CAUSAL TRANSFORM CODING, GENERALIZED MIMO LINEAR PREDICTION, AND APPLICATION TO VECTORIAL DPCM CODING OF MULTICHANNEL AUDIO

David Mary, Dirk T. M. Slock

Département de Communications Mobiles
Institut Eurécom
2229 Route des Crêtes, 06904 Sophia Antipolis, FRANCE
{David.Mary,Dirk.Slock}@eurecom.fr

ABSTRACT

The optimal causal (prediction based) decorrelating scheme is applied to the frameworks of transform coding, coding of vectorial signals (multichannel audio), and vectorial DPCM coding. We analyze the effects of backward adaptation uppon the prediction operations and compare the expressions of the coding gains under infinite and high resolution assumption. We generalize the MIMO (Multiple Input/Multiple Output) prediction by organizing differently the samples in the vectorial signals, which corresponds to different degrees of non-causality of the intersignals predictors. An extreme case is the triangular MIMO prediction, for which ”causality” becomes processing the channel in a certain order. The high resolution coding gain suggests an optimal strategy in the choice of the interband predictors. For two-dimensionnal vectorial sources (such as stereo signals) we show the superiority in terms of coding gain of the triangular MIMO predictor over the classical MIMO prediction. A theorem is established which concerns the optimal ordering of the signals for the triangular MIMO predictor. When finite prediction orders are used to perform the intersignal decorrelation, we show that the optimal positioning of a finite number of taps is fairly straightforward. [1]

1. INTRODUCTION

In a recent work, a new coding technique has been introduced ([1, 4]). A natural and usefull application was found in the coding of multichannel audio signals. In this scheme, an optimal causal transform is applied to the data before the quantization stage. This scheme is optimal in the sense that it totally decorrelates the data. The performance of this transform is evaluated in terms of coding gain and described in the following frameworks.

In the transform coding case, the optimal causal transform is a lower triangular and unit diagonal matrix, which corresponds to a (Lower-Diagonal-Upper) factorization of the autocorrelation matrix of the signal. The rows of this matrix are optimal prediction filters for the corresponding component of the vector to be coded. The transformed coefficients are optimal prediction errors. The optimal causal transform is shown to have the same coding gain as the best unitary transform, the Karhunen-Loeve Transform. Such a causal transform coding scheme was independently described in [5]. However, as in classical ADPCM, the transformation may be backward adapted, that is computed on the basis of the previously quantized samples. And as in ADPCM, we show that a quantization noise feedback occurs. Under high resolution assumption (white independent quantization noise), close form expressions for the coding gains are presented. The optimal causal transform coding under infinite and high resolution assumptions is described in the second section.

In the third section, we apply the optimal causal transform to the coding of vectorial signals (for example subband, stereo or multichannel audio signals). By considering vectors of infinite size, one can get frequential expressions for the coding gains. In this case, the optimal causal decorrelating scheme can be described by means of a prediction matrix whose entries are optimal prediction filters. The diagonal filters are scalar intrasignal prediction filters. The off-diagonal predictors are Wiener filters performing the intersignal decorrelation. We show in this paper that this decorrelating procedure leads to the notion of generalized MIMO prediction, in which a certain degree of non causality may be allowed for the off-diagonal prediction filters. In the case of non causal intersignal filters, the optimal MIMO predictor is still lower triangular, and hence ”causal”, in a wider sense. The notion of causality is generalized : the causality between channels becomes processing the channels in a certain *order*. Some signals may be coded using the coded/decoded versions of the ”previous” signals. This scheme represents a generalization to the vectorial case of the classical (scalar) ADPCM coding technique.

An interesting result in [1] is that if the quantization noise feedback is taken into account, the efficiency of the interband decorrelation depends on the order in which the decorrelation between the signals is processed. We present in the fourth section of this paper a new theorem concerning the optimal ordering of the signals for a triangular ”causal” MIMO predictor, namely the ordering which minimizes the quantization noise feedback.

The last part of this paper deals with optimal triangular MIMO prediction with finite prediction orders. Despite the non causality in the classical sense of this approach, the optimal triangular MIMO prediction is well suited for frame based audio coding, which allows a certain degree of non causality in the coding procedure. When FIR filters are used to perform the intersignal decorrelation, we show that the optimal positioning of a finite number of taps is fairly straightforward.

