The LOT: Transform Coding Without Blocking Effects

HENRIQUE S. MALVAR, MEMBER, IEEE, AND DAVID H. STAELIN, FELLOW, IEEE

***Abstract*—The Lapped Orthogonal Transform (LOT) is a new tool for block transform coding with basis functions that overlap adjacent blocks [1]–[3]. The LOT can reduce the blocking effect to very low levels. In this paper, an exact derivation of an optimal LOT is presented. The optimal LOT is related to the discrete cosine transform (DCT) in such a way that a fast algorithm for a nearly optimal LOT is derived. Compared to the DCT, the fast LOT requires about 20–30 percent more computations, mostly additions. An image coding example demonstrates the effectiveness of the LOT in reducing blocking effects. Unlike earlier approaches to the reduction of blocking effects, the LOT actually leads to slightly smaller signal reconstruction errors than does the DCT.**

I. Introduction

ONE of the most efficient methods of data compression of images and other random signals is transform coding [4], [5]. The incoming signal is generally subdivided into blocks of N samples; each block is projected into a particular basis by means of an orthogonal transform, and the coefficients of the transformation are quantized and transmitted. At the receiver, the coefficients are reconstructed and the inverse transformation applied, so that an approximation of the original block is obtained.

The effects of the quantization error are minimized when the transform basis functions are the set of eigenvectors of the autocovariance matrix of the input signal, these vectors define the Karhunen–Loève transform (KLT) [4]–[6]. The KLT packs most of the signal variance, or energy, into the minimum number of coefficients for any desired error level, and thus it leads theoretically to the minimum bit rate [5]. In practice, the discrete cosine transform (DCT) is preferred over the KLT, since the DCT is signal independent, it is a good approximation to the KLT for a large class of signals with low-pass spectra, and can be computed by means of fast algorithms [4], [7].

DCT-based transform coding has been a popular method of image and speech compression [5], and most of its recent advances concentrate on adaptive quantization strategies [8], [9]. One of the basic problems of transform coding at low bit rates, which has not been efficiently solved yet, is the so-called "blocking effect." The blocking effect is a natural consequence of the independent processing of each block. It is perceived in images as visible discontinuities in features the cross block boundaries [10] (in the interframe image coding with motion-compensated frame prediction, blocking effects are not so disturbing, but are still noticeable [11]). In transform coding of speech, blocking effects are perceived as extraneous tones [5].

Some methods for the reduction of blocking effects have been previously suggested [10], [12], [13]. In [10], two methods were proposed: *overlapping*, also discussed in [12], and *filtering*. In the overlapping method, the blocks overlap slightly, so that redundant information is transmitted for the samples in the block boundaries. The receiver averages the reconstructed samples from the neighboring blocks, in the overlapping areas. The disadvantage of this approach is the increase in the total number of samples to be processed, and thus an increase in the bit rate.

In the filtering method, the coding process at the transmitter is unchanged, and at the receiver a low-pass filter is applied only to the boundary pixels. Although this method does not increase the bit rate, it blurs the signal across block boundaries. In [13], the filtering method avoids blurring by incorporating a prefilter at the transmitter. In [14], the short-space Fourier transform (SSFT) is used instead of the DCT. Although the SSFT is intrinsically free from blocking effects, because the SSFT of a block depends on the whole signal, it introduces ringing around edges.

A new class of transforms for blocking signal coding, introduced in [1] and [2], has the same benefits of the overlapping method cited above, but *without* an increase in the bit rate. These new transforms (collectively referred to as the "lapped orthogonal transform," or LOT, after [2]) are characterized by the fact that each block of size N is mapped into a set of N basis functions, each one being longer than N samples. In this paper, we present the LOT definition and show that its basis functions can be derived as the solution to an eigenvalue problem, instead of by means of the possibly noisy iterative numerical procedure in [1] and [2]. We review the basic properties of LOT's in Section II, and an optimal LOT is derived in Section III. A fast-computable approximation to the LOT

Manuscript received July 23, 1987; revised June 27, 1988. This work was supported in part by the Center for Advanced Television Studies. The work of H. S. Malvar was supported in part by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (Brazil) under Grant 200.832-82.

H. S. Malvar was with PictureTel Corporation, One Intercontinental Way, Peabody, MA 01960. He is now with the Department of Electrical Engineering, Universidade de Brasília, 70910 Brasília, Brazil.

D. H. Staelin is with the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139.

IEEE Log Number 8826101.

0096-3518/89/0400-0553$01.00 © 1989 IEEE

is presented in Section IV. The LOT efficiency in reducing the blocking effects is verified in Section V.

II. Basic Properties of the LOT

In this section, we review those properties of the LOT [1], [2] necessary for its analytical derivation in Section III. We assume that the signals to be processed are unidimensional: extension to two or more dimensions is easily achieved by defining separable transforms based on the unidimensional profile; this is a standard approach to multidimensional transform coding [4], [5].

