14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

H.264 VIDEO CODING FOR LOW BIT RATE ERROR PRONE CHANNELS: AN APPLICATION TO TETRA SYSTEMS

S. Colonnese, A. Piccoli, C. Sansone, G. Scarano

INFOCOM Department, University of Rome “La Sapienza”
Via Eudossiana 18, 00184, Rome, Italy
(colonnese, sansone, scarano)@infocom.uniroma1.it

ABSTRACT

This work investigates a H.264 coding scheme for video transmission over links characterized by heavy packet losses and low available bitrate. The H.264 resilient coding tools such Flexible Macroblock Ordering, Redundant Slices and Arbitrary Slice Ordering are here tuned in order to adapt the application layer coding parameters to the physical layers characteristics. Due to the limited bandwidth, the tools are differentiated on a Region Of Interest. Moreover, the Redundant Slices tool is integrated by suitable application level interleaving to counteract the bursty nature of the errors. The performances of the coding scheme choices are assessed on a TETRA communication channel, that is quite challenging due to both limited bandwidth and severe error conditions. However, the illustrated codec design criteria can be adopted in different low bit-rate, error prone channels.

1. INTRODUCTION

Development of new digital coding standards enhancing compression and coding efficiency, such as the H.264 video codec [1] developed by ITU-T and MPEG Joint Video Team, enables supporting video services also on channels that were originally designed for voice and limited data services. H.264 can achieve much higher compression efficiency than previous standards, saving up to 60% of the bit-rate at the same video quality [6], and it offers many error resilience features based on the Slice syntax unit. Small sized slices increase error resilience capability and improve error concealment performance, since the effect of a data error/loss is confined in a limited area. Flexible Macroblock Ordering (FMO) allows to separate adjacent macroblocks in different slices, and Arbitrary Slice Ordering (ASO) is allowed, too. These features reduce the visual impact of losses and errors at the expense of coding efficiency.

In low bitrate error prone channels the quality of video sequences decoded after transmission can still be unsatisfactory, unless the error resilient coding options cope with the channel characteristics. This work follows a cross-layer approach, designing coding and multiplation strategies to enable video transmission on systems severely affected by errors. We will suppose that radio RLC retransmissions are not allowed and that the cross-layer interactions between the application and multiplation layers is performed at the session setup.

We will show how, in low bit-rate error prone channels, the H.264 FMO and Redundant Slice (RS) tools should be carefully designed. Due to the limited bandwidth, the RS tool is applied only on a Region Of Interest (ROI). Second, the RS tool requires suitable application level interleaving to counteract the bursty nature of the errors. In particular, the application level interleaving depth is determined by the packet channel error statistics.

To assess the performances of the codec design choices in terms of PSNR and visual quality of the decoded stream we refer here to the TETRA communication architecture. This context is quite challenging due to both limited bandwidth and severe error conditions. However, the codec design criteria that will be illustrated can be adopted in different low bit-rate, error prone channels.

This paper is organized as follows. Section 2 describes the design of a resilient H.264, with regards to parameter settings and introduced tools; Section 3 describes the main issues of video transmissions on TETRA systems; Section 4 reports the experimental results with the simulated TETRA channel and finally Section 5 gives our conclusions.

2. DESIGN OF A H.264 RESILIENT VIDEO CODEC FOR LOW BIT-RATE CHANNELS

In low bit rate error prone channels, the worsening of the decoded video quality caused by transmission on physical channels can be very noticeable and the resulting images could be even unrecognizable. Due to the limited bandwidth resources, the coding options relative to the H.264 error resilience tools must be carefully designed, in order to match the application layer output to the physical layer features**.** As long as the flexible size slice is concerned, recent investigations show that in absence of RLC retransmission it is highly recommended that the slice size matches to the RLC frame size [5].