[1] Eurécom’s research is partially supported by its industrial partners: Ascom, Swisscom, France Télécom, La Fondation CEGETEL, Bouygues Télécom, Thales, ST Microelectronics, Motorola, Hitachi Europe and Texas Instruments.

2. OPTIMAL CAUSAL TRANSFORM CODING

Let us consider the generalization of the classical DPCM coding scheme applied to a vector $X = [x_1 ... x_N]^T$, see Figure 1. A

Figure 1: Vectorial DPCM coding scheme.

matricial transformation L is applied to the vector $X : Y = LX = X - \overline{L}X$, where $\overline{L}X$ is the reference vector. The difference vector $Y = [y_1 ... y_N]^T$ is then quantized using a set $\mathbf{Q}$ of quantizers Q_i. The output X^q is $Y^q + \overline{L}X$. Note that the quantization error $\tilde{X}$ equals the reconstruction error $\tilde{Y}$:

$$\tilde{X} = X - X^q = X - (Y^q + \overline{L}X) = X - \overline{L}X - Y^q = Y - Y^q = \tilde{Y}, \quad (1)$$

as in the unitary case. The constraint imposed on the transformation is here the causality, which imposes a lower triangular structure. The unitary aspect of the transform appears in the unicity of the main diagonal ($\overline{L} = I - L$ is hence strictly lower triangular and represents the degrees of freedom of the transformation). The notion of causality could be generalized by working with the permuted components of X and Y, which gives $\mathcal{P}Y = L\ \mathcal{P}X$ or $Y = (\mathcal{P}^T L \mathcal{P})X$, where $\mathcal{P}$ is a permutation matrix. On one hand, the coding gain for a transformation L is

$$G_{TC}(L) = \frac{E\|\tilde{X}\|^2_{(I)}}{E\|\tilde{X}\|^2_{(L)}} = \frac{E\|\tilde{X}\|^2_{(I)}}{E\|\tilde{Y}\|^2_{(L)}}, \quad (2)$$

where I is the identity matrix (which corresponds to the absence of transformation), and the notation $\|\tilde{X}\|^2_{(T)}$ denotes the variance of the quantization error on the vectorial signal X, obtained for a transformation T. The second equality in (2) follows from the equality (1), as in the unitary case. On the other hand, the SNR is defined for a transformation L as

$$SNR(L) = \frac{E\|X\|^2}{E\|\tilde{X}\|^2_{(L)}} = \frac{E\|X\|^2}{E\|\tilde{Y}\|^2_{(L)}} = \frac{E\|X\|^2}{E\|Y\|^2_{(L)}} \frac{E\|Y\|^2_{(L)}}{E\|\tilde{Y}\|^2_{(L)}} \quad (3)$$

where the first factor represents the gain of the transformation. We now should determine the optimal transformation L and bit assignment which maximizes the coding gain. For a given bit assignment, we should find

$$L = \arg\max_L G_{CT}(L) = \arg\max_L SNR(L) = \arg\min_L E\|\tilde{X}\|^2_{(L)} \quad (4)$$

2.1. Ideal case

In a first step, we neglect the quantization error on the reference signal, and we suppose an optimal bit assignment. A quantizer Q_i introduces a white noise $\tilde{y}_i$ on the component y_i, of variance $\sigma^2_{\tilde{y}_i} = c\, 2^{-2R_i} \sigma^2_{y_i}$, where R_i is the number of bits assigned to the quantizer Q_i, and c is a constant depending on the probability density function of the signal to be quantized (one should assume a Gaussian distribution, linear transform invariant).
For a given L, the optimal bit assignment has to minimize $E\|\tilde{Y}\|^2_{(L)} = \sum_{i=1}^N \sigma^2_{y_i} c 2^{-2R_i}$ under the constraint $\sum_{i=1}^N R_i = NR$, where R is the average number of bits assigned to the N quantizers Q_i. Using well-known techniques [6], and abstracting the fact that the R_i are integer and non negative, one shows that

$$\sigma^2_{\tilde{y}_i} = c\, 2^{-2R_i} \sigma^2_{y_i} = c\, 2^{-2R} \left(\prod_{i=1}^N \sigma^2_{y_i} \right)^{\frac{1}{N}} \quad (5)$$

