International Journal of Applied Earth Observation and Geoinformation 26 (2014) 387–398

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

A spatial–spectral approach for deriving high signal quality eigenvectors for remote sensing image transformations

Derek Rogge[a,*], Martin Bachmann[a], Benoit Rivard[b], Allan Aasbjerg Nielsen[c], Jilu Feng[b]

[a] German Remote Sensing Data Centre, DLR, Munchnerstr. 20, D-82234, Germany
[b] Department of Earth and Atmospheric Sciences, University of Alberta, Edmonton T6G 2E3, Canada
[c] Technical University of Denmark, National Space Institute, DK-2800 Kgs. Lyngby, Denmark

ARTICLE INFO

Article history:
Received 3 July 2013
Accepted 13 September 2013

Keywords:
Hyperspectral imaging
Spatial and spectral processing
Eigenvector transformations

ABSTRACT

Spectral decorrelation (transformations) methods have long been used in remote sensing. Transformation of the image data onto eigenvectors that comprise physically meaningful spectral properties (signal) can be used to reduce the dimensionality of hyperspectral images as the number of spectrally distinct signal sources composing a given hyperspectral scene is generally much less than the number of spectral bands. Determining eigenvectors dominated by signal variance as opposed to noise is a difficult task. Problems also arise in using these transformations on large images, multiple flight-line surveys, or temporal data sets as computational burden becomes significant. In this paper we present a spatial–spectral approach to deriving high signal quality eigenvectors for image transformations which possess an inherently ability to reduce the effects of noise. The approach applies a spatial and spectral subsampling to the data, which is accomplished by deriving a limited set of eigenvectors for spatially contiguous subsets. These subset eigenvectors are compiled together to form a new noise reduced data set, which is subsequently used to derive a set of global orthogonal eigenvectors. Data from two hyperspectral surveys are used to demonstrate that the approach can significantly speed up eigenvector derivation, successfully be applied to multiple flight-line surveys or multi-temporal data sets, derive a representative eigenvector set for the full image data set, and lastly, improve the separation of those eigenvectors representing signal as opposed to noise.

1. Introduction

Most airborne hyperspectral systems today typically comprise 100's of contiguous bands ranging from 0.4 to 2.5 μm and produce significant quantities of data, commonly 100's of megabits in size. With the advent of satellite hyperspectral sensors (e.g. Hyperion (Pearlman et al., 2001), EnMap (Stuffler et al., 2007), PRISMA and HyspIRI (Buckingham and Staenz, 2008), HISUI (Kawashima et al., 2010)), the amount of data available for analysis increases substantially. Processing this type of data, from acquisition through to map generation, requires numerous computationally intensive steps (e.g. radiometric correction, atmospheric correction, geometric correction, classification, validation). Thus, there is a real need to develop cost-effective algorithm implementations to satisfy time-critical remote sensing applications (Plaza et al., 2009) that also produce quality, physically accurate results that the end user can use in specific applications (Rogge et al., 2012). Improvements in computing power, memory, parallel processing and advancing existing algorithms for more efficient processing of large data sets can help to achieve these needs. However, any improvements with respect to reducing computational load must be done such that there is little, or preferably no, loss in quality.

Spectral decorrelation (transformations) methods have long been used in remote sensing, such as principal component analysis (PCA) (Ready and Wintz, 1973; Singh and Harrison, 1985), maximum noise fraction (MNF) (Green et al., 1988) and singular value decomposition (SVD) (Danaher and O'Mongain, 1992), which are all based on strong mathematical foundations. Uses include pre- and post-processing steps, such as data compression (Du et al., 2009), correcting for spectral smile effects (Dadon et al., 2010), eigenvector based approaches to virtual dimensionality (VD) estimation (Chang and Du, 2004), or endmember extraction (Boardman et al., 1995). Transformations are commonly used to reduce the dimensionality of hyperspectral images as the number of spectrally distinct signal sources composing a given hyperspectral scene is generally much less than the number of spectral bands. Transformation of the image data onto eigenvectors can give a user quick results that show the spatial distribution of the majority of spectrally distinct

* Corresponding author. Tel.: +49 8153283364; fax: +49 8153281458.
E-mail addresses: derek.rogge@dlr.de (D. Rogge), martin.bachmann@dlr.de (M. Bachmann), benoit.rivard@ualberta.ca (B. Rivard), aa@space.dtu.dk (A.A. Nielsen), jfeng@ualberta.ca (J. Feng).

0303-2434/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jag.2013.09.007

materials in the scene. Image data derived from atmospherically corrected multi- or hyperspectral sensor systems have variable quality with respect to the ratio of signal and noise (SNR). If SNR is defined simply as the mean target signal of a spatially homogeneous material divided by the signal standard deviation, then the signal portion from an image with good SNR will be controlled primarily by the spectral characteristics of the target reflectance rather than from other factors (e.g. internal sensor characteristics). In this case, eigenvectors with highest eigenvalues derived from an image that have the highest eigenvalues should represent the highest variance in the data and subsequently the physically meaningful spectral properties of the scene. Generally, a limited number of eigenvectors are expected to contain most of the signal information, whereas the remaining eigenvectors comprise little or no signal and variance is dominated by noise. A problem with remote sensing imagery is that some image materials have weak signals and contribute little energy to the eigenvalues. Determining eigenvectors that contain important physically meaningful spectral properties (signal), as opposed to just noise, is a difficult task (Chang and Du, 2004). Thus, poor separation of the two can result in loss of significant information.

Problems arise in using PCA, SVD and MNF transformations on large images, multiple flight-line surveys, or temporal data sets as computational burden becomes significant. Spatial subsampling (e.g. every *n*th pixel) can be used to reduce data size to derive representative eigenvectors. However, this approach can result in the potential loss of important information related to specific surface materials. Specifically, a subsample needs to be relatively small in order to significantly reduce computational load, but must also be sufficiently large so as to adequately capture the statistical variation in the data set (Du and Fowler, 2008). Image transformations are also scene dependent, meaning that the derived eigenvectors are based on the input data. Thus, producing mosaics of spatially adjacent or temporal scenes that are transformed separately is not possible, meaning eigenvectors derived from one scene may not be appropriate for another. Scene to scene correlation of transformed data can be accomplished by using the same eigenvectors to transform each scene with the assumption that those eigenvectors represent all of the spectrally distinct signal sources composing the different scenes. As image size increases, or multiple flight-lines (spatially or temporally) comprise a given project, the ability to obtain a representative set of eigenvectors for the complete data set also becomes a difficult task.

A solution to deriving a global set of eigenvectors for large data sets is provided by eigenspace merging. For a given data set an eigenspace can be defined that includes m p-dimensional observations, the eigenvectors and eigenvalues, and the mean vector of the population. Eigenspace merging allows for adding of new information to a given sample set, for example as a single observation or with the addition of a whole new data set (Hall et al., 2000; Franco et al., 2002). In many cases eigenspace merging was developed for applications, such as face recognition, where the input data comes from potentially different sources (e.g. sensors). The idea of eigenspace merging is also important when using parallel processing techniques to reduce computational load and increase processing speed (e.g. El-Ghazawi et al., 2002; Yang et al., 2008). Yang et al. (2008) presented an approach specifically designed for hyperspectral imagery. Their aim was to increase processing speed using parallel processing by splitting an image into spatial subsets, solving eigenspaces independently and finally merging the spaces using the approach given by Franco et al. (2002). In their approach partial eigenvector decomposition was also done, such that each subset region was represented by a limited set of eigenvectors. For merging, the method of Franco et al. (2002) considers the mean values of each eigenspace, which is important in classification problems where the mean represents the center of a cluster of observations in a given class (Hall et al., 2000).

Eigenspace merging, thus, represents a practical approach to deal with large, multiple flight-line or temporal hyperspectral data sets. The current study adds to the body of published work by demonstrating that merging of spatially contiguous subset eigenspaces can be considered a form of spatial and spectral subsampling that is inherently a noise reducing eigenvector decomposition, which can be done without the need to estimate image noise as in the case with the well known MNF transform (Green et al., 1988). In addition, it can also be demonstrated that the approach does not require the mean eigenspace values, which allows us to develop a simplified merging approach as opposed to the more mathematically complex version given in Franco et al. (2002). This simplified approach also addresses the problems outlined above when working with large spatial and/or temporal image data sets. This is accomplished by analyzing the data (single image or images) in contiguous equal sized spatial subsets, where within each subset region a limited number of eigenvectors can be derived that comprise most of the signal information that best represent the local subset spectral diversity. These subset eigenvectors can be compiled together to form a new vector data set, which subsequently can be used to derive a global set of orthogonal eigenvectors that represent the original full data set.

