

DISCRETE WAVELET TRANSFORM EFFECT ON HYPERSPECTRAL IMAGE CLASSIFICATION PERFORMANCE USING LOSSLESS COMPRESSION ALGORITHMS WITH COLOR & QUALITY SCALABILITY UNDER JPEG2000

David Negrete, Vikram Jayaram, Bryan Usevitch
Signal Processing & Communication Research Group
The University of Texas at El Paso
500 W. University Ave., El Paso, TX 79968-0523
[\[Negrete, Jayaram, Usevitch\]@ece.utep.edu](mailto:{Negrete, Jayaram, Usevitch}@ece.utep.edu)

ABSTRACT

In this paper classification performance of Hyperspectral images using two different data decorrelating transforms is investigated. The tradeoff between the discrete wavelet transform (DWT) and the principal component analysis (PCA) is measured using different compression algorithms found in the JPEG2000 standard. These cross-component transforms can be evaluated in order to obtain a comparison of how well each transform can decorrelate data, reorder data optimally, speed of implementation, and effect on mineral classification performance. In addition, two scalability options defined by JPEG2000 were used: color and quality. Compression algorithms were carried out for the DWT and the PCA using the same set of parameters; the same Hyperspectral image, the same classification algorithm—spectral angle mapper (SAM), and the same type of compression: lossless compression. The purpose of this comparison is to be able to determine how much partial decompression is needed in order to achieve a 90% hit rate when classifying several minerals. This evaluation is focused primarily on two types of Mica called Aluminum-rich and Aluminum-poor Mica. Our past research using PCA has shown extremely good overall classification performance. However, this transform has the disadvantage of higher computational complexity. This prompted us to try out an alternate lower complexity transform such as the DWT. The classification results were found poor as compared to PCA. The complex nature of our classification does not allow us come to a complete judgment of this finding. Thus, other data sets involving low complex nature of classification shall be dealt in our future work. The simulation results conclude that the DWT is seven times faster than PCA and allows parallel processing huge data sets.

INTRODUCTION

Hyperspectral images are large data sets which are used by geologists to detect and classify minerals. These images are obtained by airborne or satellite-borne satellites by a method called remote sensing. Every mineral or plant reflects and/or emits a unique amount of electromagnetic energy distributed among many frequency bands. Remote sensing works by measuring this energy and then using special software to decode this “spectral signature” and assign it to its proper source, be it a plant or mineral. The Hyperion (Hyperspectral Imager) provides a high resolution Hyperspectral image capable of resolving 220 spectral bands (0.4 - 2.5 μm) with a 30 meters pixel resolution. The data used in these experiments is the Mt. Fitton scene in southern Australia. The layers or bands of this Hyperion data are highly correlated to each other in the spectral direction. Decorrelation of these bands allows a more compact representation of an image by packing the energy into fewer numbers of bands. Part 2 of the JPEG2000 standard recommends the use of either the PCA or the DWT for decorrelation. Although PCA is optimal transform when it comes to efficiently packing band energies in fewer dimensions nevertheless its computationally inefficient nature leads us to employ DWT for decorrelating the bands. The DWT can be implemented by either a convolution or a lifting scheme. We have used the lifting scheme implementation of DWT due to its lesser computations and memory requirements. In our experiments the 5/3 Daubechies filter is used as we are performing lossless compression.

The Hyperion instrument can image a 7.5 km by 100 km land area per image and provide detailed spectral mapping across all 220 channels with high radiometric accuracy. The Hyperion instrument, in conjunction with advanced land imager (ALI) instrument which provides accurate radiometric calibration on orbit using precisely controlled amount of solar irradiance, forms the standalone camera system on the Earth Observatory-1 (EO-1) satellite [7]. The analysis of Hyperspectral Imagery by means of such spectrometers (Hyperion) exploits the fact that each material radiates different amount of electromagnetic energy throughout all the spectra. This unique

characteristic of the material is commonly known as the spectral signature which can be read using such airborne or space borne-based detectors. The data used for experiments was collected by the Hyperion over the semiarid Mount Fitton area which is located in the Northern Flinders Ranges of South Australia. Centered at $-29^{\circ} 55' S$ and $139^{\circ} 25' E$, about 700 kilo-meters NNW of Adelaide. This Hyperion image consists of 194 atmospherically corrected contiguous spectral bands. Each band has dimension of $6702 \times 256 \times 2$ bytes, which is equal to 634.857 MB in memory utilized by the whole data cube. The most obvious problem with using data cubes this big is that storage becomes a problem when managing multiple files. In addition, moving these files is not easy due to their size. Therefore, an appropriate algorithm is needed to compress files that are larger than usual. On the other hand, the compression algorithm should minimize the difference, or error, between original and decompressed images when using lossy algorithms.

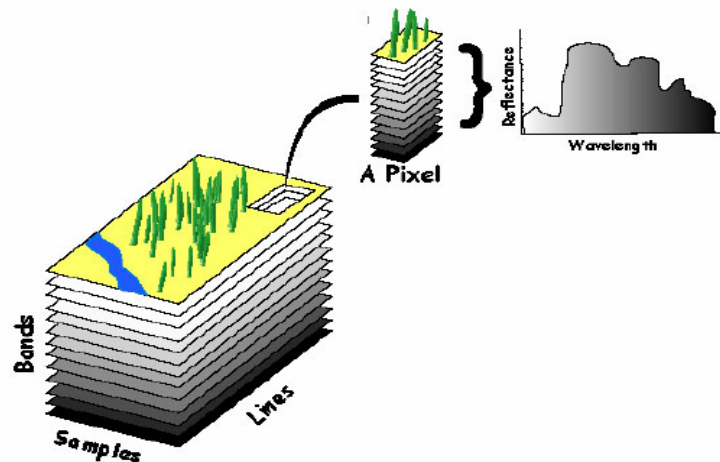


Figure 1. Hyperspectral remote sensing.

