Optimal Transforms for Multispectral and Multilayer Image Coding

Daniel Tretter and Charles A. Bouman, *Member, IEEE*

Abstract— **Multispectral images are composed of a series of images at differing optical wavelengths. Since these images can be quite large, they invite efficient source coding schemes for reducing storage and transmission requirements. Because multispectral images include a third (spectral) dimension with nonstationary behavior, these multilayer data sets require specialized coding techniques.**

In this paper, we develop both a theory and specific methods for performing optimal transform coding of multispectral images. The theory is based on the assumption that a multispectral image may be modeled as a set of jointly stationary Gaussian random processes. Therefore, the methods may be applied to any multilayer data set which meets this assumption. Although we do not assume the autocorrelation has a separable form, we show that the optimal transform for coding has a partially separable structure. In particular, we prove that a coding scheme consisting of a frequency transform within each layer followed by a separate KL transform across the layers at each spatial frequency is asymptotically optimal as the block size becomes large. Two simplifications of this method are also shown to be asymptotically optimal if the data can be assumed to satisfy additional constraints. The proposed coding techniques are then implemented using subband filtering methods, and the various algorithms are tested on multispectral images to determine their relative performance characteristics.

I. Introduction

A WIDE variety of modern signal processing applications involve the use of multidimensional data. These can range from medical imaging to remote sensing applications to simple video sequences. Many multidimensional data sets can be characterized as an image consisting of vector-valued pixels. The data can be modeled as a stationary random process across the image, where each pixel is a random vector with correlated components. A common example of such a data set is color imagery, where each pixel is characterized by three scalars representing its red, green, and blue components. A second example is multispectral imagery gathered from remote sensing platforms. These systems gather data from a single area of the earth's surface at a number of optical wavelengths. For some systems each pixel can consist of as many as two hundred spectral components [1].

A primary concern when dealing with data of this sort is the large size of a typical data set. A color image is three times the size of the corresponding monochrome image in raw (uncompressed) form. Multispectral images can easily be an order of magnitude or more larger. In order to store and transmit this data in an efficient manner, a coding scheme must be devised that takes advantage of the unique structure of such data sets.

Since each spectral wavelength of a multispectral image may be thought of as a monochrome image, we will first review methods used for monochrome image coding. Monochrome image compression systems often employ a two dimensional transform [2]. The Karhunen-Loève (KL) transform is optimal for the square error distortion measure [3]. Unfortunately, the KL transform is both data dependent and computationally expensive. Since the transform is data dependent, the transform parameters must be transmitted along with each coded image. In some instances this overhead information can require a significant number of bits.

In practice, 2-D frequency transforms are often used to approximate the optimal KL transform [3]–[5]. If the data can be assumed to be spatially stationary, these transforms are asymptotically optimal. In fact, this assumption is quite reasonable for most images since there is no *a priori* reason to assume the image statistics should be dependent on absolute spatial location. For example, in a randomly chosen image there is little reason to expect the center of the image to behave differently from the other image regions.

Two types of frequency transforms are commonly used. Block frequency transforms, such as the discrete cosine transform (DCT) [6], split the data set into blocks and perform a separate frequency transform on each data block. Alternatively, the QMF subband decomposition [7] uses multirate filtering to divide a data set into a sequence of frequency subbands. One added advantage of subband decomposition is that it avoids the blocking artifacts which are often visible in the reconstructed data [7].

For the coding of multispectral images, a 3-D frequency transform could be used in place of the data dependent KL transform [8]. Unfortunately, this approach is no longer theoretically justifiable since the multispectral data is *not stationary in the spectral dimension.* For example, some spectral wavelengths will behave differently than others due to the physical characteristics of the scene and the intervening atmosphere.

The problem of reducing the necessary data requirements for multispectral imagery has been considered by different researchers for a number of years. Wintz and his associates considered both lossless and lossy compression techniques

Manuscript received December 21, 1992; revised December 10, 1993. This work was supported by an AT&T Bell Laboratories Ph.D. Scholarship and an NEC Faculty Fellowship. The associate editor coordinating the review of this paper and approving it for publication was Prof. William A. Pearlman.

D. Tretter is with the Hewlett-Packard Corp., Palo Alto, CA USA.

C. A. Bouman is with the School of Electrical Engineering, Purdue University, West Lafayette, IN 47907-1285 USA.

IEEE Log Number 9408195.

1057–7149/95$04.00 © 1995 IEEE

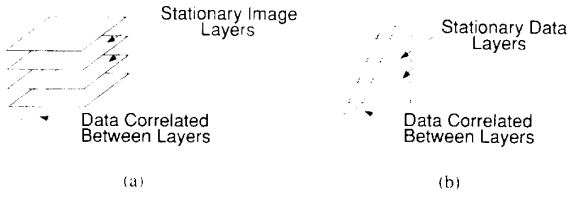


Fig. 1. (a) 3-D data model; (b) 2-D analogous model.

[9], [10]. Other proposed coding methods used data clustering [11], [12] or reduced the dimensionality of the image [13], [14]. Mailhes, Vermande, and Castanie [15] suggest a spectral compression technique based on a parametric characterization of each vector-valued pixel.

Saghri and Tescher [16] and Epstein *et al.* [17] have suggested a method in which the spectral bands are first decorrelated by performing a 1–D KL transform across them. Each transformed band is then treated as a monochrome image and is coded accordingly. This method of coding is similar to that commonly used for color image compression in that it consists of a data transformation followed by monochrome image coding on the transformed data layers [18].

More recently, predictive vector quantization has been studied as a method for multispectral image coding [8], [19]–[21]. This approach does not require that the data behave according to any predefined model, but does require some technique for generating a codebook and computing the prediction coefficients. Since a codebook search is required for each transmitted vector, this method can prove to be quite time-consuming for large codebooks.

In this paper, we implement multispectral image coders based on a novel theory of optimal transforms for multispectral data. Our theory is applicable to the optimal coding of any 3-D Gaussian data which is only stationary in two dimensions. The analysis makes few assumptions about the data and, in particular, remains valid when the data is not separable.

The proposed transform is composed of two steps. First, a block spatial frequency transform is computed for each spectral band. Next, at each spatial frequency, a separate KL transform is computed across spectral bands. We prove that the rate-distortion performance of this transform converges to that of the optimal 3–D KL transform as the block size becomes large. We also show that for the special case of a separable process, this proposed transform is equivalent to the previous methods of [16] and [17].

Our implementation of the multispectral coder uses subband filters in place of the block frequency transform of the theorem. A separate KL transform is then computed for each subband. We choose subband filters to avoid the block coding artifacts that could be undesirable in applications of multispectral data such as segmentation. The subband filters also allow the frequency domain to be divided into bands of various sizes.

We implemented two new multispectral coders. The first coder employs a subbanding scheme that divides the image layers into complex subbands [22]. This is required in order to represent both the magnitude and phase associated with each spatial frequency.

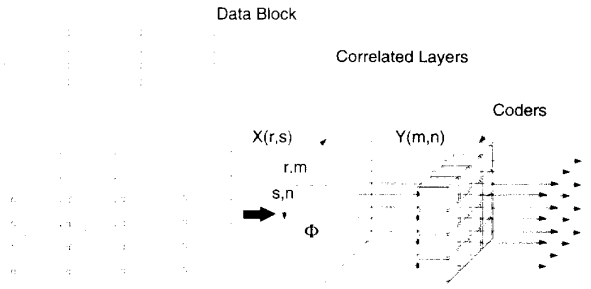


Fig. 2. Data blocks are sent through the block transform sequentially, and each coefficient sequence is sent through a vector coder.

The second coder uses conventional real subband filters instead of the complex ones. This algorithm is shown to result from more restrictive assumptions about the data behavior; it is not appropriate when spectral bands are misregistered spatially. However, the second coder results in less computation and is found to outperform the more general algorithm on real data due to the effects of nonideal subband filters.

We present experimental results which indicate that this approach can improve performance over existing methods without significantly increasing computation. We also look at the effect of coding on classification accuracy of a multispectral image. Since this is a common use of multispectral data, this measurement helps to justify the use of quality measures such as mean square error.

II. Optimal Compression of Multilayer Data

We would like to code a 3-D data set that can be modeled as N_2 layers, where each layer is a 2-D image. The pixel statistics of each layer are assumed to be independent of location, but they may vary from layer to layer, as illustrated in Fig. 1(a). Thus, the model is a 3–D random process that is stationary in two of the dimensions but possibly nonstationary in the third. Although the actual data is 3–D, we will analyze the 2-D analog of the model illustrated in Fig. 1(b). This will reduce the notational complexity.

Suppose the data to be compressed consists of a real 2-D second-order Gaussian random field with zero mean. As illustrated in Fig. 2, the data set is divided into blocks across the first dimension as a prelude to block transform coding. Denote a single N_1-length block of data by $X(r,s)$ $r = 0 \cdots N_1 - 1$, $s = 0 \cdots N_2 - 1$. We will assume that the data layers of $X(r,s)$ are jointly stationary. Thus, $X(r,s)$ is stationary in the index r, but not necessarily in the index s. Therefore, the autocorrelation function for X has the form

$$R_x(r_1 - r_2, s_1, s_2) \equiv E[X(r_1, s_1)X(r_2, s_2)].$$

We note that since R_x does not have a product form, X is not, in general, separable.

Using a square error distortion measure, the optimal coding scheme can be implemented as in Fig. 2 by first applying the linear block KL transform

$$Y(m,n) = \sum_{s=0}^{N_2-1} \sum_{r=0}^{N_1-1} \Phi(m,n,r,s)X(r,s).$$

The coefficients $\Phi(\cdot)$ of the transform can be computed using the eigenvector equation

$$\sum_{s_2=0}^{N_2-1}\sum_{r_2=0}^{N_1-1} R_x(r_1-r_2,s_1,s_2)\Phi(m,n,r_2,s_2) = \lambda(m,n)\Phi(m,n,r_1,s_1) \quad \forall m,r_1 = 0\cdots N_1-1, \; n,s_1 = 0\cdots N_2-1 \quad (1)$$

where the values $\lambda(m,n)$ are the eigenvalues and $\Phi(m,n,\cdot,\cdot)$ are the corresponding eigenvectors of the autocorrelation R_x.

