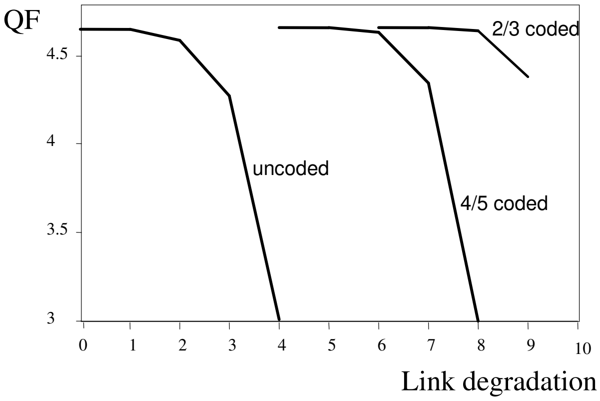

Figure 5. QF versus link degradation for different data coding rates at the bit rate of 8 Mbps. No DRS is applied.

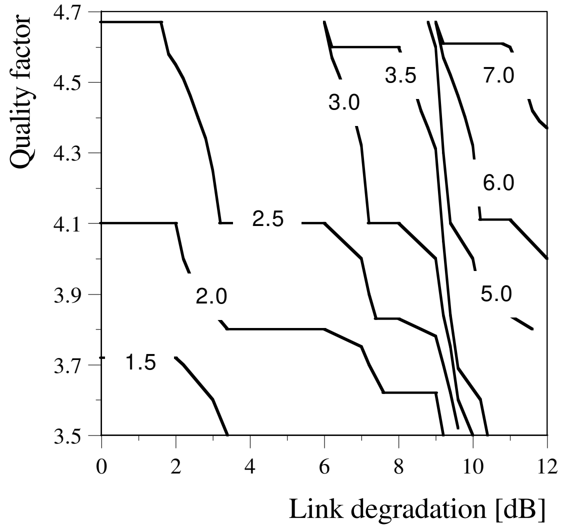

Figure 6. QF versus link degradation for various values of the bandwidth occupied.

On the other hand, if we start from 2 dB of degradation on the same curve, a compression of 50% allows a 1.35 dB gain. Starting from 7 dB of degradation with a flux 4/5 coded, a further 0.7 dB of degradation is sufficient to make QF go to 3.45, with a gain in the compression case of 2.3 dB. A problem with the method proposed in [1], which does not use any rate shaping of the video stream, is that the number of available bit and coding rates is usually small for a satellite modem. Even with the ample choice of our modem, there are attenuation regions where a trade-off between video quality and bandwidth occupancy is not possible. In Figures 6 b), d), and f) the three bold (upper) lines show the bandwidth occupancy for QFs not lower than 4.5, 4, and 3.5, respectively. For each attenuation value, the curves are obtained by using the combination of coding and bit rates which yields the smallest bandwidth.

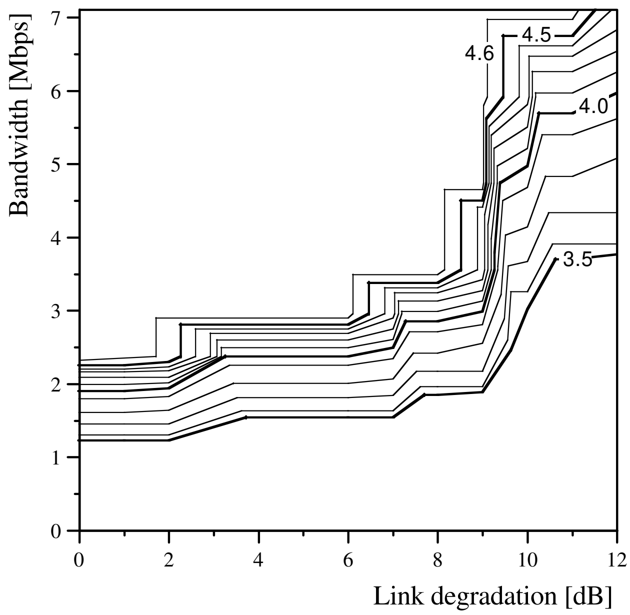

Figure 7. Bandwidth occupied by the video stream versus link degradation for various QFs.