JPEG2000 FRAMEWORK

The recent JPEG2000 standard has many advantages compared to earlier compression standards. Such advantages include superior low bit rate performance, bit rate scalability and progressive transmission by quality or resolution. The coder and decoder for JPEG2000 use a wavelet based compression method which implements the embedded block coding with optimized truncation (EBCOT) [5] scheme. The use of wavelets to decompose a signal, unlike the old standard JPEG that relied on the discrete cosine transform (DCT) to approximate a signal, has proved to have numerous advantages. For example, wavelets provide localized support which is more useful to approximate high frequency components in an image. In addition, with appropriate coding algorithms, such as embedded zero-tree wavelet (EZW) by Shapiro [6] or Set Partitioning in Hierarchical Trees (SPIHT) by Said and Pearlman [9], it has been shown that the wavelet transform yields better compression than the cosine transform [9]. These algorithms are useful since they provide an ordering of data according to its importance. Specific details about the DWT and the ordering of data will be explained later.

Quality scalability is achieved by dividing the wavelet transformed image into code-blocks. After each code-block is encoded, a post-processing operation determines where each code-block's embedded stream should be truncated in order to achieve a predefined bit-rate or distortion bound for the entire data. This bit-stream rescheduling module is referred to as Tier 2. It establishes a multi-layered representation of the final bit-stream, guaranteeing an optimal performance at several bit-rates or resolutions. Figure 2 illustrates the general flowchart of the JPEG2000 architecture. The Tier 2 component optimizes the truncation process, and tries to reach the desired bit-rate while minimizing the introduced distortion or mean squared error (MSE), utilizing Lagrangian rate allocation principles. This following procedure is known as post compression rate-distortion (PCRD) optimization [4] and the basic principles behind it are extensively discussed in [5].

The new image coding standard JPEG2000 does not only offer better compression performance than the previous one, but it also offers more options when decompressing an image. The four types of scalability presented are: spatial, resolution, color, and quality scalability. The first two take advantage of the two-dimensional wavelet transform to decompress a portion of the image or reduce its dimensions respectively. Specific details about these types of scalability can be found in [5]. However, the performance comparison between the PCA and the one-

dimensional DWT used to decorrelate the image components do not take advantage of these two scalability options. Therefore, is not pursued in the scope of this paper.

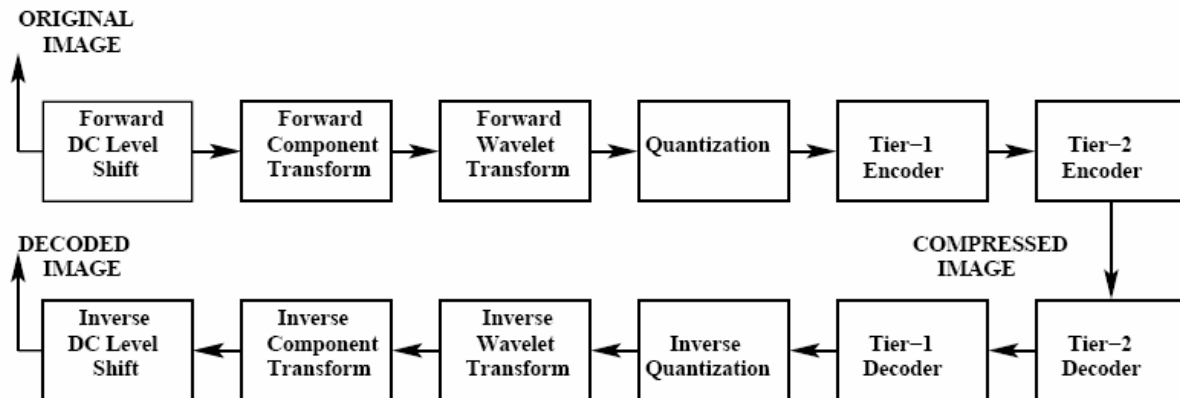


Figure 2. JPEG2000 architectural flowchart.

The other key advantage of the JPEG2000 standard in HSI compression is that due to its ordering of data, smaller files can be obtained from a compressed file by truncating the bit stream where needed. In other words, a compressed file using 4 bits per pixel of an image could be truncated in half to obtain a file using 2 bits per pixel. Due to these advantages and the availability of color and quality scaling, which will be explained later, the JPEG2000 standard was chosen as the method to compress the Hyperspectral images. The JPEG2000 standard has been established and its software implementations have started to become available. However, a three-dimensional transform has not been determined nor specified by the standard for multi-dimensional electro-optical images, which are those that have more than three color components. Therefore, it is now required to use a transform that will remove redundancy in the spectral direction. HSI compression has relied mainly on the use of the PCA transform to transform images in the spectral direction [1]. Its main advantage is that it not only allows the most decorrelation possible among spectral bands, but also the best energy compaction, which in turn results in better mineral classification. On the other hand, a main disadvantage of the PCA is that the covariance matrix, which is used to decorrelate among frequency bands, has to be calculated, and therefore is data-dependent. In addition, its implementation is computationally expensive and does not help to decrease the number of time and computations needed to classify materials. On the contrary, decompression times can be increased by several times.

USE OF DECORRELATING TRANSFORMS

Color and quality scalability are options that can be applied to HSI or its 3-D transformed version. Before explaining color and quality scalability, an important property of the PCA and DWT needs to be considered. After transformation of the data cube in the spectral direction, an interesting result is obtained. Figure 3 shows a hypothetical distribution of energy after the ten components of an image cube undergoes transformation.

It can be seen that most of the energy in the data cube is compacted into the first few bands. In addition, each band has less energy than the ones before. Therefore, data is compacted and ordered by the transform. Color scalability takes advantage of this order by decreasing the number of color component bands that are used to decompress an image. In other words, the image decoder only decompresses a predefined fewer number of bands, assuming the rest to be zero, and performs the inverse transform to approximate the original image.

One of the advantages of color scalability is that it can reduce the amount of time spent on decompressing an image since only a few components are used. Quality scalability relies on the energy compacting property of the transform. Instead of using the first bands of a transformed data cube like color scalability, quality scalability assigns larger amounts of the bit budget to those bands that are believed to be more important. This means that by distributing the bit budget accordingly, a measure of error can be optimally decreased.