Let us assume that the incoming discrete-time signal is a large segment of MN samples, where N is the block size. In traditional transform coding M blocks of length N would be independently transformed and coded. In matrix notation, if we call $\boldsymbol{x}_o$ the original input vector of length MN, the vector $\boldsymbol{y}_o$ containing the transform coefficients of all blocks is given by

$$\boldsymbol{y}_o = \boldsymbol{T}'\boldsymbol{x}_o, \tag{1}$$

where $\boldsymbol{T}'$ is the transpose of an $MN \times MN$ block-diagonal matrix, in the form

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{D} & & & 0 \\ & \boldsymbol{D} & & \\ & & \ddots & \\ 0 & & & \boldsymbol{D} \end{bmatrix}, \tag{2}$$

where $\boldsymbol{D}$ is a matrix of order N, whose columns are the basis functions that define the transform of each block.

With the LOT, each block has L samples, with $L > N$, so that neighboring blocks overlap by $L - N$ samples. The basic operation of the LOT is thus similar to the overlapping method of [10]. A fundamental difference is that the LOT maps the L samples of each block into N transform coefficients. With the number of transform coefficients being equal to the block size there is no increase in the data rate. The LOT can be defined as in (1), with $\boldsymbol{T}$ given by

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{P}_1 & & & & 0 \\ & \boldsymbol{P}_0 & & & \\ & & \ddots & & \\ & & & \boldsymbol{P}_0 & \\ 0 & & & & \boldsymbol{P}_2 \end{bmatrix}, \tag{3}$$

where $\boldsymbol{P}_o$ is an $L \times N$ matrix that contains the LOT basis functions for each block. We have assumed $L \leq 2N$, i.e., the length of each basis function is at most twice the block size. This choice will be justified later. The matrices $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ are introduced because the first and last blocks of a segment have only one neighboring block, and thus the LOT for the first and last blocks must be defined in a slightly different way, to guarantee that none of the basis functions extends beyond the segment boundaries. We shall concentrate on $\boldsymbol{P}_0$ for now.

We note that the LOT of a single block is not invertible, since $\boldsymbol{P}_0$ is not square. Nevertheless, in terms of reconstructing the whole segment $\boldsymbol{x}_o$, all we need is invertibility of $\boldsymbol{T}$. Orthogonality of $\boldsymbol{T}$ is also a desirable property, as with all transforms in traditional transform coding, since it guarantees good numerical stability. In order for $\boldsymbol{T}$ to be orthogonal, the columns of $\boldsymbol{P}_0$ must be orthogonal,

$$\boldsymbol{P}_0'\boldsymbol{P}_0 = \boldsymbol{I}, \tag{4}$$

and the overlapping functions of neighboring blocks must also be orthogonal,

$$\boldsymbol{P}_0'\boldsymbol{W}\boldsymbol{P}_0 = \boldsymbol{P}_0'\boldsymbol{W}'\boldsymbol{P}_0 = \boldsymbol{0}, \tag{5}$$

where $\boldsymbol{I}$ is the identity matrix, and the shift operator $\boldsymbol{W}$ is defined by

$$\boldsymbol{W} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{I} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}. \tag{6}$$

The identity matrix above is of order $L - N$, and we have assumed $L \leq 2N$. As we will see later, a good choice for L is $L = 2N$. We will say that an LOT matrix $\boldsymbol{P}_0$ is *feasible* if it satisfies (4) and (5). The set of feasible LOT's is clearly a superset of nonoverlapping transforms.

Besides the required orthogonality conditions above, we should expect additional properties to hold for a good LOT matrix $\boldsymbol{P}_0$, based on our knowledge of the DCT and KLT. If a feasible LOT is to exhibit good energy concentration, its basis functions should have properties similar to those of the DCT and KLT functions. Two of these properties seem to be the most relevant.

First, we recall that the DCT is a good substitute for the KLT because the DCT functions approximate the eigenvectors of the autocorrelation matrix $\boldsymbol{R}_{xx}$ of a first-order Gauss–Markov process [5], [15]

$$\boldsymbol{R}_{xx} = \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^L \\ \rho & 1 & \rho & \cdots & \rho^{L-1} \\ \vdots & & \vdots & & \vdots \\ \rho^{L-1} & & \rho & 1 & \rho \\ \rho^L & \cdots & \rho^2 & \rho & 1 \end{bmatrix}, \tag{7}$$

where ρ is the intersample correlation coefficient. Since the above matrix is symmetric and Toeplitz, its eigenvectors (which define the KLT) are either symmetric or antisymmetric [16], [17], i.e.,

$$\boldsymbol{R}_{xx}\boldsymbol{y} = \lambda\boldsymbol{y} \Rightarrow \boldsymbol{J}\boldsymbol{y} = \boldsymbol{y} \text{ or } \boldsymbol{J}\boldsymbol{y} = -\boldsymbol{y}, \tag{8}$$

where $\boldsymbol{J}$ is the "counter-identity"

$$\boldsymbol{J} = \begin{bmatrix} 0 & \cdots & & 0 & 1 \\ 0 & \cdots & 0 & 1 & 0 \\ \vdots & & & & \vdots \\ 1 & 0 & & \cdots & 0 \end{bmatrix}. \tag{9}$$

It turns out that half of the eigenvectors of $\boldsymbol{R}_{xx}$ are symmetric, i.e., $\boldsymbol{J}\boldsymbol{y} = \boldsymbol{y}$, and the other half are antisymmetric,

$Jy = -y$ [16]. It is reasonable to expect that the LOT should also have this kind of symmetry, i.e., it should be formed by $N/2$ symmetric (or even) vectors and $N/2$ antisymmetric (or odd) vectors. The DCT functions have this even-odd symmetry.

