AN ALGORITHM FOR TRANSFORM CODING FOR LOSSY PACKET NETWORKS

Francesco Palmieri, Dario Petriccione

Dipartimento di Ingegneria dell'Informazione, II Università di Napoli
Real Casa dell'Annunziata, Via Roma 29 - 81031 Aversa, CE, Italy
Email: francesco.palmieri@unina2.it ; dariopetr@supereva.it

ABSTRACT

We propose an algorithm to compute a modification of the classical discrete Karhunen-Loeve Transform (KLT) useful when some of the coefficients are randomly unavailable for reconstruction. Such a scheme can provide Multiple Description Coding (MDC) for signals and images transported by lossy packet links. The modification of the KLT is based on a "correlating" block that, from knowledge of the channel erasure statistics, is optimized with a gradient algorithm to provide minimum average reconstruction error. A set of simulations show appreciable improvements over standard schemes.

1. TRANSFORM CODING WITH ERASURES

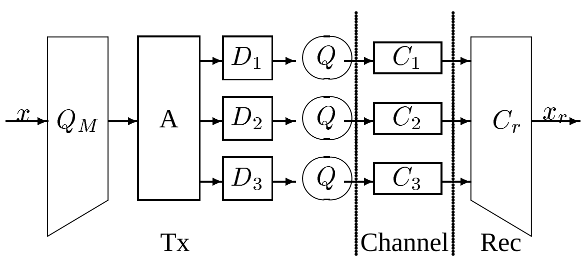

Fig. 1. The coder-channel-decoder cascade

Consider an N-dimensional source emitting independent random vectors $\mathbf{x}$. Transform coding is based on computation and quantization of linear projections of $\mathbf{x}$. Figure 1 shows the coding-decoding cascade. The first block maps the N-dimensional input vector $\mathbf{x}$ onto the M-dimensional subspace of the M first eigenvectors of $E[\mathbf{x}\mathbf{x}^T]$ (Principal components). The coefficients after such a transformation are known as the Discrete Karhunen-Loeve Transform (KLT).

Suppose now that a KLT is to be used for coding real-time signals or images to be sent in a lossy packet network (such as IP). Typically, the transform coefficients are quantized, loaded into packets and sent over the channel. At the receiver some packets may not be available for recontruction due to network congestion or excessive delays. Therefore the introduction of a protection mechanism may be appropriate to ease the receiver operation in restoring the lost pieces. A number of strategies have been proposed in the literature and they can be subdivided into two main categories: a) redundancy coding with codes such as Reed-Solomon or similar [1] (channel coding); b) source-channel transform coding in which the source transform, together with the compression task, has the role of robustifying the coefficients against random losses [5][6]. In this paper we take approach b) by proposing an algorithm for computing the structure of an MxM linear block A included before transmission. The total transform matrix will be clearly AQ_M. If we ignore for now the other blocks of Figure 1 and suppose that the M-dimensional vector $\mathbf{z} = A\mathbf{y}$ is transmitted on an erasure channel that at each channel use removes some of its components, we would like to be able to design A to minimize the degradation on the final reconstruction.

This framework fits well in a packet switching scenario where $\mathbf{z}$ is partitioned into coefficient subsets loaded into separate packets and where erasures may happen packet-wise. Such a strategy is related to what is becoming popular in the signal processing literature as Multiple Description Coding (MDC), since the various packets that code a frame could be considered as different descriptions of the original source (see [5] and [6] for more references).

Note that matrix A plays the critical role of "correlating" the coefficients coming from the KLT block with the objective of providing better protection againt the effects of the erasures on the reconstruction. Clearly in a real transmitter the transform coefficients have to be quantized. To concentrate better on the role of matrix A, we start just considering the transform blocks, delaying the inclusion of quantization to a later section.

