Utah State University

DigitalCommons@USU

1-1-2013

Using Physically-Modeled Synthetic Data to Assess Hyperspectral Unmixing Approaches

Mathew Stites

Jacob Gunther

Todd Moon

Gustavious Williams

Follow this and additional works at: https://digitalcommons.usu.edu/sdl_pubs

Utah State University
MERRILL-CAZIER LIBRARY

Article

Using Physically-Modeled Synthetic Data to Assess Hyperspectral Unmixing Approaches

Matthew Stites [1],*****, **Jacob Gunther** [2], **Todd Moon** [2] **and Gustavious Williams** [3]

[1] C4ISR Systems Division, Space Dynamics Laboratory, 1695 North Research Park Way, North Logan, UT 84341, USA

[2] Department of Electrical and Computer Engineering, Utah State University, Logan, UT 84341, USA; E-Mails: jake.gunther@usu.edu (J.G.); todd.moon@usu.edu (T.M.)

[3] Department of Civil and Environmental Engineering, Brigham Young University, Provo, UT 84341, USA; E-Mail: gus.p.williams@byu.edu

***** Author to whom correspondence should be addressed; E-Mail: matthew.stites@sdl.usu.edu; Tel.: +1-435-713-3398.

Abstract: This paper considers an experimental approach for assessing algorithms used to exploit remotely sensed data. The approach employs synthetic images that are generated using physical models to make them more realistic while still providing ground truth data for quantitative evaluation. This approach complements the common approach of using real data and/or simple model-generated data. To demonstrate the value of such an approach, the behavior of the FastICA algorithm as a hyperspectral unmixing technique is evaluated using such data. This exploration leads to a number of useful insights such as: (1) the need to retain more dimensions than indicated by eigenvalue analysis to obtain near-optimal results; (2) conditions in which orthogonalization of unmixing vectors is detrimental to the exploitation results; and (3) a means for improving FastICA unmixing results by recognizing and compensating for materials that have been split into multiple abundance maps.

1. Introduction

Hyperspectral imaging is a remote sensing approach that simultaneously collects both spatial and spectral data. Spectral data are collected in hundreds of narrow contiguous bands that may cover the visible, near-infrared, and short-wave infrared (0.4–2.5 μm), the mid-wave infrared (3–5 μm), and/or the long-wave infrared (8–14 μm). Although the size of a pixel on the ground varies, spatial measurements typically consist of hundreds of pixels in both spatial dimensions. Such images contain a wealth of information and have found application in a broad range of fields such as food safety [1], agriculture [2], mineralogy [3], ecology [4], planetary exploration [5], and target detection [6], as well as many others.

There are a number of methods for exploiting hyperspectral image data to generate useful products. One common method is spectral unmixing. This process refers to one or both of two fundamental operations. The first is the identification of spectra that are representative of the distinct materials in the scene. These spectra are referred to as endmembers and the problem of identifying them as endmember extraction. It is possible that an endmember spectrum may not be found in an image pixel, even though the associated material is present in the scene. This occurs when the material associated with that endmember does not completely fill any single pixel in the image. In that case, which is not uncommon in real data, the endmember spectrum will only be present in an image pixel in combination with other endmember spectra. Because an endmember is uniquely associated with a specific material, the terms endmember and material are used interchangeably throughout the remainder or this paper.

The second spectral unmixing operation is abundance quantification, which entails determining the proportion of each endmember in each pixel of the image. Abundance maps provide useful visualizations of hyperspectral data, showing where each endmember is located in an image and how completely each pixel is filled by that endmember. Depending on the algorithm and the application, endmembers may be determined first and subsequently utilized for abundance quantification, the endmembers and abundances may be found simultaneously, or abundances may be computed without any prior endmember information [7].

There are a wide variety of algorithms that have been developed to unmix hyperspectral data. A recent survey article classified algorithms into one of four categories: (1) geometric; (2) statistical; (3) sparse regression based; and (4) spatial-spectral contextual [8]. Independent component analysis (ICA) is a statistical unmixing approach that does not assume a specific distribution for the data [9]. This approach attempts to unmix the data by finding maximally independent abundances. A variety of ICA algorithms have been applied to hyperspectral unmixing including contextual ICA [10], joint cumulant-based ICA [11], joint approximate diagonalization of eigenmatrices (JADE) [12], and FastICA [12–15]. ICA has also been employed as a hyperspectral classification approach [16,17]. There are still questions regarding the utility of ICA as a hyperspectral unmixing approach. A common opinion is that while ICA can produce interesting and useful results, it is common for some materials to be incorrectly unmixed [12,14,15]. Because of these lingering questions, the behavior of the FastICA algorithm [18] is examined more closely later in this paper. FastICA was selected over other ICA algorithms because of its wide use and straightforward implementation.

Whenever spectral unmixing algorithms are assessed, two types of experiments are typically performed. In the first, synthetic images are created according to a simple generative model—usually the

linear mixing model. The complexity of these images varies, but they are typically composed of 2–10 endmembers whose spectra are obtained from a real hyperspectral image or from a spectral reference library. In many cases spatial contiguity is incorporated using abundance maps consisting of simple square or circular regions. These kinds of test images are fairly common in the spectral unmixing literature [19–23]. Since many spectral unmixing approaches, including ICA, do not consider spatial context, synthetic images can also be produced using randomly generated abundances that adhere to some probability distribution. In these cases a generative model is used that incorporates other interesting behavior, such as topographic variation and endmembers with spectral variability [12,24]. In the majority of these cases the endmembers are generated in relative proportion to each other. That is, there is no single material that dominates the scene spatially and no material that is present in only a very small fraction of pixels. These images are useful because they are relatively simple to generate, and because complete ground truth data are available, including abundance maps accurate to small fractions of a pixel. Spectral unmixing results can then be compared against the ground truth data to provide quantitative assessments of algorithms.

The second type of experiment tests an algorithm by unmixing a real hyperspectral data set. The results of the unmixing are often assessed visually by recognizing landmarks in the original image and in the unmixed data [11,14]. In some cases ground truth data are available and can be compared with the unmixing results [10,20]. Unfortunately, these ground truth data often only provide information for a subset of the materials in the scene and may be incomplete for certain areas or materials in the image. They do not provide the fine abundance resolution of synthetic images and are not available for every image that might be of interest.

Both of the experimental approaches described above are useful and even essential in assessing the effectiveness and behavior of a hyperspectral algorithm. There is, however, a third approach that can be viewed as something of a middle ground between the two. This approach utilizes synthetic images that more closely approximate real data by modeling scene geometry, material properties, sensor behavior, atmospheric contributions, and so forth. Complex scene geometry is desirable because it produces images that have regions of spatial contiguity, topographic variation, and endmember spectral variability. This approach also leads to broad variations in the spatial coverage of individual materials. Because the images are synthetic, complete ground truth data are still available. Such an approach is not intended to be a replacement for the existing methods described above. Instead, it should be treated as a complementary approach, allowing for exploration of unique insights and observations.

This complementary approach could be employed to explore a variety of hyperspectral unmixing algorithms. However, throughout the remainder of this paper, it is used to assess the behavior of FastICA. This exploration is warranted to confirm existing assertions regarding FastICA and also to provide further insight into the behavior of the algorithm.

The remainder of the paper proceeds as follows. Section 2 provides a basic overview of ICA and the FastICA algorithm. It also outlines the ICA data model and compares it with the linear mixing model used to describe hyperspectral data. Section 3 explains the approach taken to generate synthetic—but realistic—hyperspectral data cubes. Examples of both image data and abundance maps are shown. Section 4 describes the experiments performed, presents the results of those experiments,

and provides insight into those results. Finally, Section 5 contains a few concluding observations and remarks.

2. Independent Component Analysis

Independent component analysis (ICA) is an approach for performing blind source separation (BSS). The generalized BSS problem is modeled as

$$\mathbf{x}(t) = f\left(\mathbf{s}(t)\right) \tag{1}$$

where $\mathbf{x}(t) = \begin{bmatrix} x_1(t) & x_2(t) & \dots & x_K(t) \end{bmatrix}^T$ is the K-dimensional observed data vector, $\mathbf{s}(t) = \begin{bmatrix} s_1(t) & s_2(t) & \dots & s_L(t) \end{bmatrix}^T$ is an L-dimensional vector of the sources of interest, and $f(\cdot)$ describes the mixing process that operates on the sources to create the observed data. The observations and sources are indexed by t, which depending on the application may represent time, spatial location, or some other quantity. In the case of hyperspectral unmixing, t is used to index spatial location, *i.e.*, individual pixels. The goal of blind source separation is to estimate the original sources from the observed data with limited or no knowledge of either $f(\cdot)$ or $\mathbf{s}(t)$. The estimation process is often referred to as unmixing. Blind source separation has found application in many varied areas including biomedical signal processing [25,26], telecommunications [27,28], and finance [29,30].

ICA is an approach that attempts to perform BSS by exploiting the statistical independence of the original sources. While this can be accomplished in a number of ways, many ICA algorithms invoke the central limit theorem [31], observing that the distribution of mixed random variables tends toward a Gaussian distribution. Hence, sources can be separated by optimizing a cost function that reflects some measure of Gaussianity. Commonly used cost functions include kurtosis and negentropy. Other ICA approaches include minimization of mutual information [32], and joint diagonalization of eigenmatrices [33].

Although nonlinear ICA methods exist [34,35], linear mixing is most commonly assumed. In this case the mixing is represented by

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \quad t = 1 \dots T \tag{2}$$

where $\mathbf{A}$, is the $K \times L$ mixing matrix and T is the total number of observations (pixels). Stacking the observed and source data as $\mathbf{X} = \begin{bmatrix} \mathbf{x}(1) & \mathbf{x}(2) & \dots & \mathbf{x}(T) \end{bmatrix}$ and $\mathbf{S} = \begin{bmatrix} \mathbf{s}(1) & \mathbf{s}(2) & \dots & \mathbf{s}(T) \end{bmatrix}$, the model becomes

$$\mathbf{X} = \mathbf{A}\mathbf{S} \tag{3}$$

with the $K \times T$ observation matrix, $\mathbf{X}$, and $L \times T$ source matrix, $\mathbf{S}$.

