FORWARD ERROR CORRECTION FOR VIDEO CODING
A common solution

João Tavares and Antonio Navarro
Telecommunications Institute, Aveiro University, University Campus, Aveiro, Portugal

Keywords: Wireless video, Video common protection, Joint source-channel coding, MPEG-4, H.263

Abstract: Joint source-channel coding solutions has proved to provided better performance than dealing with both coding processes separately. Such improvement is achieved by jointly minimizing at the source encoder the channel effects. In this paper, we present an adaptive scheme for forward error protection of any video coding standard. The channel coding rate changes according to the channel bit error rate (BER). The results are impressive. For instance, a PSNR gain of about 16.7 dB is obtained at BER=10^{-2} for "Foreman" video sequence, encoded either by H.263 or by MPEG-4 and protected using the Common scheme in comparison to the unprotected case. As our proposed scheme is common to all video standards, it obviously provides some video quality degradation but still acceptable. We have assessed the quality degradation of the Common solution in comparison to the optimal scheme of protection, which uses a finite number of channel codes. Additionally, we propose a protection solution tailored to H.263/MPEG-4 video coding with an average PSNR improvement of about 0.2 dB relatively to the above mentioned Common solution.

1 INTRODUCTION

Nowadays, commercially available mobile multimedia terminals based on GSM/GPRS are not able to provide video based real time multimedia applications with accepted quality of service. Although, the 3rd and 4th mobile network generations overcome the bit rate limitation, wireless video transmission offers other important negative aspects namely caused by the deep fast fading effects. The random behavior of the quality of the received signal in different channel conditions will cause enormous damages on the reconstructed video. Compressed video is extremely vulnerable to transmission errors since a single erroneous bit may lead to a considerable number of frames to be incorrectly decoded. This behavior channel was expected in face of a source rate close to or even greater to the channel capacity as stated by Claude Shannon in 1948 (Shannon, 1948). However, in practical narrow band-limited channels, the capacity is low and consequently the source bit rate would be required to be less than its entropy. Therefore, the rate-distortion is a possible solution (Ortega and Ramchandran, 1988) which basically consists of discarding some source information. The source distortion is minimized if it is discarded the less important video source information. Besides, the communication system performance can be further increased by joint design of source-channel coding (Daut and Modestino, 1983). It is sometimes worth moving source bit rate to the channel to increase the overall communication performance.

Video coding standards were developed for error free transmission since prediction and statistical coding techniques have been adopted. Some papers published in the recent years have investigated both compatible standard and non-standard schemes (Wang, 2000) to improve error robustness of video coding. They can be classified as Forward Error Correction (FEC), Error Concealment (EC), Automatic Request (ARQ), Synchronization Markers (SM) or a combination of them. FEC schemes are usually part of the multiplexing standards such as H.223 and H.323.

This paper proposes some solutions for the problem of joint source channel coding in the context of FEC considering video communications over wireless environments. In this context, Frossard and Verscheure (Frossard and Verscheure, 2001), and Stuhlmuller et al (Stuhlmuller et al, 2000) proposed a solution for the joint source-channel coding problem for MPEG-2 and H.263 video, respectively. Gnavi et al (Gnavi et al, 2002) optimally designed some coding parameters for H.264 algorithm. More recently, in the case of image transmission, Grangetto et al (Grangetto et al, 2004)

Navarro A. and Tavares J. (2004).
FORWARD ERROR CORRECTION FOR VIDEO CODING - A common solution .
In *Proceedings of the First International Conference on E-Business and Telecommunication Networks*, pages 331-336
DOI: 10.5220/0001401603310336

also solved a code rate allocation problem and demonstrated a solution employing SPIHT and JPEG2000 as source codecs. In our research work, we follow a slight different approach from the above mentioned literature. As channel codes should be applied dynamically according to instantaneous channel characteristics, equal error protection schemes may reach very close performance of unequal error protection (Tavares, 2001, Navarro, 2002) and at the same time avoiding high complex codecs and providing a common solution covering several video coding standards.

The paper is organized as follows. Section 2 formulates the problem. Section 3 presents the test conditions mainly concerning with source coding parameters. In Section 4, we devise a protection common solution as well as an H.263/MPEG-4 tailored solution. Finally, Section 5 concludes this paper.