The objective of this paper is to demonstrate benefits of a spatial–spectral approach to eigenvector derivation where this method can be applied to: (1) significantly speed up processing without parallel processing by using spatial–spectral subsampling; (2) multiple flight-line surveys or multi-temporal data sets; (3) derive eigenvectors that are a good representation of the spectral diversity for the full image data set; and, lastly (4) show that the method possesses an inherent ability to reduce the effects of noise and allows for an improved separation of those eigenvectors that include physically meaningful signal as opposed to those dominated by noise. Two hyperspectral scenes, the first being the well known AVIRIS Cuprite scene, and the second being 7 HyMap hyperspectral flight-lines from Sokolov, Czech Republic, will be used to demonstrate the approach. The remainder of this paper begins with the rationale of the approach followed by a description of the details of the spatial–spectral approach, a description of the data sets used, the evaluation methods, results and lastly, a discussion and conclusions.

2. Spatial–spectral eigenvector derivation

2.1. Relation to MNF

PCA is an eigenvector transformation method that is commonly used to determine the underlying statistical dimensionality of an image (Richards, 1993). This is accomplished by choosing uncorrelated linear combinations of the variables (spectral bands), such that the first few combinations comprise the majority of variance (spectral) with each subsequent combination encompassing a smaller variance. What is key to PCA is that the transformation maximizes variance of image data where maximum variance is subject to the constraint that it is uncorrelated with lower order principal components (PC's). This analysis does not take into account the spatial nature of an image.

Although PCA is effective at decorrelating image data, where compression of image information onto the low order PC's shows a steadily decreasing signal-to-noise ratio (SNR), Green et al. (1988) noted that in some data PCA does not always behave in this manner. This led to the development of MNF, which differs in that it can be used to maximize SNR rather than variance. For example, given an

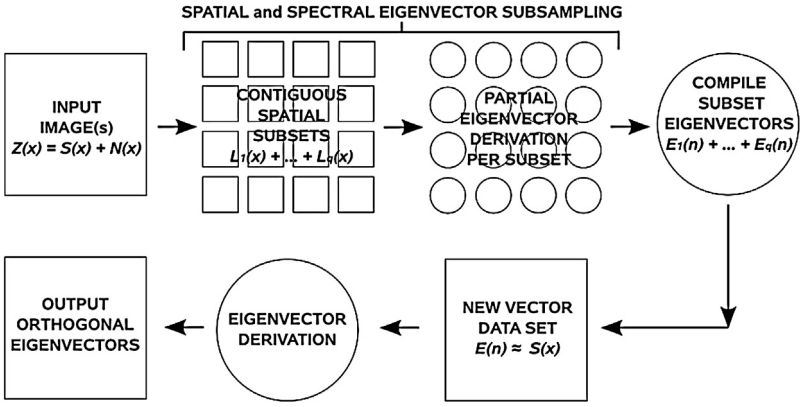

Fig. 1. Flow chart of SSEVD, clockwise from top left.

image $Z(x)$ comprising p bands, where x gives the coordinates of pixel vectors, we can assume that:

$$Z(x) = S(x) + N(x) \tag{1}$$

where $S(x)$ and $N(x)$ are the uncorrelated signal and noise components of $Z(x)$. This gives:

$$cov\{Z(x)\} = \sum = \sum_S + \sum_N \tag{2}$$

where Σ_s and Σ_N are the covariance matrices of $S(x)$ and $N(x)$. MNF requires knowledge of both Σ and Σ_N, where Σ_N is commonly estimated from the data. If this can be determined, MNF can minimize the noise fraction relative to the total variance. Thus, the key to MNF is estimating the noise fraction (Σ_N), which was designed to use some type of spatial filtering of each input band (Green et al., 1988), such as minimum/maximum autocorrelation (MAF) (Switzer and Green, 1984) that assumes signal at any point is strongly correlated with signal at neighboring pixels. In the approach presented in this paper, we also make use of spatial information, in the form of contiguous spatial subsets, for the derivation of eigenvectors. However, the key difference with MNF is that the given approach is designed to maximize $S(x)$ only, with no estimation of $\boldsymbol{\Sigma_N}$ required. Details of how this is possible using a subset approach are explained below and further demonstrated in Section 5.

2.2. Algorithm details

Fig. 1 shows the main steps of the spatial–spectral eigenvector derivation approach (from here forth referred to as SSEVD). The first step is to take an image $Z(x)$, as described above, and divide it into q equal sized contiguous image subsets.

$$Z(x) = L_1(x) + L_2(x) + \ldots + L_q(x) \tag{3}$$

where each $L_i(x)$ will also comprise uncorrelated signal and noise components,

$$L_i(x) = S_i(x) + N_i(x) \tag{4}$$

For each $L_i(x)$, we can derive a partial eigenvector decomposition, such that the subset region is represented by a limited set of eigenvectors $E_i(v)$ that explain the majority of variance. Here v gives the order of the eigenvectors for the given subset i. If we make the assumption that the majority of variance is controlled by signal, which is the case for remote sensing images with good SNR, then we can also make the assumption that:

$$E_i(v) \approx S_i(x) \tag{5}$$

If we compile all $E_i(v)$ for each subset i we then end up with a new data set of size $p \times n$ (n is the total number of eigenvectors retained from all subsets). Where,

$$E(n) \approx S(x) \tag{6}$$

As $E(n)$ is in itself just a set of vectors, we can then used it to derive a final set of orthogonal eigenvectors that does not require an estimation of the noise component $N(x)$.

Deriving a limited set of eigenvectors from spatially contiguous subsets can essentially be considered a spatial and spectral subsampling of an image. For example, if we were to calculate the average spectra for each subset we can significantly reduce our total sample set (spatial subsampling). However, averaging the spectra for a given subset we are limited to first order statistics, in the form of the average spectra. This is useful in homogeneous regions, but will not give a good spectral representation of the subset in heterogeneous regions. To deal with this limitation the spatial subsampling concept is expanded to use eigenvectors, where the advantage is that we now include second order statistics that are more representative of the total variance in each subset. This concept is specific to image data, which includes a high degree of spatial autocorrelation that is reflected in the data variance. If we choose to use relatively small subset sizes we can expect the spectral diversity to be small. Thus, the variance should be controlled by a limited number of materials and, in turn, can be represented by a limited number of eigenvectors. In addition, as $E(n)$ is considered a noise reduced spatial and spectral subsampling of the original image $Z(x)$, there is no need to consider the subset eigenspace mean values, as is the case with the method presented by Franco et al. (2002).

In SSEVD, the method to determine the number of eigenvectors to retain per subset is an important consideration. Many different approaches have been suggested, some objective (data driven) and others subjective, with a summary of some of the spectral based approaches given in Umberto et al. (2009). It is well known that with remote sensing images with good SNR, the major image components can be represented by simply retaining the top few eigenvectors based on a percent of spectral variance. Generally this assumption can be made for all images regardless of image materials and distribution, image size or the number of bands. However, this approach will commonly result in missing important spectral information that is not significant at the global scale, but is distinct locally. This constraint is particularly problematic for large images, but it is less subject to failure as the size of the image decreases and the expected spectral diversity shrinks. Thus, working with small subsets we can assume here that setting an eigenvector cutoff threshold ($t_{EV}{}^{\%}$) based on percent of spectral variance the resulting eigenvectors should successfully explain the majority of variance, and thus, the full spectral diversity of the given subset.

SVD, like PCA, can derive eigenvectors from image data which are very efficient in explaining most of the spectral variability of a given scene (e.g. Thai et al., 1999). The mathematical basis of SVD and its relationship to PCA is well established in the literature (e.g. Golub and Van Loan, 1996) and not discussed further in this paper. For our SSEVD approach we have chosen to work with SVD to derive eigenvectors, as it is known to be numerically stable. In addition, it can provide a solution even in the case of highly spectrally correlated data where a "singularity" problem can be encountered and methods such as PCA may fail (Press et al., 1992; Jolliffe, 1986). For the approach given in this paper a similar implementation of PCA could also be developed. It is also noted here that MNF could be used to derive the subset eigenvectors. However, this is considered redundant as MNF would still need to estimate the noise fraction based on local subset statistics.

For the SSEVD approach there are only two input parameters required, subset size and the SVD cutoff threshold ($t_{SVD}^{\%}$). The choice and rational of the input parameters are given in Section 4, which will be preceded by a description of the image data sets used.

3. Data sets

To demonstrate the new approach and assess the results we make use of two data sets that differ in 3 important characteristics: (1) the two data sets were acquired from two well-known sensor systems with different characteristics (e.g. SNR); (2) the physically meaningful spectral properties of each data set are different; and, (3) the second data set comprises multiple flight lines, which allows us to test the approach to reach the 2nd objective of using multiple flight-line surveys or multi-temporal data sets.

3.1. AVIRIS Cuprite

The AVIRIS Cuprite imagery (http://aviris.jpl.nasa.gov/) was acquired on June 19th, 1997, and has 224 channels covering the 0.37–2.51 μm spectral range with an average band Full Width Half Maximum of ~10 nm and a Ground Instantaneous Field of View of ~20 m. Of the 224 atmospherically corrected channels, we used 167 after removal of channels associated with H_2O absorption features near 1.4 and 1.9 μm. For this evaluation a 500 × 500 pixel region that includes both eastern and western hydrothermal alteration zones was used (Fig. 2). The Cuprite area is arid with limited vegetation cover, and has excellent mineral exposures comprising alteration zones characterized by the occurrence of key indicator minerals (Clark et al., 2003).