The transformed data $Y(\cdot)$ is uncorrelated, and the variance of each $Y(m,n)$ is the corresponding eigenvalue $\lambda(m,n)$. Note that since $X(r,s)$ is a real random process, the orthonormal transformation $\mathbf{\Phi}$ is real and each of the $\lambda(\cdot)$'s is real and nonnegative.

Each of the transformed components is sent to a separate coder, which is assumed to do optimal quantization and coding. Since the statistics of the data do not vary from block to block, the transformation is the same for each data block. The corresponding coefficients from a sequence of blocks are then quantized and coded as a vector.

The rate-distortion characteristic of this system can be written parametrically as [3]

$$R_{\mathrm{KL}}(\theta,N_1,N_2) = \frac{1}{N_1N_2}\sum_{m=0}^{N_1-1}\sum_{n=0}^{N_2-1}\max\left\{\frac{1}{2}\log\left(\frac{\lambda(m,n)}{\theta}\right),0\right\} \quad (2)$$

$$D_{\mathrm{KL}}(\theta,N_1,N_2) = \frac{1}{N_1N_2}\sum_{m=0}^{N_1-1}\sum_{n=0}^{N_2-1}\min\{\lambda(m,n),\theta\}$$

where each term of the sum represents the contribution due to the (m,n)th coder. This means that for a given desired distortion D_{KL} the bitrate can be made arbitrarily close to R_{KL} through proper coder design. However, (2) can also be shown to represent the rate-distortion bound for the data set $X(r,s)$ at the given block size N_1. This result is based on the 1-D expression derived by Kolmogorov [23] and also presented in Berger [24], which can be extended to multiple dimensions as shown by Hayes, Habibi, and Wintz [25].

Unfortunately, this direct KL transform approach is impractical for 3-D multispectral data. This is because the determination of the KL transformation $\mathbf{\Phi}$ requires the solution of a system of equations dependent on the autocorrelation of the data. Since the autocorrelation is data dependent, the transform matrix cannot be computed and stored in advance. Therefore, the receiver must be sent the transformation matrix in order to reconstruct the original data; this requires considerable overhead for the 3-D case.

Our approach is to exploit the partially stationary nature of the data model by using a partially separable transform structure, as shown in Fig. 3. This approach first transforms $X(r,s)$ along the r coordinate to form $U_c(m,s)$ and $U_s(m,s)$ corresponding to the sine and cosine components of the signal. Next, a series of 1-D KL transforms is applied to both the sine and cosine components along the s coordinate to form $Z(m,n)$.

We will use the discrete time Fourier series (DTFS) for the first transformation $\mathbf{B}$. To simplify notation the block size N_1 is assumed even, although the results hold for arbitrary N_1. The random fields $U_c(m,s)$ and $U_s(m,s)$ are then defined by the equations

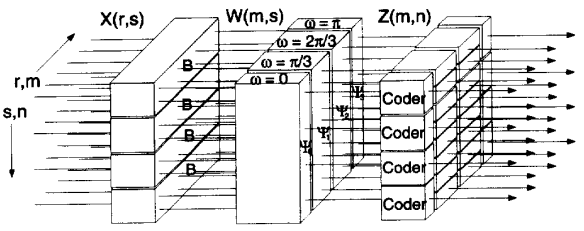


Fig. 3. Separable coding scheme for $N_1 = 6$, $N_2 = 4$, where $\omega = \frac{2\pi m}{N_1}$. $\mathbf{B}$ and Ψ are both chosen to be orthonormal linear transformations.

$$U_c(m,s) = \sum_{r=0}^{N_1-1} b_c(m,r)X(r,s) \quad m = 0\cdots\frac{N_1}{2}$$

$$U_s(m,s) = \sum_{r=0}^{N_1-1} b_s(m,r)X(r,s) \quad m = 0\cdots\frac{N_1}{2}$$

where $b_c(m,r)$ and $b_s(m,r)$ are the basis functions of the DTFS.

$$b_c(0,r) = \sqrt{\frac{1}{N_1}}$$

$$b_c(m,r) = \sqrt{\frac{2}{N_1}}\cos\left(\frac{2\pi m}{N_1}r\right) \quad m = 1\cdots\frac{N_1}{2}$$

$$b_s(m,r) = \sqrt{\frac{2}{N_1}}\sin\left(\frac{2\pi m}{N_1}r\right) \quad m = 0\cdots\frac{N_1}{2}.$$

Notice that for $m = 0, \frac{N_1}{2}$, we have defined $b_s(m,r) = U_s(m,s) = 0$ to simplify notation.

We must next choose the transformation $\mathbf{\Psi}$ for use along the index s. Because the second index s is not associated with stationary behavior, we would not be justified in using a frequency transform for $\mathbf{\Psi}$. Also, since the autocorrelation function is not separable, each of the $\mathbf{\Psi_m}$'s must be computed independently. Therefore, we will choose each $\mathbf{\Psi_m}$ to be the KL transform of the data at the corresponding spatial frequency. Note that the DTFS produces both a sine and cosine term at each spatial frequency; therefore, the KL transform must use both of these terms. To accomplish this, we define the new data set W formed by concatenating the sine and cosine terms at each frequency.

$$W(m,s) = \begin{cases} U_c(m,s) & s = 0\cdots N_2-1 \\ U_s(m,s-N_2) & s = N_2\cdots 2N_2-1 \end{cases} \quad m = 0\cdots\frac{N_1}{2}.$$

We then compute the KL transform at each spatial frequency by calculating the eigenfunctions of the autocorrelation of each column of $W(m,s)$.

$$R_W(m,m,s_1,s_2) = E[W(m,s_1)W(m,s_2)], \quad m = 0\cdots\frac{N_1}{2} \quad s_1,s_2 = 0\cdots 2N_2-1$$

$$\sum_{s_2=0}^{2N_2-1} R_W(m,m,s_1,s_2)\psi_m(n,s_2) = \eta_m(n)\psi_m(n,s_1).$$

At each column index m, or equivalently at each spatial

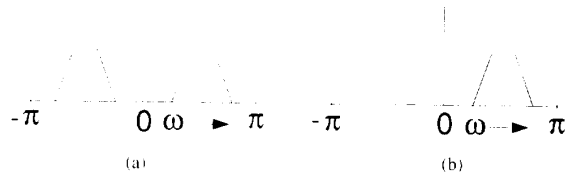


Fig. 4. Spatial frequency ranges of real and complex subbands.

frequency $\omega = \frac{2\pi m}{N_1}$, these transformations are applied to the columns of $W(m,s)$ to decorrelate the data.

$$Z(m,n) = \sum_{s=0}^{2N_2-1} v_m(n,s)W(m,s)$$

$$n = 0 \cdots 2N_2 - 1 \quad m = 0 \cdots \frac{N_1}{2}.$$

Notice that for $m = 0$ and $m = \frac{N_1}{2}$ the values of $Z(m,n), n \geq N_2$ may be assumed to be zero. This is due to the fact that the corresponding values of $W(m,n)$ are all zero.

Since $Z(m,n)$ is zero mean and Gaussian, the rate-distortion characteristic achieved by this separable implementation can be expressed in terms of $\sigma^2_{m,n} \equiv R_z(m,m,n,n) \equiv E[Z(m,n)^2]$.

$$R_{\mathrm{ST}}(\theta, N_1, N_2) = \frac{1}{N_1 N_2} \sum_{m=0}^{\frac{N_1}{2}} \sum_{n=0}^{2N_2-1} \max\left\{\frac{1}{2}\log\left(\frac{\sigma^2_{m,n}}{\theta}\right), 0\right\} \tag{3}$$

$$D_{\mathrm{ST}}(\theta, N_1, N_2) = \frac{1}{N_1 N_2} \sum_{m=0}^{\frac{N_1}{2}} \sum_{n=0}^{2N_2-1} \min\{\sigma^2_{m,n}, \theta\}$$

where $\log(0)$ is taken to be negative infinity.

In Appendix A, we prove the key theorem that the partially separable transform's performance approaches the optimum as the block size N_1 grows large. This result only depends on two technical conditions which disallow covariance functions with unusually slow decay rates.

Theorem 1: Given the second-order Gaussian random process $X(r,s)$ and the transformations and notation defined above, assume that $\forall s_1, s_2 = 0 \cdots N_2 - 1$

$$\sum_{q=-\infty}^{\infty} |R_r(q, s_1, s_2)| < \infty$$

$$\limsup_{q \to \infty} q|R_r(q, s_1, s_2)| < \infty$$

Then the rate-distortion performance of the partially separable transform converges to the optimum performance.

$$\lim_{N_1 \to \infty} R_{ST}(\theta, N_1, N_2) = \lim_{N_1 \to \infty} R_{KL}(\theta, N_1, N_2)$$

$$\lim_{N_1 \to \infty} D_{ST}(\theta, N_1, N_2) = \lim_{N_1 \to \infty} D_{KL}(\theta, N_1, N_2)$$

In order to better interpret this result, some historical context is useful. It has long been recognized that frequency transforms tend to decorrelate stationary processes [26]. Later, it was shown that frequency transforms have asymptotically optimal performance in a rate-distortion sense [27], [28], [24], [3]. This second statement is much stronger since, for example, an orthogonal transformation based on long random sequences will decorrelate data, but will not approach optimum performance in a rate-distortion sense.

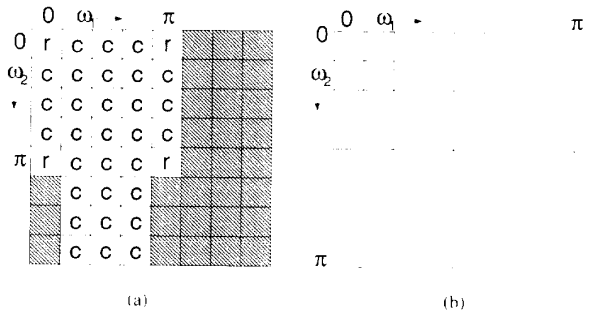


Fig. 5. Spatial frequency decomposition using: (a) complex subbands (the shaded subbands are completely determined by the unshaded ones and need not be computed); (b) real subbands.

The asymptotic optimality of frequency transforms may be proven by analyzing the Toeplitz covariance matrices which must result from stationary processes. Grenander and Szegö showed that eigenvalues of these Toeplitz matrices may be approximated by the eigenvalues of their circulant equivalents [27], [28]. Since the KL transform for any circulant matrix is simply a DFT, it is not surprising that frequency transforms approximate the KL transform for stationary processes. Later, Sakrison extended this result to multidimensional stationary processes, which for the 2-D case have Toeplitz block Toeplitz covariance matrices [29].

