MULTIPLE DESCRIPTION CODING WITH ERROR CORRECTION CAPABILITIES: AN APPLICATION TO MOTION JPEG 2000

Tammam Tillo, Marco Grangetto, Gabriella Olmo

CERCOM - Center for Multimedia Radio Communications
Dipartimento di Elettronica - Politecnico di Torino
Corso Duca degli Abruzzi 24 - 10129 Torino - Italy
Ph.: +39-011-5644077 - FAX: +39-011-5644099
tammam.tillo@polito.it, marco.grangetto@polito.it, gabriella.olmo@polito.it

ABSTRACT

In the multiple description paradigm, a controllable amount of redundancy is inserted among descriptions, in order to help estimating those ones which are possibly lost due to network congestion. This redundancy can also be exploited in order to correct errors at bit level. In this paper, we propose a novel technique to generate multiple descriptions of video encoded with Motion-JPEG 2000, which exploits the inserted extra redundancy also to guarantee error protection in case all descriptions are received, but are possibly affected by bit errors. This method yields excellent performance, since it guarantees not only protection of video information transmitted over non prioritized networks subject to independent packet erasure processes, but also resilience towards the corruption at bit level. Moreover, the generated streams are fully compatible with the Part 3 of the JPEG 2000 standard.

1. INTRODUCTION

Multiple description coding (MDC) is recognized as an effective method to protect multimedia information transmitted over non prioritized networks. In the MDC approach, two or more non hierarchical representations of the same data are generated, yielding mutually refinable information. In this paper we address the practical situation of generating only two descriptions.

Perhaps the most popular MDC methods stem from the pioneering MD scalar quantization (*MDSQ*) [1], where the index information of a standard quantizer is structured in order to meet the MDC paradigm. In [2], this principle is applied to video coding employing a predictive multiple description quantizer. Another class of methods employ pairwise correlating transform (*MDPCT*) operating on the coefficients in order to introduce a controlled amount of redundancy among the descriptions [3]. This approach was applied to motion compensated video in [4].

In the MDC context, the network over which transmission occurs is usually assumed to be subject to independent packet erasure process. In this case, the quality of the recovered signal depends only on the number of received descriptions, and not on the specific loss pattern. Whereas this model is adequate to represent wired networks, it is not suitable for most wireless networks; in fact, deep fading can cause the loss of entire packets, but also the situation of errors at bit level must be accounted for. As an example, in case of real time multimedia applications, the UDP protocol with only IP checksum is typically used at transport layer [5]; therefore, the packet payload can be delivered with residual bit errors to the application layer. In such a case, we must assume that a description can be either lost, or received affected by bit errors. Generally speaking, the first problem can be solved by MDC techniques, whereas the second one is treated by means of channel coding, which, on the other hand, introduce further redundancy in the encoded bit stream. However, since MDC already introduces a controlled amount of redundancy among the descriptions, it is worth questioning if this latter can be also exploited to correct those descriptions which are received affected by bit errors, so achieving a larger overall efficiency.

In this paper, we address the problem of generating two descriptions of video frames, employing Motion-JPEG 2000 (M-JPEG 2000) as the co-decoding engine; the objective is to guarantee both resilience towards packet losses, and protection in case the two descriptions are received affected by bit errors. The same principle could be applied to MDPCT, where the coefficients are paired following a specific scheme [3] to form the two descriptions. The pairing scheme being known, can be used in order to retrieve errors; however, the error recovering scheme is not trivial. On the other

This work was partially supported by the Italian Ministry of Education and Research under the CERCOM (Center for Multimedia Radio Communications) and PRIMO (Reconfigurable Platforms for Wideband Wireless Communications) grants.

0-7803-8554-3/04/$20.00 ©2004 IEEE.

hand, systems based on MDSQ, where quantization levels are represented by a pair of indexes, can exploit the fact that only some combinations of index pairs are valid [6]. However, to the best of our knowledge, the proposed technique is the first MDC proposal for JPEG 2000 with bit error correction capability.

2. MULTIPLE DESCRIPTION GENERATION

The proposed algorithm aims at creating two balanced[1] descriptions of a video source encoded with M-JPEG 2000. The core technology of M-JPEG 2000 is based on the baseline JPEG 2000 part 1, and each frame is Intra coded with no motion prediction/compensation. In JPEG 2000, the target rate is achieved by coding the frame code-blocks CBs one bit-plane at a time; then, a post-processing operation determines the extent to which each CB bit-stream should be truncated [8]; this means that a CB stream encoded at low bit rate is completely embedded in a higher rate one. As a consequence, the bit-stream of a frame encoded at rate $R_2 < R_1$ is entirely included in that encoded at rate R_1; this property can be exploited in order to guarantee some sort of unequal error protection.