Here we will investigate the design of FMO, and RS, for low bit rate error prone channels. As long as FMO and RS are concerned, the lack of available resources for video transmission suggests performing an unequal spatial distribution of resources dedicated to protection, in order to maintain a good coding efficiency. Let us assume that each frame of the sequence can be divided into a ROI and a background. Interestingly enough, although many mature pre-processing segmentation techniques are known in literature [7-9], the ROI can be extracted as a by product of the coding process. In fact, the background can be built by the macroblocks that are classified as skipped by the coding algorithm, or, in a simpler approach, ROI can be *a priori* selected. In the following, we will assume that the macroblocks of the background are more coarsely quantized than those in the ROI. Moreover, the FMO is more resilient on the ROI. This can be performed without explicit shape signalling by resorting to the H.264 Slice Group Syntax Level. Experimental optimization of coding cost and achieved resilience shows that a good trade-off is achieved adopting two slice groups for the ROI area and two slice groups for the background area. This choice allows splitting the ROI in chessboard map, and the background in a less resilient but more efficient row map, as shown in Fig. 1.

RS are introduced exclusively for ROI slices, adding for each of them one or two redundant slices. The cost of RS will be paid reducing image quality with a stronger quantization. Hence, the ROI RS can be considered as a kind of application level repetition code that reduces the slice loss probability according to visual perception of end user.

In absence of radio retransmissions, contrasting channel errors by repetition requires a preliminary characterization of the distribution of losses during the transmission.

At the application level, the slice losses are bursty in nature. This is not only caused by burst of bit errors, but also by the mechanism of mapping the packets into radio frames. Frames containing errors are discarded, and if an RLC/RLP frame is lost, all the application packets that were totally or partially carried by the frame are discarded. Therefore, if original and redundant slices are adjacent in the output bitstream, a radio frame error can cause the loss of both, wasting protection resources.

We will show in the experimental results that the RS slice introduction is not fully exploited unless a suitable Slice Interleaving is adopted.

Figure 1 – Explicit FMO: 4 slice groups, two for ROI (1, 2), two for background (0, 3).

Hence, block slice interleaving for outgoing slices should be introduced, in order to fix a “security distance” between packets to be protected and protection packets used for it, as shown in Fig. 2. Therefore, interleaving depth at application level should be chosen to avoid the presence of original and redundant slices on the same radio frame. More in general, the distance between the original slice and its repetition should take into account the average packet error burst length. The slice interleaving is allowed by Arbitrary Slice Ordering (ASO) provided by H.264 standard and so is full standard compliant.

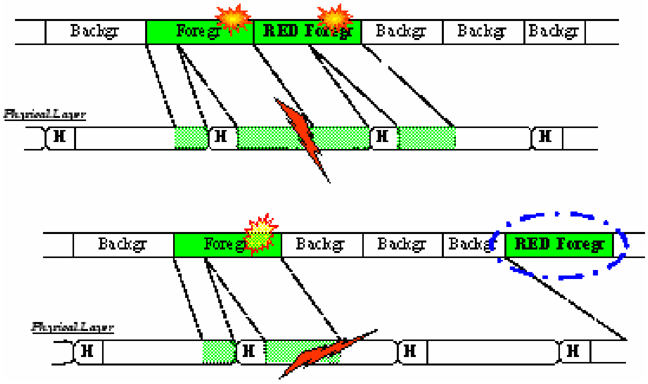


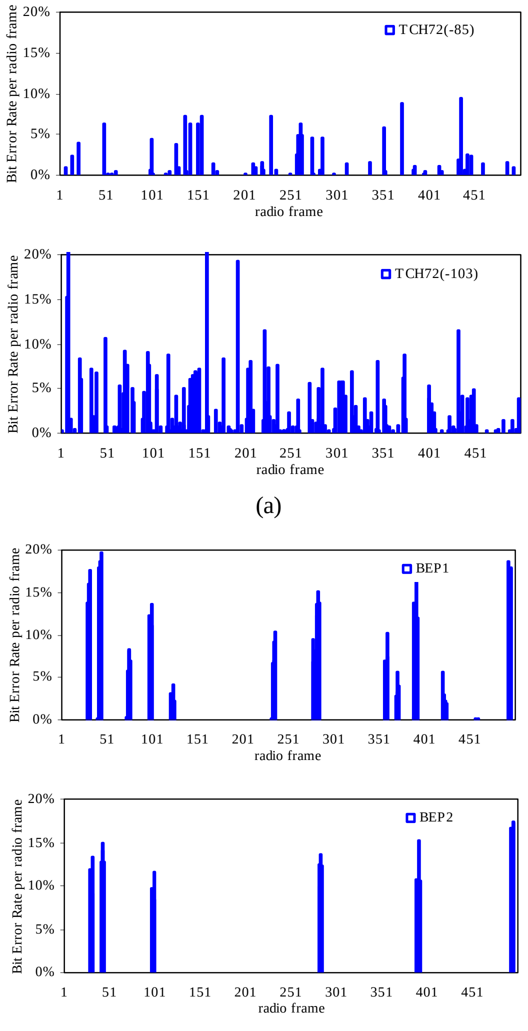
Figure 2 – Slice Interleaving effect on ROI redounded slices.