The general model that follows supposes that at each observation a number of components of $\mathbf{z}$ may be unavailable for reconstructing $\mathbf{x}$, i.e. N_e ($\leq M$) random *erasures* have happened. Describing the erasure process with the random binary vector $\mathbf{e}^T = (e_1, e_2, ..., e_M)$, with $e_i = 0$, if the i-th component is erased and $e_i = 1$ otherwise, a compact description of the erasure process can be done by defining the residual vector $\mathbf{z}_e$ containing the $N - N_e$ survivor components kept in the same order. The operation can be described via an $(M - N_e) \times M$ permutation matrix $P(\mathbf{e})$:

$$\mathbf{z}_e = P(\mathbf{e})\mathbf{z}.$$

The formulation is easily visualized with the help of the following example: consider $M = 6$ and an erasure event that for a given $\mathbf{x}$ cancels the second and the fourth components of $\mathbf{z}$. The vector $\mathbf{z}_e$ is obtained as

$$\begin{bmatrix} z_{e1} \\ z_{e2} \\ z_{e3} \\ z_{e4} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \end{bmatrix}. \quad (1)$$

Note that the permutation matrix P has zeros in the columns corresponding to the erasures and has the property that $P(\mathbf{e})P^T(\mathbf{e}) = I_{N_e}$ and $P^T(\mathbf{e})P(\mathbf{e}) = diag(\mathbf{e})$.

A first step in our investigation is devoted to establish the performance of the communication chain when no correlating block A is included $(A = I_M)$, i.e. when classical transform coding is used.

Note that if some coefficients are erased by the channel, and the receiver knows which ones, it can build its best mean square reconstruction of $\mathbf{x}$ from $\mathbf{e}$ and $\mathbf{z}_\mathbf{e}$. Standard Wiener filter theory [3], establishes that the best reconstruction is just obtained by using the surviving coefficients "re-propagated" back into the projection matrix Q_M. This is easily understood because the coefficients $\{y_i\}$ are uncorrelated. Therefore surviving coefficient cannot carry information about other coefficients that have been erased. This suggests immediately that a transform that protects the information against the erasures must be a "correlating" transform.

As we discuss the inclusion of the matrix in the comunication chain in the following, let us evalutate first the performance of the KLT in the presence of erasures. The mean squared error for each erasure is:

$$\mathcal{E}(\mathbf{e}) = \sum_{i=M+1}^{N} \lambda_i + \sum_{i=1}^{M} (1 - e_i)\lambda_i, \quad (2)$$

that averaged over all possible erasure events gives:

$$\mathcal{E} = E_\mathbf{e}[\mathcal{E}(\mathbf{e})] = \sum_{i=M+1}^{N} \lambda_i + \sum_{i=1}^{M} (1 - E[e_i])\lambda_i = \mathcal{E}_c + \mathcal{E}_e, \quad (3)$$

where λ_i are the eigenvalues of R_x, $\mathcal{E}_c$ is the distortion due to compression and $\mathcal{E}_e$ the distortion due to the erasures.

2. THE CORRELATING BLOCK

The question we are addressing now is: can we improve the average reconstruction error, if we know something about the erasure process? In other words, if the erasure process can be statistically characterized, can we find a matrix A that leads to better reconstructions? The answer is affirmative as the "correlating" matrix increases the number of degrees of freedom in the design of the coder and could be considered a sort of "pre-emphasis" filter bank. The idea was proposed in [5] and [6] where it has been suggested that we build the matrix with a sequence of "lifting steps." The receiver block in Figure 1, at each realization $(\mathbf{x}, \mathbf{e})$, knows which erasures have happened, and gives its best mean square error [3] reconstruction as:

$$\mathbf{x}_r = C_r^T(\mathbf{e})\mathbf{z}_e, \quad (4)$$

with $\mathbf{z}_e = P(\mathbf{e})\mathbf{z}$, and $C_r(\mathbf{e}) = E[\mathbf{z}_e\mathbf{z}_e^T]^{-1}E[\mathbf{z}_e\mathbf{x}^T]$, or

$$\begin{aligned} C_r(\mathbf{e}) &= (PA^TQ_M^TR_xQ_MAP^T)^{-1}PA^TQ_M^TR_x \\ &= (PA^T\Lambda_MAP^T)^{-1}PA^T\Lambda_MQ_M, \end{aligned} \quad (5)$$

where $\Lambda_M = diag(\lambda_1, ..., \lambda_M)$ and we have omitted the argument of P for simplicity of notation. The distortion due to a specific erasure is:

$$\begin{aligned} &\mathcal{E}_e(\mathbf{e}) = \sum_{i=1}^{M} \lambda_i \\ &-tr[R_xQ_MAP^T(PA^T\Lambda_MAP^T)^{-1}PA^TQ_M^TR_x]. \end{aligned} \quad (6)$$