If two streams are generated, encoded at rates R_1 and $R_2 < R_1$ respectively, two unbalanced description of the frame are obtained. The key problem is how to convert such unbalanced descriptions into balanced ones. To this end, for each layer of either bit stream, the CBs of each subband are taken alternatively from either original stream; the LL band is duplicated in both descriptions (see Fig. 1). This procedure results in almost balanced descriptions, approximately encoded at rate $(R_1 + R_2)/2$.

At the decoder side, if both descriptions are received, these latter are pre-processed and merged in a single stream, where, for each frame and layer and for each CB, the best representation is selected; the resulting stream is then M-JPEG 2000 decoded. On the other hand, if a description is lost, the received one is simply M-JPEG 2000 decoded yielding inferior, yet still acceptable quality.

As usual, the performance of the proposed MDC technique can be evaluated in terms of the *central distortion*, i.e. the distortion when both descriptions are received, and the *side distortions* when either description is received, as a functions of the total rate on both channels. Alternatively, the peak signal-to-noise ratio (PSNR), which is related to the visual quality, can be evaluated. The same central quality delivered by two descriptions encoded at the total rate of $(R_1 + R_2)$ can be obtained by single description coding (*SDC*) at the rate R_1; as a consequence the extra rate used by MDC is R_2 and the relative redundancy is $\rho = R_2/(R_1 + R_2)$.

[1]Two descriptions are said to be balanced if they have equal rate and they result in identical distortion when decoded separately

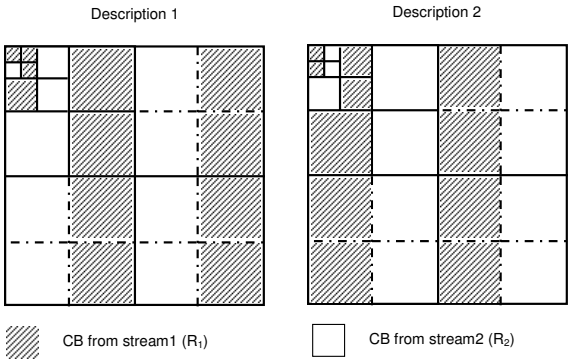

Fig. 1. Descriptions generation from two M-JPEG 2000 stream of a single frame encoded at rate R_1 and R_2.

3. ERROR DETECTION AND CORRECTION

In this section, we assume that both descriptions have been received, possibly affected by bit errors. In this situation, for each CB we have two representation available, one of which is completely embedded in the other one. This property can be exploited in order to identify the locations of bit errors in the shortest stream representing the CB. The error location can be retrieved by ex-oring the common part of the two streams representing the CB, as depicted in Fig. 2-(c), resulting in logical zeros in the error free or duplicated error positions (i.e., errors occurring in the same position on both streams), and logical ones in the locations of single errors. When the bit error rate (BER) on both channels is small and the two channels are independent of each other, as assumed in the MDC context, the probability of occurrence of duplicated errors can be neglected. Therefore, the ex-ored bit stream can be considered as an error syndrome, which can be used by the decoder to perform error correction. When an error is signaled in a given position, the decoding procedure is split into two branches, that correspond to 0 and 1 bit value respectively (Fig. 2-(d)). Branches are created each time an error is signaled, resulting into a decoding tree with a number of branches which increases exponentially with the number of syndrome errors. The decoder can sequentially exploit the standard error resilience tools offered by JPEG 2000 (e.g. segmentation markers, arithmetic coder termination at each coding pass) to select those branches corresponding to valid JPEG 2000 bit-streams. The branches which cause error detection are pruned from the tree; for memory and complexity reasons the maximum number of branches is limited by a threshold M. This procedure is able to correct errors which lie in the common part of the two CB representations. The remaining part of the CB bit-stream is concealed with the standard JPEG 2000 error resilience tools. Due to the progressivity of JPEG 2000 streams, this amounts to applying a sort of

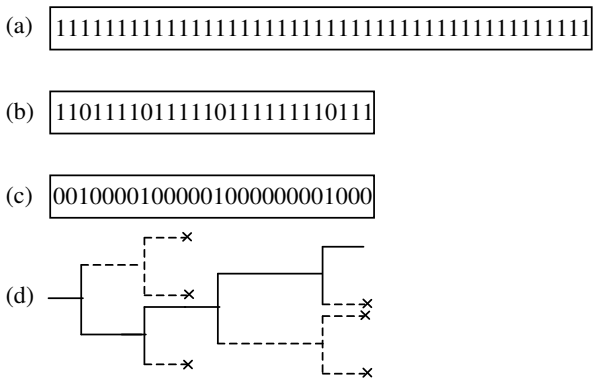

Fig. 2. Error correction when both descriptions are received; (a) stream of finer representation of CB; (b) stream of coarser representation; (c) result of ex-oring; (d) branching.

