

# SOFT DECODING OF MULTIPLE DESCRIPTIONS

*Thomas Guionnet, Christine Guillemot, Eric Fabre*

INRIA/IRISA

Campus de Beaulieu, 35042 Rennes Cedex, France

Thomas.Guionnet@irisa.fr

## ABSTRACT

A multiple description scalar quantization (MDSQ) based coding system can be regarded as a source coder (quantizer) followed by a channel coder, i.e. the combination of index and codeword assignment. The redundancy, or the correlation between the descriptions, is controlled by the number of diagonals covered by the index assignment. In this paper, we analyse dependencies between the variables involved in the MDSQ coding chain and design an estimation strategy making use of part of the global model of dependencies at each time. Inference of the hidden states of the connected models is done by applying belief propagation principles on the resulting Bayesian network. By analyzing the MDC system, we try to evidence the most appropriate form of redundancy one should introduce in the context of VLC compressed streams in order to fight against desynchronizations when impaired by channel noise.

## 1. INTRODUCTION

Multiple description coding has been introduced as a generalization of source coding subject to a fidelity criterion for communication systems that use diversity to overcome channel impairments. Several correlated coded representations of the signal are created and transmitted on different channels. The design goals are therefore to achieve the best average rate-distortion performance when all the channels work, subject to constraints on the average distortion when only a subset of channels is received correctly. Extensive work has been dedicated to finding optimal achievable rate-distortion regions [4] and to the design of practical compression systems approaching these theoretical bounds [6], [7], [1]. The design of MDSQ in [6] assumes fixed length codes and fixed codebook sizes. Significant improvements are achieved in [7] by using variable length codes (VLCs) instead of fixed length codes and by designing the quantizer under the constraint of a given entropy instead of a codebook size. However, VLCs are very sensitive to errors: their decoding can easily get "de-synchronized", with dramatic effects on the SNR of the reconstructed signal.

A coding system based on MDSQ can be seen as a source

coder (quantizer) followed by a channel coder (the combination of index assignment and codeword assignment). The amount of redundancy is controlled by the number of diagonals  $d$  covered by the index assignment (IA). The parameter  $d$  drives the trade-off between the central and side distortions obtained when respectively all or a subset of the descriptions are received correctly. The authors in [8] consider delay-constrained slow-fading channels and show that an MDSQ based system can outperform a system that uses channel coding for a given delay. The author in [5] considers the additional usage of a channel code in each channel and applies the turbo principle to the channel decoding, taking into account the index probability distributions on each channel.

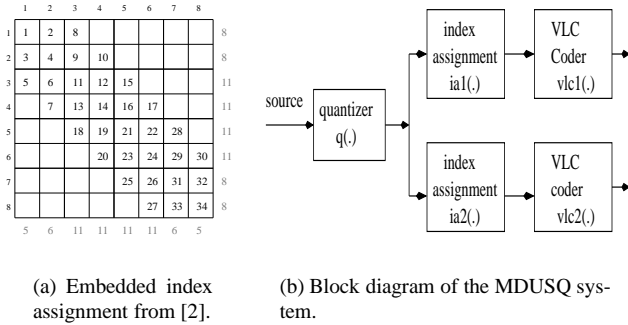
We consider here multiple description uniform quantization (MDUSQ) followed by VLCs and analyze dependencies between the different variables involved in the chain. Each element (quantized source of symbols, index assignment, VLCs) can be represented by a stochastic automaton. An accurate model of the dependencies introduced by these processes would require to consider the product of the different stochastic models, leading to a state space of unrealistic dimension. Instead of using the big product model, inference can be done making use of part of the global model at each time. Given that there is a pointwise translation of symbols into indexes and then into bits, the resulting Bayesian network follows a tree structure and is directly amenable to fast estimation algorithms. Optimal soft decoding of each description can thus be performed by using first the source coder model, assuming the sequences of indexes are white. The resulting soft information from the two descriptions are merged and used as input to a two sweep estimation that takes into account the inter-symbol correlation. The performance of the MDSQ soft decoding system is analyzed, for different  $E_b/N_0$  and for different source correlation factors.

