Research Article

Scalable Multiple-Description Image Coding Based on Embedded Quantization

Augustin I. Gavrilescu,[1] Fabio Verdicchio,[1] Adrian Munteanu,[1] Ingrid Moerman,[2] Jan Cornelis,[1] and Peter Schelkens[1]

[1] *Department of Electronics and Informatics (ETRO), Vrije Universiteit Brussel (VUB) and Interdisciplinary Institute for Broadband Technology (IBBT), Pleinlaan 2, 1050 Brussels, Belgium*

[2] *Department of Information Technology (INTEC), Universiteit Gent (UGent) and Interdisciplinary Institute for Broadband Technology (IBBT), St.-Pietersnieuwstraat 41, 9000 Ghent, Belgium*

Scalable multiple-description (MD) coding allows for fine-grain rate adaptation as well as robust coding of the input source. In this paper, we present a new approach for scalable MD coding of images, which couples the multiresolution nature of the wavelet transform with the robustness and scalability features provided by embedded multiple-description scalar quantization (EMDSQ). Two coding systems are proposed that rely on quadtree coding to compress the side descriptions produced by EMDSQ. The proposed systems are capable of dynamically adapting the bitrate to the available bandwidth while providing robustness to data losses. Experiments performed under different simulated network conditions demonstrate the effectiveness of the proposed scalable MD approach for image streaming over error-prone channels.

1. INTRODUCTION

Performing compression is of paramount importance in modern multimedia systems in order to improve the bandwidth usage and to reduce the costs associated with the signal transport. Sending image and video data in an efficient way over an ideal (error-free) channel basically consists in removing the redundancy from the input signal. Additionally, in the context of browsing through large data sets and fast access to large images transmitted over low-bandwidth channels, employing a compression scheme with progressive transmission capabilities is of critical importance. Scalable image coding technologies [1] enable the media providers to generate, in a single compression step, a unique bitstream from which appropriate subsets, producing different visual qualities, frame rates, and resolutions can be extracted to meet the preferences and the bitrate requirements of a broad range of clients. Moreover, at the decoder side it is possible to refine the image quality as more data is received.

On the other hand, in data communications over unreliable channels (e.g., mobile wireless or best-effort networks), achieving overall performance optimization is not always similar to minimizing the redundancy within the input stream. Hence, streaming data over networks involves much more than taking the output of a standard coder and writing it to the socket, justifying entirely the need for a new paradigm called robust source coding. In this context, several multiple-description (MD) coding techniques (e.g., [2–6]) have been introduced to efficiently overcome the channel impairments over diversity-based systems, allowing the decoders to extract meaningful information from a subset of the transmitted data. MD coding systems are able to generate more than one description of the source, such that (i) each description independently describes the source with certain fidelity, and (ii) when more than one description is available at the decoder, these descriptions can be combined to enhance the quality of the decoded signal.

In light of the above, we may conclude that on one hand, modern image compression systems have to provide scalability in order to meet the heterogeneous nature of networks and clients in nowadays communication systems, and on the other hand, efficient robust coding needs to be provided as well, in order to cope with the challenges posed by multimedia communication over error-prone channels. The paper

addresses this combined problem, and proposes a class of scalable MD image coding systems as a solution to this problem. The key component in these systems is given by embedded multiple-description scalar quantization (EMDSQ) [7–11]. EMDSQ belongs to the broad family of multiple-description coding approaches based on scalar quantization. The basic principle of MD based on scalar quantization (MDSQ) was firstly proposed by Vaishampayan in [3], while the optimal design of central and side quantizers has been extensively studied for the fixed-rate case in [12, 13]. The classical MDSQs are fixed-rate quantizers. This will lead to the impossibility to design embedded coding schemes having the ability to refine the image quality at the decoder side as more information is received. This problem has been solved in [14], where multiple-description uniform scalar quantizers (MDUSQ) have been proposed, simultaneously enabling multiple-description coding and a scalable encoding of the input source.

More recently, we proposed generic EMDSQ [7–11] producing double-deadzone central quantizers, known to be optimal and very nearly optimal at high and low rates, respectively, [1]. This will allow the coding systems employing EMDSQ to provide state-of-the-art coding results in data communications under realistic network conditions, and due to the embedded nature of EMDSQ, also to support progressive transmission and fine-grain rate adaptation of the output stream.

In contrast to MDUSQ, the design of EMDSQ treats the generic case of an arbitrary number of descriptions. Another important feature of EMDSQ consists in their unique ability to control not only the overall redundancy but also the redundancy, at each distinct quantization level between the descriptions.

In this paper, we propose a new scalable MD coding approach for images that couple EMDSQ and a customized version of our wavelet-based quadtree (QT) coding algorithm, originally proposed in [15, 16]. The system will be referred to as MD-QT.

In the proposed MD-QT approach, EMDSQ is used in order to produce a scalable MD representation of the wavelet coefficients, while QT coding is employed in order to encode for each description the localization information, indicating the positions of the coefficients that are found to be significant at each quantization level. However, simply coupling EMDSQ with QT coding only provides a multiple-description representations of the quantization indices. A critical design aspect is to extend the MD paradigm and to produce multiple representations of the localization information as well. This idea is followed in our design of an enhanced MD-QT image coding system. The experimental results show that in the context of transmission over error-prone packet networks, the enhanced MD-QT system provides a substantial performance increase in comparison to a simple MD-QT approach.

The rest of the paper is structured as follows. Section 2 presents the EMDSQ focusing on two main issues: Section 2.1 describes the way EMDSQs are constructed by recursive splitting of an index assignment (IA) matrix, while Section 2.2 describes the mechanism that allows for a flexible redundancy allocation for each distinct quantization level. The proposed MD-QT algorithm is detailed in Section 3.1. It is followed with Section 3.2 where an enhanced version of the algorithm extending the MD coding paradigm to all types of encoded information is presented. The experimental results are presented in Section 4, illustrating the rate-distortion performances of the proposed still-image codecs in data transmission over error-prone networks. Finally, Section 5 summarizes the conclusions of our work.

2. EMBEDDED MULTIPLE-DESCRIPTION SCALAR QUANTIZATION

2.1. Embedded index assignment

An EMDSQ is an embedded scalar quantizer designed to work in a diversity-based communication system. We can define EMDSQ as a set of embedded side quantizers generating two descriptions denoted by $Q_m^0, Q_m^1, \ldots, Q_m^P$, where m represents the description index ($m = 1, 2$) and p represents the quantization level ($0 \leq p \leq P$). The corresponding set of embedded central quantizers is denoted by $Q^0, Q^1, \ldots, Q^P$. For any p, $p < P$, the partition cells of any quantizers Q_m^p and Q^p are embedded in the partition cells of the quantizers $Q_m^P, Q_m^{P-1}, \ldots, Q_m^{p+1}$ and $Q^P, Q^{P-1}, \ldots, Q^{p+1}$, respectively. The embedded side quantizer index $\mathbf{q}_m^p = (q_m^P, q_m^{P-1}, \ldots, q_m^0)$ is the output of Q_m^p for a source sample $x \in \mathbb{R}$ [1].