Our result shows that frequency transforms may also be employed to achieve asymptotically optimal rate-distortion performance for a 2-D process when only one dimension is stationary. This is essentially equivalent to proving a result about the asymptotic eigenvalue distribution of *block Toeplitz matrices in which the blocks are not Toeplitz.*

III. Implementation

In this section, we will describe three specific implementations of a multispectral image coder based on the partially separable transformation of Section II. The first algorithm (CSM) makes the weakest assumptions about the data's behavior, but it suffers from several implementational drawbacks. The second algorithm (RSM) assumes that the cross-correlation between subbands is a symmetric function. We found this method to have the best performance on real data. The third algorithm (RSS) makes the overly restrictive assumption that the data is separable.

While still image coders often use distortion measures which are matched to human visual sensitivity [2], we choose to use a simple MSE distortion measure in this work. This distortion measure is appropriate since multispectral images are often processed by computers using a variety of application-dependent algorithms. We note that any quadratic distortion measure may also be used by simply transforming the data appropriately before coding [30], [31].

A. Complex Subbands With Multiple KL Transforms (CSM)

Although the proof of Theorem 1 assumes a block DTFS is used as the first transform, any of a variety of frequency transforms will work. We choose to use QMF subband decomposition because it removes the effects of block boundary

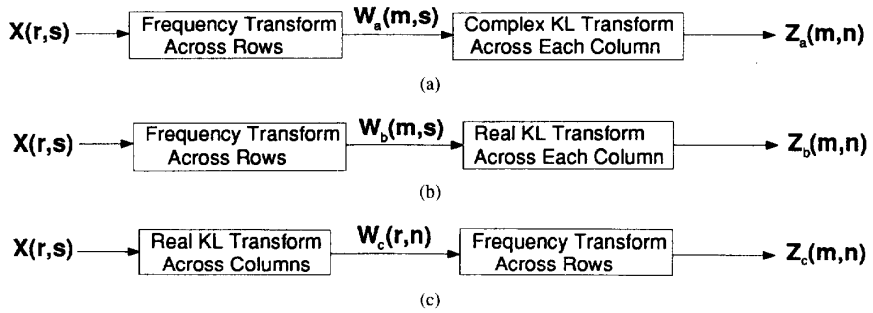


Fig. 6. (a) CSM algorithm; (b) RSM algorithm; (c) RSS algorithm.

artifacts that are common in block transform methods [7]. These block artifacts could be undesirable in the many applications which require segmentation of multispectral images.

Subbanding: The result of Theorem 1 depended critically on the use of both cosine (in-phase) and sine (quadrature) components of the signal at each spatial frequency. Intuitively, this is because the KL transform exploits the phase as well as the magnitude at each frequency. Unfortunately, conventional subband decomposition methods divide the signal into real subbands, as illustrated in Fig. 4(a), so they do not explicitly produce this phase information at the output. In order to explicitly generate both magnitude and phase components, we must use complex subband filters with asymmetric frequency response, as shown in Fig. 4(b). We will refer to this approach as the complex subbanding (CSM) method. Specifically, we use the design procedure of Swaminathan and Vaidyanathan [22] to design 2-D separable filters. This procedure employs a parallel filter bank that divides the signal into multiple subbands simultaneously. The parallel structure is required in this application to minimize phase error due to aliased components in the image subbands.

Fig. 5(a) shows how the 2-D discrete frequencies are broken up into 64 square blocks or subbands. Due to the conjugate symmetry of the frequencies, the shaded subbands are completely determined by the unshaded subbands and thus need not be computed. Of the remaining subbands, those that are purely real are designated by "r" and those that are complex by "c." Note that if each complex component is counted as two data values, the amount of data at the output of this transformation is equal to the amount at the input.

Since Theorem 1 states that the partially separable transform becomes optimal as the block size grows, the performance should improve as the subbands decrease in bandwidth. Alternatively, larger subbands can increase the amount of phase variation across a subband. This phase variation reduces the effectiveness of the KL transform which will be applied across spectral components. By choosing the subbands to have equal size at each spatial frequency, we minimize the degradation in performance in each subband.

Keeping in mind that the output of the analysis filters is subsampled by a factor of 8, a bank of 32 tap filters was chosen for the analysis stage. This leads to 176 tap synthesis filters [22]. The large synthesis filters are required to minimize reconstruction error in the image. In order to constrain the 1-D filter outputs to have the same number of samples as the inputs, we applied circularly symmetric boundary conditions at the image edges. The mean of the lowest order subband ($\omega = 0$) is computed, quantized, and subtracted, so that all subbands have mean zero. The quantized means are then included as overhead in the coded images.

KL Transform: As indicated in Fig. 6(a), the KL transform is computed across spectral bands after the subband decompositions. This is done by taking corresponding pixels in each spectral band, forming a possibly complex spectral vector, and performing the appropriate linear transformation. The spectral vectors in the subband are used to compute a sample autocorrelation matrix, which is then used to compute the KL transform for that subband. Although the KL transforms must be stored along with the image, they are only 1-D, so they do not generally represent a substantial amount of overhead data.

In general, the spectral vectors will be complex. However, because these vectors are samples from the complex envelope of jointly stationary bandpass signals, it may be shown that $E[\chi\chi^T] = 0$, where χ is a spectral vector. This result follows immediately from a similar result proven for the case of samples from the complex envelope of a single stationary bandpass process [32]. This fact allows substantial simplifications, since the KL transform can then be computed using an estimate of the complex autocorrelation matrix $E[\chi\chi^*]$ where χ^* is the conjugate transpose of χ [32], [33]. Thus, the sample autocorrelation matrix is a $N_2 \times N_2$ complex matrix rather than a $(2N_2)\times(2N_2)$ real matrix. This complex autocorrelation will have complex eigenvectors, so the KL transform will also be complex.

Quantization: We next need to determine the optimal values of the quantization step size for each KL coefficient of each subband. Since the KL transform is an orthonormal linear transformation, each KL coefficient for a given subband can be quantized using the same quantization step size, followed by entropy coding of the quantized data. Thus, we only need to consider the rate allocation among the various spatial frequency subbands. The optimal subband rate allocation for real subbands is derived by Nanda and Pearlman [34] for stationary Gaussian sources. Suppose each real subband m has quantization parameter θ_m and (one-sided) bandwidth W_m in both the horizontal and vertical directions. Nanda and Pearlman derived the optimal values for the θ_m's to code the

signal with overall quantization parameter θ. Their proof relied on the fact that for a 2-D signal the total rate R_T and distortion D_T can be written in terms of subband rate R_{θ_m} and distortion D_{θ_m} as follows:

$$R_T = \sum_{\text{subbands}} \left(\frac{W_m}{\pi}\right)^2 R_{\theta_m} \qquad D_T = \sum_{\text{subbands}} D_{\theta_m}.$$

These relationships were also derived by Antonini, *et al.* [35] for the special case of octave-based subbands (or equivalently, a wavelet transform). Using these relations, the optimal coding parameters for a real subbanding system are given by:

$$\theta_m = \left(\frac{W_m}{\pi}\right)^2 \theta.$$

This result can be extended to the case of complex subbands by noting that if the signal is decomposed into both real and complex subbands, the total rate and distortion are given by the following equations:

$$R_T = \sum_{\text{real}} \left(\frac{W_m}{\pi}\right)^2 R_{\theta_m} + \sum_{\text{complex}} \left(\frac{W_m}{\pi}\right)^2 R_{\theta_m}$$

$$D_T = \sum_{\text{real}} D_{\theta_m} + 2 \sum_{\text{complex}} D_{\theta_m}.$$

Notice that the total distortion depends on twice the distortion from the complex subbands. This results from the fact that a complex band determines both positive and negative frequency components in the reconstructed signal.

Using these revised equations, the optimal coding parameters are given by

$$\theta_m = \begin{cases} \left(\frac{W_m}{\pi}\right)^2 \theta & \text{for real subbands} \\ \left(\frac{W_m}{2\pi}\right)^2 \theta & \text{for complex subbands.} \end{cases}$$

Note that we use these coding parameters for uniform PCM quantization instead of optimal vector coders, so they are actually only suboptimal for our implementation. Since the coding parameters are proportional to the mean square distortion per sample, the quantization step size is chosen as

$$Q_m = \begin{cases} \left(\frac{W_m}{\pi}\right)\sqrt{\theta} & m \text{ a real subband} \\ \frac{1}{\sqrt{2}}\left(\frac{W_m}{2\pi}\right)\sqrt{\theta} & m \text{ a complex subband.} \end{cases}$$

Notice that an additional factor of $\frac{1}{\sqrt{2}}$ is introduced for the complex subbands. This accounts for the fact that the quantizer is applied to both the real and imaginary components of each subband.

Coding: The quantized data is coded by sending it through a multialphabet arithmetic coder. The coder is a multiplication-free implementation developed by Rissanen and Mohiuddin [36]. An arithmetic coder is used because it allows the encoder to adapt somewhat to changing data statistics. The symbol alphabet consists of multiplication factors for the coefficients that are not quantized to zero, and zero runlengths for the rest. The quantized data in each subband is ordered and the appropriate alphabet symbols are sent to the coder. The total number of bits required is computed to determine the bitrate.

B. Real Subbands With Multiple KL Transforms (RSM)

The CSM algorithm described in the previous section suffers from several implementational drawbacks. The large synthesis filters cause reconstruction to proceed quite slowly. Also, quantization error in the filter transition bands tends to cause increased MSE for this implementation.

One possible simplification is to replace the complex subbands with real subbands. The FIR filters required are considerably shorter, and the reconstruction is less sensitive to quantization errors in the frequency domain. In fact, as we show in Appendix B, if the cross-correlation between spectral images is a symmetric function of spatial offset, then the phase information is not required in order for Theorem 1 to hold. More specifically, if $R_x(r, s_1, s_2) = R_x(-r, s_1, s_2)$, then the sine and cosine components may be transformed by separate KL transforms. This assumption is often reasonable for multispectral data if the data layers are not misregistered in their spatial alignment. Intuitively, the phase information determines the amount of shifting at each spatial frequency.