2 PROBLEM FORMULATION

Signal distortion is caused by loss coding and lost/error transmission. Let $X_{i,j}$ and $\hat{X}_{i,j}$ be the original and the reconstructed macroblock (MB), respectively, where index j denote the MB number in image i. Both indices are integer numbers. The total distortion is usually a function of the mean square error (MSE) given by the series,

$$\lim_{N\to+\infty}\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}(X_{i,j}-\hat{X}_{i,j})^2 \tag{1}$$

where M and N are the number of MBs in each image and the number of images in the video sequence, respectively. Usually, the series is converted to a summation with N equal to the distance between two consecutive INTRA pictures. The problem consists of minimizing (1) constrained to a given constant or even variable total resources, R_T,

$$\sum_{i=1}^{N}\sum_{j=1}^{M}R_{i,j}\le R_T \tag{2}$$

where $R_{i,j}$ is the bit rate assigned to MB $X_{i,j}$. Nevertheless, (1) can be decomposed into a double summation of two terms, one related to source distortion and the other one to channel distortion,

$$\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}(X_{i,j}-U_{i,j})^2+\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}(U_{i,j}-\hat{X}_{i,j})^2 \tag{3}$$

where $U_{i,j}$ is the MB resulted from the quantization process. Despite the first term in (3) is well known analytically (Joshi and Fisher, 1995), the second term is a troublesome. Firstly, as the encoder makes use of statistical coding, the model of channel distortion has to take into account several error propagation phenomena. Secondly, once convolution channel codes are used, it is only known analytically the bit error bounds since its correction capability depends strongly on the source statistics. Thirdly, the right hand side terms should also model the recovering and concealment processed employed at the receiver. Finally, both terms are not completely independent as expressed in (3). Therefore, in order to increase the accuracy, we solved (1) through simulations.

3 TEST CONDITIONS

Two video coding standards, H.263 and MPEG-4, were used to access the performance of our forward error protection schemes.

The H.263 (ITU-T H.263, 1998) video coding standard is a low bit rate oriented and hybrid block-based algorithm where video frames can be encoded with either INTRA-frame or INTER-frame coding modes. The H.263 standard also includes more advanced techniques that enable better compression performance at low bit rates and better error resilience. They are introduced as optional annexes in versions 2 and 3 of the standard. All these annexes are negotiable between terminals ensuring that the decoder is able to cope with them. As the H.263 baseline decoder is mandatory, we have used it in our simulations without any optional annexes.

The MPEG-4 (ISO/IEC 14496-2, 2001) is the first multimedia standard that combines interactivity, natural and synthetic video, audio and computer graphics. The video coding algorithm is extended from the previous standards with the support for arbitrarily shaped video objects and coding tools. Additionally to the motion and texture information, the shape of an object can also be encoded.

The simulations were carried out considering four test sequences at QCIF (176*144) resolution, “Foreman”, “Coastguard”, “News” and “Container”, each containing 300 frames (only the “Foreman” sequence contains 400 frames). Each sequence was coded at 15 frames per second and the rate control has been enabled to produce constant bit rate