Consider an IA matrix $\mathbf{M}$. The matrix $\mathbf{M}$ is recursively split along each dimension $m = 1, 2$ for a number of P levels. Then, for every level p, with $0 \leq p \leq P$, $\mathbf{M}$ can be considered as a block matrix of the form

$$\mathbf{M} = [\mathbf{B}_{j_1, j_2}^p]_{1 \leq j_m \leq J_m^p}, \quad \forall p, \, 0 \leq p \leq P, \tag{1}$$

and it is represented as follows:

$$\mathbf{M} = \begin{pmatrix} \mathbf{B}_{11}^p & \mathbf{B}_{12}^p & \cdots & \mathbf{B}_{1J_2^p}^p \\ \vdots & \vdots & & \vdots \\ \mathbf{B}_{J_1^p 1}^p & \mathbf{B}_{J_1^p 2}^p & \cdots & \mathbf{B}_{J_1^p J_2^p}^p \end{pmatrix}. \tag{2}$$

The corresponding side quantizers for the dimension m at level p will be Q_m^p, containing a number of J_m^p cells. In order to obtain balanced descriptions at each distinct quantization level p, the condition $J_1^p = J_2^p = J_p$ for any p, $0 \leq p < P$, has to be satisfied. Further, we consider an arbitrary block $\mathbf{B}_{ij}^{p+1}$ within $\mathbf{M}$ and the corresponding cells $C_{\mathbf{q}_m^{p+1}}$ within the side quantizers Q_m^{p+1}. The block $\mathbf{B}_{ij}^{p+1}$ at level p is split into $L_1^p \times L_2^p$ blocks at the lower level p, and thus it can be defined by its component blocks as $\mathbf{B}_{j_1 j_2}^{p+1} = [\mathbf{B}_{i_1 i_2}^p]_{1 \leq i_m \leq L_m^p}$ for any $1 \leq j_m \leq J_{p+1}$. It is noticeable that in order to have balanced descriptions at each quantization level p, $0 \leq p \leq P$, the condition $L_1^p = L_2^p = L_p$ has to be satisfied. The resulting blocks $\mathbf{B}_{i_1 i_2}^p$ correspond to cells of the form $C_{\mathbf{q}_m^{p+1} q_m^p}$ within the side quantizers Q_m^p. To any source sample $x \in C_{\mathbf{q}_m^{p+1} q_m^p}$ one associates the quantizer index $\mathbf{q}_m^p$ obtained by concatenating $\mathbf{q}_m^{p+1}$ with q_m^p, that is, $\mathbf{q}_m^p = (\mathbf{q}_m^{p+1}, q_m^p)$. This allows us
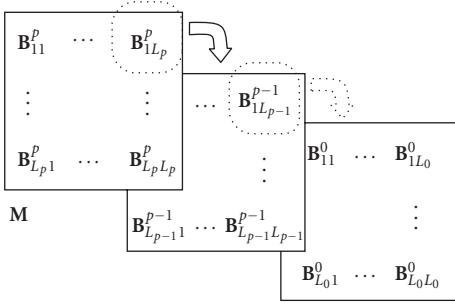
FIGURE 1: Recursive block-matrix decomposition.

to obtain the indices of lower-rate quantizers by discarding the components of higher-rate quantizers, similar to embedded scalar quantization [1]. Hence, $C_{\mathbf{q}_m^p} = C_{\mathbf{q}_m^{p+1} q_m^p}$ are embedded in $C_{\mathbf{q}_m^{p+1}}$ for any p and m, and the side-quantizer indices are embedded. This shows that the recursive splitting of $\mathbf{M}$ generates embedded side quantizers. Consequently, the central quantizers are embedded as well. The relation between the numbers of cells contained in any Q_m^p and L_p is

$$J_p = \prod_{i=p}^{P} L_i. \tag{3}$$

One can conclude that by splitting $\mathbf{B}_{j_1 j_2}^{p+1}$, a number $L_p \times L_p$ of blocks $\mathbf{B}_{i_1 i_2}^p$ will result at the lower level p. The recursive splitting of $\mathbf{M}$ along each dimension is illustrated in Figure 1.

Furthermore, following the recursive splitting of a balanced embedded IA $\mathbf{M}$ as described above, we define a strictly increasing sequence of natural numbers b_p as follows:

$$b_p = \begin{cases} 1, & p = 0, \\ \prod_{i=0}^{p-1} L_i, & p > 0. \end{cases} \tag{4}$$

Given b_p, any output index n_m of Q_m^0 for a source sample $x \in \mathbb{R}$ can be uniquely represented in the Smarandache general numeration base [17] as follows:

$$n_m - 1 = \left(\overline{q_m^P, q_m^{P-1}, \ldots, q_m^0} \right)_{(SG)} = \sum_{p=0}^{P} q_m^p b_p, \tag{5}$$

where $0 \le q_m^p \le L_p - 1$. It is noticeable, as described above, that $\mathbf{q}_m^p = (q_m^p, q_m^{P-1}, \ldots, q_m^p)$ is an embedded quantizer index representing the output of the embedded side quantizer Q_m^p.

As previously indicated in the literature on scalar quantization, an entropy-constrained uniform quantizer is very nearly optimal for input sources with smooth PDFs (probability density functions) [18, 19]. For embedded quantization, a notable example where all embedded quantizers can be optimal is the uniform case [1]. The conditions that need to be satisfied by an embedded IA defined as $\mathbf{M} = [\mathbf{B}_{j_1, j_2}^p]_{1 \le j_m \le J_p}$ to yield embedded uniform central quantizers are as follows:

(1) the corresponding central quantizer for the highest rate (Q^0) has to be uniform;
(2) for all p, $0 \le p \le P$, in each $\mathbf{B}_{j_1 j_2}^p \ne [\mathbf{0}]$ ($1 \le j_m \le J_p$), a constant number of consecutive nonzero indices are mapped (by $[\mathbf{0}]$ we denote the zero matrix).

The proof of the theorem from which the above conditions are derived is given in [8].

We define the operator $nnz(\mathbf{M})$ determining the number of nonzero elements contained in a matrix $\mathbf{M}$. The number of blocks $\mathbf{B}_{j_1 j_2}^p \ne [\mathbf{0}]$ contained in $\mathbf{M}$ is determined via the nonzero-blocks operator $nzb(\mathbf{M}, p)$. It is noticeable that $nzb(\mathbf{M}, 0) = nnz(\mathbf{M})$. Pursuant to this definition and based on the conditions above, for any of the block matrices $\mathbf{B}_{j_1 j_2}^{p+1} = [\mathbf{B}_{i_1 i_2}^p]_{1 \le i_m \le L_p}$, the number of blocks $\mathbf{B}_{j_1 j_2}^p \ne [\mathbf{0}]$ is constant and is given by $N_p = nzb(\mathbf{B}_{j_1 j_2}^{p+1}, p)$. Additionally, at level P, $N_P = nzb(\mathbf{M}, P)$ represents the number of blocks $\mathbf{B}_{j_1 j_2}^P \ne [\mathbf{0}]$ within $\mathbf{M}$. The total number of indices mapped in each block $\mathbf{B}_{j_1 j_2}^p$ at level p is then $nnz(\mathbf{B}_{j_1 j_2}^p) = \prod_{i=0}^{p} N_p$. It results that the number of cells contained in Q^p is $N = \prod_{i=p}^{P} N_p$.