3. VIDEO COMMUNICATIONS ON TETRA SYSTEM

TETRA system [2] was developed by ETSI to satisfy the need of a flexible and efficient communication network for Private Mobile Radio (PMR), such as for public safety and security, military, etc. TETRA supports both bearer services and teleservices, circuit mode or packet oriented mode, offering individual call (point-to-point), group call (point-to-multipoint) and broadcast call. Supplementary services are also supported, such as allocation of access priority. There are some interest in providing video services in the first release of TETRA system [2], which was originally designed only for voice and limited data services.

The modulation technique used is ð/4 DQPSK. TETRA is based on Time Division Multiple Access (TDMA) with four time slots per frame, offering up to 28.8 kbit/s for mixed voice and data transmissions. Each time slot provides 7.2 kbit/s for user data rate, representing the physical cannels available. These physical channels are shared between a number of hierarchical logical channels for both traffic (Traffic CHannel, TCH) and signalling information (Broadcast Control CHannel, Common Control CHannel, Associated Control CHannel, Access Assignment CHannel, Signalling CHannel). Moreover, it provides a protected mode transmission that uses Rate-Compatible Punctured Convolutional coding scheme (RCPC) coding, reducing the available bit-rate to only 19.2 kbit/s [2][3].

In typical working conditions, TETRA channels may be affected by high bit error rates, resulting in high application level packet loss rate. Two sample TETRA BEPs obtained via air interface measurement are illustrated in Fig. 3.a, where the fraction of erroneous bit per radio frame is plotted versus the radio frame index. The BEPs are taken at two different power levels, namely at the nominal input level (-85 dBm) and at the dynamic sensitivity for the mobile station (-103 dBm). For comparison sake, Fig. 3.b plots two reference UMTS BEPs [5], which are considered not suitable for real time video communication due to the relatively high BER. From a visual comparison of the TETRA and UMTS BEPs, it is clear that the TETRA environment is more hostile, and therefore it requires a more accurate design of the resilient coding options.



(a)

(b)

Figure 3. Fraction of erroneous bit per radio frame. TETRA BEPs (a), and UMTS BEPs (b).

The TETRA BEPs in Fig. 3.a result in high slice loss rate at the application level, of about 20% for BEP TCH72(-85) and up to 60% for TCH72(-103).

Moreover, radio frame losses occur in bursts. Measurement of loss burst lengths after a series of different simulations shows a clear prevalence of bursts of 2 or 3 packets for both the used error patterns (Fig. 4). In such a hostile environment, it is clear the impact of RS and the need for proper slice interleaving.

The application layer interleaving depth in this environment can be conveniently set to 3 slices, according to the physical layer packet loss burst length that, as shown in Fig. 4, is usually 2 or 3. In this way almost one between original and duplicated slices is saved from losses, going out of the current loss burst. Although deeper interleaving could reduce effects of all longer bursts, in particular with TCH72(-103) error pattern, it is good to keep lowest interleaving depth to have maximum efficiency for short bursts, the most frequent and harmful. The total slice number to interleave in each interleaving block must be settled taking attention to maximum additional delay due to receiving buffering before decoding, needed to perform de-interleaving.

Without loss of generality, we will assume the herein described TETRA channel as a framework for assessing the performances of the coding criteria described in Section II. Nevertheless, these criteria can be adopted in different error prone low bit rate channels.