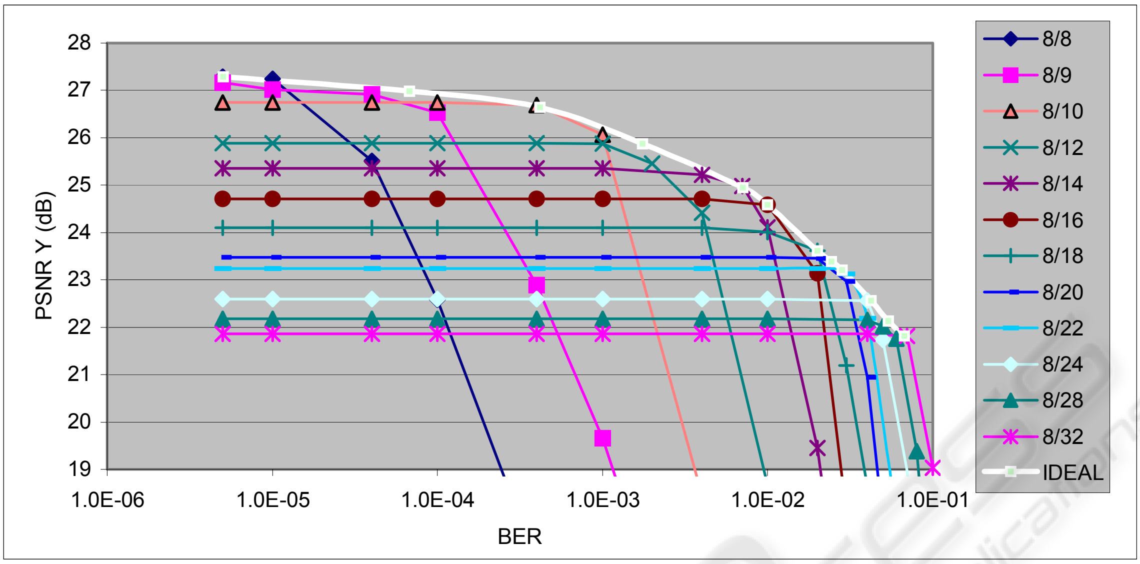


Figure 1: PSNR-BER curves for sequence “Foreman”, H.263 coded.

ranging from 16 to 64 kbps. One INTRA frame/VOP was coded every 2 seconds (30 frames). Therefore, in (1), N=30 and M=99. In H.263 coding, GOB start codes are used by our decoder as resynchronization markers and are searched in case of synchronization loss. In MPEG-4 coding, the packet length was adjusted according to the bit rate in order to have about 9 packets per VOP. Some of the error resilience tools in the MPEG-4 standard such as video packets, data partitioning, HEC and Reversible VLC have been considered. The RVLCs are used in the encoding process but not used in the decoding since are decoded in only one direction, forward direction. In MPEG-4 coding, each sequence was coded in a single scalability layer and objects are considered rectangular coincident with frames. Table 1 summarizes MPEG-4 encoding parameters.

The H.263 and MPEG-4 coded bit streams were then channel encoded with rate-compatible punctured convolutional (RCPC) codes obtained from a rate 1/4 - memory 4 mother code (Hagenauer, 1988), according to the proposed schemes as described in the following sections and transmitted over a binary symmetric channel at a constant bit rate of 64 kbps. As the channel code rate is decreased the source code rate must be decreased too in order to maintain constant the transmission bit rate. The advantage of RCPC codes is that each code in a family gives a different level of channel protection and all of them can be decoded with the same Viterbi decoder. These different levels of protection may be obtained from a given mother code using different puncturing tables and hence the level of channel protection may be adapted to channel conditions with minimal complexity. A hard decision Viterbi decoder is used to decode the received data. It is assumed that the Viterbi decoder is aware of the puncturing tables and the code rate. Also, a perfect reception of Video Object Layer header as well as time information (Temporal Reference in the case of H.263 and VOP modulo time and time increment for MPEG-4) is assumed since they can be protected by a sufficient powerful channel code. This is perfectly possible because these fields occupy just a few tens of bytes.

Table 1: MPEG-4 encoding parameters.

Source resolution	QCIF (176*144)
Source frame rate	30 Hz
Number of video objects	1
Number of layers	1 (single layer coding)
VOL frame rate	15 Hz
Target bit rate	16-64 kbps
Rate control	VM5+
Bits per pixel	8
INTRA period	30 frames
B-VOPs	Not used
Slice resynchronization	Enabled
Resynchronization packet size	Adjustable
Data partitioning	Enabled
RVLC	Enabled but not used
AIR	Disabled
NEWPRED	Disabled

Performances are evaluated in terms of image PSNR of the luminance (Y) component in function of channel BER, i.e. distortion-BER curves. The PSNR is averaged over the frames of the video test sequence. In order to obtain more reliable results, the PSNR results of several simulations with different noise seeds have been averaged.

4 JOINT SOURCE-CHANNEL CODING SOLUTIONS

Most video source reference coders are optimized to give an error-free reconstructed best quality at a certain source coding rate assuming that the coded bit stream is received correctly. Likewise, the channel coders are designed for a particular channel and BER regardless of the source error sensitivities. On one hand, the greater the source rate, the better video quality is achieved. On the other hand, the lower the channel code rate, the less errors occur and the higher transmission rate is required. However, if the channel has limited capacity, there must be a tradeoff between the two rates, source and channel.

