

This work was written as part of one of the author's official duties as an Employee of the United States Government and is therefore a work of the United States Government. In accordance with 17 U.S.C. 105, no copyright protection is available for such works under U.S. Law.

Public Domain Mark 1.0

<https://creativecommons.org/publicdomain/mark/1.0/>

Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

Journal of
Applied Remote Sensing

**Progressive band selection for
satellite hyperspectral data
compression and transmission**

Kevin Fisher
Chein-I Chang

Progressive band selection for satellite hyperspectral data compression and transmission

Kevin Fisher^a and Chein-I Chang^b

^a NASA Goddard Space Flight Center, 8800 Greenbelt Road,
Greenbelt, Maryland, 20771, USA

^b University of Maryland, Baltimore County, Department of Computer Science and Electrical
Engineering, 1000 Hilltop Circle, Baltimore, Maryland, 21250, USA

Abstract. Efficient data transmission is an important part of satellite communication, particularly when large data volumes need to be downlinked to the Earth. One general approach to dealing with this dilemma is data compression, either lossless or lossy. For hyperspectral data, compression is specifically crucial due to its high inter-band spectral correlation, resulting from the use of hundreds of high spectral resolution bands for data collection. This paper develops a new approach, called progressive band selection (PBS), to achieve both data compression and data transmission, in the sense that data can be compressed and transmitted progressively. First, PBS prioritizes each spectral band by assigning a priority score based on its information content measured by a certain criterion. Then, bands are selected progressively according to the priority scores assigned to each spectral band. Consequently, data can be compressed and transmitted in a progressive fashion to meet the application's requirements; this task cannot be accomplished by most data compression techniques. Most importantly, PBS can be implemented in two opposite manners. One is forward progressive band selection (FPBS), which starts with a low number of bands and gradually improves data quality by including more bands progressively, based on their priority scores, until data quality is satisfactory. The other is backward progressive band selection (BPBS), which begins with a high number of spectral bands and progressively removes them in accordance with their priority scores, until data quality falls below a given tolerance level. In order to determine the lower and upper bounds on the number of bands used for FPBS and BPBS, we use a recently developed concept called virtual dimensionality (VD). We demonstrate the utility of PBS in compression and transmission for satellite communication with an experiment in land use and cover classification, which uses a dataset collected by the Hyperion instrument aboard NASA's EO-1 satellite.

Keywords: Backward progressive band selection (BPBS), Forward progressive band selection (FPBS), Progressive band selection, Virtual dimensionality (VD).

1 INTRODUCTION

One of the most important processes in satellite communication is data transmission, which requires downlinking data to the Earth. Due to the limited bandwidth of satellite transceivers, a general approach is to perform onboard data processing to reduce data volumes, so as to achieve efficient data transmission. A general approach is to perform data compression prior to data transmission. Data compression can be carried out in two ways, depending on how the desired information is preserved. One is lossless compression, which removes data redundancy without the loss of any information. This type of compression is generally performed by data compaction, an example of which is the reduction of high dimensional data to low dimensional data by transforms while retaining all information. The other is lossy compression, which removes unwanted or undesired information with little loss of information.

Hyperspectral imaging is becoming a more prevalent technique to collect data for remote sensing of the Earth. Due to its hundreds of contiguous spectral bands, hyperspectral datasets are enormous. Processing and transmitting such vast amounts of data without pre-processing is very costly. So, when hyperspectral data is used for satellite communication, it is highly desirable to perform data compression before the data is downlinked. In general, two types of hyperspectral data compression can be carried out. One is dimensionality reduction, which applies a specific transform to reduce the number of dimensions of the space in which the data resides, so that the data can be more manageable. Principal components analysis (PCA) is a widely used example of this approach. The other approach is band selection, which selects an appropriate subset of image bands to fulfill the same applications the full image can, to some extent. The difference between these two approaches is that dimensionality reduction uses information from all image bands to perform data compaction, with or without loss of information, whereas band selection discards the information provided by all the bands which are not selected, relying on only the selected bands to perform required tasks. A common issue arising in dimensionality reduction and band selection is to determine the number of dimensions, q that must be retained after reduction, and the number of bands, p which must be selected. Thus far, this issue still remains unsolved since this number is dependent on the target application. This paper takes the latter approach, using band selection for hyperspectral data compression.

Generally speaking, band selection takes advantage of high band correlation to remove band redundancy in order to accomplish many functions, including data compression, data storage, data transmission, target detection, and classification. When band selection is implemented, two crucial issues arising in band selection must be addressed: first, the number of bands that should be selected, p , and second, what criterion should be used to select the bands. Although many techniques have been proposed for band selection ([1-6], to name just a few), the general approach is to first fix p a priori, then select the appropriate bands in accordance with a custom-designed criterion. This paper develops a new approach to band selection, referred to as progressive band selection (PBS), to mitigate this dilemma by relating the dependency on p to a particular application.

The idea of PBS is to design a criterion based on a specific application at the user's discretion. Once a criterion is selected, it will be used to measure and calculate the amount of information content provided by each of the spectral bands to produce priority scores for all the bands. The significance of each spectral band is then ranked and prioritized in accordance with its assigned priority score. Finally, band selection is performed by selecting spectral bands based on this band prioritization in terms of the priority scores. The selected bands can be formed into a reduced dataset to be provided as input to an image analysis application, or packaged as a compressed version of the original image. This compression (carried out by deleting the lowest scoring bands) is lossy, but since the deleted bands are by definition the ones of lowest value to the users of the image and their applications, very little useful information is lost. The compressed image can be transmitted from a satellite in less time, or archived with fewer data storage resources, than the original.

There are two ways to carry out PBS. One is called forward progressive band selection (FPBS), which selects a small number of bands to begin with, then progressively adds new bands according to the band prioritization by priority scores, until the reduced image contains a sufficient amount of information for the target application. The process resulting from FPBS is called progressive band expansion (PBE). As a complete opposite approach to FPBS, PBS can be performed backwards, referred to as backward progressive band selection (BPBS). More specifically, BPBS first selects a high number of bands according to their priority scores, then gradually removes bands with lower priority scores until data quality has deteriorated to some extent, at which point the band selection process is terminated. This process is called progressive band reduction (PBR). By virtue of PBS there is no need to know the number of bands, p required for band selection since the FPBS can always begin

with $p = 1$, and BPBS can start with p equal to the total number of spectral bands, L . However, setting $p = 1$ or L is not realistic in practice. Fortunately, a concept called virtual dimensionality (VD), recently developed in [7-8], can be used to estimate the lower and upper bounds on the number of bands to select. Using VD, the FPBS can start with $p = \text{VD}$ and begin to add bands to the reduced image. On the other hand, according to studies in [6], it seems that twice VD can be used as an upper bound for BPBS. In this case, BPBS can start with $p = 2\text{VD}$ and gradually remove bands. In order to demonstrate the utility of FPBS and BPBS in real applications, a data set collected by the space-borne EO-1 Hyperion instrument is used for experiments in land cover/use applications.