Note here that the optimal quantization errors variances $\sigma^2_{\tilde{y}_i}$ are equal (independentof i).
Optimization of L : we should consider $\min_L (\Pi_{i=1}^N \sigma^2_{y_i})^{\frac{1}{N}}$, where the $\sigma^2_{y_i}$ depend on the rows L_i of L : $\sigma^2_{y_i} = \sigma^2_{y_i}(L_{i.})$. The problem is hence separable, and minimizing $\left(\prod_{i=1}^N \sigma^2_{y_i}\right)^{\frac{1}{N}}$ with respect to L is the same as minimizing $\sigma^2_{y_i}$ with respect to $L_{i,1:i-1}$. The components y_i appear clearly as the prediction errors of x_i with respect to the past values of X, the $X_{1:i-1}$, and the optimal coefficients $-L_{i,1:i-1}$ are the optimal prediction coefficients. In other words, L is such that

$$LR_{XX}L^T = R_{YY} = D = diag\{\sigma^2_{y_1}, ... \sigma^2_{y_N}\}, \quad (6)$$

where $diag\{...\}$ represents a diagonal matrix whose elements are $\sigma^2_{y_i}$. Since each prediction error y_i is orthogonal to the subspaces generated by the $X_{1:i-1}$, the y_i are orthogonal, and D is diagonal. It follows that

$$R_{XX} = L^{-1} R_{YY} L^{-T}, \quad (7)$$

which represents the factorisation LDU of R_{XX}.

$$L = \begin{bmatrix} 1 & & & \\ \star & \ddots & \mathbf{0} & \\ \vdots & \ddots & \ddots & \\ \star & \cdots & \star & 1 \end{bmatrix} = I - \overline{L}$$

where the $\star$ represent the prediction coefficients. Refering to (2), the coding gain can be written as

$$G^{(0)}_{TC}(L) = \left(\frac{\det\left[diag(R_{XX})\right]}{\det\left[diag(LR_{XX}L^T)\right]} \right)^{\frac{1}{N}} \quad (8)$$

where $diag(R)$ denotes here the diagonal matrix which corresponds to the diagonal of the matrix R.

2.2. Quantization effects on the coding gain

Let us now inspect the case where the tranformation is not based on the original signal but on its quantized version. In this case, the output vector becomes

$$Y = X - \overline{L}X^q = X - \overline{L}(X - \tilde{X}) = LX + \overline{L}\tilde{Y}. \quad (9)$$

Y does now not only represent the prediction error LX of X, but also the quantization error $\tilde{Y}$ filtered by the optimal predictor $\overline{L}$. In this case again, the optimal bit assignment has to minimize the sum of the $\sigma^2_{\tilde{y}_i}$. It follows that the variances of the quantization noises are $\sigma^2_{\tilde{y}_i} = c 2^{-2R} (\prod_{i=1}^N \sigma^2_{y_i})^{\frac{1}{N}} = \sigma^2_{\tilde{y}_1}$, independentof i. The autocorrelation matrix of the noise is hence $R_{\tilde{Y}\tilde{Y}} = \sigma^2_{\tilde{y}_1} I$.
To optimize L, one should consider $\min_L \left(\det\left[diag(R_{YY})\right]\right)$, with this time $R_{YY} = LR_{XX}L^T + \sigma^2_{\tilde{y}_1} \overline{L}\overline{L}^T$. One can show that the resolution of the normal equations leads to the following expression of the coding gain $G^{(1)}_{CT}(L)$, taking into account the first order of the perturbations

$$G_{TC}^{(1)}(L) \approx \left(\frac{\det\left[diag(R_{XX})\right]}{\det\left[diag(LR_{XX}L^T + \sigma^2_{\tilde{y}_1}\overline{LL}^T)\right]} \right)^{\frac{1}{N}} \quad (10)$$

with $LR_{XX}L^T = D$ and $\sigma^2_{\tilde{y}_1} = c\,2^{-2R}(\det D)^{\frac{1}{N}}$ where D is the diagonal matrix of the non perturbated prediction error variances, and L and $\overline{L}$ are also non perturbated quantities. This expression is established under the high resolution assumption ($\sigma^2_{\tilde{y}_1}I$ is small in comparison with R_{XX}). On can further show that the coding gain may also be written as

$$G_{TC}^{(1)}(L) \approx G_{TC}^{(0)}(L)\left(1 - \frac{\sigma^2_{\tilde{y}_1}}{N}\left(\Sigma_{i=1}^N \frac{1}{\lambda_i} - \Sigma_{i=1}^N \frac{1}{\sigma^2_{y_i}}\right)\right) \quad (11)$$

where $\{\lambda_i\}$ denote the eigenvalues of R_{XX}.

