Unequal loss protection and multiple description coding: a performance comparison

Unequal loss protection and multiple description coding: a performance comparison

Enrico Baccaglini
Dipartimento di Elettronica
Politecnico di Torino
Corso Duca degli Abruzzi 24 - 10129 Torino - Italy
enrico.baccaglini@polito.it

Tammam Tillo
Department of Electrical and Electronic Engineering
Xi'an Jiaotong-Liverpool University
111 Ren Ai Road - Suzhou 215123 - P.R. China
tammam.tillo@xjtlu.edu.cn

Gabriella Olmo
Dipartimento di Elettronica
Politecnico di Torino
Corso Duca degli Abruzzi 24 - 10129 Torino - Italy
gabriella.olmo@polito.it *

ABSTRACT

In this paper we present a performance comparison between multiple description coding and unequal loss protection as tools to encode a layered source. We address a rate-distortion-based multiple description coding scheme and a state-of-the-art unequal loss protection algorithm based on Reed Solomon FEC codes. The comparison is performed using as a case study JPEG 2000 coded images transmitted over lossy packet networks. Complexity aspects are also considered. The simulation results show that both schemes allocate the same amount of redundancy for any given encoding output rate to protect the transmitted information. Whereas MDC, besides being computationally less intensive, achieves a smoother performance degradation, the ULP scheme yields superior performance in terms of the expected PSNR.

1. INTRODUCTION

In modern multimedia scenarios, users may want to download or stream multimedia data using terminals equipped with different capabilities in terms of power consumption, memory, computational resources and visual resolution. The contents may be accessed using broadband networks such as DSLs, optical cable or WiMax, but also full mobility GPRS/UMTS or beyond 3G networks. Consequently, scalability is a key feature. The multimedia data quality should be matched to the visual and computational resources of the terminal at hand. Moreover, the signal received after transmission on unreliable networks should exhibit graceful degradation capabilities, so as to enable the decoder to achieve different quality levels, depending on the amount of correctly received information.

Let us consider the case where data are transmitted on a non prioritized network, with all packets encompassing the same loss probability, and retransmission is unfeasible due to delay constraints or network flooding problems (such as in broadcast or multicast applications). In such a situation, it is important that all the received packets can be exploited at the application level. This is not the case of packets built on top of a layered source, where the loss of a base layer packet prevents one from exploiting the subsequent ones, even though correctly received. In multiple description coding (MDC) [3], non hierarchical representations of the source are generated, yielding mutually refinable information. The quality at the receiver side only depends on the number of successfully received packets, and not on the particular subset or arrival order. Such nice features reduce the compression efficiency, due to the need of inserting redundancy among the descriptions. This redundancy is generally measured in terms of the extra rate required by the MDC scheme, compared to a single description reference system achieving the same performance. However, a more sensible comparison should address other error resilient systems, implementing different policies to trade redundancy for robustness, e.g. unequal loss protection (ULP) using forward error correction (FEC) codes.

In this paper we compare the rate-redundancy-distortion performance of MDC and FEC-based methods. This problem is not trivial, due to the high number of involved variables and parameters. We select a sensible case study, i.e. the transmission of progressively encoded images over packet lossy networks, and address for comparison state-of-the-art MDC and ULP methods, suitable for layered sources.

The ULP scheme of [6] is one of the most popular algorithms to implement the graceful degradation concept. It is based on Reed-Solomon (RS) codes, which are unequally allocated to data fragments obtained from a progressive or layered source. The code rates to be employed are obtained as the result of an optimization procedure, aiming at maximizing the expected peak signal-to-noise ratio (PSNR) at the receiver side, given the packet (description) loss probability and the source rate-distortion (RD) characteristics. The principle in [6] has been applied in several similar works. For example, in [4] a low complexity system is proposed, which uses equal loss protection (ELP) to guarantee a basic quality level, and then applies the rate allocation of [6] on the rest of the bit stream, achieving a trade off between ELP and ULP performance. As the code allocation procedure in [6] is very demanding, a feasible, yet slightly sub-optimal

*This work has been supported in part by the European Union under the SEA project.

Mobimedia '09, September 7-9, 2009, London, UK

algorithm has been proposed in [9]. The same algorithm is applied in [8], where it is recognized that ULP is very sensitive to variations of the estimated loss rate. A cross-layer control mechanism is implemented so as to make the optimization task aware of the actual network conditions, even though the complexity of the optimization task is likely to prevent this mechanism from working in real time. We then decided to adopt the algorithm in [9] as a term of comparison of ULP and MDC, for both the rate-redundancy-distortion performance and the computational complexity.

