ABSTRACT
Packetized image communication over lossy channels presents a reconstruction problem at the decoder. To date, reconstruction algorithms have been developed for fixed coding techniques. This paper examines the dual problem — a block-based coding technique, namely a family of lapped orthogonal transforms (LOTs), is designed to maximize the reconstruction performance of a fixed reconstruction algorithm. Mean-reconstruction, in which missing coefficient blocks are replaced with averages of their available neighbors, is selected for simplicity and implementation ease. A reconstruction criterion is defined in terms of mean-squared error, and is used to define the reconstruction gain. A family of LOTs is then designed to consider both reconstruction gain and coding gain. Reconstruction capability and compression are exchanged, and the LOT family consists of transforms that provide increasing reconstruction capability with lower coding gain. A transform can therefore be selected based on channel loss characteristics, desired reconstruction performance, and desired compression.

1. INTRODUCTION
Packetized image communication over non-guaranteed channels such as a wireless link or an ATM network presents a reconstruction problem at the decoder, to deal with lost packets and hence lost visual information. Standard techniques for data recovery such as forward error correction and automatic retransmission query protocols become unwieldy at the high data rate required by image transmission [1,2]. However, unlike raw data, visual data contains a great deal of redundancy, and lost information can be reconstructed using lossy signal processing techniques at the decoder. In the case of traditional non-overlapping block-based coding algorithms, previous work has produced reconstruction algorithms to be implemented at the decoder, given that a particular coding algorithm is used (e.g., the DCT) [3-5]. The resulting algorithms are adaptive and require assumptions about the structure of the image data, and can produce poorly reconstructed blocks when the data does not match the assumptions.

To alleviate this problem and to simplify reconstruction, this paper presents and solves the dual problem — a block-based coding technique, namely a family of lapped orthogonal transforms (LOTs), is designed to maximize the reconstruction performance of a fixed reconstruction algorithm while also considering the compression performance. Reconstruction capability and compression are exchanged, and the LOT family consists of transforms that provide increasing reconstruction capability with lower coding gain. The reconstruction-optimized LOT family provides excellent reconstruction capability, and a particular transform from the family can be selected based on the loss characteristics of the channel, the desired reconstruction performance, and the desired compression.

The organization of this paper is as follows. Section 2 outlines the block-based reconstruction problem. Design of the reconstruction-optimized LOTs is presented in Section 3, and Section 4 describes transform performance in both compression and reconstruction. The paper is concluded in Section 5.

2. BLOCK-BASED RECONSTRUCTION
When a block-based transform coded image is lossily compressed and transmitted over a network with non-guaranteed transmission, coefficient blocks can be lost. In one dimension, the block-based reconstruction problem is stated as follows: given the location of a lost coefficient block \( C_z \) and the available adjacent blocks \( C_{adj} \), reconstruct the lost block from the adjacent blocks as \( \hat{C}_z = f(C_{adj}) \), such that an error measure \( d(C_z, \hat{C}_z) \) is minimized. To limit computational complexity at the decoder, \( f \) is restricted to be a linear operator.

Mean-reconstruction is selected for simplicity and implementation ease, and the error measure is the mean-squared error (MSE) of individual reconstructed coefficients within a coefficient block. In one dimension, a coefficient block is reconstructed as \( \hat{c}(n) = (1/2) (c(n-1) + c(n+1)) \), where each \( c \) variable represents an \( N \)-length vector of coefficients, and the coefficient block locations are indexed by \( n \). In two dimensions using a separable transform, this corresponds to averaging the blocks above, below, to the left, and to the right of the lost block.

The linear restriction on \( f \) implies that all reconstruction algorithms will produce the same results for traditional, non-overlapping block-based transforms. Each coefficient block contains information only about an image block in the same location, regardless of the transform used. Without transmitting side information, the only way to modify the performance of such a reconstruction technique is to generate coefficient blocks that correspond to not independent, but overlapping image blocks. A LOT with 50% overlap generates such coefficient blocks, and in two dimensions, any signal value is included in the computation of four coefficient blocks. Such an LOT consists of \( N \) basis functions, each of length \( 2N \). In this paper, LOTs with \( N = 8 \) are considered because the resulting two-dimensional coefficient blocks of size \( 8 \times 8 \) correspond to the size of those used in the JPEG still image compression algorithm.
3. RECONSTRUCTION-OPTIMIZED LOT DESIGN

This section describes the design procedure used to obtain a LOT optimized to minimize the MSE of mean-reconstructed coefficients. The transform is designed one-dimensionally, and is then performed on images independently in both the horizontal and vertical directions. In the following sections, the reconstruction criterion is described and the reconstruction gain is defined. Transform design is then described.