## 2. MULTIPLE DESCRIPTION CODING

Let us consider the two-channel MDSQ-based communication system shown in Fig. 1(b) The system is composed of a quantizer  $q(\cdot)$  mapping the sequence  $\mathbf{S} = \mathbf{S}_1 \dots \mathbf{S}_K$  of

real-valued source samples into a sequence  $S = S_1 \dots S_K$  of integer-valued source symbols taking their values in a finite alphabet  $C$  composed of  $M_C$  symbols. In the experiments we will assume that  $S$  is obtained by quantizing a Gauss-Markov sequence with variance  $\sigma^2$  and correlation  $\rho$ . The sequence  $S = S_1 \dots S_K$  is assumed to be a Markov chain. Each source symbol is then mapped to two sets of components (index) that will be coded with VLCs and transmitted over the channels. These two set of indexes take their values in finite alphabets  $I$  and  $J$  composed respectively of  $M_I$  and  $M_J$  symbols.  $C$  is a subset of  $I \times J$ . These mappings are referred to as the index assignments  $ia^I(\cdot)$  and  $ia^J(\cdot)$  and are defined by a 2-dimension matrix as shown in Fig 1(a). The index assignment produces two correlated sequences of indexes  $I = I_1 \dots I_K$  and  $J = J_1 \dots J_K$  taking their values in a finite alphabet composed of  $M_I = M_J = 2^q$  symbols. The example of Fig 1(a) considers balanced descriptions with indexes taking their values in the same finite alphabet.

A MDSQ-based decoder comprises three inverse quantizers which map the received indexes  $(i, j)$  onto reconstruction levels  $\hat{x}_k$ ,  $\hat{x}_i$ , and  $\hat{x}_j$ , taking values respectively in the codebooks  $\hat{X}_C = \{\hat{x}_c, c \in C\}$ ,  $\hat{X}_I = \{\hat{x}_i, i \in I\}$  and  $\hat{X}_J = \{\hat{x}_j, j \in J\}$ . The performance of the whole system is very much dependent on the index assignment. The redundancy is controlled by choosing the number of diagonals covered by the index assignment. Let  $R$  be the number of bits used to encode each index, and  $2 \times d + 1$  the number of diagonals covered by the index assignment. The rate of an equivalent channel code is given by  $R_c = \frac{R_{eff}(d, R)}{2R}$ , where  $R_{eff}(d, R) = \log_2(N(d, R))$  with  $N(d, R)$  the number of cells covered by the index assignment.



**Fig. 1.** Index assignment ( $k = 2$ ) and MDUSQ.

### 3. MODELING DESCRIPTION DEPENDENCIES

Let the stationary probability density function of the source be denoted by  $p_S$ . Its probability matrix is denoted by  $\mathbb{P}_S = [P_0(n)]$ , where the components  $P_0(n)$ , probability that the quantized source symbol falls within the interval  $n$ , are defined by  $P_0(n) = \int_{d_{n-1}}^{d_n} p_S(x) dx$ . The sequence of quantized symbols  $S = S_1 \dots S_K$  is then mapped into two se-

quences of indexes which can also be regarded as two Markov chains. The index assignment establishes a one-to-one correspondence between the model of the sequence of source symbols and the model of the two sequences of indexes. The stationary probability matrices  $\mathbb{P}_S^I = [P_s^I(r)]$  and  $\mathbb{P}_S^J = [P_s^J(r)]$  of the two sequences of indexes can therefore be derived from the stationary probabilities of the sequence of source symbols by

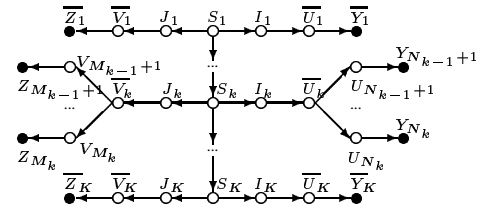
$$P_s^I(r) = \sum_{n:ia^I(n)=r} P_0(n), r = 1, 2, \dots, M_I, \quad (1)$$

where  $\{n:ia^I(n) = r\}$  denotes the set of quantizer cells  $n$  which are being mapped to the index  $r$  by the corresponding projection in the 2-D index assignment matrix (similarly for  $\mathbb{P}_S^J$ ). The term  $ia^I$  denotes the projection on the index set  $I$ .