The mixed data must satisfy two important conditions for ICA to be a valid unmixing approach. First, since ICA attempts to unmix the data by exploiting the independence of the sources, the sources must be independent. Second, because the methods of separation utilized by ICA algorithms attempt to maximize non-Gaussianity (based on the central limit theorem), no more than one source may be Gaussian distributed [18].

2.1. FastICA

FastICA is an ICA algorithm that assumes the linear mixing model in Equation (3) with the additional constraint that the number of observations must match the number of sources, *i.e.*, $K = L$, making the mixing matrix $\mathbf{A}$ square. The unmixing model then becomes $\mathbf{Y} = \mathbf{BX}$, where $\mathbf{Y}$ contains the estimates of the original sources. Defining the unmixing matrix to be

$$\mathbf{B} = \begin{bmatrix} \mathbf{b}_1^T \\ \mathbf{b}_2^T \\ \vdots \\ \mathbf{b}_K^T \end{bmatrix} \tag{4}$$

a single independent component can be obtained as

$$y_i(t) = \mathbf{b}_i^T \mathbf{x}(t) \tag{5}$$

or equivalently,

$$\mathbf{y}_i^T = \mathbf{b}_i^T \mathbf{X} \tag{6}$$

Since neither reordering nor scaling of the estimates affects their independence, ICA outputs are subject to scale ambiguity and order uncertainty. Because of this, any result of the form $\mathbf{y}_i^T = \gamma \mathbf{s}_j^T$, where γ is a constant scalar value, is generally considered a success.

Prior to performing any source separation the observed data are whitened where $\mathbf{z}(t) = \mathbf{Vx}(t)$, such that $E[\mathbf{z}] = \mathbf{0}$, and $E[\mathbf{zz}^T] = \mathbf{I}$. Incorporating the whitened data, the unmixing model becomes $\mathbf{Y} = \mathbf{WZ} = \mathbf{WVX}$, and $\mathbf{B} = \mathbf{WV}$, where $\mathbf{W}$ is comprised of stacked vectors as $\mathbf{B}$ in Equation (4).

As part of the whitening process the dimension of the observed data is reduced via principal component analysis (PCA). Unless specified by the user, the number of dimensions is determined automatically from the relative magnitudes of the eigenvalues of the covariance matrix of the observed data. This dimension reduction step is an attempt to estimate the number of sources and make the mixing matrix square, as required by the FastICA model.

After whitening and dimension reduction, the source separation is achieved by using a simple fixed-point algorithm to maximize a cost function. Thus, the source separation problem becomes

$$\max_{\mathbf{w}_i} E\left\{ G\left(\mathbf{w}_i^T \mathbf{z}(t) \right) \right\}, \quad i = 1, \ldots, K \tag{7}$$

Typically, $G(\cdot)$ in Equation (7) is defined to be

$$G_1(y) = y^4 \tag{8}$$

$$G_2(y) = \frac{1}{a_1} \log \cosh(a_1 y) \tag{9}$$

or

$$G_3(y) = -\frac{1}{a_2} \exp(-a_2 y^2/2) \tag{10}$$

The derivatives of these functions are

$$g_1(y) = 3y^3 \tag{11}$$

$$g_2(y) = \tanh(a_1 y) \tag{12}$$

and

$$g_3(y) = y \exp(-a_2 y^2/2) \tag{13}$$

The first function is an approximation of the kurtosis of y. Incorporating either of the other two functions gives an approximation of the negentropy of y.

Because the whitening step effectively orthogonalizes the observed data, the unmixing matrix, $\mathbf{W}$, is constrained to be an orthogonal matrix with $\mathbf{W}\mathbf{W}^T = \mathbf{W}^T\mathbf{W} = \mathbf{I}$. This constraint is enforced at each iteration of the cost function optimization in one of two ways. If the components are extracted one at a time, deflationary orthogonalization is performed. This approach updates a single unmixing vector using the gradient optimization algorithm. That vector is then made orthogonal to all of the previously computed unmixing vectors:

$$\widetilde{\mathbf{w}}_i = \mathbf{w}_i - \sum_{j=1}^{i-1} (\mathbf{w}_i^T \widetilde{\mathbf{w}}_j)\widetilde{\mathbf{w}}_j \tag{14}$$

The unmixing vector is then normalized as

$$\widetilde{\mathbf{w}}_i = \widetilde{\mathbf{w}}_i / \|\widetilde{\mathbf{w}}_i\| \tag{15}$$

Alternatively, if all of the components are estimated simultaneously, then symmetric orthogonalization is performed. In this case all L unmixing vectors are updated and subsequently orthogonalized using the update formula

$$\widetilde{\mathbf{W}} = \left(\mathbf{W}\mathbf{W}^T\right)^{-1/2} \mathbf{W} \tag{16}$$

2.2. Application to Hyperspectral Data

One approach to modeling the radiance of a single pixel in a hyperspectral image is the linear mixing model [36]. This model is typically formulated as

$$\mathbf{x}(t) = \sum_{l=1}^{L} \mathbf{m}_l a_l(t) + \mathbf{n}(t) = \mathbf{M}\mathbf{a}(t) + \mathbf{n}(t) \tag{17}$$

In this model $\mathbf{x}(t)$ is the observed $K \times 1$ pixel where K is the number of spectral bands of the sensor. As described previously, the index t is used to indicate the spatial location of the pixel. The $K \times 1$ vector $\mathbf{m}_l$ represents an endmember spectrum and $a_l(t)$ is the fractional abundance of that endmember in the pixel. The total number of endmembers is L. Instrument noise and model error are represented by $\mathbf{n}(t)$. The $K \times L$ matrix $\mathbf{M}$ is the endmember matrix and contains the L individual endmembers in its columns. The $L \times 1$ abundance vector, $\mathbf{a}(t)$, is formed by stacking the relative abundances. The relative abundances are subject to two constraints:

$$a_l(t) \geq 0, \quad l = 1, \ldots, L \tag{18}$$

and

$$\sum_{l=1}^{L} a_l(t) = 1 \tag{19}$$

These constraints impose the physically meaningful requirements that the fractional abundances be nonnegative and sum to one. This model is valid only when the materials in the pixel are well-partitioned from one another [36,37]. Even though this is not always the case in nature, this model is still widely used.

The pixels in the observed cube can be indexed in row-scanned order so that each spectral band is represented as a one-dimensional vector, rather than a two-dimensional image. Then, the terms on both sides of Equation (17) can be stacked as

$$\mathbf{X} = \mathbf{MA} + \mathbf{N} \tag{20}$$

where $\mathbf{X}$ and $\mathbf{N}$ are $K \times T$ matrices, $\mathbf{A}$ is an $L \times T$ matrix, and T is the total number of pixels in the image. In this arrangement a column of $\mathbf{X}$ is the spectrum of a specific pixel in the image and a row of $\mathbf{X}$ contains all of the pixels from one spectral band of the data, in row-scanned order. Similarly, a column of $\mathbf{A}$ describes the fractional abundances for every material in a single pixel while a row of $\mathbf{A}$ contains the fractional abundance in every pixel of a single material, again in row-scanned order.

Figure 1. Histograms of synthetically-generated abundance maps for (**a**) a sparse material; and (**b**) a dense material. Both of these are distributed in a way that is clearly non-Gaussian. Notice the change of scale in (a) required to display the non-zero abundance values. The left-most bin corresponding to zero actually extends above 16,000 pixels.

(**a**)

(**b**)
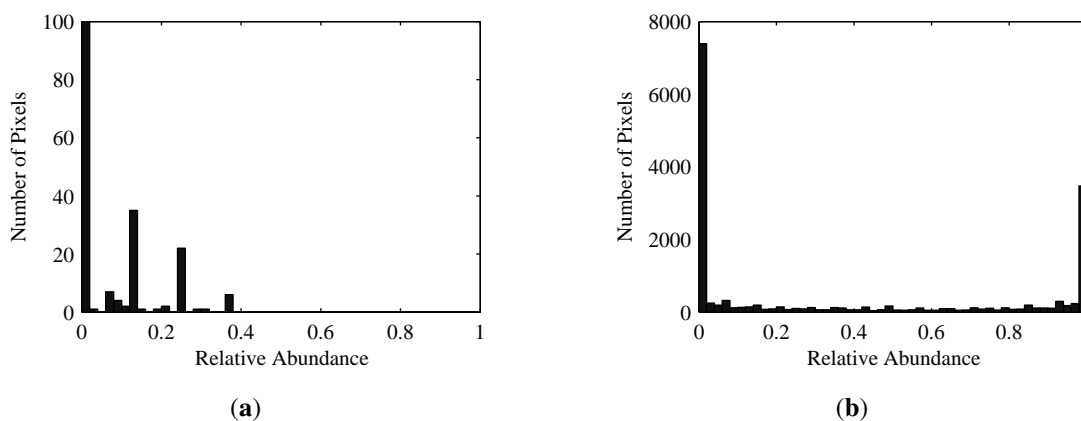
The hyperspectral mixing model in Equation(20) is structurally similar to the linear ICA model in Equation(3). The endmember matrix is analogous to the mixing matrix and the abundance matrix corresponds to the source matrix. The one difference is the addition of noise in the hyperspectral model. If the signal-to-noise ratio (SNR) is sufficiently large, the noise contribution may be safely ignored, in which case the models are identical. Otherwise, the noise effects could be minimized by smoothing, dimension reduction, or some other preprocessing step. Recall that the ICA model requires the sources to be non-Gaussian, implying that the fractional abundances for each material must not have a Gaussian distribution. This requirement is satisfied as abundance values tend to accumulate near zero or one depending on their spatial coverage and have a predominantly one-sided distribution. This behavior is illustrated in Figure 1, which shows histograms for abundance maps of two different materials generated from a three-dimensional model of a real-world scene. The other requirement imposed by ICA is that

the sources be independent. For the hyperspectral data model the abundance of each material is required to be independent of every other material. This requirement is violated by the additivity constraint in the linear mixing model Equation(19). Although this is a violation of the ICA assumptions, as the number of endmembers and/or signature variability increases, the statistical dependence of the sources decreases and ICA performance improves [12].