As for MDC, we have selected the algorithm proposed in [11], due to its good performance, flexibility in generating any number of descriptions, and compatibility with standard co-decoding tools. Moreover, also this algorithm is designed so as to maximize the expected PSNR at the receiver side, given the description loss probability and the source RD characteristics; therefore, it is directly comparable to [9]. In [11], the algorithm is applied to Gaussian data. Here, we modify it in order to make it suitable for JPEG 2000 data, obtaining a practical MDC coding tool, named *RD-aware multiple description coding* (RDMC).

From this brief discussion, it can be noticed that few MDC algorithms are suitable for a sensible comparison with ULP. Nevertheless, some effort in this direction has already been spent. In [5], a comparison between MDC and FEC is presented, using a memoryless Gaussian source and addressing rate-distortion (RD) performance bounds. The authors come to the conclusion that MDC outperforms ULP in case delay constraints are present, and a feedback is available on the channel conditions. However, the generalization of such results to real-world data is not trivial, as MDC performance bounds are known not to be strict [12]. In [2], a method is proposed to generate two unbalanced descriptions of video streams, and some general features of ULP are highlighted, such as the presence of a cliff effect. However, no numerical performance comparison is presented.

The rest of this paper is organized as follows. In Sec. 2 we describe the RDMC encoder and decoder. In Sec. 3, we present experimental results focusing on the comparison of RDMC and ULP applied to JPEG 2000 data. Finally, in Sec. 4 we draw our conclusions.

2. THE RDMC ALGORITHM

In the MDC method proposed in [11], the data RD curve is exploited to generate an arbitrary number N of descriptions. The data source is first encoded at N rates taken from a encoding rate vector $\mathbf{R} = [R_1 \geq R_2 \geq \cdots \geq R_N]$ thus generating N streams. Such rates, which are variables of the optimization problem, are subject to the constraint $\sum_{i=1}^{N} R_i = R_t$, with R_t being the total available rate. Then, the encoder groups the data source into an arbitrary number of so-called subsets (data sets) so that each subset will be available at a different coding rate among descriptions. The rates themselves are selected so as to maximize the expected PSNR at the receiver side and the constraint on the total available rate. The interested reader can find further implementation details in [11].

In this section, we adapt the subset selection procedure in [11] to the JPEG 2000 case. In particular, we use JPEG2000 codeblocks (CBs) as the basic data units. Let us assume that we are able to define $N!$ CB subsets $\mathbf{S}_l$, $l = 1, .., N!$, which are balanced in the RD sense (i.e. when subsets are encoded at the same rate, each of them yields the same contribution to the recovered image distortion). Then, subset $\mathbf{S}_1$ is encoded along descriptions using a permutation π_1 of rates $\mathbf{R}$, and so on for subset $\mathbf{S}_2$ (π_2), ..., $\mathbf{S}_{N!}$ ($\pi_{(N!)}$).

In the case of JPEG 2000 data, an issue to be solved is the identification of subsets of CBs that are equivalent from the RD standpoint. Even though the data RD curve is made available by the JPEG 2000 encoder, and consequently, in principle, explicit RD evaluation and exact CB classification is possible, this would lead to a computationally intensive procedure, which could hardly be integrated in a real time system. Therefore, we avoid explicit RD computation, and instead we rely on some assumptions, which allow us to identify a static pattern allocation of CBs to the descriptions. Our basic assumption here is that, for each level of wavelet decomposition, the detail subbands are made of CBs having roughly similar RD characteristics; as a consequence, CBs belonging to such subband groups can be considered equivalent from the RD standpoint. The same assumption is made for the lowest frequency subband. This simplifies the rate allocation procedure to a great extent, as CBs belonging to equivalent sets can be simply split into the descriptions, according to a permutation of the encoding rate vector $\mathbf{R}$. Fig. 1 shows an example for $N = 4$ descriptions.