The sequences of indexes are coded into two sequences of useful bits  $U = U_1 \dots U_N$ , and  $V = V_1 \dots V_M$ , by means of VLC. We denote by  $\bar{U}_k$  the codeword corresponding to  $I_k$ , and  $\bar{V}_k$  the codeword corresponding to  $J_k$ . So  $\bar{U} = \bar{U}_1 \dots \bar{U}_K$  and  $\bar{V} = \bar{V}_1 \dots \bar{V}_K$  represent the bitstreams  $U$  and  $V$  segmented into codewords. The entities  $K$ ,  $N$  and  $M$  are assumed to be known. The lengths  $N$  and  $M$  of the descriptions bit streams are random variables, function of  $S$ . The bitstreams  $U$  and  $V$  are sent over a memoryless channel and received as measurements  $Y$  and  $Z$ .  $Y$  and  $Z$  model the output of the two channels supposed to be memoryless, and  $Y_1^N = Y_1 \dots Y_N$  and  $Z_1^M = Z_1 \dots Z_M$  denote the particular values of  $Y$  and  $Z$  observed at the output of the channels. The general MAP decoding problem consists then in estimating  $S$ , given the observed values  $Y_1^N$ , and  $Z_1^M$ , i.e.

$$\hat{S}_1 \dots \hat{S}_K = \arg \max_{s_1 \dots s_K} \mathbb{P}(s_1 \dots s_K | Y_1^N, Z_1^M). \quad (2)$$

In the case of VLCs, the difficulty in the estimation comes from the lack of synchronization between the hidden states (symbols  $S_k$  to estimate) and the pointwise noisy measurements  $Y_1^N = Y_1 \dots Y_N$  and  $Z_1^M = Z_1 \dots Z_M$  on the sequences of useful bits. In other words, the segmentation of the two bitstreams  $U$  and  $V$  into codewords is random. It is shown



**Fig. 2.** Bayesian network for MS+MDUSQ+VLC.

in the following that the estimation can be performed as follows : the first step consists in estimating separately the bitstreams  $U$  and  $V$  using the intra-codeword redundancy and assuming that the symbols  $S_k$  are independent (hence that

indexes  $I_k$  are independent and similarly for  $J_k$ ; the soft information obtained on the bit clock source coder models must then be translated into soft information on the symbol clock model; the resulting soft information is then merged via the index assignment; then the symbol stream can in turn be estimated using the inter-symbol correlation.

The model of dependencies between symbols derives directly from the assumption that  $S$  is a Markov chain. To build a model of correlations within the codewords one can model the VLC encoder as a stochastic automaton, where the state variables  $X_n$  denote nodes on the codeword tree. However, in the case of VLCs, knowing the bit index  $n$  is not sufficient to determine the rank  $k$  of the symbol being constructed. This information is needed to translate soft information on bits into soft information on symbols. Therefore this information must be available jointly with the state variable  $X_n$ . For this we augment the state variable  $X_n$  with a counter  $K_n$  of the number of achieved symbols at time  $n$ . This amounts to defining a Markov chain distribution on pairs  $(X_n, K_n)$ . The transition probability from  $(X_n, K_n)$  to  $(X_{n+1}, K_{n+1})$  is thus determined by  $\mathbb{P}(\bar{U}_{K_{n+1}}|\bar{U}_{K_n})$  for the  $X$  component, and for the  $K$  part, one has  $K_{n+1} = K_n + 1$  each time  $X_{n+1}$  reaches a new leafnode, otherwise  $K_{n+1} = K_n$ . Similarly, the Markov chain on symbols  $S = S_1 \dots S_K$  is augmented with two bit counters  $N_k$  and  $M_k$ , corresponding to the sequences of bits generated by the VLC encoding of the two sequences of index  $I$  and  $J$ . See [3] for more details. Fig. 2 shows the dependencies between the variables involved in the model.