3. GENERALIZED MIMO PREDICTION

3.1. Ideal case

Let us now consider the case where X is made of a succession of samples of a vectorial signal $\underline{x}_k = [x_{1,k} \cdots x_{M,k}]^T$. Some particular cases of scalar signals x_i are subband signals, stereo or multichannel audio signals. $X_k = [\underline{x}_0^T\ \underline{x}_1^T \cdots \underline{x}_k^T]^T$, and one can also write $Y_k = [\underline{y}_0^T\ \underline{y}_1^T \cdots \underline{y}_k^T]^T$ with $\underline{y}_k = [y_{1,k} \cdots y_{M,k}]^T$. For these vectorial signals , it is interesting to consider the limit case where the dimension k goes to infinity, and where the signal $\underline{x}_k$ is stationary. In this case, the optimal transform L will lead to a signal $\underline{y}_k$, asymptotically stationary too, since L will become block Toeplitz (with blocks of size $M \times M$). We obtain in this case

$$G_{TC}^{(0)}(L) = \lim_{k\to\infty}\left(\frac{\det\left[diag(R_{X_kX_k})\right]}{\det\left[diag(LR_{X_kX_k}L^T)\right]}\right)^{\frac{1}{Mk}} \quad (12)$$

$$= \left(\frac{\det\left[diag(R_{\underline{x}_k\underline{x}_k})\right]}{\det[diag(R_{\underline{y}_k\underline{y}_k})]}\right)^{\frac{1}{M}} = \left(\frac{\prod_{i=1}^M \sigma^2_{x_i}}{\prod_{i=1}^M \sigma^2_{y_i}}\right)^{\frac{1}{M}} \quad (13)$$

where $y_{i,k}$ is the optimal prediction error of infinite order of $x_{i,k}$, based on $\{\underline{x}_{-\infty:k-1}, x_{1:i-1,k}\}$. One will continue to denote by L_i (now of infinite dimension) the vector of the corresponding prediction coefficients.

There exists a frequential expression for $\prod_{i=1}^M \sigma^2_{y_i}$. Writing the prediction operation in the frequency domain, and using the fact that y_k is a totally decorrelated signal (its power spectral density can be written as $S_{\underline{yy}}(f) = R_{\underline{yy}} = diag\{\sigma^2_{y_1}, \ldots, \sigma^2_{y_M}\}$), one can show that

$$\prod_{i=1}^M \sigma^2_{y_i} = e^{\int_{-\frac{1}{2}}^{\frac{1}{2}} \ln[\det(S_{\underline{xx}}(f))]\,df} \quad (14)$$

The prediction operation can also be described by optimal MIMO predictor $L(z)$ with filters of infinite length. Let us consider two extreme cases of different $L(z)$, for $M = 2$. In the classical MIMO Prediction [2] the predictor $L(z)$ is

$$L(z) = \begin{bmatrix} L_{11}(z) & L_{12}(z) \\ L_{21}(z) & L_{22}(z) \end{bmatrix} \ \ with\ \ L_0 = \begin{bmatrix} 1 & 0 \\ l_{21} & 1 \end{bmatrix},$$

in order to keep the structure (temporally) causal. Another causal (in a wider sense) transform [3] is

$$T(z) = \begin{bmatrix} 1 & 0 \\ 0 & T_{22}(z) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ W_{21}(z) & 1 \end{bmatrix} \begin{bmatrix} T_{11}(z) & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} T_{11}(z) & 0 \\ T_{22}(z)W_{21}(z)T_{11}(z) & T_{22}(z) \end{bmatrix} \quad (15)$$

where $T_{ii}(z)$ are predictors for scalar signals, and $W_{21}(z)$ is a Wiener filter (non constrained to be causal) estimating $x_{2,k}$ from the whitened version $y_{1,k}$ of $x_{1,k}$. The loss in degrees of freedom occuring with the loss of one interband predictor (L_{12} in the classical MIMO transformation) is balanced by the non causality of the remaining unique interband predictor W_{21}. Under infinite resolution, the product of the variances of the subband signal is constant, no matter which causal transform we use. The coding gain $G_{TC}^{(0)}$ is invariant by permutation [4]. Each permutation leads to another causal decorrelation of the components of one vector. For a stationary signal, this means that there exists more that one way to decorrelate the scalar signals which compose this signal.