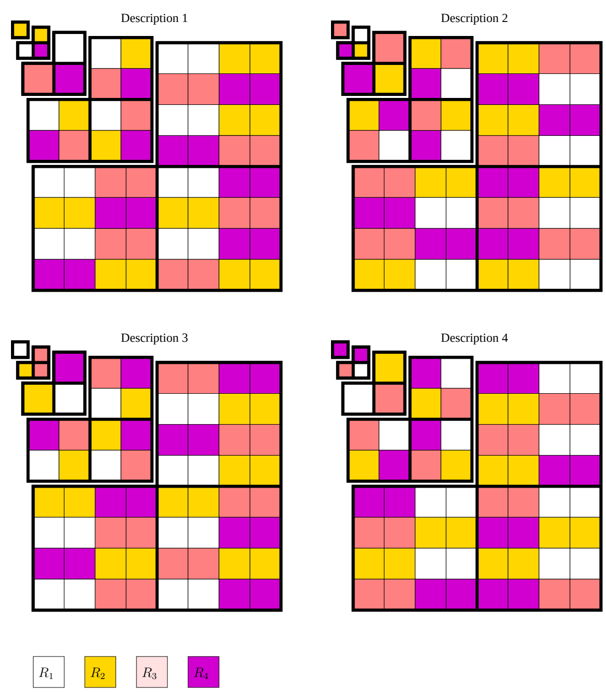

Figure 1: RDMC allocation for $N = 4$ descriptions.

If a subband group does not contain enough CBs, not all the rate permutations will be employed. In general, the description generation can be summarized as in Procedure I.

Clearly, whereas the assumption of equivalent CBs is very reasonable for the highest decomposition levels, especially for smooth images, it becomes less sound for lower frequency subbands, and especially for the lowest LL subband. Another sub-optimality arises when the number of CBs for each subband group is too low to implement all the $N!$ rate permutations. For all these reasons, this heuristic alloca-

Procedure I: Description generation

Given the number of DWT decomposition levels, L
Given the number of descriptions, N
For $i = L$ **downto 1**
  Consider the 3 higher frequency subbands
   OR (**if** $i = 1$) the residual LL subband
  Identify the number of CBs Y
  Evaluate the subset cardinality $X = Y/N!$
  If $X \geq 1$
    1. Group X CBs into subsets (e.g. in raster scan order, band by band)
    2. Assign to the N descriptions the subsets encoded according to the rate permutations so that each description contains a different representation (in terms of rate) of any subset
  If $X < 1$ or $Y/N! \neq \lfloor Y/N! \rfloor$
    Assign to the N descriptions the yet not assigned CBs encoded according to subsets of the $N!$ rate permutations
end for

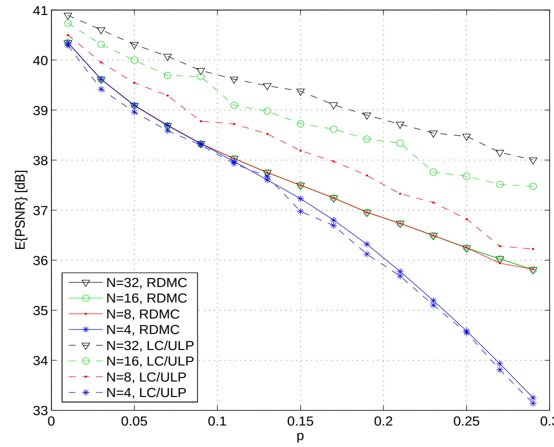

Figure 2: Expected PSNR [dB] vs. p(loss); RDMC and LC/ULP

tion algorithm does not guarantee that the descriptions are strictly balanced. The divergence of our heuristic allocation algorithm from the theory in [11], with the consequent unbalance, may be more evident when dealing with a large number of descriptions, along with a small number of CBs (stemming from either small image dimensions or large CB size, or both). The balance characteristics of the system are not discussed here for brevity. It can be shown that, although in some situations the heuristic allocation actually leads to unbalanced description sets, normally the unbalance is slight and can be neglected for all practical purposes.

At the decoder side, all the received descriptions are merged into a single bitstream, where, for each CB, the best representation is selected, which can be identified by simply determining the CB length. The resulting stream is then JPEG 2000 decoded. If all the descriptions are received, all the CBs representations at the best possible quality, identified by rate R_1, are available; thus, the quality in this case (central quality) is simply given by the distortion of a JPEG 2000 stream encoded at rate R_1. It is worth noticing that a standard JPEG 2000 decoder, not equipped by any preprocessing (merging) capability, is still able to decode any single description.