3.1 Reconstruction Criterion

The input signal $x$ is modeled as a Markov-1 signal, with correlation given by $E \{ x_m x_{m+k} \} = \sigma^2 \rho^k$, where the subscripts indicate the time indices of individual values (the signal variance $\sigma^2$ is normalized to 1 in the following). The reconstruction criterion is formulated as a minimization of the MSE in each reconstructed coefficient block. The MSE of the entire transform is therefore the key to designing LOTs with differing reconstruction characteristics is to distribute the error differently across the transform coefficients, and hence across the signal locations.

If a basis function is designed to maximize the expected energy in the corresponding coefficient (for good compression performance) while minimizing the reconstructed coefficient MSE (for good reconstruction performance), the two requirements oppose because the effect of minimizing the MSE is to minimize the energy in the coefficient. The combination of maximizing expected energy while minimizing reconstructed coefficient MSE works to equally distribute the expected energy across all coefficients.

3.2 Reconstruction Gain

Typically, only coding gain is considered in transform design, given by

$$G_{TC} = \left( \frac{1}{N} \sum_{i=0}^{N-1} \sigma_i^2 \right)^{1/N} \prod_{i=0}^{N-1} \sigma_i^2$$

where $\sigma_i^2$ is the $i$th diagonal entry of the transformed input autocovariance matrix, $Z^T R_s Z$, and $Z$ represents any transform matrix. Optimality in terms of compression performance is often argued as $\rho \to 1$ [6], in which case $G_{TC} \to \infty$. However, in the case of reconstruction, as $\rho \to 1$, MSE($\hat{c}_j(n)$) $\to 0$, and this case provides a degenerate reconstruction condition. Therefore, for reconstruction-optimized transforms, “optimality” in the traditional sense of considering performance as $\rho \to 1$ is not appropriate. However, reconstruction performance can be quantified by the reconstruction gain. Because an equal error distribution is desirable, the reconstruction gain is given as the reciprocal of (6) with $R_1$ replaced by $R_{err}$:

$$G_R = \left( \prod_{i=0}^{N-1} T_i^T R_{err} T_i \right)^{1/N} \left( \frac{1}{N} \sum_{i=0}^{N-1} T_i^T R_{err} T_i \right).$$

A transform’s overall performance can therefore be gauged by considering both coding gain $G_{TC}$ and reconstruction gain $G_R$.

3.3 Transform Design

Reconstruction gain and coding gain are simultaneously considered in LOT design by using the sequential design procedure presented in [7]. A modified objective function incorporating both coding gain and reconstruction gain is formed as

$$G_{TC,R} = T_i^T (R_1 - \alpha R_{err}) T_i,$$

where the reconstruction parameter $\alpha$ controls the emphasis placed on reconstruction capability. Therefore, by selecting appropriate values of the reconstruction parameter $\alpha$, reconstruction capability is included in the design of the basis functions. Because the overall reconstruction error is fixed and the basis functions are designed sequentially, each to minimize the reconstruction error, the LOTs designed have a reconstruction error that is more equally distributed across the coefficients.

4. TRANSFORM PERFORMANCE

Four length-16 LOTs ($N = 8$) were designed for $\alpha = \{0.6, 0.7, 0.8, 0.9\}$ with $\rho = 0.95$, and will be denoted T6, T7, T8, and T9, respectively. Transform performance in both compression and reconstruction is presented.
4.1 Compression Performance

The compression performance of the designed LOTs is evaluated using the transform coding gain (as defined in (6)) and by compressing 8-bit $512 \times 512$ luminance images from the USC database (couple, lake, lax, lena, mandrill, peppers) using a modified JPEG algorithm, where only the transform and quantization matrix are modified for each LOT. Coded blocks are assumed to be interleaved to avoid large contiguous areas of loss. Because the transforms were designed to maximize visual performance, compression performance is evaluated by comparing resulting bit rates for the different transforms for a constant image quality.

The coding gains for the four designed transforms are given in Table 1, along with the DCT-LOT, a LOT constructed using the technique presented in [6] (the DCT-LOT provides good compression performance and will be used as a reference transform in the following discussions). The transform coding gains monotonically decrease as the reconstruction parameter $\alpha$ increases, and are reduced by approximately 42% across the transform range.

