University of Waterloo
Department of Electrical & Computer Engineering
Waterloo, Ontario, Canada, N2L 3G1
Technical Report UW-E&CE#2001-3

Sequential Vector Decorrelation Technique

Farshad Lahouti and Amir K. Khandani
March 22, 2001

Sequential Vector Decorrelation Technique

F. Lahouti, A. K. Khandani*
Dept. of E&CE, University of Waterloo

April 18, 2000

Abstract

In this manuscript, we introduce a technique for sequential decorrelation of vector sources. This is useful for efficient quantization of vector sources where there is memory between the components. This method uses a triangular transform matrix for optimal decorrelation of source samples. It is shown that this technique is equivalent to the case where for each vector component all the previous components are used for prediction and the prediction error is quantized instead. The new method, which we refer to as Sequential Vector Decorrelation Technique is employed for scalar quantization of LSF parameters in Linear Predictive Coding of speech. Substantial improvements over DPCM and scalar quantization techniques are achieved.

*The authors wish to thank Dr. E. Jernigan for his valuable comments on an initial version of this work.

1 Introduction

Suppose we have a random vector process $\mathcal{X}$ and would like to efficiently quantize it with a certain quota of bits. Each sample of this process is an N-tuple vector $\mathbf{x} = \{x_1, x_2, \ldots, x_N\}$. In general, different components of vector $\mathbf{x}$ are not independent. Consequently, separate quantization of each component would be an inefficient approach. Different methods such as transform coding and vector quantization have been introduced in the literature for efficient quantization of such a source. The idea of transform coding is to remove the source correlation by performing a linear transform $\mathbf{A}$ on the input vector $\mathbf{x}$.

$$\mathbf{y} = \mathbf{A}\mathbf{x} \tag{1}$$

The output vector $\mathbf{y}$ is a vector with the same size as $\mathbf{x}$ with its components, called *transform coefficients*, much less correlated or uncorrelated. Having in this sense removed the redundancy, it is believed that we will be able to quantize the components of $\mathbf{y}$ more efficiently. In general, this process results in a non-uniform distribution of energy between the transform coefficients. The KL transform is optimum in the sense that it completely decorrelates the vector components [1]. This transform is based on eigenvector decomposition and is data dependent.

2 The Structure

In this work, we are interested in a linear transformation that optimally decorrelates the vector components in a causal manner. We are particularly interested in a causal form since it can be easily incorporated in the structure of some advanced vector quantizers, such as Trellis-Based Quantizers [4] or Lattice-Based Quantizers. We call this technique the Sequential Vector Decorrelation Technique or SVDT and we will explain it briefly below.

Consider the linear transform $\mathbf{A}$ defined as

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \\ \vdots \\ \mathbf{a}_N \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & 0 & 0 & \dots & 0 \\ a_{31} & a_{32} & a_{33} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & a_{N3} & a_{N4} & \dots & a_{NN} \end{bmatrix}, \tag{2}$$

where $\mathbf{a}_k$ is a row vector and we have:

$$y_k = \mathbf{a}_k.\mathbf{x} = \sum_{j=1}^{k} a_{kj} x_j, \qquad 1 \le k \le N \tag{3}$$

The quantizer then encodes y_k to $\hat{y}_k$. Our objective in determining the components of transform matrix $\mathbf{A}$, is to remove the correlation between different components of the input vector $\mathbf{x}$ and produce optimally decorrelated transformed coefficients $\mathbf{y} = \{y_1, y_2, \dots, y_N\}$, i.e.,

$$r_{\mathcal{Y},ij} = E[y_i y_j] = 0, \qquad i \neq j, \tag{4}$$

$$r_{\mathcal{Y},ii} = E[y_i y_i] > 0, \qquad 1 < i, j < N. \tag{5}$$

or equivalently, the autocorrelation matrix of $\mathbf{y}$ is diagonal,

$$\mathbf{R}_{\mathcal{Y}} = E[\mathbf{y}^T \mathbf{y}] = Diagonal(r_{y,11}, r_{y,22}, \dots, r_{y,NN}) \tag{6}$$

We begin here with a simple example.