Two robust decoders, H.263 and MPEG-4, were used to determine the PSNR-BER curves for each channel code described in the previous section. Following this approach, we plot the objective video quality calculated in terms of video reconstructed PSNR versus channel BER for all available channel code rates, and adjusting the source code rate accordingly in order to maintain the transmission rate constant. Figure 1 presents the H.263 PSNR-BER curves for 12 channel codes assuming a constant transmission rate of 64 kbps and “Foreman” sequence encoded by the H.263 encoder. Each PSNR-BER curve is associated to a particular channel code and is composed of two straight lines, one (horizontal) at the low BER values and the other (slant) at the high BER values and a curvature in the middle BERs. The switching points between straight and curvature lines are then defined as starting and end curvature points. The curves intercept each other at cross points. Cross points indicate the BERs where a channel code should be changed in order to maximize the PSNR. For instance, if the channel code rate is 8/14, it should be changed to 8/16 at cross point 10^{-2} as the BER increases. The optimal PSNR curve is then obtained by joining all PSNR curves at cross points. The ideal PSNR curve passes through all corners, i.e., the highest curvature points. The ideal curve represents the maximum possible PSNR and is not achievable in practice. Nevertheless, as the number of codes approaches infinity, the optimal curve tends to the ideal. In this situation, unequal error protection schemes are useless since there will be a better channel code that can correct all transmission errors when the previous channel code with higher code rate starts to produce residual errors (Navarro, 2002). The lower the number of available channel codes, the higher the mean distance will be between the optimal and the ideal curves. Despite being independent of the number of channels codes to be used, the ideal curve is dependent on the particular sequence and coding algorithm. Since, in practice, only a finite set of code rates is available, the performance will, of course, be inferior to the ideal.

In this section, we firstly propose a joint source-channel coding solution with 5 channel codes to be used for both standards, H.263 and MPEG-4. This solution is called H.263/MPEG-4. Secondly, we devise a common solution which holds complete independence and therefore can be used with any source video coding algorithm. We evaluate both solutions in relation to the 5-codes optimal case.

Figure 2 presents the code rates of the 5 channel codes-optimal solution as function of the BER. The code rate changes at BERs corresponding to the interception (cross) points between PSNR-BER curves as discussed before. The locations of cross points vary with the video sequence and coding algorithm.

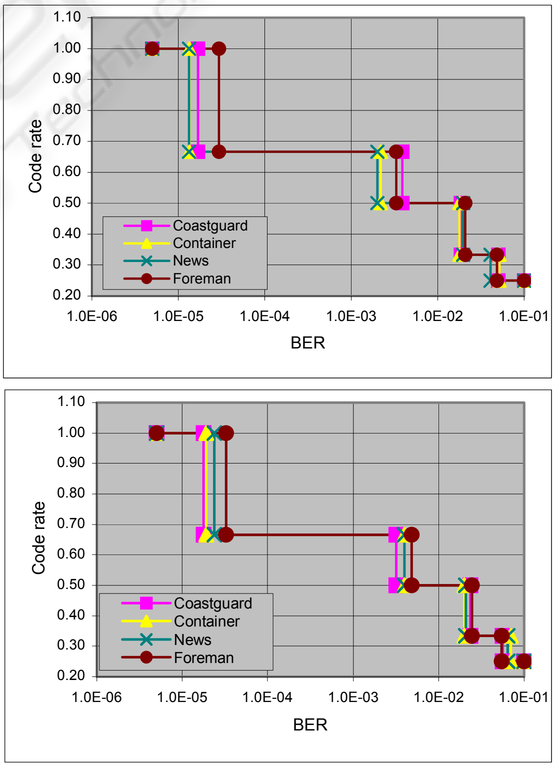


Figure 2: Optimal code rates obtained with 5 codes for H.263 (top) and MPEG-4 (bottom).

We averaged them and attained the H.263/MPEG-4 solution depicted in Figure 3. Figure 3 also shows the channel code switching BERs for the Common solution for which the switching points (BERs) are the curvature starting points on the PSNR-BER curves. In the Common solution the switching points depend almost on the correction capability of channel code. The curvature starting points are always on the left of the cross points, therefore the switching BERs for the Common solution occur before those of H.263/MPEG-4 solution as can be seen in Figure 3.

