Abstract

Future space-based, remote sensing systems will have data transmission requirements that exceed available downlinks, necessitating the use of lossy compression techniques for multispectral data. In this paper, we describe several algorithms for lossy compression of multispectral data which combine spectral decorrelation techniques with an adaptive, wavelet-based, image compression algorithm to exploit both spectral and spatial correlation. We compare the performance of several different spectral decorrelation techniques, including wavelet transformation in the spectral dimension. The performance of each technique is evaluated at compression ratios ranging from 4:1 to 16:1. Performance measures used are visual examination, conventional distortion measures, and multispectral classification results. We also introduce a family of distortion metrics that are designed to quantify and predict the effect of compression artifacts on multispectral classification of the reconstructed data.

1. Introduction

In space-based, remote sensing systems, the limited ability to transmit sensor data to the ground places a major constraint on system feasibility. Available relay systems and direct downlink capabilities are not scaled to the data-transmission requirements for wide-area, high-resolution remote sensing systems envisioned for sensor systems of the year 2000 and beyond. Assuming data rates on the order of gigabits/sec for an advanced multispectral remote sensor system and a 600Mbps ATDRSS relay link, compression ratios on the order of 5-15:1 are required to transmit sensor output in real time. Since lossless compression techniques are not expected to achieve average compression ratios greater than 2.5:1, there is clearly a need to develop lossy compression techniques for multispectral data.

Previous work in the area of lossy multispectral compression has investigated a variety of different techniques. Baker and Tse1 evaluated the performance of predictive coding, transform coding, and several vector quantization (VQ) techniques. In this work, only spectral correlations were exploited. The majority of other VQ techniques reported use VQ to exploit spatial correlations, and use predictive techniques (linear2, nonlinear3, feature4, and polynomial5) to exploit spectral correlations. In the wavelet transform-based techniques that have been reported6,7, a Karhunen-Loeve (KLT)8 transform or an approximation to it is performed prior to wavelet transformation to remove spectral redundancy in the data.

In this work, we use the wavelet transform in combination with several spectral decorrelation techniques to exploit both spectral and spatial correlation. Although the KLT is the optimum transform for the removal of spectral redundancy, it has historically been considered too computationally complex for real-time, on-board spacecraft implementation. In a previous paper9, we studied the performance of several prediction schemes to remove spectral redundancy. In this paper we examine the use of the wavelet transform to remove both spectral and spatial re-
dundancy. Both the prediction schemes and wavelet transform techniques are amenable to real-time implementation.

In addition, of greatest importance for multispectral remote sensing systems is the requirement that the compression process minimize the degradation of spatial and spectral fidelity to ensure that data exploitation is not compromised. Therefore, evaluation of lossy multispectral data compression techniques should include data exploitation simulations. However, comparison of exploitation performance is time consuming and is often impractical for compression algorithm development or parameter optimization. Conventional distortion measures (such as MSE or SNR) are not application sensitive and often do not accurately measure the effect of distortions on data exploitation. What is desired are quantitative degradation measures for exploitation algorithm performance characterization and prediction.

To address the need for meaningful image quality metrics, we introduce a set of metrics designed to quantify and predict the effect of compression artifacts on the performance of multispectral classification algorithms. These metrics, known as the Spectral Covariance Measures, are derived from the covariance matrices of the original, decompressed, and/or residual multispectral images. The goal of such metrics is to provide consistent predictive relationships between the quantitative distortion measure and a given application, such as Maximum Likelihood Multispectral Classification. Results are provided for the most promising of these measures, known as the Sum Delta Covariance Measure.

We simulate the performance of each compression algorithm on four multispectral (MS) images at compression ratios ranging from 4:1 to 16:1. An MS image consists of 8 co-registered 512x512 images, each representing a spectral band ranging from the Visible (Band 1) to the Near IR (Band 8). Performance measures used to evaluate the decompressed imagery are visual examination, conventional distortion measures (Mean Square Error), the Sum Delta Covariance Measure, and the results of Maximum Likelihood multispectral classification. We use these measures to determine the best spectral decorrelation technique, and to evaluate how well the Sum Delta Covariance Measure predicts multispectral classification performance.

The major contributions of this paper are simulation and performance evaluation of several different spectral decorrelation techniques, and preliminary results on the correlation between the Sum Delta Covariance Measure and Maximum Likelihood multispectral classification performance.

