Improving the performance of multiple description coding based on scalar quantization

Tammam Tillo, *Member, IEEE,* Gabriella Olmo, *Senior Member, IEEE*

***Abstract*— We propose an algorithm that improves the performance of rate-distortion based multiple description coding (RD-MDC). The gain is particularly significant in the high redundancy region, where RD-MDC suffers a major performance penalty with respect to MDC bounds. The improvement is obtained with negligible additional computational cost, by exploiting the coarse information also at the central decoder. The proposed method can be generalized to all MDC schemes that use scalar quantization, without modifying the quantizer structure. This feature guarantees the generation of descriptions that can be decoded without any modification of the decoder.**

***Index Terms*— Multiple description coding, scalar quantization, JPEG2000**

I. Introduction

Multiple description coding (MDC) is recognized as a useful tool for error resilient image and video transmission [1]. In MDC, two or more representations of the data are generated (*descriptions*), which yield mutually refinable information and can be independently decoded. Opposite to layered coding, no hierarchy exists among descriptions, and the reception of a specific base layer is not required. On the other hand, MDC implies an extra rate, necessary to encode redundant information. This is beneficial in case of losses, but impairs the rate-distortion (RD) performance when all descriptions are received. Actually, MDC outperforms layered coding when the network conditions are harsh, so that the allocation of extra rate is worthy [2].

When dealing with multimedia data, it is important that descriptions are compliant with standard co-decoding tools. In RD-MDC [3], data is encoded at two different rates, and divided into two subsets having similar RD characteristics; descriptions are then generated by properly combining subsets encoded at either rate. The algorithm has been integrated in JPEG2000, and outperforms other state-of-the-art methods. A similar approach is proposed in [4] for video; motion compensated temporal filtering and 2-D spatial wavelets are used to generate a spatio-temporal subband cube. Subbands are then divided into non-overlapping regions, which are coded at different rates in the different descriptions. In [5] the redundant slice option of H.264/AVC is used to implement an MDC scheme that merges coarse and fine representations of video slices.

The main disadvantage of these schemes is that, when a new description is received, the fine representation of a data block overrides the coarser one, which is then discarded. Actually, in [3] it is demonstrated that descriptions that carry duplicated information are always suboptimal, as the extra rate devoted to duplicated information cannot be exploited at the central decoder. To overcome this inefficiency, we propose a modification of RD-MDC (also applicable to similar algorithms based on quantization, such as [4], [5]). The key idea is to address a proper quantization stage avoiding duplication of information, and to modify the decoder so as to enable joint de-quantization of the coarse and fine data at the central decoder.

The rest of this paper is organized as follows. In Sec. II, we provide some background on RD-MDC. In Sec. III, we describe the proposed enhancement. Section IV describes the practical implementation issues of the proposed method in JPEG2000. In Sec. V we provide experimental results for both a Gaussian source and real images, while in Sec. VI we draw some conclusions.

The authors are with Department of Electronics - Politecnico di Torino, Corso Duca degli Abruzzi 24 - 10129 Torino - Italy - Ph.: +39-011-5644094 - FAX: +39-011-5644149 - E-mail: (tillo,olmo)@polito.it. Corresponding author: tammam.tillo@polito.it.

II. Background

RD-MDC assumes that the data consist of independent blocks. Two RD optimized streams are generated at rates R_1 and $R_2 < R_1$ bps, and distortion $D(R_1)$ and $D(R_2)$, $D(R)$ being the source RD function. Two balanced descriptions, encoded at rate $(R_1+R_2)/2$ bps, are generated by combining blocks encoded at either rate.

The central decoder selects the best representation of each block, discarding the coarsest one. It yields a central distortion $D_0 = D(R_1)$, equal to that delivered by a single description coding scheme at rate R_1. The rate $R_2 = R_t - R_1$, where R_t is the total rate, is the redundancy (extra rate) of the RD-MDC scheme. The side distortion can be evaluated as $D_1 = D_2 = [D(R_1) + D(R_2)]/2$.