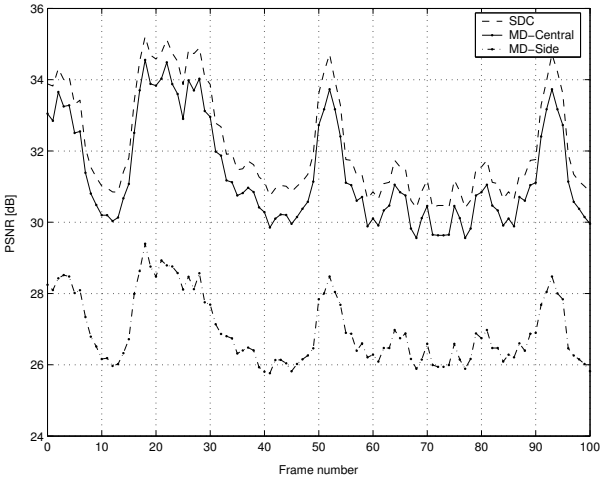

Fig. 3. SDC, Side and central PSNR (luminance component) vs. frame number; *Teen* sequence; $\rho = 0.2$.

Unequal Error Protection to the video source.

4. EXPERIMENTAL RESULTS

The proposed algorithm has been tested employing 100 frames of sequence *Teen* in CIF-YUV 4:2:0 format. The frame rate is 15fps, and the overall bit rate close to 700 kbps. The M-JPEG 2000 configuration parameters are 4 resolution levels and 32×32 code blocks; since the approximately balanced descriptions are generated by alternatively taking CBs from two streams encoded at different rates, decreasing the CB size would result in more balanced descriptions, but also in an extra overhead due to the extra CB markers. In order to achieve different trade off between central and side quality, two redundancy values are considered, namely $\rho = 0.2$ and $\rho = 0.4$. Headers are assumed to be protected by a forward error correction (FEC) code with rate 0.5; this extra information has been accounted for in terms of rate overhead.

In order to validate the error detection and correction capabilities, we simulate a binary symmetric channel (BSC) with BER of 10^{-3} and 10^{-4} respectively. The error resilience tools offered by JPEG 2000, i.e. the termination of the arithmetic coder for each pass and the insertion of four symbols code (SEGMARK) before the first new coding pass in each bit-plane, have been activated so that the proposed algorithm can determine the correct branching. The error correction results are obtained with a maximum number of branches $M = 64$.

The error free performance of the MDC scheme with $\rho = 0.2$ is reported in Fig. 3 for the whole sequence, in terms of side and central PSNR in dB (luminance component) versus the frame number. For the sake of comparison, the performance is also reported of single description M-JPEG 2000 coding (SDC), encoded at the same total rate. In Fig. 4, the central PSNR in dB is reported as a function of the frame number, for $\rho = 0.4$, BER = 10^{-3}. The performance yielded by SDC M-JPEG 2000, with the error resilient SEGMARK and arithmetic coder termination, is reported for comparison purposes. Not surprisingly, SDC outperforms MDC as for the central quality in the error free case. However, in the error prone situation, MCD almost always outperforms SDC, which, moreover, yields highly variable performance with several deep PSNR fades of more than 15 dB. On the other hand, the central MDC quality is smoother and close to the error-free value.

In Tab. 1 the average PSNR, computed as the linear average of MSE over 1000 frames, is reported and compared to SDC for the same overall rate, in the error free, BER=10^{-4} and BER=10^{-3} situations. It is worth noticing that the side decoding performance is obtained with a standard error resilient M-JPEG 2000 decoder. It can be observed that, in the error free case, MDC guarantees a fairly good side quality (in excess to 26 dB), with a small central degradation with respect to SDC. The advantage offered by the proposed MDC bit error correction strategy is apparent. In fact, for BER=10^{-4}, the MDC central decoding outperforms SDC by about 5 dB; this improvement reaches 10 dB in the case of the harsher channel with BER=10^{-3}. It is clear that increasing the redundancy improves the side performance, and is beneficial for high BER situations. Finally, it is worth pointing out that the side and SDC decoders yield approximatively the same performance in the presence of bit errors; in fact, for medium to high BER, the decoded quality depends only marginally on the coding rate and it is mainly affected by the BER value.