2. Compression Algorithm Description

A block diagram of the compression algorithms evaluated in this paper is shown in Fig. 1. The compression algorithms consist of a spectral decorrelation stage, a wavelet transformation stage, a rate allocation stage, a quantization stage, and an entropy coder stage. Each of these stages is described below.

2.1 Spectral Decorrelation Stage

We evaluated six different spectral decorrelation techniques: 1) Spatial-only (i.e., no spectral decorrelation), 2) Karhunen-Loeve transform (KLT), 3) Prediction with Two Reference Bands, 4) Band-to-band successive subtraction, 5) One dimensional wavelet transformation, and 6) Three dimensional wavelet transformation. In the Spatial-only technique, no spectral decorrelation is performed. Our purpose in evaluating this technique is to determine how much compression improvement (as measured by image quality and exploitability) can be obtained by
Fig. 1 Block Diagram of The Multispectral Compression Algorithms.

exploiting band-to-band spectral correlation. We also included the KLT in our evaluations so that its performance could be used as a reference to evaluate the performance of the other spectral decorrelation techniques.

Techniques 3 and 4 are differential schemes, in which the pixel values of a spectral band are replaced by the difference between the pixel values of the band and a predicted pixel value. In both schemes, the predicted value is obtained by using the value of a pixel at the same location, but in a different spectral band (known as the reference band). The motivation for these techniques is that because of spectral correlation, the predicted pixel value should be a reasonable estimate of the actual pixel value. The resulting differential band will have a lower entropy than the original band and will, therefore, be easier to compress. In the Two Reference Band approach (Technique 3), the predicted values for Bands 1, 2, and 4-6 are obtained by using the values of Band 3, and for Band 8, the predicted values are those of Band 7. In this technique, the values of the reference bands (Bands 3 and 7) are not changed. In the Successive subtraction approach (Technique 4), the reference band is just the next adjacent spectral band. For example, the reference band for Band 8 is Band 7, the reference band for Band 7 is Band 6, etc. In this technique, the pixel values of Band 1 are not changed. To improve the performance of these two techniques, a normalization is performed prior to subtraction: the mean of each band is subtracted and the band variances are made identical by multiplication by a scaling factor.

In Techniques 5 and 6, we use the wavelet transform as a spectral decorrelation technique. In Technique 5, we perform a one dimensional wavelet transform on each multispectral pixel, prior to performing a two dimensional wavelet transform on each decorrelated band. In Technique 6, we perform a three dimensional wavelet transform to simultaneously remove both spectral and spatial redundancy. In both techniques, the wavelet filters used in the spectral dimension are the Haar (or Daubechies 1) filters. We use these filters because their implementation requires only two filter taps, which, with 8 spectral bands, permits a three level transform in the spectral dimension. As in the prediction schemes described above, prior to performing the wavelet transform, we subtract the mean of each spectral band and make the variances of the bands equal - in this case equal to the maximum variance of the bands.

All of the spectral decorrelation techniques mentioned above are reversible - the original pixel values can be obtained from the spectrally decorrelated values. With the possible exception of the KLT, these techniques are also amenable to real-time implementation since they involve relatively few computations per multispectral pixel.

2.2 Wavelet Transformation Stage

After the spectral decorrelation stage (except in Technique 6 above), we apply a two-dimensional discrete wavelet transform (DWT) to the decorrelated spectral bands to reduce
pixel-to-pixel spatial redundancy. The wavelet transform is a subband decomposition, in which a bank of bandpass filters splits an image into a number of separate, spatial frequency components, called subbands. The motivation for this decomposition is that the subbands can be encoded more efficiently than the original image. Typically, different bit rates and even different coding techniques are used for each subband to take advantage of the statistical properties of the subband and to control or shape the coding errors in an optimal fashion.