2.1 A simple second-order case

Consider 2-dimensional vector $\mathbf{x}$, output of a random source:

$$\mathbf{x} = (x_1, x_2)^T \tag{7}$$

We assume that the samples of this random vector are correlated and the autocorrelation matrix is given by $\mathbf{R}_{\mathcal{X}}$:

$$\mathbf{R}_{\mathcal{X}} = \begin{bmatrix} r_{\mathcal{X},11} & \rho \\ \rho & r_{\mathcal{X},22} \end{bmatrix} \qquad r_{\mathcal{X},11}, r_{\mathcal{X},22} > 0 \tag{8}$$

in which we have benefited form the fact that the autocorrelation matrix is symmetric and non-negative definite. Now assume that we want to employ transform **A** to decorrelate the two samples and we are particularly interested in a causal (lower triangular) form for the matrix **A**. The transform matrix **A** has the form:

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix} \tag{9}$$

The equation which needs to be solved is as follows:

$$\mathbf{R}_{\mathcal{Y}} = \mathbf{A}\mathbf{R}_{\mathcal{X}}\mathbf{A}^T = \begin{bmatrix} r_{\mathcal{Y},11} & 0 \\ 0 & r_{\mathcal{Y},22} \end{bmatrix} \tag{10}$$

in which diagonal elements of $\mathbf{R}_{\mathcal{Y}}$ are required to be nonnegative so that, this matrix is non-negative definite. Expanding the equation (10) and solving for a_{11}, a_{12} and a_{22}, we find out that there is only one constraint on the choice of the unknown parameters. A solution to the problem can be the addition of two more constraints to normalize the basis vectors. We choose here, the parameters a_{11} and a_{22} to be 1 and solve for a_{21}. This results in the following solution:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ \frac{-\rho}{r_{\mathcal{X},11}} & 1 \end{bmatrix} \tag{11}$$

and the following autocorrelation matrix $\mathbf{R}_{\mathcal{Y}}$:

$$\mathbf{R}_{\mathcal{Y}} = \begin{bmatrix} r_{\mathcal{X},11} & 0 \\ 0 & r_{\mathcal{X},11}r_{\mathcal{X},22} - \rho^2 \end{bmatrix} \tag{12}$$

We know from matrix theory that the determinant of an autocorrelation matrix (which is non-negative definite) is always nonnegative. This along with (12) guarantees that $\mathbf{R}_{\mathbf{Y}}$ is nonnegative definite and is a true solution. Figure (1) shows the result of a numerical example. The sample vectors plotted in the 2-D plane are uniformly distributed in the area shown and there is a correlation between the two components. Both KLT and the transform

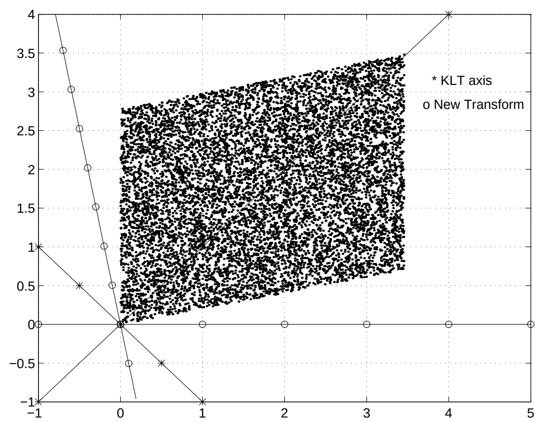

Figure 1: Decorrelation in KLT and the proposed transform

in (11) were used to decorrelate these components. Each of these transforms choose their own basis vectors which have been shown in this figure. Both figure (1) and equation (11) show that the basis vectors of the proposed transform are *not* orthogonal. We will discuss this problem further in the next section.

2.2 Problem of Orthogonality

The assumption of orthogonality is vital in the theory of transform coding. It is the base of the variance-preserving property of linear transforms. This property is used to model the bit allocation problem and to subsequently derive the maximum achievable coding gain equation [2]. When this property holds, the average variance of the quantization noise introduced while quantizing a vector of transform coefficients is equal to that of the error produced while reconstructing the input vector from these coefficients.

In the last section, we observed that the proposed method, although diagonalizes the autocorrelation matrix $\mathbf{R}_{\mathcal{X}}$, but chooses non-orthogonal basis vectors for the transformation. The special form of the transformation chosen in equation (2) guarantees the independence of basis vectors and hence prevents from information loss ($a_{ii} \neq 0, 1 \leq i \leq N$). However, non-orthogonality can cause the propagation of quantization error to other dimensions while reconstructing the data. To avoid this problem, we benefit from special form of the transformation matrix. This matrix has a lower triangular form which allows us to transform and hence encode the source samples sequentially. In other words, the only components of vector $\mathbf{x}$ used to determine the transform coefficient number k are the k'th component and the ones before that (see equation (3)).This allows us to employ a closed-loop-DPCM-based scheme to encode the transform coefficients. This scheme will take into consideration the noise produced in the quantization of elements up to a certain time and hence prevents from propagation of the noise in the elements to be quantized next. This will transform the equation (3) to:

$$y_k = \sum_{j=1}^{k-1} a_{kj} \hat{x}_j + a_{kk} x_k, \qquad 1 \leq k \leq N \tag{13}$$

in which $\hat{x_j}$ is the reconstructed version of x_j. This is a very interesting feature which allows this technique to benefit from other effective properties of closed-loop DPCM schemes as well. However, one can observe that by replacing equation (3) by (13) and considering the sequential encoding of the vector components, this technique will not be a transform any more.[1] The optimum decorrelation property still holds if the vector components are quantized finely enough[2].

2.3 The General N'th order case

A similar approach as presented in the above example, was taken to find the transforms of higher orders. According to Equation (3), y_k is given by

$$y_k \; = \; \mathbf{a}_k.\mathbf{x} \; = \; \sum_{j=1}^{k} a_{kj} x_j, \qquad 1 \leq k \leq N. \tag{14}$$

We derive the transform matrix such that the transform coefficients $\mathbf{y} = [y_1, y_2, \ldots, y_N]$ are decorrelated, i.e.,

$$r_{\mathbf{y},ij} \;= E[y_i y_j] \;= 0, \qquad i \neq j, \tag{15}$$

$$r_{\mathbf{y},ii} \;= E[y_i y_i] \;> 0, \qquad 1 < i,j < N, \tag{16}$$

[1] We will continue to use the term transform for the matrix given in equation (2).

[2] see fine quantization for DPCM systems in [2]

or equivalently, the autocorrelation matrix of $\mathbf{y}, \mathbf{R_y}$, is diagonal. It is straight forward to see that (15) holds if the transform coefficients and the source samples are orthogonal, i.e.,

$$E[y_i x_j] = 0, \qquad 1 \le j < i \le N. \tag{17}$$

Equations (14) and (17) provide the necessary means to calculate the matrix $\mathbf{A}$. Since the number of unknowns are larger than the number of constraints, we assume

$$a_{jj} = 1, \qquad 1 \le j \le N. \tag{18}$$

To calculate the remaining unknowns, we multiply Equation (14) by x_i, $1 \le i < k$, and compute the expectation. This results in $k-1$ equations for each corresponding k. We have

$$E[y_k x_i] = \sum_{j=1}^{k} a_{kj} E[x_i x_j], \qquad 2 \le k \le N, 1 \le i < k. \tag{19}$$

Considering Equations (17) and (18), this will result in

$$E[x_{ik}^2] = -\sum_{j=1}^{k} a_{kj} E[x_i x_j], \qquad 2 \le k \le N, 1 \le i < k \tag{20}$$

and subsequently,

$$\begin{bmatrix} a_{k1} \\ a_{k2} \\ \vdots \\ a_{k\,k-1} \end{bmatrix} = \begin{bmatrix} r_{x,11} & r_{x,12} & \cdots & r_{x,1\,k-1} \\ r_{x,21} & r_{x,22} & \cdots & r_{x,2\,k-1} \\ \vdots & \vdots & \ddots & \vdots \\ r_{x,k-1\,1} & r_{x,k-1\,2} & \cdots & r_{x,k-1\,k-1} \end{bmatrix}^{-1} \begin{bmatrix} r_{x,1k} \\ r_{x,2k} \\ \vdots \\ r_{x,k-1\,k} \end{bmatrix}, 2 \le k \le N, \tag{21}$$

where $r_{x,ij} = E[x_i x_j], 1 \le i, j \le N$. Considering the Orthogonality Principle [1] [6], Equation (17) also result in the minimization of the transform coefficients' power in Equation (16), or the maximization of the corresponding gains (see next section). We can observe that the transform coefficients y_k are, in fact, the linear prediction errors when all the components of vector $\mathbf{x}$ prior to x_k are employed to predict the current component x_k. We also note that these derivations are fundamentally different from the case where we intend to design a predictive coder for a causal sequence [6], since those derivations rely on the stationarity of the source and attempt to exploit the temporal correlations, however, in our case we are interested in exploiting the spatial correlations where the assumption of stationarity no longer exists.

In general, there are two different approaches to determine the matrix **A**; the closed-loop approach and the open-loop approach. In the open-loop scenario, the matrix is calculated using the training database. In the closed-loop scenario, an iterative scheme, which includes quantization of the source samples of the training database and calculating the matrix based on the quantized database, is employed. In this work, we consider the open-loop scheme for simplicity.

