THE EVOLUTION OF MULTIPLE DESCRIPTION CODING BASED ON SCALAR QUANTIZATION

Fabio Verdicchio, Adrian Munteanu, Jan Cornelis and Peter Schelkens*

Vrije Universiteit Brussel (VUB) – Interdisciplinary Institute for Broadband Technology (IBBT)
Department of Electronics and Informatics (ETRO)
Pleinlaan 2, B-1050, Brussels, Belgium

*E-mail: fverdicc@etro.vub.ac.be

ABSTRACT

Multiple Description Coding (MDC) allows for error resilient coding of data. This tutorial paper discusses one of the most popular MDC technologies, namely MDC based on scalar quantization, and its application in scalable coding of multimedia. This concept is illustrated by a fully scalable video coder, based on motion compensated temporal filtering, which incorporates the MDC principle for the encoding of both texture and motion vector data. For such system, we derive the principles of channel-aware MDC rate-allocation, which enables adjusting both the bit-rate and the resilience to errors of the output stream. Such adaptation is performed on-the-fly depending on the channel condition. Experimental results prove the advantages of MDC and the effectiveness of MDC system when transmitting video streams over unreliable networks.

1. INTRODUCTION

To improve bandwidth usage, multimedia systems typically incorporate compression to remove the redundancy from the input signal. However, in the case of communication over unreliable channels (e.g. mobile wireless or best-effort networks), achieving overall performance optimization is not always similar with minimizing the redundancy within the input stream. In fact, real-time transmission of multimedia over packet-based networks is highly sensitive to delay hence retransmission of lost packets is not as option. These factors badly reinforce the need for robust source coding. Multiple Description Coding (MDC) has 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. MDC 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 descriptions are available at the decoder, they are combined to enhance the quality of the decoded signal.

In this tutorial paper, we review one of the most popular MDC techniques, multiple description scalar quantizers (MDSQ). In section 2 we illustrate the principle of MDC and the basics of MDSQ. In section 3 we describe the extension of MDSQ towards embedded coding, resulting in embedded multiple description scalar quantizers (EMDSQ). The distinguishing features of EMDSQ are reviewed in section 3.1 and 3.2, in which the ability of EMDSQ to trim the redundancy in the transmitted stream is highlighted. In section 4 we apply EMDSQ in the context of scalable video coding obtaining a fully scalable system which grants the unique capability of streaming videos through unequally reliable links (e.g. wired/wireless) performing a single coding stage followed by a client-specific selection of the coded data. Finally conclusions are drawn in Section 5.

2. THEORETICAL BOUNDS AND PRACTICAL IMPLEMENTATIONS

The general Multiple Description Coding problem was initially posed to the scientific community at the 1979 Shannon Theory Workshop. In its original formulation, the MDC problem reads as follows: *"if an information source is described by two separate descriptions, what are the concurrent limitations and qualities of these descriptions taken separately and jointly?"* Over the following two decades, a number of researchers have addressed the above question from an information theoretic perspective and contributed to define the performance bounds for MDC. In the same time, triggered by the growing demand for reliable communication over error-prone channels, a number of practical implementations of the MDC concept have been proposed in the attempt to achieve, as closely as possible, such theoretical bounds. In the following we review the key theoretical findings of MDC and one of the most effective technologies enabling practical MDC of data.

2.1 Achievable performance for MDC

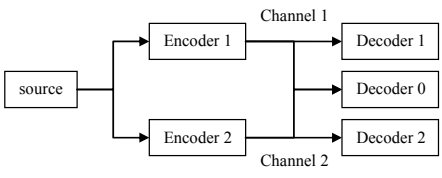

Figure 1. MDC problem for two channels and three receivers.