Wavelets are a recently developed class of subband filters in which the impulse response of the filters are orthogonal to one another and are all scaled versions of a single function known as the wavelet. The subbands produced by the transform have good redundancy removal properties, are orientation specific, and contain multiresolution information on both the location and scale of features, particularly edges or discontinuities in the image. The ability to efficiently represent image features (particularly edges) is one of the reasons that wavelet-based compression schemes provide reconstructed images with good visual quality. The 2D DWT used in this paper is equivalent to a pyramid subband decomposition, where the bandwidths of the subbands are related by powers of two and represent an octave-based frequency decomposition. The transform is implemented using two finite impulse response filters which are applied recursively to the lowest frequency subband. In this paper, the 2D wavelet transformation stage consists of a 6-level DWT, using the Daubechies 9-7 biorthogonal, linear phase filters. Symmetric edge reflection is used to avoid the introduction of discontinuities due to image boundaries.

In our implementation of a three dimensional wavelet transform, we use the Haar filter in the spectral dimension and the Daubechies 9-7 biorthogonal filters in the spatial dimensions, with symmetric edge reflection at the data boundaries in all three dimensions. The 3D transform consists of 6 levels: 3 levels performed on all three dimensions, and 3 levels performed only on the spatial dimensions.

2.3 Rate Allocation Stage

The purpose of the rate allocation stage is to select the rate (in bits/coefficient) of the wavelet subbands so that the desired compression ratio is achieved with minimum distortion in the reconstructed images. The general approach is to allocate higher rates to subbands that contain more information. Subbands allocated higher rates will be quantized with less distortion or error (the difference between the coefficient value and its quantized value). In a previous paper, we examined the performance of four different rate allocation techniques. In three of these techniques, rate allocation is performed in two stages. In the first stage which occurs after spectral decorrelation, rate is allocated among the decorrelated bands in the spatial domain. The decorrelated bands are then treated as separate, independent images in the second stage, which allocates rate among the different wavelet subbands. In the fourth technique, all of the spectral bands are treated as a single dataset and rate allocation is performed in a single stage after spectral decorrelation and wavelet transformation. Our simulation results indicate that the fourth approach has the best performance. Use of any of the two stage rate allocation techniques results in significantly poorer performance. Thus, we use the single stage technique exclusively in this analysis.

After spectral decorrelation and wavelet transformation, we determine a bit rate/subband using the following formula which allocates rate based on the variance of the subband:

$$\text{Rate} = \frac{\text{Variance}}{\text{Coefficient}}$$
\[ R_i = \Theta + \frac{1}{2} \log_2 \left( \frac{\sigma_i^2}{\prod_{k=1}^{S} \sigma_k^{2N_k}} \right)^{1/N} \]  

where \( R_i \) is the allocated rate for subband \( i \), \( \sigma_i^2 \) is the variance of subband \( i \), \( \Theta \) is the desired average rate for the dataset, \( S \) is the number of subbands, \( N_k \) is the number of coefficients in subband \( k \), and \( N \) is the total number of coefficients in the dataset (equal to the number of bands times the number of pixels in the band). For spectral decorrelation Techniques 1-5, the number of subbands \( S \) is equal to 152 (8 spectral bands times 19 subbands/spectral band) and for Technique 6, the number of subbands is 31.

Eq. 1 is the rate allocation formula found in [13] that we have modified to account for the different sizes of the wavelet subbands. One problem with this formula is that if the variance of a subband is too small compared to the geometric mean of all of the subband variances, then this formula will result in a negative rate for the subband. In this case, we remove from Eq. 1 those subbands allocated a negative rate in the previous calculation and recalculate the \( R_i \). This process generally requires at most 2-3 iterations to converge. The subbands that have been removed are not coded. All of the coefficient values in these subbands are set to zero.

2.4 Quantization Stage

The quantization stage consists of two parts: stepsize selection and uniform quantization. The purpose of the stepsize selection process is to determine a quantizer stepsize for each subband so that the quantized subband will be entropy coded at the allocated bit rate. We use a search algorithm that iteratively selects a stepsize, quantizes the subband, and then measures the first order entropy of the quantized subband to determine if the quantized subband meets its allocated rate, which indicates the suitability of the selected quantizer stepsize. After a stepsize is selected for each subband, the wavelet coefficients of the subband are quantized by dividing the coefficient value by the stepsize and rounding to the nearest integer.

Currently the iterative search algorithm used to determine quantizer stepsize is too computationally intensive for real-time implementation. A future effort is to replace the iterative search algorithm with a table lookup approach, developed through training, that selects quantizer stepsize based on the desired rate and variance of the subband.