Second, it is reasonable to assume that at least the low-order vectors (responsible for most of the energy concentration) should be slowly varying sequences, e.g., sampled sinusoids with low frequencies. The eigenvectors of R_{xx} in (7) are exactly sampled sinusoids [4], for any value of ρ. The DCT basis functions are also sampled sinusoids.

Based on the discussion above, we will assume that half of the basis functions that compose a good LOT matrix P_0 are even, and half are odd. Furthermore, we should look for a set of feasible functions where the lower order functions are as smooth as the DCT or KLT functions. It is interesting to note that the orthogonality of shifted functions in (5) automatically forces the samples of the low-order basis functions to decay toward zero at their ends (otherwise, a zero inner product would not be attained) [3]. This is the key to reducing blocking effects.

III. An Optimal LOT

An optimal LOT should minimize the bit rate for any reconstruction error level. Assuming that the Markov model of (7) is applicable, this is equivalent to maximizing the "energy compaction" measure (also called maximum transform coding gain [5])

$$G_{TC} = \frac{\frac{1}{N}\sum_{i=1}^{N} \sigma_i^2}{\left(\prod_{i=1}^{N} \sigma_i^2\right)^{1/N}}, \tag{10}$$

where σ_i^2 is the ith diagonal entry of the matrix

$$R_0 = P_0' R_{xx} P_0. \tag{11}$$

Cassereau [1], [2] obtained optimal LOT's by means of an iterative optimization technique that searches for a maximum of G_{TC}. At each step, a new basis function (i.e., a column of P_0) is obtained. This approach has the disadvantage of being highly sensitive to numerical errors, even with double-precision computations. Also, the optimal P_0 may not be easily factorable so that a fast algorithm may not exist.

We present here a direct approach [3] for the derivation of an optimal LOT when $L = 2N$, i.e., the basis functions of neighboring blocks overlap by N samples. Our approach is virtually insensitive to numerical errors, and it also leads to a better understanding of the LOT, so that a fast algorithm can also be derived. The key point is to start with a feasible LOT matrix P that is not necessarily optimal. Then, the matrix

$$P_0 = PZ \tag{12}$$

is also a feasible LOT for any orthogonal Z, since

$$P_0' P_0 = Z' P' PZ = Z' Z = I, \tag{13}$$

$$P_0' W P_0 = Z' P' WPZ = 0. \tag{14}$$

We can define a feasible LOT from the DCT, by

$$P = \frac{1}{2}\begin{bmatrix} D_e - D_o & D_e - D_o \\ J(D_e - D_o) & -J(D_e - D_o) \end{bmatrix}, \tag{15}$$

where D_e and D_o are the $N \times N/2$ matrices containing the even and odd DCT functions, respectively [3]. It is easy to verify the feasibility of P above. This particular choice will be justified in Section IV.

With P as in (15), what we need to obtain an optimal LOT is to find an optimal Z in (12). Substituting (12) into (11), we obtain

$$R_0 = Z' P' R_{xx} PZ. \tag{16}$$

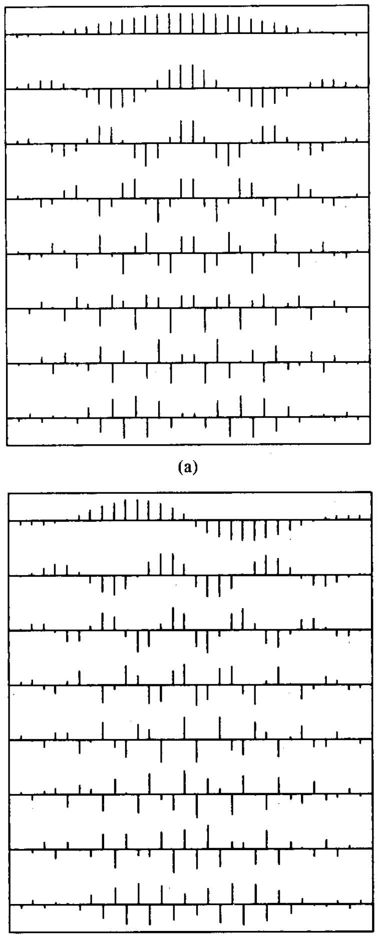
With P and R_{xx} fixed, it is clear that G_{TC} is maximized when R_0 is diagonal, i.e., when the columns of Z are the eigenvectors of $P' R_{xx} P$. With such a Z, the LOT matrix P_0 is optimal.

It is important to point out that our optimization approach leads to an optimal LOT that is tied to the choice of the initial matrix P. Since each column of P has L elements, with $L > N$, they span an N-dimensional subspace of $\mathbb{R}^L$. For any Z, the matrix PZ will always belong to that subspace, and so will the optimal LOT. However, there may exist a feasible LOT $\hat{P}$ that does not belong to the subspace spanned by the columns of P, i.e., it cannot be generated by (12).