2 CRITERIA FOR BAND SELECTION

Many criteria have been developed for band selection. [1-6] In this section, several popular criteria will be introduced and discussed briefly.

2.1 Statistics-Based Criteria

Since we assume no prior knowledge about the image, the criteria to be considered for band selection must be based solely on image data characteristics specified by statistics. In this case, each spectral band is treated as a random variable X with all image samples considered as its realizations. As a result, statistics of various orders can be derived and used as criteria to prioritize all spectral bands.

2.1.1 Variance (second-order central moment)

$$\sigma^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (1)$$

where x_i is the value of pixel i , \bar{x} is the mean value of all pixels in the band, and N is the number of pixels.

2.1.2 Skewness (third-order central moment)

$$\text{skewness} = \frac{1}{N-1} \sum_{i=1}^N \frac{(x_i - \bar{x})^3}{\sigma^3} \quad (2)$$

where σ is the standard deviation of the values of all pixels in the band.

2.1.3 Kurtosis (fourth-order central moment)

$$\text{kurtosis} = \frac{1}{N-1} \sum_{i=1}^N \frac{(x_i - \bar{x})^4}{\sigma^4} \quad (3)$$

2.2 Infinite-order statistics

Two criteria can be derived from infinite-order statistics: entropy and information divergence.

2.2.1 Entropy

Entropy is the classic measure of information content in a data set, defined as $H(X)$:

$$H(X) = -\sum_{i=1}^N p(x_i) \log_2 p(x_i) \quad (4)$$

where $\mathbf{p}(\mathbf{x}) = (p(x_1), p(x_2), \dots, p(x_N))^T$ is the probability mass function calculated from the histogram of pixels in the image X , and N is the number of different pixel values in the image. This equation calculates $H(X)$ in bits per pixel.

2.2.2 Information divergence

Information divergence [9] is an alternative measure of information content. It measures the discrepancy between two probability distributions. For our purpose, let $\mathbf{p}(\mathbf{x})$ denote the probability distribution of a spectral band image and let $\mathbf{g}(\mathbf{x}) = (g(x_1), g(x_2), \dots, g(x_N))^T$ be a Gaussian distribution with the same mean and variance as the image characterized by $\mathbf{p}(\mathbf{x})$. The more the distribution of data in the image deviates from the Gaussian model, the higher its information divergence (D), and the more information it is presumed to contain.

$$D(p; g) = \sum_{i=1}^N p(x_i) \log(p(x_i) / g(x_i)) + \sum_{i=1}^N g(x_i) \log(g(x_i) / p(x_i)) \quad (5)$$

2.3 Control criteria

This section presents two criteria that can be considered extreme opposite cases of the statistics-based criteria described in Section 2.2: uniform band selection (UBS) and random band selection (RBS). UBS is a deterministic band selection method which has nothing to do with image sample statistics, while RBS assumes image sample statistics to be fully random.

2.3.1 Uniform criterion

Uniform band selection (UBS) is a common practice in band selection which selects spectral bands uniformly over the entire set of bands. In this case, it assumes that there is no particular preference for any of the spectral bands. That is, for a given number of selected bands, p and total number of spectral bands, L , the bands to be selected start with any band j with $j < p$, band $j + \lceil L/p \rceil$, band $j + 2\lceil L/p \rceil$, band $j + 3\lceil L/p \rceil$, and so on until $j + p\lceil L/p \rceil$. If $j + p\lceil L/p \rceil > L$, then band $j + p\lceil L/p \rceil$ is replaced by band L . (Note that $\lceil x \rceil$ is defined as the largest integer $\leq x$). In other words, all bands in the band interval $\lceil L/p \rceil$ are considered to be equally important. In this case, we can select any set of bands as long as any two consecutive selected bands (except the first and last bands) are p bands apart.

2.3.2 Random band selection

In complete opposition to UBS, which can be considered a deterministic band selection method, random band selection (RBS) selects bands in a completely non-deterministic manner. Its criterion uses a pseudorandom number generator to produce a random priority score between 0 and 1 to each spectral band. All the spectral bands are then ranked and selected according to their assigned priority scores.

2.4 Sensor-based criterion

The sensor-based criterion uses basic information about the image sensor. This criterion assigns a zero score to bands known to be uncalibrated, masked, or of otherwise low quality, then gives random scores to the remaining bands. This optimization adds very little complexity to the random selection algorithm.

3 PROGRESSIVE BAND SELECTION

The criteria described in Section 2 are used to create a ranking among the image's spectral bands. The band with the highest priority score receives the highest rank, followed by the one with the next highest score, and so on through the band which received the lowest score. The system creates an initial subset image, then adds or removes bands from it. In general, the algorithm adds or removes the marginal band, either adding the band with the highest score which is not already part of the subset image, or removing the band already in the subset image which has the lowest score. This process is called progressive band selection (PBS). Since PBS is designed to tailor the number of bands progressively, it is not necessary for users to know the exact number of bands to be selected. PBS will search for an appropriate number of bands to retain to meet an application-specific task. Three different strategies to implement PBS are described in the following sections.

3.1 Forward progressive band selection (FPBS)

The idea of forward progressive band selection (FPBS) is to start with any number of bands, say a , and then gradually add more bands. Bands are added in the order indicated by their priority scores to increase and improve the data quality, until a stopping rule determined by a specific application is met. In this case, the final band number is denoted by a^* . The only issue arising in FPBS is to determine the value of a , the initial band count. Fortunately, it has been shown in [6] that virtual dimensionality (VD), a new concept recently developed in [7-8], provides a reasonable lower bound on the number of bands to be selected. In this case, the value of a can be set to VD.

3.2 Backward progressive band selection (BPBS)

Compared to FPBS, backward progressive band selection (BPBS) works in a completely opposite manner. It begins with a high number of bands, say b , and then gradually removes bands with the lowest priority scores. Bands are removed until overall image quality has deteriorated to some threshold level. The process is then terminated, and the final band count is denoted by b^* . As in FPBS, VD can be also used to establish the value of b as two times VD. In general, after FPBS and BPBS, the desired number of bands, p should be between a^* and b^* , i.e., $a^* \leq p \leq b^*$.