3.2. HyMap Sokolov

The HyMap airborne data of Sokolov, Czech Republic, was acquired by HyVista and the Deutsches Zentrum für Luft-und Raumfahrt (DLR) (German Aerospace Center) on August 21, 2010. The Sokolov data has 125 bands covering the 0.45–2.48 μm spectral range with an average band Full Width Half Maximum of 15 nm and a Ground Instantaneous Field of View of ~4 m. The data has been converted from radiance to reflectance and atmospherically corrected using ATCOR 4 (Richter, 2010). The data is geometrically corrected, but no cross-track illumination or BRDF correction was applied. Of the 125 atmospherically corrected bands, 14 of those occurring near the main 1.4 and 1.9 μm atmospheric water absorption features and noisy bands longer than 2.45 μm were removed from the data leaving 110 bands. A total of seven lines were flown, each having 512 samples and ranging from 2498 to 5301 lines for a total of 15,962,624 pixels. A geometrically corrected image of the data is shown in Fig. 3. Sokolov is located near the Czech Republic's western border with Germany. The area comprises a mixture of forest, lakes, rivers, agricultural land, open pit coal mines and remediation areas, and numerous small urban centers.

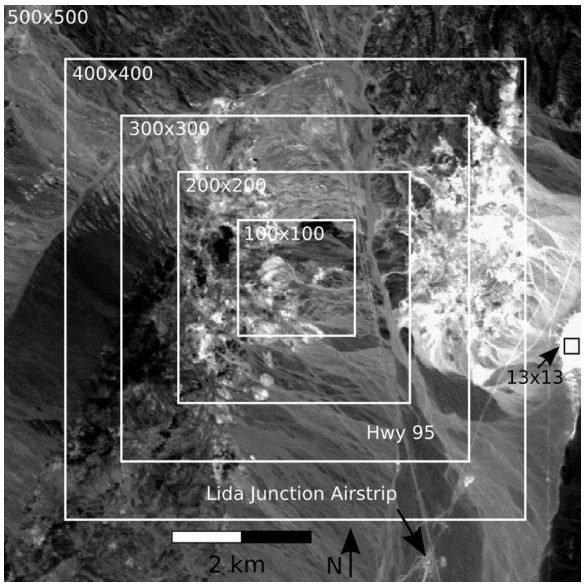

Fig. 2. AVIRIS Cuprite hyperspectral data. White squares show the 500 × 500, 400 × 400, 300 × 300, 200 × 200 and 100 × 100 regions to test processing speed of FI and SSEVD approaches with respect to image size. The small black square (approximately center right edge) is the subset size (13 × 13) used in the SSEVD approach.

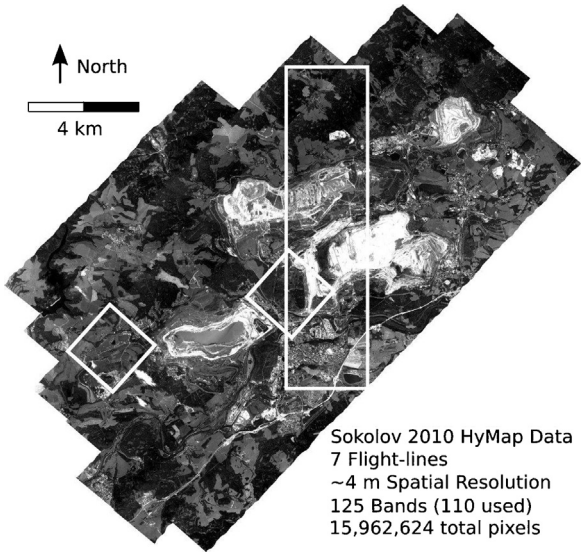

Fig. 3. Sokolov, Czech Republic, HyMap hyperspectral data. Small white boxes indicate 512 × 512 pixel regions used in the spatial correlation evaluation (see Section 5.2). Large vertical rectangle region is used to show examples of transformed data for 6 of the 7 flight-lines (see Section 5.2).

4. Evaluation methods

For the evaluation of the SSEVD approach we divide it into two main parts. The first will use the AVIRIS Cuprite data set to compare eigenvector derivation applied to a full image (from here forth referred to as FI) with that of the SSEVD approach. The second evaluation applies SSEVD to the multiple flight-line Sokolov HyMap hyperspectral scene. All processes were run with code generated within the IDL7.0 – ENVI 4.5 environment with no parallel processing implemented. The SVD process used in this

paper is calculated using the SVDC routine available in IDL, which is based on the routine svdcmp described in Press et al. (1992).

4.1. Cuprite data evaluation

Three criteria are assessed with the Cuprite data: (1) processing speed of FI and SSEVD; (2) comparison of the resulting eigenvectors; and, (3) assessment of the level of signal versus noise in the resulting transformed images with respect to eigenvalue rank using spatial correlation statistics. For processing speed we test FI and SSEVD using 5 different input image sizes (500×500, 400×400, 300×300, 200×200, 100×100) (see Fig. 2) and compare the processing time required for both methods. In this experiment the subset size for SSEVD is set to 13×13 (see Fig. 2). This is based on the minimum subset size to calculate the SVD for the Cuprite data, where the number of samples (m) in the subset must be greater than $p+1$. Using a small subset size is important as we wish to limit the spectral variance within a subset, such that only a few eigenvectors are required. The number of eigenvectors retained per subset will be based on the percent of cumulative variance of the eigenvalues, or eigenvector cutoff threshold ($t_{EV}^{\%}$). For the processing speed test, a $t_{EV}^{\%}$ of 99% will be used.

Next, the resulting eigenvectors from FI and SSEVD are compared for the largest image size (500×500). To assess the impact of using different subset $t_{EV}^{\%}$, three values were tested (95, 98, and 99 percent of spectral variance). The final number of orthogonal eigenvectors for SSEVD, derived from the second SVD run on the compiled eigenvectors, is equal to the number of bands, which is the same as FI. Here a visual and quantitative comparison, and an assessment of spectral variance per eigenvector is made in order to reveal differences in the approaches. Lastly, spatial correlation of the transformed images is assessed, where transformed images that comprise signal should show a high degree of spatial correlation, whereas those that are dominated by noise should have very low (or no) spatial correlation. To quantitatively assess the spatial correlation we use two well known spatial autocorrelation measures, the Geary's and Moran's methods (Griffith, 1987), which are included in the ENVI 4.5 environment. Geary's and Moran's methods are approximately inversely related. However, Moran's is considered a more global indicator, whereas Geary's is more sensitive to differences in small neighborhoods. For both methods spatial correlation is assessed using a 3×3 pixel window. For the Geary's and Moran's indexes high spatial correlation is associated with values near 0 and 1, respectively. For low correlation the values are near 1 and 0, respectively. Negative correlation for Geary's are represented by values greater than 1 and for Moran's values below 0.

4.2. Sokolov data evaluation

One of the key advantages of SSEVD to be demonstrated is its ability to run on multiple large images (spatial or temporal) and produce a single set of representative eigenvectors that can then be use to transform each data set. To demonstrate this we use the 7 flight-line Sokolov hyperspectral data set, where the subset size is set to 11×11 (110 band minimum) and the $t_{EV}^{\%}$ is set to 90, 95, 98% of variance. As the Sokolov data comprises 7 flight-lines, subset eigenvectors from all flight-lines can be simply compiled into a single file just as if they were from a single image. The final set of orthogonal eigenvectors is used to transform each flight-line to assess the spatial similarities and differences using Geary and Moran's spatial autocorrelation. A comparison of SSEVD to FI using all 7 flight-lines was not possible because of computational load and memory constraints required for the FI approach. Thus, to compare spatial correlation on the Sokolov data we run both approaches on two 512×512 test regions (see Fig. 3). Spatial autocorrelation for all 7 flight-lines is only shown for the SSEVD approach.

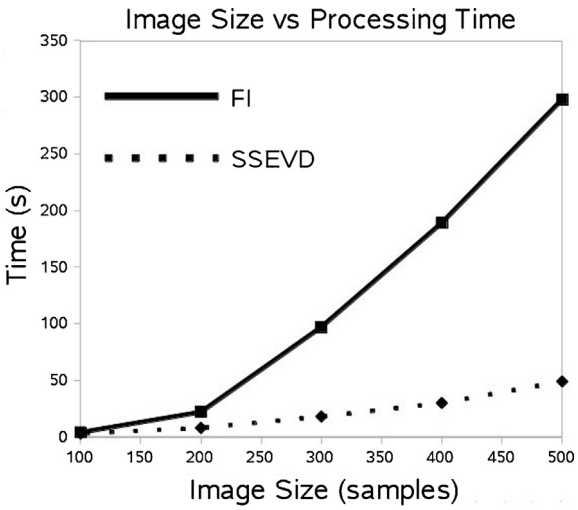

Fig. 4. Image size (given as number of samples for square image data) versus time (s) for calculating the eigenvectors of the Cuprite data.