2.2. Redundancy control

Denote by R_m the rates and by $D_m(R_m)$ the corresponding side-description distortions. Also, denote by D_0 the central distortion. In single-description (SD) coding, one minimizes D_0 for a given rate R_0. The redundancy is the bitrate sacrificed by MD coding compared to SD coding in order to achieve the same central D_0 distortion:

$$\rho = R_1 + R_2 - R_0. \tag{6}$$

In the embedded case, the overall rate is the cumulated value of rates corresponding to each distinct quantization level, and it can be written analytically as $R_0 = \sum_{p=0}^{P} R_0^p$, where R_0^p represents the rate at level p. In the same way for the case of embedded MD coding, we can write for each of the side descriptions $R_m = \sum_{p=0}^{P} R_m^p$, where R_m^p represents the side rate at level p. This will result in the expression of the redundancy per quantization level: $\rho_p = \sum_{p=0}^{P}(R_1^p + R_2^p - R_0^p)$. From this, the overall redundancy between the descriptions can be interpreted as the sum of the redundancies at each distinct quantization level: $\rho = \sum_{p=0}^{P} \rho^p$. In our case, $R_m = \log_2(\prod_{i=p}^{P} L_i)$ and $R_0 = \log_2(\prod_{i=p}^{P} N_i)$. This leads to $\rho_p = 2\log_2(\prod_{i=p}^{P} L_p) - \log_2(\prod_{i=p}^{P} N_p)$ which can be written as

$$\rho_p = \sum_{p=0}^{P} \frac{(L_p)^2}{N_p}. \tag{7}$$

This simple formulation shows that for any EMDSQ instantiation, the redundancy is directly dependent on the level p. In addition, the redundancy can be controlled at each distinct quantization level by changing the ratio N_p/L_p (via the N_p and L_p parameters).

It is noticeable that the redundancy is rate-dependent and ranging in $[0, R_0]$. In order to have a rate-independent

analytical expression of redundancy (which is more appropriate for measurement purposes), we rely on the normalized redundancy, lying in the range $[0, 1]$:

$$\rho_P = \frac{R_1^p + R_2^p}{R_0^p} - 1. \tag{8}$$

In progressive coding relying on embedded quantization, the information within the stream is prioritized according to its impact on the overall rate-distortion performance. Hence, in order to improve the protection provided by an MD coding system, the redundancy in the layers corresponding to the coarser quantization levels should be higher than the redundancy corresponding to the finer levels. In other words, the redundancy has to increase with level p as

$$\rho_0 \leq \rho_1 \leq \cdots \leq \rho_P. \tag{9}$$

3. MULTIPLE-DESCRIPTION IMAGE CODING

3.1. *Multiple-description quadtree image codec*

In this section, we present a still-image MD coding system, enabling progressive transmission over unreliable channels. The proposed MD quadtree (MD-QT) coding approach is a wavelet-based system derived from our single-description square-partitioning (SQP) codec proposed in [15, 16]. SQP employs successive approximation quantization (SAQ) of the wavelet coefficients, followed by quadtree coding of the significance maps and adaptive arithmetic entropy coding. Other quadtree-based embedded image coding techniques have been proposed in the past, including for instance the nested quadtree splitting (NQS) algorithm of [20], the wavelet-based quadtree (WQT) codec of [15], and the set-partitioned embedded block coding (SPECK) approach of [21]. A later improvement of SQP is the QT-L codec of [22, 23], providing comparative compression performance against state-of-the-art codecs, such as JPEG2000 [1] or SPIHT [24]. The choice of the underlying single-description (SD) coding system is thus justified by the proven fact [22, 23] that intraband coding based on quadtree coding of the significance maps provides competitive compression performance against the state of the art.

In the proposed MD-QT approach, the classical SAQ is replaced by EMDSQ. This allows for producing more than one description of the input data. Additionally, EMDSQ retains the capability to provide fine-grain rate adaptation and progressive transmission of each description. Finally, the employed EMDSQ provides embedded deadzone quantization not only for the central quantizer, but for the side quantizers as well. Due to this feature, the proposed codec provides competitive rate-distortion characteristics for both central and side decoders.

At each quantization level p, $0 \leq p \leq P$, the proposed MD-QT coding algorithm performs the same coding passes as our SQP codec of [15]: the *significance pass*, in which the significant wavelet coefficients w, $T^p \leq |w| < T^{p+1}$, exceeding a certain significance threshold of the current level T^p

are localized and quantized, and the *refinement pass*, during which the quantization accuracy of the already significant coefficients (i.e., w: $T^{p+1} \leq |w|$) is refined. With the exception of the full-redundancy case, a different side quantizer is employed for each side description. Hence, it is natural to allocate for each distinct description a different set of significance thresholds, denoted as T_m^p, with m representing the description index ($m = 1, 2$ in the following). Once the positions and the corresponding symbols of the significant coefficients are encoded, p is set to $P - 1$ and the significance pass is restarted to identify the new significant coefficients. At every quantization level p, $p < P$, only the significance of the previously nonsignificant coefficients is encoded, and the corresponding quantization step is applied.

In order to detail the algorithm, we begin by defining the QT partition rule used in the significance pass and introduce the related notations. Consider the wavelet-transformed image $\mathbf{W}$ as a matrix of dimension $V_1 \times V_2$ and denote by $w(\mathbf{l})$ a wavelet coefficient contained in $\mathbf{W}$ at coordinates $\mathbf{l} = [\, l_1 \; l_2 \,]$, with $0 \leq l_1 < V_1$ and $0 \leq l_2 < V_2$. We denote by $\mathrm{QT}(\mathbf{k}, \mathbf{v})$ a quadrant with top-left coordinates $\mathbf{k} = [\, k_1 \; k_2 \,]$ and size $\mathbf{v} = [\, v_1 \; v_2 \,]$, where v_1 and v_2 represent the quadrant's width and height, respectively. In view of simplification, we assume identical power-of-two quadrant dimensions v_1 and v_2, that is, $v_1 = v_2 = 2^r$ for some $r \in \mathbb{N}$. Hence, the quadrant $\mathrm{QT}(\mathbf{k}, \mathbf{v})$ can be considered as a square matrix containing the wavelet coefficients $w(\mathbf{l})$ and it is defined as follows:

$$\mathrm{QT}(\mathbf{k}, \mathbf{v}) = \big[w(\mathbf{l}) \big]_{k_i \leq l_i < k_i + v_i} \quad \text{with } i = 1, 2. \tag{10}$$

According to the above notations, the wavelet-transformed image can be considered as a quadrant denoted as $\mathbf{W} = \mathrm{QT}(\mathbf{0}, \mathbf{V})$, where $\mathbf{0} = [\, 0 \; 0 \,]$, and $\mathbf{V} = [\, V_1 \; V_2 \,]$ denotes the image size.