There is no trade-off possible in the ranges $0 \div 2.3$ dB and $3.7 \div 6.5$ dB, for example, because whatever the desired QF is, the channel coding rate to be used is 1/1 in the first range, and 4/5 in the second one, with the result that a higher QF than required is obtained in many cases. If, for example, we want to closely match a target QF, then it would be more efficient to have a method of continuously varying some parameters in order to get a lower QF, and a corresponding bandwidth gain, for any link attenuation. By using DRS, such a continuous variation can be obtained for very wide bandwidth ranges.

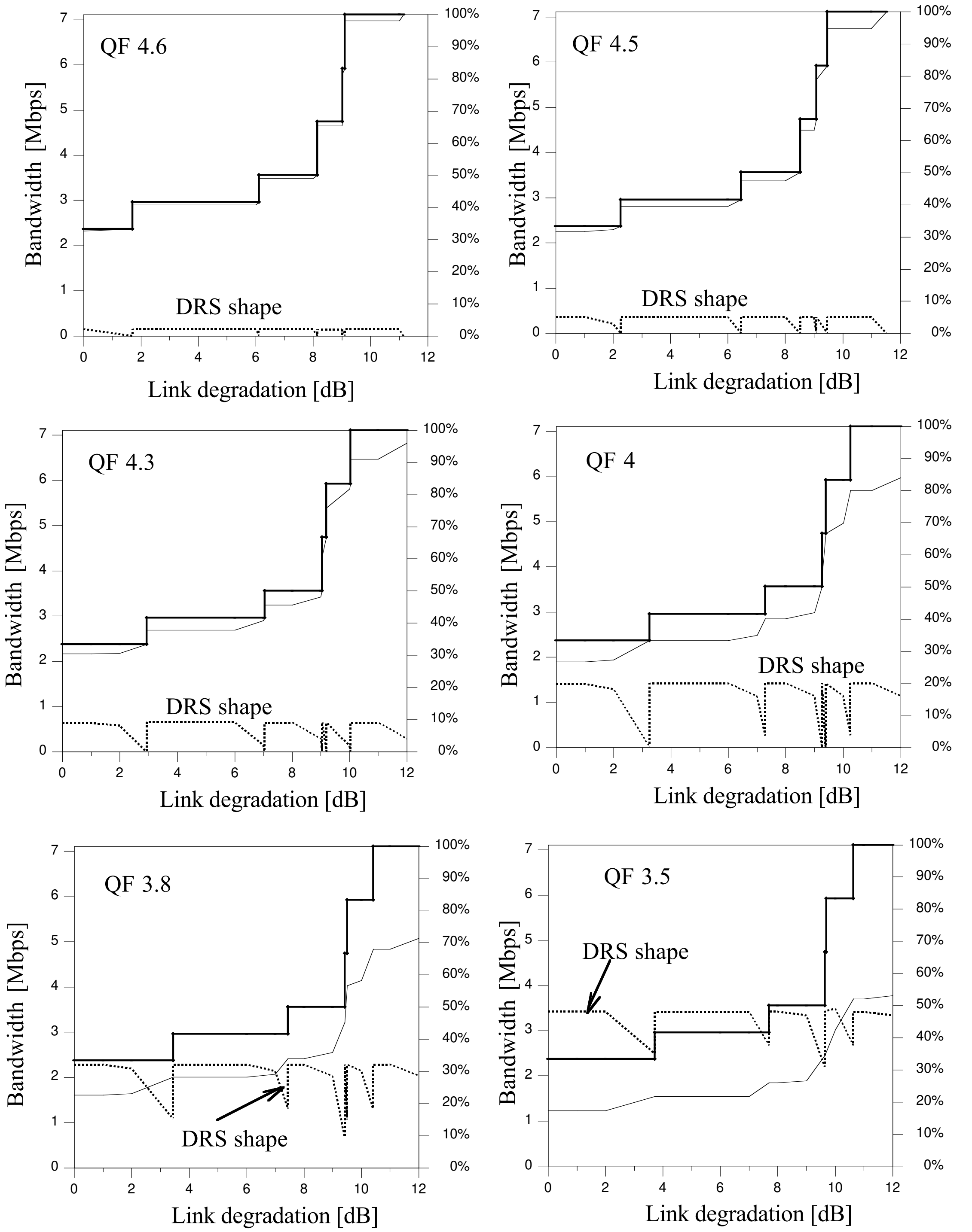

Figure 8. Bandwidth occupancy for various QFs in different fade conditions, with and without (bold lines) DRS. The rightmost scale of each figure represents the compression percentage.

