A COMPRESSED SENSING APPROACH TO FRAME-BASED MULTIPLE DESCRIPTION CODING

Teodora Petrisor[1,3], *Béatrice Pesquet-Popescu*[1] *and Jean-Christophe Pesquet*[2]

[(1)] ENST, Signal and Image Processing Department
46, rue Barrault, 75634 Paris Cédex 13, FRANCE
[(2)] IGM and UMR-CNRS 8049, Université de Marne-la-Vallée,
5, Bd Descartes, Champs sur Marne, 77454 Marne la Vallée Cédex 2, France.
[(3)] Galilee Institute, Université Paris 13 ,
99, Av. Jean-Baptiste Clement, 93430 Villetaneuse, FRANCE
Email: {petrisor,pesquet}@tsi.enst.fr,pesquet@univ-mlv.fr

ABSTRACT

In this paper, we consider a two description coding scheme based on a general frame synthesis operator. Through some approximations of the original rate-distortion problem, the design of the efficiently encoded coefficients is formulated as a convex optimization problem. We also show that there exists a close link between the proposed coding strategy and compressed sensing problems. Simulations results are provided to show the validity of our approach.

***Index Terms*—** Image coding, signal reconstruction, wavelet transforms, transform coding.

1. INTRODUCTION

Multiple Description Coding (MDC) is a useful technique for lossy networks, that exploits the existing path diversity in order to send several independently decodable streams instead of a single one. This coding strategy gives the possibility of recovering the source without involving retransmission (for instance in low-delay applications, network congestion etc.). The multiple streams, called *descriptions*, are obtained by splitting the source information after having introduced some redundancy. When channel losses occur or a whole path is not operational, the source is decoded only from the correctly received streams, with a reduced but acceptable quality of reconstruction. Among the practical directions followed to build MDC schemes which are thoroughly listed in [1], we will be mainly interested in a frame-based one [2], thus taking advantage of the inherent redundancy of the transform.

In the literature, there may be several issues of interest when designing MDC schemes, for instance efficient handling of redundancy in order to fulfill a rate-distortion tradeoff [3], [4], efficient forming of descriptions in order to get side-distortions as balanced as possible while being close to the central distortion [5], efficient reconstruction based on encoding constraints and so on. Here we present a complementary viewpoint inspired from the recent Compressed Sensing (CS) theory (see for instance [6], [7]) which gives remarkable results for analyzing signals having a sparse representation in some frame. These results basically state that such a signal can be perfectly recovered from a reduced number of arbitrary projections. This signal recovery problem can also be found under the name of sparse approximation and it has mainly been addressed by Matching Pursuit [8] or Basis Pursuit techniques [6]. Matching Pursuit has also been used in a first stage of generating multiple balanced descriptions (see [9] and references therein) for still images, in a lossy network scenario.

In our MDC framework we borrow from the CS field the idea that the encoder should determine a reduced number of components of the image in a frame representation from the observation of its pixel values. The choice of the components is grounded on a rate-distortion formulation of the MDC problem, which after some simplifications is re-expressed as a convex optimization problem. Then, we apply this approach to a two-description scheme for the transmission of still images.

The paper is organized as follows. In Section 2, we discuss the two complementary viewpoints for designing an MDC scheme from a frame representation. Then, in Section 3, we present the adopted approach relying on a synthesis frame paradigm and we formulate the considered rate-distortion problem. The proposed convex optimization algorithm is given in Section 4 and Section 5 provides a simple example and concludes the paper.

2. ANALYSIS *VS* SYNTHESIS FRAMES

We assume that the signal to be encoded belongs to a real Hilbert space $\mathcal{H}$ endowed with an inner product $\langle .,. \rangle$ and the associated norm $\|.\|$. We consider a two-description scheme based on two families of vectors $(e_{1,k})_{k\in\mathbb{K}_1}$ and $(e_{2,k})_{k\in\mathbb{K}_2}$ of $\mathcal{H}$, with $\mathbb{K}_1 \subseteq \mathbb{N}$ and $\mathbb{K}_2 \subseteq \mathbb{N}$. The union of these two families is assumed to form a frame of $\mathcal{H}$. Consider the associated decomposition operators: for all $i \in \{1, 2\}$,

$$\begin{aligned} L_i \;:\; &\mathcal{H} \to \ell^2(\mathbb{K}_i) \\ &x \mapsto (\langle x, e_{i,k}\rangle)_{k\in\mathbb{K}_i}. \end{aligned} \tag{1}$$