2.5 Entropy Coding Stage

In the entropy coding stage the quantized wavelet coefficients are mapped into a set of variable-length code words. More frequently used values are mapped to short length code words and less frequently used values to long code words. Compression is achieved because the average number of bits to represent the output codewords is less than the average number of bits used to represent the quantized wavelet coefficients.

Our entropy coder is a hybrid that combines two well known techniques: the Rice coder\(^{14}\) and an arithmetic coder\(^{15}\). We use these two techniques in a complementary fashion. The Rice coder works well on short sequences and on sequences that have a first order entropy greater than 2bits/symbol. The arithmetic coder works well on long sequences that have low first order entropies (i.e., < 2bits/symbol). In coding each subband, we select the technique based on the size of the subband and its allocated bit rate. Performance simulations of this hybrid entropy
coder demonstrate coding efficiencies within 5-10% of information theoretical performance (based on first order entropy), which is significantly better than the performance of either technique alone.

2.6 Algorithm Summary

In Table 1, we list the different compression algorithms evaluated in this paper. For each algorithm in the table, we indicate the spectral decorrelation technique that is used. We also assign to each technique a short alpha-numeric symbol that we use to identify the specific technique in the graphs and tables of this paper.

<table>
<thead>
<tr>
<th>Algorithm Symbol</th>
<th>Spectral Decorrelation Technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial-only</td>
<td>None</td>
</tr>
<tr>
<td>KLT</td>
<td>Karhunen-Loeve Transform</td>
</tr>
<tr>
<td>PRED1</td>
<td>Two Reference Band Predictor</td>
</tr>
<tr>
<td>PRED2</td>
<td>Successive Subtraction Predictor</td>
</tr>
<tr>
<td>WV1D</td>
<td>1D Wavelet Transform</td>
</tr>
<tr>
<td>WV3D</td>
<td>3D Wavelet Transform</td>
</tr>
</tbody>
</table>

Table 1. Multispectral Compression Algorithms

3. Performance Measures and Methodology

The goal of the compression schemes studied in this paper is to achieve a desired compression ratio with minimum distortion in the reconstructed MS image. One of the most common criteria used to measure distortion is the Mean Square Error (MSE):

\[
MSE = \frac{1}{NN_B} \sum_{i=1}^{N} \sum_{j=1}^{N} (X_{ij} - \hat{X}_{ij})^2,
\]

where \(N\) is the number of pixels in the spectral band, \(N_B\) is the number of spectral bands in the dataset, \(X_{ij}\) is the original pixel value of pixel \(j\) in Band \(i\) and \(\hat{X}_{ij}\) is the pixel value after compression and decompression. We also measure the MSE for individual spectral bands. To calculate the MSE/band, we use an equation similar to Eq. 2, except that the summation is only over the pixels in the band.

Another criteria that we use to evaluate performance is a visual comparison between the reconstructed and original spectral bands of the MS images. We also viewed error images of the individual bands to study the types of errors introduced. The error images are constructed by taking the difference between the original and the reconstructed image and then scaling the errors to be in the range of 0-255 for display.

3.1 Spectral Covariance Measures

As a parallel effort to compression algorithm development and evaluation, we are investigating application specific distortion metrics. The objective of such a metric is to provide a predictive mapping between metric value and the change in performance of specific data exploitation applications after any process which introduces distortion to data, such as lossy compression. If such a relationship can be identified consistently between the metric and the application, then it will only be necessary to compute the metric to predict how the distortion process will affect the application. Ideally, such a metric should be straightforward to calculate and is
particularly useful if it correlates well to several applications (albeit perhaps via different predictive relationships).

For multispectral applications we have developed and investigated a set of measures called Spectral Covariance Measures. These measures are derived from the spectral covariance matrices of the original, decompressed and/or residual images. Design of these metrics is motivated by the fact that spectral principal components are the basis of many spectral feature extractors and that spectral covariance describes the degree of linear correlation between bands. An additional motivation is that some common classifiers, such as Mahalanobis Distance and Maximum Likelihood Classifiers, explicitly rely on spectral covariance to perform classification. We have investigated whether predictive relationships exist between these metrics and Maximum Likelihood Classifier performance. The most promising of the metrics, with respect to Multispectral Classification, is called the Sum Delta Covariance (SDC) metric. The SDC metric is computed as follows:

$$ SDC = \sum_{\text{bands}} \left| \text{Cov}_{\text{original}} - \text{Cov}_{\text{compressed}} \right|, $$

where all covariances are normalized. In this work we compare how well MSE and the SDC measure predict multispectral classifier performance.