Figure 8 highlights the bandwidth gained by using DRS in the range $0 \div 12$ dB of degradation, for some given values of QF. In each of the six pictures, the stepwise curve is obtained without DRS, while the lower bandwidth curve is obtained by using the combination of bit rate, coding rate and DRS compression that yields the smallest bandwidth for each degradation level. The gain is small for target QFs close to the maximum (4.66) as the compression is by necessity small, but it becomes noticeable for lower QFs.

Figure 6 shows the best QF that can be obtained by using a constant bandwidth versus link degradation, for values of the bandwidth occupied ranging from 1.5 to 7 Mbps.

Figure 7 shows a more detailed scenario of DRS application for QFs in the range from 3.5 to 4.6. It is now possible to notice that the use of DRS, besides giving a bandwidth saving (Fig. 8), offers more flexibility as well. In fact, in practically all regions of Fig. 7 it is possible to trade bandwidth for QF for each link degradation.

4: Applications and future work

The problem now is how to assess the convenience of using the techniques presented above for sending an MPEG-2 encoded bit stream over a satellite link, for different levels of link degradation and/or traffic conditions. Using a TDMA technique on the satellite channel with variable data coding and bit rates, the occupancy of a data packet in the time space depends on the redundancy applied to the data. A major link degradation, which imposes a higher data redundancy, and thus a higher occupancy in the time space, requires an increase in resource allocation as if there were an increase in the amount of traffic. The factors that contribute to the convenience evaluation of the various solutions are: the average bandwidth occupancy, the bandwidth variability and the simplicity in system development. The second item is very important both if a traffic statistic multiplexing is performed on the channel, and if dynamic techniques for the demand assignment of the bandwidth are employed.

Let us assume that we are operating a dynamic adaptation in order to control the transmission of the data stream, and that the flux trimming is operated by an intermediate agent, integrated in the earth station, which is able to evaluate the channel conditions and make the adaptive tune-up on the data video stream. This situation is more flexible than using an end-to-end feed-back technique, where the end user evaluates the quality perceived and sends that information back to the source, which can then operate the adaptation needed to guarantee the video quality required by the user. The major drawback of end-to-end feedback is that the video regulation must be operated at the source level (which is generally separated from the satellite transmission earth station) in view of the suitable data protection (coding and bit rates) which the earth station will operate. Instead it is much more convenient to be able to play with flux trimming and data protection at the earth station, where all the information about the overall link quality (including up- and down-links) and the amount of traffic directed towards the satellite is available. Furthermore, in multicast transmissions that use multispot antennas on board the satellite, the data stream must be replicated and sent on each link, by the intermediate agent. In this case, a system like the one we are considering allows us to trim the video stream to each individual condition of the link. This, in fact, may be affected by different down-link attenuations due to different atmospheric situations in zones sufficiently far apart, as happens in multi-spot coverage. Once the video source has been coded with a given quantization step, the sending earth station is able to operate the data protection required according to the channel fade conditions.

Figure 8 shows that using DRS is always advantageous, and it allows a bandwidth saving which increases as the QF decreases. Note the utility of using the curves which give the QF and the bandwidth needed as functions of the link degradation. This in fact allows us to employ a link quality estimator, which is much cheaper and quicker in response than a QF meter. The QF values shown are averaged over the entire duration of the movie, whereas quasi-instant measurements of QF, as required by dynamic adaptive techniques, may accidentally give spurious values, which are undesirable in a control system. This happens, for example, when some bits in the header fields of a video sequence are corrupted. On the other hand, link quality estimation, like the BER, is much more stable and can be quick and accurate [2].

When considering this type of dynamic adaptive techniques, one should consider that the problem is complicated by a non-negligible delay in the control loop, and by the fact that the attenuation varies with time in a complex way [11]. Consequently, it is necessary to devise a way of using knowledge of the instantaneous attenuation for choosing the coding and DSR compression which are necessary to obtain a given QF. The problem is analogous to transmitting data with a given bit error rate [12].