Their adjoint operators are

$$\begin{aligned} L_i^* \;:\; &\ell^2(\mathbb{K}_i) \to \mathcal{H} \\ &(\xi_k)_{k\in\mathbb{K}_i} \mapsto \sum_{k\in\mathbb{K}_i} \xi_k e_{i,k}. \end{aligned} \tag{2}$$

Two different viewpoints can be adopted to design an MDC scheme from these frame operators.

In the *analysis* frame paradigm, a signal $x \in \mathcal{H}$ is decomposed by the linear operators L_1 and L_2, as given before, so as to provide two descriptions

$$c_i = L_i x, \qquad i \in \{1, 2\}, \tag{3}$$

which are quantized and transmitted separately. At the decoder side, one of these descriptions or both may be available and the problem is to reconstruct a signal $\hat{x}$ as close as possible to x. In order to improve the quality of reconstruction, the convex constraints induced by the quantization rules can be addressed within a convex optimization approach [10, 11], which leads to a nonlinear reconstruction.

In the *synthesis* frame paradigm, the operators L_1^* and L_2^* are used at the decoder side. Two sequences of quantized values $\overline{c}_1$ and $\overline{c}_2$ corresponding to the two descriptions are transmitted. The reconstructed signal at the central decoder is then given by

$$\hat{x} = L_1^*\overline{c}_1 + L_2^*\overline{c}_2. \tag{4}$$

Similarly, the i-th side decoder computes

$$\hat{x} = \widetilde{L}_i^*\overline{c}_i, \qquad i \in \{1, 2\}. \tag{5}$$

where $\widetilde{L}_i^* \;: \ell^2(\mathbb{K}_i) \to \mathcal{H}$ is a given reconstruction operator. We see that one of the advantages of this approach is that the decoders take a simple linear form. At the encoder, the problem is however to generate the sequences $\overline{c}_1$ and $\overline{c}_2$ in the best way in a rate-distortion sense, taking into account the channel characteristics. The design of the encoding rule then yields a nonlinear optimization problem which is formulated in the next section.

3. SYNTHESIS FRAME APPROACH

3.1. Rate-distortion problem

Let $R(\overline{c}_i)$, $i \in \{1, 2\}$, denote the number of bits required to transmit the sequence of quantized values $\overline{c}_i$. We aim at minimizing the global bitrate

$$R_{\text{global}} = R(\overline{c}_1) + R(\overline{c}_2) \tag{6}$$

under a distortion constraint which is expressed as

$$\begin{aligned} D = {}& \alpha_{1,2}\|x - L_1^*\overline{c}_1 - L_2^*\overline{c}_2\|^2 \\ & + \alpha_1\|x - \widetilde{L}_1^*\overline{c}_1\|^2 + \alpha_2\|x - \widetilde{L}_2^*\overline{c}_2\|^2 \le D_{\max}. \end{aligned} \tag{7}$$

Hereabove, $D_{\max}$ is the maximum distortion allowed whereas $\alpha_{1,2}$, α_1 and α_2 are some positive weighting factors. Within a probabilistic setting, one can choose $\alpha_{1,2} = P_{\mathcal{D}_1+\mathcal{D}_2}$, $\alpha_i = P_{\mathcal{D}_i}$, $i \in \{1, 2\}$, where $P_{\mathcal{D}_1+\mathcal{D}_2}$ (resp. $P_{\mathcal{D}_i}$) is the probability that both descriptions are received (resp. only the i-th description is received). Other considerations e.g. perceptual quality can however be taken into account in the choice of the constants $\alpha_{1,2}$, α_1 and α_2.

The determination of $\overline{c}_1$ and $\overline{c}_2$ minimizing (6) subject to the constraint (7) is a difficult global nonconvex optimization problem. Note also that the upper bound $D_{\max}$ should be chosen large enough to guarantee the existence of a solution to the optimization problem. We will now bring some simplifications to this problem so as to be able to solve it using convex programming techniques.