The significance of any quadrant $\mathrm{QT}(\mathbf{k}, \mathbf{v}) \in \mathbf{W}$, $\mathbf{v} \neq [\, 1 \; 1 \,]$, with respect to the applied threshold T_m^p is determined via the significance operator defined as follows:

$$\sigma^p \big(\mathrm{QT}(\mathbf{k}, \mathbf{v}) \big) \big|_{\mathbf{v} \neq [1 \; 1]}$$
$$= \begin{cases} 1 & \text{if } \exists w(\mathbf{l}) \in \mathrm{QT}(\mathbf{k}, \mathbf{v}) : \big| w(\mathbf{l}) \big| \geq T_m^p, \\ 0 & \text{if } \forall w(\mathbf{l}) \in \mathrm{QT}(\mathbf{k}, \mathbf{v}) : \big| w(\mathbf{l}) \big| < T_m^p. \end{cases} \tag{11}$$

The binary matrix $\mathrm{QT}_m^p(\mathbf{k}, \mathbf{v})$, which indicates the significance $\sigma^p(w(\mathbf{l}))$ of each coefficient $w(\mathbf{l}) \in \mathrm{QT}(\mathbf{k}, \mathbf{v})$ with respect to the applied threshold T_m^p, is defined as

$$\mathrm{QT}_m^p(\mathbf{k}, \mathbf{v}) = \big[\sigma^p \big(w(\mathbf{l}) \big) \big]_{k_i \leq l_i < k_i + v_i} \quad \text{with } i = 1, 2. \tag{12}$$

Finally, we define a partitioning rule that divides a significant quadrant and the relative matrix $\mathrm{QT}_m^p(\mathbf{k}, \mathbf{v})$ into four adjacent minors, as follows:

$$\mathrm{QT}_m^p(\mathbf{k}, \mathbf{v}) = \left[\mathrm{QT}_m^p \left(\mathbf{k} + \frac{\mathbf{v}}{2}\boldsymbol{\alpha}, \frac{\mathbf{v}}{2} \right) \right], \quad \text{if } \sigma^p \big(\mathrm{QT}(\mathbf{k}, \mathbf{v}) \big) = 1. \tag{13}$$

where $\boldsymbol{\alpha} = \big[\begin{smallmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{smallmatrix} \big]$ with $\alpha_i = 0, 1$ for $i = 1, 2$ and $\mathbf{k} + (\mathbf{v}/2)\boldsymbol{\alpha} = [\, k_1 + \alpha_1(v_1/2) \;\; k_2 + \alpha_2(v_2/2) \,]$.

Next we describe, for a number of two description, the *significance and refinement passes* performed by the encoder. Starting with the coarsest quantization level $p = P$, the significance pass is activated first and the significance of the wavelet image $\mathbf{W}$ is determined with respect to the threshold T_m^P as in (11). If $\sigma^P(\mathbf{W}) = 1$, a significance symbol is emitted and the significance map $QT_m^p(\mathbf{0}, \mathbf{V})$ is split into four quadrants $QT_m^p(\mathbf{V}\boldsymbol{\alpha}/2, \mathbf{V}/2)$ according to the partitioning rule (13). Then, following a depth-first technique [25], the descendent quadrants are further tested for significance and only the significant ones are iteratively spliced as in (13). The recursive process ends when all the 4×4 leaf nodes, containing at least one significant coefficient, are isolated and quantized, applying the proper EMDSQ side quantizer and the output symbols are added to the stream.

Conversely ($\sigma^P(\mathbf{W}) = 0$), all coefficients belonging to nonsignificant quadrants need not be explicitly quantized and a single nonsignificance symbol suffices to map all elements in a quadrant to the side quantizer deadzone.

Thus, the significance pass (i) records the positions $\mathbf{l}$ of all the leaf nodes newly identified as significant, using a recursive tree-structure of quadrants, and (ii) quantizes the values of the coefficients contained in the significant leaf nodes. The tree structure of matrices, produced by the partitioning rule, can be represented by the corresponding tree structure of significant or nonsignificant symbols. The employed depth-first scanning procedure allows the encoder to map the tree structure to a one-dimensional stream of symbols. Similarly, by inverting such mapping, the decoder reconstructs, at each quantization level, the significance matrices from the received stream.

It is noticeable that for an individual wavelet coefficient, we no longer apply the significant operator σ^p, and instead we use the quantizer-index allocation operator, which we denote by $\delta(w(\mathbf{l}))$, determining in which partition cell the wavelet coefficient is contained. To do so, the wavelet coefficients have to be compared with the partition-cell boundary points. Consider that an arbitrary partition cell at level p is divided into a maximum number L_{p-1} of partition cells at level $p - 1$. In the index allocation process, the operator δ determines the symbol associated to each quantized coefficient as follows:

$$\delta(w(\mathbf{l})) = \begin{cases} S_{L_{p-1}}, & w(\mathbf{l}) \in C_{L_{p-1}}^{p-1}, \\ \vdots \\ S_2, & w(\mathbf{l}) \in C_2^{p-1}, \\ S_1, & w(\mathbf{l}) \in C_1^{p-1}, \end{cases} \tag{14}$$

where C_n^{p-1} with $1 \leq n \leq L_{p-1}$ denote the partition cells of Q_m^{p-1} in which a given partition cell C_k^p of Q_m^p is divided. All the coefficients contained in significant leaf nodes are stored in a *significance list*; we denote by $w^p(\mathbf{l})$ the coefficients contained in the *significance list* at level p.

Subsequently, the significance pass is followed by the refinement pass. This pass is activated at every level p, $p < P$, in order to refine the quantization accuracy for the coefficients recorded in the *significance list*, which were already coded as significant at a previous quantization level q, $p < q \leq P$. The coefficients stored in the *significance list* are refined by applying the quantizer index allocation operator δ that corresponds to the current coding pass. Before applying the operator δ, all the coefficients contained in the *significance list* must be rescaled according to a refinement threshold $T_r = T_m^p$. The rescaling is performed by subtracting the value of the significance threshold from the coefficient magnitude till the resulting value is smaller than the refinement threshold. This can be expressed analytically as $w^p(\mathbf{l}) = w^{p+1}(\mathbf{l}) - n \cdot T_m^p$, where $w^p(\mathbf{l})$ is the rescaled value obtained from $w^{p+1}(\mathbf{l})$, and n is the minimum integer for which $w^{p+1}(\mathbf{l}) - n \cdot T_m^p \leq T_m^p$.

Last, p is decremented and the described procedure is repeated, until the finest quantization accuracy is achieved, that is, $p = 0$, or the target bitrate is met.

Finally, we illustrate the manner in which the significance thresholds are computed, in the particular case of an EMDSQ instantiation [10] with $M = 2$, as depicted in Figure 2. For the coarsest quantization accuracy, $p = P$, the starting thresholds corresponding to each channel are $T_1^P = 2T$ and $T_2^P = T$, respectively. Since it is not desirable that the quantizer is characterized by an overload region, the T value is related to the highest absolute magnitude $w_{\max}$ of the wavelet coefficients as

$$T = 3^{\lfloor \log_3(w_{\max}/3) \rfloor + 1}. \tag{15}$$

Hence, the maximum number of quantization levels is $P = \lfloor \log_3(w_{\max}/3) \rfloor + 1$. In general, for $p < P$, the significance thresholds used for each channel $m = 1, 2$ are given by

$$T_m^p = \frac{T_m^P}{3^{P-p}}. \tag{16}$$

3.2. Enhanced MD-QT image codec

This section introduces an improved version of the MD-QT approach presented in Section 3.1. To start with, it is important to observe that the MD-QT coder described above applies the MD coding paradigm only at the level of producing multiple quantized (and coded) representations of the wavelet coefficients. However, the output bitstreams are not composed solely of this type of information, but also of additional localization (or QT) information. Practically, the QT information is the data stored in the output stream generated by encoding the locations of the significant leaf nodes, as obtained by (11) for each applied threshold. Motivated by this observation, we extend in this section the MD paradigm to all types of information generated by the coder, thus providing an enhanced version of MD-QT coding. As demonstrated experimentally, in case of erasures, this will lead to a systematically better rate-distortion performance in comparison to the MD-QT coder presented in Section 3.1.