3. Experimental Data Description

In order to perform the kind of complementary experiments described earlier, a means of producing realistic images and the associated ground truth data is needed. This section describes the tool employed to produce the synthetic data that were incorporated into the experiments described in subsequent sections of this paper.

The Digital Image and Remote Sensing Image Generation (DIRSIG) software is a physics-based image simulation tool developed at the Rochester Institute of Technology (RIT) [38]. The tool allows the user to describe complex scene geometry, viewing geometry, and the spectral and thermal properties of materials in a scene. The user can also describe a variety of sensor properties including sensor type, scan behavior, focal length, detector layout, and spectral and spatial response. MODTRAN [39] is incorporated to simulate realistic atmospheric behavior from user-provided atmospheric and weather information. Incorporating all of this information, the software employs thermal and radiometric models along with a ray tracer to compute radiance fluxes at specific points [40]. The approach is used to generate realistic remote sensing images. Additionally, DIRSIG can also export the ground truth associated with each image.

For our experiments, two test images were generated using DIRSIG. Both images incorporate the "MegaScene" geometric scene description, which models a 0.6 square mile area of Rochester, New York. A pushbroom spectrometer model that incorporates a spectral response between 0.4 μm and 2.5 μm with 224 bands was used. The spectral response is similar to the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) [41]. The altitude of the sensor was 2 km. With these settings in place, 1,024 $\times$ 1,024 pixel cubes and truth maps were generated with a ground sampling distance (GSD) of 0.25 m. These were then binned spatially to produce 128 $\times$ 128 pixel radiance cubes and truth maps with a GSD of 2.0 m. The binning was performed to produce data with the desired linear mixing behavior. Randomly-generated Gaussian noise was added to the data to produce cubes with a variety of signal-to-noise ratios (SNRs).

The first radiance cube generated is referred to as "Mega1" because of its location within the first tile of the MegaScene. The scene is dominated by two large buildings surrounded by a parking lot. At the top of the image is a residential road with homes on either side that are mostly obscured by trees. Three tennis courts are located at the bottom of the image. The remainder of the scene is grass. There are 43 unique materials in this scene. The second radiance cube comes from the fourth MegaScene tile and is aptly named "Mega4." This scene contains ten large industrial tanks surrounded by some buildings and parking lots. Around the periphery of the scene are areas of trees and grass. This scene contains 21 unique materials. Examples of the synthetic data are shown in Figure 2.

A list of the materials contained in each scene is provided in Section 5. These materials are sorted by the number of pixels in which they appear and are loosely segregated into four categories based on their spatial coverage in the image. Super-sparse materials are those with a combined coverage of less than one pixel. Materials in the sparse category typically are present in 1% or less of the image pixels and cover less than 0.5% of the image. They may or may not appear in the image as pure pixels. Dense materials appear in over half of the pixels in the image and consequently also constitute a large number of pure pixels. Materials falling between the sparse and dense categories are classified as intermediate materials. This categorization is used to analyze how materials of varying spatial distribution are affected in the spectral unmixing process. This is an example of the type of assessment that is not usually made in the two most common experimental scenarios described in Section 1.

Figure 2. Examples of the test images generated in DIRSIG. (**a**) Grayscale image of Mega1; (**b**) Grayscale image of Mega4; (**c**) Mega1 abundance map for "Roof, Gravel, Gray"; (**d**) Mega4 abundance map for "Roof, Gravel, Gray".

(**a**)

(**b**)

(**c**)

(**d**)
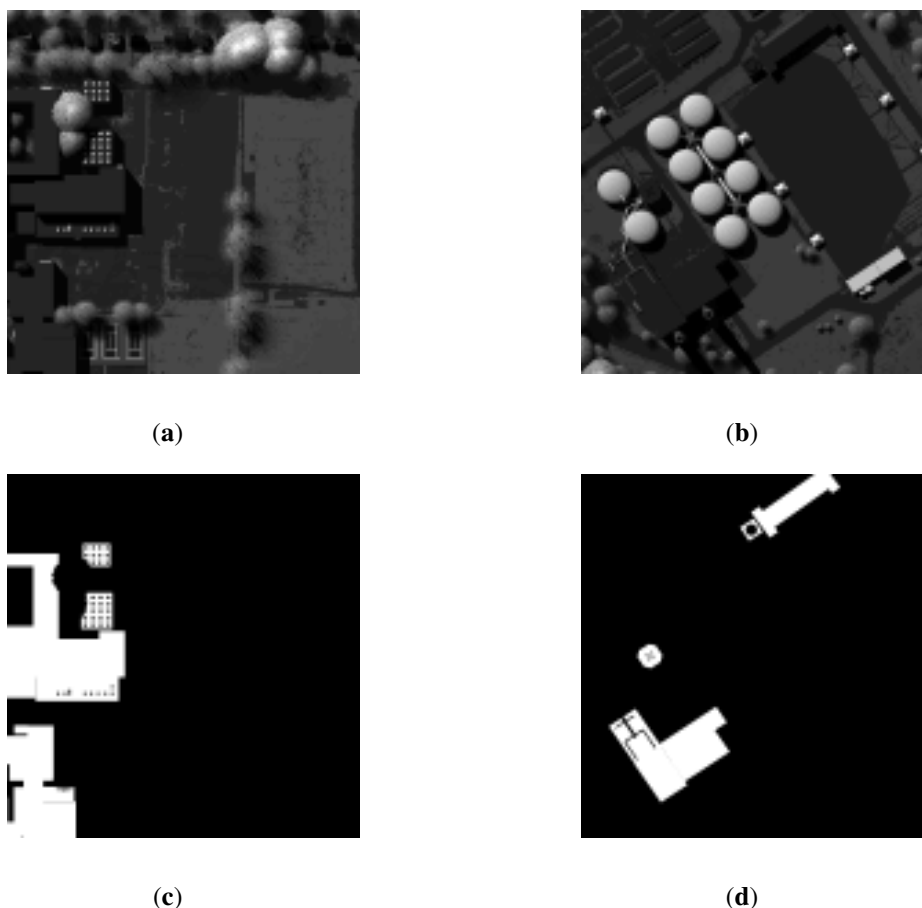
4. Experimental Results

Three sets of experiments were performed to characterize the utility of FastICA as a hyperspectral unmixing approach. The first set of experiments examined the impact of dimension reduction on the best-case unmixing scenario. Second, the effects of orthogonalization were explored, again considering a best-case unmixing scenario. Because dimension reduction and orthogonalization are not unique

to FastICA, these two experiments are of interest beyond the scope of FastICA. In the final set of experiments, unmixing was performed using FastICA. The results of these experiments are quantified by comparing estimated material abundances with corresponding abundance ground truth. The quality of endmember extraction was not considered in these experiments. Some observations are made in the following narrative on the effects of adding noise to the synthetic images, but complete characterization of the impact of noise on the unmixing process is beyond the scope of this paper.

For the remainder of this paper, whenever performance is plotted versus material, *i.e.*, the x-axis is "Material Number", the materials are numbered according to the lists in the appendix. The first (left-most) material in the plot is the most sparse and the last (right-most) is the most dense. Markers are used to denote the four categories of material spatial coverage. A circle (○) is used to identify super-sparse materials, a cross (×) for sparse materials, a diamond (◇) for intermediate materials, and a square (□) for dense materials.

4.1. Computation of Optimal Estimates

Because complete ground truth abundance maps are available, the optimal, linear unmixing vector and corresponding abundance estimate can be calculated for each material. This was done prior to performing any experiments. These results constitute a best-case unmixing scenario, *i.e.*, the best result FastICA could produce, and provide a baseline against which experimental results can be compared. A common metric used in such comparisons is mean-square error (MSE),

$$MSE \equiv \frac{1}{T} \sum_{t=1}^{T} (\hat{a}(t) - a(t))^2 \tag{21}$$

where $\hat{a}(t)$ is an estimated abundance and $a(t)$ is the ground truth abundance. However, MSE is not invariant to scaling, which is essential when considering ICA outputs, since they are subject to scale ambiguity. Thus, a preferred metric to MSE is the correlation coefficient, defined as

$$r(\hat{\mathbf{a}}, \mathbf{a}) \equiv \left(\frac{\hat{\mathbf{a}} - \mu_{\hat{\mathbf{a}}}}{\|\hat{\mathbf{a}} - \mu_{\hat{\mathbf{a}}}\|} \right)^T \left(\frac{\mathbf{a} - \mu_{\mathbf{a}}}{\|\mathbf{a} - \mu_{\mathbf{a}}\|} \right) \tag{22}$$

The absolute value of this metric is invariant to scaling of the arguments, as desired. Conveniently, this value also always falls in the range $[0, 1]$. It is used throughout the remaining experiments to quantify performance.

The unmixing formula Equation (5) in combination with the linear mixing model for hyperspectral data Equation (17) provides a formula for extracting individual abundances, $\hat{a}_i(t) = \mathbf{b}_i^T \mathbf{x}(t)$. Stacking this result to eliminate the spatial indexing yields $\hat{\mathbf{a}}_i^T = \mathbf{b}_i^T \mathbf{X}$. The unmixing vector that maximizes $r(\hat{\mathbf{a}}, \mathbf{a})$ is given by

$$\check{\mathbf{b}} = \left[(\mathbf{X} - \mu_{\mathbf{X}})(\mathbf{X} - \mu_{\mathbf{X}})^T \right]^{-1} (\mathbf{X} - \mu_{\mathbf{X}}) \frac{(\mathbf{a} - \mu_{\mathbf{a}})}{\|\mathbf{a} - \mu_{\mathbf{a}}\|} \tag{23}$$

The optimal abundance estimate is then

$$\check{\mathbf{a}}^T = \check{\mathbf{b}}^T \mathbf{X} \tag{24}$$

The optimal unmixing vectors and abundance estimates were calculated according to Equations (23) and (24), respectively, for every material in both of the test cubes. In the absence of noise, as shown

in Figure 3, the maximum correlation coefficient, $r(\check{\mathbf{a}}_i, \mathbf{a}_i)$, is very high overall. It can be seen that the correlation coefficient tends to improve with an increase in spatial coverage. The fact that the correlation coefficient is not exactly one for every material in the scene stems from illumination, endmember, and atmospheric variability in the DIRSIG-generated cubes. Figure 4 provides a visual comparison between ground truth and optimal estimates from Mega1 for one material from each of the four material coverage classifications. From these images it can be seen that material locations can be clearly discerned for values of $|r| \geq 0.8$. Below this threshold, the material locations are less clear and background artifacts become more obvious. Depending on the spatial coverage and congruency of a material, correlation coefficient values as low as 0.5 may be useful.