Thus, an optimal LOT derived by the procedure above may not be the globally optimal LOT, in the sense of maximizing the energy compaction. However, as we will see later, our choice for P in (15) is good enough since we have obtained the same energy compaction as Cassereau's functions [1], [2], which are designed to be globally optimal (actually, we have obtained slightly higher G_{TC}'s; this is probably due to some error propagation in Cassereau's algorithm).

For a Markov model with $\rho = 0.95$, the columns of the optimal P_0 are shown in Fig. 1. The functions are not very sensitive to variations in ρ, so that the results for $\rho = 0.8$, for example, a virtually the same as those in Fig. 1. We note in Fig. 1 that the orthogonality constraints for the functions belonging to neighboring blocks led to basis functions that decay toward zero at their boundaries. The first basis function, for example, has a boundary value that is 5.83 times lower than its value at the center. So, the discontinuity from zero to the boundary value is much lower than that of the standard DCT functions, and this is one of the main reasons why blocking effects are reduced.

There are two basic properties of the LOT functions in Fig. 1 that are a direct consequence of the choice $L = 2N$. First, if the lowest order basis functions for a group of consecutive blocks are superimposed, the resultant sequence has a constant dc value. This is an important and desirable characteristic, since it implies that a flat field can be reproduced with only one transform coefficient per block. If L were smaller than $2N$, perfect dc reconstruction with only one coefficient would necessarily lead to a loss of smoothness in the central portion of the first basis



(a) (b)

Fig. 1. (a) An optimal LOT for $N = 16$, $L = 32$, and $\rho = 0.95$, even basis functions. (b) An optimal LOT for $N = 16$, $L = 32$, and $\rho = 0.95$, odd basis functions.

function. Second, the fact is that the right boundaries of the basis functions for block r are immediately adjacent to the left boundaries of the functions for block $r + 2$. If L were smaller than $2N$, there would be two different positions in block $r + 1$ where discontinuities from block r and block $r + 2$ might occur. It seems, therefore, that $L = 2N$ is a good choice for the length of the functions.

The factor $\boldsymbol{Z}$ of the optimal LOT matrix $\boldsymbol{P}_0 = \boldsymbol{PZ}$ may not be factorable in $N \log (N)$ butterfly stages. This is exactly the same deficiency of the optimal KLT for block coding without overlapping. In the next section, we discuss an approximation to the optimal LOT that can be implemented through a fast algorithm, just as the DCT is a fast-computable approximation to the KLT.

IV. Fast Computation of the LOT

The key to a fast algorithm for the LOT is the approximation of the matrix $\boldsymbol{Z}$ by a product of a few simple factors. Actually, this is the main reason why we have chosen the DCT basis functions in the definition of $\boldsymbol{P}$ in (15). Such a definition will allow us to obtain a useful expression for the matrix $\boldsymbol{P}' \boldsymbol{R}_{xx} \boldsymbol{P}$. In order to simplify notation, let us refer to the Gauss–Markov autocorrelation matrix in (7) as $\boldsymbol{R}(2N, \rho)$, where the first parameter represents the matrix order. We can relate $\boldsymbol{R}(2N, \rho)$ to $\boldsymbol{R}(N, \rho)$ by

$$\boldsymbol{R}(2N, \rho) = \begin{pmatrix} \boldsymbol{R}(N, \rho) & \boldsymbol{B} \\ \boldsymbol{B} & \boldsymbol{R}(N, \rho) \end{pmatrix}, \tag{17}$$

where $\boldsymbol{B} = \rho \boldsymbol{J} \boldsymbol{r} \boldsymbol{r}'$ and $\boldsymbol{r} = [1 \ \ \rho \, \rho^2 \cdots \rho^N]'$.

Combining (15) and (17), we obtain, after a few manipulations,

$$\boldsymbol{R}_0 = \begin{pmatrix} \boldsymbol{R}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{R}_2 \end{pmatrix}, \tag{18}$$

where the diagonal blocks $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ are given by

$$\boldsymbol{R}_1 = \boldsymbol{D}_e' \boldsymbol{R}(N, \rho) \boldsymbol{D}_e + \boldsymbol{D}_o' \boldsymbol{R}(N, \rho) \boldsymbol{D}_o + \rho \boldsymbol{D}_e' \boldsymbol{r} \boldsymbol{r}' \boldsymbol{D}_e + \rho \boldsymbol{D}_o' \boldsymbol{r} \boldsymbol{r}' \boldsymbol{D}_o, \tag{19}$$

and

$$\boldsymbol{R}_2 = \boldsymbol{D}_e' \boldsymbol{R}(N, \rho) \boldsymbol{D}_e + \boldsymbol{D}_o' \boldsymbol{R}(N, \rho) \boldsymbol{D}_o - \rho \boldsymbol{D}_e' \boldsymbol{r} \boldsymbol{r}' \boldsymbol{D} e - \rho \boldsymbol{D}_o' \boldsymbol{r} \boldsymbol{r}' \boldsymbol{D}_o. \tag{20}$$