3.3 Binary bisection band selection (BBBS)

FPBS and BPBS search iteratively for the right number of bands to retain. By contrast, binary bisection band selection (BBBS) performs a binary search. The BBBS begins with an output

image containing the highest-ranking 50% of bands from the input image (the midpoint between an empty image and the full dataset). If the application performance is above the threshold, the lowest-ranking half of these bands is removed (leaving the highest ranking 25%). However, if performance was inadequate, half of the bands that were excluded from the first dataset are added, such that the output image consists of the highest-ranking 75% of input bands. Performance is measured again, and if necessary the number of bands is adjusted by a factor of 1/8 (12.5%). The process continues until performance is within an acceptable range. BBBS can also be constrained by multiples of VD rather than zero and the number of bands in the full image. Figure 1 illustrates the three search strategies.

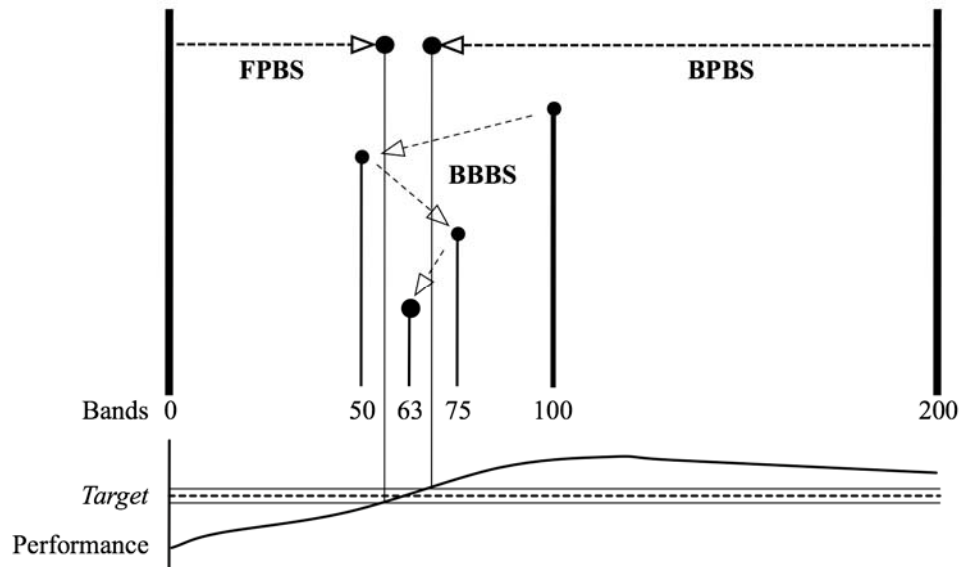


Fig. 1. FPBS and BPBS search iteratively for the correct number of bands, while BBBS employs a binary search.

4 EXPERIMENT

4.1 Image to be studied

The image to be used in our experiments, shown in Fig. 2, is a small subset of an image taken by the Hyperion instrument onboard NASA's Earth Observing One (EO-1) spacecraft.

The image depicts a suburban and mountainous forest area near Tuscon, Arizona on June 17, 2003 at 10:10:15am local time [10]. The original image has 5585 rows of pixel vectors, however for this experiment we use a subset with only 512 rows, centered over the suburban area. Both images have 256 spatial columns, and each pixel vector covers a 30m x 30m area. There are 242 spectral bands in the image, covering wavelengths between 400 and 2500 nm. Atmospheric correction of the image is not necessary since all the training data needed for the classification algorithms are taken directly from the image itself. The same atmospheric effects affect all training and test pixels.

Figure 3 shows the major features in the dataset. A suburb covers the center of the image; a grid of paved streets encircles residential areas and has shopping malls and parking lots at its intersections. The residential areas consist of houses with white reflective roofs and grassy lawns along narrow side streets.

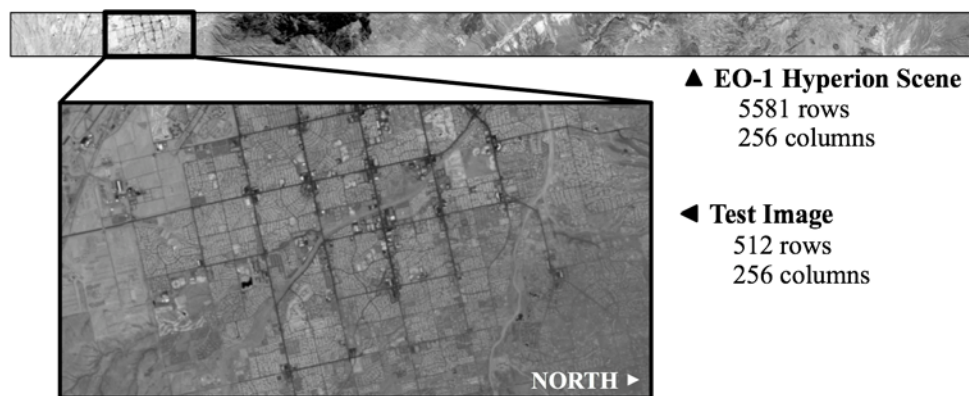


Fig. 2. The test image is a small subset of a scene generated by EO-1 Hyperion. Both are rotated 90 degrees clockwise for display.

The largest expanse of concrete is the runway at the airfield, south of the suburbs, and the largest expanses of grass are found at golf courses dotted throughout the area (the largest is just north of the airfield). A forest fire rages in the mountains to the north, and while the fire itself is not included in the test image, the upper-right corner is shrouded by smoke from the blaze. The native soil is a mix of salt and sand. Some of it is loose, and some has been packed to serve as a large parking area ("boneyard") for airplanes. More packed sand and salt can be found in the dry river beds snaking west to east through town.