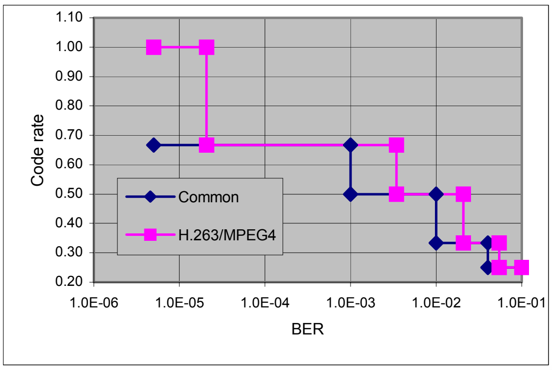


Figure 3: Code rates for H.263/MPEG-4 and Common solutions.

We now investigate which degradations both solutions achieve in comparison to the 5-codes optimal case. Figures 4 and 5 show the PSNR losses for the H.263/MPEG-4 and the Common solutions, respectively.

As expected, the PSNR loss is lower in the former solution. The highest peak loss is 3.3 dB and was obtained for "Container" video sequence encoded with MPEG-4 using the Common solution at BER about $6x10^{-2}$. This video sequence showed greater loss at BERs close to 10^{-1}. The mean PSNR losses calculated over the entire BER range are 0.06 dB (Fig. 4-top), 0.07 dB (Fig. 4-bottom), 0.12 dB (Fig. 5-top) and 0.37 dB (Fig. 5-bottom). Thus the degradation introduced by both solutions is quite small in comparison to the 5 codes optimal solution and can be used for any video sequence. The Common solution can also be used for any video coding standard.

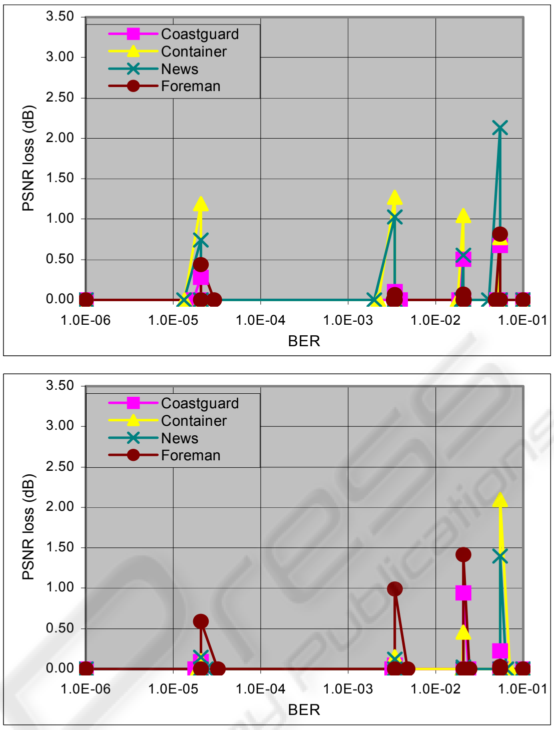
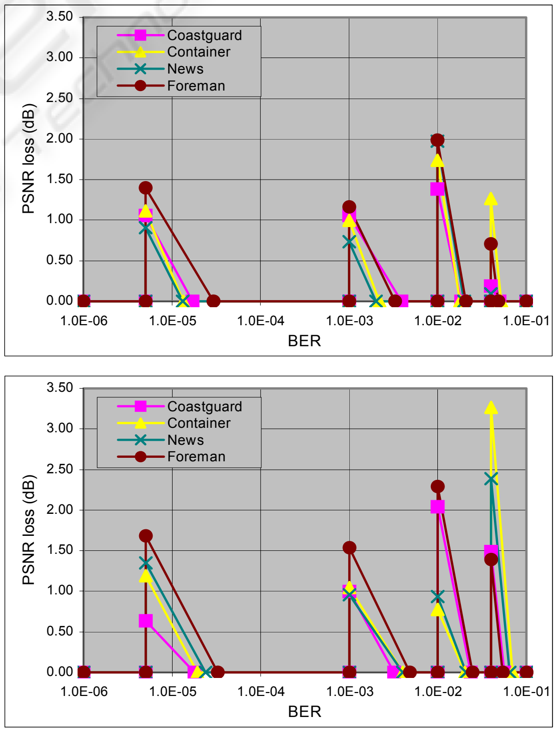


Figure 4: PSNR loss of the H.263/MPEG-4 solution in relation to the 5 codes-optimal solution for H.263 (top)



and MPEG-4 (bottom).