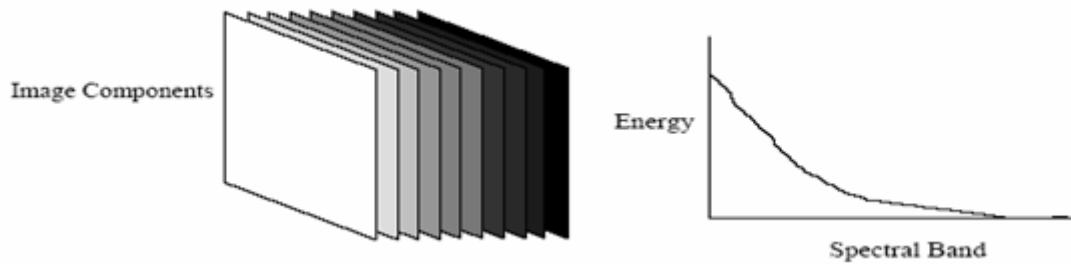


Figure 3. Energy plot of an image with ten components after transformation.

COMPRESSION PERFORMANCE

Since JPEG2000 supports compression of images with more than three color components, but does not define the transform in the color component direction, the most important aspect that has to be addressed is that of decorrelation among the bands, which will result in higher compression ratios. The cross-component transform had to be written and implemented separately from current JPEG2000 compression software in order to use off the shelf software. In addition, once a Hyperspectral image has been transformed, it is divided into its color components. Each file contains a single color component band, which results in a total of 194 files that can then be compressed independently. Lossless compression algorithms were used for all experiments. The method is carried backwards when decompressing—component bands are first merged and then inverse transformed to obtain the original image.

The PCA is defined as the transform that achieves the most decorrelation possible among bands. Since compression is directly related to decorrelation, the PCA transform is also the one that achieves most compression. Therefore, the first item to be compared is compression performance. Previous results show that when using the PCA transform and lossless compression, the resulting 194 files had a combined size of 271,022,749 bytes, or 258.467 MB. The compressed file is only 40.71% of the original file. However, since the decoder requires the transformation matrix used for the PCA, two more files have to be included. The first file contains the mean vector and its size is 776 bytes. In addition, the second file of size 304,842 bytes contains the transformation matrix. This results in a total file of size 271,328,367 bytes, equivalent to reducing the original file to 40.76% of its size. The next step was to repeat the procedure, with the only change being the use of the DWT to decorrelate component bands. The compression algorithm was followed exactly. Now, the 194 files had a combined size of 298,548,385 bytes, or 284.718 MB. This yields a compression percentage to about 44.84% of the original size. This means that the DWT—a transform that is significantly faster, is not data-dependent and does not require additional files for decompression—has a compression performance of about 4% less than that of the PCA.

COLOR SCALABILITY USING JPEG2000

Cross-component transforms, such as the DWT and PCA, have the objective of compacting energy into as few component bands as possible. By using only the most important bands of an image, some compression can be achieved. This is what is referred to as color scalability. Using a cross-component transform makes it possible to distinguish which bands are more important in an image.

Just like the PCA, the DWT tries to compact as much energy into the first bands as possible. However, the DWT does not achieve as good data ordering as the PCA. In other words, after using the DWT, the component bands do not necessarily decrease in energy. Using the first component bands to decompress an image might result in a reduction of mineral classification performance due to the fact that these bands do not contain the most information. Therefore, two methods were followed to decompress an image using color scalability. The first involves using the order of the component bands after implementing the Discrete Wavelet Transform. This means, that the first bands are always used. The second method requires using another parameter to measure the importance of the component bands. This results in having the bands reordered. Using the order of component bands after applying the DWT is required for this method. The algorithm followed is shown in Figure 5.

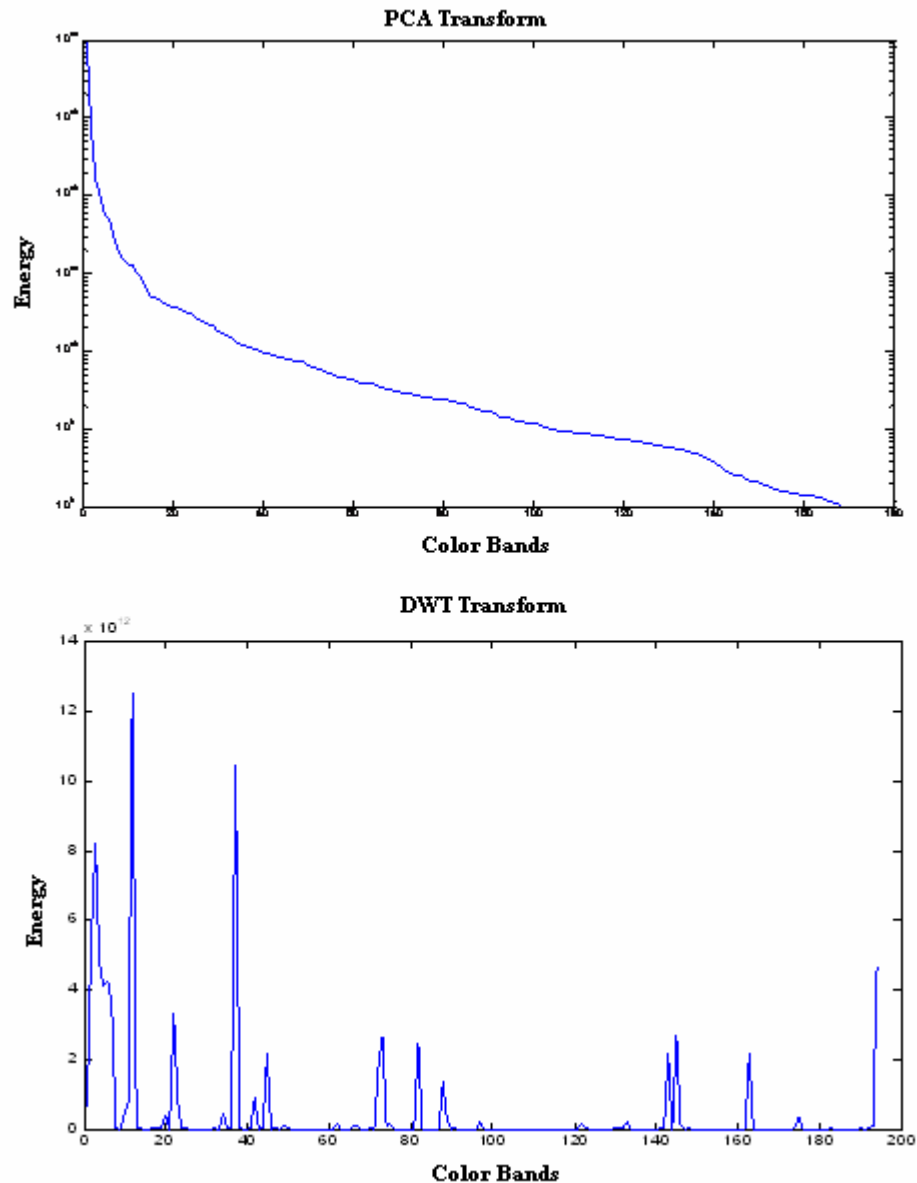


Figure 4. Comparison of energy distributions among bands after PCA and DWT transforms.