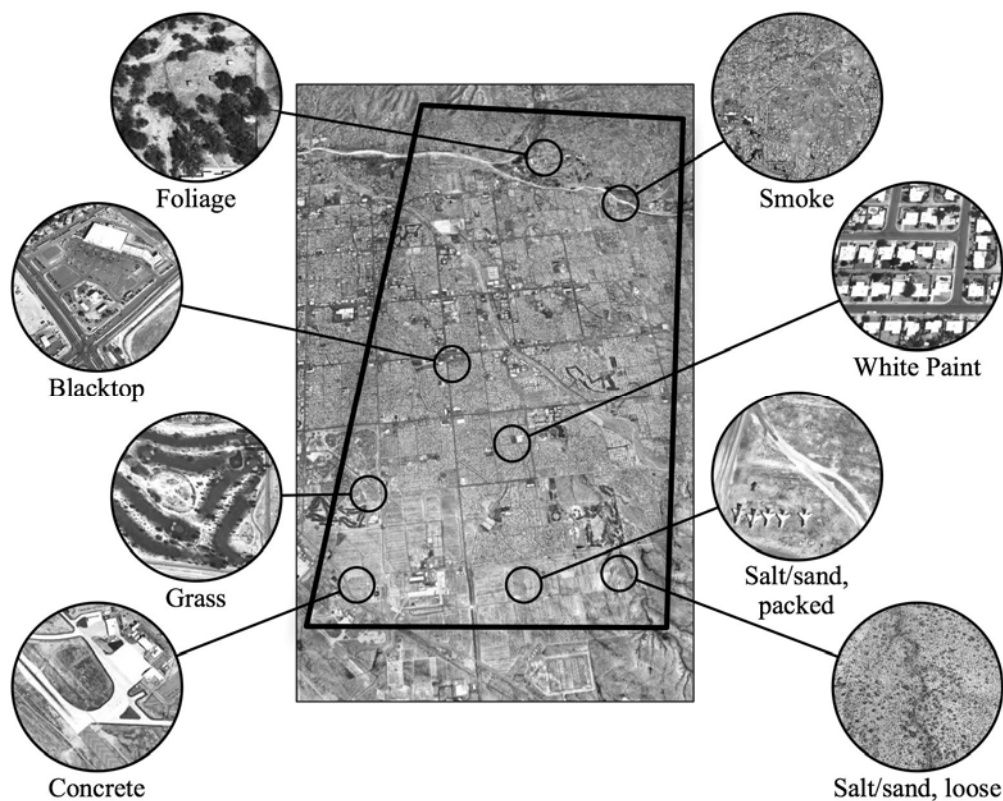


Fig. 3. Major materials and features in the test image.

4.2 Experiment

Three band selection techniques, FPBS, BPBS, and BBBS, coupled with various scoring criteria were used to generate reduced images for our experiments. Data variance, skewness, and information divergence were selected as band prioritization criteria to represent second-order statistics, high-order statistics, and infinite-order statistics respectively. Added to these are the control criteria: UBS, RBS, and sensor-based band selection. Since land cover/use is of major interest, our experiment focuses on a general approach to land cover classification: linear spectral unmixing. It assumes a finite number of basic signature constituents, $\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_p$ are present in the data, using them to form a signature matrix $\mathbf{M} = [\mathbf{m}_1 \mathbf{m}_2 \dots \mathbf{m}_p]$. It then unmixes all the data samples \mathbf{r} into abundance fractions, $\alpha_1, \alpha_2, \dots, \alpha_p$ of $\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_p$ via a linear mixing model: $\mathbf{r} = \mathbf{M}\mathbf{a} + \mathbf{n}$, where $\mathbf{a} = (\alpha_1, \alpha_2, \dots, \alpha_p)^T$ is an abundance vector and \mathbf{n} is a model error or a measurement error. There are two physical constraints on the abundance fractions: the abundance sum-to-one constraint (ASC), $\sum_{j=1}^p \alpha_j = 1$ and the abundance non-negativity constraint (ANC), $\alpha_j \geq 0$ for $1 \leq j \leq p$.

Three popular techniques have been used to perform spectral unmixing for classification, depending upon which abundance constraints are imposed. The least squares orthogonal subspace projection (LSOSP) [11] technique is the simplest one; it imposes no abundance constraints on the linear mixing model. LSOSP has been shown to be equivalent to the Gaussian maximum likelihood classifier [12-13]. The abundance non-negativity constrained least squares (NCLS) technique only imposes ANC on the linear mixing model. It has been shown effective in signal detection [14]. The abundance fully constrained least squares (FCLS) technique is the most commonly used spectral unmixing algorithm because it satisfies physical constraints by imposing both ASC and ANC [15]. In the following experiments, FCLS was chosen for this reason to evaluate the effectiveness of our band selection method.

Since FCLS requires knowledge of all signatures $\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_p$ to form a linear mixing model, we used ground truth provided by Google Earth satellite imagery to select $p = 8$ basic signature constituents of the image scene. Shown in Fig. 3, the classes chosen for spectral unmixing are concrete, pavement, packed salt/sand, loose salt/sand, grass, foliage, white paint, and smoke. The VD estimate for this particular scene is $n_{VD} = 22$, calculated by the Harsanyi-Farrand-Chang (HFC) method in [16]. It should be noted that although VD can be used to determine the number of spectrally distinct signatures in the scene [7-8], only 8 signatures can be identified as meaningful materials by visual inspection of the Google Earth imagery. The other 14 signatures are unknown signatures that are of no interest in land cover classification. Nevertheless, these 14 unknown signatures require at least 14 spectral bands to accommodate their presence for separation. Consequently, $n_{VD} = 22$ is the minimum number of bands required for band selection.

Typically, PBS uses FCLS to perform linear spectral unmixing to determine an acceptable number of bands. In order to initialize the spectral unmixing process, n_{VD} was used as an initial estimate of the number bands to be selected. PBS will then process the image cube, creating subset images with between $n_{VD} = 22$ and $2n_{VD} = 44$ bands. FCLS will use these subset images to produce fractional abundance maps of 8 land cover signatures for visual inspection. These spectrally unmixed abundance maps will serve to measure the effectiveness of PBS in land cover classification.

PBS requires a certain amount of processing time. Ideally, the time saved by processing the band-selected image in place of a larger image will more than offset the cost of using PBS. Figure 4 shows the time needed to process the test image using PBS with various band

selection criteria. The skewness and kurtosis criteria are the most computationally expensive, followed closely by information divergence, all of which use expensive mathematical operations to generate priority scores.

