Incorporating spatial information in spectral unmixing: A review

Chen Shi a, b, Le Wang a, b, *

a The State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS), Wuhan University, Wuhan, Hubei 430079, P. R. China
b Department of Geography, University at Buffalo, The State University of New York, 105 Wilkeson Quad., Buffalo, NY 14261, USA

Spectral unmixing is the process of decomposing the spectral signature of a mixed pixel into a set of endmembers and their corresponding abundances. Endmembers are spectra of the pure materials present in the image and abundances at each pixel represent the percentage of each endmember that is present in the pixel. Many spectral unmixing techniques treat a pixel as independent of its neighbors, therefore, only spectral characteristics of the image are used to address the spectral unmixing problem. However, a number of recent studies have found that spatial autocorrelation provides useful information for spectral unmixing. Combining spatial information with its spectral counterpart can lead to improvements in the unmixing results. In this paper, the unmixing methods that incorporate spatial information are termed spatial spectral unmixing, whereas those exploiting only spectral information are referred to as spectral-only unmixing. We summarize the available spatial spectral unmixing methods according to the following three categories: 1) endmember extraction, 2) selection of endmember combinations, and 3) abundance estimation. An experiment-based comparison between representative spatial spectral and spectral-only unmixing methods is also presented in order to demonstrate the advantages of spatial spectral methods. Furthermore, considerations and suggestions of the incorporation of spatial information are provided. With this review, we hope to bring spatial spectral unmixing to the attention of the remote sensing community and stimulate new research initiatives to integrate both spatial and spectral information for unmixing purposes.

© 2014 Elsevier Inc. All rights reserved.
1. Introduction

In the last decade, a large volume of multispectral and hyperspectral remote sensing data have been acquired by airborne and satellite sensors. Such data have been widely applied to environmental and urban studies (Adams et al., 1995; Atkinson, Cutler, & Lewis, 1997; Bannari, Pacheco, Staenz, McNairn, & Omari, 2006; Carpenter, Gopal, Maconber, Martens, & Woodcock, 1999; Delfries, Hansen, & Townshend, 2000; Drake, MacKin, & Settle, 1999; Hochberg & Atkinson, 2003; Kruse et al., 1993; Lobell & Asner, 2004; Lu, Moran, & Batistella, 2003; McGwire, Minor, & Fenstermaker, 2000; Painter, Roberts, Green, & Dozier, 1998; Quarnby et al., 1992; Small, 2001; Wu & Murray, 2003). Nevertheless, one of the pressing challenges confronting remotely sensed data analysis is the presence of mixed pixels that are composed of a combination of distinct materials (Keshava & Mustard, 2002; Quintino, Fernandez-Manso, Shimabukuro, & Pereira, 2012). If the spatial resolution of a sensor is coarser than the scale of spatial heterogeneity of the ground surface, a mixture of disparate substances is inevitably contained in a pixel. Detailed analysis of the pure materials at the subpixel level can be achieved by using the technique called spectral unmixing. In spectral unmixing the spectral signatures of the constituent substances present in a mixed pixel are referred to as endmembers, and the fractional area coverage of each endmember in a pixel is called its abundance (Bioucas-Dias et al., 2012). As such, spectral unmixing is the process of decomposing the acquired spectrum of a mixed pixel into a set of endmembers and their corresponding fractional abundances (Keshava & Mustard, 2002). It should be noted that two alternative terms describing this process are Spectral Mixture Analysis (Gillespie, 1992; Sabol, Adams, & Smith, 1992) and Spectral Mixture Modeling (Adams, Smith, & Johnson, 1986). In this paper, we use the term spectral unmixing to represent the same general technique.

A remotely sensed image is conceptually a two-dimensional, pictorial representation of the ground surface, and pixels in the image display spectral radiance or reflectance measured by an airborne or satellite sensor. As such, both spatial and spectral attributes are inherently contained in an image. Although spectral unmixing methods have been applied to analyze remotely sensed images for more than two decades (Adams et al., 1986), many methods only exploit the spectral information, but do not address the inherent spatial arrangement of pixels. Early attempts of incorporating spatial information in spectral unmixing date back over ten years (Plaza, Martinez, Perez, & Plaza, 2002; Rand & Keenan, 2001; Roessner, Segl, Heiden, & Kaufmann, 2001; Van der Meer, 1999) and many ensuing studies have shown its great potential for refining accuracy (Plaza, Du, Bioucas-Dias, Jia, & Kruse, 2011). For simplicity we will use the term spatial spectral unmixing in the following sections to refer to those unmixing methods that incorporate spatial information as opposed to those spectral-only unmixing.