Because the separable LOTs generate coefficient blocks of size $8 \times 8$, these coefficients can be passed to a modified JPEG encoder that follows the JPEG syntax, but using the LOT coefficients rather than DCT coefficients. However, to provide good quality, different quantization matrices are required. To simplify matrix design and to use the results of previous psychovisual work, the LOT quantization matrices were designed using the JPEG recommended quantization matrix. The cosinusoidal variations in the even basis functions of the 16-point DCT correspond to those in the basis functions of DCT-8, for which the JPEG matrix is designed. A two-dimensional LOT quantization step size is selected as the JPEG quantization step size corresponding to the maximum cosinusoidal component in the horizontal and vertical LOT basis functions as determined by the DCT-16 decomposition.

Figure 1 plots the compression results for the modified JPEG algorithm on the six images. The results are normalized with respect to DCT-LOT (the DCT-LOT yields compressed file sizes 0.2-2% larger than those generated by standard JPEG compression). For all the images, T6 produces file sizes approximately 40% larger than DCT-LOT, while for T7, T8, and T9, the file size continues to monotonically increase with the expansion range approximately constant at 30% across all images. For T8 and T9, the two least active images, lena and peppers, had the worst compression performance, while the two busiest images, lax and mandrill, had the best compression performance.

4.2 Reconstruction

Table 1 lists the reconstruction gain for the designed LOTs, which increases by 50% across the transform range. Including the reconstruction criterion in transform design has the desired effect of more equally distributing the reconstruction error across the coefficients, as shown in Figure 2. As $\alpha$ increases, less error occurs in the first three coefficients, while the cumulative error is approximately equal by the fourth coefficient. Distributing the error in the coefficient domain also produces a "minimax" effect and benefit in the pixel domain — as $\alpha$ increases, the maximum expected error in the pixel domain decreases, as shown in Figure 3. With a smaller maximum expected error, the reconstructed results are expected to be visually more pleasing, as shown in the reconstructed images.

Reconstruction performance was visually evaluated by simulating loss of 25% of all coefficient blocks in a checkerboard pattern, and 10% random block loss. As expected, the equal-MSE property of the LOTs yields PSNRs that are equal, within 0.3 dB, for the reconstructed images. However, visual results differ dramatically. T6, the LOT with the least weight on the reconstruction criterion, provides visual results not much better than the DCT-LOT, but T7, T8, and T9 provide dramatic improvement. The better visual performance can be explained by the minimax pixel error. If there is a threshold below which individual pixel errors cannot be perceived, then the designed transforms exhibit smaller, though perhaps more, pixel errors above this threshold. Figure 4 compares the reconstruction performance of the DCT-LOT and T9 on peppers.

<table>
<thead>
<tr>
<th>Transform</th>
<th>Coding Gain</th>
<th>Reconstruction Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCT-LOT</td>
<td>7.70</td>
<td>0.401</td>
</tr>
<tr>
<td>T6</td>
<td>6.07</td>
<td>0.488</td>
</tr>
<tr>
<td>T7</td>
<td>5.22</td>
<td>0.545</td>
</tr>
<tr>
<td>T8</td>
<td>4.74</td>
<td>0.581</td>
</tr>
<tr>
<td>T9</td>
<td>4.47</td>
<td>0.600</td>
</tr>
</tbody>
</table>

Table 1 Coding and reconstruction gains of the designed LOTs and the DCT-LOT.

5. SUMMARY

This paper presents the design and performance of a family of lapped orthogonal transforms optimized for reconstruction of lost coefficient blocks. Though all LOTs produce the same MSE for mean-reconstruction, the coefficient and pixel error distributions differ, and hence visual results differ. As reconstruction capability increases, visual reconstruction results improve dramatically. The reconstruction-optimized LOTs exchange compression and reconstruction performance, and hence a transform can be selected based on channel loss characteristics, desired reconstruction performance, and desired compression.

REFERENCES

Figure 1  Compression performance for the designed transforms, normalized to the DCT-LOT.

Figure 2  Cumulative reconstruction MSE for coefficients for the designed LOTs and the DCT-LOT.

Figure 3  Expected error per pixel location for the designed transforms and the DCT-LOT (same legend as Figure 2).

Figure 4  Quantized, reconstructed peppers image. (a) DCT-LOT with no reconstruction, showing locations of loss, PSNR=13.0 dB, (b) reconstructed DCT-LOT, PSNR=26.8 dB (c) Reconstructed T9, PSNR=26.9 dB.