#### 4. ESTIMATION

In order to proceed with the estimation of the sequence of symbols, one has first to get  $\mathbb{P}(\bar{Y}_k, \bar{Z}_k | S_k, N_k, M_k)$  to be used as input to a two-sweep estimation algorithm on the source model. Given the tree structure of dependencies, the observations on each channel are independent conditionally to the two sequences of indexes. Therefore, this soft information can be obtained by merging soft information on processes  $(I, N)$  and  $(J, M)$  as

$$\mathbb{P}(\bar{Y}_k, \bar{Z}_k | S_k, N_k, M_k) = \mathbb{P}(\bar{Y}_k | I_k, N_k) \mathbb{P}(\bar{Z}_k | J_k, M_k) |_{S_k=(I_k, J_k)}. \quad (3)$$

It has been shown in [3] that

$$\mathbb{P}(\bar{Y}_k | I_k, N_k) \propto \frac{\dot{\mathbb{P}}_i(I_k, N_k | \bar{Y}_1^k)}{\dot{\mathbb{P}}_i(I_k, N_k | \bar{Y}_1^{k-1})}, \quad (4)$$

where  $\dot{\mathbb{P}}_i(I_k, N_k | \bar{Y}_1^k)$  is defined as  $\dot{\mathbb{P}}_i(I_k, N_k | \bar{Y}_1^k) \propto \mathbb{P}(\bar{Y}_1^k = \bar{y}_1^k, I_k, N_k)$  and denotes the quantity  $\mathbb{P}(I_k, N_k | \bar{Y}_1^k)$  for particular values  $y$  of the observations  $Y = y$  (similarly for  $\mathbb{P}(J_k, M_k | \bar{Z}_1^k)$ ). The notation  $\propto$  stands for a renormalization over  $(I_k, N_k)$ . The subscript  $i$  denotes the probability distributions under the assumption of symbol independence. One must then first derive the soft information on

each augmented process  $(I_k, N_k)$ , and  $(J_k, M_k)$  separately,  $k = 1, \dots, K$ .

Let us now consider the derivation of  $\dot{\mathbb{P}}_i(I_k, N_k | \bar{Y}_1^k)$ . Since some values for  $(X_n, K_n)$  correspond to the leaves of the codeword tree, i.e. possible symbols at time  $k$ ,  $\mathbb{P}_i(X_n, K_n | Y)$  corresponds to the desired  $\mathbb{P}_i(I_k, N_k | Y)$ . Therefore, to obtain soft information on the process  $(I, N)$  it suffices to estimate the augmented state process  $(X, K)$ . Under the assumption of symbol independence, the quantity  $\mathbb{P}_i(X_n, K_n | Y)$  can be obtained with one single sweep of the MPM estimation on the source coder model, since symbol independence induces  $\mathbb{P}_i(X_n, K_n | Y) = \mathbb{P}_i(X_n, K_n | Y_1^n)$ . Therefore, the quantity  $\dot{\mathbb{P}}_i(I_k, N_k | \bar{Y}_1^k)$  can be derived by the forward sweep of estimation on the source coder model, assuming symbols independent. The quantity  $\dot{\mathbb{P}}_i(I_k, N_k | \bar{Y}_1^{k-1})$  is derived as

$$\dot{\mathbb{P}}_i(I_k, N_k | \bar{Y}_1^{k-1}) = \mathbb{P}(I_k) \cdot \dot{\mathbb{P}}_i(N_{k-1} | \bar{Y}_1^{k-1}) \Big|_{N_{k-1}=N_k-\mathcal{L}(I_k)} \quad (5)$$