In the area of spatial spectral unmixing three directions of research have been proposed. The first is endmember extraction for which spatial context is taken into account. The earliest work appeared in Plaza et al. (2002) in which morphological operators were applied to help identify the purest pixels in an image. Such pixels then served as image endmembers. The second direction is selection of endmember combinations in which spatial information is employed to determine the number and type of endmembers as well as their corresponding spectral signatures that vary across the image. Rand and Keenan (2001) and Roessner et al. (2001) conducted the earliest studies in which a different set of endmember was derived according to the specific location where the pixel to be unmixed fell. The third is abundance estimation in which spatial characteristics of residuals and abundances are utilized. The earliest attempt was made by Van der Meer (1999) who proposed a series of spatial criteria in abundance estimation with the aim of removing spatial structure that still remained in the residuals from the spectral-only unmixing. Surrounding the three aforementioned aspects of spatial spectral unmixing numerous methods have been developed (Canham, Schlamm, Ziemann, Basener, & Messinger, 2011; Castrodad et al., 2011; Deng & Wu, 2013; Eches, Dobigeon, & Tourneret, 2011; Jordache, Bioucas-Dias, & Plaza, 2012; Li & Zhang, 2011; Zare, Gader, Bchir, & Frigui, 2013). However, compared to the extent to which spatial information is incorporated in other remote sensing analysis, such as image classification, spatial spectral unmixing is still in its infancy. There has been no study that focuses exclusively on the assimilation and synthesis of the wide range of existing methods in order to understand their merits and drawbacks. More importantly, directions on further refining these methods need to be seriously scrutinized and discussed.

Therefore, the present study aims to provide a comprehensive overview of the available methods that have been developed for spatial spectral unmixing. Specifically, the research objectives can be summarized as follows:

1) review the state-of-the-art of spatial spectral unmixing in the context of endmember extraction, selection of endmember combinations, and abundance estimation;
2) demonstrate the advantages of spatial spectral unmixing by performing comparative experiments between representative spectral-only and spatial spectral methods;
3) identify key factors that need to be further considered in refining spatial spectral unmixing.

The remainder of the paper is organized as follows. Section 2 introduces the mathematical background of spectral unmixing techniques. In Section 3, the available spatial spectral unmixing methods are thoroughly reviewed. Section 4 shows an experiment-based evaluation of various spectral-only and spatial spectral unmixing methods. Section 5 discusses some important issues involved in the incorporation of spatial information and presents suggestions for future studies. Finally, a summary is provided in Section 6.

2. Background of spectral unmixing

The existing spectral unmixing methods can be broadly categorized as linear mixture models (LMMs) and nonlinear mixture models (NLMMs) according to the mathematical formulation for describing the underlying mixing process (Bioucas-Dias et al., 2012; Keshava & Mustard, 2002). LMMs generally assume that the spectral signature of a mixed pixel is represented by the weighted sum of the endmember spectra and that the weights associated with the endmembers are given by their corresponding proportional area coverage in the pixel. It has been documented that the linear mixing effect dominates the spectral signature of a mixed pixel if the constituent materials present in the pixel appear in spatially segregated patterns and the incident light interacts with a single surface component (Keshava & Mustard, 2002). Nevertheless, ground materials usually depict an intimate mixture where the light is multiply scattered between at least two components. In this case, the spectral mixing process is generally nonlinear (Borel & Gerstl, 1994; Ray & Murray, 1996; Roberts, Smith, & Adams, 1993). Although theoretically more closely resembling the physical mixing process in many real scenarios, the NLMM has not been applied as widely as its linear counterpart. The reasons can be attributed to the intrinsic complexity of the mixture modeling or the difficulties of obtaining scene parameters (Bioucas-Dias et al., 2012; Borel & Gerstl,
be selected from the image itself, or measured in the
quacies. Note that the subscripts
in spectral unmixing: the determination of the endmembers and the
set of matrices
ability of allowing the number and type of endmembers and their corre-
spectral variability of the image. Recent studies have shown the possi-
throughout the entire scene. However, it might fail to account for the
aspect of endmember determination is to select endmember combina-

where

in Eq.(1) in order to explicitly describe the spatial

denoting the number of bands. As such, 

denoting the width and length of the image, respectively, and

denoting the number of endmembers present in the pixel. Let 

denoting the fractional abundance cube and 

deas the spectral dimension.