To justify this design approach, let us analyze the types of information generated by the proposed MD-QT algorithm. The principle behind MD-QT is to generate robust descriptions over erasure channels by producing more descriptions of the input data. The MD capability is attained by employing
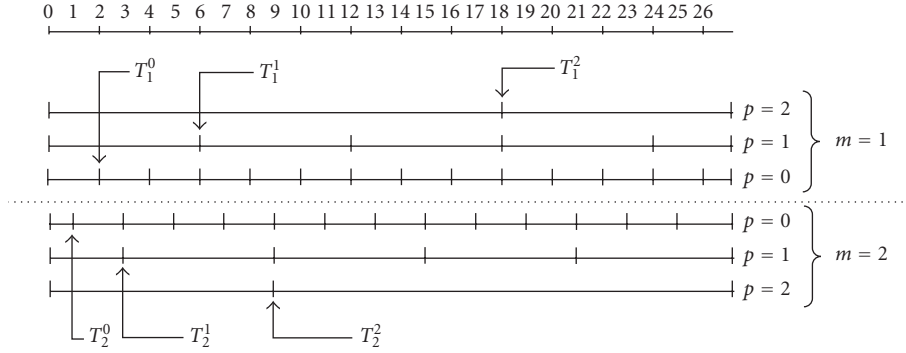
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26

FIGURE 2: Example of a three-level representation of Q_m^p for two-description EMDSQ ($M = 2, 0 \le p \le 2$) with granular region ranging from 0 to 26. The significance-map coding is performed with respect to the set of thresholds T_m^p, as defined by (16).

EMDSQ. The level of redundancy between the side quantizers' output can be adjusted via the *index allocation* (IA) of the EMDSQ [26] and is related to the channel probability of failure.

At every quantization level, each MD output stream is composed of the following encoded symbols: the quantized information, the sign information, and the QT information. The quantized and sign information are generated during the significance and refinement passes. For each quantization symbol resulting from the significance pass, we have a corresponding QT piece of information that serves to localize the position of the corresponding sample in the wavelet-transformed image.

The distribution of the symbols generated for a single coding step, corresponding to one quantization level p is illustrated in Figure 3. Additionally, Table 1 gives the relative percentage magnitude of each symbol type within the total number of symbols for each of the first five coding steps, corresponding to the quantization levels p, $P - 4 \le p \le P$.

The numbers are calculated by averaging, for each coding step, the number of symbols obtained by encoding a set of ten images. Notice that for the first coding step, there are no refinement symbols and that the number of sign symbols equals the number of quantization symbols. This experiment reveals that at least half of the amount of information contained in the output stream at each coding pass consists of QT symbols.

Now, let us consider that each of the descriptions is separately packetized and transmitted over the network. It is obvious that in the case of an erasure at the level of the quantization symbols (both the significance and refinement symbols), these symbols can be recovered with a certain fidelity from the received description at the decoder side. In the case of lost sign data for one description, the second description will provide complete recovery since the sign information is completely redundant in each description. However, the QT information needs to be protected against potential errors as well. In fact, if compared to the impact of errors occurring on quantization information, the errors that occur at the level of QT information will have a greater impact on the decoded
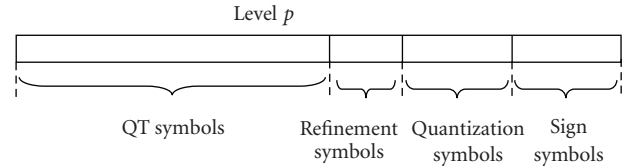
Level p

FIGURE 3: Distribution of the symbols generated in a coding pass of a MD-QT coding system.

TABLE 1: Distribution of the coding symbols in a bitstream generated in the first five coding passes of a wavelet-based MD-QT coding system.

Quantization level	P	$P - 1$	$P - 2$	$P - 3$	$P - 4$
Sign symbols (%)	13	23	21	19	16
Quantization symbols (%)	13	23	21	19	16
Refinement symbols (%)	0	5	8	10	12
QT symbols (%)	74	49	50	52	54

data. Indeed, if there is no mechanism that correlates the QT information among the different descriptions, then there is no way to enable QT data recovery from a received description.

From the above discussion, the following conclusions can be drawn.

(i) Despite the fact that multiple descriptions are provided by the system of Section 3.1, the MD principles are limited only to the quantization and sign information. The information represented by the QT symbols is nevertheless present in all descriptions, but is not correlated, thus leading to the impossibility of recovering such data in case of erasures.

(ii) The QT symbols stand for the largest part of information generated in each coding pass. Therefore, if the channel is characterized by a uniformly distributed probability of failure, there is a greater likelihood that the QT information will become erroneous.
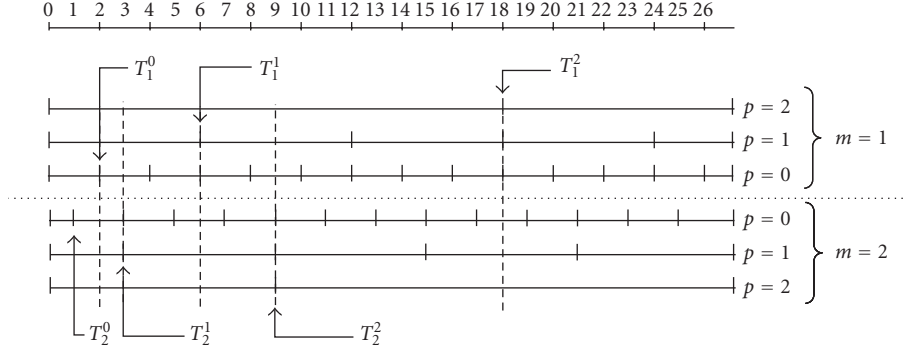
FIGURE 4: Three-level representation of Q_m^p for a two-channel EMDSQ ($M = P = 2$, $0 \leq p \leq 2$) for an example with granular region ranging from 0 to 26. The significance-map coding is performed for each of the two descriptions with respect to the set of thresholds $(T_1^P, T_2^P), \ldots, (T_1^0, T_2^0)$.

In the following, we present a new QT encoding approach that will lead to redundant QT information among the resulting descriptions. The starting point in this design is the coding scheme described in Section 3.1. Similar to this scheme, each coding step is composed of *significance* and *refinement passes*, with the exception of the first step, corresponding to the coarsest quantization level, in which only the significance pass is performed.