If we let the correlation coefficient ρ approach unity, the matrices $\boldsymbol{D}_e$ and $\boldsymbol{D}_o$ will contain the asymptotic even and odd eigenvectors of $\boldsymbol{R}(N, \rho)$, respectively, since the DCT is the limit of the KLT as $\rho \to 1$ [15]. Thus, the terms $\boldsymbol{D}_e' \boldsymbol{R}(N, \rho) \boldsymbol{D}_e$ and $\boldsymbol{D}_o' \boldsymbol{R}(N, \rho) \boldsymbol{D}_o$ are asymptotically diagonal, with positive entries. Also, as $\rho \to 1$, the vector $\boldsymbol{r}$ will have all of its entries equal to one, i.e., it will be an even vector. Thus, the term $\boldsymbol{D}_o' \boldsymbol{r} \boldsymbol{r}' \boldsymbol{D}_o$ goes to zero. Furthermore, since the vector $[1 \ 1 \cdots 1]'$ is equal to $\sqrt{N}$ times the first column of $\boldsymbol{D}_e$, it follows that

$$\boldsymbol{D}_e' \boldsymbol{r} \boldsymbol{r}' \boldsymbol{D}_e \to \begin{pmatrix} N & 0 & 0 & \cdots & 0 \\ 0 & 0 & & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & & 0 \end{pmatrix}. \tag{21}$$

Thus, it is clear that $\boldsymbol{R}_1$ will asymptotically be a diagonal matrix with positive diagonal entries. The factor $\boldsymbol{R}_2$, however, may not have a dominant diagonal because the third term in (20) is subtracted from the others. Nevertheless, we can expect the following approximation to hold as ρ gets closer to one:

$$\boldsymbol{Z} \simeq \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \tilde{\boldsymbol{Z}} \end{bmatrix}, \tag{22}$$

where $\tilde{\boldsymbol{Z}}$ is of order $N/2$. Although $\boldsymbol{R}_2$ may not have a strongly dominant diagonal, we should expect some diagonal dominance, so that $\tilde{\boldsymbol{Z}}$ should not be far from the identity matrix. In fact, in [3] it is shown that $\tilde{\boldsymbol{Z}}$ can be closely approximated by a cascade of $N/2 - 1$ plane rotations, in the form

$$\tilde{Z} = T_1 T_2 \cdots T_{N/2-1}, \tag{23}$$

where each plane rotation is defined as

$$T_i = \begin{bmatrix} I & 0 & 0 \\ 0 & Y(\theta_i) & 0 \\ 0 & 0 & I \end{bmatrix}. \tag{24}$$

The matrix $Y(\theta_i)$ is a 2×2 butterfly

$$Y(\theta_i) = \begin{pmatrix} \cos\theta_i & \sin\theta_i \\ -\sin\theta_i & \cos\theta_i \end{pmatrix}, \tag{25}$$

where θ_i is the rotation angle, and the top left identity factor in (24) is of order $i - 1$. If we apply the transpose of each T_i to $\tilde{Z}$ in the reverse order of (24), we should obtain a close approximation to the identity matrix, and this is indeed the case [3]. For $N = 16$ and $\rho = 0.95$, an appropriate set of angles is $[\theta_1 \cdots \theta_7] = [0.42\ 0.53\ 0.53\ 0.5\ 0.44\ 0.35\ 0.23\ 0.11]$ [3]. With these angles, the energy compaction is $G_{TC} = 9.32$, which is close to the value $G_{TC} = 9.49$ corresponding to the exact solution. Thus, the loss in coding gain by using the approximation in (23) is only 0.08 dB. The energy compaction for a DCT of size 16 is $G_{TC} = 8.82$, so that the LOT leads to an improvement of 0.32 dB in the rms reconstruction error.

It is important to note that the approximation of $\tilde{Z}$ by a cascade of $N - 1$ butterflies is satisfactory for small N. When $N \geq 32$, the approximation may introduce small discontinuities in the low-order basis functions, which would lead to noticeable artifacts in the reconstructed signal. The problem of finding a good approximation to $\tilde{Z}$ for large N is still unsolved.

Our fast LOT is defined by P_0 in (12), with P given by (15) and Z by (22)–(25). The resulting P_0 can also be written as

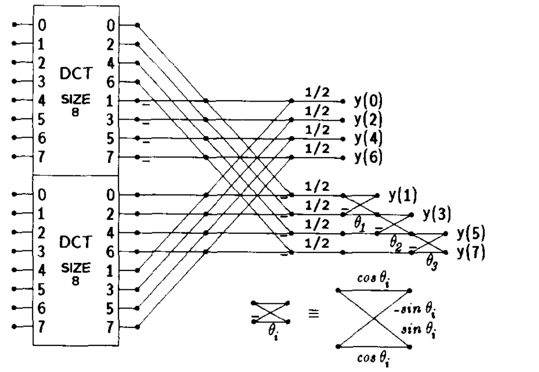


Fig. 2. Flowgraph of the fast LOT for a single block, for $N = 8$. The angles that best approximate the optimal LOT are $\theta_1 = 0.13\pi$, $\theta_2 = 0.16\pi$, and $\theta_3 = 0.13\pi$. The direct LOT is obtained by processing the data from left to right, and the inverse LOT from right to left.