3.2. Rate minimization

With little loss of generality for practical purposes, we will subsequently assume that a finite number of frame coefficients is considered, that is $\mathbb{K}_1 = \{1, \ldots, K_1\}$ and $\mathbb{K}_2 = \{1, \ldots, K_2\}$ (which implies that $\mathcal{H}$ is finite dimensional). In addition, for all $i \in \{1, 2\}$, $\overline{c}_i = (\overline{c}_{i,k})_{1\le k\le K_i}$ is assumed to be a vector of uniformly quantized values with a quantization step $q > 0$. The vector $\overline{c}_i$ can then be viewed as a realization of a random vector $\overline{C}_i = (\overline{C}_{i,k})_{1\le k\le K_i}$ taking its values in $\{\ldots - 2q, -q, 0, q, 2q, \ldots\}^{K_i}$. Thus, the entropy of $\overline{C}_{i,k}$ is given by

$$H(\overline{C}_{i,k}) = -\sum_{n\in\mathbb{Z}} P(\overline{C}_{i,k} = nq)\log_2\big(P(\overline{C}_{i,k} = nq)\big). \tag{8}$$

Instead of minimizing the global bitrate, we propose to minimize the entropy of the transmitted frame coefficients:

$$H_{\text{global}} = \sum_{k=1}^{K_1} H(\overline{C}_{1,k}) + \sum_{k=1}^{K_2} H(\overline{C}_{2,k}) \tag{9}$$

which is known to provide a lower bound of R_{global} for memoryless sources.

The coefficients $\overline{c}_{i,k}$ can be viewed as the outputs of a uniform quantizer driven with real-valued coefficients $c_{i,k}$, which are realizations of random variables $C_{i,k}$ with probability density functions $p_{i,k}$. For a fine enough quantization step, it has been proven [12] that the following relation holds between the discrete entropy of $\overline{C}_{i,k}$ and the differential entropy of $C_{i,k}$, denoted by $h(C_{i,k})$:

$$H(\overline{C}_{i,k}) \approx h(C_{i,k}) - \log_2(q). \tag{10}$$

Recall that the differential entropy is given by

$$h(C_{i,k}) = -\int p_{i,k}(\xi)\log_2\big(p_{i,k}(\xi)\big)d\xi = -\mathsf{E}\{\log_2\big(p_{i,k}(C_{i,k})\big)\}. \tag{11}$$

Let $\mathcal{S}_{i,k}$ denote the index set of coefficients in description i having the same probability distribution as $C_{i,k}$. For example, in a wavelet frame, this may correspond to a given subband. Under i.i.d. or more general classical mixing conditions, the expectation in (11) can be approximated by the sample estimate:

$$h(C_{i,k}) \approx -\frac{1}{N_{i,k}}\sum_{\ell\in\mathcal{S}_{i,k}}\log_2\big(p_{i,k}(c_{i,\ell})\big) \tag{12}$$

provided that $N_{i,k} = \mathrm{card}\mathcal{S}_{i,k}$ is large enough.

Let us now adopt a generalized Gaussian model for the probability distribution of the (centered) random variable $C_{i,k}$. This class of distributions was indeed shown to be quite flexible for modeling coefficients of sparse linear representations e.g. wavelet ones both for compression and denoising applications [13]. We have then

$$\forall\xi\in\mathbb{R},\qquad p_{i,k}(\xi) = \frac{\beta_{i,k}\omega_{i,k}^{1/\beta_{i,k}}}{2\Gamma(1/\beta_{i,k})}e^{-\omega_{i,k}|\xi|^{\beta_{i,k}}} \tag{13}$$

where $\omega_{i,k} > 0$, $\beta_{i,k} \geq 1$ and Γ is the gamma function. By injecting this expression in (12), we get

$$h(C_{i,k}) \approx \frac{\omega_{i,k}}{N_{i,k}\ln(2)}\sum_{\ell\in\mathcal{S}_{i,k}}|c_{i,\ell}|^{\beta_{i,k}} - \log_2\Big(\frac{\beta_{i,k}\omega_{i,k}^{1/\beta_{i,k}}}{2\Gamma(1/\beta_{i,k})}\Big). \tag{14}$$

From (9), we see that H_{global} is (up to a $\ln(2)$ dividing factor and an additive constant) approximately equal to

$$J(c_1,c_2) = \sum_{k=1}^{K_1}\omega_{1,k}|c_{1,k}|^{\beta_{1,k}} + \sum_{k=1}^{K_2}\omega_{2,k}|c_{2,k}|^{\beta_{2,k}}. \tag{15}$$

This suggests that J is an appropriate criterion to be minimized to control the bitrate.