The classical MDC setup for two channels and three receivers is depicted in Figure 1. We denote by R_m ($m=1,2$) the transmission rate over channel m and by D_m ($m=0,1,2$) the distortion at the decoder side ($m=0$ indicates the *cen-*

tral decoder receiving both channels). The theoretical MDC problem consists in finding the achievable values for the quintuple $(R_1, R_2, D_0, D_1, D_2)$. The achievable rate-distortion region is completely known only for memoryless Gaussian sources with squared error distortions. In the following we report this result, obtained by Ozarow in his classic paper [1]. In light of simplification we assume balanced descriptions (i.e. $R_1 = R_2$ and $D_1 = D_2$) and unit variance of the input source, thus obtaining [1]:

$$D_1 \geq 2^{-2R_1}$$
$$D_0 \geq 2^{-4R_1} \cdot \left[1 - \left(\sqrt{\Pi} - \sqrt{\Delta}\right)^2\right]^{-1} \quad (1)$$

where $\Pi = (1 - D_1)^2$ and $\Delta = D_1^2 - 2^{-4R_1}$. The distortion region defined by (1) is depicted in Figure 2. Let us first consider the extreme case of individually good descriptions, i.e. $D_1 = 2^{-2R_1}$ (d_1 in the picture) therefore $\Delta = 0$. Thus we obtain $D_0 > D_1/2$ (d_4 in the figure). Conversely, when central decoding attains optimality, i.e. $D_0 = 2^{-4R_1}$ (d_3 in the picture), by imposing $\Delta = \Pi$ we obtain $D_1 = (1 + D_0)/2$ (d_3 in the figure). The transition between those regions can be sketched assuming high rates hence obtaining $D_0 D_1 \geq 4^{-1} 2^{-4R_1}$.

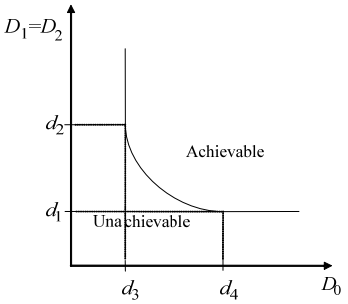

Figure 2. Achievable central and side distortions for balanced ($R_1 = R_2$ and $D_1 = D_2$) MDC of memoryless Gaussian sources with squared-error distortion.

Despite its simplicity, the above performance analysis highlights the peculiarity of MDC, in which the rate is spent to decrease distortion in central decoding and, at the same time, to attain low side distortions. A parameter commonly used to quantify the MDC trade-off, is the redundancy ρ defined as

$$\rho = R_1 + R_2 - R(D_0) \quad (2)$$

where $R(D)$ denotes the rate-distortion function of the input source, hence $R(D_0)$ is the rate need by single description coding (SDC) to attain the same distortion D_0 as the MDC central decoder. Clearly, when $\rho = 0$, D_1 reaches its minimum hence, in case of one channel impairment, the decoded quality will be poor. As ρ increases, a larger share of the rate is allocated to lower the side distortion, thus providing a better reconstruction in case of one channel failure. However, as the available total rate $R_1 + R_2$ cannot be exceeded, this means that D_0 must be lowered, thus underutilizing the available bandwidth if no error occurs. In lose terms ρ can be considered a measure of robustness to transmission errors.

2.2 Multiple description scalar quantization

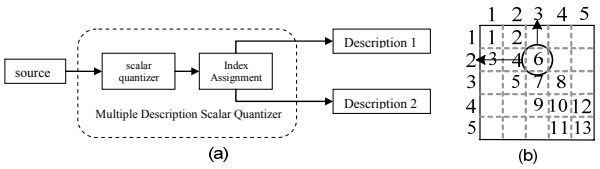

Figure 3. MDC system based on MDSQ.