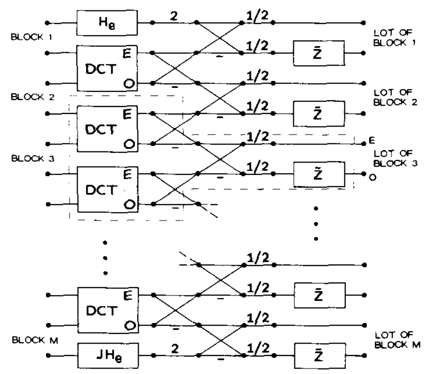
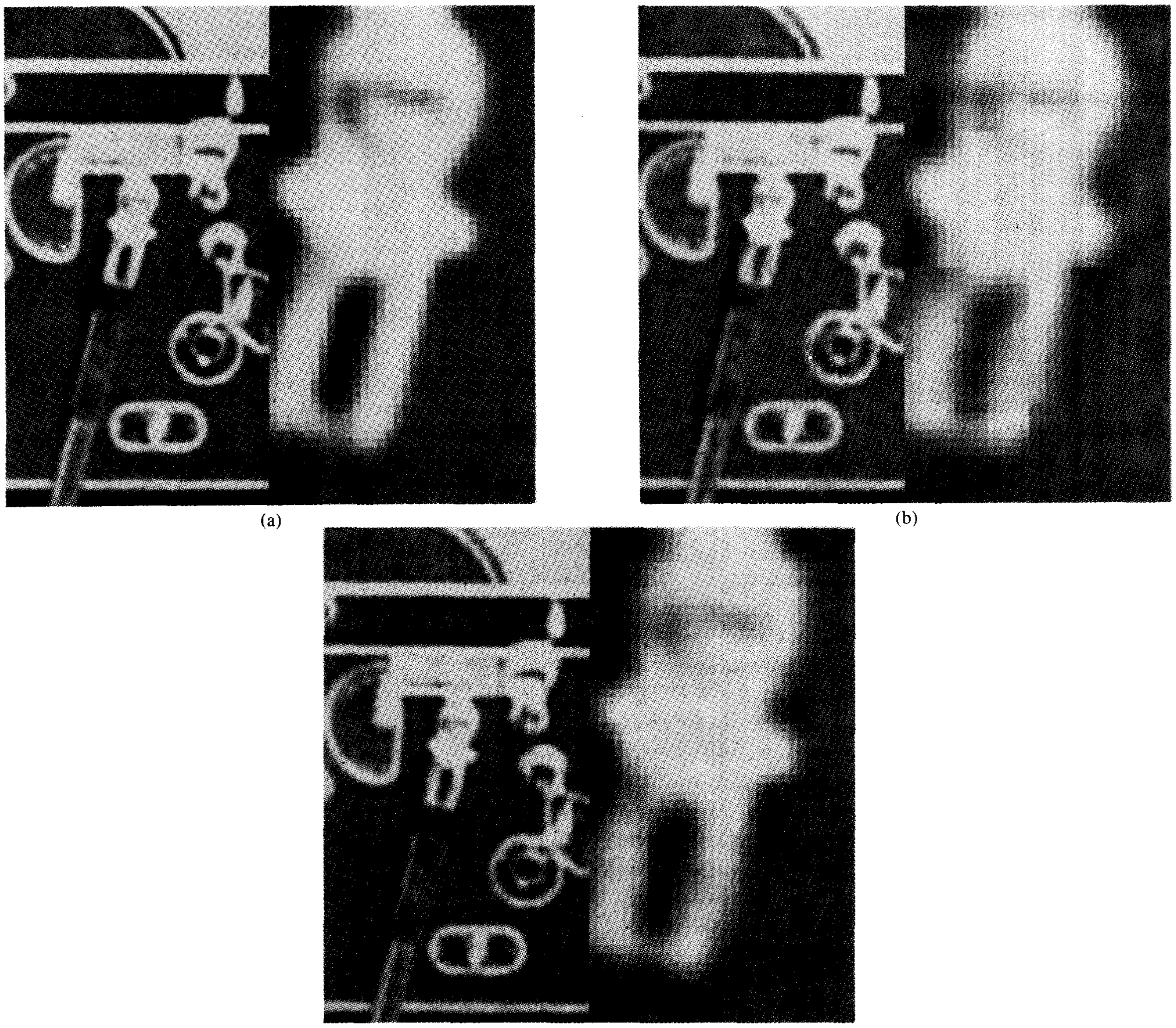


Fig. 3. Flowgraph of the fast LOT for a data segment composed of M blocks of size N. Note that each line in the flowgraph is actually a set of $N/2$ consecutive coefficients, and each butterfly represents a set of $N/2$ actual butterflies, in a simple extension of Fig. 2. Inside the dashed line is the LOT of a single block. The letters "E" and "O" represent even and odd coefficients, respectively.

$$P_0 = \frac{1}{2}\begin{pmatrix} D_e & D_o & 0 & 0 \\ 0 & 0 & D_e & D_o \end{pmatrix} \begin{pmatrix} I & I & & \\ I & -I & & 0 \\ & & I & I \\ 0 & & I & -I \end{pmatrix} \begin{pmatrix} 0 & 0 \\ I & I \\ I & -I \\ 0 & 0 \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & \tilde{Z} \end{pmatrix}. \tag{26}$$

This LOT matrix has been presented in a shorter version of this work, reported in [18]. Here we will present details of the implementation of that LOT. The flowgraph corresponding to the above matrix, for $N = 8$, is shown in Fig. 2. Note that the flowgraph can be used for both the direct and the inverse LOT's, by transposition, as with all orthogonal transforms. Although the flowgraph of Fig. 2 seems to indicate that we need to compute two DCT's of size N to obtain N LOT coefficients, this is not so, as we discuss below.