The soft information  $\mathbb{P}(S_k, N_k, M_k | \bar{Y}_k, \bar{Z}_k)$  obtained by merging the soft information on each description (Eq. 3) is used as input of a two sweep procedure that will estimate the process  $(S, N, M)$ . The difficulty in this estimation resides in the random structure of the Bayesian network connecting the different models. However, given process  $S$ , hence  $I, J, N$  and  $M$ , the tree structure is fixed : for every value  $(s_k, n_k, m_k)$  of the triple  $(S_k, N_k, M_k)$ , vectors  $\bar{Y}_1^k = Y_1 \dots Y_{N_k}$ ,  $\bar{Y}_{k+1}^k = Y_{N_k+1} \dots Y_N$ ,  $\bar{Z}_1^k = Z_1 \dots Z_{M_k}$  and  $\bar{Z}_{k+1}^k = Z_{M_k+1} \dots Z_M$  are perfectly defined. If we define  $\dot{\mathbb{P}}(S_k, N_k, M_k | \bar{Y}_1^k, \bar{Z}_1^k)$  for given observations  $y$  and  $z$  as  $\dot{\mathbb{P}}(S_k, N_k, M_k | \bar{Y}_1^k, \bar{Z}_1^k) \propto \mathbb{P}(\bar{Y}_1^k = \bar{y}_1^k, \bar{Z}_1^k = \bar{z}_1^k, S_k, N_k, M_k)$ , then the equation

$$\mathbb{P}(S_k, N_k, M_k | Y, Z) \propto \dot{\mathbb{P}}(S_k, N_k, M_k | \bar{Y}_1^k, \bar{Z}_1^k) \cdot \dot{\mathbb{P}}(\bar{Y}_{k+1}^k, \bar{Z}_{k+1}^k | S_k, N_k, M_k) \quad (6)$$

holds for those particular values  $y$  and  $z$  and provides a valid factorization for performing the estimation of  $(S, N, M)$ .

#### 5. EXPERIMENTATION RESULTS

Experiments have been performed on a first-order Gauss-Markov source, with zero-mean, unit-variance and correlation factor  $\rho = 0.5$ . The rate of the MDUSQ is 3 bpsps/desc. and the diagonal width of the index assignment is  $d = 2$ . Hence the source is quantized with a 34 levels uniform quantizer (about 5 bits), on the interval  $[-3, 3]$ . This is equivalent to a redundancy rate of 5/6 in channel coding terminology. We consider sequences of length  $K = 50$  symbols. The VLC source coder is based on two Huffman codes, designed for the stationary distributions of each description. All simulations have been performed assuming an AWGN channel with a BPSK modulation. The results are averaged over 500 realizations.

Fig. 4 provides the residual symbol error rates (SER) and the signal to noise ratio (SNR) for different channel  $E_b/N_0$ . On each plot, MDUSQ soft decoding is compared to soft decoding without redundancy adding, and to source-channel turbo decoding [3] with a recursive systematic convolutional code of rate 1/2 punctured to 5/6 (4 iterations). Relative performances of the different approaches depends on the amount of correlation present within the source. For a correlation factor  $\rho = 0.1$  (Fig. 3), and for  $E_b/N_0 \leq 3\text{db}$ , MDUSQ brings a gain between 2 and 4 dB in SNR over the other methods. When the source has little correlation left, due to prior treatments, then the most appropriate strategy is to reintroduce redundancy directly in the compressed stream. Experiments have also been performed with varying levels of added redundancy. It shows that MDUSQ soft decoding outperforms source-channel turbo decoding for low levels of redundancy ( $d \geq 2$ ).

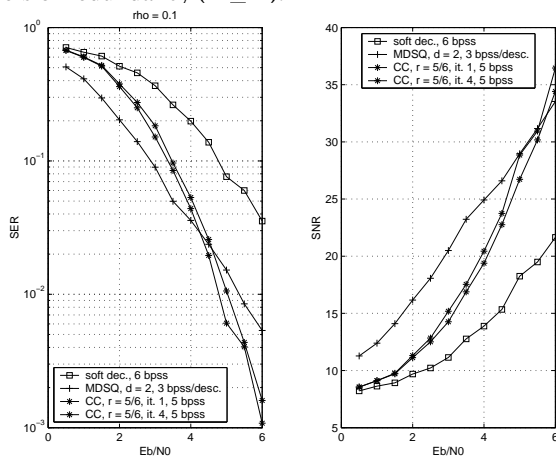


Fig. 3. SER and SNR performances ( $\rho = 0.1$ , rate 5/6).