In [1], descriptions are built starting from JPEG 2000 streams encoded at two different rates. Analytic comparisons of RDMC with [1] are not reported here for brevity. However, we have shown that, using $N = 4$ descriptions, when more than one description is received and for several redundancy levels, RDMC outperforms the two-rate scheme. We have also compared RDMC with the scheme reported in [10] for JPEG 2000. In order to enable direct comparisons with the results reported in [10], we have used the same settings, i.e. the Lenna image of dimension 256×256 pixels, the (5, 3) filter bank and 3 levels of wavelet decomposition. The redundancy of RDMC is tuned so as to obtain the same side quality (i.e. when only a subset of descriptions is received) of the benchmark algorithm (≈ 23.6 dB). Under these conditions, RDMC yields a central quality of 35.5 dB against 27.4 dB of [10] encoded at the same overall rate (≈ 1.25 bpp). From these results, we can conclude that RDMC is one of the best state-of-the-art MDC algorithm for JPEG 2000 data and $N > 2$.

3. COMPARING RDMC AND ULP

The goal of the experiments reported in this section is to validate the graceful degradation capabilities of RDMC, and to provide comparisons with the ULP scheme in [9] (labeled *LC/ULP* in the following) applied to the same layered source. The simulation settings are:

- JPEG 2000 codec engine from the OpenJPEG libraries [7], generalized to support the encoding/decoding procedure of an arbitrary number of descriptions (for the RDMC scheme).
- Lenna image of dimension 512×512 pixels @ 1.2 bpp (similar results hold for different images, resolutions and coding rates, even though they are not reported here for brevity).
- DWT with (9,7) filter bank and 4 decomposition levels.
- Header information accounted for in the total available rate.
- Both algorithms designed so as to maximize the expected PSNR at the receiver side.

Fig. 2 reports the expected PSNR achieved by RDMC and LC/ULP, as a function of the probability of description loss p, and for several values of N. The goal of this first set of simulations is to compare the average performance of the two methods, designed so as to achieve the same number of descriptions. In the case $N = 4$, the average performance of the two methods is equivalent, whereas when $N = 8$, LC/ULP exhibits an average performance gain of about 0.5 dB for most values of p. The performance of RDMC saturates for $N \geq 8$. On the other hand, LC/ULP does not exhibit this saturation phenomenon (at least, in the addressed range of N values). For $N = 32$, LC/ULP exhibits a PSNR gain with respect to RDMC of about 1 dB for low values of p, and about 2 dB in the range of moderate to high p

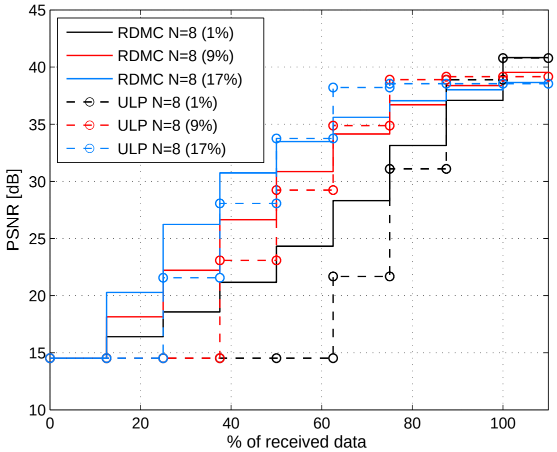

Figure 3: PSNR [dB] vs. percentage of received data. RDMC and LC/ULP, 8 descriptions

values. We can conclude that LC/ULP, on average, yields better end-to-end PSNR values than RDMC given the same p and total rate. More insight in the performance of the two algorithms can be gained considering their behaviour when not all the data are received. To this end, Fig. 3 reports the PSNR achieved by either algorithm and $N = 8$, as a function of the percentage of data received. The different curves refer to several p values, input to the algorithm to drive the optimization task. When all data are received and for the same p, both algorithms yield approximatively the same PSNR (e.g. about 40 dB when $p = 0.17$). This means that, given a probability of description loss, the two algorithms allocate approximatively the same amount of redundancy. However, the methods differ in the way that such redundancy is exploited to achieve robustness. In fact, the difference between the redundancy exploitation strategies can be appreciated by comparing the algorithms' performance when not all data is received. We can notice that the performance of RDMC increases gracefully with the percentage of received data, whereas the curves related to LC/ULP exhibit more abrupt variations. As a matter of fact, the quality achieved by RDMC steps up for each further description received, allowing for 8 different quality levels (being $N = 8$, each description amounts to 12.5% of total data). For example, when $p = 0.17$, a single description yields PSNR $\approx$ 20 dB and two descriptions yield PSNR $\approx$ 26 dB. A quality threshold of 30 dB is reached when 3 descriptions are received (37.5% of data). Even though the PSNR improvements are not constant, each description received achieves a PSNR gain of 2.5 dB on average. Notably, the first received descriptions get the highest PSNR improvements; this allows one to quickly achieve a satisfactory performance with a limited percentage of data received. In fact, each received description yields a quality improvement, and a number of quality levels equal to N is made available to enable graceful performance degradation. On the other hand, we note that LC/ULP is not able to achieve a PSNR value larger than 14 dB (corresponding to the image variance σ^2) unless at least 25% of data is received. In fact, the code optimization is such that the reception of a single description is not sufficient to guarantee the decoding of even a coarse quality version of the image. The threshold PSNR $= 30$ dB is achieved when at least 50% of data is received. This cliff effect, typical of channel coding, is even more evident if the p used to tune the encoder is as low as 0.01. In this case, LC/ULP allows for only 4 quality levels, and the coarsest one, corresponding to PSNR $\approx$ 22 dB, requires the reception of 62.5% of data (5 descriptions out of 8). The threshold PSNR $= 30$ dB is reached with 6 descriptions received (75% of data). On the other hand, RDMC designed for $p = 0.01$ achieves 8 quality levels, and still exhibits a typical graceful degradation behaviour. For example, the quality achieved with 50% of data received is about 24 dB, whereas LC/ULP yields PSNR $= \sigma^2$ in the same conditions.