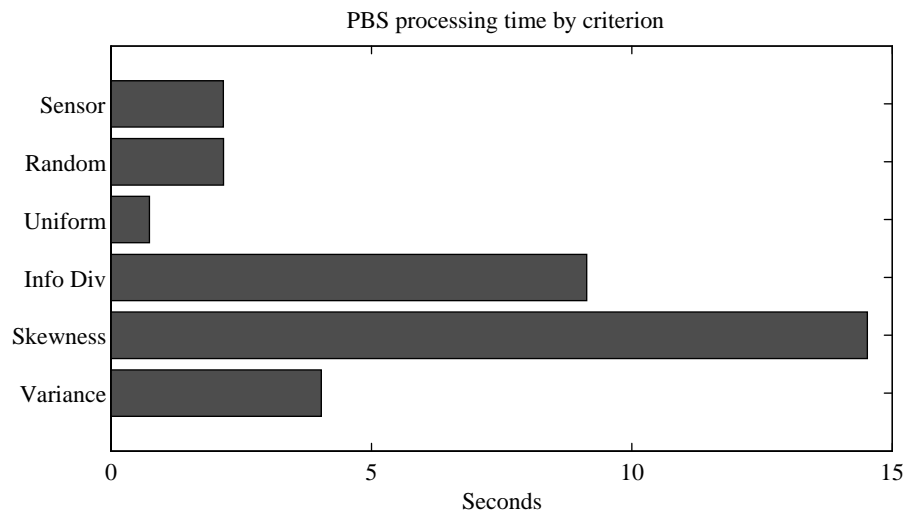


Fig. 4. Computationally complex criteria exact a steep performance penalty.

The images formed by bands chosen by PBS are provided as input to the spectral unmixing algorithm. The algorithm requires complete prior knowledge of signatures to form the linear mixing model, but clearly this is not practical for satellite imagery. Instead, we use a recently developed system for automatic signature generation called unsupervised linear spectral mixture analysis (ULSMA) [18] which uses principal components analysis (PCA) and independent component analysis (ICA) to find background and small-target signatures with no *a priori* knowledge of the image. This algorithm generated the set of 22 signatures based on the largest components of the ICA transform of the reduced image, which are taken to substantially represent all the prominent materials in the image. Each signature is of the same dimensionality (contains as many bands) as a single pixel in the reduced image. In previous experiments, we identified four principal materials in this image: pavement, sand, water and grass. Although ULSMA had no knowledge of the location or signatures of these materials, it managed to generate very similar signatures automatically. In some cases, ULSMA even discriminated subtypes of each material; for example, grass and larger-scale vegetation (foliage).

We will judge PBS performance by its effect on computational resource requirements and the quality of FCLS output.

4.2.1 Computational Resources

Although it takes an entire hyperspectral image as input, FCLS operates on each pixel individually. It attempts to solve a set of linear equations to unmix each pixel into the sum of fractional amounts of each of the eight material signatures also provided as input. The process generates an estimate of the abundance fractions of each material, then refines it until the residual error (the fraction of the pixel's energy that cannot be attributed to any of the eight materials) is reduced below a threshold. If the pixel is noisy or does not contain enough information to clearly separate the contributions of each material, the refining process may require several iterations.

As such, we can use the amount of processing time FCLS requires to unmix an image as an estimate of the quality of that image. Figure 5 shows FCLS processing time as a function of the number of bands in the subset image. While images with more bands contain more data (and would therefore seem to require more time to load the image, and generate larger matrices to manipulate), they also contain more information, which helps FCLS finish in fewer iterations.

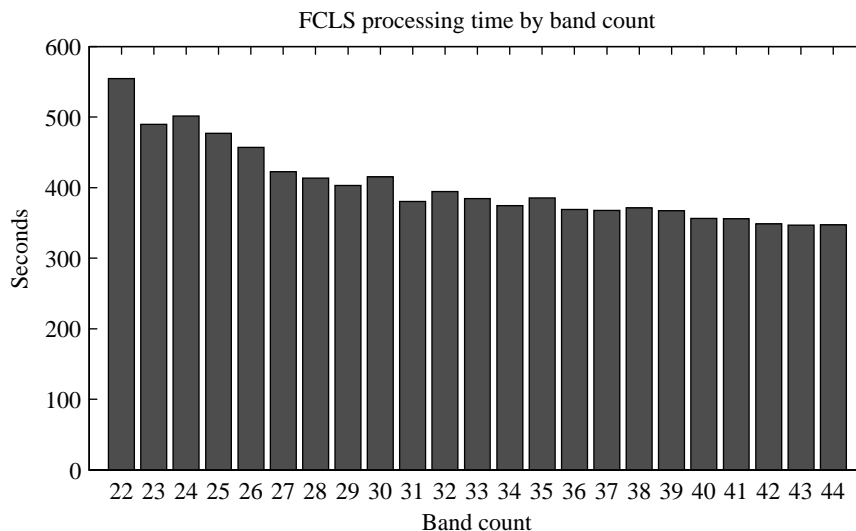


Fig. 5. Additional bands provide more information to accelerate unmixing.

The mark of an effective PBS criterion is to generate subset images that contain more information useful to FCLS, contained in fewer bands. Figure 6 shows that variance is the most effective criterion for this application: for any given number of bands, the variance criterion selects bands that help FCLS finish faster.

On the test system (a MacBook Pro with a 2GHz Intel Core Duo processor and two gigabytes of main memory), the variance criterion selects bands and produces a subset image in less than five seconds. Meanwhile, FCLS processes that subset image 36 seconds faster, on average, than a subset image with the same number of bands chosen randomly.

The processing time statistics converge as the number of bands increases. It appears that when more than 40 bands can be selected, all the criteria (except information divergence) find roughly all the relevant data for the unmixing task. The value of choosing the right criterion is in being able to package this information in a subset image with fewer bands. In this example, the variance criterion seems to include all the needed information in a subset image with only 27 bands.

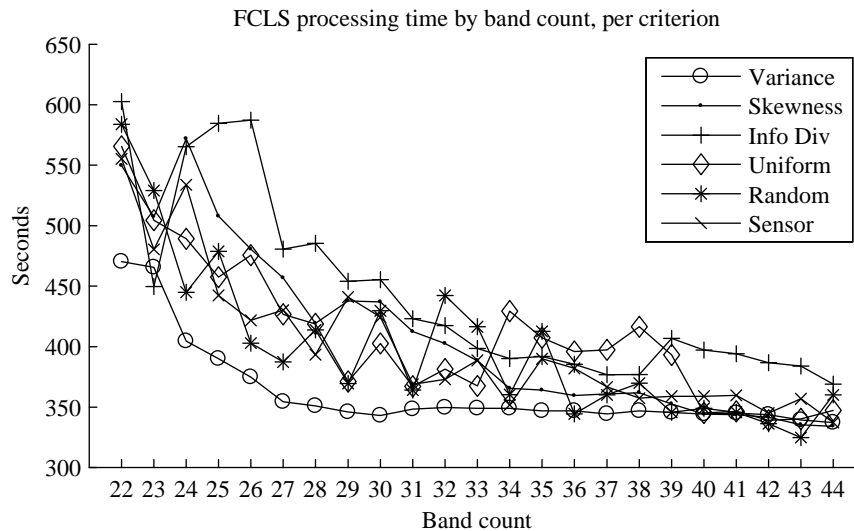


Fig. 6. Variance typically outperforms other criteria at selecting bands for FCLS.