Fig. 6(b) shows this algorithm, which we will refer to as RSM (real subbands with multiple KL transforms). The implementation uses a real octave-based subbanding scheme similar to that proposed by Gharavi and Tabatabai [37]. Since subband phase is not used in this method, it was found that subbands of nonuniform bandwidth gave superior performance. Fig. 5(b) shows the structure of the decomposition. The QMF filter prototype is an eight-tap FIR filter designed by Johnston [38], and edge effects were handled by using the symmetric extension of Smith and Eddins [39].

For each subband, a sample covariance matrix is computed across spectral bands. Since the subbands are real, the covariance matrices and their associated KL transforms are also real. The remaining quantization and coding steps for this RSM algorithm are implemented in the same way as for the CSM algorithm described above.

C. Real Subbands With Single KL Transform (RSS)

Fig. 6(c) illustrates the RSS (real subbands with single KL transform) algorithm obtained by performing *a single KL transform* across spectral bands. Notice that, unlike the previous methods, the KL transform is performed before the spatial subbanding. This approach has been proposed by several researchers [16], [17] and is equivalent to assuming that the data is separable along the dimension of the multispectral bands.

Specifically, if the 3-D autocorrelation function has the separable form $R_x(r, s_1, s_2) = f(|r|)g(s_1, s_2)$, then this method is asymptotically optimal. This is easily shown by using separable eigenvectors of the form $\Phi(m, n, r, s) = \Phi_1(m, r) \cdot \Phi_2(n, s)$, and decomposing the double sum of (1) into a product of sums.

In general, this assumption is less realistic for multispectral data. For example, some spectral bands may have different spatial frequency content, leading to different sample covariance matrices for each subband. In this case, the data cannot be made separable by spatially registering the spectral bands.

The RSS algorithm is implemented using the same subbanding, quantization, and coding as used with the RSM algorithm. Since only one KL transform is used, the RSS method requires less overhead data than RSM or CSM. However, we have not found overhead to significantly affect the total bitrate.

Computationally, there is little difference between RSS and RSM, while both require much less computation than CSM. The RSM and RSS algorithms require the same computation for forming the sample covariance matrices, but RSM requires that 9 additional eigenvalue decompositions be performed per image.

IV. Simulation Results

In this section, we compare the proposed algorithms with two additional compression algorithms. The first algorithm takes advantage of spectral redundancy only by performing a single KL transform across spectral bands for the entire image and quantizing the resulting data. The second additional algorithm performs the octave-based subbanding scheme on each spectral band separately, and thus only takes advantage of the spatial redundancy in an image. These two additional implementations serve as performance baselines to show the advantages of taking into account all redundancy in the image.

The distortion measure used in the following plots is peak signal-to-noise ratio (PSNR) rather than MSE. Since all of our test data uses 8-bit spectral bands, this is computed from MSE by the formula

$$\mathrm{PSNR(dB)} = 10 * \log_{10}\left(\frac{255^2}{\mathrm{MSE}}\right)$$

where the 255^2 represents the peak signal energy for 8-bit data.

A. Performance Comparison

The algorithms were tested and compared on four different multispectral images. The results for the first three are in this subsection, while the fourth is used to test classification results. The first data set is a two-band synthetic image, which has been designed to illustrate an instance in which the simplified algorithms are not justified. The first spectral band is an 8-bit image consisting of white Gaussian noise with mean 127 and variance 1600 (standard deviation of 40). The second band is identical to the first except that the entire image is circularly shifted by one row. This is an extreme example of a situation in which the phase associated with each spatial frequency is critical. The cross-correlation function is clearly not symmetric since the correlation coefficient corresponding to a shift of a single row in one direction is unity, while any other correlation coefficient is zero. As a result, we expect the CSM algorithm to outperform the simplified versions. Although this image is a synthetic image designed to illustrate the drawbacks of the schemes involving real subbands, this sort of phase-dependent behavior could exist in real data sets. A multispectral image in which one or more bands is misregistered would, for example, behave in a similar fashion.

The rate versus distortion curves in Fig. 7(a) were generated by running each algorithm on the synthetic image multiple times. The original algorithm gives the best performance, as expected. The algorithms that perform only spatial or only spectral compression were not run for this image since neither would achieve significant compression at high quality levels.

Fig. 7(b) shows the rate-distortion characteristics achieved by the five algorithms on a three band SPOT image, where each spectral band contains 512 × 512 pixels. The spectral bands are well registered for this image, and the KL transform across spectral bands is also largely independent of spatial frequency, so the autocorrelation can be modeled as symmetric and separable. As a result, we expect the two simplified algorithms to perform better than the CSM method, which has the aforementioned implementational drawbacks.

From Fig. 7(b) we see that the RSM and RSS algorithms outperformed the rest. Thus, for this particular image, the implementational drawbacks of the complex subbanding scheme outweigh its theoretical advantages. Between the two simplified algorithms, the RSM approach produces marginally better results than the RSS method. The multiple KL transforms require more overhead information, but this method decorrelates the data better, leading to a slight improvement overall. Since the computation involved is essentially the same for the two techniques, the one using multiple transforms may be preferable for this image.

Fig. 8 shows part of a single spectral band from the reconstructed images of several of the algorithms at a bitrate of 0.8 bpp. As indicated by the rate-distortion curves, the RSM algorithm produces the highest quality output. The algorithm that only performs spectral compression performs quite poorly, producing many contouring artifacts, while the technique that only performs spatial compression blurs the image. These effects are especially evident around the cloverleaf near the center of the image.

Fig. 7(c) shows the performance results on a three-band thematic mapper (TM) image consisting of bands 1, 4, and 6. Band 6 of TM is a thermal band with a larger IFOV (and hence lower resolution) than the other bands. Thus, although the autocorrelation function for this image is nearly symmetric, it is definitely not separable. As a result, we expect the RSM algorithm, which uses multiple KL transforms, to outperform the RSS algorithm, which uses only a single KL transform.

As illustrated in Fig. 7(c), the RSM algorithm gives the best performance, as expected. As with the SPOT image, the implementational drawbacks of the complex subbanding degrade the CSM algorithm's performance considerably. It is interesting to note that the algorithm using spatial compression only outperforms the RSS algorithm, which precedes spatial compression with a single KL transform. The single KL transform across spectral bands actually degrades coding performance by adding high-frequency energy to a band that initially has no high-frequency components.

B. Classification Results

We would like to know how well the MSE (or, equivalently, PSNR) distortion measure correlates with classification accuracy for a multispectral image. Multispectral data is often used to divide a scene into distinct classes in the spatial dimension, where each class is distinguished by the statistical behavior

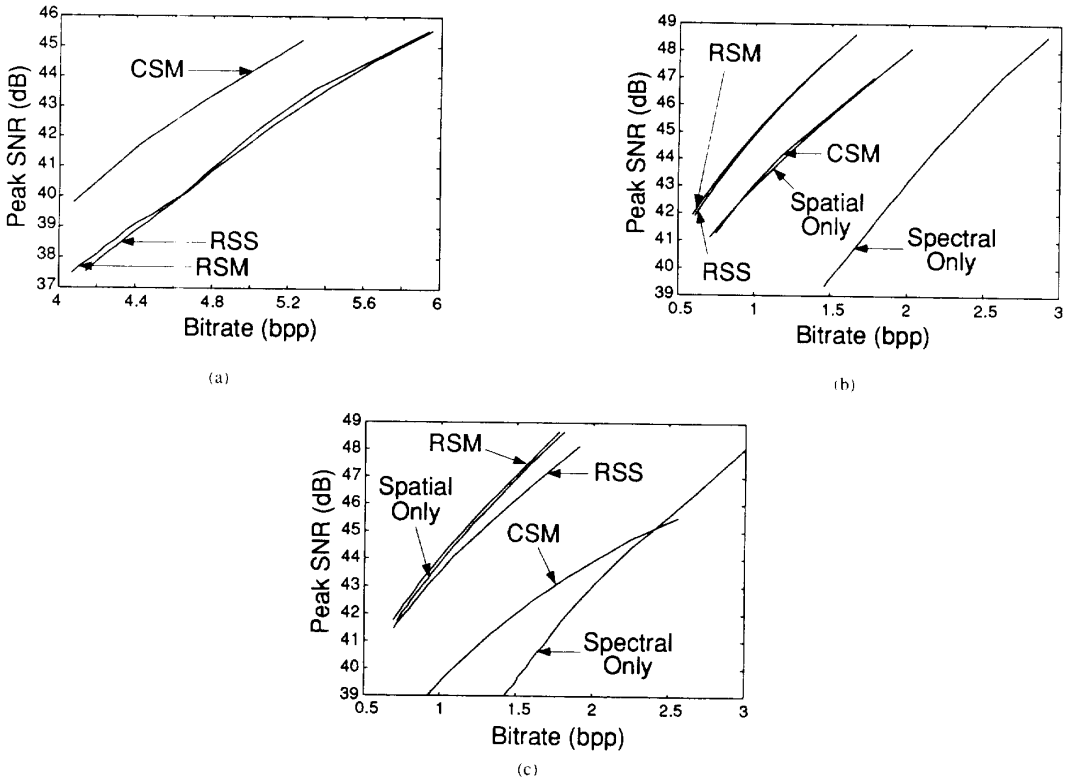


Fig. 7. Coding performance achieved by the various algorithms for: (a) white noise image; (b) SPOT image; (c) TM image.

of the data across spectral bands. Thus, classification error rates are often a more appropriate measure of distortion than MSE. We have quantified the relationship between these two measures for a particular multispectral image by computing classification accuracy at a variety of compression ratios.

Fig. 9(a) illustrates the rate-distortion functions achieved on a multispectral image designated *flightline c1*. *Flightline c1* is a 12-band multispectral image with 949 × 220 pixels in each spectral band. As expected, the RSM algorithm yields the best performance, with the RSS performing slightly worse.

Flightline c1 was gathered by an $M-7$ airborne scanner and consists primarily of farmland for which ground truth is available. Approximately half a dozen ground cover types were identified and labeled in the image for classification purposes. The multispectral image analysis program *Multispec* [40] was used to generate classification error rates for the various compression schemes at a variety of PSNR's via a maximum likelihood (ML) algorithm.

Fig. 9(b) contains a graph of classification accuracy versus PSNR. Since an ML classifier does not use spatial information when classifying the pixels, any operation that tends to spatially smooth the image can actually increase the classification accuracy. Thus, only the spectral compression algorithm gives monotonically decreasing accuracy as PSNR drops. Note that although the spatial compression scheme retains higher accuracy for lower PSNR's, the required bitrate is still larger than for RSM or RSS. The CSM algorithm produces lower classification accuracy than the two real algorithms, which is consistent with the rate-distortion performance.