## 6. CONCLUSION

We have presented a soft decoding procedure of multiple descriptions for robust transmission of Gauss-Markov sources over error prone environments. The method proposed is exact, i.e., leads to the true values of  $\mathbb{P}(S_k|Y, Z)$ . Experiments have shown that MDSQ brings a significant gain for low correlated sources transmitted on channels with high error rates, compared to the usage of classical error correction codes. Notice that in the context of joint source-channel coding, the sources to be considered have already undertaken pre-treatments (via e.g. transforms) and that the resulting sources to be encoded are characterized by low correlation factors. Current work is dedicated to MDUSQ soft decoding of arithmetic encoded sources for inclusion into a state-of-the-art image codec.

## 7. REFERENCES

[1] V.K. Goyal and J. Kovacevic. Optimal multiple description transform coding of gaussian vectors. *Proc. IEEE Data Compression Conf., DCC'98*, pages 388–397, March 1998.

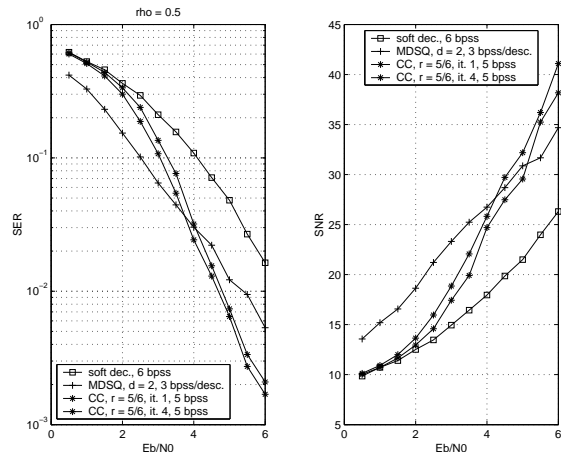


Fig. 4. SER and SNR performances ( $\rho = 0.5$ , rate 5/6).

- [2] T. Guionnet, C. Guillemot, and S. Pateux. Embedded multiple description coding for progressive image transmission over unreliable channels. In *Intl. Conf. on Image Processing, ICIP'2001*, Oct. 2001.
- [3] A. Guyader, E. Fabre, C. Guillemot, and M. Robert. Joint source-channel turbo decoding of entropy coded sources. *IEEE Jour. on Selected Areas in Communication*, Sept. 2001.
- [4] L. Ozarow. On a source coding problem with two channels and three receivers. *Bell Syst. Tech. J.*, 59:1909–1921, Dec. 1980.
- [5] M. Srinivasan. Iterative decoding of multiple descriptions. In *Proc. IEEE Intl Conf. on Data Compression, DCC'98*, May 1998.
- [6] V.A. Vaishampayan. Design of multiple description scalar quantizers. *IEEE Trans. On Information Theory*, 39(3):821–834, May 1993.
- [7] V.A. Vaishampayan and J. Domaszewicz. Design of entropy constrained multiple description scalar quantizers. *IEEE Trans. On Information Theory*, 40:245–251, January 1994.
- [8] S-M. Yang and V.A. Vaishampayan. Low-delay communication for rayleigh fading channels. *IEEE Trans. On Communications*, 43(11):2771–2783, Nov. 1995.

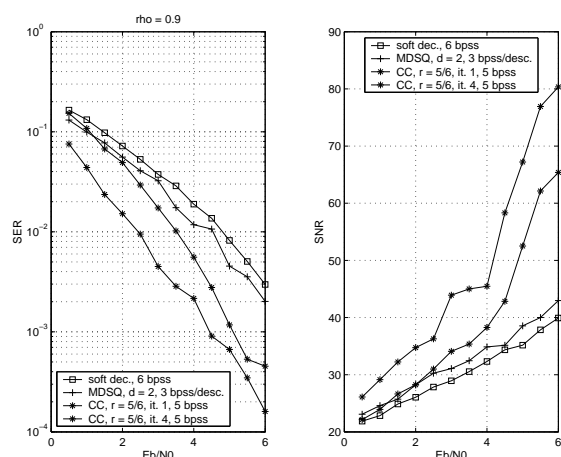


Fig. 5. SER and SNR performances ( $\rho = 0.9$ , rate 5/6).