4.2.2 Linear unmixing application (FCLS) algorithm output

While CPU utilization data is a useful numerical proxy for application performance, the best indicator is the actual application output; in this case, a set of eight abundance maps showing the distribution of each material in the image. Each abundance map has the same spatial dimensions and resolution as the original image. Figure 7 shows the abundance maps produced when the full image is given as input.

4.2.2.1 PBS criteria functions

The difference in performance of the three PBS criterion functions (variance, skewness, and information divergence) is clearly illustrated in Fig. 8. The variance criterion scores each band with a low-order statistic, which is a good predictor of the amount of information the band contains about large-scale features. This is helpful in land classification scenarios, where the goal is to identify large expanses of spectrally-similar landscape. FCLS yields output with only 25 bands chosen by the variance criterion (shown in Fig. 9) that is quite similar to that from the full 242-band image. The only major difference is the lack of separation between loose and packed sand/salt, which have very similar spectra.

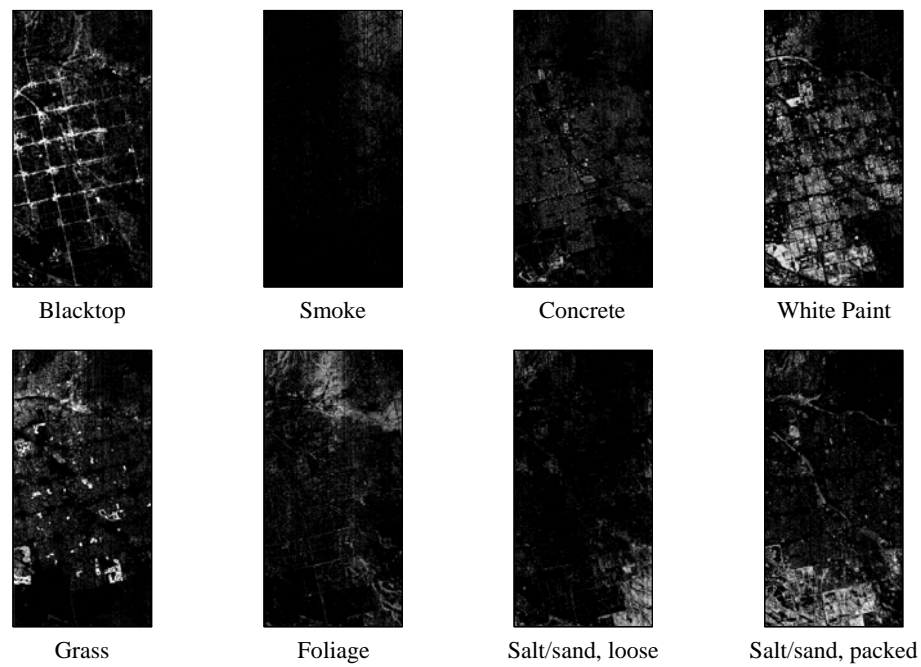


Fig. 7. Abundance maps produced by FCLS when given the full image as input.

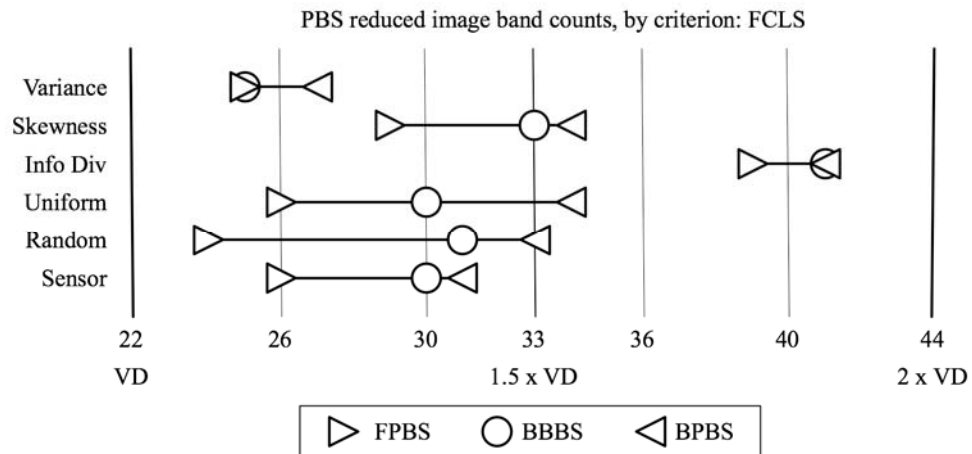


Fig. 8. Variance and information divergence are clear outliers from the other criterias' middling performance.

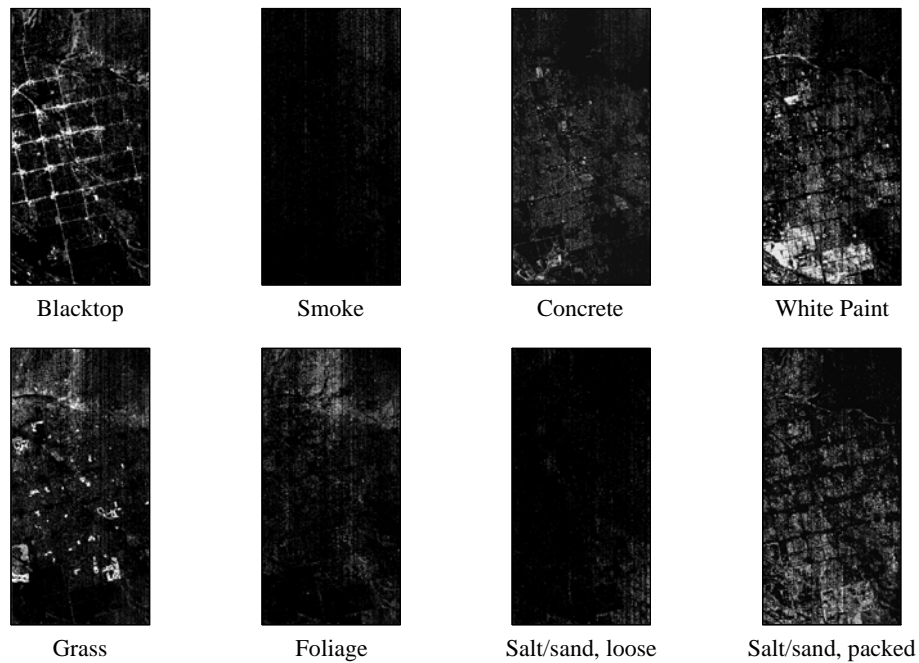


Fig. 9. FCLS output from a 25-band subset chosen with the variance criterion.