Applying the cross-component transform is the first step required. As described previously, this transform is implemented separately of current JPEG2000 implementations. Once this is done, each component band is saved in a new file, therefore creating 194 files, which are compressed independently. This is where color scalability takes place: only a few bands are decompressed. Let the number of component bands used be N . Obviously, N can only be an integer less than or equal to 194. These bands are then merged, and the rest of the component bands are assumed to be zero. The inverse transform is performed, which results in an approximation of the original HSI. Finally, mineral classification is performed and results are taken. Several compression rates were obtained by using a set of few numbers of component bands. This means that this method was followed when using 7, 13, 25, 49, and 97 component bands to approximate the original image. One of the advantages of color scalability is the considerable reduction of time required to decompress an image, since only a few bands are decompressed. In addition, this method is also easily implemented since additional files are not required. For instance, a transformation matrix, such as the one used for the PCA, is not needed. However, it is expected that this method will result in slightly worse classification performance because of the energy compaction properties of the DWT.

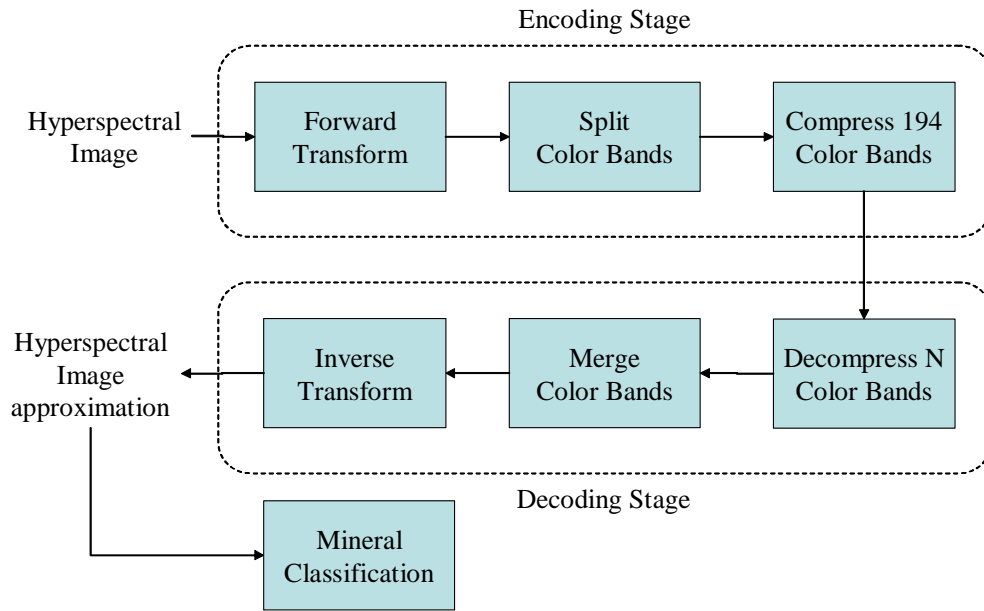


Figure 5. Color scalability algorithm.

Alternative Method: Ordered by Variance

This method requires the use of a parameter to measure component band importance. One important assumption is made here: the variance of a component band is directly related to the amount of information it contains. Therefore, the variance of each component band was used to measure its importance. In other words, bands with higher variances are more important and are decompressed first than bands with smaller variances. The procedure followed for this method is almost the same as the first one. The only differences include having to calculate the variance of each component band and creating a file with the order of the bands according to their variances.

The variance of each band is calculated once the Hyperspectral image has been split into 194 files. Their use is to distinguish which component bands contain more information. Therefore, a file with the order of the variances is created. In other words, the number of the component band with highest variance is saved first, followed by the number of the second highest variance band, and so on. This file is not strictly required by the decoder, but it is created in order to avoid having to recalculate variances each time a file needs to be decompressed. The file is used to know the order in which component bands need to be decompressed. When the required number of component bands has been reached, the procedure described previously is followed exactly. Although this method is expected to yield better classification performance, many of the advantages for the previous method are still present. For instance, decompression times are also reduced. The only additional computations come from the calculation of the variance for each component band, but this is done only one time for each image. In addition, the overhead created by this method is very small, almost negligible.

QUALITY SCALABILITY USING JPEG2000

Quality scalability works by finding optimal bit allocations among different sources of information, such as component bands. There exist numerous techniques to find this optimal bit allocation. The objective of this allocation is to minimize a measure of error, subject to the total bit rate constraint. MSE is commonly used as the error parameter because of many mathematical advantages that it has, such as being a continuous function that can be differentiated when needed. In this manner, a technique called Lagrange multipliers is used to obtain this bit allocation among the component bands, which is detailed in [5] and [2]. Lagrange multipliers are used because a set of equations is subject to a constraint, which in this case is the available bit budget. An advantage of using Lagrange multipliers is the simplicity involved. The set of equations that solve the problem are usually simple, and their calculation generally involves few computations. Let R be the available bit rate that will be distributed among the N

component bands. The constraint in this case will be

$$R = \frac{1}{N} \sum_{k=1}^N R_k \text{ and}$$

$$g(x) = 0 = R - \frac{1}{N} \sum_{k=1}^N R_k .$$

Here, R_k is the bit rate assigned to each component band. Using MSE to measure error, the distortion introduced in the k th band, σ_{rk}^2 is directly related to the quantizer error, and it is defined as

$$\sigma_{rk}^2 = \varepsilon^2 2^{-2R_k} \sigma_k^2 ,$$

where ε^2 is a constant dependent on the input distribution and quantizer, which will later cancel out. This means that this technique is appropriate for any type of distribution or quantizer. The input variance of each component band is denoted by σ_k^2 . The function to be minimized is the sum of these errors given by

$$f(x) = \sigma_r^2 = \sum_{k=1}^N \sigma_{rk}^2 = \sum_{k=1}^N \varepsilon^2 2^{-2R_k} \sigma_k^2 .$$