V. Conclusion

In this paper, we have developed a model for multispectral or multilayer data sets and have used this model to develop three coding algorithms for multispectral data. The first algorithm, CSM, is based on a theory of asymptotically optimal transforms for well behaved Gaussian random processes which are stationary in two out of three dimensions. The second algorithm, RSM, additionally assumes spatial symmetry of the cross-covariance functions between spectral bands. The third algorithm, RSS, assumes that the multispectral image is separable.

In general, the RSM algorithm gave the best coding performance on real multispectral images with essentially the same computation as the RSS algorithm. While the RSM algorithm is less general than the CSM algorithm, it suffered from fewer implementational coding losses. Given these results, the RSM algorithm may be the preferred choice for lossy multispectral image compression.

Appendix A

Here we prove Theorem 1 by showing

$$\lim_{N_1\to\infty} |R_{\mathrm{KL}}(\theta, N_1, N_2) - R_{\mathrm{ST}}(\theta, N_1, N_2)| = 0$$
$$\lim_{N_1\to\infty} |D_{\mathrm{KL}}(\theta, N_1, N_2) - D_{\mathrm{ST}}(\theta, N_1, N_2)| = 0.$$

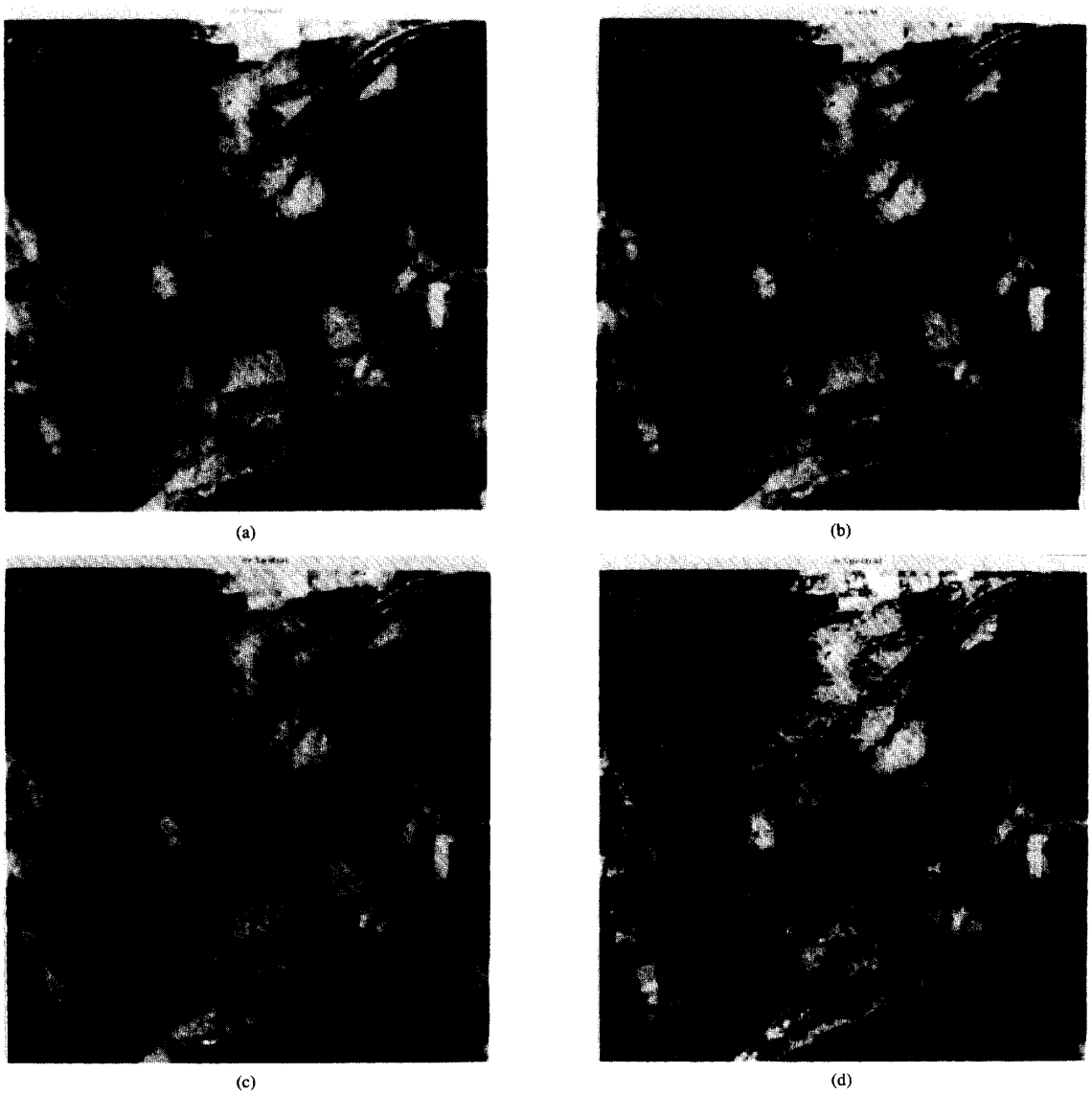


Fig. 8. Band 3 of SPOT image after 10:1 compression: (a) original image; (b) RSM algorithm: (c) spatial compression only; (d) spectral compression only.

We first simplify (2) by reindexing the eigenvalues with a single index.

$$R_{\mathrm{KL}}(\theta, N_1, N_2) = \frac{1}{N_1 N_2} \sum_{k=0}^{N_1 N_2 - 1} \max\left\{\frac{1}{2}\log\left(\frac{\lambda(k)}{\theta}\right), 0\right\} \tag{4}$$

$$D_{\mathrm{KL}}(\theta, N_1, N_2) = \frac{1}{N_1 N_2} \sum_{k=0}^{N_1 N_2 - 1} \min\{\lambda(k), \theta\}.$$

Similarly, the nonzero $\sigma_{m,n}^2$'s can be reindexed to simplify (3).

$$R_{\mathrm{ST}}(\theta, N_1, N_2) = \frac{1}{N_1 N_2} \sum_{k=0}^{N_1 N_2 - 1} \max\left\{\frac{1}{2}\log\left(\frac{\sigma_k^2}{\theta}\right), 0\right\} \tag{5}$$

$$D_{\mathrm{ST}}(\theta, N_1, N_2) = \frac{1}{N_1 N_2} \sum_{k=0}^{N_1 N_2 - 1} \min\{\sigma_k^2, \theta\}.$$

Note that the eigenvalues and variances of (4) and (5) can be indexed in any order.

Since for positive θ, α, and β

$$\left|\max\left(\frac{1}{2}\log\left(\frac{\alpha}{\theta}\right), 0\right) - \max\left(\frac{1}{2}\log\left(\frac{\beta}{\theta}\right), 0\right)\right| \leq \frac{1}{2\theta}|\alpha - \beta|$$

the difference in rates is bounded by

$$|R_{\mathrm{KL}}(\theta, N_1, N_2) - R_{\mathrm{ST}}(\theta, N_1, N_2)| \leq \frac{1}{N_1 N_2} \sum_{k=0}^{N_1 N_2 - 1} \frac{1}{2\theta}|\lambda(k) - \sigma_k^2|.$$

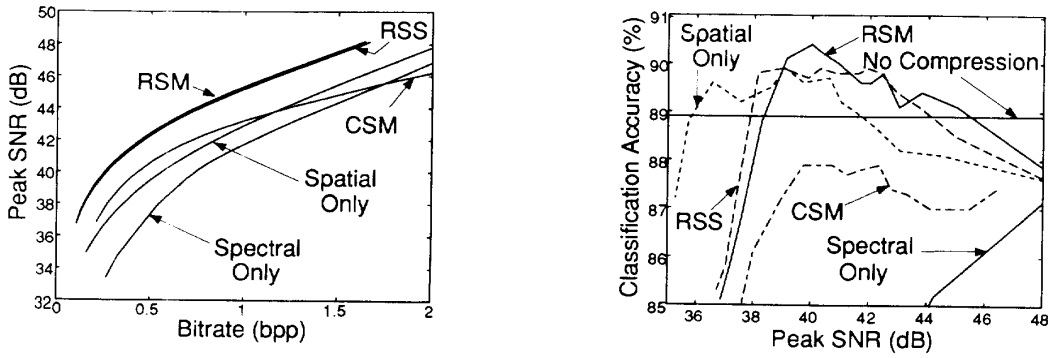


Fig. 9. (a) Coding performance for flightline c1; (b) classification accuracy for compressed images.

Since the l^1-norm is bounded above by the l^2-norm times the square root of the number of components:

$$|R_{\mathrm{KL}}(\theta, N_1, N_2) - R_{\mathrm{ST}}(\theta, N_1, N_2)| \leq \frac{1}{2\theta}\frac{1}{\sqrt{N_1 N_2}}\left(\sum_{k=0}^{N_1 N_2 - 1}(\lambda(k) - \sigma_k^2)^2\right)^{\frac{1}{2}}. \quad (6)$$

Similarly, the difference in distortion can be bounded by

$$|D_{\mathrm{KL}}(\theta, N_1, N_2) - D_{\mathrm{ST}}(\theta, N_1, N_2)| \leq \frac{1}{\sqrt{N_1 N_2}}\left(\sum_{k=0}^{N_1 N_2 - 1}(\lambda(k) - \sigma_k^2)^2\right)^{\frac{1}{2}}. \quad (7)$$

Define the real Toeplitz matrix $\mathbf{R}_x^{(s_1,s_2)}$ with the elements

$$(\mathbf{R}_x^{(s_1,s_2)})_{r_1,r_2} = R_x(r_1 - r_2, s_1, s_2).$$

Let $\gamma = \frac{N_1}{2} - 1$, and define the function

$$R_x(q, s_1, s_2) = \begin{cases} R_x((q+\gamma)_{\mathrm{mod}N_1} - \gamma, s_1, s_2) & s_1 \geq s_2 \\ R_x((-q+\gamma)_{\mathrm{mod}N_1} - \gamma, s_2, s_1) & s_1 < s_2. \end{cases}$$

We now define the circulant matrices $\tilde{\mathbf{R}}_x^{(s_1,s_2)}$ with elements

$$(\tilde{\mathbf{R}}_x^{(s_1,s_2)})_{r_1,r_2} = \tilde{R}_x(r_1 - r_2, s_1, s_2).$$

Note that the elements of $\tilde{\mathbf{R}}_x^{(s_1,s_2)}$ are equivalent to those of $\mathbf{R}_x^{(s_1,s_2)}$ at every coordinate (r_1, r_2) where $|r_1 - r_2| < \frac{N_1}{2}$.