Figure 3. Correlation coefficient between optimal abundance estimates and corresponding ground truth abundances. (**a**) Mega1 results; (**b**) Mega4 results. Note that Mega1 contains twice as many materials as Mega4.
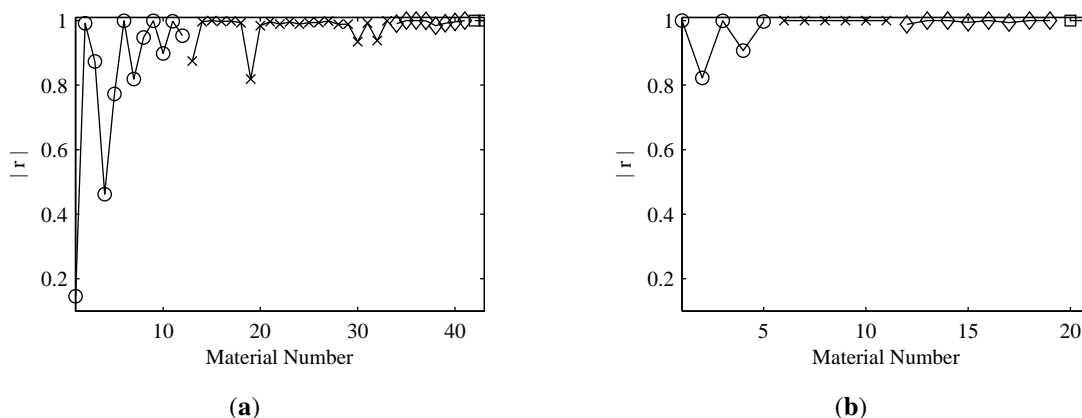

(**a**)

(**b**)

Figure 4. A comparison of material truth maps (first row (**a**–**d**)) with their maximum correlation estimates (second row (**e**–**h**)). All images come from the Mega1 scene. (a) and (e) Material 4, Siding, Cedar, Stained Dark Brown, Fair, $r = 0.4617$; (b) and (f) Material 19, Roof Shingle, Asphalt, Eclipse Sample Board, Twilight Gray, $r = 0.8185$; (c) and (g) Material 38, Tree, Norway Maple, Leaf, $r = 0.9840$; (d) and (h) Material 43, Grass, Brown and Green w/ Dirt, $r = 0.9999$.
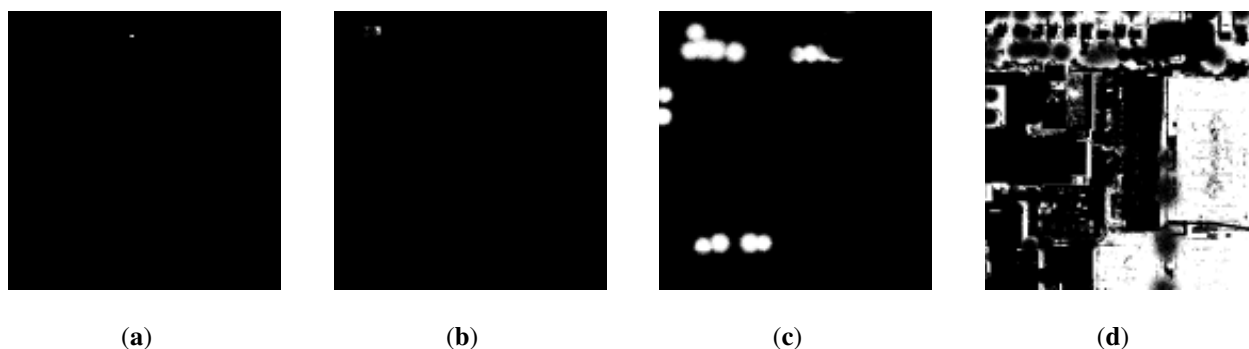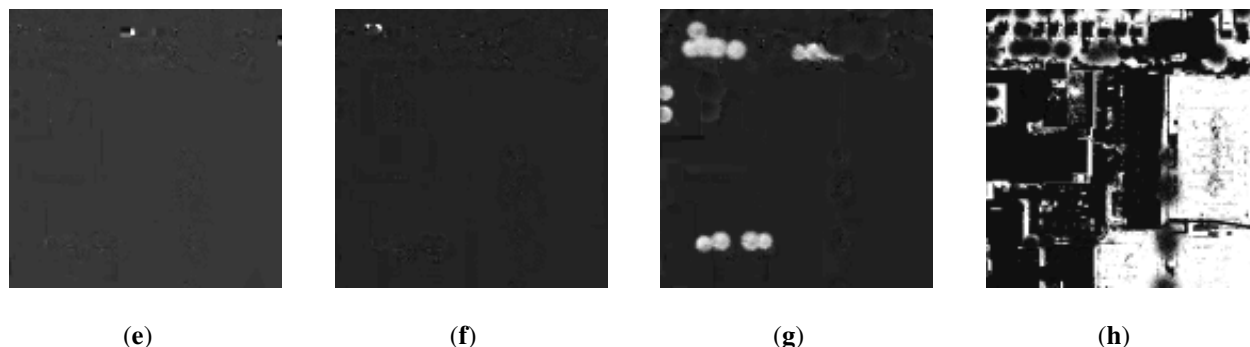

(**a**)

(**b**)

(**c**)

(**d**)

Figure 4. *Cont*.

(e)   (f)   (g)   (h)

4.2. Dimension Reduction

Because it is typically used as a preprocessing step in a variety of spectral unmixing approaches, including FastICA, an experiment was performed to examine the effect of dimension reduction on the best-case unmixing scenario. To do this, the maximum correlation abundance estimates were calculated using dimension-reduced data obtained from PCA. The same maximum correlation formulas Equations (23) and (24) were used, replacing $\mathbf{X}$ with the dimension-reduced data, $\mathbf{X}_N$, given by $\mathbf{X}_N = \mathbf{V}_N^T\mathbf{X}$, where $\mathbf{V}_N$ is the $K \times N$ whitening matrix associated with the N most energetic principal components of $\mathbf{X}$.

The results of this experiment are shown in Figure 5 and Table 1. The plots in Figure 5 demonstrate how the correlation coefficient of the optimal estimate with the ground truth decreases as the dimensionality of the data is reduced. The correlation coefficient (y-axis) in these plots is normalized by the correlation coefficient obtained when there has been no dimension reduction. The slope of each curve illustrates the contribution of individual principal components to the correlation coefficient of the optimal estimates for a specific material. It is clear from the sharp jumps in the correlation for the dense and intermediate materials that they are well described by the first several principal components. What is also clear is that there is no similar jump for the sparse and super-sparse materials. The information associated with these materials appears to be almost uniformly scattered across all of the principal components. For this reason, a relatively large number of dimensions must be retained to achieve near-optimal estimates of these materials. Table 1 underscores this conclusion, showing the average number of dimensions that must be kept to obtain 95% and 75% levels of the correlation coefficient obtained when no dimension reduction was performed.

One approach to determining the number of dimensions that should be retained when performing PCA is to keep as many dimensions as are needed to retain some percentage of the total variance in the image. Retaining 99.9% of the total variance in the Mega1 and Mega4 images requires only six and five dimensions, respectively. Based on the results in Table 1, that would allow only the dense materials to be extracted at near-optimal levels.

Figure 5. Normalized correlation coefficient of the maximum correlation estimates obtained using dimension reduced data. The first row shows the Mega1 results and the second shows the results for Mega4. (**a**) Mega1 super-sparse materials; (**b**) Mega1 sparse materials; (**c**) Mega1 intermediate materials; (**d**) Mega1 dense materials; (**e**) Mega4 super-sparse materials; (**f**) Mega4 sparse materials; (**g**) Mega4 intermediate materials; (**h**) Mega4 dense materials.
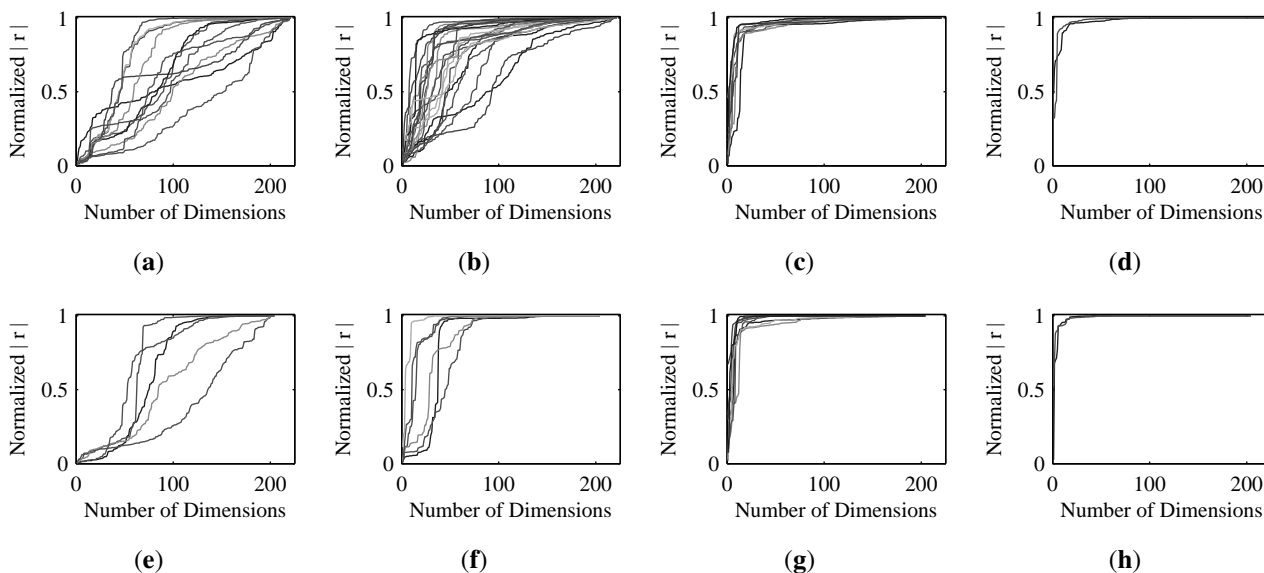

(a)   (b)   (c)   (d)

(e)   (f)   (g)   (h)

Table 1. Number of dimensions necessary to obtain 95% and 75% levels of optimal correlation, by material classification.