Using the property $tr[AB] = tr[BA]$, the expression can be re-written as:

$$\mathcal{E}_e(\mathbf{e}) = \sigma_c^2 - tr[PA^T\Lambda_M^2AP^T(PA^T\Lambda_MAP^T)^{-1}]. \quad (7)$$

Clearly if $P = I_M$ (no losses), the second term becomes $tr[\Lambda_M]$ independently of A. Also if A were a diagonal matrix, there would be no difference with respect to the KLT.

The problem of non trivial optimal choice for A is then formulated as:

$$\begin{cases} A_o = argmax_A\ \mathcal{C} \\ \mathcal{C} = E_\mathbf{e}[tr[PA^T\Lambda_M^2AP^T(PA^T\Lambda_MAP)^{-1}]]. \end{cases} \quad (8)$$

Note that the performance measure depends only on the eigenvalues and the statistics of the erasure process. We have computed the gradient of $\mathcal{C}$ with respect to A using techniques from matrix differential calculus [4]:

$$\begin{cases} \nabla_A\mathcal{C} = 2E_\mathbf{e}[\Lambda_M^2AP^TB_1^{-1}P - \Lambda_MAB_1^{-1}B_2B_1^{-1}P^T] \\ B_1 = PA^T\Lambda_MAP^T \\ B_2 = PA^T\Lambda_M^2AP^T \end{cases} \quad (9)$$

For space reasons the details of the computation cannot be reported here, but will be included in a larger paper. They are in any case available on request (use our e-mail address indicated in the title). A search for the optimal matrix A can be done with a gradient ascent algorithm as $\Delta A = \mu \nabla_A \mathcal{C}$. It appears hard to infer in general from the structure of the cost function what is the theoretical lower limit for the recostruction error. However we have found experimentally that in all cases our gradient searches give solutions for A that largely improve over the standard KLT scheme. Current investigation is devoted to the inclusion of appropriate constraints on A.

Note that the gradient expression requires evaluation of the average over all the possible erasure events and we need to know or estimate the probability of each configuration. This may be exponentially complex if the number of configurations is large. In our applications, however, it is likely that the coefficient set $\mathbf{z}$ are to be divided into a small number of subgroups (packets, in our experiments only three), that can exhibit a manageable number of loss configurations.

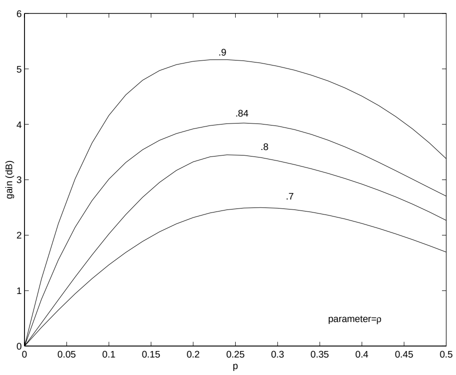

Fig. 2. Computed theoretical gain on the KLT (no quant.)

A simulation without quantization: We have performed a large number of simulations on synthetic and natural signals finding very consistent results [2]. We report here the case of transform coding of a Markov sequence with autocorrelation $r(i) = \rho^{|i|}$. This is a good test for our system because this model may approximate rather well the autocorrelation structure of signals of practical interest such as image blocks and some speech segments. In this experiment the projections are supposed to be sent on the lossy channel without quantization. The window size is $N = 160$ and only $M = 90$ components are kept after KLT. The values of ρ range from 0.7 to 0.9 which in the absence of losses correspond to a compression mean square error $\mathcal{E}_c$ that ranges from about -10dB to -16 dB, respectively. The coefficients are partitioned into three groups of dimension 30 and independent Bernoulli loss of packets with various loss probabilities p are simulated for the channel. We have computed the average mean square error after losses when the KLT coefficients are grouped in three packets. It is easy to see that the grouping for the KLT is irrelevant since losses are independent and the coefficients uncorrelated. This distorsion has been compared to the distorsion obtained after the inclusion of the correlating matrix A. Matrix A is computed with the gradient ascent algorithm from knowledge of the eigenvalue set and the loss probability. The gain is shown in Figure 2 with the highest improvement, about 5 dB, obtained when the process is highly correlated $\rho = 0.9$ and the loss probability is about 0.25. However, as it can be seen from the figure, the improvement is appreciable in most practical cases.