In order to assess the performance gap of RD-MDC with respect to MDC bounds, let us consider a memoryless, first order stationary Gaussian process with zero mean and variance σ^2, encoded into two balanced descriptions using RD-MDC. The RD curves of this system can be analytically evaluated as

$$\begin{aligned} D_0 &= C_q(R_1)\sigma^2 2^{-2R_1} \\ D_1 &= \frac{\sigma^2}{2}\left[C_q(R_1)2^{-2R_1} + C_q(R_2)2^{-2R_2}\right] \end{aligned} \quad (1)$$

where $C_q(R)$ depends on quantization [6].

The RD-MDC best possible central distortion, which assumes ideal quantization, is given by (1) with $C_q(R_i) = 1$. On the other hand, the performance of a practical RD-MDC scheme can be evaluated assuming entropy-constrained quantization, i.e. $C_q(R_i) \approx \frac{e\pi}{6}$ [7]. In Fig. 1, both these situations are compared to the Ozarow bound D_b [8]. The

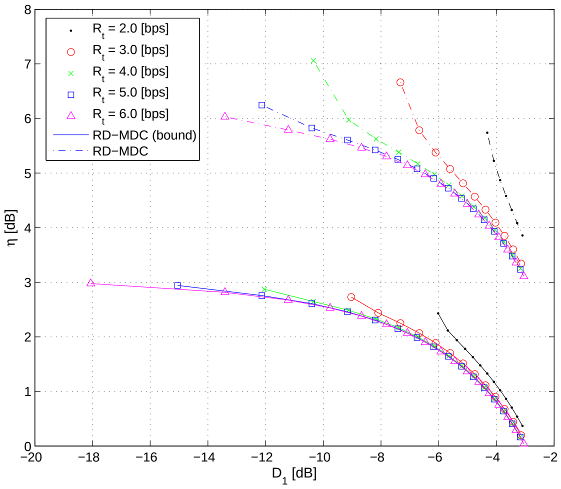

Fig. 1. Performance gap η versus D_1, for various R_t and $R_2 \in [0, R_t/2]$.

performance gap $\eta = 10\log_{10}(D_0/D_b)$ [dB] is reported versus D_1, for $R_t = 2-6$ bps. The redundancy range $[0\,,\,R_t/2]$ has been scanned while keeping R_t constant. We can notice that, with ideal quantization and in the low redundancy region, RD-MDC tends to approach the theoretical curve, whereas in the high redundancy region the gap increases up to 3 dB. In case of entropy constrained quantization, the performance gap is even larger, although it is mostly due to the suboptimal quantization scheme and not to the MDC algorithm. In any case, there is a significant margin of performance enhancement.

III. *Improving RD-MDC*

Our objective is to modify the way redundancy is inserted in RD-MDC, so as to make it exploitable also for central decoding. In other words, coarse data, which at present are discarded at the central decoder, should help improving fine quality to some extent. The idea is to address proper quantizers so that, by interlacing the two quantized versions of the same coefficient, it is possible to derive a joint de-quantization stage able to enhance the fine decoding.

For the sake of simplicity, and without loss of generality, let us focus on uniform scalar quantization (a generalization will be discussed in Sect. III-A). The high and low resolution versions are obtained with quantization steps Δ_f and Δ_c respectively; for simplicity, we assume that $\frac{\Delta_c}{\Delta_f}$ is integer. Let Q_{fc} denote the quantizer configuration, with f and c referring to fine and coarse quantizers, and $f, c \in \{T, R\}$, where T and R label mid-tread and mid-rise quantizers.

When a fine quantization interval is overlapped with two coarse intervals, the reconstruction uncertainty may be possibly reduced. For example, for Q_{TT} and $\frac{\Delta_c}{\Delta_f} = 2n$, two fine intervals per each coarse interval can be refined. On the other hand, if $\frac{\Delta_c}{\Delta_f} = 2n+1$, all fine intervals are embedded in the coarser interval and no refinement is possible. We want to assess the possible performance improvement of a refinement stage. To this end, we evaluate the *refined central distortion* d_r, under the high rate assumption [9], i.e. assuming that the probability density function (pdf) of each quantized coefficient is uniform within each quantization interval.