	Mega1		Mega4	
	95%	75%	95%	75%
super-sparse	150	110	138	102
sparse	116	53	46	29
intermediate	36	10	23	8
dense	16	6	12	4

4.3. Orthogonalization

A second experiment examined the effect of constraining the unmixing vectors to be orthogonal. Because the PCA and whitening step decorrelates the observed data, it is expected that the unmixing vectors for the whitened data should be orthogonal. In the FastICA implementation, this constraint is enforced on the unmixing vectors at the end of each iteration of the cost function optimization.

To apply the orthogonality constraint to the optimal unmixing vectors requires a minor modification to the orthogonalization formula, since the optimal vectors were not calculated using whitened data. When the data are not whitened, the formulas for deflationary orthogonalization Equations (14) and (15) become:

$$\widetilde{\mathbf{b}}_i = \mathbf{b}_i - \sum_{j=1}^{i-1} (\mathbf{b}_i^T \mathbf{C_x} \widetilde{\mathbf{b}}_j) \widetilde{\mathbf{b}}_j \tag{25}$$

and

$$\widetilde{\mathbf{b}}_i = \frac{\widetilde{\mathbf{b}}_i}{\sqrt{\widetilde{\mathbf{b}}_i^T \mathbf{C_x} \widetilde{\mathbf{b}}_i}} \tag{26}$$

respectively, where $\mathbf{C_x}$ is the covariance matrix of $\mathbf{X}$. The symmetric orthogonalization formula (16) changes in a similar way:

$$\widetilde{\mathbf{B}} = \left(\mathbf{B} \mathbf{C_x} \mathbf{B}^T \right)^{-1/2} \mathbf{B} \tag{27}$$

These changes result from the fact that orthogonality of the unmixing vectors of whitened data is equivalent to $\mathbf{B} \mathbf{C_x} \mathbf{B}^T = \mathbf{I}$, where $\mathbf{B}$ contains the unmixing vectors of the unwhitened data.

Figure 6. Normalized correlation coefficient of estimates obtained by orthogonalizing the optimal unmixing vectors for Mega1 (first row) and Mega4 (second row). (**a**) Mega1 symmetric orthogonalization; (**b**) Mega1 deflationary orthogonalization (sparse to dense); (**c**) Mega1 deflationary orthogonalization (dense to sparse); (**d**) Mega4 symmetric orthogonalization; (**e**) Mega4 deflationary orthogonalization (sparse to dense); (**f**) Mega4 deflationary orthogonalization (dense to sparse).
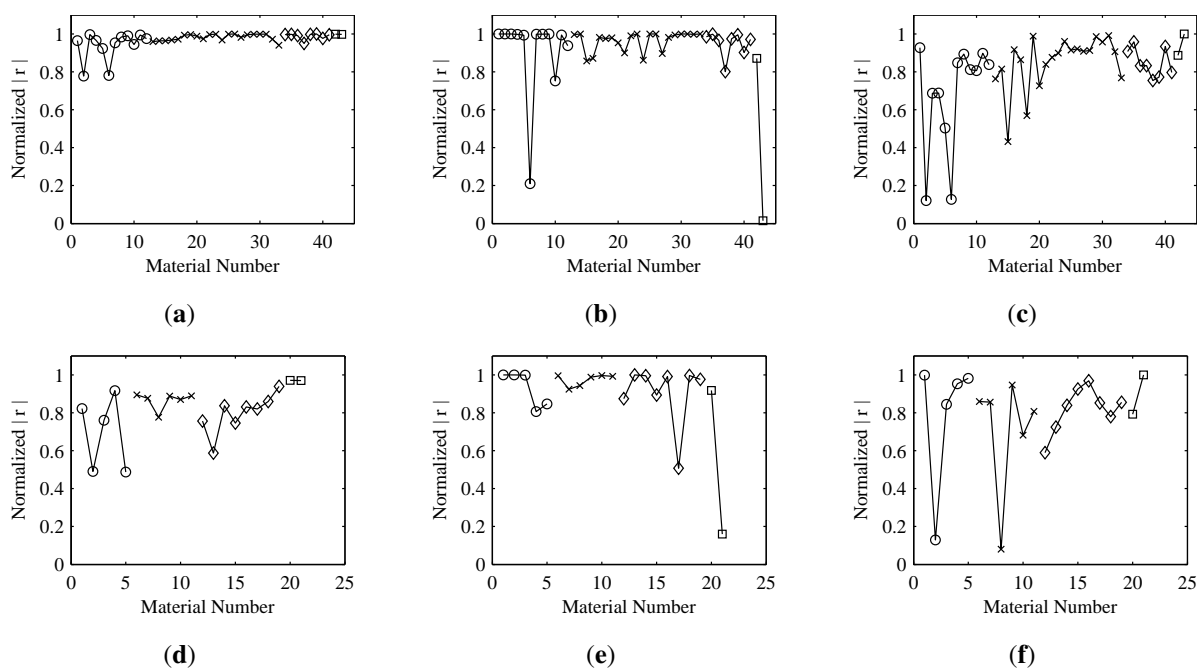
The optimal unmixing vectors calculated by Equation (23) were forced to be orthogonal using the formulas above. Abundance estimates were then calculated from the orthogonalized vectors. The effect on the correlation coefficient of the estimates due to orthogonalization is shown in Figure 6. Because the deflationary orthogonalization approach is sequential, the ordering of the vectors matters. The deflation was performed in both ascending and descending material order (most sparse to most dense and *vice versa*). As would be expected, the results show that better estimates are obtained for those materials that are used earlier in the deflation process. Thus, to obtain better estimates of a

material, it would be desirable for the cost function optimization algorithm to extract the unmixing vector corresponding to that material before any others. The results also show that deflating the estimates for more sparse materials first has less of an effect on the more dense materials than deflating in the opposite order. The symmetric approach is something of a compromise, balancing the negative effects of the orthogonalization across all of the materials.

Figure 7. An image representation of the correlation coefficient of the optimal unmixing vectors for Mega1. Off-diagonal bright spots indicate correlation between the vectors, despite whitening. Notice the dark area in the bottom-right of the image due to the negative correlation between the dense materials.
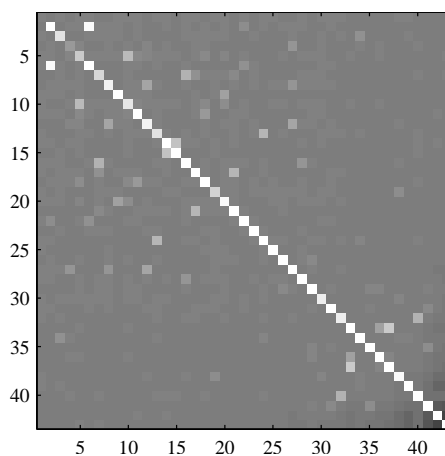
The results show that, in most cases, orthogonalization does not cause significant degradation of the estimates. This is true even in the presence of additive noise. There are a few exceptions, however, where the degradation is noticeable. Obvious examples of this are materials 2 and 6 in the Mega1 results. When symmetric orthogonalization is used, both show an appreciable decrease from the optimal correlation. When the ascending deflationary approach is used, material 2 is unaffected, but material 6 shows significant loss. Both are affected when the deflation is performed in descending material order. This behavior implies that there must be some information shared between the two materials. Thus, if material 2 is extracted first, it leads to a degradation when extracting material 6 and vice versa. Both experience degradation when the symmetric approach is used. This pattern can be explained by looking at an image representation of the matrix $\mathbf{BC_x B}^T$, shown in Figure 7. If the materials were truly uncorrelated when whitened, then the image would be that of a diagonal matrix with white pixels on the diagonal and the remainder black. However, the off-diagonal bright spots in Figure 7 indicate correlation between the optimal unmixing vectors, even when the data are whitened. In the case of this image, material 2 only shows up in one pixel and material 6 only shows up in two pixels, one of which is shared with the lone material 2 pixel. Wherever there is a drop in correlation due to orthogonalization, similar results are found, *i.e.*, a more sparse material shows up entirely in a subset of the pixels containing a more dense material. In these cases the additivity constraint in Equation(19) leads to stronger correlation than for those materials that share pixels with many different materials. Therefore, while it is true that as the number of endmembers in the data increases, the statistical dependence among

sources decreases and ICA performs better [12], co-located materials with limited spatial coverage will still be poorly estimated.

4.4. FastICA Performance

As a final experiment, FastICA was used to generate abundance maps for the Mega1 and Mega4 data. Each of the three cost functions in Equations (8)–(10) was considered, as well as both symmetric and deflationary orthogonalization. The number of components was left to be determined by the algorithm. In each case the algorithm was initialized with a random matrix. As noted earlier, there is a scale ambiguity associated with the FastICA outputs. To be useful in abundance quantification these outputs should fall in the range $[0, 1]$. The best method of rescaling the outputs is not explored in this paper. Instead, a metric that is invariant to scale is used to assess the results. A simple approach to rescaling ICA-produced abundances is suggested in [15].

Because the number of components was not specified, more independent components were generated than there are materials in the scene. For this experiment the normalized correlation coefficient of every independent component with every material ground truth was calculated, and the maximum was retained for each material. These results are shown in Figure 8. Average performance across material classifications is shown in Table 2. Generally, it appears that no single cost function or orthogonalization approach is vastly superior to any other. For extracting dense materials, it appears that the pow3 cost function should be avoided and that deflationary orthogonalization usually outperforms symmetric. This might imply that dense materials tend to be found earlier than materials from other categories. For sparse and super-sparse materials only the gauss cost function combined with symmetric orthogonalization gave consistently poor results.