3. THE TRANSFORM WITH QUANTIZATION

The next obvious step in this investigation is to establish if the introduction of the correlating matrix A leads to improvements also when scalar quantization on the coefficients is included.

In the standard KLT scheme if scalar quantization is sufficiently fine to be modeled as additive independent gaussian noise the total reconstruction mean square error at the receiver can be estimated to be [1]: $\mathcal{E} = \sigma_t^2 + (p-1)\sum_{i=1}^{M} \lambda_i + (1-p)\mathcal{E}_q = \mathcal{E}_c + \mathcal{E}_e + \mathcal{E}_q$, where $\mathcal{E}_q = M \cdot C \cdot 2^{-2\overline{R}} (\prod_{i=0}^{M-1} \lambda_i)^{1/M}$, with $\overline{R} = \frac{N_b}{M}$, $N_b =$ *number of bits used by the quantizer*, C a constant and Bernoulli losses with loss probability p. The quantizer is the optimal gaussian quantizer for each coefficient based on the knowledge of the relative variance (eigenvalue) [1].

The quasi-optimal performance of the scalar quantizer is to attribute to the fact that the coefficients are uncorrelated. Therefore in our modified scheme, when a correlating block A is introduced, a loss of performance may happen if we perform scalar quantization on the coefficients as they are [5][6]. Therefore, we introduce within each partition subset a decorrelating linear block D_i as shown Figure 1. In this way in each set we have coefficients that are globally correlated to the others, but uncorrelated among themselves. By stretching the assumption on the additive independent noise model on each coefficient in this modified scheme, we can re-work the calculation done above with $\mathbf{x}_r = C_r^T(\mathbf{e}, \eta)\mathbf{z}_e$, $\mathbf{z}_e = P(\mathbf{z}' + \eta)$ and $\mathbf{z}' = DA^T Q_M^T \mathbf{x}$, where we have included in the vector η the independent additive noise components, and in the block-diagonal matrix D the decorrelating blocks. Using Wiener filter theory, the best recostruction matrix, which depends also on the autocorrelation R_η, is $C_r(\mathbf{e}) = E[\mathbf{z}_e\mathbf{z}_e^T]^{-1}E[\mathbf{z}_e\mathbf{x}^T]$, with mean square error: $\mathcal{E}(\mathbf{e}) = \sigma_t^2 - tr\left\{E[\mathbf{x}\mathbf{z}_e^T]E[\mathbf{z}_e\mathbf{z}_e^T]^{-1}E[\mathbf{z}_e\mathbf{x}^T]\right\}$. The expectation is computed over all the possible erasure events $\mathbf{e}$ and the quantization process η. By using steps similar to the ones used for the unquantized case we have:

$$\mathcal{E}(\mathbf{e}) = \sigma_t^2 - tr[PD^T A^T \Lambda_M^2 ADP^T (PD^T A^T \Lambda_M ADP^T + PR_\eta P^T)^{-1}]. \quad (10)$$

The problem of non trivial optimal choice for A is formulated as:

$$\begin{cases} A_o = argmax_A\, \mathcal{C} \\ \mathcal{C} = E_{e,\eta} tr[PD^T A^T \Lambda_M^2 ADP^T \\ \quad (PD^T A^T \Lambda_M ADP^T + PR_\eta P^T)^{-1}]. \end{cases} \quad (11)$$

The gradient of $\mathcal{C}$ with respect to A is:

$$\begin{cases} \nabla_A(.) = 2E_{e,\eta}\left[\Lambda_M^2 ADP^T B_1^{-1} PD^T - \right. \\ \quad \left. -\Lambda_M ADP^T B_1^{-1} B_2 B_1^{-1} PD^T\right], \\ B_1 = PD^T A^T \Lambda_M ADP^T + PR_\eta P^T \\ B_2 = PD^T A^T \Lambda_M^2 ADP^T \end{cases} \quad (12)$$

A search for the optimal matrix A can be done, like in the analysis without quantization, with a gradient ascent algorithm. Unfortunately, the cost function includes the quantization error autocorrelation matrix, which in turn depends on A and on how the scalar quantizers and the decorrelating blocks in D are designed.