Let us focus on Q_{TT} and $\frac{\Delta_c}{\Delta_f} = 2n$. For the two fine intervals that can be refined, as they are overlapped with two coarse intervals, an equivalent quantization step size $\frac{\Delta_f}{2}$ can be defined, yielding a contribution to the quantization distortion $\frac{1}{12}\left[\frac{\Delta_f}{2}\right]^2 = \frac{1}{4}d_f$, where $d_f = \frac{\Delta_f^2}{12}$ is the distortion of a uniform quantizer with step Δ_f. The probability that the sample falls in one of the overlapped intervals, conditioned by the sample belonging to a given coarse interval, is $\frac{\Delta_f}{\Delta_c}$. On the other hand, the remaining fine intervals cannot be refined, and each of them yields a contribution d_f to the overall distortion; the conditional probability that the sample falls in one of the non overlapped interval is $1 - \frac{\Delta_f}{\Delta_c}$. The refined central distortion d_r^{TT} can be then evaluated as

$$\begin{cases} d_f\left(1 - \frac{\Delta_f}{\Delta_c}\right) + \frac{\Delta_f(\Delta_f/2)^2}{12\Delta_c} = d_f\left(1 - \frac{3}{4}\frac{\Delta_f}{\Delta_c}\right) & ; \frac{\Delta_c}{\Delta_f} = 2n \\ d_f & ; \frac{\Delta_c}{\Delta_f} = 2n+1 \end{cases} \quad (2)$$

Following similar reasoning, one can work out that d_r^{RT} has the same expression as (2), but with the refinement terms obtained for $\frac{\Delta_c}{\Delta_f} = 2n+1$. Moreover, $d_r^{TR} = d_f(1 - \frac{3}{4}\frac{\Delta_f}{\Delta_c})$ $\forall \frac{\Delta_c}{\Delta_f} \in \mathbb{Z}$ and $\mathbb{Z}$ denoting strictly positive integers. Finally, no joint dequantization of two mid-rise quantizers is possible, as no fine interval is overlapped with a coarse one in this case. These results can be generalized to non integer $\frac{\Delta_c}{\Delta_f}$, defining γ as the overall fraction of fine intervals that can be refined. It can be easily shown that $d_r \geq d_f(1 - \frac{3}{4}\gamma)$, with equality holding when each refineable interval is equally overlapped by two coarser intervals.

TABLE I

GAIN IN CENTRAL SNR [DB] VERSUS γ; GAUSSIAN SOURCE

γ	1/10	1/8	1/4	1/3	1/2	2/3	1
SNR gain [dB]	0.34	0.43	0.90	1.25	2.04	3.01	6.02

In Tab. I, we report the gain in terms of central SNR versus γ, with respect to RD-MDC, for a Gaussian source and assuming entropy constrained quantization. It can be noticed that the maximum gain is obtained when $\gamma \rightarrow 1$, which represents the case of a mid-tread and a mid-rise uniform quantizer with the same step size. In fact, this situation is equivalent to assigning one more bit to the fine quantizer, leading to the well known 6dB SNR gain.

A. Scalar quantization with deadzone

The idea of quantization refining can be easily generalized to schemes such as non uniform or entropy constrained scalar quantization. As an example, we address two scalar quantizers with deadzone. These are mid-tread quantizers, characterized by a zero bin width larger than the other bins. Their interest arises from the fact that scalar quantization with deadzone is widely used in transform-based image and video coding such as JPEG2000 and H.264/AVC [10].

Let us assume that the zero-bin width is $\eta\Delta_t$, with $t \in \{f, c\}$ and $\eta \geq 1$. In many applications η is set to 2, as this choice yields reasonably good performance for image applications [10]. For the sake of simplicity, let us assume that $\frac{\Delta_c}{\Delta_f} = n \in \mathbb{Z}$. Under this assumption, it can be easily understood that the fine deadzone quantizer is fully embedded in the coarse one for $\eta = \frac{2n}{k}$, $k, n \in \mathbb{Z}$ and $k < n+1$; as a consequence, no improvement can be obtained with joint dequantization in this case. On the other hand, if $\eta = \frac{2n-1}{k}$, $k, n \in \mathbb{Z}$ and $k < n+1$, each coarse interval can help refining one fine interval, leading to an equivalent quantization step size $\frac{\Delta_f}{2}$. In this situation, the refined central distortion d_r can be evaluated as:

$$d_r^{dd} = d_f^{dd} - \frac{3}{4}\frac{\Delta_f^2}{12}P \tag{3}$$

where d_f^{dd} represents the distortion of the fine deadzone quantizer, and P represents the probability that a fine quantization interval uniformly overlaps two coarse intervals. P can be easily worked out as $P \leq \frac{\Delta_f}{\Delta_c}$, with equality holding for $\eta = 1$, i.e. uniform quantization.

IV. Practical implementation using JPEG2000

In JPEG2000, the image is first DWT transformed, and the generated coefficients are quantized by a high rate quantizer. Then, the coefficients of each subband are divided into non-overlapping rectangular areas (codeblocks - CBs), which are bit-plane encoded in the so-called Tier-1 module. The bitstream is organized by the rate allocator (Tier-2 module) into a sequence of layers, each layer containing contributions from each CB. The block truncation points associated with each layer are optimized in the RD sense [11]. Truncating the CBs stream is equivalent to using a quantizer with a scaled version of the subband original step size, where the scaling factor is a power of two.

The enhanced RD-MDC procedure starts generating four streams, at rates R_1 and $R_2 < R_1$ bpp and using the JPEG2000 standard quantizer with step sizes Δ_1 and Δ_2. CBs are divided into two sets, equivalent in the RD sense as in [3]. Description 1 is then obtained combining the first CB set encoded at $(R_1,\ \Delta_1)$ with the second set encoded at $(R_2,\ \Delta_1)$. Conversely, description 2 is generated from the first CB set encoded at $(R_2,\ \Delta_2)$ and the second set encoded at $(R_1,\ \Delta_2)$. This procedure yields almost balanced descriptions, each of which is encoded at rate $(R_1 + R_2)/2$, and is JPEG2000 compliant stream. The coarse and fine quantizers are designed so as to be compatible with JPEG2000 - Part 1, which uses a mid-tread uniform quantizer with central deadzone. In our experiments, we have selected $\frac{\Delta_1}{\Delta_2} = 1.5$, which can be shown to yield the maximum possible overlapping of these quantizers. This configuration leads to $\gamma \leq \frac{1}{3}$, with the equality holding when $R_2 \rightarrow R_1$ (high redundancy region).

At the decoder side, if a description is lost, the received one is JPEG2000 decoded yielding the side quality. When both descriptions are received, the actual (possibly refined) quantization interval of each coefficient must be identified. To this end, the Tier-1 decoding operations are performed twice, working on the coarse and fine quantized coefficients respectively. This is carried out by forcing three further decoding passes for each CB in the arithmetic decoder in the Tier-1 module. The first additional coding pass is the coding pass immediately following the last coding pass of a non enhanced RD-MDC JPEG 2000 decoder. Opposite to standard Tier-1 decoding passes, which evaluate the dequantized value, in the enhanced decoder the two decision intervals of the coefficient are worked out. Finally, the central decoder evaluates the intersection of the fine and coarse quantization intervals, which is used to reconstruct a possibly refined coefficient. It is worth noticing that the extra complexity of these operations is negligible.

A. Comment

In [12], a modified design of MD scalar quantization (MDSQ) [13] is proposed, making use of two staggered quantizers with bins $\Delta/2$ apart, yielding the coarse quality descriptions. A refinement stage is implemented by further dividing the equivalent joint quantizer into finer bins, which are entropy coded and splitted into the descriptions. The algorithm has been incorporated in a wavelet based image coder, and comparisons are made with MDSQ used in the same context. Notably, the results yielded by this algorithm, as reported in [12], are almost identical to those obtained with RD-MDC [14], even though this latter, besides being compatible with JPEG2000, addresses a much simpler quantization stage.