The method of Lagrange multipliers indicate that such problem can be solved by

$$\nabla f(x) + \lambda \nabla g(x) = 0$$

$$\frac{\partial}{\partial R_k} \left(\sum_{k=1}^N \varepsilon^2 2^{-2R_k} \sigma_k^2 + \lambda \left[R - \frac{1}{N} \sum_{k=1}^N R_k \right] \right) = 0 .$$

Solving this equation for R_k will eventually lead to the next result we arrive at

$$R_k = R + \frac{1}{2} \log_2 \frac{\sigma_k^2}{\left[\prod_{j=1}^N (\sigma_j^2) \right]^{1/N}} ,$$

which displays the relationship between the variance of each component band and the geometric mean of the variances to the fraction of the bit budget assigned to it. In other words, just like when using component scalability, component bands with higher variances are believed to be more important than others. The difference now is that component bands might not be decompressed completely; instead, truncated bit streams defined by the JPEG2000 standard will be decoded. These shorter bit streams will always be close to being optimal for lossy compression when using multiple layers during encoding.

Using Log of Variances

The method of Lagrange multipliers reaches a very simple set of equations as a result. Each component band's bit rate will be the sum of the desired target bit rate and half of the \log_2 of the ratio of the input variance to the geometric mean of all the variances. The latter part of the equation is what gives name to the result of using Lagrange multipliers: *log of variances*. Implementing these equations is relatively straightforward and does not require much processing time. The only computationally-intensive task that has to be performed before solving these equations is calculating the variance of each component band. The exact same procedure was followed when using color scalability, with the difference that the method of Lagrange multipliers will require the use of the actual values of each variance. In this manner, the variance of each component band is calculated independently and then saved for later use. This creates an additional file, which is very small in size. In fact, using a 32-bit coefficient to represent each variance will be enough. This results in an additional file of size $4 \times 194 = 776$ bytes, which is less than 1 KB, thus creating relatively no overhead.

The actual procedure followed is somewhat similar to the one used when doing color scalability. Figure 6 shows the algorithm and presents some differences. The actual bit rate allocation has to be calculated each time a new compression rate is desired. A problem of using Lagrange multipliers is that the equations to be solved are not

forced to result in non-negative numbers. In other words, the optimal bit allocation might require some component bands to have negative bit rates, which is obviously not possible. Bands not being decompressed will have a bit rate of zero, which will result in an increase of the target rate. Therefore, an iterative process is followed to prevent going over the target rate. Bands with higher variances are assigned their bit rate according to the equations given by the method of Lagrange multipliers. Bands with negative bit rates are assigned a zero bit rate instead. The new bit rate mean (calculated in bits per pixel per band) is calculated. If the target rate has been reached the process stops, if not, the next component band is assigned and the process is repeated. Once the process is stopped, component bands are decompressed and merged together. The inverse transform is performed and the image is classified. This algorithm was repeated for several bit rates: 0.25, 0.5, 1.0, 1.5, and 2 bits per pixel per band.

The purpose of using quality scalability is to minimize distortion introduced when quantizing or compressing coefficients. Therefore, this approach might result in slightly better mineral classification performance than using color scalability. On the other hand, one of its disadvantages is that optimal results will not be necessarily obtained when using Lagrange multipliers. However, its simple implementation represents an alternative to avoiding excessive computations.

RESULTS

One of the objectives of this paper is not only to measure the compression performance of the DWT, but also to evaluate how much information is lost when using lossy algorithms. In this manner, the Hyperspectral images were classified and the results were used to compare the DWT to PCA—which achieves most image compression possible and results in better classification performance.

Mineral classification was focused on two similar materials: Aluminum-Rich and Aluminum-Poor mica. Software used to classify minerals use data of a Hyperspectral image and assign a mineral to each of its pixels. This is achieved by assigning each pixel vector of length 194 to one of the spectral signatures used for classification. The problem is that Al-Rich and Al-Poor mica have almost identical spectral signatures. Therefore, the classification of both types of mica represents a good parameter to measure classification performance. In addition, other minerals have a similar spectral signature, which increases the complexity of the problem. Figure 6 shows the algorithm followed to classify minerals and compare them.

There exist numerous methods to classify Hyperspectral images. However, in order to maintain consistency with our past results and the classifier's popularity, the usage of SAM with PCA was recommended. The SAM classifier requires only one parameter, which is the threshold (in degrees.) This parameter was kept at 5° for all cases. When an image has been classified, it has to be compared to a file containing correct mineral classification. This file is called ground truth and was obtained by classifying the original image. Ground truth was compared to each classification of compressed images, which results in three different possibilities when comparing two pixels. Table 6.1 displays the events possible and their meaning when comparing images.

Table 1. Hit, Miss, and False Alarm definitions

Event	Ground Truth	Compressed Image	Definition
Hit	Mica	Mica	Mica is classified correctly
Miss	Mica	Other	Mica was not detected
False Alarm	Other	Mica	Mica was erroneously found

It is obvious that hit rates and miss rates will always add up to 100%. This is why only hit and false alarm rates will be shown and described in this chapter. The purpose of each algorithm used was to obtain a 90% hit rate and have a false alarm rate as low as possible.

Classification Performance using Color Scalability (First Method)

Previous results showed that using the PCA as the cross-component transform produced hit rates over 90% when using only 30 of the 194 bands. Since the DWT does not achieve as good of an ordering of data as the PCA, it is expected that classification performance will be somewhat degraded. Figure 6 shows the comparison between the PCA and DWT when classifying Al-Rich Mica.