5. Results

5.1. Cuprite results

Fig. 4 demonstrates the decrease in processing time required for SSEVD compared with FI using different image sizes (see Fig. 2). As image size increases the difference in processing speed becomes significant. Note, changing the $t_{EV}^{\%}$ does not affect the time required to derive the local eigenvectors. However, the total number of compiled vectors does change (Table 1), which effects the time required to run the second SVD. For the Cuprite data this difference is minimal and requires only 1 additional second.

The remaining results for the Cuprite data are given for only the 500×500 image region. The number of subsets for this image size is 1444. Table 1 shows the number of local eigenvectors retained for all subsets using different $t_{EV}^{\%}$. Using t_{SVD}^{99} up to 4 vectors per subset were retained, but with the majority being 1–3 vectors. This decreases for t_{EV}^{98} and t_{EV}^{95}, where for t_{EV}^{95} essentially only 1 vector per subset is retained.

Fig. 5 shows the difference in percent variance for the top ranked eigenvectors derived from SSEVD and FI. In this figure, and for subsequent figures, the eigenvector with the highest rank have an x-axis label of 1. For FI the variance drops off very quickly before leveling off. For SSEVD, using t_{SVD}^{95}, shows a similar profile as FI, whereas for t_{EV}^{98} and t_{EV}^{99} a more gradual decrease is observed. Table 1 and Fig. 5 emphasize two key points: (1) as fewer eigenvectors per subset are retained, SSEVD gives similar results, with respect to percent variance, as using FI; and, (2) the top eigenvectors derived using t_{SVD}^{98} and t_{SVD}^{99} threshold values comprise more spectral variance as a function of eigenvector rank than compared to those derived from FI, with the exception of the first eigenvector.

Table 1
Cuprite eigenvectors retained for all subsets for t_{EV}^{99}, t_{EV}^{98}, t_{EV}^{95}.

# of subsets retaining	SSEVD t_{EV}^{99}	SSEVD t_{EV}^{98}	SSEVD t_{EV}^{95}
1 Eigenvector	1444	1444	1444
2 Eigenvector	1125	542	20
3 Eigenvector	443	35	0
4 Eigenvector	19	0	0
Total Eigenvectors	3031	2021	1464

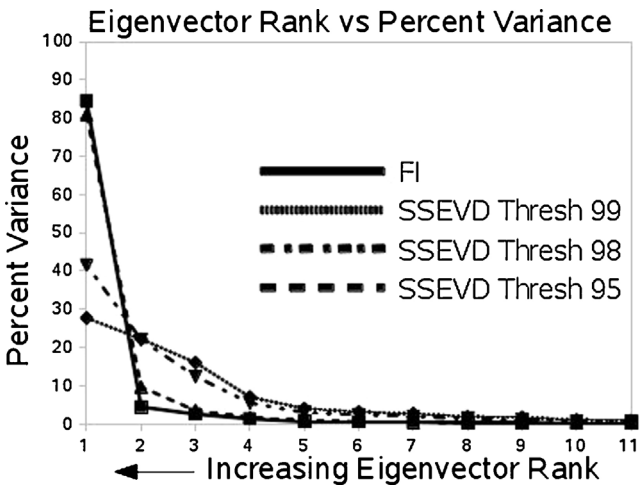

Fig. 5. Eigenvector percent variance as a function of eigenvector rank for FI and SSEVD approaches using Cuprite data. Only eigenvectors with greater than 99% variance are shown.

Fig. 6 shows the complete set of eigenvectors for SSEVD (t_{EV}^{99}) and FI, with Fig. 7 showing the spectral angle (Price, 1994) and root mean squared error (RMS) difference between them. Fig. 8 gives a visual of the profiles for the top 10 eigenvectors. These three figures demonstrate that the top few eigenvectors are similar, but this similarity decreases with rank. Closer examination of Fig. 6 shows that the lower rank vectors show different characteristics with respect to noise, where for FI specific band ranges have high noise, whereas for SSEVD this is less evident and the noise appears more uniform across the bands.

The transformed data for the top 10 FI and SSEVD (t_{EV}^{99}) eigenvectors is shown in Fig. 9. As with the vector profiles in Figs. 7 and 8, there are strong similarities for the first few vector transformations (1–5) with differences appearing more obvious with increasing rank (6–10). To better assess the differences we make use of spatial correlation indexes calculated for the FI and SSEVD (t_{EV}^{99}, t_{EV}^{98} and t_{EV}^{95}) transformed data (Fig. 10). It is evident from Fig. 10 that the lower ranked transformed bands for SSEVD using t_{EV}^{95}, t_{EV}^{98} and t_{EV}^{99} have index values closer to zero and negative compared with FI. In addition, the FI values show a higher degree of fluctuation. The implication here is that lower ranked eigenvectors for SSEVD appear to have minimal spatial correlation and are dominated by noise, whereas FI has a higher degree of spatial correlation suggesting potential signal.

5.2. Sokolov results

SSEVD was successfully run on the 7 flight-line Sokolov HyMap image data (15,962,624 total pixels), where the processing times required using t_{EV}^{98}, t_{EV}^{95} and t_{EV}^{90} were 1384, 1337 and 1299 s, respectively. As the total number of compiled subset eigenvectors for the Sokolov data is significantly higher compared to the Cuprite example, there is a more noticeable time difference for each threshold value. However, this difference is within about 6%. Fig. 11 shows eigenvector percent variance as a function of rank for the two 512 × 512 test regions (see Fig. 3) and for all 7 flight-lines (SSEVD only). The results in each plot show similar characteristics between the pixel regions and for all 7 flight-lines. In addition, the

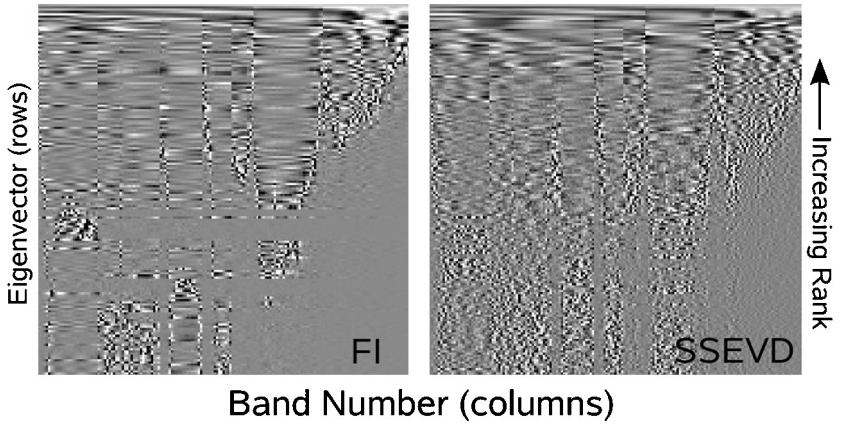

Fig. 6. Eigenvectors derived for the FI (left) and SSEVD (right) (t_{EV}^{99}) for the Cuprite data. Each row is an eigenvector (brightness reflects higher values), where rank decreases from top to bottom.

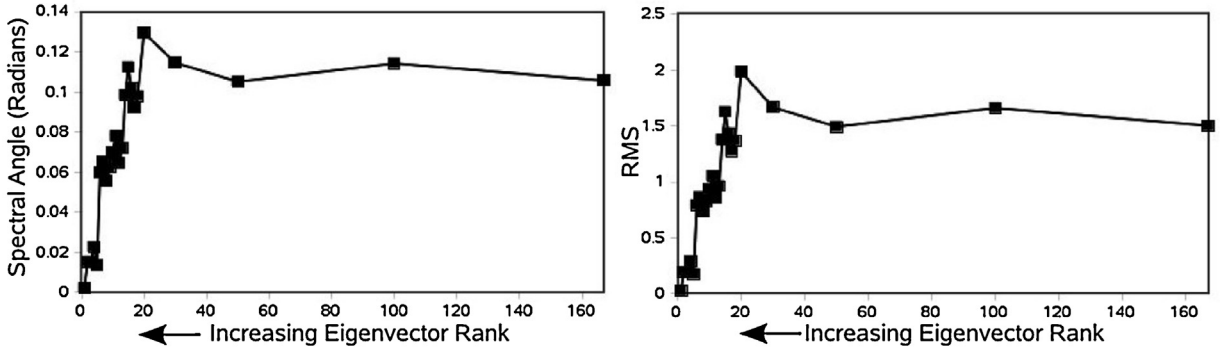

Fig. 7. Spectral angle and RMS differences between eigenvectors derived from FI and SSEVD (t_{EV}^{99}) for the Cuprite data.