Fig. 1 shows a schematic diagram of the spectral unmixing process. 
As shown in the figure, there are two fundamental procedures involved 
in spectral unmixing: the determination of the endmembers and the 
estimation of the abundances (Bioucas-Dias et al., 2012; Keshava & 
Mustard, 2002). First, endmember determination is employed to popu-
late the matrix 
the LMM. The spectral signatures of endmembers can be selected from the image itself, or measured in the field or laboratory (Adams, Smith, & Gillespie, 1993). Many studies have been dedicated to extracting endmembers from the image. Detailed information about the endmember extraction techniques are provided in Section 3.1. Another aspect of endmember determination is to select endmember combina-
tions for each pixel. Conventionally, a fixed set of endmembers is used 
throughout the entire scene. However, it might fail to account for the 
spectral variability of the image. Recent studies have shown the possi-
bility of allowing the number and type of endmembers and their corre-
sponding spectral signatures to vary on a per-pixel basis (Poll et al., 
1998; Somers, Asner, Titis, & Coppen, 2011; Zare & Ho, 2014). 

The constant matrix of endmember spectra 
in Eq. (1) is thus converted to a set of matrices 
adaptive to each pixel in the image. It has been 
reported that the accuracy of unmixing is largely dependent on the 
selection of the exact number of endmembers for each pixel (Sabol 
et al., 1992). Too few endmembers will force the fractions of undefined 
endmembers to be allocated to those considered for unmixing, cre-
ating fraction errors; too many endmembers will make Eq. (1) 
underdetermined, resulting in the model sensitive to instrumental 
oise, atmospheric contamination and spectral variability. The selection 
of endmember combinations is discussed extensively in Section 3.2.

Second, after the endmembers are determined for the examined 
pixel, the abundances are then estimated by inverting the equation sys-
tem of Eq. (1). The least squares method is often used to provide the 
optimal estimate of the abundance vector 
by minimizing the squared error between the observed spectrum 
and the approximated spectrum 
, i.e.,

1° in order to obtain the physically meaningful abundances, two 
constraints are often imposed on the objective function of Eq. (2) 
(Shimabukuro & Smith, 1991). First, the nonnegativity constraint en-
sures that each abundance element 
in the vector 
be nonnegative (see Eq. 3); Second, the full additivity constraint requires all abun-
dances to sum to one so as to account for the entire composition of 
a mixed pixel (see Eq. 4).

Along with the full additivity constraint, each spectral band provides a 
linear equation between the measured spectrum of a mixed pixel and the 
endmember spectra (see Eq. 1), therefore, the theoretical maximum 
number of endmembers used to decompose a pixel is equal to the 
number of spectral bands plus one (Atkinson et al., 1997; Keshava & 
Mustard, 2002). The relatively low number of spectral bands in multi-
spectral images limits detailed analysis of the constituent materials 
present in a pixel. In contrast, hyperspectral images with hundreds of 
spectral channels are more suitable to conduct spectral unmixing (Bioucas-Dias et al., 2012).

The closed-form solution of Eq. (1) can be provided if no constraint 
or only the full additivity constraint is included (Keshava & Mustard, 
2002). The optimization procedure becomes complex when the 
nonnegativity constraint is incorporated alone or jointly with the 
full additivity constraint (Heinz & Chang, 2001). Due to such mathemat-
ical difficulties, many spectral unmixing studies implement only the 
full additivity constraint and then apply the nonnegativity constraint to the 
result (Silvan-Cardenas & Wang, 2010a). In some cases, however, nei-
ther of the two constraints are imposed, as the fractions estimated by the 
unconstrained model can be used as a criterion of evaluating the 
selection of endmembers (Roberts et al., 1998). In addition to the least 
squares method, the equation system of Eq. (1) can also be solved by 
Singular Value Decomposition (Boardman, 1989; Gong, Miller, & 
Spanner, 1994) and Gram–Schmidt Orthogonalization (Shimabukuro, 
Carvalho, & Rudorff, 1997). More discussion about abundance estima-
tion can be seen in Section 3.3.