It is important to observe that in order to produce similar QT (i.e., localization) information for all descriptions, we need to apply a common set of significance thresholds when we construct the descriptions. A simple solution to obtain such a set is to combine the different sets of significance thresholds used to generate each distinct description in the MD-QT scheme of Section 3.1. Consider for example the EMDSQ instantiation depicted in Figure 2. In this case, the first set of thresholds is $T_1^P, T_1^{P-1}, \ldots, T_1^0$ and the second set of thresholds is $T_2^P, T_2^{P-1}, \ldots, T_2^0$, where T_m^p is given by (16). The common set of thresholds will be of the form $(T_1^P, T_2^P), (T_1^{P-1}, T_2^{P-1}), \ldots, (T_1^0, T_2^0)$. Figure 4 depicts the same EMDSQ instantiations as given in Figure 2. It is noticeable that in the case of enhanced MD-QT, two thresholds correspond to each quantization level. This results into two significance passes and one refinement pass for each coding step. In general, by following this approach of merging the sets of significance thresholds, the algorithm will perform M significance passes and one refinement pass for every coding stage p, $p < P$.

4. EXPERIMENTAL RESULTS

In this section, we present experimental results testing different aspects of our proposed approach. For all the experiments presented in this section, the MDC system is providing a number of two descriptions.

The first set of experiments focuses on (i) the redundancy control mechanism and (ii) the comparative performance assessment between instantiations of EMDSQ and the state-of–the-art MDUSQ of [14]. In order to demonstrate the redundancy control mechanism, three instantiations of EMDSQ at

different overall redundancy levels are employed in the MD-QT coding system. The MD-QT yielding $\rho = 0.9$ employs an EMDSQ instantiation that corresponds to a two-diagonal embedded IA [26, 27]. It is noticeable that the EMDSQ cells are not disconnected in this case. The MD-QT yielding a total redundancy of $\rho = 0.4$ employs an EMDSQ instantiation that corresponds to a two-diagonal embedded IA for all quantization levels except the finest one ($p = 0$) [8, 11]. For the finest quantization level, we allocate no redundancy in between the two descriptions. Following the notations from Section 2.2, this can be written as $N_0 = (L_0)^2$. It is noticeable that for the last quantization level, EMDSQ employs disconnected cells [8, 11]. Finally, for the last employed EMDSQ instantiation we allocate no redundancy at all for the final two quantization levels, that is, $N_0 = (L_0)^2$ and $N_1 = (L_1)^2$. This case will result in a total redundancy value of $\rho = 0.3$. [8, 11].

Figure 5 depicts the rate distortion behavior of the MD-QT central decoder employing the above-mentioned EMDSQ instantiations. The experiments demonstrate that both the overall redundancy and redundancy per each quantization level can be controlled. Additionally, (9) suggests that one can adapt the EMDSQ in order to provide an unequal error protection scheme where we can tune the redundancy according to the importance of the layer being encoded—for example, more redundancy can be allocated to the base layer (corresponding to the coarser quantization levels) and less to the enhancement layers, corresponding to the finer quantization levels.

Additionally, Figure 5 shows the rate-distortion results obtained with MD-QT incorporating MDUSQ at an overall redundancy $\rho = 0.9$. It is important to notice that for EMDSQ, it is possible to control the redundancy at each distinct quantization level, while the MDUSQ do not feature this important control mechanism. These results show also that at the same redundancy level, EMDSQ outperforms the state of the art on the whole range of rates.

Finally, Figure 5 depicts the rate-distortion results of the corresponding SD coder in order to allow for a comparison with the corresponding MD coder in the case of error-free channels. It is noticeable that in applications that require a
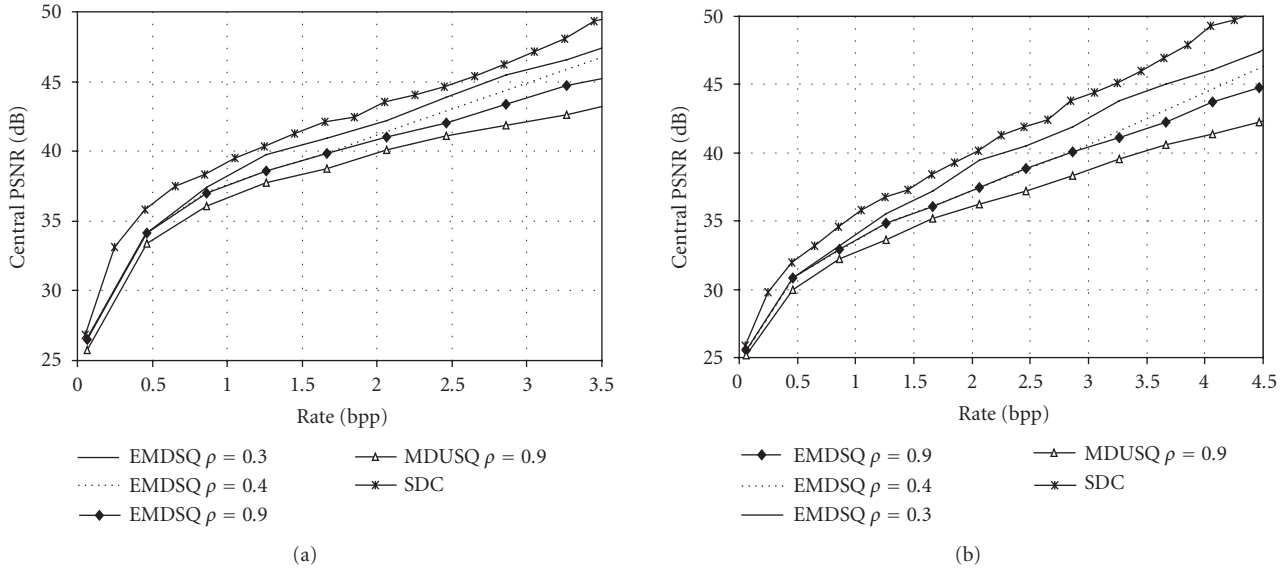
FIGURE 5: Comparative central rate-distortion performance for SDC and MD-QT incorporating the EMDSQ at different overall redundancies and MDUSQ of [7] applied on the (a) Lena and (b) Goldhill images.

lesser protection level (which will be reflected in less redundancy between the two descriptions), the rate-distortion gap between the SD and MD coders can be narrowed.

The goal of the second sets of experiments from this section is to evaluate the performance of the proposed MD-QT codecs operating in a realistic data communication scenario. Transmission systems are increasingly using packetization techniques. Therefore, we will asses how our error-resilient MD-QT systems are able to cope with packet losses in comparison to the equivalent single-description coding (SDC) system based on QT coding.

For this purpose, the output stream is divided into packets and is transmitted via the channel. We assume that the employed packetization method provides packets with a number of payload bytes representing the coded data plus sufficient header information. The header information will allow detecting the lost packets at the decoder side, and thus the correct sequencing of the remaining packets could be maintained. In the following experiments, we chose a packet payload of 640 bytes. For each probability of loss, all the possible erasure patterns are explored and the resulting PSNR values are averaged, as described next.

Let us consider a certain number of packets N that are to be transmitted, and let p_L be the average packet-loss probability. The average number of packets being lost is $k = p_L \cdot N$, and there will be $\binom{k}{N}$ combinations to lose k packets out of N. The average distortion is then calculated by measuring and averaging the MSE over all possible combinations.