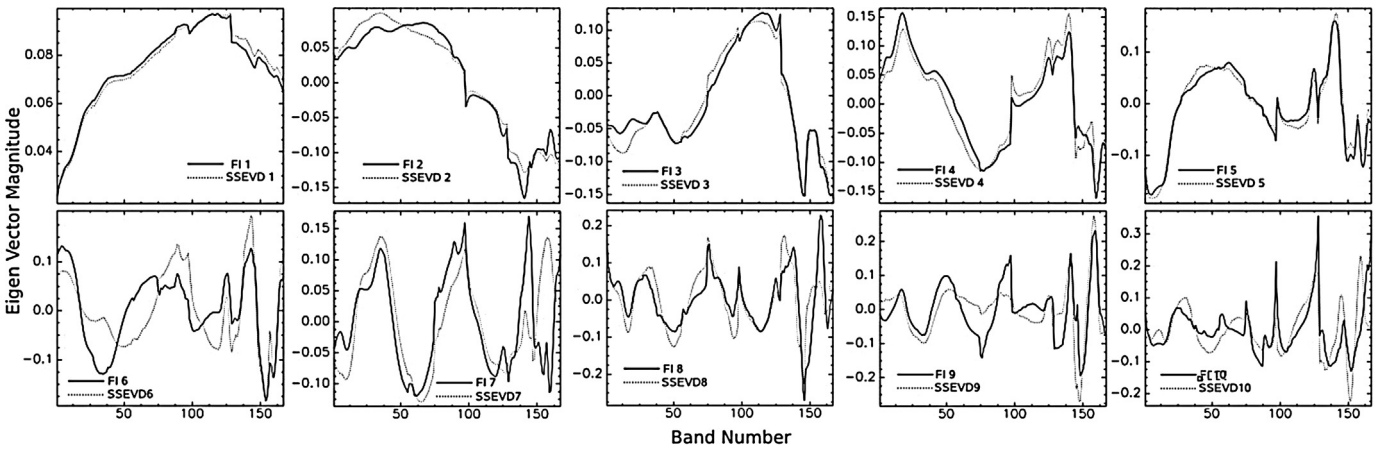

Fig. 8. Top 10 eigenvector profiles for the 500 × 500 Cuprite image size showing comparison of results from the FI versus SSEVD ($t_{EV}{}^{99}$) approaches.

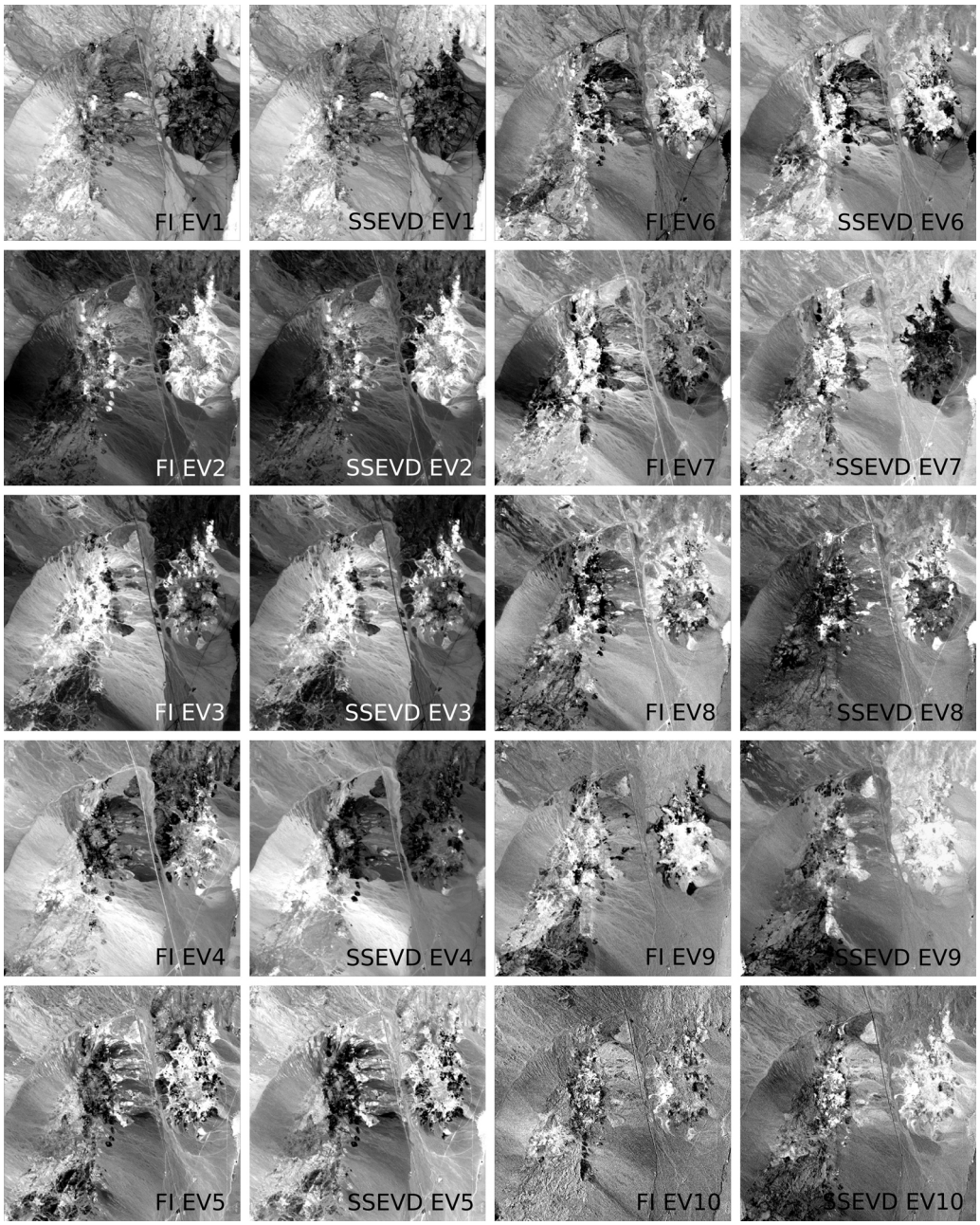

Fig. 9. Top 10 eigenvector transformation images comparing results from the FI versus SSEVD ($t_{EV}{}^{9}$) approaches for the Cuprite data.

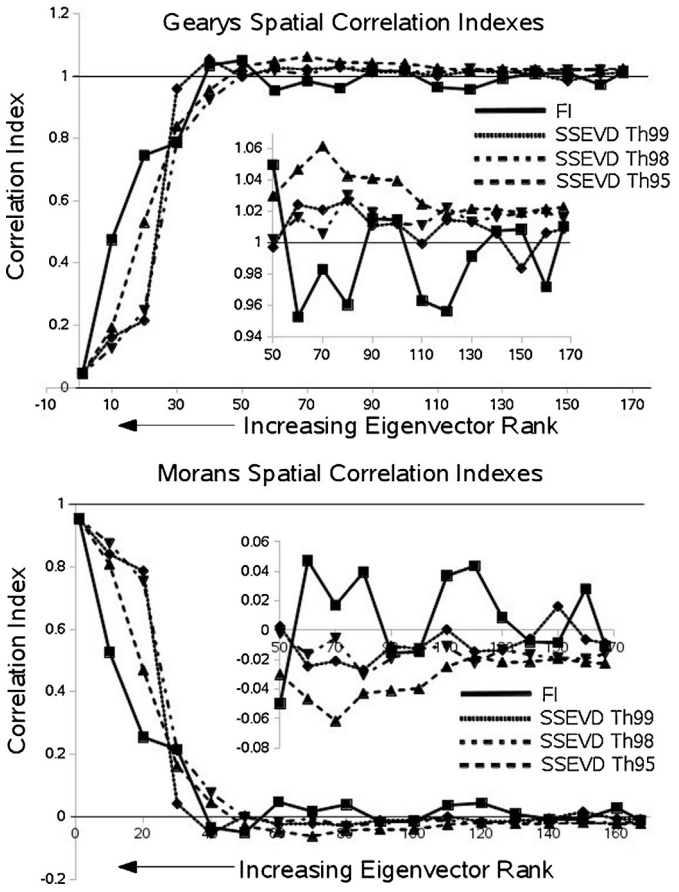

Fig. 10. Cuprite (500 × 500 pixel region) Geary's (top) and Moran's (bottom) spatial correlation indexes for FI and SSEVD (t_{EV}^{95}, t_{EV}^{98}, and t_{EV}^{99}) versus eigenvector rank transformed images (every 10th rank shown). Insets show values for transformed images with rank 50–167. Degree of spatial correlation: Geary's – positive = 0, no = 1, negative = 2; Moran's – positive = 1, no = 0, −1 negative (Griffith, 1987).

results are also similar to that observed with the Cuprite example, with respect to FI and SSEVD using different threshold values. Thus, we can assume that FI would show similar characteristics if it was applied to the 7 flight-lines.

For the Sokolov data the top eigenvectors derived using each $t_{EV}^{\%}$ are similar, with differences increasing with rank (Fig. 12). This indicates that using different $t_{EV}^{\%}$ (and SSEVD as a whole) does not have a large impact on the top eigenvectors. The main difference lies in those eigenvectors that contain more subtle signal and may be intermixed with those that represent noise variance. To assess this we must again consider the spatial correlation results for the two test regions and for all 7 flight-lines (Fig. 13). For A and B test regions FI shows higher spatial correlation and more fluctuation for the lower ranked eigenvectors (>40) compared with SSEVD at t_{EV}^{90} and t_{EV}^{95}. On the other hand t_{EV}^{98} shows only slightly better results than compared to the FI results. Based on the results given in

Table 2
Sokolov eigenvectors retained for all subsets for t_{EV}^{99}, t_{EV}^{98}, t_{EV}^{95}.