3.3. Distortion bound

Let $\epsilon_i = (\epsilon_{i,k})_{1\leq k\leq K_i}$, $i\in\{1,2\}$, denote the vector of quantization errors defined by $\overline{c}_i = c_i + \epsilon_i$. The distortion constraint (7) can be rewritten as:

$$D = \alpha_{1,2}\|x - L_1^*(c_1+\epsilon_1) - L_2^*(c_2+\epsilon_2)\|^2 + \alpha_1\|x - \widetilde{L}_1^*(c_1+\epsilon_1)\|^2 + \alpha_2\|x-\widetilde{L}_2^*(c_2+\epsilon_2)\|^2 \leq D_{\max}. \tag{16}$$

Let us now focus on the quadratic term corresponding to the distortion at the side decoder i. We have

$$\begin{aligned}&\|x-\widetilde{L}_i^*(c_i+\epsilon_i)\|^2\\ =&\|x-\widetilde{L}_i^*c_i\|^2 + 2\langle x-\widetilde{L}_i^*c_i, \widetilde{L}_i^*\epsilon_i\rangle + \|\widetilde{L}_i^*\epsilon_i\|^2\\ =&\|x-\widetilde{L}_i^*c_i\|^2 + 2\langle \widetilde{L}_i(x-\widetilde{L}_i^*c_i), \epsilon_i\rangle + \|\widetilde{L}_i^*\epsilon_i\|^2.\end{aligned} \tag{17}$$

We can consider that ϵ_i and $\zeta_i = \widetilde{L}_i(x-\widetilde{L}_i^*c_i) = (\zeta_{i,k})_{1\leq k\leq K_i}$ are realizations of random vectors $E_i = (E_{i,k})_{1\leq k\leq K_i}$ and $Z_i = (Z_{i,k})_{1\leq k\leq K_i}$. For a fine enough quantization step, the quantization errors $E_{i,k}$ can be assumed to be zero-mean, i.i.d. and independent of all the other random variables, in particular Z_i. We have then $\mathsf{E}\{E_{i,k}Z_{i,k}\} = 0$. The law of large numbers can be invoked to assert that

$$K_i^{-1}\sum_{k=1}^{K_i}E_{i,k}Z_{i,k} \xrightarrow{P} 0 \tag{18}$$

as $K_i\to\infty$. This means that, when K_i is large, the inner product term $\langle\zeta_i,\epsilon_i\rangle = \sum_{k=1}^{K_i}\zeta_{i,k}\epsilon_{i,k}$ can be neglected in (17). By exploiting the assumption of independence between E_1 and E_2, similar arguments can be used to approximate the first quadratic term in (16) corresponding to the distortion at the central decoder. Thus, we obtain $D \approx G(c_1,c_2) + D_\epsilon$, where

$$G(c_1,c_2) = \alpha_{1,2}\|x - L_1^*c_1 - L_2^*c_2\|^2 + \alpha_1\|x-\widetilde{L}_1^*c_1\|^2 + \alpha_2\|x-\widetilde{L}_2^*c_2\|^2 \tag{19}$$

$$D_\epsilon = \alpha_{1,2}\|L_1^*\epsilon_1 + L_2^*\epsilon_2\|^2 + \alpha_1\|\widetilde{L}_1^*\epsilon_1\|^2 + \alpha_2\|\widetilde{L}_2^*\epsilon_2\|^2. \tag{20}$$

The latter term can be evaluated from the second-order statistics of the quantization noise.