Three sets of experiments are performed on two standard images, that is, 512×512 gray-scale *Lena*, and *Barbara* images, which have been compressed using (i) the simple and (ii) enhanced MD-QT codecs employing EMDSQ, and (iii) the equivalent single-description SQP coder, employing successive approximation quantization.

In a first round of experiments, we transmit $N = 14$ packets, corresponding to a coding rate of 0.27 bit per pixel (bpp). For this test, we use a channel model with a loss rate varying between 0% and 35%. Although 35% may seem high for current networks, such high loss rates commonly occur with wireless networks and on the Internet at peak times. The results given in Figure 6 show that in case of no error, the SDC system (i.e., SQP) provides better performance. On the other hand, in the presence of errors, even with a small probability of failure, the SDC system experiences a large drop in performance. This justifies the need for MD coding and demonstrates the robustness of the proposed approach over a broad range of packet-loss rates.

These results also demonstrate that generating common QT information for both descriptions by using a common significance threshold set comes at practically no cost in the error-free case, and significantly improves the performance in the error-prone case. Finally, it is important to point out that these results are systematically observed on a broad set of test images [26].

The results for the second experiment are presented in Figure 7 and demonstrate comparative progressive transmission capability (or quality scalability) in the context of communication over error-prone channels for the MD-QT versus SQP. The vertical axis represents the average PSNR obtained for all the possible loss patterns, while the horizontal axis represents the number of received packets.

For this second experiment, we send a total of 26 packets to represent the Lena image (corresponding to a coding rate of 0.51 bpp) and 40 packets for the Barbara image (corresponding to 0.78 bpp). In Figure 7, a probability of failure around 4% is considered.
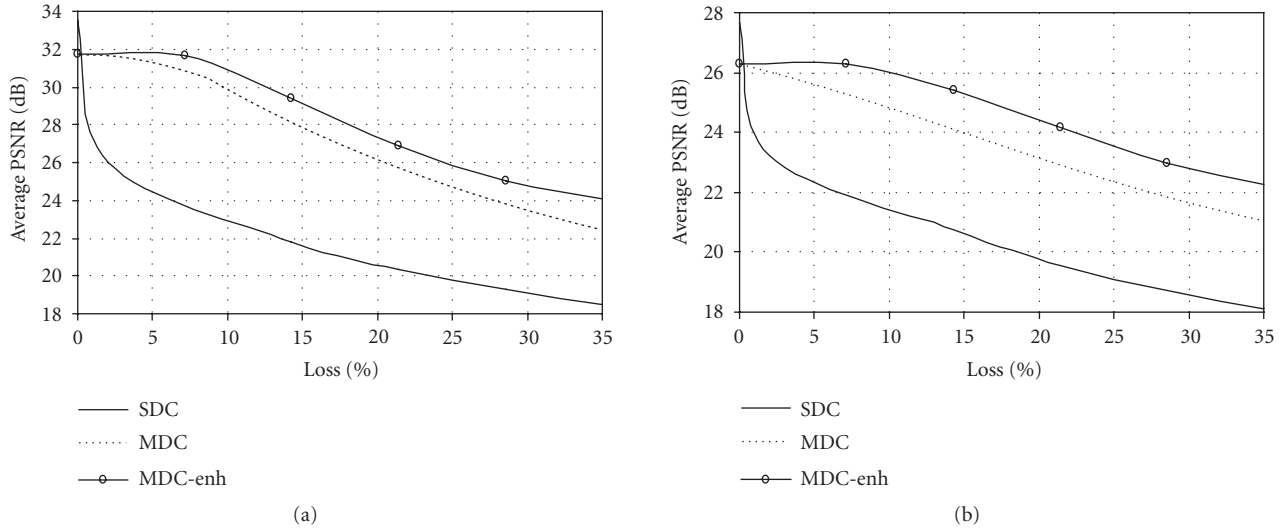
FIGURE 6: Effect of packet loss on average PSNR for all loss patterns for a number of 14 transmitted packets and probability of loss varying between 0% and 35% for (a) Lena and (b) Barbara. The employed coders are the simple MD-QT (MDC) and enhanced MD-QT (MDC-enh) and the corresponding SDC version, SQP.
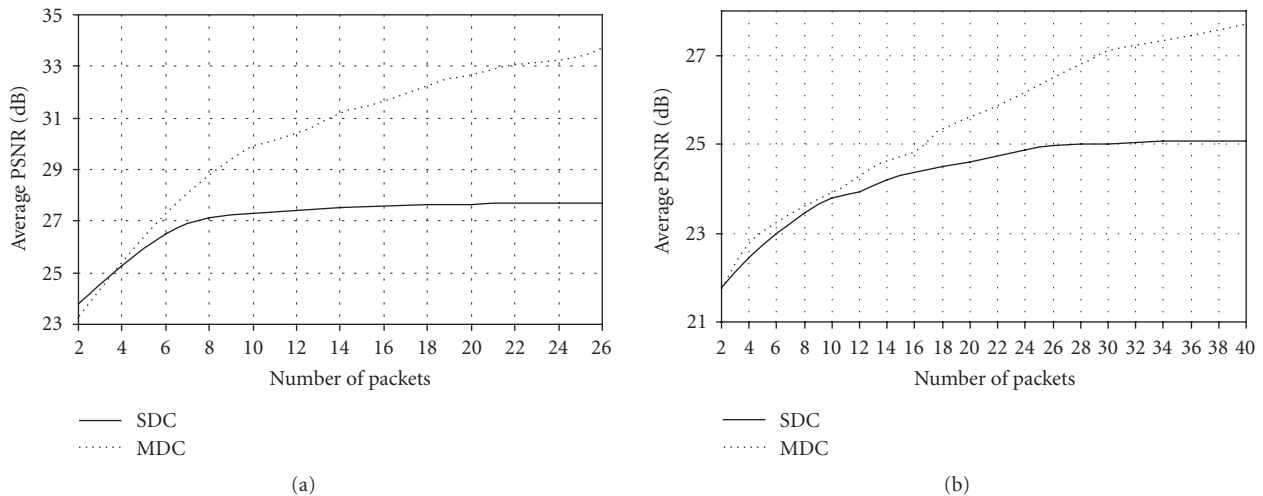

FIGURE 7: Effect of 4% packet loss on average PSNR for progressive transmission of (a) Lena and (b) Barbara. The employed coders are the MD-QT (MDC) and the corresponding SDC version, SQP.

Several conclusions can be drawn from this experiment. First, we can notice that an MD-QT coding system is a robust progressive transmission system since it allows for an increase in image quality with each packet received at the decoder side, in the context of erasures. Second, it is noticeable that even for a small error rate and a small number of received packets, the MD-QT coder outperforms the SQP coder by a large margin. Finally, it is noticeable that for MD-QT, the distortion is monotonically decreasing with the number of received packets, while the one of SQP is characterized by a *plateau effect*. This effect indicates that the average SDC performance can no longer be improved as the

probability of correctly receiving an increasing number of packets diminishes drastically.