In order to choose the most suitable technique to transmit video data over faded satellite channels, we must stress that, at least when the rate of corrupted bits is very low, the optimisation of the average satellite bandwidth occupied by the video stream is reached with a VBR sequence obtained with a quantization step that is suitable for achieving the target QF. This is only applicable to1 end-to-end feedback systems, which have the drawbacks we have pointed out. In systems like the one we have presented, where the data protection is chosen at the earth station level for each individual link, significant advantages may be obtained by handling scalable video sequences such as a double flux of data (as shown in [1]) and/or the adoption of DRS techniques. The reduction in bandwidth, obtained by using DRS, may have practical advantages because it introduces more flexibility in the bandwidth occupation. On the other hand it does not produce any actual gain, in the sense that the QF resulting from a DRS application is lower than the one obtained with a quantization step that allows the same bandwidth occupation as the one of the compressed stream. However, if the quantisation step has been fixed by the source, the adoption of DRS, used in addition to the channel coding techniques, leads to more flexibility in the way a video data stream can be sent, because it allows the intermediate agent to continuously calibrate the bandwidth within a certain range. If the goal of using adaptive techniques is to keep the QF as close as possible to a desired value, a certain amount of bandwidth is certainly saved, because the bandwidth versus link degradation curve is smoother than if only the code and bit rate step variations are adopted, as in [1]. Furthermore, we have seen that when DRS is used on video sequences, which are subsequently heavily corrupted by noise, the video bandwidth reduction may allow a better signal protection, which gives a better QF as a result.

Several factors contribute to the choice of the most suitable video transmission technique. We must consider, for example, that a satellite link is generally in a non-fade state most of the time, that the bandwidth variability is significantly reduced when DRS is applied, and finally that the DRS technique is easy to implement. Future work will be to investigate the DRS applied to the VBR part of a double flux of data in the scalable case of MPEG-2 video coding, already considered in [1].

References

[1] N. Celandroni, E. Ferro, F. Potortì, A. Chimienti, M. Lucenteforte, and R. Picco, “MPEG-2 coded video traces transmitted over a satellite link: scalable and non-scalable solutions in rain fading condition”, *Multimedia Tools And Applications*, Vol. 10, N. 1, pp. 73-97, January 2000.

[2] N. Celandroni, E. Ferro, and F. Potortì, “Quality estimation of PSK modulated signals”, *IEEE Communications Magazine*, Vol. 35, No. 7, pp. 50-55, July 1997.

[3] Roger L. Freeman, “Reference Manual for Telecommunications Engineering”, Second Edition, 1995 update.

[4] P. N. Tudor, and O.H. Verner, “Real-time transcoding of MPEG-2 video bit streams”, *Proc. of International Broadcasting Convention*, 12-16 September 1997, Amsterdam (NL).

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

[6] Y. Ho, C. Basile, and A. Miron, “MPEG-based video coding for digital simulcasting”, *Proc. of International Workshop on HDTV*, Kawasaky, Japan, 1992.

[7] A. Webster, C. Jones, M. Pinson, S. Voran, S. Wolf, and “An objective video quality assessment system based on human perception”, *Proc. of SPIE Human Vision, Visual Processing and Digital Display IV*, 1-4 February 1993, San Jose, California, Vol. 1913, pp. 15-26, 1993.

[8] D. Anastassiou, and A. Eleftheriadis, “Meeting Arbitrary QoS Constraints Using Dynamic Rate Shaping of Coded Digital Video”, *Proc. of 5th International Workshop on Network and Operating System Support for Digital Audio and Video*, Durham (USA), pp. 95-106, 1995.

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

[10] ISO/IEC/JTC1/SC29/WG11, Coded representation of picture and audio information - Test Model 5, April, 1993.

[11] N. Celandroni, and F. Potortì “Modeling Ka band scintillation as a fractal process”, IEEE J. Select. Areas Commun., Vol.17 No.2, February 1999.

[12] N. Celandroni, and F. Potortì “Fade Countermeasure Using Signal Degradation Estimation for Demand-Assignment Satellite Systems”, to appear on the Journal of Communications and Networks in August.