Hence, a generalized MIMO prediction in the case of M scalar signals can be defined as a classical MIMO prediction on $\underline{x}_k^{'} = [x_{1,k}\ x_{2,k+d_1} \ldots x_{M,k+d1+\ldots+d_{M-1}}]^T$, where d_i are delays representing the degree of non-causality of the interband predictors. The two previous examples present (for $M = 2$) two extreme cases on an infinity of variantes, which are parametrized by the degrees of non causality (in the classical sense) of the interband predictors.

The triangular "causal" MIMO predictor is an extreme case of $d_i \to \infty,\ i = 1, \ldots, M-1$. The "causality" of the triangular predictor matrix becomes processing the channel in a certain *order*, that is, the signals get decorrelated in this order. Note that the notion of causality remains unchanged for the diagonal predictors.

3.2. Quantization effects on the coding gain

If we now consider the effects of the quantization in the closed loop, the gain $G_{CT}^{(1)}(L)$ can be expressed as

$$G_{TC}^{(1)}(L) \approx \lim_{k\to\infty}\left(\frac{\det[diag(R_{X_kX_k})]}{\det[diag(LR_{X_kX_k}L^T + \sigma^2_{\tilde{y}_1}\overline{LL}^T)]}\right)^{\frac{1}{Mk}} \quad (16)$$

which leads to

$$G_{TC}^{(1)}(L) \approx G_{TC}^{(0)}(L)\left(1 - \sigma^2_{\tilde{y}_1}\frac{1}{M}\sum_{i=1}^M \frac{\|L_i\|^2 - 1}{\sigma^2_{y_i}}\right). \quad (17)$$

As in the ideal case, one can derive a frequential expression for $G_{CT}^{(1)}(L)$

$$G_{TC}^{(1)} \approx G_{TC}^{(0)}\left[1 + \frac{\sigma^2_{\tilde{y}_1}}{M}\left(-\int_{-\frac{1}{2}}^{\frac{1}{2}} tr\left(S_{\underline{xx}}^{-1}(f)\right)df + \sum_{i=1}^M \frac{1}{\sigma^2_{y_i}}\right)\right] \quad (18)$$

where, comparing with equation (17), the term $\int_{-\frac{1}{2}}^{\frac{1}{2}} tr\left(S_{\underline{xx}}^{-1}(f)\right)df$ corresponds to $\sum_{i=1}^M \frac{\|L_i\|^2}{\sigma^2_{y_i}}$. For filters of infinite length, the backward adapted triangular "causal" MIMO predictor has a direct application to ADPCM coding : the intersignal prediction of the "next" signal can be based on the coded/decoded version of the "previous" signal.

4. OPTIMAL ORDERING OF THE SUBSIGNALS FOR VECTORIAL DPCM WITH TRIANGULAR MIMO PREDICTION

Comparing $G_{TC}^{(1)}$ above with the infinite resolution case (where the coding gain is independent of the delays d_i), the different variances produced by the different decorrelation approaches induce now different sums. Hence, the coding gain $G_{TC}^{(1)}$ depends on a carefull choice of the decorrelation procedure. In the case $M = 2$, maximizing the coding entails making the variances as different as possible. Thus, the subsignal of greater variance should be processed first, and all the degrees of freedom of the interband decorrelator should be used to decrease the variance of the subsignal of lower variance. The triangular MIMO predictor is in this case superior to the classical MIMO predictor, since W_{12} defined above is the most efficient interband predictor.
For $M > 2$, the following theorem holds :

Theorem : Optimal ordering of the subsignals for triangular MIMO prediction . The optimal ordering of the subsignals in a stationary vectorial signal for maximizing the high-resolution coding gain $G_{TC}^{(1)}$ of vectorial DPCM with triangular MIMO prediction is obtained by processing the signals in order of decreasing variance.