V. Experimental results

A. Gaussian source

Simulations have been carried out on vectors of length $m = 2400$ samples of a zero mean and unit variance memoryless Gaussian random process. We have employed two entropy constrained quantizers, obtained with uniform scalar quantizers followed by an arithmetic coder. We have considered R_t in the range $3-6$ bps. The redundancy has been tuned in $[0, R_t/2]$. For each value of the redundancy, 400 trials have been run.

Fig. 2 reports the central distortion D_0 [dB] versus the side distortion D_1 [dB], for both the basic and the enhanced RD-MDC schemes, using Q_{TT}. The performance of RD-MDC when $\frac{\Delta_c}{\Delta_f}$ is integer is also reported for clarity. These latter curves are obtained by scanning a range of different values of Δ_f and Δ_c, while maintaining $\frac{\Delta_c}{\Delta_f}$ constant. From these results we can appreciate the performance improvements due to the enhancement stage, especially significant in the high redundancy region (low side distortion). As stated by Eq. 2, no improvement is obtained when $\frac{\Delta_c}{\Delta_f}$ is odd.

Similar results (not reported here for brevity) have been worked out using Q_{TR}, Q_{RT} and Q_{RR}. As expected, simulations show that using two mid-rise quantizers represents the worst possible configuration. It is also worth noticing that the simulation results for Q_{TR} and Q_{RT} have shown that in the region Δ_c less than $4\Delta_f$, the best performance in terms of side and central quality is obtained for $\Delta_c = \Delta_f$; a 6 dB gain of the proposed method over the RD-MDC can be appreciated at this operational point. Consequently, if one has to operate in such a region, the best possible choice is to use quantizers with the same step size.

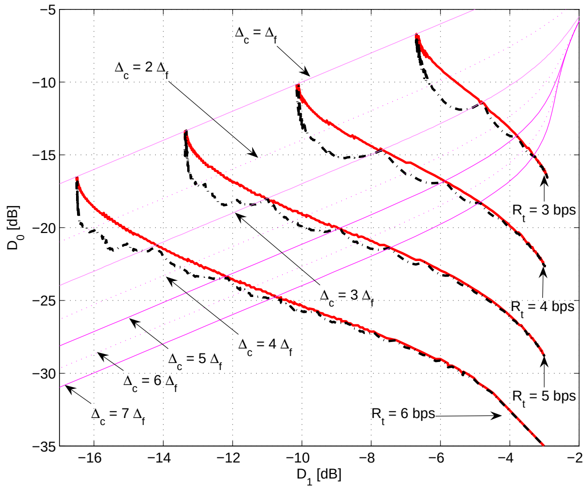

Fig. 2. Basic RD-MDC (solid thick curves) and enhanced RD-MDC (dash-dotted thick curves) central distortion [dB] versus side distortion [dB] along with the performance curves when $\frac{\Delta_c}{\Delta_f}$ is integer. Q_{TT}. Zero mean iid Gaussian random process.

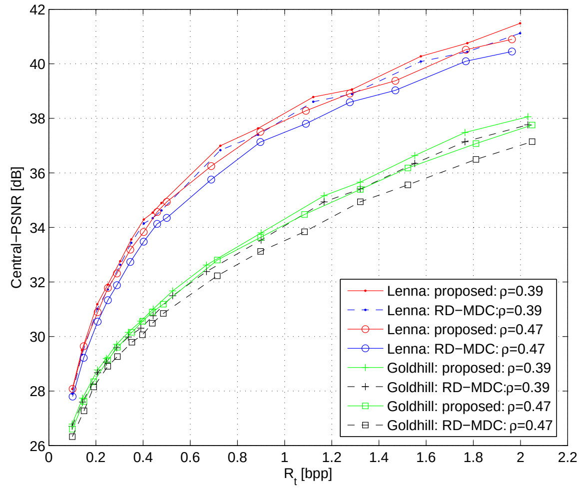

Fig. 3. Central PSNR versus the total rate for the basic and enhanced RD-MDC JPEG2000 algorithm; Lenna and Goldhill images.