Multiple description scalar quantizers (MDSQ) were introduced by Vaishampayan in [2]. MDSQ can be conceptually seen as a pair of independent scalar quantizers that output two descriptions of the same input sample. In a nutshell, MDSQ consist of two main components: a scalar quantizer and an index assignment (IA), as illustrated in Figure 3(a). Assuming that two descriptions are generated from an N level (central) quantizer, the IA is defined as a mapping function $\delta : N \to J^2$. Assuming that N is finite, the mapping δ can be thought of as a matrix of size $J \times J$, in which only N locations, corresponding to the central quantizer indices, are occupied. In this way, to each central quantizer index mapped in the IA matrix correspond a pair of numbers, the column and row indices. This pair represents the side quantizer indices, constituting the two mutually refinable descriptions of the input source. An IA instantiation is shown in Figure 3(b). We remark that the fundamental MDC trade-off is embodied in MDSQ by the design of the IA. On one extreme, an IA resulting in a "filled" matrix (i.e. $N = J^2$) generates individually poor descriptions, as the side quantization cells become disconnected, while it provides an efficient central decoder, as all the possible J^2 output symbols are used. The other extreme is an IA when only $N = J$ elements are filled (on the main diagonal) which results in two very accurate but identical side quantizers. The asymptotic analysis of MDSQ and a comparison with the theoretical bounds (1) was first performed in [3] while [4] analyzed the IA design as a combinatorial problem of arranging numbers in a matrix.

Classical MDSQ produce fixed-rate quantizers. This prevents the design of MDC schemes having the ability to adapt the output rate to varying channel capacity. This problem is addressed in Section 3.

3. EMBEDDED MULTIPLE DESCRIPTION SCALAR QUANTIZERS

In the attempt to extend MDSQ to variable-rate quantization, multiple description uniform scalar quantizers (MDUSQ) were presented [5], and later superseded by more generic embedded multiple description scalar quantizers (EMDSQ) which attain better coding performance for the same level of redundancy [6]. In contrast to MDUSQ, EMDSQ allow pro-

ducing an arbitrary number of descriptions. This feature is used to tune the overall redundancy. Moreover a new family of EMDSQ has been proposed [7] which grants the capability to control the redundancy at each distinct quantization level.

3.1 EMDSQ with *M* channels

We denote the set of embedded central quantizers as Q^p, with $p = 0,1,\ldots,P$ where $p=0$ and $p=P$ correspond to the finest and coarsest quantization levels respectively. Similarly, we denote the side quantizers as Q_m^p with $m \in \mathcal{M} = \{1,2,\ldots,M\}$. Partition cells of central and side quantizers at any level p are denoted respectively as C^p and S_m^p. In both cases, the partition cells of level $p < P$ are embedded in the ones at level $p+1$, i.e. $C^p \subseteq C^{p+1}$ and $S_m^p \subseteq S_m^{p+1}$. One instantiation of such quantizers is depicted in Figure 4 with $M=4$ and $P=1$.

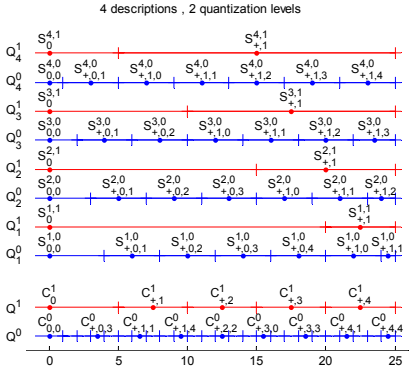

Figure 4. Example of EMDSQ with *M=4* and *P=1*.