Fig. 4 reports the PSNR achieved by the algorithms and $N = 32$, as a function of the percentage of data received. We can draw considerations similar to those made for $N = 8$. In particular, RDMC yields 32 quality levels, whereas LC/ULP yields a variable number of quality levels (8 for $p = 0.17$, only 4 for $p = 0.01$). With RDMC, the reception of one further description yields an average PSNR improvement of 0.78 dB. If p is 0.17 (0.01 respectively), the threshold of PSNR= 30 dB is achieved when 50% (75%) of data are received, corresponding to 16 and 24 descriptions. The performance of LC/ULP exhibits a steeper behavior. If p is 0.17 (0.01 respectively), the threshold PSNR $= 30$ dB is achieved when about 60% (90%) of data are received, corresponding to 19 and 29 descriptions.

From the reported performance comparisons, we can draw some conclusions. First of all, LC/ULP is significantly more efficient than RDMC if the expected PSNR performance is considered as quality metric. However, it is subject to a cliff effect, which impairs its graceful degradation performance. This is due to the "all-or-nothing" decoding property of erasure codes, and makes the system behavior very sensitive to variations of the p parameter with respect to the value used to setup the encoder. This fact may be detrimental in highly non stationary environments such as wireless networks.

On the other hand, RDMC exhibits graceful degradation properties to a larger extent than LC/ULP. In fact, it matches the MDC paradigm that each received description yields an improvement to the quality of the decoded data. When all data are received, both algorithms achieve the same performance, meaning that the total amount of redundancy is almost the same for a given total rate.

Clearly enough, other LC/ULP systems can be devised, which are optimized not in the expected PSNR sense, but in order to achieve a number of intermediate quality levels. We can expect that the behavior of such a LC/ULP scheme be closer to that of an MDC system, i.e. the expected PSNR is impaired and graceful degradation is achieved to a larger extent. However, it is worth noticing that neither RDMC nor LC/ULP are optimized in this sense. The study and comparison of MDC and ULP schemes with optimization function other than the expected PSNR is left to future research. From a practical standpoint, one could ask whether it is meaningful to have a large number of quality levels for image coding applications. The answer depends on the application details (e.g. transmission conditions, overall bandwidth, image resolution etc.). As a rule of thumb, whereas as many as 32, hardly distinguishable quality levels may do not make much sense, especially for small image resolutions, 4 or 8 quality levels are surely beneficial to enable graceful

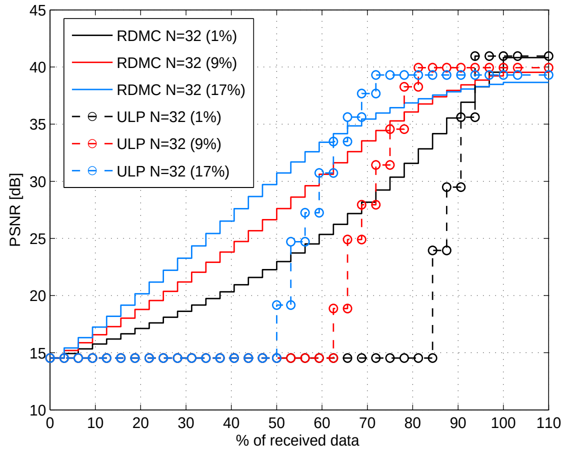

Figure 4: PSNR [dB] vs. percentage of received data. RDMC and LC/ULP, 32 descriptions

degradation when packets are received from many different sources (as in peer-to-peer networks) and the network scenario is heterogeneous. When $p = 17\%$, 50% of data is successfully received, and 8 descriptions are generated, both RDMC and LC/ULP guarantee a quality of approximately 34 dB. When the probability of description loss is lower (p=1% or 9%), with $N = 8$, and only a subset of data is received, RDMC achieves better performance.