Figure 5: PSNR loss of the Common solution in relation to the 5 codes-optimal solution for H.263 (top) and MPEG-4 (bottom).

5 CONCLUSIONS

This paper proposes two sub-optimal adaptive solutions for video coded error protection at a transmission rate of 64 kbps. Given a channel BER, the solutions derive a RCPC code. One of the solutions is a video coding algorithm independent solution, i.e. a Common solution. The other one, named H.263/MPEG-4, is dependent on the number of channel codes used and was tailored to H.263 and MPEG-4 coding standards and therefore providing better performance than the Common solution for both standards. These two solutions were assessed in comparison to the optimal scheme which is dependent on the video coding algorithm, channel codes and video test sequence. On average, the H.263/MPEG-4 solution with 5 channel codes provides an improvement of 0.06 dB-H.263 / 0.3 dB-MPEG-4 and a degradation of 0.06 dB-H.263 / 0.07 dB-MPEG-4 in comparison to the Common solution and to the optimal case, respectively. These results show that both solutions are very close to the 5 codes optimal scheme. In the future, we consider solving the same problem but for other channels with memory. We also intend to extend the bit rate range.

REFERENCES

Shannon, C.E., 1948. A mathematical theory of communication. *The Bell Sys. Tech. J.*, vol.27, pp. 379-423.

Ortega, A. and Ramachandran, K., 1988. Rate-distortion methods for image and video compression. *IEEE Signal Proc. Mag.* , vol. 15, pp. 23-50.

Daut, D. G. and Modestino, J. W., 1983. Two-dimensional DPCM image transmission over fading channels. *IEEE Trans. Communications*, vol.31, pp. 315 - 328.

Wang, Y., Wenger, S., Wen, J., Katsaggelos, A. K., 2000. Error resilient video coding techniques. *IEEE Signal Processing Magazine*, vol. 17, July, pp. 61-82.

Frossard, P. and Verscheure, O., 2001. Joint source/FEC rate selection for quality-optimal MPEG-2 video delivery. *IEEE Trans. on Image Processing*, vol. 10, N. 12, pp. 1815-1825.

Stuhlmuller, K., Farber, N., Link, M., and Girod, B., 2000. Analysis of video transmission over lossy channels. *IEEE JSAC*, vol. 18, N. 6, pp. 1012-1737.

Gnavi, S. et al, 2002. Rate allocation for video transmission over lossy correlated networks. *Electronics Letters*, vol. 38, N. 20, pp. 1171-1172.

Grangetto, M., Magli, E. and Olmo, G., 2004. Ensuring quality of service for image transmission: Hybrid loss protection. *IEEE Trans. on Image Processing*, vol. 13, N. 6, pp. 751-757.

Tavares, J. and Navarro, A., 2001. Error statistical Analysis of H.263 Syntactical Elements. In *Proc. SPIE Vol. 4310 - Visual Communications and Image Processing*. San Jose, California-USA, pp. 382-391.

Navarro, A. and Tavares, J., 2002. Adaptive Joint H.263-Channel Coding for Memoryless Binary Channels. In *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*. Lisbon, Portugal.

Joshi, R.L. and Fischer, T.R., 1995. Comparison of generalized Gaussian and Laplacian modeling in DCT image coding. *IEEE Signal Processing Letters,* vol. 13, N. 6, pp. 751-757.

ITU-T Recommendation H.263, 1998. *Video coding for low bitrate communication*.

ISO/IEC 14496-2, 2001. *Information Technology - Coding of Audio-Visual Objects: Part 2 – Visual*.

Hagenauer, J., 1988. Rate-Compatible punctured convolutional codes (RCPC Codes) and their applications. *IEEE Transactions on Communications*, vol. 36, no. 4, pp. 389-400.