3. Incorporating spatial information in spectral unmixing

This section provides a comprehensive overview of the methods that 
incorporate spatial information in spectral unmixing. According to dif-
ferent stages in which spatial information is accounted for, we grouped 
all the prevalent methods into three categories: endmember extraction, 
selection of endmember combinations, and abundance estimation.

![Fig. 1. Schematic diagram of the spectral unmixing process.](image-url)
Table 1 lists major spatial spectral unmixing methods for each category. A detailed description of the three categories and their corresponding subcategories is provided in the remainder of this section.

3.1. Incorporating spatial information in endmember extraction

Endmembers are a collection of constituent spectra corresponding to distinct ground substances (Keshava & Mustard, 2002). The accuracy of spectral unmixing is highly dependent on the selection of appropriate, representative endmembers (Tompkins, Mustard, Pieters, & Forsyth, 1997). The spectral signatures of endmembers can be extracted directly from the remotely sensed images (referred to as image endmembers) or measured in the field or laboratory by using spectroradiometers (referred to as reference endmembers) (Adams et al., 1993). Image endmembers are obtained at the same spatial scale as the image to be analyzed, whereas reference endmembers, given its proximity to objects, are usually collected under different atmospheric condition from that of airborne or satellite imagery, and at a different spatial scale. In addition, it is very labor intensive to acquire sufficient endmember spectra for all important ground components in the field or laboratory. Therefore, it is often convenient to use image endmembers in spectral unmixing if they exist in the image and can be accurately extracted.

Conventionally, the theory of convex geometry is often utilized by endmember extraction methods, e.g., pixel purity index (PPI) (Boardman, 1993; Boardman, Kruse, & Green, 1995), orthogonal subspace projection (OSP) (Harsanyi & Chang, 1994), N-FINDR (Winter, 1999), iterative error analysis (IEA) (Neville, Staenz, Szeredi, Lefebvre, & Hauff, 1999), vertex component analysis (VCA) (Nascimento & Dias, 2005), etc. Viewing remotely sensed data as a point cloud in a multidimensional space where the number of the dimensions is identified as the number of spectral bands, these methods aim to find the vertices of a simplex fit to the convex hull of the point cloud. The pixels corresponding to these vertices are identified as endmembers. These geometrical based methods differ in the strategies adopted in searching for the vertices of the data simplex. For example, PPI repeatedly projects all pixels in an image onto a large number of randomly generated vectors in spectral space. For each pixel, the number of times it falls at the extremes of the vectors is counted. The pixels that give counts above a cutoff threshold are considered endmember candidates and are further investigated to manually select a desired number of endmembers. Different from PPI, N-FINDR searches for the set of pixels that define the simplex with the largest volume in spectral space and considers these pixels as endmembers. These two well-known spectral-only endmember extraction algorithms were selected for our comparative experiments (Section 4.3).

In spite of their broad applications, the geometrical based methods exploit the spectral property of remotely sensed images alone. However, the integration of both spatial and spectral characteristics of the data can be beneficial for endmember extraction (Parente & Plaza, 2010; Plaza, Martin, & Zortea, 2009b). Spatial information, in combination with spectral information, has been used to develop 1) endmember extraction methods and 2) preprocessing modules.

3.1.1. Endmember extraction methods

In order to address spatial aspects in endmember extraction, several endmember extraction methods have been developed. Among them, the automated morphological endmember extraction (AMEE) algorithm is one of the first attempts of incorporating spatial information in endmember extraction (Plaza et al., 2002). This method integrates both spatial and spectral properties of a remotely sensed image in a simultaneous manner by extending conventional grayscale morphology to the multi/hyperspectral imagery. The key to this extension is to define a spectral purity ordering scheme that allows for the determination of the pixels associated with the maximum and minimum spectral purity in a spatial window. The spectral purity of a pixel can be measured by calculating its distance, e.g., the spectral angle distance or SAD (Kruse et al., 1993), to the mean spectrum of the pixels lying within the spatial window. In order to alleviate its sensitivity to the size of spatial window, AMEE is designed as an iterative process with different window sizes. Running on the full data cube with no dimensionality reduction, it begins with the minimum window size. As the window moves through each pixel in a scene, the purest pixel and the most mixed pixel are respectively identified by the two morphological operations, i.e., dilation and erosion, in each spatial window. Then a morphological eccentricity index (MEI) value, defined as the distance (e.g., SAD) between the pixel of maximum purity and the pixel of minimum purity, is assigned to the

### Table 1

A summary of major spatial spectral unmixing methods.