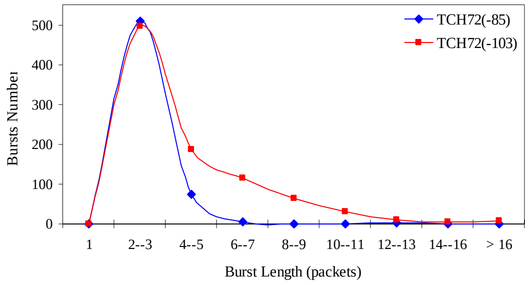


Figure 4. Distribution of losses burst length.

4. EXPERIMENTAL RESULTS

The simulation scenario models a packet-oriented video communication on TETRA connection.

The application level employs a H.264 codec developed on the base on the JM codec [10], and extended in order to perform ROI selection, Redundant Slice generation, and application level interleaving. The test video sequences Silent, Akiyo, Foreman, QCIF format (176x144 pels), are coded at 5 frames per second. Quantization Parameters (QPs) were chosen in order to get the best luminance PSNR maintaining bitrate below 24 kbit/s. The encoded sequence is I – P – P – P – P – P. Adaptive Intra Refresh is performed, with full MB refresh of 10 randomly chosen macroblocks per coded frame. In order to cope with the radio frame payload size of 640 bits, the slice size is limited to 45 bytes.

The user plane protocol stack shown in Fig. 5 has been implemented. The simulator, developed on the base of the Mobile IP SW [11], performs framing at the PDCP/PPP level, adding overhead of RTP/UDP/IP compressed header, scheduling and splitting application packets to RLC/RLP frames. The Bit Error Pattern (BEP) representing the Physical Layer output is mapped to RLC/RLP frames. Radio frames containing errors are discarded, and if an RLC/RLP frame is lost the application packets carried by it (partially or as a whole) are discarded.

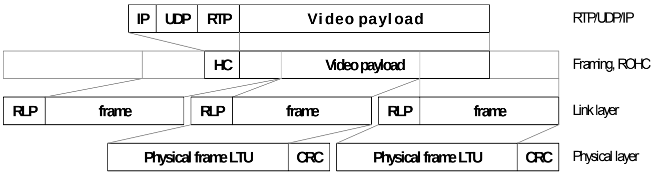


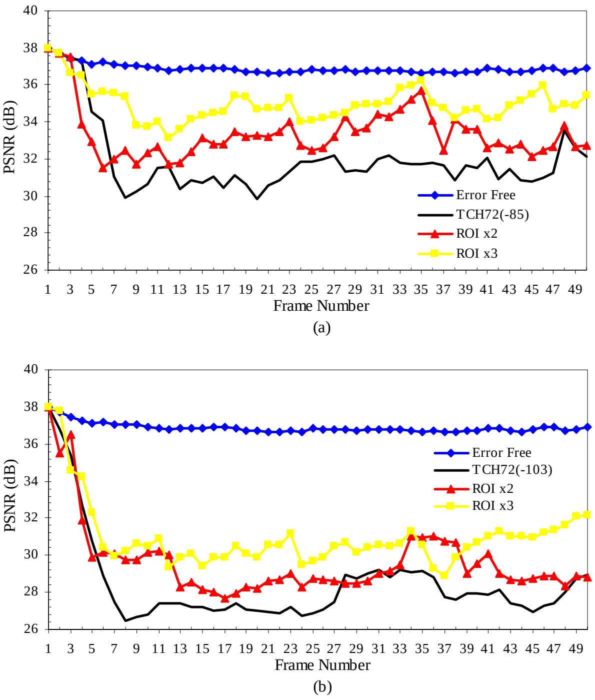
Figure 5. Packetization of application packets through the protocol stack.

The TETRA physical channel is simulated by the BEPs obtained via air interface measurement shown in Fig. 3.a, that will be referred in the following as *TCH72(-85)* and *TCH72(-103)*, respectively.

