Joint Source/Channel Coding of Multispectral Imagery

Arthur L. A. da Cunha[†], Weiler A. Finamore[†], and Eduardo A. B. da Silva[††]

[†] Pontifícia Universidade Católica do Rio de Janeiro
Centro de Estudos em Telecomunicações
Rua Marquês de São Vicente, 255 (CETUC)
22453-900, Rio de Janeiro, RJ - Brazil
{arthur,weiler}@cetuc.puc-rio.br

[††] Universidade Federal do Rio de Janeiro
PEE/COPPE/DEL/EE
Caixa Postal 68504
21945-970, Rio de Janeiro, RJ - Brazil
eduardo@lps.ufrj.br

Abstract—**In this paper we investigate the problem of compressing and reliably transmitting multispectral imagery over binary symmetric channels. A novel source/channel multispectral image coding scheme is proposed. The basic engines of our scheme is a powerful de-correlating transform both in spectral and spatial domains, an optimal bit allocation procedure, block classification and a robust quantizer to cope with the noise introduced by the channel. SNR results show that our coder is very competitive with other approaches when the channel bit error rate is zero, while still attaining robustness to channel errors when the channel is noisy. Furthermore, graceful degradation is also observed as the channel cross increases.**

I. Introduction

Remote sensing satellite (RSS) imagery is of interest for a wide variety of applications including geology, meteorology, military surveillance, cloud recognition and environment monitoring. These images are often transmitted to earth over a bandlimited channel. Such channel imposes limits on the amount of bits that can be reliably transmitted and since multispectral imagery contains large amounts of information, compression is often needed. Two kinds of compression are commonly employed, lossy and lossless. While lossless coding can render images without reconstruction distortion, it often achieves only a small compression ratio (2:1 at best), not enough to solve the problem of compressing and transmitting multispectral data over a bandlimited channel. The second approach can achieve much higher compression ratios but at the price of reconstruction distortion. In the recent years, several algorithms have been proposed for lossy image coding. An example is the recent JPEG2000 standard, which employs the wavelet transform to de-correlate image pixels prior to quantization and entropy coding. Typical RSS images are often composed of several image bands, not only the well-known RGB bands used to represent natural images and efficiently handled, for example, by the JPEG2000 encoder. Therefore a more elaborate approach must be used to exploit not only spatial correlation but also spectral correlation among the various bands. Recently, several ingenious solutions have been proposed to handle the problem of compressing multispectral and hyperspectral images [1], [2], [3], [4], [5]. Most of the schemes proposed in the literature use two main techniques for spectral de-correlation – DPCM and transform coding. The shortcoming with the DPCM approach is that it would perform poorly in a channel polluted with noise since errors would propagate during reconstruction. Transform coding, either the Karhunen Loève transform (KLT) or the faster discrete-cosine transform (DCT) is a better approach when a noisy channel is being considered.

Although a handful of solutions exist for lossy compressing multispectral images, to the authors knowledge none of them considered the problem of transmitting the images over a noisy channel. In contrast, several schemes for compressing natural images to be transmitted over a binary symmetric channel (BSC) have been proposed [6], [7], [8]. Two approaches are often used to ameliorate or even eliminate completely the noise introduced by the channel: (1) *tandem* coding, i.e, compression followed by some form of forward error correction (FEC) and (2) joint source/channel coding where source and channel coding are carried out in a combined fashion.

While the best results are often obtained with the *tandem* scheme, joint source/channel coding schemes are often less wasteful requiring fewer computational and memory resources. Such resources, due to cost constraints, are often limited in RSS applications. This motivates the search for efficient joint source/channel schemes.

Classical examples of source/channel coding are the channel-optimized quantizers, either vector (COVQ) or scalar (COSQ) [9], [10]. The COVQ can be seen as a generalization of the Lloyd's algorithm to take the account the effect of the channel. In practice, the COVQ shrinks the partition so as to minimized the distortion caused by a wrong received index. Although the mean square error (MSE) performance of the COVQ is comparable, and for certain cases even surpasses the *tandem* scheme, it suffers from unpleasant impulsive noise. Such noise appears in images as a salt-and-pepper effect.