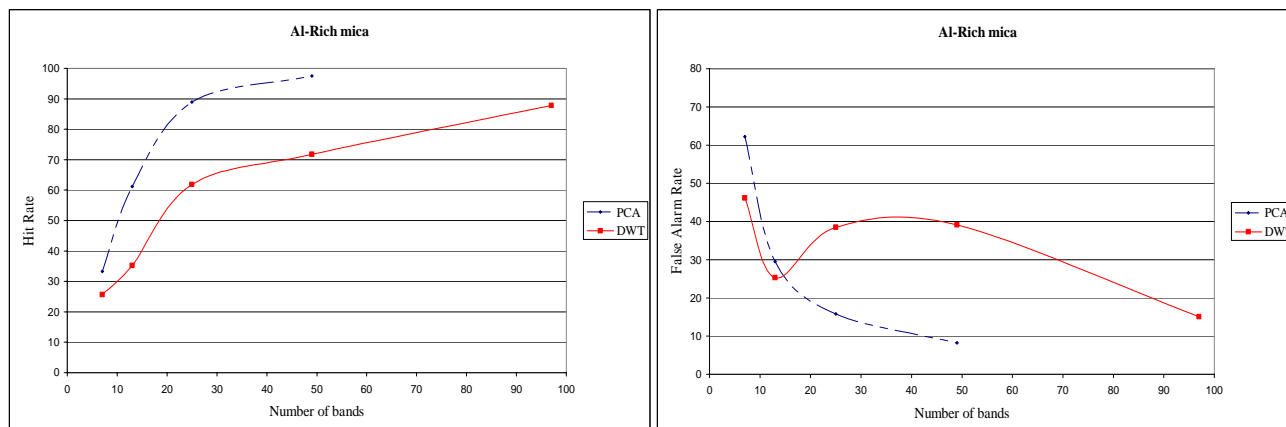


Figure 6. PCA and DWT hit rate and false alarm plot using Color Scalability (1st Method)—Al-Poor Mica.

Classification performance were found different for both minerals. Hit rate increases slowly, while false alarm rate will also increase sometimes. The only difference worth noting is that the 90% hit rate mark was achieved when using 97 component bands. The actual hit rate was 93.42% having a false alarm rate of 10.01%. The difference between the PCA and DWT performances was somewhat expected due to the data ordering achieved by both transforms. If the assumption made that input variance is directly related to the amount of information contained in a component band, then the alternative method will achieve better classification performance.

Classification Performance using Color Scalability (Alternative Method)

Sorting the component bands is expected to yield better classification performance. Since the PCA transform is ideal in the sense that all the component bands will always be ordered optimally, then sorting these bands will only decrease classification performance. Therefore, this method had not to be performed for the PCA transform, and the results are equal to those of the previous section. Therefore, only the comparison between this method and the previous will be described.

It can be seen that this method starts with worse classification performance. On the other hand, this method begins to reach the same performance when using about 37 component bands, and actually performs better after 49 component bands. Figure 7 also shows a problem of using the DWT: using wavelets has a tendency to increase false alarm rates at specific numbers of component bands. For example, when using 49 bands the false alarm rate for both cases is over 39%, but using the alternative method yields a hit rate of 88%, which is almost 17% more than the first method. The plot in Figure 7 also shows that 90% hit rate can be achieved when using about 57 component bands—a significant improvement over the first method, which did not even reach this rate when using 97 bands. The performance comparison when classifying Al-Poor Mica is somewhat similar.

Once again, the classification performance is deteriorated but starts increasing very rapidly. In fact, this method starts doing better than the first when using 25 component bands, with the only exception that false alarm rate is still high at this point. On the other hand, using 49 component bands decreases false alarm rate to 6.36%, which is very close to the value of 3.92% when using the PCA. In addition, this method achieves 87.87%, only about 10% less, hit rate than the PCA. Of course, this loss has its advantages: a faster and data-independent transform. In other words, decompressing the same file will take less time when using the DWT. Moreover, the exact same transform will work for all compressed files, which is not the case for the PCA.

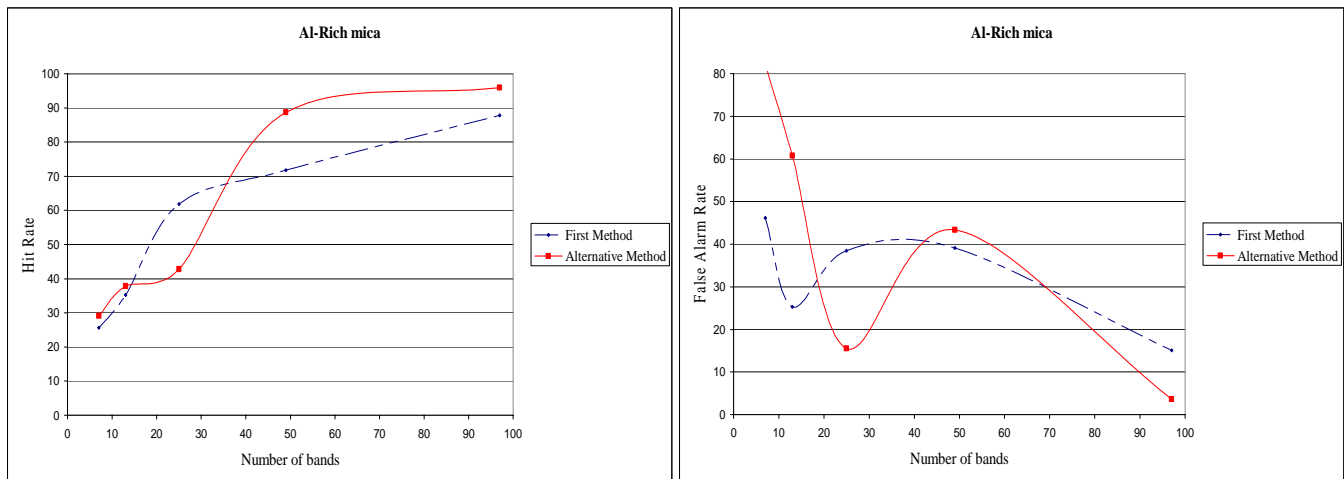


Figure 7. DWT hit rate and False Alarm plot using Color Scalability (Both Methods)—Al-Rich Mica.