In summary, the original rate-distortion problem can be recast as: Find c_1 and c_2 minimizing $J(c_1,c_2)$ subject to the quadratic inequality constraint

$$G(c_1,c_2) \leq G_{\max} = D_{\max} - D_\epsilon. \tag{21}$$

At this point, the strong connection between the proposed approach and the compressed sensing framework can be more easily understood. Indeed, when for all k, $\beta_{1,k} = \beta_{2,k} = 1$ and $\omega_{1,k} = \omega_{2,k} = 1$, the optimization problem is quite similar to the one addressed in compressed sensing in the presence of noise [6].

4. CONVEX OPTIMIZATION

By using a Lagrangian formulation of the convex optimization problem stated at the end of the previous section, we come up to the following saddle point problem:

$$\max_{\mu\geq 0}\min_{c_1,c_2}\Big(J(c_1,c_2) + \mu(G(c_1,c_2) - G_{\max})\Big). \tag{22}$$

The most challenging part of the optimization process is the inner minimization. Generally, one must resort to iterative algorithms to solve this problem. We propose here to use the efficient algorithm described in [14], which is itself an extension of the methods in [15]. Starting from initial coefficients $(c_1^{(0)}, c_2^{(0)})$, the algorithm generates a sequence $(c_1^{(n)}, c_2^{(n)})_{n\geq 1}$

converging to a solution of the optimization problem. At iteration n, we set $c_i^{(n)} = (c_{i,k}^{(n)})_{1 \le k \le K_i}$, $i \in \{1, 2\}$, and we compute, for all $k \in \{1, \ldots, K_i\}$,

$$\pi_{i,k}^{(n)} = \mathrm{prox}_{\gamma\omega_{i,k}|\,.\,|^{\beta_{i,k}}}(c_{i,k}^{(n)} - \gamma g_{i,k}^{(n)}) \qquad (23)$$

$$c_{i,k}^{(n+1)} = c_{i,k}^{(n)} + \lambda(\pi_{i,k}^{(n)} - c_{i,k}^{(n)}) \qquad (24)$$

where $\gamma \in [0, \gamma_{\max}]$ is the algorithm step-size, $\lambda \in (0, 1]$ is a relaxation parameter, $\mathrm{prox}_{\gamma\omega_{i,k}|.|^{\beta_{i,k}}}$ is the proximity operator of the function $\gamma\omega_{i,k}|\,.\,|^{\beta_{i,k}}$ and

$$\begin{aligned}(g_{i,k}^{(n)})_{1 \le k \le K_i} = 2\big(&\alpha_{1,2} L_i(L_i^* c_i^{(n)} + L_{3-i}^* c_{3-i}^{(n)} - x) \\ &+ \alpha_i \widetilde{L}_i(\widetilde{L}_i^* c_i^{(n)} - x)\big). \end{aligned} \qquad (25)$$

We recall that the proximity operator of a convex function $f : \mathbb{R} \to \mathbb{R}$ is $\mathrm{prox}_f : u \mapsto \arg\min_v \frac{1}{2}(v - u)^2 + f(v)$. For the considered power functions, the proximity operator can be calculated explicitly for some values of the exponent $\beta_{i,k}$ [14], otherwise it can be easily computed numerically.

5. EXAMPLE AND CONCLUSION

We consider a simple example of a two-description scheme where, inspired from the JPEG2000 standard, L_1^* is the reconstruction from a 9-7 biorthogonal wavelet basis and L_2^* is also the reconstruction from the same wavelet basis functions but shifted by 1 pixel in each spatial direction. In this case, a natural choice for the side decoders is $\widetilde{L}_i^* = 2L_i^*$, $i \in \{1, 2\}$. A 3-resolution level dyadic filter bank structure is applied for the MDC encoding of the 512×512 standard Lena image. The weighting factors in the distortion constraint have been chosen here as $\alpha_{1,2} = 0.8$, $\alpha_1 = \alpha_2 = 0.1$. The frame coefficients are synthesized by the optimization approach described in the previous section where the parameters of the generalized Gaussian model have been estimated by an iterative Maximum Likelihood method. The quantization step q for each rate has been optimized and the JPEG2000 algorithm has been employed to encode the two quantized descriptions. Fig. 1 shows the evolution of the PSNR w.r.t. the global bitrate for the central and side decoders. We provide for comparison, the results corresponding to the direct application of the JPEG2000 encoder at half the bitrate i.e. at the same bitrate as a description since each one is a whole basis representation. As expected, the proposed scheme provides better results for the central decoder while showing a good performance for the two side decoders. It is worth noticing that better results could be expected by using more sophisticated frames.