In this work we investigate the compression and transmission of multispectral imagery over binary symmetric channels. In particular, we propose a novel joint source/channel coding algorithm for compressing multispectral images and transmitting over a BSC. The main features of the proposed scheme are:

1. Transform coding among bands (DCT or KLT) for intraband de-correlation.
2. Lapped Transform (LT) within each transformed band

for inter-band or spatial de-correlation.
3. Classification prior to quantization to exploit image inherent non-stationarity.
4. A near optimal bit allocation approach based on the steepest descent.
5. Phase Scrambling followed by channel-optimized quantizers to handle the impulsive noise introduced by the channel.

Each of the features of the proposed scheme provide excellent performance thus guaranteeing an overall good performance. In particular, we choose to use high performance LTs given its block nature and its comparable performance to the more memory demanding wavelet transform. This block nature also allows us to use efficient block classification algorithms, thus providing good reconstruction quality for high compression ratios.

The paper is organized as follows. In section II we discuss the proposed scheme highlighting its main features and the overall encoding engine. In sections III-V we present in detail the main features of our coder. Results are presented in section VI and concluding remarks in section VII.

II. Proposed Scheme

Fig. 1 displays the various stages of the proposed codec. The image spectral bands are normalized to unit power prior to encoding so as to maintain equal energy among the various bands. First, the spectral covariance matrix $K_{xx} = E[(x_i - \mu)(x_i - \mu)^T]$ is estimated by

$$\hat{K}_{xx} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu})(x_i - \hat{\mu})^T \quad (1)$$

where x_i denotes the column vector with entries the pixels of equal spatial location from each spectral band and $\hat{\mu}$ is the estimated mean vector of the sample set. The average is thus taken over the available ensemble. Given the estimated covariance matrix the KLT matrix is calculated. The KLT matrix is quantized to 16-bit precision and transmitted as side information. Following KLT transformation across spectral bands, each band is segmented into 8×8 blocks and each block lapped transformed. The transformed image with largest variance is then classified into two classes. The classification map generated is transmitted as side information. The 8×8 matrices containing the standard deviation of coefficients within the same subband in each class for each transformed image are calculated to be used in the bit allocation module. These matrices are also quantized with 16-bit precision and transmitted as side information. After the blocks are assigned to the corresponding classes and the bit allocation matrices calculated by the allocation module, each sequence, formed by coefficients sharing the same subband in a given class, is robust quantized with the quantizer selected according to the number of bits assigned in the allocation matrix for the corresponding class. The side information represents only a small amount of data and can be protected by any FEC code, for example, a simple repetition code just as done in [6].

The robust quantizer works by first scrambling the phase of each sequence and then scalar quantizing the scrambled coefficients. Decoding is simply the reverse process. After the side information is decoded (classification map + KLT matrix + standard deviation matrices + bit allocation matrices), the centroids corresponding to the received indices are phase-descrambled. After de-normalization, an lapped synthesis transform is computed followed by the inverse KLT.

III. Lapped Transforms

Lapped Transform were first introduced to reduce blocking artifacts inherent to block transform such as the DCT. In a LT, each transform block of M samples is computed by a $M \times LM$ linear operator where the LM input samples constitutes the M samples from the current signal block plus $(L-1)M$ samples borrowed from adjacent blocks. For example, with $L = 2$, denoting by P_a and P_s^T (T for matrix transposition) the analysis and synthesis operators respectively, we have that for perfect reconstruction it is required that [11], [12]

$$P_s^T P_a = I \text{ and } P_s^T W P_a = I \quad (2)$$

where W is the one-block-shift operator [12] and I is the identity matrix.

A popular LT for image compression, that satisfies (2) with $P_a = P_s$ is the Lapped Orthogonal Transform (LOT) introduced by Malvar and Staelin in [13], which can be understood as a post-processing of the DCT coefficients. This allows for a fast algorithm, i.e, a fast DCT plus some plane rotations, which adds aproximately 20% to the computational complexity of the DCT. While the LOT reduces significantly the blocking or Gibbs phenomenon it does not eliminate it completely.