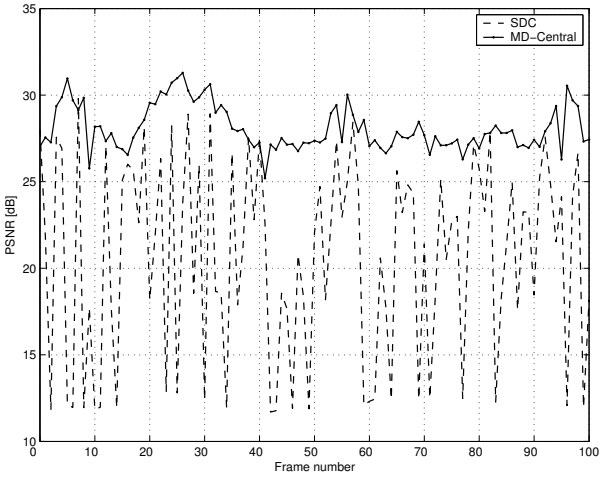

Fig. 4. SDC and central PSNR vs. frame number; BSC channel with BER = 10^{-3}; *Teen* sequence; $\rho = 0.4$.
.

Table 1. Average PSNR for *Teen* sequence; SDC and MDC performance as a function of BER and MDC redundancy ρ.

BER	SDC	$\rho = 0.2$		$\rho = 0.4$	
		side	central	side	central
Error free	32.11	26.61	30.16	27.84	28.80
10^{-4}	24.61	23.11	29.66	23.57	28.70
10^{-3}	17.02	16.82	27.30	16.91	27.89

5. CONCLUSION

In this paper, we have proposed an MDC scheme based on M-JPEG 2000, which can be efficiently exploited also for error correction, in case both descriptions are received affected by bit errors. The extra redundancy of the MDC scheme can be easily controlled, in order to match the network and transmission conditions, in term of packet loss rate and BER. The proposed scheme largely outperforms the plain M-JPEG 2000 equipped by its standard error resilience tools. We have also shown that a base quality level can always be guaranteed, corresponding to the quality delivered by the bit stream encoded at the lower rate R_2. Moreover, the error detection and correction scheme is easy to implement. Usually, when comparisons are made between MDC and layered coding, three different protection scenarios are considered: no error protection, Automatic Repeat reQuest (ARQ) based error protection and FEC based error protection [7]. Exploiting the MDC redundancy in order to retrieve and correct bit errors improves the performance of MDC compared to layered coding. Moreover, the generated descriptions are M-JPEG 2000 part 3 compliant, meaning that any M-JPEG 2000 decoder (part 3 compliant) can decode the descriptions and output a valid sequence, without provoking any decoder crash. This makes the proposed algorithm a good choice for applications requiring backward compatibility.

Future developments will encompass the treatment of duplicated errors occurring at the same location in the two description, the use of resilient AC equipped with forbidden symbols for more efficient error detection, the creation of rate distortion optimized descriptions from JPEG 2000 streams.

6. REFERENCES

[1] V. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Trans. on Information Theory*, vol. 39, no. 3, pp. 821-834, May 1993.

[2] V. Vaishampayan, S. John, "Balanced interframe multiple description video compression," *International Conference on Image Processing (ICIP)*, vol. 3, pp. 812-816, October 1999.

[3] Y. Wang, M. T. Orchard, V. Vaishampayan, A. R. Reibman, "Multiple description coding using pairwise correlating transforms," *IEEE Trans. on Image Processing*, vol. 10, no. 3, pp. 351-367, March 2001.

[4] A.R. Reibman, H. Jafarkhani, Yao Wang, M.T. Orchard, R. Puri, "Multiple description coding for video using motion compensated prediction," *International Conference on Image Processing (ICIP)*, vol. 3, pp. 837-841, October 1999.

[5] L.A. Larzon, M. Degermark, S. Pink, "Efficient use of wireless bandwidth for multimedia applications," *IEEE International Workshop on Mobile Multimedia Communications*, pp. 187-193, Nov. 1999.

[6] V. Vaishampayan, S. John, "A Viterbi based decoding algorithm for multiple description variable length codes," *International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, pp. 2573-2576, June 2000.

[7] C. Lee, J. Kim, Y. Altunbasak, R. M. Mersereau "Layered coded vs. multiple description coded video over error-prone networks," *Signal Processing: Image Communication*, May, 2003.

[8] C. Christopoulos, A. Skodras, T. Ebrahimi, "The JPEG2000 still image coding system: an overview," *IEEE Trans. on Consumer Electronics*, Vol. 46, no. 4, pp. 1103-1127, Novemeber 2000.