3 Quantization of LSF parameters

In this section, we report the result of applying the SVDT to quantization of LSF parameters for speech coding applications. For more information on LSF parameters, their properties and applications refer to [3] [4] [7]. The application of this technique to our formerly proposed Block-based Trellis Quantizer [4][3] is given in [5]. One immediate advantage of the causal structure of SVDT for quantization of LSF parameters is the fact that, the quantization operation is performed sequentially and hence the quantized parameters can be reconstructed right away. Therefore, the stability of the filter can be easily verified. This will keep the overhead complexity due to the use of the SVDT at a very low level. However, this is not the case in transform coding.

3.1 Scalar Quantization with SVDT

Figure 2, compares the gains ($G = 10log_{10}\frac{\sigma^2_{x_k}}{\sigma^2_{y_k}}dB$) achieved by employing SVDT with those achieved using differential encoding of LSF parameters where the differences of consecutive LSF parameters are encoded, i.e.

$$y_1 = x_1, \tag{22}$$

$$y_k = x_k - x_{k-1}, \qquad 2 \leq k \leq N. \tag{23}$$

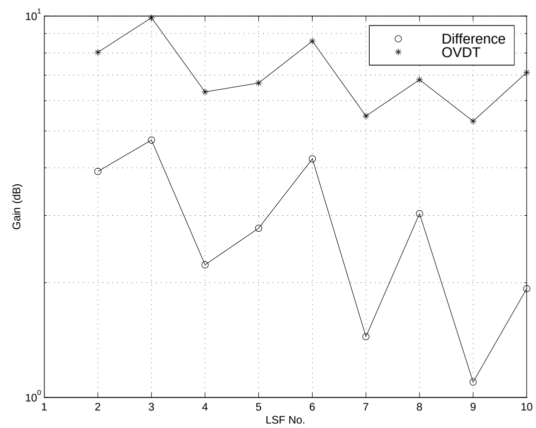

Figure 2: The Gain achieved using (1) Differential Encoding (2) Sequential Vector Decorrelation Technique

The gains achieved by SVDT shown in figure 2 are the *maximum* achievable by any linear function of the LSF parameters. Figure 1 demonstrates the performance of 3 schemes for quantization of LSF parameters. The first scheme is plain scalar quantization of LSF parameters. The second scheme is a closed-loop DPCM scheme using scalar quantization. The third one is scalar quantization using the SVDT. It is seen that the scalar quantization employing SVDT achieves the best performance among others and this comply with what we expected based on the gains in figure 2. The bit allocation

Performance	SD [dB]			2dB OL [%]		
Bit-rate	26	28	30	26	28	30
Scalar	1.75	1.59	1.41	28.83	18.51	9.39
DPCM	1.31	1.21	1.03	6.20	4.11	1.88
SVDT	1.28	1.10	0.99	5.71	2.82	1.55

Table 1: Average spectral distortion and 2 dB outliers percentage for (1) scalar quantization (2) DPCM and scalar quantization (3) SVDT and scalar quantization

between different LSF parameters in all the schemes are performed based on their variance. The higher the variance the more bits are allocated to them.

4 Conclusions

In this work, a technique for optimal decorrelation of random vector sources is described. The application of this technique to quantization of LSF parameters in CELP speech coding using scalar quantizers is studied. It is observed that the bit rate at which the transparent quantization is achieved is substantially reduced.

References

[1] A. Gresho and R. M. Gray, *Vector Quantization and Signal Compression,* Boston: Kluwer Academic Publishers, 1992.

[2] N. S. Jayant and P. Noll, *Digital Coding of Waveforms,* Bell Telephone Laboratories Inc, 1984.

[3] F. Lahouti, A. K. Khandani, "Intra-frame and Inter-frame Coding of Speech LSF Parameters Using A Trellis Structure", Technical report ECE-99-12, Dept. of E&CE, University of Waterloo, April 1999.

[4] F. Lahouti, A. K. Khandani, "Quantization of line spectral parameters using a trellis structure," *Int. Conf. on Acoustics, Speech and Signal Processing,* Istanbul, Turkey, June 2000.

[5] F. Lahouti, A. K. Khandani, "Quantization of LSF Parameters Using Trellis Modeling," Technical report ECE-2001-4, Dept. of E&CE, University of Waterloo, March 2001.

[6] A. Papoulis, *Probability, Random Variables and Stochastic Processes,* McGraw-Hill, 1991.

[7] W.B. Kleijn and K. K. Paliwal, Eds., *Speech coding and synthesis,* Elsevir Science B.V., 1995.