<table>
<thead>
<tr>
<th>Category</th>
<th>Subcategory</th>
<th>Reference</th>
<th>Method in abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endmember extraction</td>
<td>Endmember extraction methods</td>
<td>Plaza et al. (2002)</td>
<td>AMEE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mei et al. (2010)</td>
<td>SPEE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rogge et al. (2007)</td>
<td>SPA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Zhang et al. (2008)</td>
<td>HEEA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Li and Zhang (2011)</td>
<td>SPP</td>
</tr>
<tr>
<td></td>
<td>Preprocessing methods</td>
<td>Zordea and Plaza (2009)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Thompson et al. (2010)</td>
<td></td>
</tr>
<tr>
<td>Selection of endmember combinations</td>
<td>Per-pixel methods</td>
<td>Zare et al. (2013)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maselli et al. (2001)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Per-field methods</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abundance estimation</td>
<td>Zare et al. (2013)</td>
<td></td>
</tr>
<tr>
<td>Spatial characteristics of residuals</td>
<td></td>
<td>Van der Meer (1999)</td>
<td>ISU</td>
</tr>
<tr>
<td>Spatial characteristics of abundances</td>
<td></td>
<td>Jia and Qian (2007)</td>
<td>SSCBSS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Song et al. (2010)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Castrodad et al. (2011)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Iordache et al. (2012)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Eches et al. (2011)</td>
<td></td>
</tr>
</tbody>
</table>
purest pixel in the window. This process is iterated with increasing window size until a predefined maximum window size is achieved. The MEI value associated with each selected pixel is accumulated as the window moves and the window size is increasing. An MEI threshold value is applied to the final grayscale MEI image to obtain endmember candidates. A region-growing procedure is then conducted to incorporate the neighboring pixels that are spectrally similar to the candidates and a final set of endmembers is the mean spectra calculated for the resulting regions. Although AMEE has shown the capacity to extract accurate endmembers compared to other endmember extraction methods (Plaza, Martinez, Perez, & Plaza, 2004), its computational time is increased significantly when the maximum window size becomes large. Another limitation of AMEE is that only one pixel per spatial window receives the MEI value.

Different from Plaza et al. (2002), Mei, He, Wang, and Feng (2010) developed a Spatial Purity based Endmember Extraction (SPEE) algorithm with an aim of detecting pure spatial neighborhoods instead of the purest pixel lying within each spatial neighborhood. The purity of spatial neighborhood is determined based on image gray-scale values or secondary features calculated from the gray-scale values, such as those derived from principal component analysis. Based on moving windows, the spatial neighborhood purity index is calculated and assigned to the center pixel. A predefined threshold is then applied to the spatial neighborhood purity values. The initial endmember candidates are obtained by averaging pixels in the derived pure spatial neighborhood. These endmember candidates will then be further refined based on spatial adjacency using a graph-based spatial refinement process. As such, an unsupervised clustering technique is used to group the spatially refined endmembers according to their spectral similarity. Finally, the mean spectra of endmember clusters are chosen for follow-up unmixing. Compared to AMEE with an increasing window size, SPEE largely improves computational efficiency as no iterative process is involved (Mei et al., 2010). Nevertheless, it should be noted that both AMEE and SPEE involve a thresholding process of gray-scale images, although it can be automated by using the well-known Otsu’s method (Otsu, 1979).