In [11] the bi-orthogonal LT is introduced. By allowing different synthesis and analysis operators (i.e, bi-orthogonality), extras degrees of freedom permits better designs. Recently, research efforts have led to a wide variety of new orthogonal and bi-orthogonal LT's filters with comparable performance to the wavelet popular 9/7 filters [14], [11], [15]. For example the Generalized Bi-orthogonal Transform introduced in [14] can attain coding gains of up to 9.96dB. In contrast, a three-stage wavelet decomposition with 9/7 filters has a coding gain of 9.45dB.

Throughout this work we will be using the 8×16 Generalized Bi-orthogonal transform presented in [14], [16] for its good performance versus complexity tradeoff.

IV. Block Classification and Bit Allocation

Block classification is crucial to ensure a good overall performance. Since images have a non-stationary nature, coding the pixels without appropriately classification is often wasteful and better results can be achieved with judicious classification schemes, such as those presented in [17]. We have used the equal mean-normalized standard deviation ratio (EMNSD) classification criterion. The figure of merit in the classification procedure is the block coding gain g_i

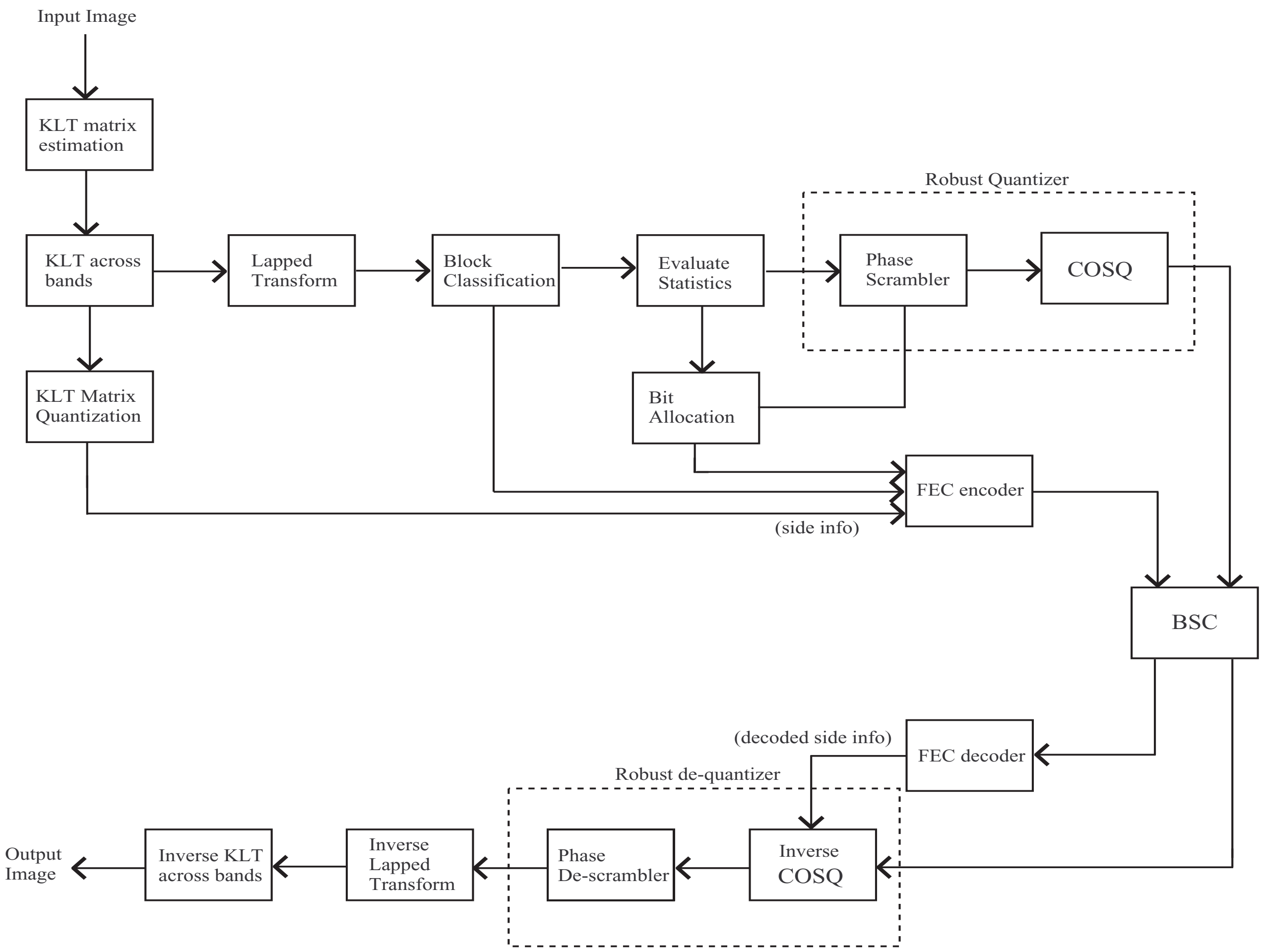


Fig. 1. Proposed scheme building blocks.

which can written as

$$g_i^2 = \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} G_i^2(k,l) - G_i^2(0,0) \tag{3}$$

where the index i denotes one of wh/M^2 blocks of an image with h rows and w columns. The classification algorithm attempts to preserve an equal σ_i/μ_i ratio for $i = 1 \dots J$ and a J number of classes. An algorithm to find the number of blocks N_i, $i = 1 \dots J$ for each class is described in [17]. We have used two classes and one single map for all the spectral band. One map for each band would probably render a better performance, but at cost of an excessive overhead which needs to be protected by a proper FEC code.

The bit allocation is done jointly for both classes. The scheme used here is a similar version of the steepest descent bit allocation algorithm of [18], especially suitable for channel-optimized quantizers. Given a set of N sequences of transform coefficients (we will drop the double index (k,l) for simplicity), the algorithm attempts to find a set of bit-rates $\mathcal{B} = \{b_0, b_1, \dots, b_{N-1}\}$ such that the overall distortion is minimized. The average distortion, as a function of the set of rates, is taken as the average

$$D(\mathcal{B}) = \frac{1}{N} \sum_{i=0}^{N-1} d_i(r_i) \tag{4}$$

where $d_i(r_i)$ denotes the average distortion of the i-th sequence of coefficients as a function of r_i. The distortion D will depend on the codeword assignment as well as on the quantizer levels c_i. Denoting by $\mathcal{I}$ the set of indices produced by the encoder (quantizer), the distortion $d(r)$ for each source, generally denoted by 'x', can be evaluated as follows:

$$d(r) = \sum_{i,j \in \mathcal{I}} \text{Prob}(i,j) \int_{S_i} (x - c_i)^2 dx \tag{5}$$

where $\text{Prob}(i,j)$ denotes the probability that the indices $i, j \in \mathcal{I}$ be transmitted and received respectively. Also, S_i represents the partition in which c_i is the centroid.

The source corresponding to a given b_k belongs to one of the classes $i \in \{1, 2\}$. For each bit in the bit budget, the algorithm searches for the source which most diminishes the overall distortion over all sources and assign a bit for it, until the bits budget is exhausted. The search can be speeded by testing only the adjacent sequences in a giving subband, i.e, for the subband corresponding to position (i, j), the algorithm tests for instance $(i+1, j)$, $(i+1, j+1)$ and $(i, j+1)$ only.

V. Robust Quantization

In order to cope with the errors introduced by the BSC we have used the robust quantization scheme also employed in [6]. Such robust quantizer is highlighted in Fig. 1. The phase scrambling consists of adding a reference pseudo-noise sequence to the phase of each sequence of normalized coefficients. To this end, a Discrete Fourier Transform (DFT) of the sequence is computed and divided into magnitude and phase. After adding the pseudo-sequence to the phase of the DFT, an inverse DFT is computed and the coefficients obtained fed into the COSQ. In the absence of quantization, the process can be reversed by subtracting the pseudo-noise sequence from the phase of the DFT. The seed used in the pseudo-sequence generation must be transmitted, so that the receiver can reverse the process. The phase scrambler has two actions: (1) it re-shapes the histogram of the input coefficients and (2) it spreads the impulsive noise introduced by the channel. Re-shaping the histogram can improve quantization performance since for Generalized Gaussian sources, a scalar quantizer would have better performance for greater shape parameters. It turns that the histogram of the output of the phase scrambler has a Gaussian-like histogram, thus with greater shape parameter with respect to the typically Laplacian LT coefficients. We have noted in our experiments that a typical 1dB improvement over plain scalar quantization for each quantized source can be obtained with the robust quantizer for zero bit error rate.

VI. Results

We have simulated the proposed scheme with four-band, 180-m resolution images captured by a LEO RSS, provided by the Brazilian Space Agency (INPE). The figure of merit in our evaluation is the signal-to-noise ratio written as

$$SNR = 10 \log_{10} \left(\frac{E[x^2]}{MSE} \right) \quad (6)$$

where MSE stands for the average mean square error. We use the SNR instead of common peak-SNR since the latter's results can often be misleading. For instance, we have observed in our experiments reconstructed images with PSNR as large as 40dB, but yet quite different from the original image. Typically, SNR of 16dB yields images almost indistinguishable from the original, while, say 10dB, characterize images with acceptable distortion for a wide variety of applications.

Fig. 2 displays rate-distortion performance of the proposed scheme compressing a typical RSS image in the absence of channel noise. Also in the figure, is the performance of the proposed scheme with the KLT replaced by the DCT. As a benchmark, we use the SPIHT coder applied to each band separately. As can be seen, the proposed scheme outperforms the SPIHT compressor by a wide margin. This margin (of appox. 2dB for rates above 0.5bpp) was also obtained with the 3-D SPIHT encoder proposed in [3], although for a different set of images. We point out that the performance of our scheme as a compressor for noiseless channels can be further improved by the addition of entropy coding.

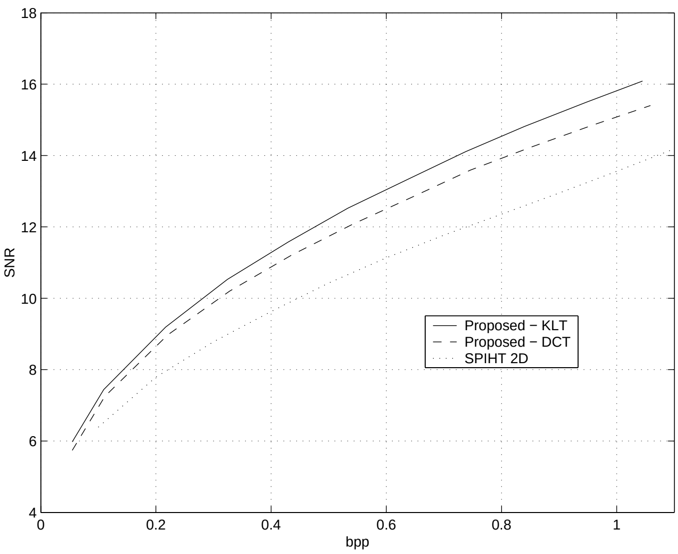


Fig. 2. Rate-distortion performance of the proposed scheme in comparison with the SPIHT compressor compressing each band separately.

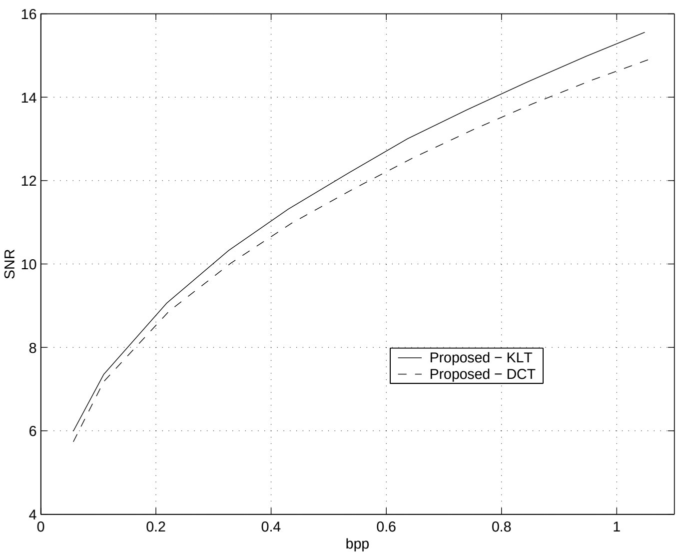
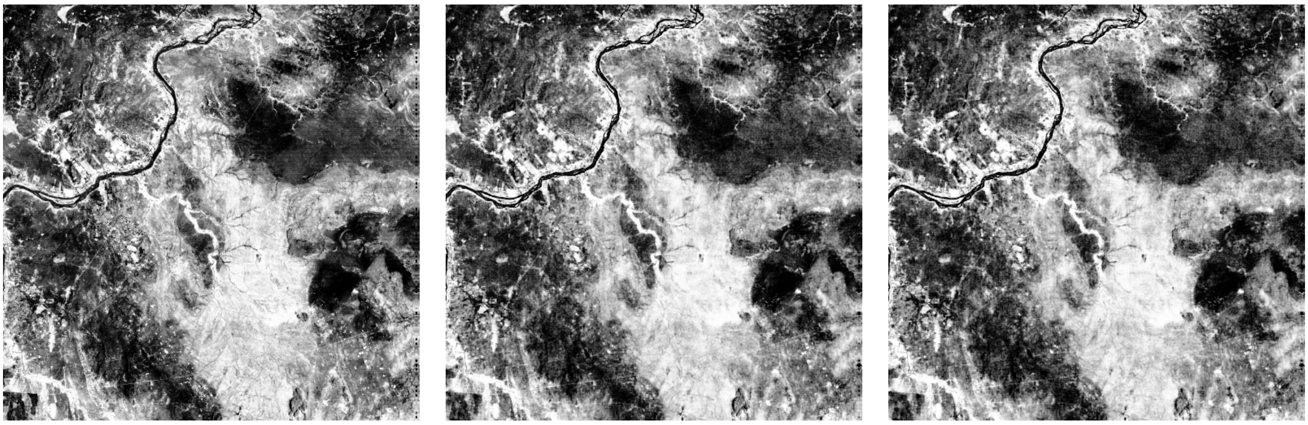


Fig. 3. Rate-distortion performance of the proposed scheme in comparison over BSC with bit error probability $P_e = 10^{-3}$.

In Fig. 3 the rate-distortion performance of the proposed schemes (DCT and KLT) over a BSC with cross probability $P_e = 0.001$ is illustrated (averages taken over ten experiments). As can be seen, the performance of the DCT is



(a) (b) (c)

Fig. 4. Example of typical decoded images. The original image (left) was compressed at 0.54bpp and transmitted over BSC with $P_e = 0.001$ (middle), SNR = 11.82dB and $P_e = 0.01$ (right), SNR=10.90dB.

very close to the KLT when noise is introduced, thus justifying its use when computational resources are limited. Fig. 4 displays examples of decoded images. Only very little artifacts remain on the decoded images. Even for $P_e = 0.01$ the quality is still acceptable for some applications. The robust quantizer alone mitigates the impulsive noise bringing significant perceptual improvement.

VII. Conclusion

We have presented a novel source/channel multispectral robust image compression scheme. The proposed scheme relies on the KLT or DCT for inter-band de-correlation and a high performance LT for intra-band de-correlation. The classification scheme coupled with the steepest descent bit allocation and scalar quantization produce very acceptable distortion even for low bit rates. In addition, when transmitted over the BSC, the encoder is robust producing small degradation for moderate cross probability. This is attained only by the nature of the robust quantizer, without the need of channel coding which considerably increases the complexity and delay. As a bonus, the phase scrambler provides security since it relies on a pseudo-noise sequence which can only be generated by the corresponding seed.

References

[1] G. P. Abousleman, M. W. Marcellin, and B. R. Hunt, "Compression of hyperspectral imagery using the 3-d dct and hybrid dpcm/dct," *IEEE Trans. Geosciences and Remote Sensing*, vol. 33, no. 1, pp. 26–34, January 1995.

[2] John A. Saghri, Andrew G. Tescher, and John T. Reagan, "Practical transform coding of multispectral imagery," *IEEE Signal Proc. Magazine*, pp. 32–43, January 1995.

[3] Pier L. Dragotti, Giovanni Poggi, and Arturo R. P. Ragozini, "Compression of multispectral images by three-dimensional spiht algorithm," *IEEE Trans. Geosciences and Remote Sensing*, vol. 38, no. 1, pp. 416–428, January 2000.

[4] Arto Kaarna, Pavel Zemcik, Heikki Kälviäinen, and Jussi Parkkinen, "Compression of multispectral remote sensing images using clustering and spectral reduction," *IEEE Trans. Geosciences and Remote Sensing*, vol. 38, no. 2, pp. 1073–1082, March 2000.

[5] Giacinto Gelli and Giovanni Poggi, "Compression of multispectral images by spectral classification and transform coding," *IEEE Trans. on Image Processing*, vol. 8, no. 4, pp. 476–489, April 1999.

[6] Qing Chen and Thomas R. Fischer, "Image coding using robust quantization for noisy digital transmission," *IEEE Trans. on Img. Proc.*, vol. 7, no. 4, pp. 496–505, April 1998.

[7] Tuyet-Trang Lam, Glen P. Abousleman, and Lina J. Karam, "Image coding with robust channel-optimized trellis-coded quantization," *IEEE JSAC*, vol. 18, no. 6, pp. 940–951, June 2000.

[8] P. G. Sherwood and K. Zeger, "Progressive image coding on noisy channels," *Proc. IEEE Data Compress. Conf.*, pp. 72–81, March 1997.

[9] N. Farvardin, "A study of vector quantization for noisy channels," *IEEE Transactions on Information Theory*, vol. 36, no. 4, pp. 799–809, July 1990.

[10] N. Farvardin and V. Vaishampayan, "Optimal quantizer design for noisy channels: An approach to combined source-channel coding," *IEEE Trans. Information Theory*, vol. 33, no. 6, pp. 827–838, November 1987.

[11] H. S. Malvar, "Biorthogonal and nonuniform lapped transforms for transform coding with reduced blocking and ringing artifacts," *IEEE Trans. Signal Processing*, pp. 1043–1053, April 1998.

[12] Henrique S. Malvar, *Signal Processing with Lapped Transforms*, Artech House, Boston, MA, USA, 1992.

[13] H. S. Malvar and D. H. Staelin, "The lot: Transform coding without blocking effects," *IEEE Trans. on Signal Processing*, vol. 37, no. 4, pp. pp. 553–559., April 1989.

[14] T. D. Tran, R. L. de Queiroz, and T. Q. Nguyen, "Linear phase perfect reconstruction filter bank: lattice structure, design, and application in image coding," *IEEE Trans. on Sig. Proc.*, vol. 48, pp. 133–147, Jan. 2000.

[15] R. L. de Queiroz, T. Q. Nguyen, and K. R. Rao, "The GenLOT: Generalized linear-phase lapped orthogonal transforms," *IEEE Trans. on Sig. Proc.*, vol. 44, pp. 497–507, April 1996.

[16] Ricardo L. de Queiroz and T. D. Tran, "Lapped transforms for image compression," in *The Handbook on Transforms and Data Compression*, K. R. Rao and P. Yip, Eds., chapter 5. CRC Press, Oct. 2000.

[17] R. Joshi, H. Jafarkhani, J. Kasner, T. Fischer, N. Farvardin, M. Marcellin, and R. Bamberger, "Comparison of different methods of classification in subband coding of images," *IEEE Trans. on Img. Proc.*, vol. 6, no. 11, pp. 1473–1486, 1997.

[18] V.A. Vaishampayan and N. Farvardin, "Optimal block cosine transform image coding for noisy channels," *IEEE Transactions on Communications*, vol. 38, no. 3, pp. 327–336, March 1990.