

HYPERSPECTRAL IMAGE CODING USING 3D TRANSFORMS

Dmitry Markman and David Malah

Technion-Israel Institute of Technology
Department of Electrical Engineering
Technion City, Haifa 32000, Israel

Email: dmarkman@tx.technion.ac.il and malah@ee.technion.ac.il

ABSTRACT

This work considers the efficient coding of hyperspectral images. The *shape-adaptive* DCT is extended to the three-dimensional case. Both the 3D-SA-DCT and the conventional 3D-DCT are combined with either of two alternative techniques for coding the transform coefficients. The proposed schemes are compared with two state of the art coding algorithms, which serve as benchmarks, and are found to have better performance (in terms of PSNR). Particularly good performance is shown by the coder based on the 3D-SA-DCT in a classification task applied to the coded images.

1. INTRODUCTION

A hyperspectral image cube (HIC) is a set of hundreds of images, where each image in the set is the result of sensing the same scene at a different spectral band (i.e., at different wavelengths). It can be also defined as a 2D array of "spectral voxels". Color images are a particular case of this type of images, where only three spectral regions are analyzed. Modern remote sensing systems are able to collect more than 200 visible and near-infrared wavebands. Such images cover in great detail a wide spectral window and can provide valuable information about the land cover of the area under investigation. Applications dealing with hyperspectral imagery usually relate to spatial and spectral information by means of image analysis (classification, material detection etc.).

A major issue in hyperspectral imaging is the huge amount of data to be transmitted and/or stored. Such images usually have both high spatial resolution and high spectral resolution (a large number of bands). For example, a hyperspectral image cube of size 2000x2000 pixels containing 200 spectral bands, with 8 bits per pixel, occupies 763Mbytes! To address this problem data compression must be applied.

Hyperspectral image compression algorithms must cope with high compression ratios and sufficient reconstruction fidelity in terms of the subsequent analysis needed for a given application (for example, material classification). It is obvious that employment of standard methods for still images compression, like JPEG, would not meet such demands. Abousleman *et al.* in [1] applied hybrid DPCM/DCT and a three-dimensional DCT (3D-DCT) followed by entropy-constrained trellis-coded quantization (ECTCQ) schemes, while Canta and Poggi in [2] proposed a Kronecker product Gain-Shape VQ (KP-GS-VQ) algorithm for hyperspectral image compression. These techniques take advantage of both spatial (intra-band) and spectral (inter-band) redundancies, inherently existing in hyperspectral imagery, thus achieving good results, while requiring only a moderate coding complexity.

In an attempt to improve on those schemes results, we examined a coding scheme which is based on the 3D-DCT but employs a *shape-adaptive* coding approach to high activity blocks. The new scheme is termed 3D-SA-DCT and is described in Section 2. Two different algorithms for quantizing and efficient coding of the resulting transform coefficients are introduced in Section 3. Simulation results of the proposed scheme, in comparison to those obtained by the two benchmark algorithms: KP-GS-VQ and the hybrid DPCM/DCT coder with TCQ, are presented in Section 4. Section 5 concludes the paper.

2. PROPOSED 3D-SHAPE-ADAPTIVE DCT-BASED SCHEME

Transform coding techniques, and in particular the DCT, are very attractive for image coding, mainly due to their good energy compaction characteristics. Since in a HIC both spatial and spectral redundancies are present, a 3D-DCT is a natural choice. In the following discussion a block refers to a 8x8x8 spatio-spectral cube of pixels from the original HIC.

Although the 3D-DCT has good energy compaction characteristics and the encoding of the transformed blocks is usually efficient, it still fails in compacting the energy of high activity blocks. This is mainly due to the fact that high activity in a spatio-spectral block (because, for example, an edge in the block) results in relatively high AC coefficients - even at high frequencies. These coefficients either decrease the compression ratio that can be achieved or, if quantized coarser, reduce the quality of the reconstructed bands. The main idea in the proposed scheme is to determine these high activity blocks and encode them separately in an efficient manner, while the remaining blocks are to be encoded using the conventional 3D-DCT. The following subsections describe the building blocks of the proposed scheme.

2.1. Determination of High Activity Blocks

High block-activity may result from two sources. Either the particular block contains an edge or it contains coarse texture. Since we are dealing with three-dimensional blocks, the variations in pixel values may occur in all three dimensions. We have chosen the value of the block variance as a measure of its activity level. Thus, high activity blocks are determined by comparing the variance of each block to a threshold value. This approach, although simple, was found in simulations to provide better results than gradient-based criteria. Since we do not have a model for the variance of blocks in HICs, threshold values are set empirically.

2.2. Block Splitting Algorithm

The main idea in the proposed scheme is to split each high activity block into a number of low activity regions with smaller variations in pixel values. Applying the shape-adaptive DCT to each one of such regions is expected to result in more efficient coding of transform coefficients. In order to decrease the amount of encoded information (one SA block of transform coefficients is produced for each region) and to reduce the partitioning side-information, each high activity block was split just into 2 lower activity regions.

Since hyperspectral imagery has high spectral resolution, the variance along the spectral axis in a block (just 8 bands) is not expected to be high. On the other hand, if two adjacent voxels are from different materials (i.e., having different spectral signatures), they typically will have quite different intensity values, thus causing large spatial variations. The problem of clustering the 64 voxels (of length 8, each) in a block into two clusters, is solved by applying a 2-class *Vector Quantization* (VQ) clustering algorithm, using the well known LBG algorithm. Each class corresponds to a different spectral signature in the highly active block. The resulting partition provides for each block a 8x8 segmentation bitmap in which entries corresponding to the voxels belonging to one of groups are assigned '0' and the others are assigned '1'. Such a segmentation and its effect on the way the shape-adaptive DCT is performed, are illustrated in Figure 1.

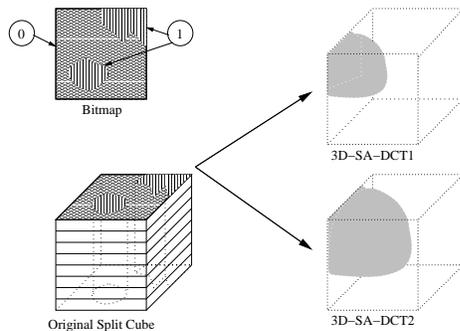


Fig. 1. Block of size 8x8x8 is split by a 2-class VQ: segmentation bitmap and cube splitting (left) and 2 sets of 3D-SA-DCTs (right).

2.3. Lossless Coding of Side Information

To reduce the amount of side information, each 8x8 segmentation bitmap has to be encoded. Lossless compression is required since the decoder needs the exact partition map of voxels in order to reconstruct correctly the data and to perform the corresponding 3D-SA-IDCT for the two regions.

We examined in our work four techniques for encoding the binary data: Fixed and variable run-length coding, followed by Huffman encoding tables; Elias coding, and block-based Huffman coding. In addition, different scanning techniques were used to convert each 8x8 binary bitmap to a 64 element vector of zeros and ones: Hilbert scan and 2 different “snake” scans (horizontal/vertical) as depicted in Figure 2, allowing also block-adaptive scan direction.

In our experiments, the preferred technique for lossless bitmap coding was found to be the variable run-length Huffman coding

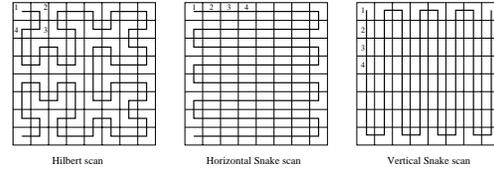


Fig. 2. Hilbert, horizontal and vertical Snake scans.

method preceded by a horizontal snake scan. It achieved an average bitmap compression ratio of 2 (i.e., 32 code bits per block). More details can be found in [3]. Note that the segmentation bitmap describes the partition of the 8x8x8 block into two classes, hence the side-information rate is $32/512=1/16$ bit/pixel.

2.4. Applying a 3D Shape-Adaptive DCT to Split Cubes

Shape-adaptive 2D-DCT has been proposed for image coding as a way to improve the coding efficiency of blocks on object boundaries [4]. Here we extended this technique for coding 3D data of cube partitions of arbitrary 3D shapes. In our particular case, the partition is defined by a 2D bitmap and hence is uniform along the spectral axis (see Figure 1), simplifying the SA-DCT in the last stage to just a 1D-DCT of fixed size (8). In the other stages, the 2D-SA-DCT is implemented via 1D-DCT's of variable lengths, as explained in [3]. The following important features of the 3D-SA-DCT are worth mentioning:

- Denoting the discrete transform coordinates by (u, v, w) (each in the range of 0 to 7), the DC coefficient is located at the upper left corner ($u = v = w = 0$).
- The 3D-SA-DCT is performed twice on each high activity block - once for each of the two cube partitions. The total number of transform coefficients remains 512 (as the number of pixels in the original cube) and they are concentrated in the top left corner of each transformed cube (depicted by the shadowed region in each cube in the right side of Figure 1). The remaining coefficients have a value of zero.
- A very important feature of the SA-DCT is its ability to handle non-contiguous regions. This means that it can handle non-connected regions, like in the situation shown in Figure 1 or when a block is textured.

3. TRANSFORM COEFFICIENTS CODING

The last stage of the proposed scheme is the encoding of transform coefficients. In this work we examined two alternative algorithms:

3.1. Quantization Table (QT) Algorithm

In this scheme each of the 8x8x8 DCT coefficient-cubes is first quantized using a 3D quantization table. Then, the quantized cube-coefficients are rearranged into a vector which is losslessly encoded as in JPEG.

The 3D quantization table generation, as well as the coefficients rearrangement (differently from the zigzag scan of JPEG), are based on [5] where they were introduced for the compression of image sequences (video). We were able to use the model proposed in [5] for coding HICs by adjusting the parameters of the quantization function [3].

According to the model, the *shifted hyperboloid* defined by the curved surface $(u + 1)(v + 1)(w + 1) > C$, where C is a parameter, captures the less significant coefficients, while the *shifted complement hyperboloid* defined by the curved surface $(u + 1)(v + 1)(w + 1) \leq C$ captures the most dominant coefficients, as shown in Figure 3. The following exponential function was utilized to generate the quantization table:

$$q(u, v, w) = \begin{cases} 255 \left(1 - \frac{e^{-\beta_{in} s(u, v, w)}}{e^{-\beta_{in}}} \right) + 1 & s(u, v, w) \leq C \\ 255 \left(1 - e^{-\beta_{out} s(u, v, w)} \right) & s(u, v, w) > C \end{cases}$$

where $s(u, v, w) = (u + 1)(v + 1)(w + 1)$, and β_{in} and β_{out} are parameters controlling the exponential decay in the dominant (inside) and less significant (outside) regions, respectively. C controls the border between the regions and u, v, w are in the range 0 to 7, each (for an 8x8x8 quantization table). The parameters were fitted as explained in [5] to give a good overall performance (in terms of peak-SNR (PSNR) and compression-ratio (CR)).

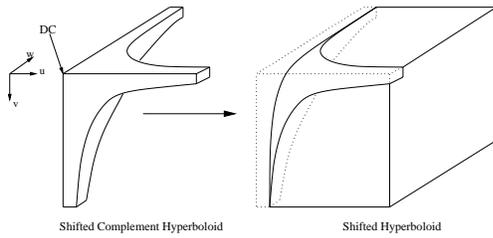


Fig. 3. The 3D-DCT dominant coefficients are concentrated along the major axes and captured by a *shifted complement hyperboloid*.

As mentioned above, the quantized coefficients are rearranged into a vector according to the scan order proposed in [5] - for the pursuant lossless coding. This ordering was chosen since it was found in our simulations to give better results than a 3D zigzag-scan or a 3D Hilbert-scan. Each vector of quantized coefficients is of length 512 (two vectors for each split active block in SA approach). The first element of the vector is the quantized DC coefficient and the remaining 511 quantized AC coefficients are located afterwards, with the dominant coefficients closer to DC, while most of the coefficients in the tail of the vector are usually zero valued. The lossless encoding scheme is based on the JPEG entropy coding scheme. The Huffman tables for the AC and DC coefficients were adjusted to the statistics of hyperspectral imagery. The complete lossless coding scheme is detailed in [3].

3.2. Quantization with a Set of Vector Quantizers (VQ)

We introduce here an alternative algorithm for encoding the 3D-DCT/3D-SA-DCT coefficients. It is based on using a set of *vector quantizers*. The main idea is to partition the vector of quantized coefficients, discussed in the previous section, into sub-vectors and code each sub-vector by a separate codebook. The size of each codebook is determined by the number of bits allocated to each sub-vector. Since the ordering process (of cube-coefficients into a vector) groups high energy coefficients at the head of the vector, both the partitioning of the vector into sub-vectors and the bit allocation are based on coefficient variances, as measured from training data. To simplify matters the full vector is partitioned into

8 sub-vectors of lengths: 4,4,8,16,32,64,128,256. The sizes of the codebooks, N_i , are then determined by the number of bits, R_i , allocated to each sub vector through $N_i = 2^{R_i}$, $i = 1, \dots, 8$. The bit allocation was determined for different rates on the basis of the sub-vector variances in the training data. The codebook design was done by the LBG algorithm.

4. SIMULATION RESULTS

The simulations were conducted on the coding scheme described in the previous sections. Four HICs were used in our study, three for training and one for testing. The HICs consist of 120 spectral bands (with spatial resolutions of 128x256 and 256x256), covering the visible and near-infrared spectral window (wavelengths from 400[nm] to 1000[nm]). Each pixel in each band has 8 bits of radiometric information. As a quality measure we used the average PSNR over all 120 spectral bands, but also examined the results of a classification test.

For comparison purposes, the Kronecker-product gain-shape VQ (KP-GS-VQ) and the hybrid DPCM/DCT/TCQ (slightly modified) coders were also implemented and served as benchmarks.

The proposed scheme was tested for five different values of the variance threshold parameter θ : 500, 700, 1000, 1500 and 2000 and, in addition, with a very high threshold value - thus applying the conventional 3D-DCT to all the blocks. Since split blocks are encoded twice, more bits were allocated to these blocks.

We applied the shape-adaptive approach to active blocks and compared its results with those obtained when no block is split. The results obtained for 3D-SA-DCT with either the VQ or QT techniques are shown in Figure 4. The compression ratios include the side-information overhead. The performance of the conventional 3D-DCT (without SA) scheme for the same encoding rates (or CRs) is given by the dotted lines - for both algorithms.

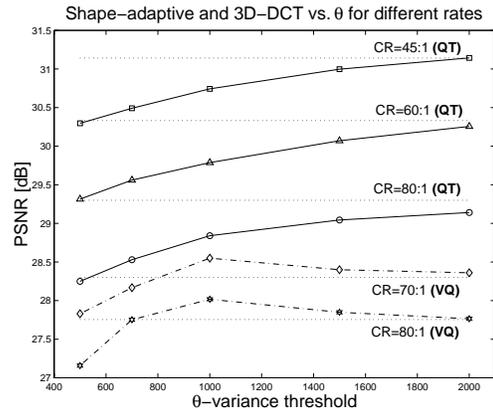


Fig. 4. Performance of the 3D-SA-DCT approach with QT/VQ quantization algorithms.

As we hoped, when the VQ scheme follows the 3D-SA-DCT of high activity blocks, it performs better than when it follows a conventional 3D-DCT, with a suitable choice of the threshold value. In Figure 4 this value is $\theta_{opt} = 1000$. For this threshold value, 14% of blocks are split and transformed by 3D-SA-DCT. The reason for the relatively small advantage is the high overhead which is characteristic to this approach, stemming from the need to code two sets of transform coefficients and the bitmap. While the

encoding of two lower energy regions of a split block is typically more efficient than using the conventional 3D-DCT for coding a high activity block, the bitmap side-information which consists of 32 bits per each split block (on average) almost cancels the advantage achieved by the splitting.

For the QT coding technique the above advantage disappears and at present there is no advantage in using the shape-adaptive approach with QT. Nevertheless, there is, in principle, a case for studying further this approach, because of the evidence to its superiority over direct 3D-DCT encoding of high activity blocks. In particular, improvements in the splitting criterion and reduction of the side-information overhead of the segmentation bitmap are needed.

A performance comparison of the two benchmark coders with the 3D-DCT, followed by VQ or QT, is presented in Figure 5.

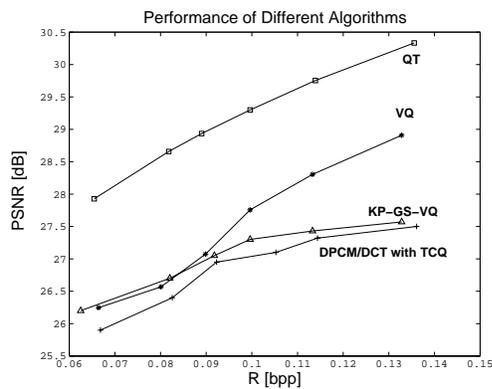


Fig. 5. Comparative performance summary without SA coding.

The figure shows the average PSNR obtained for each algorithm versus the average bit-rate in bits per pixel. The rate axis corresponds to compression ratios (CR) between 60:1 (0.13[bpp]) and 130:1 (0.06[bpp]), with 0.1[bpp] corresponding to CR = 80.

It is seen from the figure that both the VQ and the QT techniques achieve better average PSNR, for the same bit-rate, than KP-GS-VQ and DPCM/DCT with TCQ algorithms, in almost all the range of bit-rate values examined. It should be noted that only in the QT scheme lossless compression was applied, since it is a built-in feature of the QT approach. But even the VQ method, which does not apply lossless coding (only little gain was achieved by applying it), achieves better results than the benchmarks.

It is interesting to note that the application of a standard H.261 video coder to the set of hyperspectral images, resulted, at best, in a lower PSNR (by 0.5 to 1.5 dB) than the 3D-DCT with the VQ scheme. This, in a way, justifies the development of a specific coder for this type of images.

Table 1 summarizes the classification error rates of a simple 3-class Maximum Likelihood (ML) classifier [6], implemented as a subsequent image analysis operation and applied to both the original and the coded images (see [3] for details). The error rate of the classifier applied to the original test HIC is 10.3%. It can be observed that the SA scheme achieved lower error rates than other algorithms and even lower than the classification error obtained for the original image. This result gives additional motivation for utilizing and further studying the shape-adaptive approach.

R [bpp]	[1]	[2]	3D-DCT		3D-SA-DCT	
	scheme	scheme	QT	VQ	QT	VQ
0.13	12.6%	11.5%	12.1%	10.5%	11.1%	10.3%
0.1	13.2%	14%	14.2%	11.1%	9.1%	9.5%
0.08	15.7%	14.9%	16%	15%	10.2%	10.6%

Table 1. Classification error summary for different coders.

5. CONCLUSION

The attractive features of the shape-adaptive DCT and the relevance of its three-dimensional extension to efficient encoding of high activity blocks in hyperspectral images encouraged us to apply this technique. Due to its ability to code arbitrary shaped regions, the 3D-SA-DCT enables the encoder to overcome the inefficient coding by conventional 3D-DCT of high activity blocks - due to both edges and coarse texture. To code the transform coefficients (resulting from either a 3D-DCT or a 3D-SA-DCT), two techniques - QT and VQ, along with a special coefficient reordering, were adopted and found to give better coding performance (in PSNR) than two benchmark algorithms examined.

Because of the added side-information, of the segmentation bitmap needed by the shape-adaptive (SA) approach, it failed to improve the results obtained by the conventional 3D-DCT when both are followed by the QT scheme. Yet, in a classification task the 3D-SA-DCT provided better results than the conventional 3D-DCT when both were followed by either the QT or VQ schemes. Thus, there is room for further study of the 3D-SA-DCT approach for hyperspectral image coding, in particular with respect to the splitting criterion and efficient coding of the segmentation bitmap.

6. REFERENCES

- [1] G. P. Abousleman, M. W. Marcellin and B. R. Hunt, "Compression of hyperspectral imagery using 3-d DCT and hybrid DPCM/DCT", IEEE Trans. Geosci. Remote Sensing, Vol. 33, pp. 26-34, Jan. 1995.
- [2] G. R. Canta and G. Poggi, "Kronecker-product gain-shape vector quantization for multispectral and hyperspectral image coding", IEEE Trans. on Image Processing, Vol. 7, No. 5, pp. 668-678, May 1998.
- [3] D. Markman, "Hyperspectral Image Coding Using 3D Transforms", M.Sc. Research Thesis, Technion - Israel Institute of Technology, Haifa, Israel, July. 2000, http://www-sipl.technion.ac.il/Dima_thesis.ps
- [4] T. Sikora and B. Makai, "Shape-Adaptive DCT for generic coding of video", IEEE Trans. Circuits Syst. Video Technol., Vol. 5, No. 1, pp. 59-62, Feb. 1995.
- [5] M. C. Lee, R. K. W. Chan and D. A. Adjero, "Quantization of 3D-DCT Coefficients and Scan Order for Video Compression", Journal of Visual Communication and Image Representation, Vol. 8, No. 4, pp. 405-422, 1997.
- [6] D. Landgrebe, "Some Fundamentals and Methods for Hyperspectral Image Data Analysis", SPIE Photonics West, San Jose CA, Jan. 1999.