To show the theorem, consider a recursive argument. First of all, the theorem is clearly true for the case of two channels. Now consider $n-1$ channels that we have ordered in order of decreasing variance. When we add a nth channel, the question is in which position it should be put w.r.t. the other channels. Assume in a first scenario that we put the channel in a position such that all n channels are in order of decreasing variance. Assume in a second scenario that we insert the nth channel at another position. Then we can evolve from the first to the second scenario by a sequence of permutations of two consecutive channels. In one such permutation operation, assume that the channels involved in the permutation are in positions i and $i+1$. Then the channels $1, \ldots, i-1$ are unaffected in the triangular MIMO prediction approach. The channels $i+2, \ldots, n$ are also unaffected by the order in which channels i and $i+1$ are put since in any case they get orthogonalized w.r.t. the signals in those channels. So the only effect of the permutation between channels i and $i+1$ is on the prediction error variances of those channels i and $i+1$. In other words we are reduced to the two channel case, in which case we know that we should put the channels in order of decreasing variance. So, as we move from scenario one to scenario two by a succession of permutations of two consecutive channels, we decrease the coding gain in each permutation. Hence, the optimal ordering is in order of decreasing variance.

5. OPTIMAL TRIANGULAR MIMO PREDICTION WITH FINITE PREDICTION ORDERS

So far we have assumed that all filters involved are of infinite length. In the classical MIMO linear prediction, a finite number of prediction coefficients is typically used in a way that is a straightforward extension from the scalar case. Namely, the MIMO prediction order is limited to a finite order, resulting in a desired number of prediction coefficients (from the point of view of complexity or performance or both). In the triangular predictor case, it is more straightforward to assign a finite number of coefficients in an optimal fashion. The diagonal terms in the MIMO prediction filter correspond to classical scalar predictors, so the number of coefficients assigned will simply determine the prediction order as usual. However, for the non-causal off-diagonal terms, the filters are Wiener filters of unconstrained structure, except that we wish to use a finite number of taps. The problem then becomes the optimal positioning of those taps. We shall assume that the diagonal scalar predictors are of sufficient order for the whitened versions of the signals to be considered as effectively white. In that case, the design of the off-diagonal terms in a row of the MIMO prediction filter corresponds to an issue of estimating a signal x on the basis of uncorrelated variables y_i. Due to the uncorrelatedness of the y_i, the estimation in terms of the y_i decouples and the contribution of each y_i can be considered separately. In particular, the variance of the estimation error becomes

$$R_{\tilde{x}\tilde{x}} = R_{xx} - \sum_i \frac{(R_{xy_i})^2}{R_{y_iy_i}} \qquad (19)$$

where $R_{xy} = \mathrm{E}\, X\, y$ is the correlation. So, those variables y_i should be used for which the ratio $\frac{(R_{xy_i})^2}{R_{y_iy_i}}$ is the largest. Within a subset of the y_i that are samples of a certain whitened signal, $R_{y_iy_i}$ is independent of i due to stationarity and hence it suffices to use those samples y_i for which $|R_{xy_i}|$ is largest. So the optimal positioning of a finite number of taps in the off-diagonal filters is fairly straightforward.

6. REFERENCES

[1] D. Mary and Dirk T.M. Slock, “Vectorial DPCM Coding and Application to Widebanb Speech Coding,” in ICASSP 2001, Salt Lake City, USA, May 2001.

[2] B. Maison and L. Vanderdorpe, “About the Asymptotic performance of Multiple-Input/Multiple-Output Linear Prediction of Subband Signals,” IEEE Signal Proc. Letters, vol. 5, no. 12, December 1998.

[3] P. W. Wong, “Rate Distortion Efficiency of Subband Coding with Crossband Prediction,” IEEE Trans. on Inf. Theory, Januar 1997.

[4] D. Mary and Dirk T.M. Slock, “Codage DPCM Vectoriel et Application au Codage de la Parole en Bande Élargie,” in CORESA 2000, Poitiers, France, October 2000.

[5] Y.-P. Lin and S. M. Phoong, “Prediction-based Lower Triangular Transform,” IEEE trans. on Sig. Proc., vol. 48, no. 7, July 2000.

[6] N.S. Jayant and P. Noll, Digital Coding of Waveforms, Signal Processing. Prentice Hall, 1984.