3.2 Multispectral Classification Methodology

The fourth criteria used to evaluate the performance of the compression algorithms is to compare how well the compressed/decompressed imagery can be classified compared to the original multispectral (MS) images. A signature database defines the statistical characteristics of the proposed classes and is generated via training with representative MS data. The signature database is subsequently used by the MS classifier in conjunction with a decision rule to classify MS pixels. In general training may be supervised or unsupervised. For this study, unsupervised training is performed, due to lack of available ground truth. Both training and MS Classification are performed within the ERDAS GIS (Geographic Information Systems) and Image Processing environment. Unsupervised training is performed by the ISODATA clustering algorithm, and actual MS classification is performed using a Maximum Likelihood Decision Rule. Visual examination and measured signature divergence are used to iteratively edit and merge signatures derived from the original training images, yielding the final signature database.

In general we would like to use as much training data as possible to develop the signature databases, however for this effort we have a limited set of calibrated, registered MS images representing the spectral bands of immediate interest (Visible to Near IR). Specifically, this analysis is based upon 4 calibrated, co-registered MS images: 2 from each of 2 MS bandsets. These datasets are referred to as Airfield 1, Airfield 2, Urban 1 and Urban 2. Thus two signature databases are required for this analysis - one for each bandset. Eight spectral bands from each image were used. For this initial work, all eight bands were used for MS Classification. Future tasks will identify band subsets best suited for specific classification schemes and perform compression and exploitability analysis on these selected band subsets.

Each original MS image contains approximately 1000 X 700 MS pixels. For compression analysis, a 512X512 MS subimage was extracted from each image. The original (1000 X 700) images were used for classifier training. Thus each MS bandset's signature database is derived from two 1000 X 700 MS images. The image calibration data is used to "radiance normal-
size" the data prior to training, such that within an MS bandset, the mapping from digital count to radiance is consistent.

In order to evaluate the impact of several compression algorithms on MS classification, each of the normalized uncompressed 512X512 MS images is submitted to the MS Classifier, using the appropriate signature database. This classification is treated as "truth" and becomes the basis for comparing classification results after compression. Data compression is performed on imagery which has not been radiance normalized, because raw sensor data which is input to an on-board compressor is typically unnormalized. After reconstruction, the compressed MS image is normalized using the same calibration and normalization factors which have been applied to the corresponding uncompressed MS image. This data is then submitted to the MS Classifier, using the appropriate signature database. The number of correctly classified pixels after compression is computed, yielding the percent correct classification results. This is done for each compression algorithm, at each compression ratio, for each 512X512 image.

4. Simulation Results

4.1 Compression Algorithm Performance

In Fig. 2 we compare the performance of the different spectral decorrelation techniques. In these two graphs, we display Mean Square Error as a function of compression ratio. Fig. 2a contains the results for dataset Airfield 1 and Fig. 2b contains the results for dataset Urban 1. From both of these graphs, it is clear that the KLT spectral decorrelation technique results in the best (i.e., lowest MSE) performance. For both datasets, the performance of the Two Reference Band technique and the 1D wavelet technique are comparable and have performance close to that of the KLT technique. For the Airfield 1 dataset, the performance of the Successive Subtraction technique and 3D wavelet technique are comparable and are better than not exploiting spectral decorrelation (the Spatial-only approach). However, in the Urban 1 dataset, the Successive Subtraction technique is actually worse than the Spatial-only approach, while the 3D wavelet technique still results in better performance.

![Fig. 2. Comparison of Different Spectral Decorrelation Techniques. (a) Dataset Airfield 1 and (b) Dataset Urban 1.](image-url)
In Fig. 3 we show the results of performing multispectral classification on the reconstructed MS images. In both Figs. 3a and 3b, we display the percentage of MS pixels that are correctly classified as a function of compression ratio. As in Fig. 2, the KLT spectral decorrelation technique results in the best performance for both datasets. For dataset Airfield 1, the prediction schemes have similar classification performance, and both prediction techniques perform better than either the 1D or 3D wavelet-based techniques, which is a different relative performance ranking than the ranking obtained by comparing MSE performance. For dataset Urban 1, the classification performance of the 1D wavelet technique is almost as good as the KLT and significantly better than the prediction techniques or the 3D wavelet technique.