The autocorrelation of the transform output $Z(\cdot)$ is given by

$$\begin{aligned} R_z(m_1, m_2, n_1, n_2) &\equiv E[Z(m_1, n_1)Z(m_2, n_2)] \\ &= E\left[\sum_{s_1=0}^{2N_2-1}\psi_{m_1}(n_1, s_1)W(m_1, s_1) \times \sum_{s_2=0}^{2N_2-1}\psi_{m_2}(n_2, s_2)W(m_2, s_2)\right] \\ &= \sum_{s_1=0}^{2N_2-1}\sum_{s_2=0}^{2N_2-1} R_W(m_1, m_2, s_1, s_2) \times \psi_{m_1}(n_1, s_1)\psi_{m_2}(n_2, s_2). \end{aligned}$$

Evaluating this expression at $m_1 = m_2$,

$$\begin{aligned} R_z(m_1, m_1, n_1, n_2) &= \sum_{s_1=0}^{2N_2-1}\sum_{s_2=0}^{2N_2-1} R_W(m_1, m_1, s_1, s_2) \times \psi_{m_1}(n_1, s_1)\psi_{m_1}(n_2, s_2) \\ &= \eta_{m_1}(n_2)\sum_{s_1=0}^{2N_2-1}\psi_{m_1}(n_1, s_1)\psi_{m_1}(n_2, s_1) \\ &= \begin{cases} \eta_{m_1}(n_2) & n_1 = n_2 \\ 0 & n_1 \neq n_2 \end{cases}. \end{aligned} \quad (8)$$

Define the covariance function

$$R_{cs}(m_1, m_2, s_1, s_2) = E[U_c(m_1, s_1)U_s(m_2, s_2)]$$

and define R_{cc}, R_{sc}, and R_{ss} similarly. Then using the definition of W, we have

$$\begin{aligned} R_z(m_1, m_2, n_1, n_2) &= \sum_{s_1=0}^{N_2-1}\sum_{s_2=0}^{N_2-1}\{\psi_{m_1}(n_1, s_1)\psi_{m_2}(n_2, s_2)R_{cc}(m_1, m_2, s_1, s_2) \\ &\quad + \psi_{m_1}(n_1, s_1)\psi_{m_2}(n_2, s_2 + N_2)R_{cs}(m_1, m_2, s_1, s_2) \\ &\quad + \psi_{m_1}(n_1, s_1 + N_2)\psi_{m_2}(n_2, s_2)R_{sc}(m_1, m_2, s_1, s_2) \\ &\quad + \psi_{m_1}(n_1, s_1 + N_2)\psi_{m_2}(n_2, s_2 + N_2) \\ &\quad \times R_{ss}(m_1, m_2, s_1, s_2)\}. \end{aligned}$$

Now define the vectors

$$\mathbf{b}_c^{(m)} \equiv [b_c(m, 0) \cdots b_c(m, N_1 - 1)]^T$$
$$\mathbf{b}_s^{(m)} \equiv [b_s(m, 0) \cdots b_s(m, N_1 - 1)]^T.$$

Then we can write the autocorrelation function of $Z(\cdot)$ as

$$\begin{aligned} R_z(m_1, m_2, n_1, n_2) &= \sum_{s_1=0}^{N_2-1}\sum_{s_2=0}^{N_2-1}\{\psi_{m_1}(n_1, s_1)\psi_{m_2}(n_2, s_2)(\mathbf{b}_c^{(m_1)})^T \\ &\quad \times \mathbf{R}_x^{(s_1,s_2)}\mathbf{b}_c^{(m_2)} + \psi_{m_1}(n_1, s_1)\psi_{m_2}(n_2, s_2 + N_2) \\ &\quad \times (\mathbf{b}_c^{(m_1)})^T\mathbf{R}_x^{(s_1,s_2)}\mathbf{b}_s^{(m_2)} + \psi_{m_1}(n_1, s_1 + N_2) \\ &\quad \times \psi_{m_2}(n_2, s_2)(\mathbf{b}_s^{(m_1)})^T\mathbf{R}_x^{(s_1,s_2)}\mathbf{b}_c^{(m_2)} \\ &\quad + \psi_{m_1}(n_1, s_1 + N_2)\psi_{m_2}(n_2, s_2 + N_2) \\ &\quad \times (\mathbf{b}_s^{(m_1)})^T\mathbf{R}_x^{(s_1,s_2)}\mathbf{b}_s^{(m_2)}\}. \end{aligned}$$

Also define an analogous function $\tilde{R}_z(\cdot)$ as follows:

$$\begin{aligned}\tilde{R}_z(m_1,m_2,n_1,n_2) &= \sum_{s_1=0}^{N_2-1}\sum_{s_2=0}^{N_2-1}\Big\{\psi_{m_1}(n_1,s_1)\psi_{m_2}(n_2,s_2)(\mathbf{b}_c^{(m_1)})^T \\ &\times \tilde{\mathbf{R}}_x^{(s_1,s_2)}\mathbf{b}_c^{(m_2)} + \psi_{m_1}(n_1,s_1)\psi_{m_2} \\ &\times (n_2,s_2+N_2)(\mathbf{b}_c^{(m_1)})^T\tilde{\mathbf{R}}_x^{(s_1,s_2)}\mathbf{b}_s^{(m_2)} \\ &+ \psi_{m_1}(n_1,s_1+N_2)\psi_{m_2}(n_2,s_2)(\mathbf{b}_s^{(m_1)})^T \\ &\times \tilde{\mathbf{R}}_x^{(s_1,s_2)}\mathbf{b}_c^{(m_2)} + \psi_{m_1}(n_1,s_1+N_2)\psi_{m_2} \\ &\times (n_2,s_2+N_2)(\mathbf{b}_s^{(m_1)})^T\tilde{\mathbf{R}}_x^{(s_1,s_2)}\mathbf{b}_s^{(m_2)}\Big\}. \end{aligned} \tag{9}$$

We will next show that for $m_1 \neq m_2$ $\tilde{R}_z(m_1,m_2,n_1,n_2) = 0$. Define

$$\mathbf{e}_m \equiv \left[1\; e^{j\frac{2\pi m}{N_1}} \cdots e^{j\frac{2\pi m}{N_1}(N_1-1)}\right]^T.$$

We can write the $\mathbf{b}_c^{(m)}$'s and $\mathbf{b}_s^{(m)}$'s in terms of the $\mathbf{e}_m$'s as follows:

$$\begin{aligned}\mathbf{b}_c^{(m)} &= L_m(\mathbf{e}_m + \mathbf{e}_m^*) \\ \mathbf{b}_s^{(m)} &= M_m(\mathbf{e}_m - \mathbf{e}_m^*)\end{aligned}$$

where L_m and M_m are appropriate complex constants.

It is well known that the $\mathbf{e}_m$'s are the eigenvectors of any real $N_1 \times N_1$ circulant matrix [41]. Since the circulant matrix $\tilde{\mathbf{R}}_x^{(s_1,s_2)}$ is real, if $\mathbf{e}_m$ is an eigenvector with eigenvalue ν_m, $\mathbf{e}_m^*$ is also an eigenvector with eigenvalue ν_m^*.

$$\begin{aligned}&(\mathbf{b}_c^{(m_1)})^T\tilde{\mathbf{R}}_x^{(s_1,s_2)}\mathbf{b}_c^{(m_2)} \\ &\quad= L_{m_1}L_{m_2}\big(\nu_{m_2}(\mathbf{e}_{m_1})^T\mathbf{e}_{m_2} + \nu_{m_2}(\mathbf{e}_{m_1}^*)^T\mathbf{e}_{m_2} \\ &\qquad+ \nu_{m_2}^*(\mathbf{e}_{m_1})^T\mathbf{e}_{m_2}^* + \nu_{m_2}^*(\mathbf{e}_{m_1}^*)^T\mathbf{e}_{m_2}^*\big).\end{aligned} \tag{10}$$

Note that for $m_1 \neq m_2$ and $0 \leq m_1, m_2 \leq \frac{N_1}{2}$, $(\mathbf{e}_{m_1})^T\mathbf{e}_{m_2} = (\mathbf{e}_{m_1})^T\mathbf{e}_{m_2}^* = 0$. This implies that (10) is zero for $m_1 \neq m_2$. Similar expressions may derived for the three other terms in (9). Since this must hold for all $s_1, s_2 = 0 \cdots N_2 - 1$, this implies that

$$\tilde{R}_z(m_1,m_2,n_1,n_2) = 0 \quad \text{for} \quad m_1 \neq m_2. \tag{11}$$

Define the $(N_1N_2)\times(N_1N_2)$ matrices $\mathbf{R}_x$ and $\tilde{\mathbf{R}}_x$ as follows.

$$(\mathbf{R}_x)_{N_1s_1+r_1,N_1s_2+r_2} \equiv (\mathbf{R}_x^{(s_1,s_2)})_{r_1,r_2} \tag{12}$$

$$\begin{aligned}(\tilde{\mathbf{R}}_x)_{N_1s_1+r_1,N_1s_2+r_2} &\equiv (\tilde{\mathbf{R}}_x^{(s_1,s_2)})_{r_1,r_2} \\ &\qquad r_1,r_2 = 0\cdots N_1-1 \\ &\qquad s_1,s_2 = 0\cdots N_2-1.\end{aligned} \tag{13}$$

Then $\mathbf{R}_x$ will be a symmetric matrix made up of Toeplitz blocks, where the blocks themselves are arranged in a symmetric (but not necessarily Toeplitz) fashion. The matrix $\tilde{\mathbf{R}}_x$ will be similar except that the blocks will each be circulant instead of Toeplitz.