Embedded quantization allows each side-decoder to progressively refine the reconstruction accuracy of a sample x. Upon reception of symbol x_p, the source sample, so far mapped to the cell $S^{m,p+1}_{\boldsymbol{x}_{p+1}}$ with $\boldsymbol{x}_{p+1} = x_P, x_{P-1}, \ldots, x_{p+1}$, is refined to the cell $S^{m,p}_{\boldsymbol{x}_{p+1},x_p}$. Assuming that all descriptions are received up to the same level p, $0 \le p \le P$, the central-quantizer cell is recovered as $C_x^p = \bigcap_{m\in\mathcal{M}} S^{m,p}_{\boldsymbol{x}_p}$ and its centroid is chosen to reconstruct the sample $\hat{x}$. The concept of embedded central and side-decoding has been extended in [8] taking into account the fact that, in general, the side descriptions may not be received at the same quantization level p. In fact, assuming the realistic scenario of best-effort packet networks, any description spanning more packets could be incompletely received as a result of packet losses. Embeddedness allows decoding partly received descriptions provided that causality is preserved. Thus in general we can determine the central cell associated to possibly-different quantization levels $\mathbf{p} = [p_1\ p_2 \ldots p_M]$ as $C_x^{\mathbf{p}} = \bigcap_{m\in\mathcal{M}} S^{m,p_m}_{\boldsymbol{x}_{p_m}}$. Remarkably this allows deriving a more generic formulation for the central quantizer as

$$Q^{\mathbf{p}} \equiv Q^{\mathbf{p}=[p_1\ p_2\cdots p_M]},\ p_m \in \{0,1,\ldots,P+1\}\ \forall m \in \mathcal{M}. \quad (3)$$

Assuming the convention that quantization level $P+1$ of any description m comprises a single cell $S^{m,P+1}$, which spans the entire granular region, the above notation represents as special cases the central $Q^{[p_1=p\ p_2=p\ldots p_M=p]} \equiv Q^p$ and side quantizers $Q^{[p_1=P+1\ldots p_m=p\ldots,p_M=P+1]} \equiv Q_m^p$.

3.1.1 Controlling the redundancy

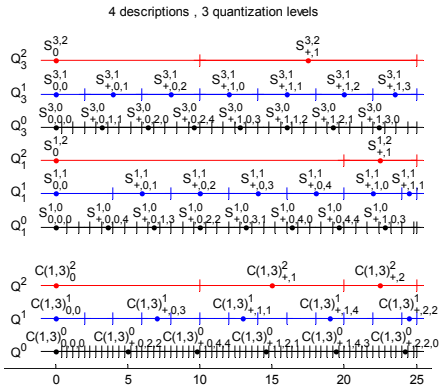

Figure 5. Example of EMDSQ with $\mathcal{M}'=\{1,3\}$ and *P=2*.

It was shown in [6] that by varying the number of generated descriptions M it is possible to change the redundancy $\rho = \sum_{m=1}^{M} R_m - R(D_0)$. Similarly, as shown in [8], streaming only a subset $\mathcal{M}' \subseteq \mathcal{M}$ of the generated descriptions is a mean of controlling the redundancy in the transmitted data *after* quantization has been performed. The subset of retained quantizers corresponds to $Q^{\mathbf{p}}$ as given by (3) with $p_m = P+1 \quad \forall m \notin \mathcal{M}'$ and attains redundancy $\rho' = \sum_{m\in\mathcal{M}'} R_m - R(D_0)$, as no information is transmitted for the discarded descriptions. Given the embedded nature of the quantizers, the accuracy of the reconstruction is not sacrificed as additional (finer) quantization levels of the retained descriptions can be can be transmitted. An example of the above concept is provided in Figure 5, which depicts the quantizers obtained retaining the subset $\mathcal{M}'=\{1,3\}$ out of the $M=4$ descriptions shown in Figure 4. We conclude that, once an appropriate vale of M is chosen based on worst-case estimates, it is possible, when more reliable links are involved, to reduce the redundancy in the coded stream without repeating the coding stage, thus trading error-robustness for compression efficiency. This concept is further discussed in Section 4

3.2 Generalized index allocation for EMDSQ

The EMDSQ described in Section 3.1 are quantizers with connected partitions cells, hence they generate individually

good, yet highly correlated, descriptions. In this section, we review the design of generalized embedded IA that allows tuning the redundancy between the descriptions for each quantization level. In the following we restrict our attention to the case of $M = 2$ descriptions.