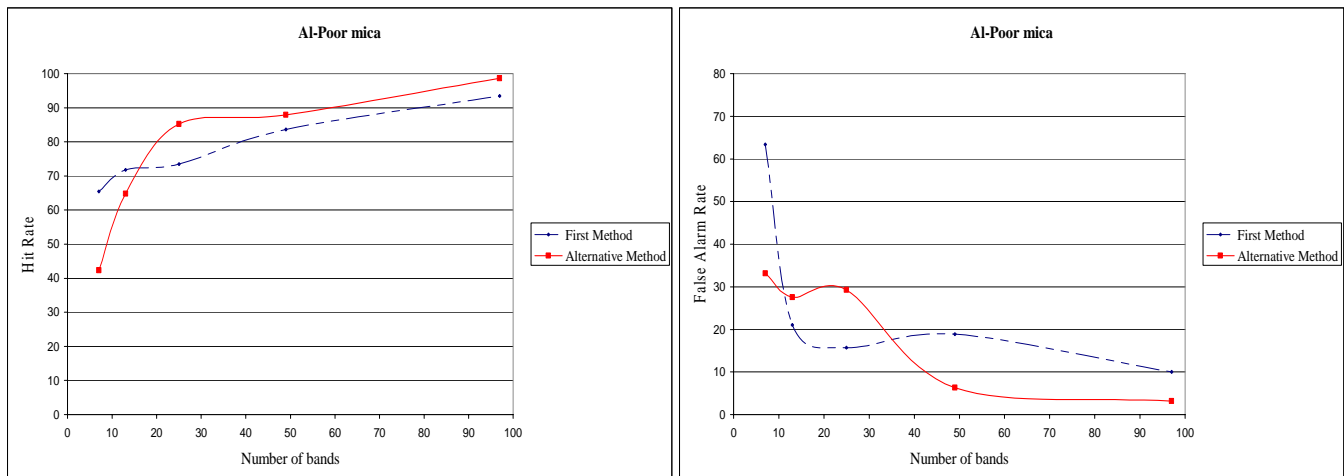


Figure 8. DWT false alarm rate plot using Color Scalability (Both Methods)—Al-Poor Mica.

Classification Performance using Quality Scalability

Classification results were not available for this type of scalability; therefore, the procedure was followed when using the PCA and the Discrete Wavelet Transform. Figure 9 shows the comparison between the two transforms when classifying Al-Rich Mica, at several bit rates, in bits per pixel per band (bpppb.)

The comparison demonstrates that the performance of the DWT is usually about 15% less than the PCA. In addition, the DWT reaches 90% hit rate at about 1.8 bpppb. Similarly, the false alarm rate has similar behavior to that of the PCA when using above 1.25 bpppb, which is less than 10%. Therefore, a bit rate of about 1.8 bpppb might be enough to have a classification performance as good as that of the PCA. This bit rate yields a compression ratio of 1.8 bpppb to the original 16 bpppb, which is equivalent to 11.25%. The classification algorithm was repeated for Al-Poor Mica. Figure 10 illustrates the performance of the DWT and PCA.

Figure 10 also shows that better results are achieved when classifying Al-Poor Mica. For example, the hit rate when using the DWT starts at about 20% less, but decreases constantly. In fact, the 90% hit rate mark is achieved under 1.25 bpppb, approximately at about 1.125 bpppb. Moreover, the false alarm when using the DWT is always very close to that obtained when using the PCA. Bit rates above 0.6 bpppb have false alarm rates smaller than 10%, which is very close to PCA results. A comparable performance between the DWT and the PCA can be obtained when using 1.125 bpppb, which is equivalent to a compression ratio of 7.03%. Quality scalability has the advantage

that it is capable of achieving better classification performance than color scalability. This is because quality scalability has the objective of minimizing a measure of distortion, which in this case is the MSE. In addition, the implementation of the method of Lagrange multipliers does not represent a computationally expensive task, yet good results are obtained.

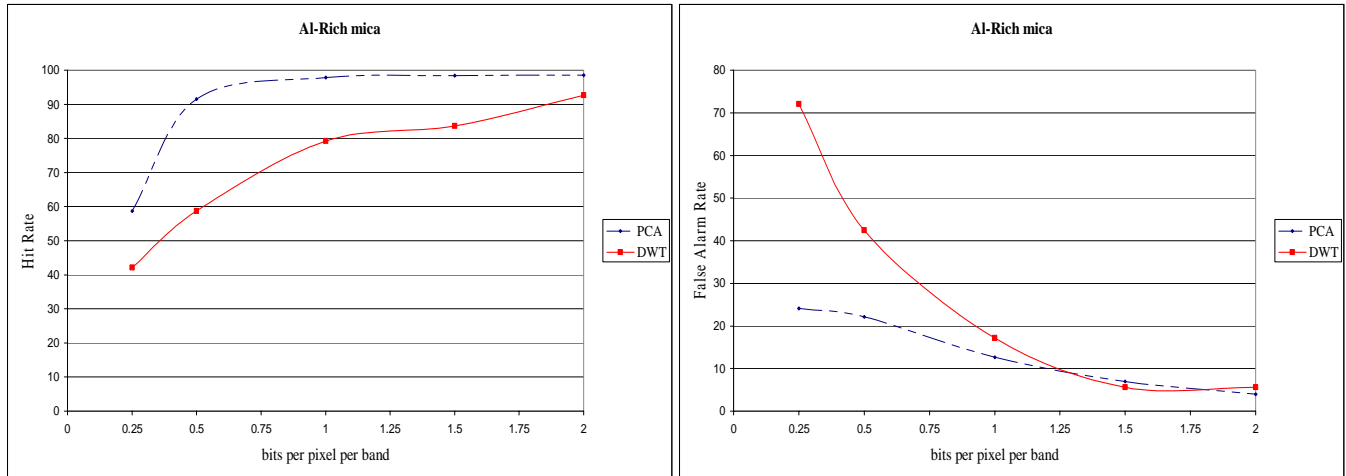


Figure 9. PCA and DWT false alarm rate plot using Quality Scalability—Al-Rich Mica.