Information divergence, by contrast, eschews low-order statistics in favor of infinite-order statistics, which are useful in identifying small single- or sub-pixel targets. The quality of the abundance maps produced from its output varies widely with the number of bands in the subset image. Figure 10 shows the output of FCLS given a 38-band subset image. Note how the grid of paved streets is visible in the abundance maps of blacktop, grass, and foliage.

Skewedness provides generally middle-of-the-road performance. This high-order statistic does not identify large image features as well as variance, however BPBS found FCLS performance with a 34-band subset assembled using skewness to be acceptable. The salt/sand abundance maps in Fig. 11 are distorted (faded) by a small number of outlier pixels.

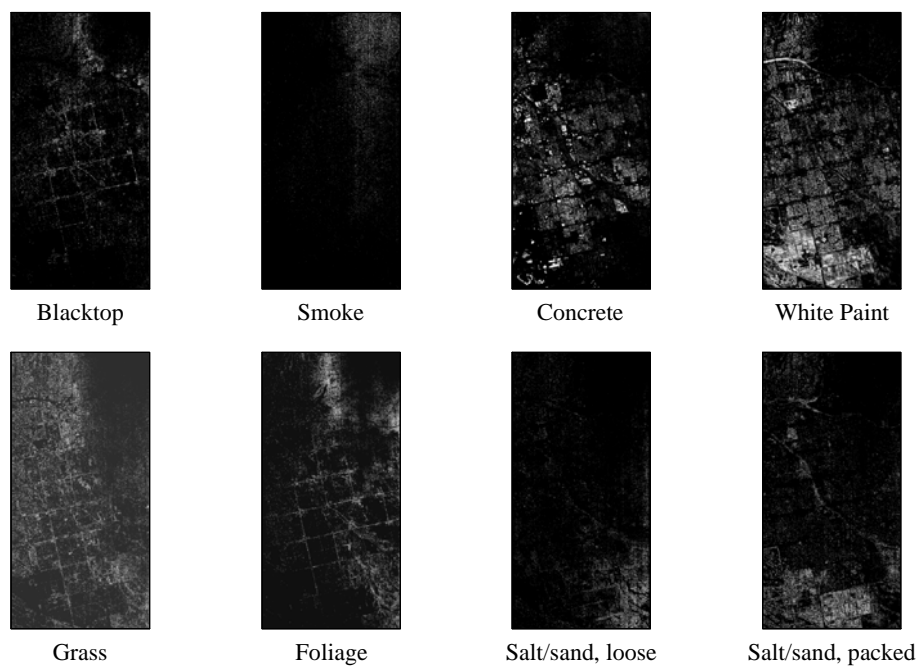


Fig. 10. Information divergence fails to provide enough information in a 38-band subset to differentiate blacktop from grass and foliage.

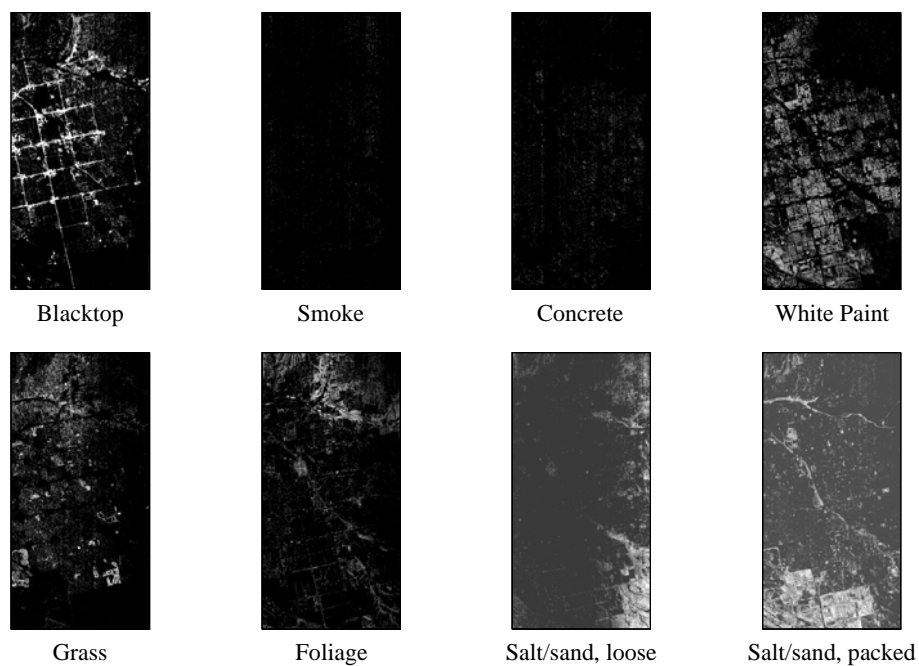


Fig. 11. Skewedness performance lies between the other PBS criteria; shown here is the output of FCLS given 34 bands.

4.2.2.2 Control criteria functions

While the control criteria possess none of the smarts of the other PBS criterion functions, they still yield reasonable performance. The range of band counts chosen by FPBS, BPBS, and BBBS are wider for these criteria because their performance varies greatly from one band count to the next. RBS and sensor band selection choose bands largely at random, therefore there is little overlap between the sets of bands selected for, say, the 23 and 24-band subset images. While FCLS produced surprisingly good output from the random selection of 24 bands (shown in Fig. 12), it could do nothing with the 23-band subset, producing a set of null abundance maps.

Sensor-based band selection selects bands randomly as well, but it is programmed to always avoid the worst bands. As expected, it performs roughly the same as RBS, although BPBS was able to iterate to a lower number of bands. This is because RBS has a higher chance of selecting an especially poor band from the original image, which would reduce FCLS output below an acceptable level and cause BPBS to stop evaluating output from subsets with fewer bands.

UBS, the criteria that selects a uniformly-spaced set of bands across the frequency spectrum, performs similarly because it functions like a deterministic pseudo-random band selector. Like RBS and sensor-based band selection, the set of bands present in two subset images with "neighboring" band counts will be almost completely different, unless a certain band has an identification number with multiple divisors. (For example, if UBS were to generate subset images from every other image band and every third band, both subsets would contain band number 6.) UBS has no capacity to avoid the bands that cause poor FCLS performance, so BPBS and FPBS stop at high and low band numbers, respectively.