The generalized IA matrix $\mathbf{I}$ is defined, at any quantization level p, $0 \le p \le P$, as a block matrix

$$\mathbf{I} = \left[\mathbf{A}^{p}_{j_1 j_2} \right]_{1 \le j_m \le J^p_m} \tag{4}$$

with each block $\mathbf{A}^{p}_{j_1 j_2}$ corresponding to the central quantization cell given by the union of the indices mapped into the block. Similarly, the union of the blocks $\mathbf{A}^{p}_{j_1 j_2}$ along each row (or column) corresponds to one cell of either side quantizer. Embeddedness implies recursive splicing of each block $\mathbf{A}^{p}_{j_1 j_2}$ at any level p, $0 < p \le P$, into $l^{p-1}_{j_1} \times l^{p-1}_{j_2}$ subblocks as follows:

$$\mathbf{A}^{p}_{j_1 j_2} = \left[\mathbf{A}^{p-1}_{i_1 i_2} \right]_{1 \le i_m \le l^{p-1}_{j_m}}, \quad 1 \le j_m \le J^{P}_{m}. \tag{5}$$

3.2.1 Controlling the redundancy

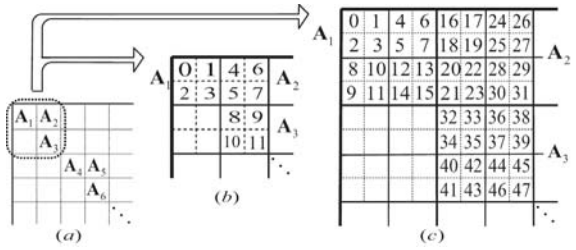

Figure 6. EMDSQ IA from [7]. The top quantization level (a) uses connected cells, while one (b) and two (c) finer quantization levels employ disconnected cells.

The rationale behind the IA proposed in [7] is to produce embedded side quantizers with connected cells for the coarser quantization levels, thus granting high resilience to the most important layers, and, for finer quantization levels (i.e. less important layers), to generate embedded side quantizers with disconnected partition cells, thus maximizing the usage of the available rate. Therefore, for the K coarsest quantization levels p, $P-K < p \le P$, $\mathbf{I}$ has the form (4) where the blocks $\mathbf{A}^{p}_{j_1 j_2} \ne [\mathbf{0}]$ are clustered around the main diagonal and consecutive indices are mapped into consecutive blocks. This results in connected cells as in Figure 6(a). Conversely, at any finer quantization level p, $0 \le p \le K$, more nonempty subblocks (5) are obtained and more indices are mapped into each subblock, thus generating disconnected cells for the side quantizers, as in Figure 6(b) and (c). The performance of such approach is assessed in the context of still-image coding and reported in Figure 7. In these experiments EMDSQ with connected cells at any level yields $\rho = 0.9$, while employing fully disconnected cells at the finest and the two finest levels results in $\rho = 0.4$ and $\rho = 0.3$ respectively. Results show how the IA of EMDSQ can adjust the redundancy level and favour central decoding quality, thus approaching the SDC performance (reported for comparison) when error-free channels are involved. Finally, we remark that, at the same redundancy level, EMDSQ outperform MDUSQ [5] on the whole range of rates.

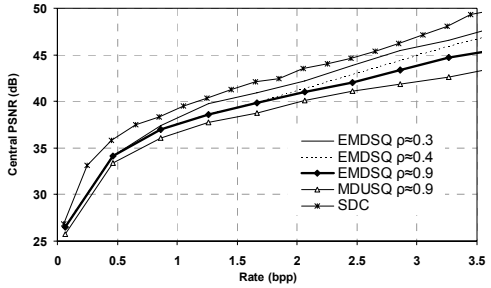

Figure 7. Comparative central rate-distortion performance for SDC and MDC applied to the *Lena* image. The MDC codec incorporates the EMDSQ at different redundancies and the MDUSQ of [5].