# of subsets retaining	SSEVD t_{EV}^{99}	SSEVD t_{EV}^{98}	SSEVD t_{EV}^{95}
1 Eigenvector	130,226	130,226	130,226
2 Eigenvector	122,907	77,679	1503
3 Eigenvector	29,214	26,646	20
4 Eigenvector	1635	0	0
5 Eigenvector	13	0	0
Total Eigenvectors	546,995	234,551	131,749

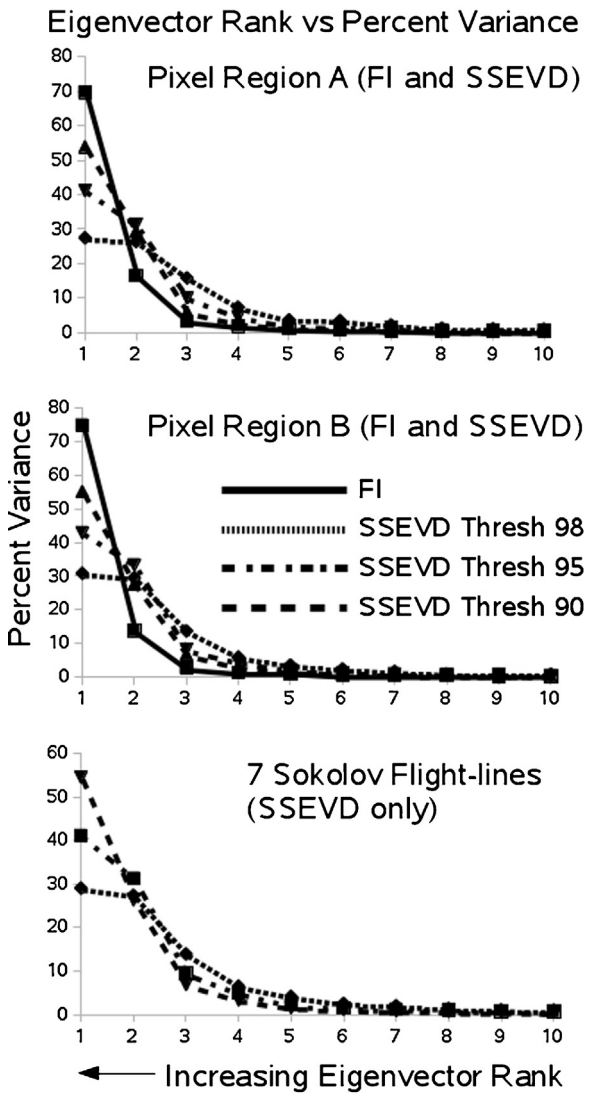

Fig. 11. Sokolov data eigenvector percent variance as a function of eigenvector rank for FI and SSEVD approaches. Top 10 eigenvectors shown.

Table 2 and Figs. 11 and 13, a threshold less than t_{EV}^{98}, but greater than t_{EV}^{95} may be more optimal for this data.

Fig. 14 shows a section crosscutting 6 of the 7 flight-lines (see Fig. 3) and demonstrates how SSEVD can effectively work on multiple flight-lines where spatial correlation of the transformed flight-lines continues across all lines. This figure includes a true color RGB reflectance image for reference (Fig. 14, left) with three other RGB combinations representing the top 9 eigenvector transformations. These combinations highlight a high level of spectral variability within forested, agricultural, mine, and urban areas. However, bidirectional reflectance distribution function (BRDF) effects visible in the reflectance image also show up to some extent in the transformed data, which is not unexpected.

6. Discussion

6.1. Input parameters

For SSEVD there are two input parameters, subset size and $t_{EV}^{\%}$. For the subset size, as noted above, the minimum number of samples must be greater than $p+1$. In this paper the minimum subset size was use for both data sets (13×13, 11×11), both of which

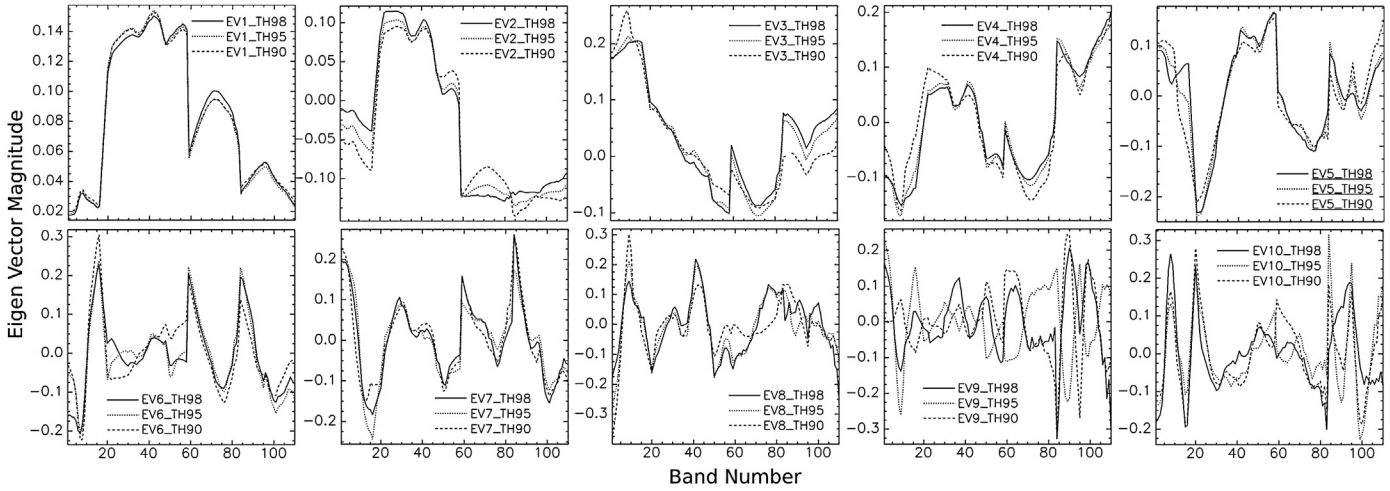

Fig. 12. Top 10 eigenvector profiles for the 7 flight-line Sokolov data showing comparison of results for the SSEVD approach using $t_{EV}{}^{95}$, $t_{EV}{}^{98}$, and $t_{EV}{}^{99}$.

produced good results. However, subset size could also reflect the spatial characteristics and the distribution of the various materials in the scene. For example, a larger subset size could be used in the analysis of images characterized by large homogeneous patterns (e.g. agricultural region). On the other hand, more spatially complex images (e.g. urban regions) may benefit from using the minimum subset size. In our two examples we use hyperspectral data, which results in a large minimum subset size compared with imaging systems with a limited number of bands. For multispectral data it is likely that using the minimum size may not be appropriate (e.g. 3×3) as the limited number of pixels may not allow for a large enough statistical distribution to obtain representative local eigenvectors. This, and other aspects of applying SSEVD to multispectral systems, is an avenue of future research. However, it is noted that the spatial implementation of the SSEVD approach makes it particularly attractive to process spatially large multispectral scenes.

The maximum eigenvector cutoff threshold ($t_{EV}{}^{100}$) will retain as many vectors as bands per subset. However, this is not physically realistic as the number of surface material within the given subset size is likely much smaller. Thus, the majority of the vectors can be