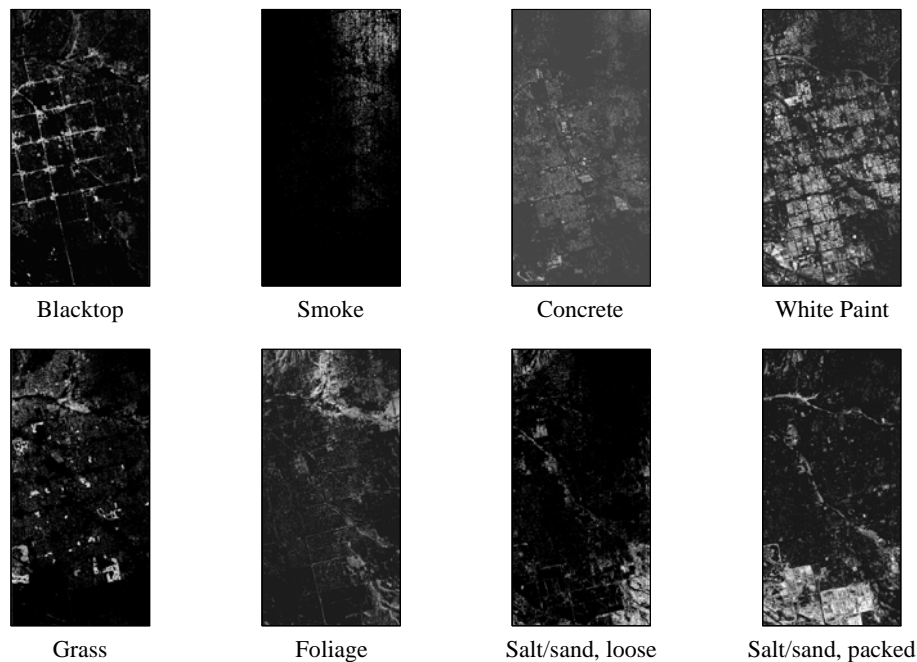


Fig. 12. FCLS output from 24 bands selected randomly; performance was deemed adequate by FPBS.

5 CONCLUSION

Satellite communication generally requires fast downlink transmission to the Earth. Hyperspectral data, which represents enormous data volumes, may require a significant amount of downlink time. This paper presents an effective band selection method which selects spectral bands according to their band prioritization. These prioritized bands can be further used for progressive data transmission during satellite communication. To the authors' best knowledge, the idea of compressing data while transmitting it progressively has never been explored in the literature. In our method, all the spectral bands of a hyperspectral image are first ranked by their priority scores, which can be calculated and obtained by a specific application-designed criterion. This is then followed by progressive band selection (PBS), carried out in either forward or backward manner, referred to as forward PBS (FPBS) and backward PBS (BPBS), according to their assigned band prioritization. In order to determine the initial number of bands FPBS or BPBS should select, virtual dimensionality (VD) is used to provide a reasonable estimate on lower and upper bounds on the number of bands needed to be selected. Finally, a data set collected by Hyperion was used for land cover/use experiments to demonstrate the utility of PBS.

References

- [1] P. W. Mausel, W. J. Kramber, and J. K. Lee, "Optimum band selection for supervised classification of multispectral data," *Photogramm. Eng. Remote Sens.* **56**, 55-60 (1990).
- [2] C. Conese and F. Maselli, "Selection of optimum bands from TM scenes through mutual information analysis," *ISPRS J. Photogramm. Remote Sens.* **48**, 2-11 (1993) [doi: 10.1016/0924-2716(93)90059-V].
- [3] S. D. Stearns, B. E. Wilson, and J. R. Peterson, "Dimensionality reduction by optimal band selection for pixel classification of hyperspectral imagery," *Proc. SPIE* **2028**, 118-127 (1993) [doi: 10.1117/12.158622].
- [4] C.-I Chang, Q. Du, T. S. Sun, and M. L. G. Althouse, "A joint band prioritization and band decorrelation approach to band selection for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.* **37**, 2631-2641 (1999) [doi: 10.1109/36.803411].
- [5] R. Huang and M. He, "Band selection based feature weighting for classification of hyperspectral data," *IEEE Geosci. Remote Sens. Lett.* **2**, 156-159 (2005) [doi: 10.1109/LGRS.2005.844658].
- [6] C.-I Chang and S. Wang, "Constrained band selection for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.* **44**, 1575-1585 (2006) [doi: 10.1109/TGRS.2006.864389].
- [7] C.-I Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*, Kluwer Academic/Plenum Publishers, New York (2003).
- [8] C.-I Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.* **42**, 608-619 (2004) [doi: 10.1109/TGRS.2003.819189].
- [9] T. Cover and J. Thomas, *Elements of Information Theory*, Wiley, New York (1991) [doi: 10.1002/0471200611].
- [10] *EO-1 Hyperion image EO1H0360372003184110KW*, Data available from the U.S. Geological Survey.
- [11] C.-I Chang, X. Zhao, M. L. G. Althouse and J.-J. Pan, "Least squares subspace projection approach to mixed pixel classification in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.* **36**, 898-912 (1998) [doi: 10.1109/36.673681].
- [12] J. J. Settle, "On the relationship between spectral unmixing and subspace projection," *IEEE Trans. Geosci. Remote Sens.* **34**, 1045-1046 (1996) [doi: 10.1109/36.508422].

- [13] C.-I Chang, "Further results on relationship between spectral unmixing and subspace projection," *IEEE Trans. Geosci. Remote Sens.* **36**, 1030-1032 (1998) [doi: 10.1109/36.673697].
- [14] C.-I Chang and D. Heinz, "Constrained subpixel detection for remotely sensed images," *IEEE Trans. Geosci. Remote Sens.* **38**, 1144-1159 (2000) [doi: 10.1109/36.843007].
- [15] D. Heinz and C.-I Chang, "Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.* **39**, 529-545 (2001) [doi: 10.1109/36.911111].
- [16] J. Harsanyi, W. Farrand, and C.-I Chang, "Determining the number and identity of spectral endmembers: An integrated approach using Neyman–Pearson eigenthresholding and iterative constrained RMS error minimization," *Proc. 9th Thematic Conf. Geologic Remote Sens.* (1993).
- [17] C.-C. Chang and C.-J. Lin, "LIBSVM : a library for support vector machines," (2001). Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- [18] X. Jiao, Y. Du, and C.-I Chang, "Unsupervised linear spectral mixture analysis for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, under review (2009).