Figure 8. Normalized correlation coefficient of estimates obtained using FastICA for Mega1 (first row) and Mega4 (second row). The deflationary orthogonalization results are shown with a solid line, symmetric orthogonalization with a dotted line. (**a**) Mega1 results using cost function "pow3" described by Equations (8) and (11); (**b**) Mega1 results using cost function "tanh" described by Equations (9) and (12); (**c**) Mega1 results using cost function "gauss" described by Equations (10) and (13); (**d**) Mega4 results using cost function "pow3" described by Equations (8) and (11); (**e**) Mega4 results using cost function "tanh" described by Equations (9) and (12); (**f**) Mega4 results using cost function "gauss" described by Equations (10) and (13).
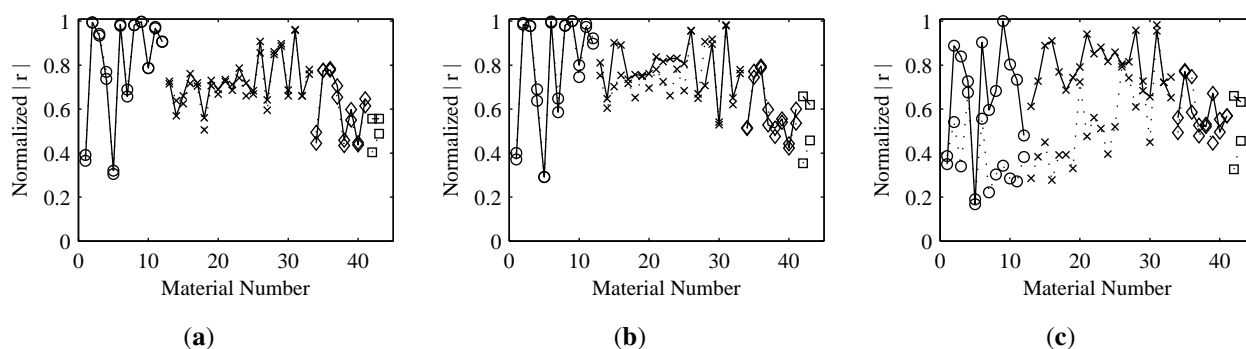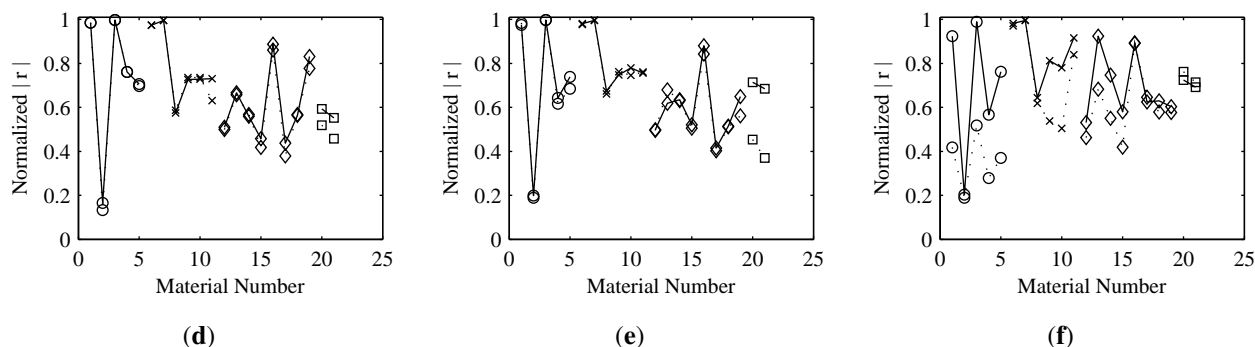

(**a**)   (**b**)   (**c**)

Figure 8. *Cont.*

(**d**)   (**e**)   (**f**)

Table 2. Average normalized correlation coefficient FastICA based on material classification.

	pow3		tanh		gauss	
	defl	**symm**	**defl**	**symm**	**defl**	**symm**
Mega1, super-sparse	0.8027	0.8095	0.7916	0.8023	0.6808	0.3745
Mega4, super-sparse	0.7229	0.7146	0.7106	0.6943	0.6893	0.3547
Mega1, sparse	0.7303	0.7191	0.7871	0.7435	0.7966	0.5463
Mega4, sparse	0.7882	0.7765	0.8222	0.8174	0.8551	0.7446
Mega1, intermediate	0.6055	0.5945	0.5841	0.5837	0.5876	0.5745
Mega4, intermediate	0.6140	0.5909	0.5915	0.5787	0.6912	0.6012
Mega1, dense	0.5566	0.4660	0.6390	0.4060	0.6467	0.3910
Mega4, dense	0.5729	0.4881	0.6996	0.4120	0.7081	0.7384
Mega1, all	0.7192	0.7084	0.7437	0.7145	0.7184	0.4964
Mega4, all	0.6858	0.6636	0.6960	0.6586	0.7392	0.5965

Three ground truth images as well as the independent components most strongly correlated with them are shown in Figure 9. The correlation coefficients of the truth maps and estimates are, from left to right, $|r| = 0.5054$, $|r| = 0.7443$, and $|r| = 0.8853$. These values are not normalized by the best-case coefficients. The images give an idea of the quality of the unmixed data for a range of correlation coefficients.

The images in the first row of Figure 10 show two independent components obtained using FastICA to unmix the Mega1 data. They illustrate two interesting features that have been frequently noticed in the FastICA output. First, there is an intensity gradient across the horizontal dimension of the images. The DIRSIG tool uses a push broom sensor model to generate these data with the sensor moving from bottom to top. Based on this, it appears that FastICA is extracting information that is associated with the view angle of the sensor and/or path length. This is interesting behavior considering that FastICA does not consider spatial organization of pixels. Further examination of the associated endmember and atmosphere data is needed to determine exactly what is being highlighted in this gradient. ICA has been shown to extract components corresponding to solar angle effect [5]. This may be a comparable result.

Figure 9. Material truth maps from Mega1 (first row) and the independent components most correlated with them (second row). (**a**) Tree, Norway Maple, Leaf truth map; (**b**) Sheet Metal, White, Fair truth map; (**c**) Brick, Brampton Brick, Old School, Brown, truth map; (**d**) Tree, Norway Maple, Leaf best estimate, $|r| = 0.5054$; (**e**) Sheet Metal, White, Fair best estimate, $|r| = 0.7443$; (**f**) Brick, Brampton Brick, Old School, Brown best estimate, $|r| = 0.8853$.

(**a**)　　　　　　　　　　　(**b**)　　　　　　　　　　　(**c**)

(**d**)　　　　　　　　　　　(**e**)　　　　　　　　　　　(**f**)
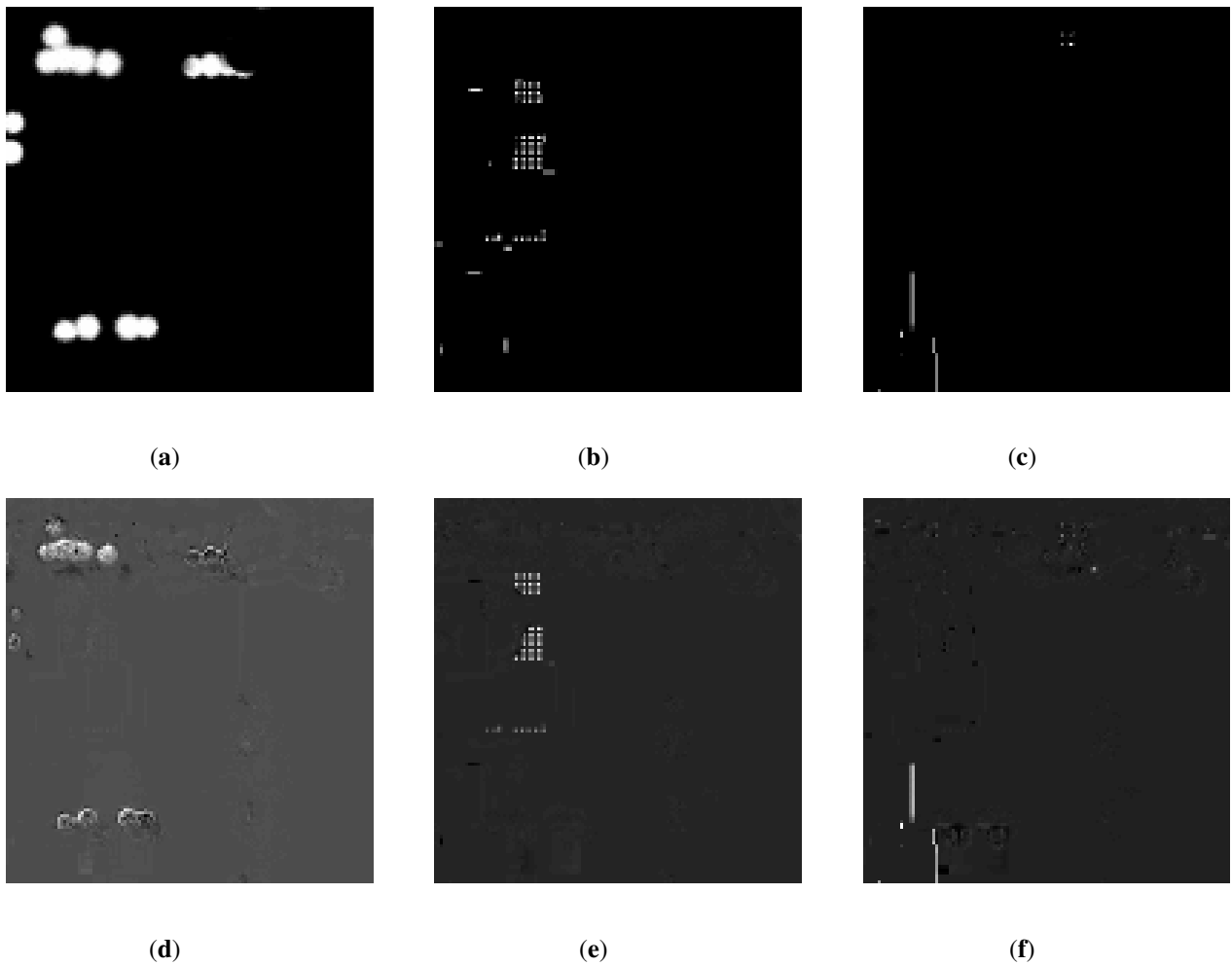
The second observation is that these two components are both strongly correlated to the same material. The correlation coefficient of the first with the truth map is $|r| = 0.6011$. For the second, $|r| = 0.4673$. A linear combination of the two can be used to produce the image in Figure 10(d) for which $|r| = 0.7606$. This splitting of a single material into two components seems to occur frequently. The clustering of independent components of hyperspectral images has been examined [42], but an attempt to automate and optimize the process based on results from synthetic data remains a future research effort.