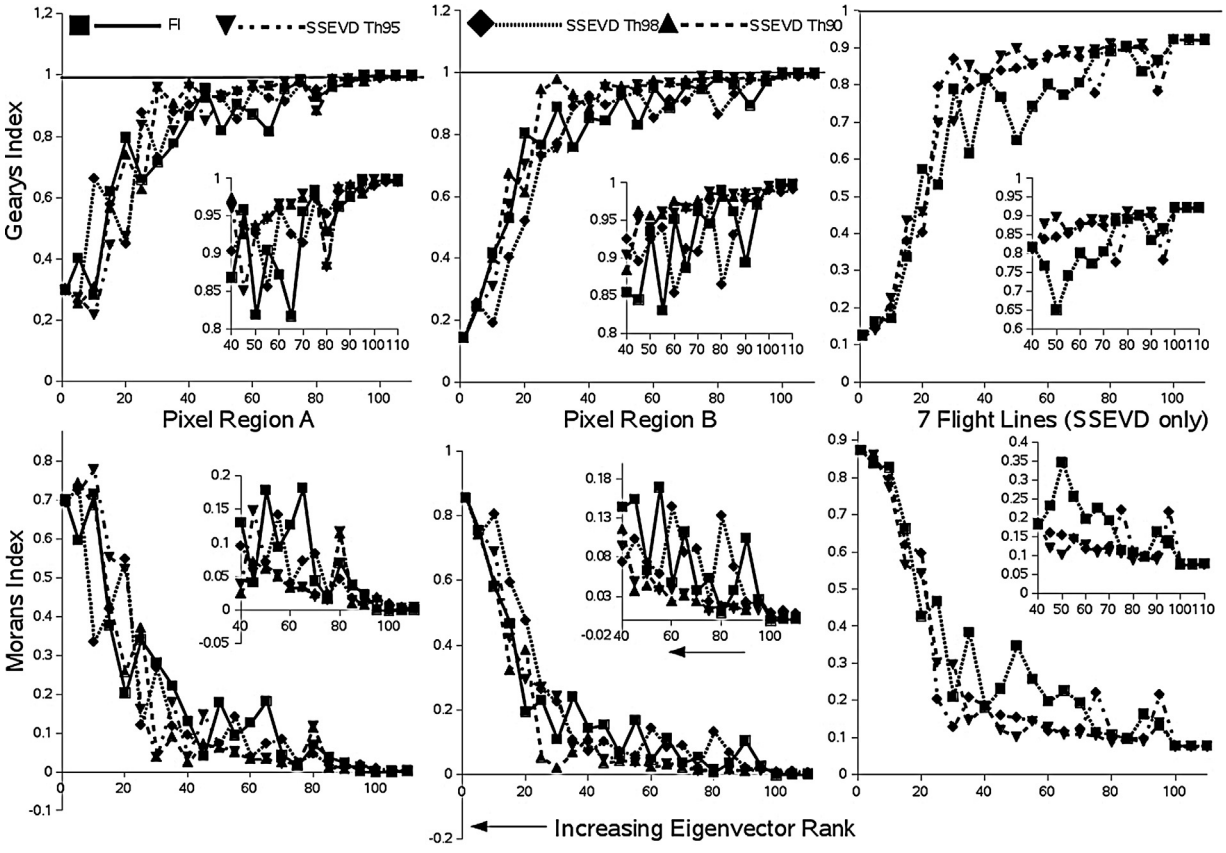

Fig. 13. Sokolov (pixel regions A, B and for all 7 flight-lines) Geary's and Moran's spatial correlation indexes for FI and SSEVD ($t_{EV}{}^{90}$, $t_{EV}{}^{95}$, and $t_{EV}{}^{8}$) versus eigenvector rank transformed images (every 5th rank shown). Insets show values for transformed images with rank 40–110. Degree of spatial correlation: Geary's – positive = 0, no = 1, negative = 2; Moran's – positive = 1, no = 0, −1 negative (Griffith, 1987).

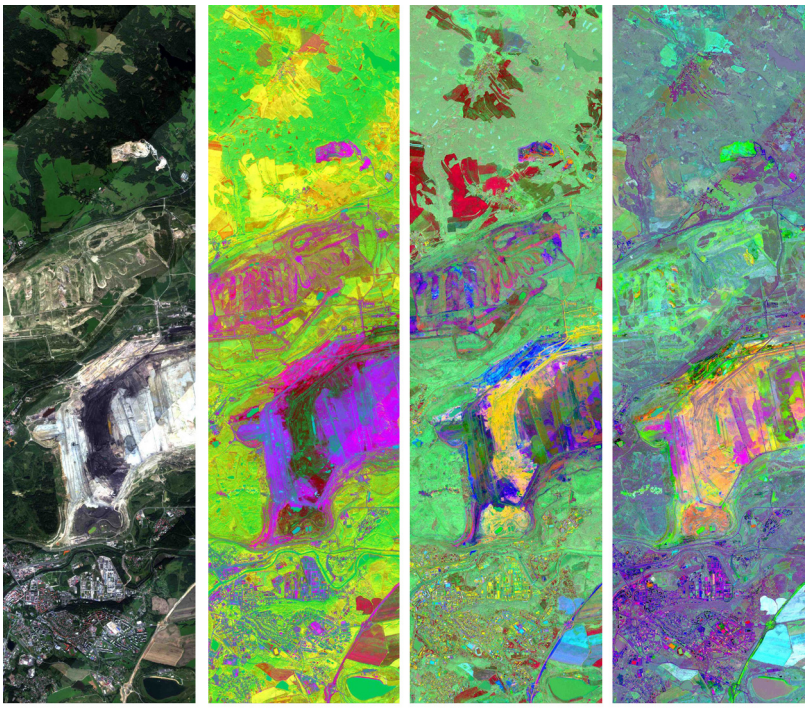

Fig. 14. Sokolov SSEVD eigenvector transformation (t_{EV}^{95}) RGB combinations for a select region of 6 of the 7 flight-lines (see Fig. 3) with left image showing true color RGB.

attributed to noise and will have an adverse effect on the process. A minimum threshold will result in only one vector per subset. This was essentially the case for t_{EV}^{95} and t_{EV}^{90} used for the Cuprite and Sokolov data, respectively. In both cases the results showed similar characteristics as FI, with respect to the percent spectral variance attributed to each eigenvector. This is attributed to the fact that the first eigenvector for each subset region is commonly related to the relative albedo of the dominant surface component. In addition, with this type of data the first eigenvector commonly has a similar shape to the average spectra for the same subset region. Thus, in the case where only the first eigenvector is retained in each subset, the dominant characteristic of the compiled eigenvector space is controlled by albedo, which is also the case for the reflectance image.

With respect to spatial correlation in the transformed images, retaining one eigenvector per subset did result in lower and more consistent spatial correlation index for lower rank bands. This indicates a good separation between eigenvectors dominated by signal and those with by noise. However, retaining only one vector per subset may result in the loss of less dominant surface materials. Thus, the use of higher $t_{EV}^{\%}$ will improve the prospect of retaining all surface materials, yet is still effective at separating signal and noise as shown in the given experiments.

In the case of the Cuprite and Sokolov data sets the optimal $t_{EV}^{\%}$ were not the same. For Cuprite a value of t_{EV}^{99} or t_{EV}^{98} showed better results, whereas for Sokolov the value was t_{EV}^{95}. This may reflect the characteristics of the images, such as the spatial distribution and spectral variability of the surface materials. However, to note, in each case the more optimal $t_{EV}^{\%}$ resulted in about the same number of eigenvectors per subset retained (1–3) for subset sizes of 13×13 and 11×11. This may indicate that given a small subset window it is likely that only a small number of eigenvectors is required to explain the local spectral variability, which was an initial assumption of SSEVD. Determining the optimal $t_{EV}^{\%}$ for each data set can be accomplished through the use of initial tests, where applying SSEVD on a selection of small test regions, as was done with the Sokolov example, can give the user a quick idea of an appropriate $t_{EV}^{\%}$. In this demonstration we choose to use percent of spectral variance, which is quite a simple approach. Other, more advanced methods, such as older, but well known Akaike information criteria (AIC) (Akaike, 1974) and minimum description length (MDL) (Rissanen, 1978), or the more recent Harsanyi–Farrand–Chang (HFC) (Chang and Du, 2004) and eigenvalue likelihood maximization (ELM) (Andreou and Karathanassi, 2013) could also be used and may further improve the results of SSEVD. However, integration must consider a balance between increased complexity and computational load versus improvement.

6.2. Implications on estimating dimensionality and endmembers

The final set of eigenvectors derived from SSEVD are similar to those derived using FI, specifically for those with high rank, even though the associated percent variance attributed to the vectors are different. This has implications for methods such as endmember extraction or estimating dimensionality. For example, if one makes use of the percent variance attributed to each eigenvector to determine the number of endmembers or dimensionality within a given scene, SSEVD will result in a higher dimensionality or number of endmembers (assuming more than one eigenvector per subset is retained). This may indicate that the existing methods that use percent variance may be underestimating the dimensionality or the number of endmembers in a given image. If this is the case one may also speculate that SSEVD will result in a better representation of the various surface materials in the scene. If more advanced dimensionality estimation methods are used (e.g. HFC or ELM) this may or may not be the case, and thus, this is avenue for further research. One endmember extraction method has additional implications based on the work presented in this paper. The spatial spectral endmember extraction tool (SSEE), developed by similar authors, uses the same approach to derive subset eigenvectors, which are used to project the data and select candidate endmembers (Rogge et al., 2007, 2012). Thus, the derivation of subset eigenvectors shows that the inherent design of SSEE makes it

less susceptible to noise, which was not discussed in the two earlier papers on SSEE.

6.3. Processing load

Overall SSEVD was shown to be significantly faster than FI, where this difference increases with larger images. For SSEVD the total time required is controlled by the derivation of local eigenvectors. If $t_{EV}^{\%}$ is increased, the number of compiled local vectors increases as well. However, even in this case deriving the final set of eigenvectors requires minimal additional time. We did not demonstrate the effect of using larger subset sizes in this paper. However, based on earlier work with SSEE (Rogge et al., 2007, 2012), which generates local vectors in the same fashion, subsets up to twice the minimum size have minimal effect on processing time. If significantly larger subsets are used then processing time can increase substantially, but as subset size increases, the advantages of the spatial approach are lost.