The relatively poor performance of the three dimensional wavelet transform approach may be due to the fact that there is a significantly smaller number of subbands (approximately a factor 5) in this approach than in any of the other approaches. The smaller number of subbands means that the subbands are larger than in the other approaches and, therefore, the bit rate allocation and quantization are more coarse. In other words, because the other techniques group the transform coefficients into a larger number of smaller groups, there is more flexibility in rate allocation and quantizer design. This additional flexibility translates into better performance.

4.2 Sum Delta Covariance vs. MSE Metric Performance Comparison

Because multispectral classification is applied to radiance normalized data, all MSE values used for metric evaluation are computed after radiance normalization of original and compressed imagery. Similarly, SDC is computed from radiance normalized data. Fig. 4 illustrates SDC vs. CR and MSE vs. CR for each compression algorithm for Airfield 1. When compared to Fig. 3a we see that neither SDC nor MSE consistently corresponds to the relative performance of the compression algorithms (as defined by classification accuracy).

In order to assess whether SDC shows promise as the basis of a predictive metric of classification accuracy, we have examined the correlation of both SDC and MSE to classification
accuracy. This is illustrated in Fig. 5 for each of the individual images. In these and the following figures, results are derived from 11 wavelet-based compression algorithms, including the six algorithms described in this paper and five algorithms described in a previous paper\(^9\). For any given image, SDC has only a slightly higher linear correlation to classification accuracy than
MSE. More important however, is what occurs when this correlation is examined over all images from both bandsets, as is illustrated in Fig. 6. When analyzed over both bandsets SDC has a notably higher correlation to classification accuracy than MSE. It appears that SDC is less sensitive than MSE to scene, sensor, and spectral variations. Thus it is possible that a refinement of the SDC measure will provide a useful predictive measure of classification accuracy.

![Graphs](a) SDC vs. Classification for Airfield 1, Airfield 2, Urban 1 and Urban 2 (b) SDC vs. Classification for Airfield 1, Airfield 2, Urban 1 and Urban 2

5. Conclusions

In this paper, we have evaluated the performance of a number of wavelet-based multispectral compression algorithms. All of the algorithms use the wavelet transform to reduce pixel-to-pixel spatial redundancy. The difference in the compression algorithms lies in the techniques used to reduce band-to-band spectral correlation. Simulations of each of the compression algorithms was performed on four 8-band multispectral images at four different compression ratios. Visual examination, Mean Square Error, the Sum Delta Covariance Measure, and the results of multispectral classification of the decompressed images were the criteria used to evaluate the performance of the different algorithms.

As expected, the results of the simulations indicate that the Karhunen-Loeve transform is the best spectral decorrelation technique. Good performance is obtained with either a one dimensional spectral wavelet transform or a simple prediction scheme in which the pixel values of one of two bands is used to predict the pixel values in the remaining spectral bands. The performance of the three dimensional wavelet transform and that of the Successive subtraction prediction scheme were, in general, better than not exploiting spectral redundancy, but were significantly poorer than the other spectral decorrelation techniques.

We have implemented and evaluated a spectral covariance based metric called the Sum Delta Covariance. This metric correlated to multispectral classification accuracy more strongly than MSE and appears to be less sensitive than MSE to scene, sensor, and spectral variations. Thus this measure shows promise as the basis of a metric which can be used to predict multispectral classification accuracy.

In future directions of this research, we will concentrate on three areas: 1) development of improved compression algorithms, 2) an examination of sensor systems issues and their im-
pact on compression algorithm design and performance, and 3) development of improved compression evaluation techniques. Our focus in developing better compression algorithms is to evaluate different quantization schemes. For each processing stage we will tune algorithmic parameters and approaches for real-time on-board spacecraft implementation. Sensor systems issues that we plan to investigate are the effects on compression performance due to spectral band misregistration and detector nonuniformities. In the area of compression evaluation techniques, we plan to refine our classification techniques, the spectral covariance measures, and develop other application-specific image quality measures.

Acknowledgment

Special thanks to Gloria Yang for her support in performing the MS Classification analysis.

References