Figure 10. Two independent components, (**a**) and (**b**), that are strongly correlated to the same truth map, shown in (**c**); A linear combination of the two, (**d**), provides an improvement to the correlation coefficient.

(**a**)

(**b**)

(**c**)

(**d**)
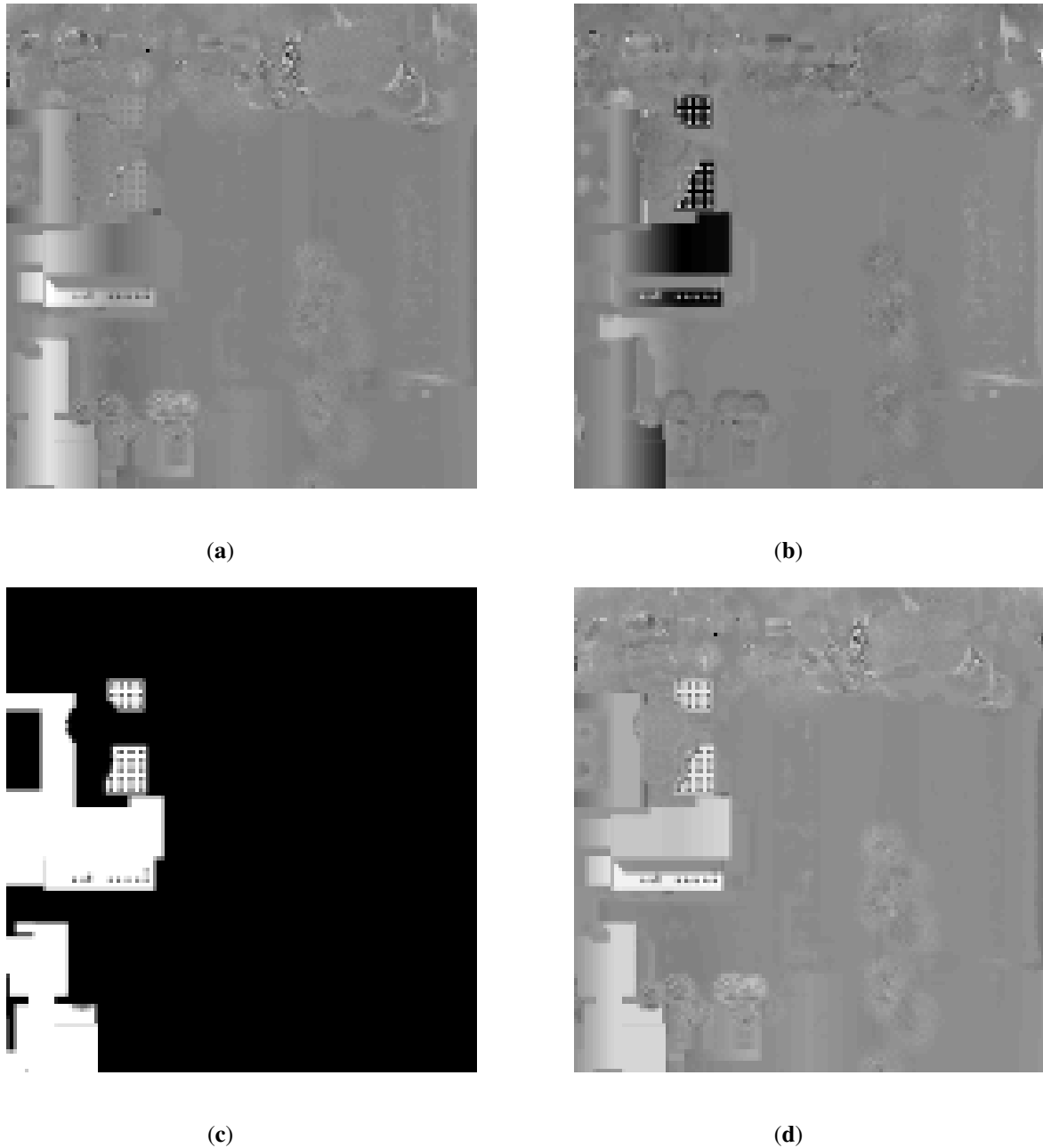
5. Conclusions

In this paper, the utility of realistic but synthetic data to assess spectral unmixing approaches was demonstrated using two hyperspectral images generated by DIRSIG. Three distinct but related experiments were performed to demonstrate this utility. The first experiment quantified the effect of dimension reduction using PCA and demonstrated that to achieve near-optimal results, more dimensions need to be retained than would be expected based on an analysis of eigenvalues. The number of additional dimensions that are necessary depends on the spatial distribution of the materials of interest, but is

approximately an order of magnitude greater for sparse materials. The second experiment considered the impact of orthogonalization, which was found to reduce the correlation coefficient by less than 10% except in the case where sparsely distributed materials were found to be consistently co-located. The method of orthogonalization as well as the order of material extraction determines the severity of the effect for those materials. The final experiment showed that FastICA is effective at unmixing some, but not all, materials. This complementary experimental approach allowed for the identification of a splitting behavior in which FastICA produces multiple outputs containing distinct pieces of a common material. It was shown that these outputs can be merged in a way that produces improved results, increasing the correlation coefficient by 30%–60%. An approach to automatically identify and merge these outputs is an area of future research.

References

1. Kim, M.S.; Lefcourt, A.M.; Chao, K.; Chen, Y.R.; Kim, I.; Chan, D.E. Multispectral detection of fecal contamination on apples based on hyperspectral imagery: Part I. Application of visible and near-infrared reflectance imaging. *Trans. Am. Soc. Agric. Eng.* **2002**, *45*, 2027–2038.
2. Moran, M.; Inoue, Y.; Barnes, E. Opportunities and limitations for image-based remote sensing in precision crop management. *Remote Sens. Environ.* **1997**, *61*, 319–346.
3. Kruse, F.A.; Boardman, J.W.; Huntington, J.F. Comparison of airborne hyperspectral data and EO-1 Hyperion for mineral mapping. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1388–1400.
4. Schmid, T.; Koch, M.; Gumuzzio, J.; Mather, P.M. A spectral library for a semi-arid wetland and its application to studies of wetland degradation using hyperspectral and multispectral data. *Int. J. Remote Sens.* **2004**, *25*, 2485–2496.
5. Moussaoui, S.; Hauksdottir, H.; Schmidt, F.; Jutten, C.; Chanussot, J.; Brie, D.; Doute, S.; Benediktsson, J. On the decomposition of Mars hyperspectral data by ICA and Bayesian positive source separation. *Neurocomputing* **2008**, *71*, 2194–2208.
6. Manolakis, D.; Marden, D.; Shaw, G.A. Hyperspectral image processing for automatic target detection applications. *Lincoln Lab. J.* **2003**, *14*, 79–116.
7. Keshava, N.; Kerekes, J.P.; Manolakis, D.G.; Shaw, G.A. Algorithm taxonomy for hyperspectral unmixing. *Proc. SPIE* **2000**, *4049*, 42–63.
8. Bioucas-Dias, J.; Plaza, A.; Dobigeon, N.; Parente, M.; Du, Q.; Gader, P.; Chanussot, J. Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5*, 354–379.
9. Cichocki, A.; Ichi Amari, S. *Adaptive Blind Signal and Image Processing*, 1st ed.; Wiley: Chichester, UK, 2002.
10. Bayliss, J.; Gualtieri, J.A.; Cromp, R.F. Analyzing hyperspectral data with independent component analysis. *Proc. SPIE* **1997**, *3240*, 133–143.
11. Zhang, X.; Chen, C.H. New independent component analysis method using higher order statistics with application to remote sensing images. *Opt. Eng.* **2002**, *41*, 1717–1728.
12. Nascimento, J.; Dias, J. Does independent component analysis play a role in unmixing hyperspectral data? *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 175–187.

13. Tu, T. Unsupervised signature extraction and separation in hyperspectral images: A noise-adjusted fast independent component analysis approach. *Opt. Eng.* **2000**, *39*, 897–906.

14. Foy, B.R.; Theiler, J. Scene analysis and detection in thermal infrared remote sensing using independent component analysis. *Proc. SPIE* **2004**, *5439*, 131–139.

15. Wang, J.; Chang, C. Applications of independent component analysis in endmember extraction and abundance quantification for hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2601–2616.

16. Shah, C.A.; Arora, M.K.; Robila, S.A.; Varshney, P.K. ICA Mixture Model Based Unsupervised Classification of Hyperspectral Imagery. In Proceedings of the 31st Applied Imagery Pattern Recognition Workshop, Washington DC, USA, 16–18 October 2002; pp. 29–35.

17. Chiang, S.; Chang, C.; Ginsberg, I. Unsupervised Hyperspectral Image Analysis Using Independent Component Analysis. In Proceedings of the Geoscience and Remote Sensing Symposium, Honolulu, USA, 24–28 July 2000; Volume 7, pp. 3136–3138.

18. Hyvarinen, A.; Karhunen, J.; Oja, E. *Independent Component Analysis*, 1st ed.; Wiley-Interscience: New York, NY, USA, 2001.

19. Plaza, A.; Martinez, P.; Perez, R.; Plaza, J. Spatial/spectral endmember extraction by multidimensional morphological operations. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 2025–2041.