Fig. 6 shows the luminance PSNR for the Akiyo sequence. The PSNR loss in case of BEP TCH72(-85) is about 5 dB, with negative peaks up to 9 dB. With TCH72(-103) PSNR losses are even higher, about 9 dB in average. Corresponding measured Packet Loss Rates are about 20% for BEP TCH72(-85) and up to 60% for TCH72(-103). Received images are very disturbed, almost unrecognizable. We can see the effects of Selective Redundancy on PSNR, with doubled and trebled ROI. Results are not fully satisfying, because of redundancy inefficiency caused by burst losses.
Fig. 7 shows PSNR enhancement of interleaved redundant approach compared to the simple redundancy case: average PSNR on the entire frame is increased.
Comparison between Fig. 6 and Fig. 7 points out the improvements obtained with interleaving strategy. Best results have been reached especially for TCH72(-85) pattern. For pattern TCH72(-103) quality recovering is only partial: very high packet loss rate measured in this case makes final sequence not fully satisfying in spite of redundancy and interleaving strategies, even if the comparison with totally destroyed sequence received without protection tools shows great improvements. Fig. 8 and Fig. 9 show the visual effects of interleaved ROI redundancy strategy with the error patterns TCH72(-85) and TCH72(-103), compared with the simple ROI redundancy cases.
Let us observe that the interleaving depth is set equal to 3 slices, according to packet loss burst length that, as shown in Fig. 2, is usually 2 or 3. Longer burst losses of 10, 11, up to 16 packets appear, that are not counteracted by this design choice. Although in principle the redundancy interleaving depth could be increased, furtherly reducing the lost areas in the decoded sequence, the benefit of interleaving should be balanced with the cost of buffering and delay, which are typical constraints in the design phase.

5. CONCLUSIONS

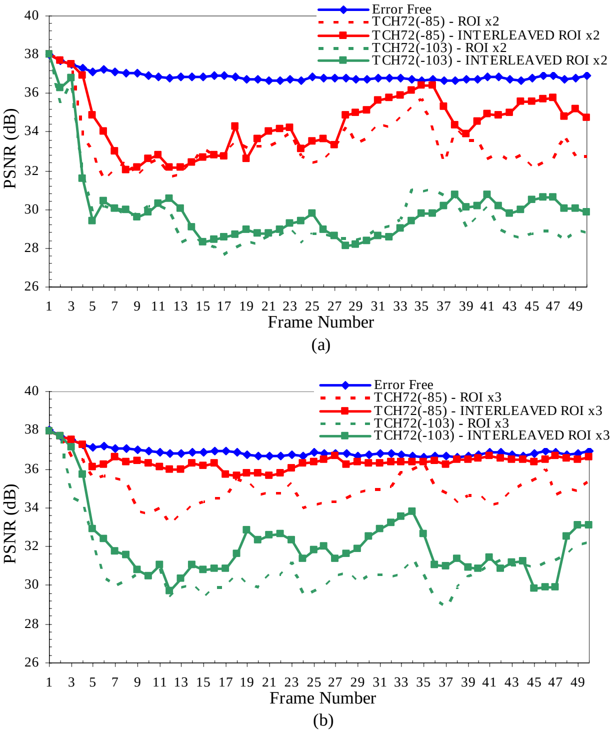
In this paper we have described a H.264 coding scheme for video communications on low bit-rate error prone channels. The work has shown that the coding tools provided at the application level should be tuned to the characteristics of the physical layer. In low bit-rate error prone channels, the H.264 FMO and Redundant Slice tools should be only on a Region Of Interest, to cope with the limited bandwidth. Moreover, a suitable application level interleaving may counteract the bursty nature of the errors. The performances of the codec in terms of PSNR and visual quality of the decoded stream are simulated for a packet oriented communication in a TETRA system in severe error conditions. Nevertheless, the illustrated codec design criteria can be adopted in different low bit-rate, error prone channels.



(a)

(b)

Figure 6. Effects of Selective Redundancy on PSNR, with doubled and trebled ROI, in comparison with error free and lossy cases. Results with TCH72(-85) (a), and TCH72(-103) (b).



(a)

(b)

Figure 7. Interleaving Enhancement on PSNR with Akiyo sequence in comparison of non-interleaved redundancy case: doubled ROI (a); trebled ROI (b).