4. SCALABLE MDC OF VIDEO

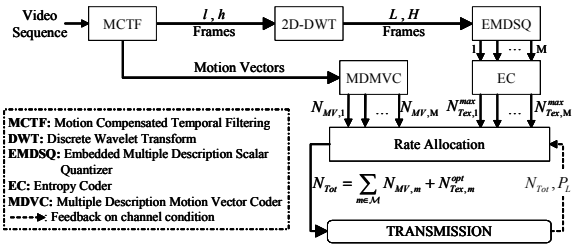

Figure 8. MDC open-loop SVC architecture [8] based on EMDSQ.

In multicast streaming the provider needs to transmit different coded version of the same content to classes of users connected via heterogeneous links. This section describes a fully scalable video codec which can adapt the bandwidth and the error-robustness of the coded output via a single coding stage. The flow diagram of the scalable erasure-resilient video coding architecture [8] is given in Figure 8. Temporal-and-spatial decomposition is performed as in classical open-loop motion compensated video coding [9], however, the resulting *L-*, *H-frames* and *MV*s are quantized using the EMDSQ of Section 3.1, thus obtaining M quality-scalable descriptions. Upon completion of the coding process, an optimized rate-allocation scheme selects for transmission only the data belonging to a proper subset $\mathcal{M}'$ of the encoded descriptions thus matching *on-the-fly* the degree of error-resilience to the channel conditions.

4.1 Channel-aware MDC rate-allocation

For any X, being a frame or *MV* field, quantized using EMDSQ, we derive the distortion-rate function in a stochastic framework accounting for source quantization and network erasures. We first derive the deterministic distortion-rate function of X reconstructing $\hat{X}$ for any quantizer $Q^{\mathbf{p}}$ given by (3) and computing the associated distortion and rates. The resulting distortion-rate points are then interpolated to obtain a regular sampling of the M rates, since it is assumed that transmission occurs by means of packets of

given size τ [8]. The ensuing discrete function is denoted as $D_\tau(X,\mathbf{N})$. Next we consider the effect of packet erasures, which result in (partial or complete) loss of one or more descriptions. Depending on the loss distribution a different approximation $\hat{X}$ is reconstructed at the decoder. Finally statistical average yields the *expected distortion* at decoding site $\bar{D}_\tau(X,\mathbf{N})$ [8].

The purpose of our rate-allocation is to determine the distribution $\mathbf{N}(X)=\left[N_1(X)\,N_2(X)\ldots N_M(X)\right]$ which, for any X, optimally distributes the available packet-budget attaining the lowest expected distortion in the decoded sequence $\hat{S}_{\mathbf{N}}$, given by [8]

$$\bar{D}_\tau(S_{\mathbf{N}})=\frac{1}{2^J G}\sum_{g=0}^{G-1}\sum_{j=1}^{J}\sum_{i=0}^{2^{J-j}-1}\sum_{X\in\mathcal{X}_{g,i}^{J,j}}\bar{D}_\tau\left(X,\mathbf{N}(X)\right) \quad (6)$$

where $\mathcal{X}_{g,i}^{J,j}=\left\{H^j_{g\cdot 2^{J-j}+i},MV^j_{g\cdot 2^{J-j}+i}\right\}$ with $1\le j<J$ and $\mathcal{X}_{g,i}^{J,J}=\left\{L^J_g,H^J_g,MV^J_g\right\}$.

The key to channel adaptation is varying the amount of redundancy used to represent any X. This is achieved by considering packet distributions which allocate rate to subset $\mathcal{M}'$ of the coded descriptions, i.e.

$$\mathbf{N}^{\mathcal{M}'}(X)\ :\ N_m(X)=0;\ m\notin\mathcal{M}'. \quad (7)$$

In general the choice of $\mathcal{M}'$ differs for any X and depends on the contribution of X to the distortion at decoding site (6) as well as on the channel statistics. We remark that such rate-redundancy-allocation does not require repeating the compression stage. Moreover it is possible to restrict the possible choices (7) thus keeping the complexity of the MDC rate-allocation comparable to the single description case [8].