Therefore, we adopt the following computing strategy: 1) divide the M coefficients into packets; 2) compute the best matrix A without quantization; 3) compute the decorrelating blocks in D using spectral decomposition; 4) calibrate a gaussian compander for each coefficient on the basis of the estimated variances; 6) distribute the total number of bits N_b among all the coefficients using an optimal bit allocation algorithm [1]; 7) estimate the quantization noise autocorrelation matrix R_η; 8) repeat the gradient algorithm using also R_η to find a new matrix A; iterate.

The computational complexity of such an algorithm can be considerable if we estimate the autocorrelation matrix R_η at each gradient iteration. However, various heuristics can be used to speed up the computation such as maintaining the same estimation for R_η if the average error has not varied too much, or performing similar strategies.

Figure 3 shows the theoretical improvement of the new method over the KLT for the Markov sequence model when $\rho = .84$. The same optimal bit allocations are used for both schemes. These results, which are consistent with many others obtained on synthetic and natural signals [2], show that the solution remains quite robust with respect to quantization. We find that when at least 2-3 bits in the average are assigned to each coefficient, there is very little difference in performance with respect to the unquantized case.

As a verification, we have also repeated the simulation of Fig. 3 by generating a Markov sequence and applying to it the optimal coders and decoders at various loss probabilities. Figure 4 shows the results for $\rho = 0.84$. The curve is very close to the expected theoretical result. This is also in strong support of the assumptions that have been necessary to model quantization.

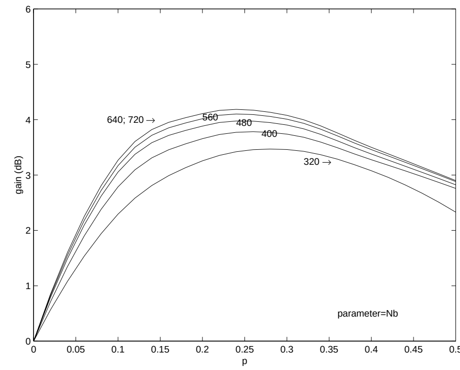

Fig. 3. Computed theoretical gain over KLT (quantization included, $\rho = 0.84$)

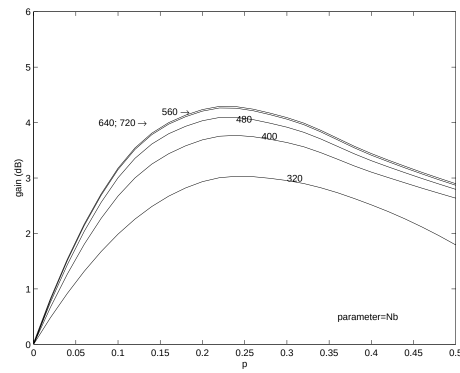

Fig. 4. Simulated gain over KLT (quantization included, $\rho = 0.84$)

4. REFERENCES

[1] A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, 1989.
[2] F. Palmieri et al. "Optimal Transform Coding for Audio and Images on Lossy Packet Channels," *Proc. of Int. Conf. Packet Video 2000*, Forte Village, Italy, March 2000.
[3] S. Haykin, *Adaptive Filter Theory*, Second Edition, Prentice Hall, 1991.
[4] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, John Wiley, 1988.
[5] J. Kovacevic and V. K. Goyal, "Multiple Descriptions, Source-Channel Coding Methods for Communications," *Proc. Int. Workshop on Multimedia Communication*, Ischia, Italy, Sept. 1999.
[6] V. Goyal, J. Kovacevic and M. Vetterli, "Quantized Frame Expansions as Source-Channel Codes for Erasure Channels," *Proc. Data Compression Conference*, Snowbird, UT, March 1999.