20. Plaza, A.; Martinez, P.; Perez, R.; Plaza, J. A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 650–663.

21. Plaza, A.; Chang, C.I. Impact of initialization on design of endmember extraction algorithms. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 3397–3407.

22. Ifarraguerri, A.; Chang, C.I. Multispectral and hyperspectral image analysis with convex cones. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 756–770.

23. Winter, M.E. N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data. *Proc. SPIE* **1999**, *3753*, 266–275.

24. Nascimento, J.; Dias, J. Vertex component analysis: A fast algorithm to unmix hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 898–910.

25. Jung, T.P.; Makeig, S.; Humphries, C.; Lee, T.W.; Mckeown, M.J.; Iragui, V.; Sejnowski, T.J. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology* **2000**, *37*, 163–178.

26. Cichocki, A.; Shishkin, S.L.; Musha, T.; Leonowicz, Z.; Asada, T.; Kurachi, T. EEG filtering based on blind source separation (BSS) for early detection of Alzheimer's disease. *Clin. Neurophysiol.* **2005**, *116*, 729–737.

27. Joutsensalo, J.; Ristaniemi, T. Learning Algorithms for Blind Multiuser Detection in CDMA Downlink. In Proceedings of The Ninth IEEE International Symposium on the Personal, Indoor and Mobile Radio Communications, Boston, MA, USA, 8–11 September 1998; Volume 3, pp. 1040–1044.

28. Ristaniemi, T.; Joutsensalo, J. On the Performance of Blind Source Separation in CDMA Downlink. In Proceedings of the International Workshop on Independent Component Analysis and Signal Separation (ICA'99), Aussois, France, 11–15 January 1999; pp. 437–441.

29. Back, A.; Weigend, A. A first application of independent component analysis to extracting structure from stock returns. *Int. J. Neural Syst.* **1997**, 8, 473–484.

30. Cha, S.; Chan, L. Applying Independent Component Analysis to Factor Model in Finance. In Proceedings of Intelligent Data Engineering and Automated Learning IDEAL 2000. Data Mining, Financial Engineering, and Intelligent Agents, Hong Kong, 13–15 December 2000; pp. 161–173.

31. Papoulis, A.; Pillai, S. *Probability, Random Variables and Stochastic Processes*, 4th ed.; McGraw Hill Higher Education: New York, NY, USA, 2002.

32. Yang, H.H.; Amari, S. Adaptive online learning algorithms for blind separation: Maximum entropy and minimum mutual information. *Neural Comput.* **1997**, *9*, 1457–1482.

33. Cardoso, J.F.; Souloumiac, A. Blind beamforming for non-Gaussian signals. *IEE Proc. F Radar Signal Process.* **1993**, *140*, 362–370.

34. Pajunen, P.; Hyvarinen, A.; Karhunen, J. Nonlinear Blind Source Separation by Self-Organizing Maps. In Proceedings of the International Conference on Neural Information Processing, Hong Kong, 24–27 September 1996; pp. 1207–1210.

35. Pajunen, P.; Karhunen, J. A maximum likelihood approach to nonlinear blind source separation. *Lect. Note. Comput. Sci.* **1997**, 1327, 541–546.

36. Keshava, N.; Mustard, J. Spectral unmixing. *IEEE Signal Process. Mag.* **2002**, *19*, 44–57.

37. Singer, R.B.; McCord, T.B. Mars: Large Scale Mixing of Bright and Dark Surface Materials and Implications for Analysis of Spectral Reflectance. In Proceedings of the 10th Lunar and Planetary Science Conference, Houston, TX, USA, 19–23 March 1979; pp. 1835–1848.

38. Rochester Institute of Technology. *The DIRSIG User's Manual*; RIT: Rochester, NY, USA, 2006.

39. Berk, A.; Bernstein, L.; Anderson, G.; Acharya, P.; Robertson, D.; Chetwynd, J.; Adler-Golden, S. MODTRAN cloud and multiple scattering upgrades with application to AVIRIS. *Remote Sens. Environ.* **1998**, *65*, 367–375.

40. Ientilucci, E.J.; Brown, S.D. Advances in Wide-Area Hyperspectral Image Simulation. *Proc. SPIE* **2003**, *5075*, 110–121.

41. Green, R.; Eastwood, M.; Sarture, C.; Chrien, T.; Aronsson, M.; Chippendale, B.; Faust, J.; Pavri, B.; Chovit, C.; Solis, M.; *et al*. Imaging spectroscopy and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). *Remote Sens. Environ.* **1998**, *65*, 227–248.

42. Rajan, S. Unmixing of Hyperspectral Data Using Independent Component Analysis. M.Sc. Thesis, Utah State University, Logan, UT, USA, 2005.

Appendix

Material Lists for Synthetic Test Images

Table A1. MegaScene 1, Tile 4 Test Image Materials.

ID	Material Name	Total Pixels Present	Total Pure Pixels	Fractional Area
	Super-Sparse Materials (indicated by ○ in plots)			
1	Sheet Metal, Maroon, Shiny, Fair	1	0	0.016
2	Tree, Maple, Trunk	1	0	0.016
3	Tree, Red Maple, Leaf	1	0	0.016
4	Tree, Dogwood, Trunk	2	0	0.031
5	Brick, Old Carolina Brick Company, Charlestowne	2	0	0.266
	Sparse Materials (×)			
6	Sheet Metal, Black, Shiny, Dirty	21	0	1.969
7	Brick, Hampton Brick, Sandmist	36	0	2.875
8	Concrete, Cinder Blocks, Textured	68	17	40.344
9	Brick, Mixed Tan and Caramel Colors	82	0	10.234
10	Brick, Old Carolina Brick Co., Savannah Gray	102	10	45.578
11	Sheet Metal, White, Fair	183	0	55.734
	Intermediate Materials (◇)			
12	Tree, Silver Maple, Leaf	206	117	165.953
13	Sheet Metal, Tan, Shiny, Fair	276	187	229.172
14	Building Roof, Painted Metal, Gray, Weathered	333	129	224.063
15	Tree, Dogwood, Leaf	370	27	175.938
16	Sheet Metal, Gray, Shiny, Dusty	660	9	101.656
17	Tree, Norway Maple, Leaf	667	182	401.734
18	Siding, Vinyl, Tan, Fair	1,115	771	938.859
19	Roof, Gravel, Gray	1,194	767	998.188
	Dense Materials (□)			
20	Grass, Brown and Green w/ Dirt	8,718	2,534	5,864.234
21	Asphalt, Black, New	10,880	4,726	7,127.125

Table A2. MegaScene 1, Tile 1 Test Image Materials.

ID	Material Name	Total Pixels Present	Total Pure Pixels	Fractional Area
	Super-Sparse Materials (indicated by ○ in plots)			
1	Siding, Mineral, Painted, Dark Green	1	0	0.016
2	Siding, Wood, Painted Off White, Fair	1	0	0.078
3	Tree, Black Oak, Bark	2	0	0.031
4	Siding, Cedar, Stained Dark Brown, Fair	2	0	0.078
5	Siding, Wood, Painted White, New, Rough	2	0	0.094
6	Brick, Old Carolina Brick Company, Charlestowne	2	0	0.453
7	Glass	3	0	0.047
8	Brick, Brampton Brick, Old School, Red	4	0	0.313
9	Siding, Vinyl, Off White, Fair	4	0	0.594
10	Roadway Surfaces, Sidewalk, Brick, Sealed, Mixed Color	4	0	0.813
11	Vinyl, Vision Pro Sample Board, Blue D-4	7	0	0.719
12	Roof Shingle, Asphalt, Mix Brown, Good	7	0	0.781
	Sparse Materials (×)			
13	Stone Siding, Apple Ridge, Buckingham Fieldstone	9	0	1.156
14	Sheet Metal, Gray, Shiny, Dusty	11	0	1.078
15	Swimming Pool (Lining and Water)	12	0	5.375
16	Siding, Wood, Planks, Brown	13	0	2.359
17	Siding, Wood, Painted Tan, Fair	15	0	2.313

Table A2. *Cont.*

ID	Material Name	Total Pixels Present	Total Pure Pixels	Fractional Area
18	Roof Shingle, Asphalt, Harmony Sample Board, Cove Gray	26	5	16.281
19	Roof Shingle, Asphalt, Eclipse Sample Board, Twilight Gray	27	0	1.859
20	Roof Shingle, Asphalt, Black, Weathered	29	16	20.516
21	Roof Shingle, Asphalt, Black, Fair	30	6	17.453
22	Roof Shingle, Asphalt, Eclipse Sample Board, Shadow Black	30	12	20.328
23	Roof Shingle, Asphalt, Dark Light, Fair	30	12	20.813
24	Roof Shingle, Asphalt, Brown and Red Blend, Fair	31	0	9.984
25	Roof Shingle, Asphalt, Eclipse Sample Board, Forest Green	35	15	24.281
26	Roof Shingle, Asphalt, Brown, Black, New	35	10	24.719
27	Brick, Siding, Mix Brown, Fair	44	13	33.750
28	Roof Shingle, Asphalt, Harmony Sample Board, Sequoia Tile	64	16	40.953
29	Brick, Brampton Brick, Old School, Brown	76	0	13.672
30	Tree, Dogwood, Leaf	77	3	30.797
31	Brick, KF Plymouth Blend, Red Brick	84	0	14.563
32	Tree, Maple, Trunk	140	0	5.406
33	Tennis Court, Playing Surface, White Line	194	0	37.625
Intermediate Materials (◊)				
34	Tree, Black Oak, Leaf	212	14	77.469
35	Sheet Metal, White, Fair	222	0	58.188
36	Tennis Court, Playing Surface, Red	250	67	155.688
37	Tennis Court, Playing Surface, Green	262	59	175.625
38	Tree, Norway Maple, Leaf	1,005	196	632.422
39	Tree, Silver Maple, Leaf	1,299	717	1,013.297
40	Tree, Red Maple, Leaf	1,360	7	611.625
41	Roof, Gravel, Gray	2,373	1,845	2,176.047
Dense Materials (□)				
42	Asphalt, Black, New	8,198	2,928	4,975.422
43	Grass, Brown and Green w/ Dirt	9,275	3,124	6,158.922