6. REFERENCES

[1] V. K. Goyal, "Multiple Description Coding: Compression meets the network," *IEEE Signal Processing Magazine*, pp. 74–93, September 2001.

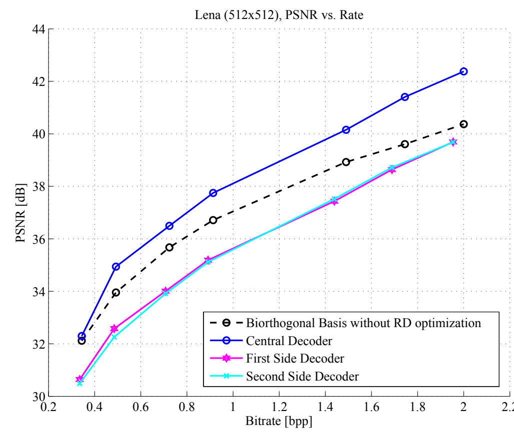

Fig. 1. Rate-distortion performance of the proposed scheme for Lena.

[2] V.K. Goyal, J. Kovačević, and J.A. Kelner, "Quantized frame expansions with erasures," *J. of Appl. and Comput. Harmonic Analysis*, vol. 10, pp. 203–233, May 2001.

[3] Y. Wang, M.T. Orchard, and A. Reibman, "Optimal pairwise correlating transforms for multiple description coding," in *Proc. IEEE Int. Conf. Image Proc.*, October 1998, vol. 1, pp. 679–683.

[4] P. L. Dragotti, S. D. Servetto, and M. Vetterli, "Optimal filter banks for multiple description coding: Analysis and Synthesis," *IEEE Transactions on Information Theory*, vol. 48(7), pp. 2036–2052, 2002.

[5] I. V. Bajic and J. W. Woods, "Domain-based Multiple Description Coding of images and video," *IEEE Transactions on Image Processing*, vol. 12(18), pp. 1211–1225, October 2003.

[6] D.L. Donoho, "Compressed Sensing," *IEEE Transactions on Information Theory*, vol. 52(4), pp. 1289–1306, 2006.

[7] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52(2), pp. 489 – 509, 2006.

[8] M. F. Duarte, M. A. Davenport, M. B. Wakin, and R.G. Baraniuk, "Sparse signal detection from incoherent projections," in *IEEE Int. Conf. on Acoustics, Speech and Signal Proc.*, 2006.

[9] I. Radulovic and P. Frossard, "Multiple description image coding with block-coherent redundant dictionaries," in *Picture Coding Symposium, China, 6pg.*, 2006.

[10] P.A. Chou, S. Mehrotra, and A. Wang, "Multiple Description decoding of overcomplete expansions using Projections onto Convex Sets," in *Data Compression Conference*, 1999, pp. 72–81.

[11] T. Petrisor, B. Pesquet-Popescu, and J.-C. Pesquet, "Wavelet-based multiple description coding of images with iterative convex optimization techniques," in *IEEE Int. Conf. on Image Processing*, 2005.

[12] H. Gish and J.N. Pierce, "Asymptotically efficient quantizing," *IEEE Transactions on Information Theory*, vol. IT-14(5), pp. 676–683, 1968.

[13] H. Masmoudi, A. Benazza-Benyahia, and J.-C. Pesquet, "Block-based adaptive lifting schemes for multiband image compression," *Wavelet Applications in Industrial Processing, SPIE Conf., Providence, Rhode Island, USA*, vol. 5266, pp. 118–128, 2003.

[14] C. Chaux, P. L. Combettes, J.-C. Pesquet, and V. R. Wajs, "Iterative image deconvolution using overcomplete representations," in *European Signal Processing Conference (EUSIPCO 2006)*, Sept 2006.

[15] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Comm. Pure Appl. Math.*, vol. 57, pp. 1413–1457, 2004.