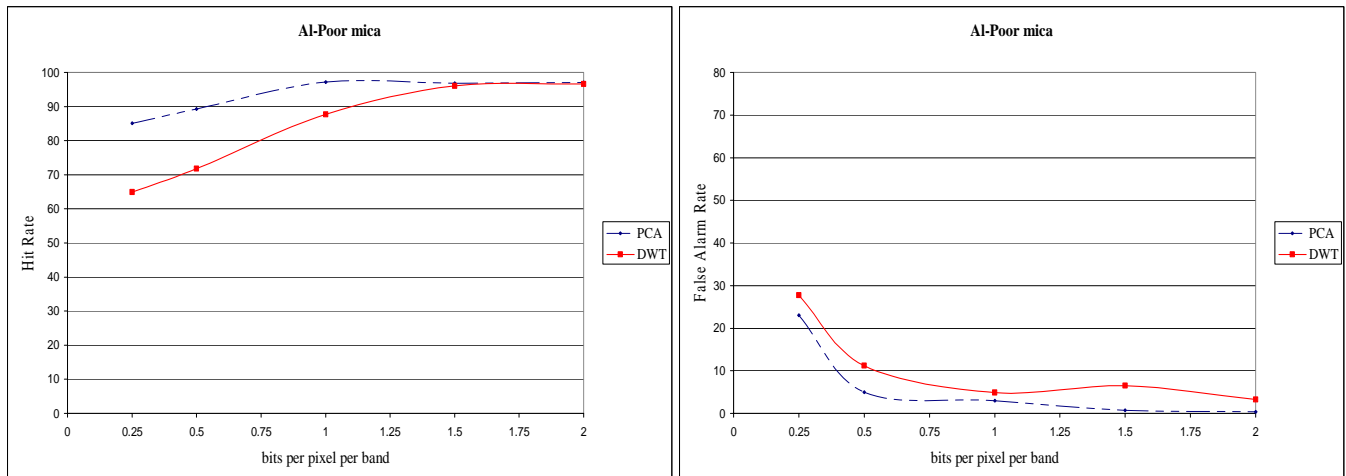


Figure 10. PCA and DWT false alarm rate plot using Quality Scalability—Al-Poor Mica.

Computational Efficiency

Previous sections detail the comparison in classification performance between the PCA and DWT. The PCA being the theoretically ideal transform is expected to have better results for both classification performance and compression. However, its biggest disadvantage is its computationally expensive implementation. Therefore, the DWT is a means of achieving faster decompression times while sacrificing classification performance. As seen previously, the use of the DWT decreases performance slightly. This section will now compare the time it takes for both transforms to be implemented. Only the actual implementation times required to transform an image were measured, since classification will require the exact amount of time for both transforms. Table 2 shows the results obtained for each transform.

Table 2. Implementation times for PCA and DWT

Transform →	PCA	DWT
Time (in seconds)	2837	378

As seen in the table, the DWT is about 7.5 times faster, which is a significant reduction in the time spent each time an image is decompressed. Furthermore, this transform does not require any overhead, such as the transformation matrix, which eliminates the need to calculate additional information each time an image is transformed. Therefore, the DWT is a very efficient and fast transform that does not require additional files to transform images.

CONCLUSIONS

The discrete wavelet transform represents a method to obtain reasonable classification performance results while maintaining a low number of computations. The results show that the DWT is capable of reaching 90% hit rate when using 50 component bands and color scalability. This number is less than double of that required by the PCA, which are 30 component bands. Similarly, when using quality scalability, the bit rate has to double or triple to achieve the 90% hit rate. However, the advantage of using the DWT is that it is faster—about 7 times faster than the PCA. Although PCA has a superior classification performance, these results demonstrate that the DWT could be used as an alternative in less complex classification problems. One other main advantage of DWT, that is allows dividing a huge data set in to parts and pre-processing each of these independent of the other. This enables parallel processing which cannot be done with PCA.

REFERENCES

- Vikram Jayaram, Bryan Usevitch and Olga Kosheleva, “Detection from Hyperspectral images compressed using rate distortion and optimization techniques under JPEG2000 Part 2,” in *Proc. IEEE 11th DSP Workshop & 3rd SPE Workshop* (cdrom), August 2004.
- Max Garces and Bryan Usevitch, “Classification performance of reduced color band Hyperspectral imagery compressed domain under JPEG2000 Part 2,” in *Proc. IEEE 11th DSP Workshop & 3rd SPE Workshop* (cdrom), August 2004.
- M. D. Pal, C. M. Brislawn, and S. P. Brumby, “Feature Extraction from Hyperspectral Images Compressed using the JPEG-2000 Standard,” in *Proceedings of Fifth IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI’02)*, Santa Fe, NM, pp. 168–172, April 2002.
- Silpa Attluri, Vikram Jayaram and Bryan Usevitch, “Optimized Rate allocation of Hyperspectral Images in compressed domain under JPEG2000 part 2” in *Proc. IEEE 2006 Southwest Symposium on Image Analysis & Interpretation (SSIAI)*, Denver, Colorado, March 2006.
- D. Taubman and M. Marcellin, *JPEG2000 Image Compression Fundamentals, Standards and Practice*. Boston, Dordrecht, London: Kluwer, 2002.
- Bryan Usevitch, “A Tutorial on Modern Lossy Wavelet Image Compression: Foundations of JPEG2000”, *IEEE Signal Processing Magazine*, Vol. 18, Issue 5, September 2001.
- T. J. Cudahy, R. Hewson, J. F. Huntington, M. A. Quigley and P. S. Barry, “The performance of the Satellite borne Hyperion Hyperspectral VNIR–SWIR imaging system for mineral mapping at Mount Fitton, South Australia,” in *Proceedings of IEEE International Geosciences and Remote Sensing Symposium IGARSS01*, Sydney, July 2001.
- Shaw Gary and Manolakis Dimitris, “Signal Processing for Hyperspectral Image Exploitation”, *IEEE Signal Processing Magazine*, Vol. 19, Issue 1, January 2002, ISSN 1053-5888.
- Said Amir and William A. Pearlman, “Signal Processing for Hyperspectral Image Exploitation”, *IEEE Signal Processing Magazine*, Vol. 19, issue 1, January 2002, ISSN 1053-5888.