(a)
(b) (c)
(d) (e)

Figure 8. Effect of BEP TCH72(-85) on Akiyo sequence (a); simple ROI redundancy x2 and x3 (b-c); Interleaved ROI redundancy x2 and x3 (d-e).

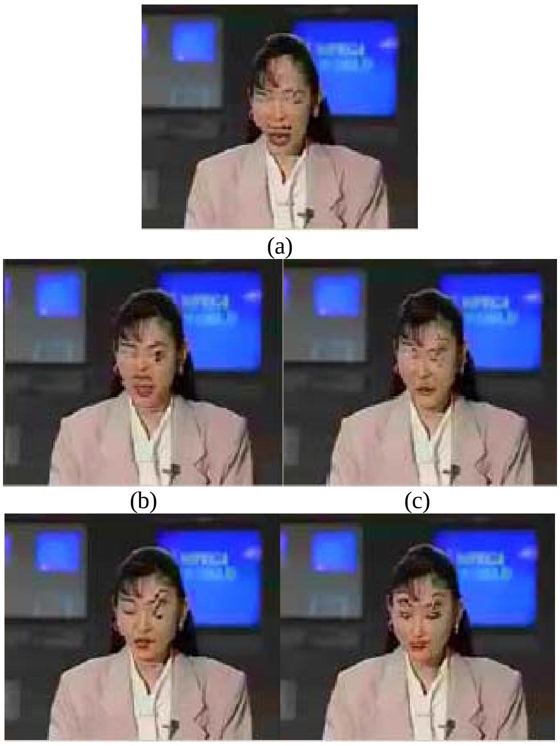

(a)
(b) (c)
(d) (e)

Figure 9. Effect of BEP TCH72(-103) on Akiyo sequence (a); simple ROI redundancy x2 and x3 (b-c); Interleaved ROI redundancy x2 and x3 (d-e).

REFERENCES

[1] T. Wiegand, G.J. Sullivan, G. Bjontegaard, A.Luthra, "Overview of the H.264 Video Coding Standard", IEEE Transaction on Circuits and Systems for Video Technology, Vol. 13, No. 7, pp. 560-576, July 2003.

[2] ETSI ETS 300 392-1 DE/RES-06001-1 Radio Equipment and Systems (RES); Trans-European Trunked Radio (TETRA); Voice plus Data (V+D); Part 1: General network design. February 1996.

[3] ETSI ETS 300 392-2 DE/RES-06001-2 Radio Equipment and Systems (RES); Trans-European Trunked Radio (TETRA); Voice plus Data (V+D); Part 2: Air interface (AI). March 1996.

[4] A. Annunziato, D Sorbara, "TETRA radio performance evaluated via the software package TETRASIM" , Mobile Networks and applications, Vol. 5, Issue 1, pp. 17-26, March 2000.

[5] T. Stockhammer, M.M. Hannuksela, T. Wiegand, "H.264 in Wireless Environment", IEEE Transaction on Circuits and Systems for Video Technology, Vol. 13, No. 7, pp. 657-673, July 2003.

[6] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, G.J. Sullivan, "Rate-Constrained Coder Control and Comparison of Video Coding Standards", IEEE Transaction on Circuits and Systems for Video Technology, Vol. 13, No. 7, pp. 688-703, July 2003.

[7] M. M. Hannuksela, Y.-K. Wang, and M. Gabbouj, "Sub-picture: ROI coding and unequal error protection," presented at International Conference on Image Processing (ICIP'02), Sep 22-25 2002, Rochester, NY, United States, 2002.

[8] M. M. Hannuksela, Y.-K. Wang, and M. Gabbouj, "Isolated Regions in Video Coding," IEEE Transactions on Multimedia, vol. 6, pp. 259-267, 2004.

[9] A. Neri, S. Colonnese, G. Russo, P.Talone "Automatic Moving Object and Background Separation", Signal Processing Vol. 66, No. 2, Aprile 1998.

[10] H.264 JM 9.3 codec available at http://iphome.hhi.de/suehring/tml/.

[11] MobileIP software available at the ITU-T archive http://ftp3.itu.ch/av-arch/video-site.