MULTISPECTRAL IMAGE COMPRESSION ALGORITHMS [1]

Tassos Markas[2]
Department of Electrical Enginnering,
Duke University,
Durham, NC 27706

John Reif
Department of Computer Science,
Duke University,
Durham, NC 27706

Abstract

In this paper we present a data compression algorithm that is capable of significantly reducing the vast amounts of information contained in multispectral and hyperspectral images. The loss of information ranges from a perceptually lossless level, achieved at 20-30:1 compression ratios, to a level that the exploitation of the images is still possible (over 100:1 ratios). The developed algorithm takes into consideration the spectral and spatial correlations found in multispectral images, and is capable of surpassing the compression/distortion performance of other reported methods. The proposed scheme utilizes a one-dimensional transform coder to remove the spectral redundancy, and a two-dimensional wavelet transform to remove the spatial redundancy of multispectral images. The transformed images are subsequently divided into active regions that contain significant wavelet coefficients. Each active block is then hierarchically encoded using multidimensional bitmap trees. We will also show that application of reversible histogram equalization methods on the spectral bands can significantly increase the compression/distortion performance. We will conclude by presenting the performance of the proposed algorithm using Landsat Thermatic Mapper data.

1 Motivation

Multispectral imaging satellites, such as LANDSAT and SPOT, which measure luminance in ten or fewer spectral bands, have been in operation for nearly twenty years. However, the trend in remote sensing is toward high resolution hyperspectral (more than 100 spectral bands) imaging systems, also known as imaging spectrometers. Current technology allows the development of high spatial/spectral resolution sensors, capturing images in up to 200 discrete spectral bands. The utility of these devices can be significantly impaired by a lack of adequate transmission bandwidth. Existing transmission resources can accommodate only a small percentage of the potentially available information. This would result in a drastic and unsophisticated form of data compression since only a small percentage of the captured data can be collected and analyzed by scientific researchers. This is true for telemetry and dissemination of data to numerous remote scientific users. Some form of image compression must be applied if the full scientific mission objectives of these advanced sensors are to be fully realized. In this paper we will give a brief overview of existing multispectral

[1]This work has been supported by the Center of Excellence in Space Data and Information Sciences under NASA/CESDIS NAS USRA subcontract 550-63 of prime contract NAS5-30428. Additional support has been provided by DARPA/ISTO Contracts N00014-88-K-0458, N00014-91-C-0114, N00014-91-J-1985, and NSF Grants 88-00282/2, NSF-IRI-91-00681

[2]Also with Center for Systems Engineering, Research Triangle Institute, RTP, NC 27709

Preprint of paper appearing in Proceedings: IEEE Data Compression Conference (DCC'93), Snowbird, UT, IEEE Computer Society Press, James A. Storer, Martin Cohn (Eds.), p.391-400, March 1993.

image compression techniques and we will present two new methods that advances the compression technology in this area.

2 Introduction

Multispectral sensors measure electromagnetic radiation in different spectral bands, including bands outside the visible spectrum. Data generated by those instruments contain a high degree of redundancy among spectral bands, due to both the characteristics of the sensing mechanism and the imaging environment. Compression algorithms for video data can achieve extremely high compression ratios that can be accomplished by removing the temporal redundancies among successive frames. Although spectral redundancy is significantly different than temporal redundancy, the main principle remains the same: only a limited portion of the spatial information changes among bands. For video data the dissimilarities are movement of objects which are functions of time, whereas in multispectral images the dissimilarities among bands are different intensities of same object. Multispectral sensors measure the reflectivity of a particular object in a given frequency window. Therefore, objects that absorb radiation in a given spectrum will appear as dark, whereas objects that reflect radiation in the same spectrum will appear as bright. The static nature of those dissimilarities between different spectral bands makes multispectral information very compressible.

A considerable amount of research in the area of lossy multispectral compression has been published over the past three years. Predictive coding, block truncation coding, Discrete Cosine Transform, and various forms of Vector Quantization (VQ) methods have been evaluated in [1]. From all those methods the mean-residual and the gain-shape VQ experienced better compression/distortion performance. An integrated system that is capable of compressing data with loss (for browsing applications), moderate loss and lossless has been reported in [2, 3]. This algorithm includes block averaging, quad-trees, and an iterative region growing scheme to form low variance regions that can be effectively represented by one value. Other multispectral VQ methods include the feature predictive and polynomial predictive VQ [4, 5] where a small subset of the image bands is encoded using VQ and the remaining bands are predicted (using linear and non-linear predictors) from the quantized ones. A KLT-wavelet scheme that utilizes transform and wavelet coding showed that 80:1 compression ratios are possible for browsing type applications [6]. Finally, vector quantization can be easily adapted to three dimensions to exploit the redundancy between different spectral bands [7]. In the first section of this paper we will present a hybrid algorithm that enhances the transform/subband coding model, described in [6], using histogram equalization to preprocess the multispectral images and a novel multidimensional bitmap encoding scheme to efficiently compact the significant transform/subband coefficients. We will refer to this algorithm as Hybrid Transform/Subband scheme. This algorithm achieves perceptually lossless compression at ratios between 20:1 and 30:1, and browsing level compression at ratios over 100:1.

3 Hybrid Transform/Subband Algorithm

The basic building blocks of the proposed algorithm, as shown in Figure 1, is a transform coder that removes the spectral redundancy, and a subband coder that removes the spatial redundancy of the spectral bands. The dynamic range of the spectral

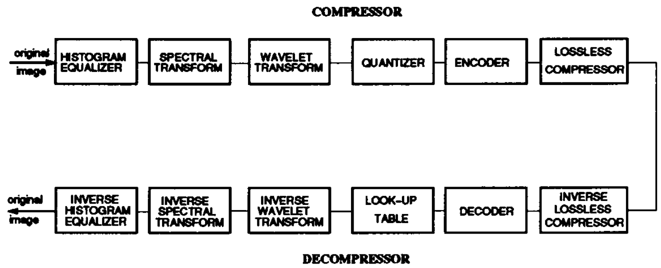

Figure 1: Hybrid Multispectral compression system

bands has been previously modified using histogram equalization. The purpose of the histogram equalizer is to minimize the variance along the spectral domain so that the spectral information can be efficiently compressed by the transform coder. The resulting images are then quantized and encoded using a novel scheme that partitions the image in active and non-active areas and encodes them hierarchically using multidimensional trees. Active areas contain high amplitude coefficients that will be used by the decoder to reconstruct the original information. The location along with the amplitude of the significant coefficients is also encoded using multidimensional trees. Finally, a lossless entropy coder removes any remaining redundancy from the encoded data stream.

3.1 Removal of Spectral Redundancy

One-dimensional transform coding will be used along the spectral domain to remove the redundancy of the spectral information. Transform coding is a lossy coding technique, where a unitary transform is applied to uncorrelate the data. In transform coding, the transform subprocess is responsible for transforming the spectral information to a different domain of representation, where a large fraction of the image information tends to reside in relatively few coefficients. Some examples of methods include the Discrete Cosine Transform (DCT) [8], and the optimal for stationary random sources Karhunen-Loeve transform (KLT) [9]. Figure 2a shows the intensity of a typical multispectral image block, containing 4x4 adjacent pixels, as a function of the spectral band. This figure shows that spatial and spectral information is highly correlated. However, spectral correlation is not exploitable in its present form. This can be seen from Figure 2b where the block shown Figure 2a has been transformed to the frequency domain using a one-dimensional KLT transform. In this case, a large number of high amplitude coefficients is required to accurately reconstruct the spectral information. The high variance of the spectral information is a result of the different dynamic range of each band. Figure 3 shows the histograms, which are often viewed as scaled Probability Density Functions (PDF), of a typical multispectral image. This Figure shows that the spectral information of each band is located in different regions of the intensity domain. This is the cause of the rather high variation of the spectral information that we saw in figure 2a. Exploitation of spectral correlation can be achieved using histogram equalization of all spectral bands.

Histogram equalization/modification is a non-linear process that equalizes the PDF $p(f)$ of a random variable f with the PDF $p_d(g)$ of a desired random variable g. Histogram modification is accomplished by identifying a monotonically nondecreasing transformation $g = T[f]$ that equalizes the cumulative distribution functions of the

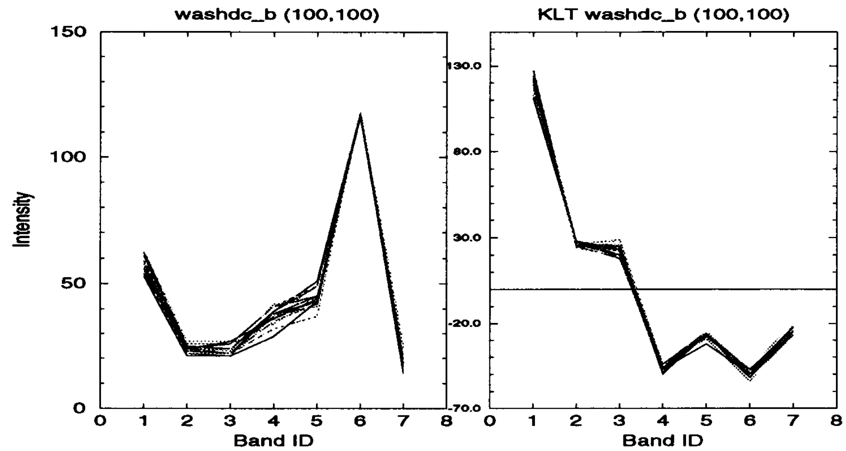

Figure 2: Spectral information as a function of the band (a), and KLT-transformed version (b)

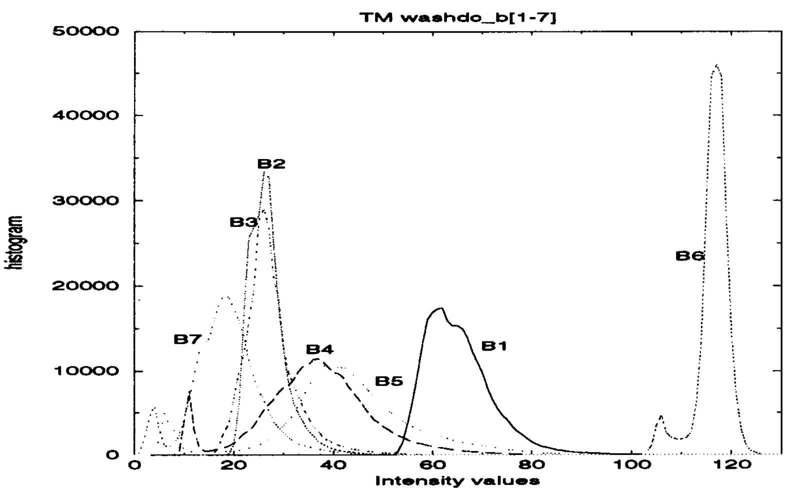

Figure 3: Histogram of the "washdc_b[1-7]" image

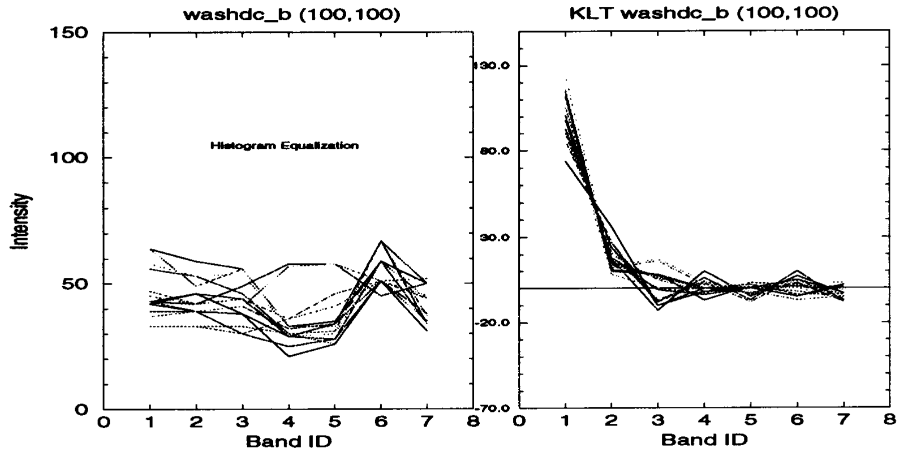

Figure 4: Histogram modified spectral information (a), and its KLT-transformed version (b)

two images [10] and applying it to the image that needs to be modified.

$$P(f) = \sum_{i}^{f} p(i) \quad equalized\ with \quad P_d(g) = \sum_{i}^{g} p_d(i)$$

The monotonically nondecreasing property of $T[f]$ guarantees that a pixel with a higher intensity in the original image will not become a pixel with lower intensity in the transformed image. In general, histogram modification is an irreversible process since there are cases where $T[f]$ is a many-to-one mapping function. By posing several constraints to the equalizing process we can define a reversible transformation that can be used to reconstruct the original image without any loss of information. The necessary constraints to achieve this is to define a monotonically increasing transformation so that two adjacent intensity values will not map at the same intensity of the transformed image, and to select a target image with a wider PDF than the source image. This way, we can guarantee that there will not be any many-to-one mapping. This will prevent information from being lost during the inverse equalization process. Figures 4a,b show the histogram modified versions of the same multispectral image block shown in Figures 2a,b. Figure 4b shows that the spectral information is mostly contained in the DC component of the spectral information. This representation can be more efficiently encoded compared to the one shown in Figure 2b.

Another advantage of the histogram equalization method is its ability to mask quantization errors in bands with narrow dynamic range. This occurs since a narrow band will be expanded during histogram equalization, resulting to a one-to-many mapping. Compression will then be performed using a representation with wider dynamic range. During reconstruction, this new representation can tolerate more quantization error since the distorted value may map in the original one.

3.2 Removal of Spatial Redundancy

After removing the spectral redundancy, the image consists of the same number of bands that contain the spectral information transformed at the frequency domain. Subband coding is then used to remove the spatial redundancy of each band. Subband coding has been a popular method of signal compression for image data [11]. During subband coding the signal is decomposed in different spectral bands which are then encoded separately using scalar or vector quantizers. Besides its good coding performance, subband coding gets excellent subresolution signals which are very useful for browsing large amounts of information. Subband coding into octave band channels can be seen as a Discrete Wavelet Transform [12, 13]. The DWT is a set of two finite impulse response filters of very short length, which are applied recursively to low frequency subresolution components. DWT can be implemented using the well known pyramid interconnect structure [14], where the lower levels of the pyramid contain high resolution data and the upper levels of the pyramid contain low resolution data.

Because of its compression performance and its excellent subresolution signals, the wavelet transform was chosen to remove the spatial redundancy of the transformed spectral bands. The first band that contains the high-energy DC-component of the spectral information is compressed at moderate rates to minimize the overall loss of information. The rest of the bands are compressed at rates that reflect their information content. Bands with high average energy are encoded at relatively low compression ratios, whereas bands with low energy are compressed at high ratios.

The transformed images are then fed to the quantizer. Uniform scalar quantization with different quantization constants was used at each level of the multiresolution pyramid. The upper levels of the multiresolution pyramid that contain important DC components of the wavelet transform are encoded using a fine quantizer, whereas the lower levels that contain high frequency components are encoded using a coarse quantizer.

3.3 Hierarchical Encoding of the Wavelet Coefficients

As was mentioned earlier, the transformation process preserves the image information in relatively few coefficients from which the original image can be efficiently reconstructed. Those coefficients tend to reside in regions where there is a significant spatial information. We will refer to those regions as "active blocks". A $x \cdot y$ block is considered active, if it contains at least one significant coefficient. Within each active block there is a number of non-zero wavelet coefficients whose amplitude along with their location with respect to the origins of the current active block needs to be encoded. Traditionally, Run-Length Encoding (RLE) has been used to encode the locations and the amplitudes of those significant coefficients. Here, we propose a different approach that utilizes multidimensional bitmap trees to either encode the locations of the active (non-zero) or the non-active (zero) coefficients within each active block. Bitmap trees are N-dimensional tree structures that partition a two-dimensional block in N^M regions, where M is the maximum depth of the tree. At each level i, the block is divided in N^i regions of size $x \cdot y/N^i$. Each node in the tree is represented by a codeword of size N indicating which subregions of the lower level contain at least one active (non-active) coefficient. Figure 5 shows a 16x16 block encoded using a quad-tree representation (N=4, M=4).

The efficiency of the multidimensional trees depends on the number of the active locations within a block. In case of sparse blocks, the described approach results in a very low bit overhead. However, in cases where there is a high percentage of active

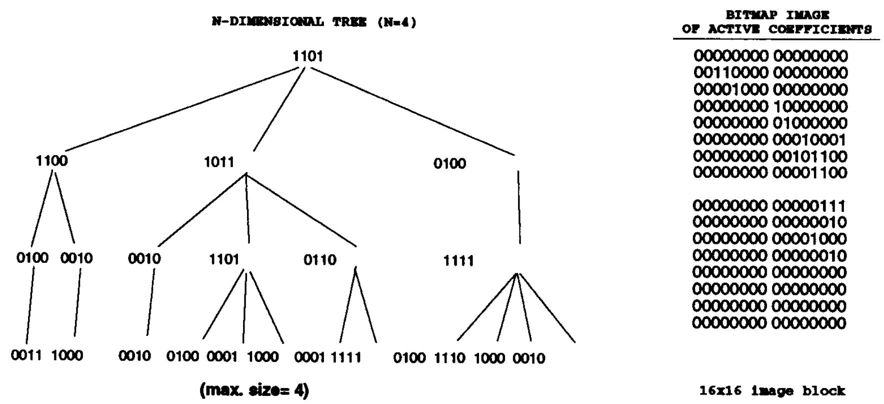

Figure 5: Multidimensional Bitmap tree

locations, the proposed approach is very inefficient. For this reason, we have devised a modified bitmap encoding that takes into consideration the percentage of the active locations within each active block. For dense blocks, where only a few locations are zeros (upper levels of the multiresolution pyramid), we encode the non-active locations instead of the active locations. Another feature of this modified scheme is that the same tree can be used to simultaneously encode active and non-active regions. For this reason an additional codebit is included in the upper levels of the tree to indicate whether active or non-active encoding will be used in all children nodes. The advantage of the multidimensional trees is that they hierarchically decompose the original image block in subblocks that contain information in a highly correlated form. For this reason, they utilize less bits to encode a given block of information compared to the RLE schemes. Simulation results of typical images have shown that independently of the percentage of active locations within a block the bitmap tree approach requires 5-10% less bits compared to the RLE scheme. It should be also noted that this savings is accomplished without significantly increasing the computational overhead.

As was mentioned earlier, the coefficients of the transformed image are enclosed in a finite set of active blocks whose locations, with respect to the original image, needs also to be encoded. The location of the active blocks within the transformed image can be also represented as a $(X/x)\cdot(Y/y)$ bitmap image, where X,Y are the dimensions of the original image, and x,y are the fixed dimensions of the image blocks. This image is also encoded using multidimensional bitmap trees. Bitmap trees can exploit the intra-level and intra-band correlation of the active blocks in multiresolution pyramids. In a multiresolution pyramid, if a block at a given level is inactive, then there is a very good possibility that all children blocks at the lower levels of the pyramid will also be inactive. The modified bitmap tree approach that was described earlier is particularly useful in this case since the lower levels of the pyramid structure contain mostly non-active blocks, whereas the upper levels of the pyramid contains mostly active blocks.

After applying transform/subband coding, quantization, and hierarchical encoding to the original images, the encoded information consists of the following components:

- Encoded bitmap of active/non-active blocks

- Encoded one-zero bitmap that represents which blocks are compressed using "zero" bitmap encoding, and which using "one" bitmap encoding
- Bitmap codes of all active blocks indicating the location of the significant coefficients within each block
- Quantized wavelet coefficients

Even this information contains a significant amount or redundancy. The probability density function of the quantized coefficients tend follow a Gaussian distribution. This means that their entropy is not maximized, which suggests that they can be further compressed using entropy coding. We removed this redundancy using entropy coders that replace data symbols with variable-length codes according to their frequency occurrence. The specific algorithm that was used for this purpose was the classic Huffman coding [15]. It should be noted that use of more efficient entropy coding methods, such as arithmetic [16] and Q-coders [17], is expected to increase even more the compression performance at a small computational cost. The bitmap codes also contain some redundancy that was removed using Huffman coding. The active/non-active and the one/zero bitmaps were fairly compact, and they were left at their present compressed form. A header containing information such as the size of the image, the wavelet taps, the quantization constants for each band, etc., was also included in the compressed file.

4 Experimental Results

The 7-band "washdc_b" Thermatic Mapper image was used to evaluate the performance of the developed algorithm. The size of the image was 512x512, 8 bits deep, and the distortion measure that was used to evaluate the loss of information was the Peak Signal-to-Noise Ratio. To remove the correlations along the spectral domain we used the KLT transform obtained from the covariance matrix of the first order stationary Markov sequence. The wavelet transform was implemented using five multiresolution levels that were separated using a 9-tap Quadrature Mirror Filter (reflection was used at the edges). The block size was set at a fixed value of 16x16 pixels, and a fixed size multidimensional bitmap tree was used to encode the locations of the significant coefficients. The size of the bitmap tree was set to four which is equivalent to a quad-tree decomposition.

The following tables show the compression performance of the "washdc_b" image at various rates. Table 1 shows the compression results without histogram equalization, whereas Table 2 shows the compression results with histogram modification. In the latter case, the bands were equalized with respect to the fifth spectral band which has the widest probability density function (Figure 3). It should be also noted that the distortion was measured against the original images and not against their equalized versions.

The results show that histogram equalization approach results in a much higher compression/distortion performance. The most of this gain is the improved quality of bands with narrow dynamic range. As was mentioned earlier, histogram equalization of bands with narrow range can tolerate more quantization error. This indicates that histogram equalization of all spectral bands to their full dynamic range may result in even higher performance. In addition the results showed perceptually lossless image quality at 20-30:1 compression ratios. At browsing quality the developed algorithm was able to compress the information at rates exceeding 100:1. The same experiments were also conducted for the "neworleans" image. The compression performance was

Compression ratio	Distortion (PSNR)							
	B1	B2	B3	B4	B5	B6	B7	AVERAGE
103.10:1	35.366	36.951	35.384	32.894	31.048	36.214	36.147	34.858
82.22:1	35.362	36.430	35.086	32.949	31.373	35.214	36.266	34.669
55.19:1	36.842	39.270	36.602	34.110	31.925	37.618	37.446	36.259
37.90:1	36.091	37.030	35.905	33.803	32.282	35.090	37.316	35.360
33.87:1	38.056	40.588	37.268	35.395	33.191	37.722	38.482	37.243
16.31:1	39.997	41.571	37.927	37.719	34.981	38.004	39.923	38.589

Table 1: Compression/Distortion performance of "washdc" image

Compression ratio	Distortion (PSNR)							
	B1	B2	B3	B4	B5	B6	B7	AVERAGE
154.20:1	35.117	39.740	36.929	31.144	29.901	43.805	35.618	36.036
97.17:1	36.205	41.012	37.756	33.038	31.111	45.213	36.550	37.269
91.47:1	36.314	41.144	37.880	32.887	31.319	45.441	36.678	37.381
74.19:1	36.583	41.517	38.236	33.884	31.625	45.217	36.948	37.751
41.48:1	38.069	42.952	39.329	34.253	32.694	45.366	37.908	38.653
26.47:1	39.097	44.248	40.342	35.616	33.483	46.120	38.683	39.656
16.44:1	40.510	45.769	41.868	37.548	34.738	46.769	39.882	41.013

Table 2: Compression/Distortion performance of "washdc" image using histogram equalization

slightly worse. Perceptually lossless image quality was achieved at ratios between 15:1 and 24:1.

One disadvantage of this algorithm, which is also present in most of the existing multispectral compression algorithms, is that the distortion is not equally distributed among all spectral bands. The above tables show that the majority of loss occurs in bands with high spatial variation (bands four and five in this case). A solution to this problem is to apply first the subband to remove the spatial redundancy, and then the transform coder to remove the spectral redundancy. This way, we can better control the maximum amount of information that is lost in each band.

5 Conclusions

In this paper, we have initially described a hybrid compression algorithm that is capable of compressing multispectral data at higher distortion/compression performance compared to other existing methods. Simulation results have shown that compression ratios of 100:1 for browsing type applications, and 20-30:1 for perceptually lossless distortion are possible. The execution time of the encoding process depends on the compression ratios and varies between 20 and 53 CPU seconds measured on a SUN 4/75 workstation. The decoding process requires approximately the same time. In addition we have offered some ideas that can improve the performance of this algorithm.

References

[1] R.L.Baker, and Y.T. Tse, *Compression of High Spectral Resolution Imagery*, Proc. of SPIE: Applications of Digital Image Processing, Aug. 1988, pp. 255-264.

[2] J.C.Tilton, *Hierarchical Data Compression: Integrated browse, moderate loss, and lossless levels of data compression*, Proc. of Intern. Geoscience and Remote Sensing Symposium, 1990.

[3] J.C. Tilton and M. Manohar, *Compression Experiments with AVHRR data*, Proc. of 1991 Data Compression Conference, Apr. 1991, pp. 411-420.

[4] S. Gupta, and A. Gersho, *Feature Predictive Vector Quantization of Multispectral Images*, Trans. on Geoscience & Remote Sensing, Vol.30, No.3, May 1992, pp.491-501.

[5] D.D.Giusto, *On the Compression of Multispectral Images*, Proc. of Intern. Geoscience and Remote Sensing Symposium, 1990, pp.1651-1654.

[6] B.R. Epstein, R. Hingorani, J.M. Shapiro, and M. Czigler, *Multispectral KLT-Wavelet Data Compression for Landsat Thematic Mapper Images*, Proceedings of Data Compression Conference (DCC'92), Mar. 1992, pp. 200-208.

[7] M. Manohar and J.C. Tilton, *Progressive Vector Quantization of Multispectral Image Data using a Massively Parallel SIMD Machine,* Proc. of the 1992 Data Compression Conference, Mar. 1992, pp. 181-190.

[8] N. Ahmed, T. Natarajan, and K.R. Rao, *Discrete Cosine Transform*, IEEE Trans. Comput., Vol. C-23, Jan. 1974, pp. 90-93.

[9] W.D.Ray, and R.M.Driver, *Further Decomposition of the Karhunen-Loeve Series Representation of a Stationary Random Process*, IEEE Trans. on Info. Theory, IT-11, Nov. 1970, pp. 663-668.

[10] J.S. Lim, *Two-Dimensional Signal and Image Processing*, Prentice Hall, Englewood Cliffs, NJ, 1990, pp. 453-459.

[11] M. Vetterli, *Multi-dimensional Sub-band Coding: Some Theory and Algorithms,* Signal Processing, Vol. 6, No. 2, pp. 97-112, Feb. 1984.

[12] O. Rioul and M. Vetterli, *Wavelets and Signal Processing,* IEEE Signal Processing Magazine, Vol. 8, No. 4, Oct. 1991, pp. 14-38.

[13] S.G. Mallat, *Multifrequency Channel Decompositions of Images and Wavelet Models*, Trans. on Accoust. Speech, and Sig. Proc., Vol.37, No.12, Dec. 1989, pp. 2091-2110.

[14] P.J. Burt and E.H. Adelson, *The Laplacian Pyramid as a Compact Image Code,* IEEE Trans. on Commun., Vol. 31, No. 4, Apr. 1983, pp. 532-540.

[15] D.A. Huffman, *A Method for the Construction of Minimum Redundancy Codes*, Proc. of IRE, Vol. 40, Sept. 1952, pp. 1098-1101.

[16] I.H. Witten, R.M. Neal, and J.G. Cleary, *Arithmetic Coding for Data Compression*, Commun. of ACM, No. 6, Jun. 1987, pp. 520-540.

[17] Special issue of the IBM Journal on Research and Development on Q-coder adaptive binary arithmetic coder, Nov. 1988.