5. DISCUSSION AND CONCLUSIONS

This paper presents a new type of scalable erasure-resilient image codecs. In the proposed approach, scalability and packet-erasure resilience are jointly provided via EMDSQ. A scalable multiple-description image coding system (MD-QT) is presented relying on quadtree coding of the EMDSQ output, along with an enhanced version of it. It is found experimentally that extending the MD paradigm by generating

common localization information across descriptions comes at practically no cost in the error-free case, and significantly improves the performance in the error-prone case.

The advantages of both MD coding systems are demonstrated in the context of image transmission over packet-lossy networks. The experimental results demonstrate that both the overall redundancy and redundancy per quantization level can be controlled. A comparative performance assessment between instantiations of EMDSQ and the state-of-the-art MDUSQ of [14], both incorporated in a common MD-QT coding system, is performed. The experimental results demonstrate that at the same redundancy level, EMDSQ outperforms the state of the art on the whole range of rates.

Finally, we notice that even for a small error rate and a small number of transmitted packets, the MD-QT codec outperforms the single-description coding equivalent (SQP) by a large margin. Moreover, for the MD-QT codec, the distortion is monotonically decreasing with the number of received packets, while the SQP codec is characterized by a plateau effect. This justifies the need for MD coding and demonstrates the robustness of the proposed approach over a broad range of packet-loss rates.

These results show that while transmitting over reliable links, the coding penalty associated with the proposed MD approaches versus single-description coding is controllable and can be reduced by reducing the overall redundancy. In other words, the "cost" of MDC can become negligible, while preserving significant benefits when transmitting over error-prone channels.

ACKNOWLEDGMENTS

REFERENCES

[1] D. Taubman and M. W. Marcelin, *JPEG2000: Image Compression Fundamentals, Standards, and Practice*, Kluwer Academic, Norwell, Mass, USA, 2002.

[2] L. Ozarow, "On a source-coding problem with two channels and three receivers," *The Bell System Technical Journal*, vol. 59, no. 10, pp. 1909–1921, 1980.

[3] V. A. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 821–834, 1993.

[4] Y. Wang, M. T. Orchard, V. Vaishampayan, and A. R. Reibman, "Multiple description coding using pairwise correlating transforms," *IEEE Transactions on Image Processing*, vol. 10, no. 3, pp. 351–366, 2001.

[5] V. K. Goyal and J. Kovačević, "Generalized multiple description coding with correlating transforms," *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2199–2224, 2001.

[6] R. Puri and K. Ramchandran, "Multiple description source coding using forward error correction codes," in *Proceedings of 33rd the Asilomar Conference on Signals, Systems, and Computers*, vol. 1, pp. 342–346, Pacific Grove, Calif, USA, October 1999.

[7] A. I. Gavrilescu, A. Munteanu, P. Schelkens, and J. Cornelis, "Embedded multiple description scalar quantizers," *IEE Electronics Letters*, vol. 39, no. 13, pp. 979–980, 2003.

[8] A. I. Gavrilescu, A. Munteanu, J. Cornelis, and P. Schelkens, "Generalization of embedded multiple description scalar quantizers," *IEE Electronics Letters*, vol. 41, no. 2, pp. 63–65, 2005.

[9] A. I. Gavrilescu, A. Munteanu, P. Schelkens, and J. Cornelis, "Embedded multiple description scalar quantizers for progressive image transmission," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '03)*, vol. 5, pp. 736–739, Hong Kong, April 2003.

[10] A. I. Gavrilescu, A. Munteanu, J. Cornelis, and P. Schelkens, "High-redundancy embedded multiple-description scalar quantizers for robust communication over unreliable channels," in *Proceedings of 5th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS '04)*, Lisboa, Portugal, April 2004.

[11] A. I. Gavrilescu, A. Munteanu, J. Cornelis, and P. Schelkens, "A new family of embedded multiple description scalar quantizers," in *Proceedings of IEEE International Conference on Image Processing (ICIP ' 04)*, vol. 1, pp. 159–162, Singapore, October 2004.

[12] V. A. Vaishampayan and J. Domaszewicz, "Design of entropy-constrained multiple-description scalar quantizers," *IEEE Transactions on Information Theory*, vol. 40, no. 1, pp. 245–250, 1994.

[13] V. A. Vaishampayan and J.-C. Batllo, "Asymptotic analysis of multiple description quantizers," *IEEE Transactions on Information Theory*, vol. 44, no. 1, pp. 278–284, 1998.

[14] T. Guionnet, C. Guillemot, and S. Pateux, "Embedded multiple description coding for progressive image transmission over unreliable channels," in *Proceedings of IEEE International Conference on Image Processing (ICIP '01)*, vol. 1, pp. 94–97, Thessaloniki, Greece, October 2001.

[15] A. Munteanu, J. Cornelis, G. Van der Auwera, and P. Cristea, "Wavelet-based lossless compression scheme with progressive transmission capability," *International Journal of Imaging Systems and Technology*, vol. 10, no. 1, pp. 76–85, 1999.

[16] A. Munteanu, J. Cornelis, G. Van der Auwera, and P. Cristea, "Wavelet image compression—the quadtree coding approach," *IEEE Transactions on Information Technology in Biomedicine*, vol. 3, no. 3, pp. 176–185, 1999.

[17] C. Dumitrescu and V. Seleacu, *Some Notions and Questions in Number Theory*, American Research Press, Rehoboth, NM, USA, 1998.

[18] N. Farvardin and J. W. Modestino, "Optimum quantizer performance for a class of non-Gaussian memoryless sources," *IEEE Transactions on Information Theory*, vol. 30, no. 3, pp. 485–497, 1984.

[19] H. Gish and J. Pierce, "Asymptotically efficient quantizing," *IEEE Transactions on Information Theory*, vol. 14, no. 5, pp. 676–683, 1968.

[20] C. K. Chui and R. Yi, "System and method for nested split coding of sparse data sets," US patent no. 5748116, Teralogic, Menlo Park, Calif, USA, 1998.

[21] A. Islam and W. A. Pearlman, "Embedded and efficient low-complexity hierarchical image coder," in *Visual Communications and Image Processing*, vol. 3653 of *Proceedings of SPIE*, pp. 294–305, San Jose, Calif, USA, January 1999.

[22] A. Munteanu, *Wavelet image coding and multiscale edge detection—algorithms and applications*, Ph.D. thesis, Department of Electronics and Information Processing (ETRO), Brussels, Belgium, 2003.

[23] P. Schelkens, A. Munteanu, J. Barbarien, M. Galca, X. Giro-Nieto, and J. Cornelis, "Wavelet coding of volumetric medical datasets," *IEEE Transactions on Medical Imaging*, vol. 22, no. 3, pp. 441–458, 2003.

[24] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 3, pp. 243–250, 1996.

[25] K. Parsaye and M. Chignell, *Expert Systems for Experts*, John Wiley & Sons, New York, NY, USA, 1988.

[26] A. I. Gavrilescu, *Multiple descriptions scalar quantizers—theory and applications*, Ph.D. thesis, Department of Electronics and Information Processing (ETRO), Brussels, Belgium, 2006.

[27] A. I. Gavrilescu, A. Munteanu, J. Cornelis, and P. Schelkens, "Generalization of embedded multiple description scalar quantizers yielding embedded uniform central quantizers," in *Proceedings of 6th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS '05)*, Montreux, Switzerland, April 2005.