Now we return to the point that what we need to compute is the LOT of the whole data segment x_o in (1), and also to the relationship among P_0, P_1, and P_2, in (3). From Fig. 2, it is clear that the DCT's used in block r can also be used in part for blocks $r - 1$ and $r + 1$, as shown in Fig. 3. The LOT of the first and last blocks, P_1 and P_2, are obtained by reflecting the data at the segment boundaries. This is equivalent to using the block labeled H_e in Fig. 3, where H_e is the matrix containing half of the sam-



(a) (b) (c)

Fig. 4. (a) Original "CAMERA" image. The right side is a magnified view of a segment from the left side. (b) "CAMERA" coded at 0.5 bits per sample with the DCT, with a block size $N = 16$; rms error = 12.1 percent. (c) "CAMERA" coded at 0.5 bits per sample with the LOT, with a block size $N = 16$; rms error = 10.5 percent.

ples of the even DCT functions, that is,

$$D_e = \begin{pmatrix} H_e \\ JH_e \end{pmatrix}, \tag{27}$$

where, as before, the columns of D_e are the even DCT functions.

We note that the LOT of a data segment of M blocks of size N can be computed by first computing the DCT of all blocks, as in traditional transform coding, and then applying the $+1/-1$ and $\tilde{Z}$ butterflies of Figs. 2 and 3.

V. An Image Processing Example

The shapes of the LOT functions in Fig. 1 suggest that the LOT can be very effective in reducing the blocking effects. As an example, consider the "CAMERA" image in Fig. 4(a), which contains 256×240 pixels, at 8 bits per pixel. The right half of the image was replaced by the magnified version of a region of the left half, so that the effects of processing over that particular area of the image could be better observed.

In Fig. 4(b), the image was coded with the DCT at an average rate of 0.5 bits per sample, with a block size $N = 16$, and a single bit pattern with scalar Max quantizers for all blocks, as described in [4]. The image quantity is somewhat low for half a bit per pixel, because the bit pattern was not adaptive. We note that the blocking effect in the magnified area is strong enough to be annoying. In Fig. 4(c), the DCT was replaced by the LOT, at the same rate of 0.5 bits per sample, with a new bit pattern, rederived in terms of the new estimated coefficient variances. The blocking effects are reduced to a level where they can barely be detected by the human eye. The coding noise pattern is virtually unaffected by the LOT, being mainly a function of the quantization process. The rms error was slightly lower with the LOT, the main reason being that the compaction G_{TC} of the LOT is somewhat larger than that of the DCT, for the same value of N.

VI. Conclusion

We have derived an optimal set of overlapping basis functions, which comprise the Lapped Orthogonal Transform, LOT. Unlike the derivation in [1] and [2], where the basis functions are obtained recursively as the solution to a series of nonlinear optimization problems, we have obtained the LOT as the solution to a simple eigenvalue problem. Therefore, we have derived a direct representation for the LOT. By approximating one of the factors of the optimal LOT by a product of plane rotations, it was possible to derive an efficient implementation for the LOT, which allows LOT-based transform coding to be realized with little computational overhead, when compared to the DCT.

A typical image processing example has shown the efficiency of the LOT in reducing the blocking effects, in agreement with the experiments reported in [1]–[3]. We believe that the fast LOT introduced in this paper allows the implementation of block coding systems at low bit rates (below 1.5 bits per sample) with much less noticeable blocking effects than traditional DCT-based transform coding.

One point about the LOT that we have not mentioned before is that the LOT can be viewed, as any block transform, as a critically sampled multirate filter bank. Unlike all nonoverlapping transforms, the basis functions of the LOT are the impulse responses of reasonably good bandpass filters. Since the conditions in (4) and (5) guarantee perfect reconstruction in the absence of coding, the LOT can also be viewed [20] as a quadrature-mirror-filter (QMF) bank [19] that has a fast algorithm.

Acknowledgment

The authors would like to thank B. Hinman, J. Bernstein, C. Clapp, and Dr. S. Ericsson for many useful comments.

References

[1] P. Cassereau, "A new class of optimal unitary transforms for image processing," S. M. thesis, Dep. Elec. Eng. Comput. Sci., Mass. Inst. Technol., Cambridge, MA, May 1985.

[2] P. M. Cassereau, D. H. Staelin, and G. de Jager, "Encoding of images based on a lapped orthogonal transform," to be published.

[3] H. S. Malvar, "Optimal pre- and post-filters in noisy sampled-data systems," Ph.D. dissertation, Dept. Elec. Eng., Mass. Inst. Technol., Aug. 1986. (Also as Tech. Rep. 519, Res. Lab. Electron., Mass. Inst. Technol., Aug. 1986.)