4.1.1 Performance evaluation

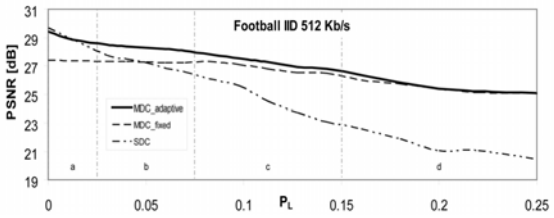

Figure 9. Sequence *Football* streamed assuming IID losses.

The efficiency of the rate-redundancy-allocation is assessed by comparing the performance of two instantiations of the same video codec. One instantiation (MDC_fixed) transmits all the encoded descriptions regardless of the channel conditions. The other instantiation (MDC_adaptive) uses the optimized MDC rate-allocation to select the appropriate descriptions to be sent to the decoder. In addition, we report the results obtained with SDC to highlight the price and the gain of MDC and show the advantages of the adaptive approach. In the experiments of Figure 9, the video sequence is streamed assuming IID packet losses occurring with probability P_L belonging to the intervals (a)-(d) in Figure 9. The MDC_adaptive codec performs the optimized rate-allocation only once for each interval (a)-(d), assuming that P_L equals the barycentre of the given interval. Results confirm the ability of the adaptive system to trim the redundancy within the streamed data thus obtaining, in the practically loss-free case (a), a quality close to SDC and, in (b)-(d), the best performance.

5. CONCLUSIONS

This tutorial paper comprises an overview of MDC techniques based on scalar quantization. MDSQ and its embedded extension EMDSQ are reviewed. The ability of EMDSQ to control the redundancy within the transmitted data is demonstrated in the context of image and video coding. Concerning the latter, a mechanism is described which allows minimizing the expected distortion at decoding site by adjusting the redundancy after compression has been performed.

ACKNOWLEDGMENTS

This work was supported in part by the Fund for Scientific Research of Flanders (FWO), Egmontstraat 5, B-1000 Brussels, Belgium (postdoctoral fellowships P. Schelkens and A. Munteanu).

REFERENCES

[1] L. Ozarow, "On a source coding problem with two channels and tree receivers," *Bell Syst. Tech. J.*, vol. 59, pp. 1909-1921, 1980.

[2] V. A. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Trans. Inform. Th.*, vol. 39, pp. 821-834, 1993.

[3] V. A. Vaishampayan and J.-C. Batllo, "Asymptotic analysis of multiple description quantizers," *IEEE Trans. Inform. Th.*, vol. 44, pp. 278-284, 1998.

[4] T. Y. Berger-Wolf and E. M. Reingold, "Index assignment for multichannel communication under failure," *IEEE Trans. Inform. Th.*, vol. 40, pp. 2656-2668, 2002.

[5] T. Guionnet, C. Guillemot, and S. Pateux, "Embedded multiple description coding for progressive image transmission over unreliable channels," Proc. ICIP, vol.1, pp. 94-97, 2001.

[6] A. Gavrilescu, A. Munteanu, P. Schelkens, and J. Cornelis, "Embedded multiple description scalar quantizers," *Electronics Letters*, vol. 39, pp. 979-980, 2003.

[7] A. Gavrilescu, A. Munteanu, J. Cornelis, and P. Schelkens, "A new family of embedded multiple description scalar quantizers," Proc. ICIP, vol.1, pp. 159-162, 2004.

[8] F. Verdicchio, A. Munteanu, A. Gavrilescu, J. Cornelis, and P. Schelkens, "Embedded Multiple Description Coding of Video," *IEEE Trans. Image Proc.*, vol. 15, pp. 3114-3130, 2006.

[9] J.-R. Ohm, "Advances in scalable video coding," *Proceedings of the IEEE*, vol. 93, pp. 42-56, 2005.