We next bound the difference between $\mathbf{R}_x$ and $\tilde{\mathbf{R}}_x$ using the Frobenius norm defined for a matrix $\mathbf{A}$ as

$$\|\mathbf{A}\|_F = \left(\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}|a_{ij}|^2\right)^{\frac{1}{2}}.$$

By construction, the difference matrix $\delta\mathbf{R}_x \equiv \mathbf{R}_x - \tilde{\mathbf{R}}_x$ is symmetric, and its components are zero for $|r_1 - r_2| < \frac{N_1}{2}$. Using the symmetry of $\delta\mathbf{R}_x$, we can write the Frobenius norm as a function of the elements for which $q = r_1 - r_2 \geq \frac{N_1}{2}$. Summing over diagonals, we get

$$\begin{aligned}\|\delta\mathbf{R}_x\|_F^2 &= 2\sum_{s_1=0}^{N_2-1}\sum_{s_2=0}^{N_2-1}\sum_{q=\frac{N_1}{2}}^{N_1-1}(N_1-q) \\ &\quad\times |R_x(q,s_1,s_2) - \tilde{R}_x(q,s_1,s_2)|^2 \\ &\leq 2\sum_{s_1=0}^{N_2-1}\sum_{s_2=0}^{N_2-1}\sum_{q=\frac{N_1}{2}}^{N_1-1}(N_1-q) \\ &\quad\times |R_x(q,s_1,s_2) - R_x(q-N_1,s_1,s_2)|^2 \\ &\leq 4\sum_{s_1=0}^{N_2-1}\sum_{s_2=0}^{N_2-1}\left\{\sum_{q=\frac{N_1}{2}}^{N_1-1}(N_1-q)R_x^2(q,s_1,s_2)\right. \\ &\quad\left.+\sum_{q=\frac{N_1}{2}}^{N_1-1}(N_1-q)R_x^2(N_1-q,s_2,s_1)\right\} \\ &\leq 4\sum_{s_1=0}^{N_2-1}\sum_{s_2=0}^{N_2-1}\left\{\sum_{q=\frac{N_1}{2}}^{N_1-1}qR_x^2(q,s_1,s_2)\right. \\ &\quad\left.+\sum_{p=1}^{\frac{N_1}{2}}pR_x^2(p,s_2,s_1)\right\} \\ &\leq 4\sum_{s_1=0}^{N_2-1}\sum_{s_2=0}^{N_2-1}\left\{\sum_{q=1}^{\infty}qR_x^2(q,s_1,s_2)\right. \\ &\quad\left.+\sum_{p=1}^{\infty}pR_x^2(p,s_2,s_1)\right\} \\ &= 8\sum_{s_1=0}^{N_2-1}\sum_{s_2=0}^{N_2-1}\sum_{q=1}^{\infty}qR_x^2(q,s_1,s_2).\end{aligned}$$

Let $S = \limsup_{q\to\infty} q|R_x(q,s_1,s_2)|$ for fixed s_1, s_2. Then, by assumption, $S < \infty$. This implies that there exists $M < \infty$ such that $q|R_x(q,s_1,s_2)| < (S+1)\forall q > M$. From this,we get

$$\begin{aligned}&\sum_{q=1}^{\infty}qR_x^2(q,s_1,s_2) \\ &\quad= \sum_{q=1}^{M}qR_x^2(q,s_1,s_2) + \sum_{q=M+1}^{\infty}qR_x^2(q,s_1,s_2) \\ &\quad\leq K_1 + (S+1)\sum_{q=M+1}^{\infty}|R_x(q,s_1,s_2)| \\ &\quad\leq K_1 + (S+1)K_2 \\ &\quad= K\end{aligned}$$

where K_1, K_2, and K are finite constants. Let $C = \sqrt{8K^*}$, where K^* is the maximum value of K over all $s_1, s_2 = 0\cdots N_2 - 1$, and substitute into the earlier expression for

$\|\delta\mathbf{R}_x\|_F^2$ to get

$$\|\delta\mathbf{R}_x\|_F^2 \le C^2 N_2^2 \Rightarrow \|\delta\mathbf{R}_x\|_F \le CN_2.$$

Let $\mathbf{A}$ be the linear transformation matrix that maps the N_1N_2 components of $X(r,s)$ to the N_1N_2 nonzero values of $Z(m,n)$. The matrix $\mathbf{A}$ is formed by the composition of the DTFS matrix $\mathbf{B}$, the KL transform matrix $\mathbf{\Psi}$, and possibly a permutation of the elements of Z. Since each of these transforms is orthonormal, $\mathbf{A}$ will also be an orthonormal matrix.

The covariance of Z is given by

$$\begin{aligned}\mathbf{R}_z &= \mathbf{A}\mathbf{R}_x\mathbf{A}^T \\ &= \mathbf{A}(\tilde{\mathbf{R}}_x + \delta\mathbf{R}_x)\mathbf{A}^T \\ &= \tilde{\mathbf{R}}_z + \mathbf{A}\delta\mathbf{R}_x\mathbf{A}^T\end{aligned}$$

where $\tilde{\mathbf{R}}_z = \mathbf{A}\tilde{\mathbf{R}}_x\mathbf{A}^T$. If we define $\mathbf{D}$ to be the diagonal part of $\mathbf{R}_z$, and $\mathbf{E}$ to be the off-diagonal part, then we have

$$\begin{aligned}\mathbf{R}_z &= \tilde{\mathbf{R}}_z + \mathbf{A}\delta\mathbf{R}_x\mathbf{A}^T \\ &= \mathbf{D} + \mathbf{E}.\end{aligned} \tag{14}$$

We want to characterize the matrix $\mathbf{E}$. By construction, the diagonal components of $\mathbf{E}$ are zero. The components of $\mathbf{R}_z$ corresponding to $m_1 = m_2, n_1 \ne n_2$ are known to be zero from (8), and since these components must lie off-diagonal, the matrix $\mathbf{D}$ is also zero at these points. Thus, $\mathbf{E}$ is also zero at these points. For the remaining components, which correspond to $m_1 \ne m_2$, it was shown that $\tilde{\mathbf{R}}_z$ is zero, so $\mathbf{E}$ equals $\mathbf{A}\delta\mathbf{R}_x\mathbf{A}^T$ at these points. Thus,

$$\|\mathbf{E}\|_F \le \|\mathbf{A}\delta\mathbf{R}_x\mathbf{A}^T\|_F = \|\delta\mathbf{R}_x\|_F \le CN_2$$

since the Frobenius norm is invariant under orthonormal transformations [42].

Finally, a theorem from linear algebra states that for symmetric matrices $\mathbf{R}_z$ and $\mathbf{D}$

$$\sum_{j=0}^{N-1} (\lambda_j[\mathbf{R}_z] - \lambda_j[\mathbf{D}])^2 \le \|\mathbf{R}_z - \mathbf{D}\|_F^2$$

where $\lambda_j[\mathbf{R}_z]$ and $\lambda_j[\mathbf{D}]$ are some ordering of the eigenvalues of $\mathbf{R}_z$ and $\mathbf{D}$, respectively [42].

Since eigenvalues are invariant under orthonormal linear transformations, the eigenvalues of $\mathbf{R}_z$ are the eigenvalues of $\mathbf{R}_x$, $\lambda(k)$. The eigenvalues of $\mathbf{D}$ are just its diagonal components, σ_k^2. Using the fact that $\mathbf{E} = \mathbf{R}_z - \mathbf{D}$ from (14) yields

$$\sum_{k=0}^{N_1N_2-1} (\lambda(k) - \sigma_k^2)^2 \le \|\mathbf{E}\|_F^2 \le C^2N_2^2$$

for an appropriate ordering of the $\lambda(k)$'s and σ_k^2's. Combining this expression with (6) and (7), we get the equations

$$|R_{\mathrm{KL}}(\theta, N_1, N_2) - R_{\mathrm{SI}}(\theta, N_1, N_2)| \le \frac{C\sqrt{N_2}}{2\theta}\frac{1}{\sqrt{N_1}}$$

$$|D_{\mathrm{KL}}(\theta, N_1, N_2) - D_{\mathrm{SI}}(\theta, N_1, N_2)| \le C\sqrt{N_2}\frac{1}{\sqrt{N_1}}.$$

Taking the limit as $N_1 \to \infty$ gives the desired result.

APPENDIX B

In this appendix, we show that when $R_x(r, s_1, s_2) = R_x(-r, s_1, s_2)$, the sine and cosine components may be transformed separately. We define a modified version of W by concatenating the sine and cosine terms for each band:

$$W(m,s) = \begin{cases} U_c(m,s) & m = 0 \cdots \frac{N_1}{2} \\ U_s(m - \frac{N_1}{2}, s) & m = \frac{N_1}{2} + 1 \cdots N_1 - 1 \\ & s = 0 \cdots N_2 - 1. \end{cases}$$

The definition of $Z(m,n)$ remains the same with the only differences being in the limits of summation. We only need to show that with these definitions of W and Z:

$$R_z(m, m, n_1, n_2) = 0 \quad \text{for} \quad n_1 \ne n_2 \tag{15}$$

$$\tilde{R}_z(m_1, m_2, n_1, n_2) = 0 \quad \text{for} \quad m_1 \ne m_2. \tag{16}$$

If these two conditions hold, then the proof of Appendix A will go through as before.