B. JPEG2000-based MD codec

In this section we want to validate the proposed MDC scheme applied to the JPEG2000. We have used the codec engine from the OpenJPEG libraries [15], generalized to support the encoding/decoding procedure of two descriptions as in [3]. In Fig 3, the central PSNR of basic and enhanced RD-MDC are reported versus the total rate, for Lenna and Goldhill 512x512 images. The relative redundancy $\rho = \frac{R_2}{R_1+R_2}$ has been set to 0.39 and 0.47. The two algorithms are tuned so as to yield the same side distortion (not reported for brevity); 4 levels of DWT decomposition have been applied.

As already discussed, the step-sizes ratio of 1.5, addressed in our experiments, leads to $\gamma \leq 1/3$. From Tab. I, we can expect a SNR gain of 1dB for $\gamma = 1/3$; this gain can be approached at high redundancy levels ($R_2 \approx R_1$). from the experimental results, we can appreciate a gain of 0.7 dB; this can be explained by the fact that zero-quantized coefficients cannot be refined, and those coefficients are more present in a deadzone quantizer, such as that adopted in JPEG2000, whereas, the results of Tab. I refers to a strict uniform scalar quantizers.

VI. Conclusions

In this paper we propose an approach to improve the performance of RD-MDC methods. The improvement is obtained with nearly no additional computational cost by exploiting the different information carried by the two descriptions at the central decoder. At the highest redundancy level it is possible to achieve a 6dB gain with respect to the RD-MDC, for a Gaussian source. Moreover, the proposed algorithm can be adopted in those standards that use scalar quantization without modifying the quantizer structure. This feature guarantees the generation of descriptions that can be decoded with standard decoders, thus maintaining the backward compatibility feature of the RD-MDC approach. Future developments include the extension of the technique to other MDC schemes, possibly with $N > 2$ descriptions.

References

[1] V.K. Goyal, “Multiple description coding: compression meets the network,” *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 74–93, Sept. 2001.
[2] R. Singh, A. Ortega, L. Perret, and W. Jiang, “Comparison of multiple description coding and layered coding based on network simulations,” in *Proc. Visual Communications and Image Processing*, Jan. 2000, vol. 3974, pp. 929 – 939.
[3] T. Tillo, M. Grangetto, and G. Olmo, “Multiple description image coding based on lagrangian rate allocation,” *IEEE Trans. on Image Processing*, Mar 2007.
[4] E. Akyol, A. Tekalp1, and M.R. Civanlar, “Scalable multiple description video coding with flexible number of descriptions,” in *Proc. of the IEEE International Conference on Image Processing*, 2005.
[5] T. Tillo, M. Grangetto, and G. Olmo, “Redundant slice optimal allocation for H.264 multiple description coding (to appear),” *IEEE Transactions on Circuits and Systems*, Jan 2008.
[6] R. M. Gray and D. L. Neuhoff, “Quantization,” *IEEE Tran. on Info. Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.
[7] V.K. Goyal, “Theoretical foundations of transform coding,” *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 9–21, Sept. 2001.
[8] L. Ozarow, “On a source-coding problem with two channels and three receivers,” *Bell Syst. Tech. J.*, vol. 59, no. 10, pp. 1909–1921, Dec. 1980.
[9] N.S. Jayant and P. Noll, *Digital Coding of waveforms*, Prentice Hall, 1984.
[10] Yanling Xu and Yuanhua Zhou, “On optimal entropy-constrained deadzone quantization,” *IEEE Transactions on Consumer Electronics*, vol. 50, no. 4, pp. 1135 – 1141, Nov 2004.
[11] ISO/IEC 15444-1, *Document ISO/IEC 15444-1:2004, “JPEG 2000 part 1”*.
[12] C. Tian and S. Hemami, “A new class of multiple description scalar quantizer and its application to image coding,” *IEEE Signal Processing Letters*, vol. 12, no. 4, pp. 329–332, April 2005.
[13] V. Vaishampayan, “Design of multiple description scalar quantizers,” *IEEE Trans. on Information Theory*, vol. 39, no. 3, pp. 821–834, May 1993.
[14] T. Tillo and G. Olmo, “A novel multiple description coding scheme compatible with the JPEG 2000 decoder,” *IEEE Signal Processing Letter*, vol. 11, no. 11, pp. 908–911, Nov. 2004.
[15] Open JPEG Libraries, “www.openjpeg.org,” .