[4] R. J. Clarke, *Transform Coding of Images*. London, England: Academic, 1985.

[5] N. S. Jaynant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

[6] A. Segall, "Bit allocation and encoding for vector sources," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 162–169, Mar. 1976.

[7] H. S. Malvar, "Fast computation of the discrete cosine transform through fast Hartley transform," *Electron Lett.*, vol. 22, pp. 352–353, Mar. 27, 1986.

[8] W.-H. Chen and C. H. Smith, "Adaptive coding of monochrome and color images," *IEEE Trans. Commun.*, vol. COM-25, pp. 1285–1292, Nov. 1977.

[9] H. G. Musmann, P. Pirsch, and H.-J. Grallert, "Advances in picture coding," *Proc. IEEE*, vol. 73, pp. 523–548, Apr. 1985.

[10] H. C. Reeve, III, and J. S. Lim, "Reduction of blocking effect in image coding," in *Proc. ICASSP 83*, Boston, MA, pp. 1212–1215.

[11] S. Ericsson, "Fixed and adaptive predictors for hybrid predictive/transform coding," *IEEE Trans. Commun.*, vol. COM-33, pp. 1291–1302, Dec. 1985.

[12] D. E. Pearson and M. W. Whybray, "Transform coding of images using interleaved blocks," *IEE Proc.*, Part F, vol. 131, pp. 466–472, Aug. 1984.

[13] H. S. Malvar, "A pre- and post-filtering technique for the reduction of blocking effects," presented at the Picture Coding Symp., Stockholm, Sweden, June 1987.

[14] B. L. Hinman, J. G. Bernstein, and D. H. Staelin, "Short-space Fourier transform image processing," in *Proc. ICASSP 84*, San Diego, CA, pp. 4.8.1–4.8.4.

[15] R. J. Clarke, "Relation between the Karhunen-Loève and cosine transforms," *IEEE Proc.*, Part F, vol. 128, pp. 359–360, Nov. 1981.

[16] A. Cantoni and P. Butler, "Eigenvalues and eigenvectors of symmetric centrosymmetric matrices," *Lin. Algebra Appl.*, vol. 13, pp. 275–288, 1976.

[17] J. Makhoul, "On the eigenvectors of symmetric Toeplitz matrices," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 868–872, Aug. 1981.

[18] H. S. Malvar and D. H. Staelin, "Reduction of blocking effects in image coding with a lapped orthogonal transform," in *Proc. ICASSP 88*, New York, pp. 781–784.

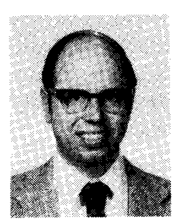
[19] M. Vetterli, "A theory of multirate filter banks," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 356–372, Mar. 1987.

[20] H. S. Malvar, "The LOT: A link between block transform coding and multirate filter banks," in *Proc. Int. Symp. Circuits Syst.*, Espoo, Finland, July 1988, pp. 835–838.

Henrique S. Malvar (S'82–M'86) was born in Rio de Janeiro, Brazil, in 1957. He received the B.S. degree in 1977 from Universidade de Brasilia, Brazil, the M.S. degree in 1979 from Universidade Federal do Rio de Janeiro, Brazil, and the Ph.D. degree in 1986 from the Massachusetts Institute of Technology, Cambridge, all in electrical engineering.

Since 1979 he has been on the Faculty of the Universidade de Brasilia, Brazil, where he is now Associate Professor of Electrical Engineering. From 1982 to 1987 he was on leave at the Massachusetts Institute of Technology, where he had been a Doctoral student from 1982 to 1986, and a Visiting Assistant Professor of Electrical Engineering for the 1986–1987 academic year. From 1985 to 1987 he had also been a consultant for PictureTel Corporation, Peabody, MA, where he was part of the design team for the C-2000 video codec. His teaching, research, and consulting interests include analog and digital filter synthesis, digital signal processing, and low-bandwidth coding of video and speech. He has several publications in these areas, and pending patent applications in the U.S.A., Canada, Europe, and Japan.

Dr. Malvar is a member of Sigma Xi. He was the recipient of the Young Scientist Award from the Marconi International Fellowship in 1981.



David H. Staelin (S'59–M'65–SM'75–F'79) received the Bachelor's, Master's, and Doctorate degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge.

He is Professor of Electrical Engineering at the Massachusetts Institute of Technology, Cambridge, where he has served on the Faculty since 1965. His current research interests involve theoretical and experimental research in remote sensing, digital signal processing, and radio and optical astronomy. His digital signal processing research involves high- and low-definition video image processing, transform methods, and acoustic pattern recognition. His current remote sensing research involves passive microwave techniques for imaging the terrestrial atmosphere and surface from satellites and aircraft. He was principal investigator for the NASA Nimbus-5 and Nimbus-6 microwave spectrometer experiments, which observed the terrestrial atmosphere at millimeter wavelengths from space, and he was co-investigator for the Nimbus-7 scanning multichannel microwave radiometer and Voyager Planetary Radio Astronomy experiments. His teaching involves electromagnetic radiation, signal processing, and systems.