Concerning the comparison of the computational complexity of RDMC and ULP, even though a thorough complexity analysis is beyond the scope of the present paper, RDMC is definitely less complex than ULP at both the encoder and the decoder side. In fact, ULP requires the implementation of a demanding code assignment optimization as in [6, 9]. At the decoder side, ULP requires RS decoding, whose complexity is quadratic with the code length. RDMC can be obtained by modifying the rate allocation and Tier-2 modules of a standard JPEG 2000 encoder. In fact, after the wavelet decomposition, N RD-optimized streams of the transformed image are generated, and then subsets are combined from such streams. The DWT and Tier-1 stages are evaluated only once for each image (as in a standard JPEG 2000 encoder), so that no modification to these stages is required. In fact, the RDMC encoder requires a JPEG 2000 encoding at the highest considered rate (lower rates are embedded in the higher ones), and a splitting of CBs into the descriptions. The sub-optimal algorithm employed in this paper adopts a deterministic splitting pattern, which does not require any optimization procedure.

4. CONCLUSIONS

We presented a performance comparison of a multiple description coding scheme with an unequal loss protection algorithm based on Reed Solomon FEC allocation. The study reveals that both schemes allocate the same total redundancy given the same total output rate. Multiple description coding, besides being computationally simpler, achieves a smoother performance degradation as the number of lost information increases, whereas LC/ULP yields superior performance in terms of expected PSNR.

5. ACKNOWLEDGEMENT

This paper is based on work performed in the framework of the Project SEA IST-214063, funded by the European Community.

6. REFERENCES

[1] E. Baccaglini, T. Tillo, and G. Olmo. A flexible R-D based multiple description scheme for JPEG 2000. *IEEE Signal Proc. Letters.*, 14(3):197–200, Mar. 2007.

[2] D. Comas, R. Singh, A. Ortega, and F. Marques. Unbalanced multiple-description video coding with rate-distortion optimization. *EURASIP Journal on Appl. Signal Proc.*, 2003:1:81–90, 2003.

[3] V. Goyal. Multiple description coding: compression meets the network. *IEEE SP Mag.*, 18(5):74–93, Sep 2001.

[4] M. Grangetto, E. Magli, and G. Olmo. Ensuring Quality of Service for image transmission: Hybrid loss protection. *IEEE Trans. on Image Proc.*, 13(6):751–757, 2004.

[5] M. Y. Kim and W. Kleijn. Comparative rate-distortion performance of multiple description coding for real-time audiovisual communication over the Internet. *IEEE Trans. on Comm.*, 54(4):625 – 636, April 2006.

[6] A. E. Mohr, E. A. Riskin, and R. E. Ladner. Unequal loss protection: graceful degradation of image quality over packet erasure channels through forward error correction. *IEEE Journal on Selected Areas in Commun.*, 18(6):819–828, Jun 2000.

[7] Open JPEG Libraries. www.openjpeg.org.

[8] R. Puri, K. Lee, K. Ramchandran, and V. Bharghavan. Forward error correction (FEC) codes based multiple description coding for Internet video streaming and multicast. *Signal Process.: Image Commun.*, 16(8):745–762, May 2001.

[9] R. Puri and K. Ramchandran. Multiple description source coding using forward error correction codes. In *Proc. of 33rd Asilomar Conf. on signals, systems and computers*, volume 1, pages 342–346, Oct 1999.

[10] K. Subbalakshmi and S. Somasundaram. Multiple description image coding framework for EBCOT. In *Proc. of IEEE Intl. Conf. on Image Proc.*, volume 3, pages 541 – 544, June 2002.

[11] T. Tillo, E. Baccaglini, and G. Olmo. A flexible multi-rate allocation scheme for balanced multiple description coding applications. In *Proc. of IEEE Intl. Workshop on Multim. Signal Proc.*, pages 1–4, Oct 2005.

[12] R. Venkataramani, G. Kramer, and V. Goyal. Multiple description coding with many channels. *IEEE Trans. on Info. Theory*, 49(9):2106–2114, Sept 2003.