The design of SSEVD allows it to run easily on data sets that include multiple images. This was demonstrated with the 7 flightline Sokolov data set. As noted in Du and Fowler (2008) subsampling of an image needs to be relatively small in order to significantly reduce computational load, but must also be sufficiently large so as to adequately capture the statistical variation in the data set. Methods that use partial sampling (e.g. random, stratified, systematic, cluster (Congalton, 1991)) physically remove data and run the risk of not having a statistical representation of the data or even missing materials altogether that have a limited distribution in a given scene. With the SSEVD approach the input data is still the full data set, where subsampling uses instead contiguous spatial subsets that are represented by a limited number of eigenvectors. This spatial and spectral subsampling approach also has implications with respect to temporal data sets where images may be constantly added over time. In this case SSEVD can archive the existing subset eigenvectors and if new images are acquired, new subset vectors can simply be generated and added to the archive to derive a new set of final eigenvectors. This significantly minimizes the need to reprocess data as new data is acquired. In addition, comparing the updated eigenvector set to the old set may be particularly useful for applications related to change detection.

The implementation of SSEVD given here did not include parallel processing. However, the spatial design of SSEVD makes it particularly suited for parallel processing, where processing speed can be further improved. Combining the SSEVD approach with parallel processing will be particularly advantageous for processing very large quantities of data from more commonly available airborne or spaceborne surveys. Implementation of the SSEVD approach using parallel processing techniques is being investigated.

7. Conclusions

This paper has presented a spatial–spectral approach to generating high quality eigenvectors (SSEVD) for image transformations. The spatial–spectral approach was shown to be significantly faster than running the same process on the full image data simultaneously. This improved processing speed is accomplished without loss of quality of data and the approach is well suited to work with very large spatial and/or temporal data sets. The results shown here also demonstrated that the approach possesses an inherent ability to reduce the effects of noise from image data, such that the resulting orthogonal eigenvectors are better separated into those that represent signal, including rare signal sources, as opposed to noise. This is accomplished without the need to estimate noise, or consider the mean value of the subset eigenspaces. This approach also has important implications for estimating intrinsic dimensionality and the number of endmembers within a scene. Lastly the approach is well suited to make use of parallel processing method so that it can work effectively with increasingly large data sets, such as those that will become available with the launch of new hyperspectral satellites.

Acknowledgements

The authors gratefully acknowledge the support from the different project partners, especially from the Czech Geological Survey and from Sokolovská Uhelná.

References

Akaike, H., 1974. A new look at the statistical model indentification. IEEE Transactions on Automatic Control AC-19, 716–723.
Andreou, C., Karathanassi, V., 2013. Estimation of the number of endmembers using robust outlier detection method. IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing (in press).
Boardman, J.W., Kruse, F.A., Green, R.O., 1995. Mapping target signatures via partial unmixing of AVIRIS data. In: Summaries, Fifth JPL Airborne Earth Science Workshop, vol. 1, JPL Publication 95-1, pp. 23–26.
Buckingham, R., Staenz, K., 2008. Review of current and planned civilian space hyperspectral sensors for EO. Canadian Journal for Remote Sensing 34, S187–S197.
Chang, C.-I., Du, Q., 2004. Estimation of number of spectrally distinct signal sources in hyperspectral imagery. IEEE Transactions on Geoscience and Remote Sensing 42, 608–619.
Clark, R.N., Swayze, G.A., Livo, K.E., Kokaly, R.F., Sutley, S.J., Dalton, J.B., 2003. Imaging spectroscopy: earth and planetary remote sensing with the USGS Tertracorder and expert systems. Journal of Geophysical Research 108 (E12), 5-1-5-44.
Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment 37, 35–46.
Dadon, A., Ben-Dor, E., Karniele, A., 2010. Use of derivative calculations and minimum noise fraction transform for detecting and correcting the spectral curvature effect (Smile) in Hyperion images. IEEE Transactions on Geoscience and Remote Sensing 48, 2603–2612.
Danaher, S., O'Mongain, E., 1992. Singular value decomposition in multispectral radiometry. International Journal of Remote Sensing 13, 1771–1777.
Du, Q., Fowler, J.E., 2008. Low-complexity principal component analysis for hyperspectral image compression. The International Journal of High Performance Computing Applications 22, 438–448.
Du, Q., Zhu, W., Yang, H., Fowler, J.E., 2009. Segmented principle component analysis for parallel compression of hyperspectral imagery. IEEE Transactions Geoscience and Remote Sensing Letters 6 (4), 713–717.
El-Ghazawi, T., et al., 2002. Parallel and adaptive reduction of hyperspectral data to intrinsic domentionality. In: Proceedings of the IEEE International Conference on Cluster Computing, October 11, Newport Beach, CA, USA, pp. 102–109.
Franco, A., Lumini, A., Maio, D., 2002. Eigenspace merging for model updating. In: Proceedings 16th International Conference on Pattern Recognition, August 11–15, vol. 2, pp. 156–159.
Golub, G., Van Loan, C., 1996. Matrix Computations. Johns Hopkins Univ Press, Baltimore.
Green, A.A., Berman, M., Switzer, P., Craig, M.D., 1988. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. IEEE Transactions Geoscience and Remote Sensing 26 (1), 65–74.
Griffith, D.A., 1987. Spatial Autocorrelation: A Primer. Association of American Geographers, Washington, DC.
Hall, P., Marshall, D., Martin, R., 2000. Merging and splitting eigenspace models. IEEE Transactions on Pattern Analysis and Machine Intelligence 22, 1042–1049.
Jolliffe, I.T., 1986. Principal Component Analysis. Springer Verlag, New York.
Kawashima, T., Narimatsu, Y., Inada, H., Ishida, J., Hamada, K., Ito, Y., Yoshida, J., Ohgi, N., Tatsumi, K., Harada, H., Kawanishi, T., Sakuma, F., Iwasaki, A., 2010. The functional evaluation model for the on-board hyperspectral radiometer. In: Larar, A.M., Chung, H.S., Suzuki, M. (Eds.), Multispectral, Hyperspectral, and Ultraspectral Remote Sensing Technology, Techniques, and Applications III. Spie-International Society for Optical Engineering, Bellingham.
Pearlman, J., Carman, S., Segal, C., Jarecke, P., Clancy, P., Browne, W., 2001. Overview of the Hyperion Imaging Spectrometer for the NASA EO-1 mission. In: Geoscience and Remote Sensing Symposium, IGARSS, vol. 7, Sydney, Australia, pp. 3036–3038.
Plaza, A., Benediktsson, J.A., Boardman, J.W., Brazile, J., Bruzzone, L., Camps-Valls, G., Chanussot, J., Fauvel, M., Gamba, P., Gualtieri, A., Marconcini, M., Tilton, J.C., Trianni, G., 2009. Recent advances in techniques for hyperspectral image processing. Remote Sensing of Environment 113, S110–S122.
Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T., 1992. Numerical Recipes in C: The Art of Scientific Computing, second ed. Cambridge University Press, Cambridge.
Price, C., 1994. How unique are spectral signatures? Remote Sensing of Environment 49, 181–186.

Ready, P.J., Wintz, P.A., 1973. Information extraction, SNR improvement and data compression in multi-spectral imagery. IEEE Transactions on Communications 21, 1123–1131.
Richter, R., 2010. Atmospheric/topographical correction for airborne imagery. In: ATCOR-4 User Guide, DLR-IB 565-02/10. DLR, Wessling, Germany.
Richards, J.A., 1993. Remote sensing Digital Analysis, An Introduction, second ed. Springer-Verlag, New York.
Rissanen, J., 1978. Modeling by shortest data description. Automatica 14, 465–471.
Rogge, D., Rivard, B., Zhang, J., Sanchez, A., Harris, J., Feng, J., 2007. Integration of spatial–spectral information for the improved extraction of endmembers. Remote Sensing of Environment 110, 287–303.
Rogge, D., Bachmann, M., Rivard, B., Feng, J., 2012. Spatial sub-sampling using local endmembers for adapting OSP and SSEE for large-scale hyperspectral surveys. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 5, 183–195.
Singh, A., Harrison, A., 1985. Standardized principal components. International Journal of Remote Sensing 6, 883–896.
Stuffler, T., Kaufmann, C., Hofer, S., Forster, K.P., Schreier, G., Mueller, A., Eckardt, A., Bach, H., Penne, B., Benz, U., Haydn, R., 2007. The EnMAP hyperspectral imager – An advanced optical payload for future applications in Earth observation programmes. Acta Astronautica 61 (1–6), 115–120.
Switzer, P., Green, A., 1984. Min/max Autocorrelation Factors for Multivariate Spatial Imagery, Tech. Rep. 6. Dept. of Statistics, Stanford University.
Thai, B., Healey, G., Slater, D., 1999. Invariant subpixel material identification in AVIRIS imagery. In: Proceedings JPL AVIRIS Workshop, JPL Publication, Pasadena, CA, 99-17.
Umberto, A., Cavalli, R.M., Palombo, A., Pignatti, S., Santini, F., 2009. Experimental approach to the selection of the components in the minimum noise fraction. IEEE Transactions Geoscience and Remote Sensing 47, 153–160.
Yang, H., Du, Q., Zhu, W., Banicescu, I., Fowler, J.E., 2008. Parallel data compression for hyperspectral imagery. In: Proceedings IGARSS, July 7–11, Boston, Massachusetts, USA, pp. 986–989.