Clearly, (15) must hold since $Z(m,n)$ is the KL transform of $W(m,s)$ across the second dimension and must therefore decorrelate the data in that dimension. To show that (16) holds consider the form of $\tilde{R}_z$

$$\begin{aligned}\tilde{R}_z(m_1, m_2, n_1, n_2) = &\sum_{s_1=0}^{N_2-1}\sum_{s_2=0}^{N_2-1} \tilde{R}_W(m_1, m_2, s_1, s_2) \\ &\times \psi_{m_1}(n_1, s_1)\psi_{m_2}(n_2, s_2)\end{aligned}$$

where $\tilde{R}_W$ may be written

$$\tilde{R}_W(m_1, m_2, s_1, s_2) = \begin{cases} (\mathbf{b}_c^{(m_1)})^T \tilde{\mathbf{R}}_x^{(s_1,s_2)} \mathbf{b}_c^{(m_2)} & \text{if } m_1, m_2 \le \frac{N_1}{2} \\ (\mathbf{b}_c^{(m_1)})^T \tilde{\mathbf{R}}_x^{(s_1,s_2)} \mathbf{b}_s^{(m_2)} & \text{if } m_1 \le \frac{N_1}{2}, m_2 > \frac{N_1}{2} \\ (\mathbf{b}_s^{(m_1)})^T \tilde{\mathbf{R}}_x^{(s_1,s_2)} \mathbf{b}_c^{(m_2)} & \text{if } m_1 > \frac{N_1}{2}, m_2 \le \frac{N_1}{2} \\ (\mathbf{b}_s^{(m_1)})^T \tilde{\mathbf{R}}_x^{(s_1,s_2)} \mathbf{b}_s^{(m_2)} & \text{if } m_1, m_2 > \frac{N_1}{2}. \end{cases}$$

Appendix A showed that most of these terms for $\tilde{R}_W$ are zero when $m_1 \ne m_2$. The only two terms in question are $\tilde{R}_W(m, m+\frac{N_1}{2}, s_1, s_2)$ and $\tilde{R}_W(m+\frac{N_1}{2}, m, s_1, s_2)$, $0 < m < \frac{N_1}{2}$. By assumption, the circulant matrix $\tilde{\mathbf{R}}_x^{(s_1,s_2)}$ will now be symmetric, and therefore have real eigenvalues. This implies

$$\begin{aligned}\tilde{R}_W\left(m, m+\frac{N_1}{2}, s_1, s_2\right) &= \tilde{R}_W\left(m+\frac{N_1}{2}, m, s_1, s_2\right) \\ &= (\mathbf{b}_c^{(m)})^T \tilde{\mathbf{R}}_x^{(s_1,s_2)} \mathbf{b}_s^{(m)} \\ &= L_m^2 \nu_m \left((\mathbf{e}_m)^T \mathbf{e}_m + (\mathbf{e}_m^*)^T \mathbf{e}_m \right. \\ &\quad \left. - (\mathbf{e}_m)^T \mathbf{e}_m^* - (\mathbf{e}_m^*)^T \mathbf{e}_m^*\right) \\ &= 0.\end{aligned}$$

ACKNOWLEDGMENT

The authors would like to thank D. Landgrebe, L. Biehl, and the Purdue Research Foundation for providing the necessary software, documentation, and technical support on the *Multispec* software package.

REFERENCES

[1] A. F. H. Goetz and M. Herring, "The high resolution imaging spectrometer (HIRIS) for EOS," *IEEE Trans. Geosci. Remote Sensing*, vol. 27, no. 2, pp. 136–144, Mar. 1989.

[2] A. N. Netravali and J. O. Limb, "Picture coding: A review," *Proc. IEEE*, vol. 68, no. 3, pp. 366–406, Mar. 1980.

[3] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Boston: Kluwer, 1992.

[4] R. J. Clarke, "Relation between the Karhunen Loève and cosine transforms," *IEE Proc.*, vol. 128, pt. F, no. 6, pp. 359–360, Nov. 1981.

[5] H. S. Hou and D. R. Tretter, "Frequency characterization of the discrete cosine transform," in *Proc. SPIE Int. Conf. Appl. Digital Image Processing* (San Diego, CA), July 10–13, 1990, pp. 31–42.

[6] N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete cosine transform," *IEEE Trans. Comput.*, pp. 90–93, Jan. 1974.

[7] J. W. Woods and S. D. O'Neil, "Subband coding of images," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP-34, pp. 1278–1288, 1986.

[8] R. L. Baker and Y. T. Tse, "Compression of high spectral resolution imagery," in *Proc. SPIE Int. Conf. Appl. Digital Image Processing*, pp. 255–264, 1988.

[9] J. R. Duan and P. A. Wintz, "Information preserving coding for multispectral data," in *Proc. Machine Processing of Remotely Sensed Data Symp.*, 1973, pp. 28–35.

[10] P. J. Ready and P. A. Wintz, "Information extraction, SNR improvement, and data compression in multispectral imagery," *IEEE Trans. Commun.*, vol. COM-21, no. 10, pp. 1123–1130, Oct. 1973.

[11] A. Habibi and A. S. Samulon, "Bandwidth compression of multispectral data," *Proc. SPIE Conf. Efficient Transmission Pictorial Inform.*, 1975, vol. 66, pp. 23–35.

[12] K. Maeda and D. G. Goodenough, "Data compression of SAR/MSS data sets," in *Proc. Machine Processing of Remotely Sensed Data Symp.*, p. 188, 1981.

[13] K. C. Gowda, "A new method for mapping multidimensional data to lower dimensions," in *Proc. Int. Symp. Remote Sensing of Environment, 2nd Thematic Conf.* (Fort Worth, TX), Dec. 6–10, 1982, pp. 765–771.

[14] C. Chen and D. Landgrebe, "Spectral feature design for data compression in high dimensional multispectral data," in *Proc. IGARSS '87 Symp.* (Ann Arbor, MI), May 18–21, 1987, pp. 685–690.

[15] C. Mailhes, P. Vermande, and F. Castanie, "Spectral image compression," *J. Optics (Paris)*, vol. 21, no. 3, pp. 121–132, 1990.

[16] J. A. Saghri and A. G. Tescher, "Near-lossless bandwidth compression for radiometric data," *Optic. Eng.*, vol. 30, no. 7, pp. 934–939, July 1991.

[17] B. R. Epstein *et al.*, "Multispectral KLT-wavelet data compression for Landsat thematic mapper images," *Proc. Data Compression Conf.* (Snowbird, UT), Mar. 24–27, 1992, pp. 200–208.

[18] C. S. Kim, M. J. T. Smith, and R. M. Mersereau, "Two-stage multirate coding of color images," *Signal Processing: Image Commun. 2*, pp. 79–89, 1990.

[19] Y. T. Tse *et al.*, "Lossy compression techniques of hyperspectral imagery," in *Proc. IGARSS '90 Symposium*, 1990, pp. 361–364.

[20] D. D. Giusto, "On the compression of multispectral images," in *Proc. IGARSS '90 Symp.*, 1990, pp. 1651–1654.

[21] S. Gupta and A. Gersho, "Feature predictive vector quantization of multispectral images," *IEEE Trans. Geosci. Remote Sensing*, vol. 30, no. 3, pp. 491–501, May 1992.

[22] K. Swaminathan and P. P. Vaidyanathan, "Theory and design of uniform DFT, parallel, quadrature mirror filter banks," *IEEE Trans. Circuits and Syst.*, vol. CAS-33, no. 12, pp. 1170–1191, Dec. 1986.

[23] A. N. Kolmogorov, "On the Shannon theory of information transmission in the case of continuous signals," *IRE Trans. Inform. Theory*, vol. IT–2, pp. 102–108, Dec. 1956.

[24] T. Berger, *Rate Distortion Theory, A Mathematical Basis for Data Compression*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1971.

[25] J. F. Hayes, A. Habibi, and P. A. Wintz, "Rate distortion function for a Gaussian source model of images," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 507–509, July 1970.

[26] H. Cramér, "On the theory of stationary random processes," *Annals Math.*, vol. 41, no. 1, pp. 215–230, Jan. 1940.

[27] M. Kač, W. L. Murdock, and G. Szegö, "The eigenvalues of Hermitian forms," *J. Rational Mech., Anal.*, vol. 2, pp. 767–800, 1953.

[28] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*. Berkeley, CA: Univ. of California Press, 1958.

[29] D. J. Sakrison, "An extension of a theorem of Kač, Murdoch and Szegö to N dimensions," *IEEE Trans. Inform. Theory*, vol. IT-15, pp. 608–610, Sept. 1969.

[30] R. A. McDonald and P. M. Schultheiss, "Information rates of Gaussian signals under criteria constraining the error spectrum," *Proc. IEEE*, vol. 52, no. 4, pp. 415–416, Apr. 1964.

[31] D. J. Sakrison, "The rate distortion function of a Gaussian process with a weighted square error criterion," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 506–508, May 1968.

[32] C. W. Therrien, *Discrete Random Signals and Statistical Signal Processing*. Englewood Cliffs, NJ: 1992.

[33] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part 3*. New York: Wiley, 1971.

[34] S. Nanda and W. A. Pearlman, "Tree coding of image subbands," *IEEE Trans. Image Processing*, vol. 1, no. 2, pp. 133–147, Apr. 1992.

[35] M. Antonini *et al.*, "Image coding using wavelet transform," *IEEE Trans. Image Processing*, vol. 1, no. 2, pp. 205–220, Apr. 1992.

[36] J. Rissanen and K. M. Mohiuddin, "A multiplication-free multialphabet arithmetic code," *IEEE Trans. Commun.*, vol. 37, no. 2, pp. 93–98, Feb. 1989.

[37] H. Gharavi and A. Tabatabai, "Sub-band coding of monochrome and color images," *IEEE Trans. Circuits Syst.*, vol. 35, no. 2, pp. 207–214, Feb. 1988.

[38] J. D. Johnston, "A filter family designed for use in quadrature mirror filter banks," in *Proc. IEEE Int. Conf. ASSP*, Apr. 9–11, 1980, pp. 291–294.

[39] M. J. T. Smith and S. L. Eddins, "Subband coding of images with octave band tree structures," *Proc. IEEE Int. Conf. ASSP*, Apr. 6–9, 1987, pp. 1382–1385.

[40] D. Landgrebe and L. Biehl, *An Introduction to Multispec*. West Lafayette, IN: Purdue Research Foundation, 1992.

[41] P. Lancaster, *Theory of Matrices*. New York: Academic, 1969.

[42] B. N. Parlett, *The Symmetric Eigenvalue Problem*. Englewood Cliffs, NJ: Prentice-Hall, 1980.



Daniel Tretter (S'87–M'95) received B.S. degrees in electrical engineering and mathematics from the Rose-Hulman Institute of Technology in 1987, and the M.S.E.E. degree from Purdue University in 1988. He received the Ph.D. degree in electrical engineering from Purdue University in 1994.

From 1989 to 1990, he was a member of the technical staff at The Aerospace Corporation. He is currently employed at Hewlett-Packard in Palo Alto, CA. His research interests include multiscale image processing and image coding.



Charles A. Bouman (S'86, M'89) was born in Philadelphia, PA. He received the B.S.E.E. degree from the University of Pennsylvania, and the M.S. degree in electrical engineering from the University of California at Berkeley in 1982. From 1982 to 1985, he was a staff member in the Analog Device Technology Group at the Massachusetts Institute of Technology's Lincoln Laboratory. He received M.A. and Ph.D. degrees in electrical engineering from Princeton UnIversity in 1987 and 1989, respectively.

In 1989 he joined the faculty of the School of Electrical Engineering at Purdue University as an Associate Professor. His research interests include statistical image modeling, multiscale processing, and the application of human visual system models to the display, printing, and processing of images.

Dr. Bouman is a member of SPIE and IS&T, and is currently an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING.