

JPEG2000 Coding Strategies for Hyperspectral Data

Justin T. Rucker, James E. Fowler, and Nicholas H. Younan
Department of Electrical and Computer Engineering
GeoResources Institute, Mississippi State ERC
Mississippi State University, Mississippi State, MS 39762

Abstract—In using JPEG2000 for the coding of multiple-component, or multiband, images such as hyperspectral imagery, one must consider spectral decorrelation and rate allocation between image components, issues that concern the design of the JPEG2000 encoder and are, consequently, outside the scope of the JPEG2000 standard. Spectral decorrelation via a wavelet transform, as well as three alternative strategies for extending to multiple components the optimal codeblock-bitstream-truncation process widely used for spatial rate allocation in JPEG2000 coding of single-component imagery, are considered. Results indicate that the strategy of simultaneously truncating all codeblock bitstreams from all codeblocks from all image components coupled with wavelet-based spectral decorrelation significantly outperforms the other techniques considered in terms of not only rate-distortion performance but also accuracy of unsupervised classification.

I. INTRODUCTION

Uncompressed hyperspectral images are very large, with a single image potentially occupying hundreds of megabytes. Compression is thus necessary to facilitate both the storage and the transmission of hyperspectral images. JPEG2000 [1–3] is an embedded, wavelet-based coder that has been increasingly considered for the coding of hyperspectral imagery as well as other types of volumetric data, such as medical imagery. JPEG2000 is attractive because of its proven state-of-the-art performance for the compression of grayscale and color photographic imagery. However, its performance for hyperspectral compression can vary greatly depending on how the JPEG2000 encoder handles multiple-component images, i.e., images with multiple spectral bands.

In effect, the JPEG2000 standard specifies the syntax and semantics of the compressed bitstream and, consequently, the operation of the decoder. The exact architecture of the encoder, on the other hand, is left largely to the designer of the compression system. For example, in [4], Varma and Bell explore tradeoffs for several parameters to a JPEG2000 encoder, such as the color space, quantization stepsize, and the number of transform levels. However, the coding of multiple-component imagery is not considered.

In deploying JPEG2000 on multiple-component images, such as hyperspectral imagery, there are two primary issues that must be considered in the implementation of the JPEG2000 encoder: 1) spectral decorrelation and 2) rate allocation between image components. The first issue arises due to the fact that there tends to exist significant correlation between consecutive bands in a hyperspectral image. In this paper, we consider spectral decorrelation, via a wavelet transform and find significant performance improvement results from its use.

The second encoder-design issue—rate allocation between image components—arises from the fact that, essentially, JPEG2000 is a 2D compression algorithm. Consequently, given a specific target bitrate of R bits per pixel per band (bpppb), the JPEG2000 encoder must determine how to allocate this total rate appropriately between bands. It is usually the case that certain bands have significantly higher energy than other bands and thus will weigh more heavily in distortion measures than the other, weaker-energy bands. Consequently, it is likely that the JPEG2000 encoder will need to allocate proportionally greater rate to the higher-energy bands in order to maximize distortion performance for a given total rate R . In this paper, we explore several rate-allocation strategies and find significant performance difference between them.

In the following sections, we first briefly overview JPEG2000 compression for single-component imagery. We then discuss the application of a wavelet transform for spectral decorrelation within a JPEG2000 encoder and describe three strategies for rate allocation between multiple image components. Subsequently, we evaluate all the considered techniques in terms of both rate and distortion performance, as well as in terms of accuracy of unsupervised classification.

II. JPEG2000 FOR SINGLE-COMPONENT IMAGES

To code a single-component image, a JPEG2000 encoder first performs a 2D wavelet transform on the image and then partitions each transform subband into small, 2D rectangular blocks called codeblocks, which are typically of size 32×32 or 64×64 pixels. Subsequently, the JPEG2000 encoder independently generates an embedded bitstream for each codeblock. To assemble the individual codeblock bitstreams into a single, final bitstream, each codeblock bitstream is truncated in some fashion, and the truncated bitstreams are concatenated together to form the final bitstream. The method for codeblock-bitstream truncation is an implementation issue concerning only the encoder as codeblock-bitstream lengths are conveyed to the decoder as header information. Consequently, this truncation process is not covered by the JPEG2000 standard.

It is highly likely that, for codeblocks residing in a single image component, any given JPEG2000 encoder will perform a Lagrangian rate-distortion optimal truncation as described as part of Taubman's EBCOT algorithm [3, 5]. This optimal truncation technique, post-compression rate-distortion (PCRD) optimization, is a primary factor in the excellent rate-distortion performance of the EBCOT algorithm. PCRD optimization is performed simultaneously across all of the codeblocks from

the image, producing an optimal truncation point for each codeblock. The truncated codeblocks are then concatenated together to form a single bitstream. The PCRD optimization, in effect, distributes the total rate for the image spatially across the codeblocks in a rate-distortion-optimal fashion such that codeblocks with higher energy, which tend to more heavily influence the distortion measure, tend to receive greater rate.

III. SPECTRAL DECORRELATION FOR MULTIPLE-COMPONENT IMAGES

The JPEG2000 standard allows for up to 16,385 image components to be included in a single bitstream; however, the standard does not specify how these image components should be encoded for best performance. Whereas Part I of the JPEG2000 standard [1] permits spectral decorrelation only in the case of three-band images (i.e., red-green-blue), Annexes I and N of Part II of the standard [2] make provisions for arbitrary spectral decorrelation, including wavelet transforms.

By applying a 1D wavelet transform spectrally, and then subsequently employing a 2D wavelet transform spatially within each component, we effectively implement a particular 3D wavelet-packet transform which has been used extensively for the coding of 3D volumes (e.g., [6, 7]). We note that many JPEG2000 implementations are not yet fully compliant with Part II of the standard. In this case, we can “simulate” the spectral decorrelation permitted under Part II by employing a 1D wavelet transform spectrally on each pixel in the scene before the image cube is sent to the Part-I-compliant JPEG2000 encoder. Such an external spectral transform as been used previously [7, 8] to implement a “2D spatial + 1D spectral” wavelet-packet transform with Part-I-compliant coders.

IV. RATE-ALLOCATION STRATEGIES ACROSS MULTIPLE IMAGE COMPONENTS

The PCRD optimization procedure of EBCOT [3, 5] produces a rate-distortion-optimal bitstream for a single-component image by optimally truncating the independent codeblock bitstreams from the component. However, there are several ways that this single-component truncation procedure can be extended to the multiple-component case, and the resulting multiple-component truncation procedure, in effect, dictates how the total rate available for coding the hyperspectral image is allocated between the individual spectral bands.

That is, for a multiple-component image, a JPEG2000 encoder will partition each component, or spectral band, into 2D codeblocks which are coded into independent bitstreams as described above in Sec. II for single-component imagery. To assemble a final bitstream, these individual codeblock bitstreams are truncated and concatenated together. Although the method for codeblock-bitstream truncation is an implementation issue concerning only the encoder and is thus not covered by the JPEG2000 standard, it is highly likely that, any given multiple-component JPEG2000 encoder with perform PCRD optimization for *at least* the codeblocks originating from a single image component. How this truncation process is extended across the multiple components may vary with encoder implementation.

Below, we describe three possible multiple-component rate-allocation strategies and evaluate each for the compression of hyperspectral data. In the following, let a hyperspectral image volume X be composed of N bands X_i , i.e., $X = \{X_1, X_2, \dots, X_N\}$. We code X with a total rate of R bpppb. Assume that $B_i = \text{JPEG2000_Encode}(R_i, X_i)$ is a single-component JPEG2000 encoder that encodes component X_i with rate R_i using PCRD optimization, producing a bitstream B_i .

A. JPEG2000-BIFR

The most straightforward method of allocating rate between multiple image components is to simply code each component independently and assign to each an identical rate. This JPEG2000 band-independent fixed-rate (JPEG2000-BIFR), strategy operates as follows:

```
JPEG2000_BIFR( $R, \{X_1, \dots, X_N\}$ )
 $B = \emptyset$ 
for  $i = 1, 2, \dots, N$ 
     $B_i = \text{JPEG2000\_Encode}(R, X_i)$ 
 $B = B \circ B_i$ 
return  $B$ 
```

where the “ \circ ” operator denotes bitstream concatenation.

B. JPEG2000-BIRA

The next method, JPEG2000 band-independent rate allocation (JPEG2000-BIRA), also codes each band independently; however, rates are allocated explicitly so that more important bands are coded with higher bitrate, and less important bands are coded at a lower bitrate.

```
JPEG2000_BIRA( $R, \{X_1, \dots, X_N\}$ )
 $B = \emptyset$ 
for  $i = 1, 2, \dots, N$ 
     $\sigma_i^2 = \text{variance}[X_i]$ 
for  $i = 1, 2, \dots, N$ 
     $R_i = \frac{\log_2 \sigma_i}{\sum_{j=1}^N \log_2 \sigma_j} \cdot RN$ 
     $B_i = \text{JPEG2000\_Encode}(R_i, X_i)$ 
 $B = B \circ B_i$ 
return  $B$ 
```

The rates, R_i , are determined so that bands with larger variances (i.e., higher energy) are coded at a higher bitrate than those with lower variances, while the total rate for the entire volume is R . This approach is, in essence, an *ad-hoc* variant of classical optimal rate allocation for a set of quantizers based on log variances (chap. 8 of [9], [10]).

C. JPEG2000-MC

The final approach, what we will call JPEG2000 multi-component (JPEG2000-MC), can be employed when the JPEG2000 encoder is capable of performing PCRD optimization across multiple bands. That is, all of the spectral bands are input to the encoder which produces codeblock bitstreams for every codeblock in every subband of every image component. Then, PCRD optimal truncation is applied to all codeblock

bitstreams from all bands simultaneously, rather than simply the codeblock bitstreams for a single band as in Sec. II. In this way, the PCRD optimization performs to the maximum of its potential, implicitly allocating rate in a rate-distortion fashion, not only spatially within each image component, but also spectrally across the multiple bands.

V. EXPERIMENTAL RESULTS

For the results of this section, all JPEG2000 coding was done with Kakadu¹ Version 4.3, with 5 levels of wavelet decomposition both spatially and spectrally and a quantization step size of 0.0000001. The popular 9-7 biorthogonal filter (included in Part I of the standard) was used for both spatial and spectral transforms. Since Kakadu is not yet fully compliant with Part II of the JPEG2000 standard, the spectral transform was applied externally as described in Sec. III and [7, 8].

All the datasets used in the experiments were collected by the Airborne Visible Infrared Imaging Spectrometer (AVIRIS), an airborne, hyperspectral sensor that collects data in 224 contiguous bands from 400 nm to 2500 nm. For the results here, we crop the first scene in each dataset to produce image cubes with the dimensions $512 \times 512 \times 224$. In all cases, the unprocessed radiance data was used.

We first examine rate-distortion performance of JPEG2000 encoding. In Fig. 1, we plot rate-distortion performance for a range of rates, while in Table I, distortion performance at a single rate is tabulated. In these results, techniques labeled as “2D” do not use any spectral transform (i.e., only 2D wavelet transforms are applied spatially), while the other techniques use the 3D wavelet-packet transform which includes a spectral transform. For each dataset, we present performance for the three rate-allocation techniques described in Sec. IV, both with and without the spectral decorrelation transform. With the exception of JPEG2000-BIFR, all the rate-allocation techniques perform significantly better when a spectral transform is performed. We see that JPEG2000-MC substantially outperforms the other techniques by at least 5–10 dB.

We next turn our attention to classification performance. All of the classification results were calculated using the ISODATA and k -means unsupervised classification techniques as implemented in ENVI Version 4.0. A maximum of ten classes were used, and accuracy of the classification was determined by applying the classification to the original dataset as well as to the reconstructed volume and comparing the classification map produced for reconstructed volume to that of the original dataset, using the classification map of the original dataset effectively as “ground truth.” Unsupervised classification results are presented for ISODATA and k -means in Table II. We see that the classification performances correlate well with SNR figures from Table I—if one technique outperforms another in the rate-distortion realm, then it will mostly likely have higher classification accuracy as well. As expected, JPEG2000-MC performs substantially better than the other techniques in terms of classification accuracy.

¹<http://www.kakadusoftware.com>

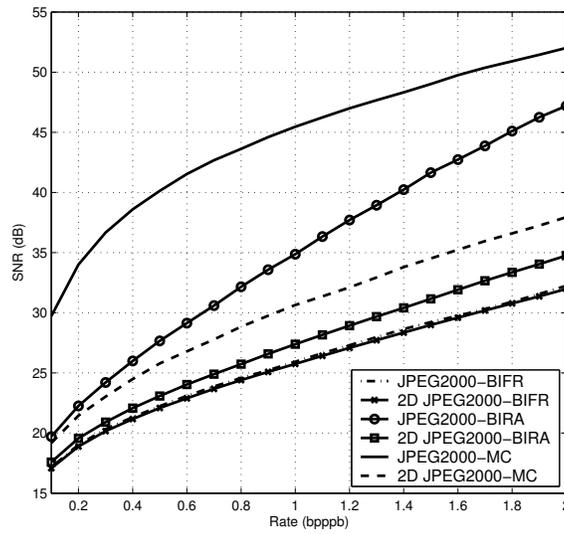
VI. CONCLUSIONS

The JPEG2000 standard covers only the decoder, so how encoders handle multiple spectral bands is left to the design of the encoder. The results in this paper demonstrate that different encoder designs may substantially affect performance. In this paper, we considered three different rate-allocation strategies for JPEG2000 and evaluated the performance of each, both with and without a spectral transform. We find that the technique that performs optimal rate-distortion truncation of bitstreams from all codeblocks from all image components—JPEG2000 multiple-component (JPEG2000-MC)—substantially outperforms the other techniques. Additionally, performance almost always benefits greatly from the application of a 1D spectral wavelet transform to remove correlation in the spectral direction. We find also that the accuracy of classification performed on the compressed image correlates well with the rate-distortion performance of the compression.

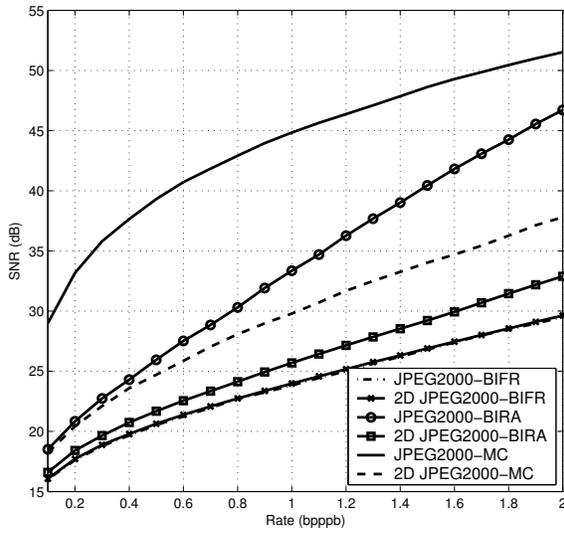
We note that both Kakadu Version 4.3 and the JPEG2000 encoder in ENVI Version 4.1 (which uses the Kakadu coder) implement JPEG2000-MC rate allocation, and neither support the use of a spectral transform since they are not fully compliant with Part II of the JPEG2000 standard. Thus, the performance of these coders is equivalent to that of 2D JPEG2000-MC approach considered here. As our results indicate, adding a spectral transform would significantly enhance the performance of these coders.

REFERENCES

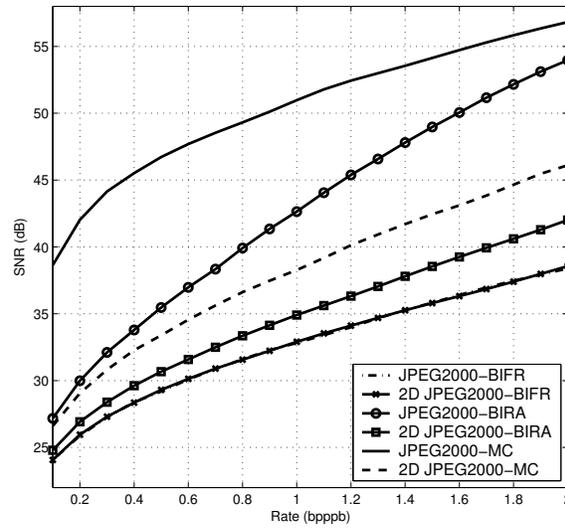
- [1] *Information Technology—JPEG 2000 Image Coding System—Part 1: Core Coding System*, ISO/IEC 15444-1, 2000.
- [2] *Information Technology—JPEG 2000 Image Coding System—Part 2: Extensions*, ISO/IEC 15444-2, 2004.
- [3] D. S. Taubman and M. W. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards and Practice*. Boston, MA: Kluwer Academic Publishers, 2002.
- [4] K. Varma and A. Bell, “JPEG2000 choices and tradeoffs for encoders,” *IEEE Signal Processing Magazine*, vol. 21, no. 6, pp. 70–75, November 2004.
- [5] D. Taubman, “High performance scalable image compression with EBCOT,” *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1158–1170, July 2000.
- [6] B.-J. Kim, Z. Xiong, and W. A. Pearlman, “Low bit-rate scalable video coding with 3-D set partitioning in hierarchical trees (3-D SPIHT),” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 8, pp. 1374–1387, December 2000.
- [7] Y. Wang, J. T. Rucker, and J. E. Fowler, “3D tarp coding for the compression of hyperspectral images,” *IEEE Geoscience and Remote Sensing Letters*, vol. 1, no. 2, pp. 136–140, April 2004.
- [8] H. S. Lee, N. H. Younan, and R. L. King, “Hyperspectral image cube compression combining JPEG-2000 and spectral decorrelation,” in *Proceedings of the International Geoscience and Remote Sensing Symposium*, vol. 6, Toronto, Canada, June 2002, pp. 3317–3319.
- [9] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Norwell, MA: Kluwer Academic Publishers, 1992.
- [10] J. J. Y. Huang and P. M. Schultheiss, “Block quantization of correlated Gaussian random vectors,” *IEEE Transactions on Communications*, vol. 11, no. 3, pp. 289–296, September 1963.



(a)



(b)



(c)

Fig. 1. Rate-distortion performance; rate is in bits per pixel per band (bpppb) and distortion is measured as signal-to-noise ratio (SNR) in dB. (a) moffett, (b) jasper ridge, (c) cuprite.

TABLE I
SNR PERFORMANCES AT 1.0 BPPPB.

| Dataset | SNR (dB) | | | | | |
|--------------|----------|------|---------|------|-------|-------------|
| | 2D BIFR | BIFR | 2D BIRA | BIRA | 2D MC | MC |
| moffett | 25.8 | 25.9 | 27.4 | 34.9 | 30.6 | 45.5 |
| jasper ridge | 24.0 | 23.8 | 25.7 | 33.4 | 29.8 | 44.8 |
| cuprite | 32.9 | 32.8 | 34.9 | 42.6 | 38.3 | 51.0 |

TABLE II
ISODATA AND *k*-MEANS CLASSIFICATION PERFORMANCES AT 1.0 BPPPB.

| Dataset | ISODATA (% accuracy) | | | | | |
|--------------|----------------------|------|---------|------|-------|-------------|
| | 2D BIFR | BIFR | 2D BIRA | BIRA | 2D MC | MC |
| moffett | 83.4 | 94.5 | 86.6 | 94.5 | 93.2 | 99.7 |
| jasper ridge | 77.3 | 75.5 | 82.2 | 93.7 | 93.9 | 99.7 |
| cuprite | 80.3 | 78.1 | 85.1 | 94.7 | 94.7 | 99.8 |

| Dataset | <i>k</i> -means (% accuracy) | | | | | |
|--------------|------------------------------|------|---------|------|-------|-------------|
| | 2D BIFR | BIFR | 2D BIRA | BIRA | 2D MC | MC |
| moffett | 75.4 | 73.2 | 79.9 | 91.7 | 89.8 | 99.6 |
| jasper ridge | 67.2 | 64.7 | 73.9 | 90.4 | 91.0 | 99.5 |
| cuprite | 71.3 | 68.3 | 77.6 | 92.2 | 92.1 | 99.6 |