From a different perspective, several studies have attempted to incorporate spatial information in conventional spectral-only endmember extraction methods, such as PPI, OSP, IEA and VCA. The spatial spectral endmember extraction (SSEE) algorithm (Rogge et al., 2007) is essentially an improvement of PPI that utilizes spatial characteristics of the remotely sensed images. The method first partitions the entire image into square-shaped, non-overlapping subsets with the same size. Singular value decomposition (SVD) is then applied for each image subset to obtain a set of local eigenvectors, which retain most of the spectral variance. Every pixel in the entire image is projected onto a compiled vector set that contains the eigenvectors of all image subsets. The pixels that lie at either end of the eigenvectors comprise the initial endmember set. Then additional endmembers are defined in the vicinity of the initial endmember set based on spectral similarity. Finally, updated endmember spectra, centered on the endmember candidate, are derived by using a spatial averaging process. This process is repeated iteratively to reduce the within-class variance of endmember spectra. The derived endmember set is then reordered according to the spectral similarity and a final set of endmembers is obtained by a manual grouping process. The shortcomings of SSEE are, however, that many input parameters need to be fine-tuned in order to obtain accurate endmembers and a manual process is involved in this method.

Along a similar direction, Zhang, Rivard, and Rogge (2008) introduced the successive projection algorithm (SPA) with a constraint on the spatial adjacency of candidate endmember pixels. Similar to such conventional spectral-only endmember extraction methods as OSP and VCA, this method also aims to determine the vertices of the simplex that is fit to the convex hull of the remotely sensed data in the spectral domain. Nevertheless, after selecting multiple pixels as endmember candidates for one vertex, SPA averages the pixels that meet both the spectral similarity and spatial adjacency criteria so as to reduce the susceptibility to outlier pixels (Zhang et al., 2008). In spite of its high computational efficiency, the selection of two thresholds for spectral similarity and spatial adjacency lacks justification.

In addition, spatial information has been exploited in the hybrid endmember extraction algorithm (HEEA) to improve IEA (Li & Zhang, 2011). Similar to SSEE, HEEA begins by partitioning the entire image into several sub-images that are approximately equal in size. At each iteration, the pixels associated with the largest unmixing errors in each sub-image are put into a candidate endmember set sorted in the descending order according to the error. Then the candidates are investigated one by one to determine whether they are lying within a homogeneous neighborhood. The authors used the recently proposed spectral-information-divergence–spectral-angle-distance (SID–SAD) metric (Du et al., 2004) to measure the spectral similarity between the candidate endmember and its neighboring pixels. The homogeneity condition is met when there exist sufficient pixels that are spectrally similar to the endmember in a local window. The candidate endmember and its neighboring pixels that are spectrally similar are then averaged to produce the final endmember used in the subsequent iterations. It is worth noting that different weights are assigned to these neighboring pixels according to their distance to the candidate endmember, similar to the weighting strategy used by Zortea and Plaza (2009). Experiments showed that HEEA is highly robust when analyzing images with low signal-to-noise ratios (SNRs). However, the determination of an optimal window size is difficult, and it often requires an empirical optimization procedure. In addition, the method requires many parameters to be predefined.

3.1.2. Preprocessing methods

Differing from the preceding endmember extraction methods, several studies have aimed to incorporate spatial information through a preprocessing step before extracting endmembers from remotely sensed images. Due to the independent nature of preprocessing modules, such methods can be easily combined with the existing endmember extraction methods.

Zortea and Plaza (2009) proposed a spatial preprocessing (SPP) method that favors the search for image endmembers in homogeneous regions with an inherent assumption that endmembers, i.e., pure pixels, are more likely to be found in such regions. In order to achieve this goal, they presented a method to adjust the original spectral signature on a per-pixel basis according to a spatial weighting factor. The weighting factor is defined to resemble the spectral similarity of pixels within a predefined spatial neighborhood. As a result of the weighting, pixels located in spatially heterogeneous areas will engage a larger amount of adjustments than those in spatially homogeneous areas. In this manner, when the preprocessed image is input to an existing endmember extraction method, pure pixels located in spatially homogeneous areas are more likely to be identified as endmembers. It should be noted that the endmember spectra will come from the original image but not the preprocessed one. Compared to AMEE and other methods that adopt equal weights in forming spatial neighborhood, the spatial pre-processing method assigns pixels in the spatial neighborhood different weights, which are inversely proportional to their square distance to the center pixel, thus, reducing the sensitivity to the size of spatial neighborhood. However, the effectiveness of this method might suffer from a lack of homogenous regions in a scene (Zortea & Plaza, 2009).

Thompson, Mandrade, Gilmore, and Castano (2010) utilized a super-pixel representation for endmember extraction. Super-pixels are homogeneous regions in an image where contiguous pixels share similar spectra. They can be derived by applying segmentation techniques to the scene. The mean spectra of each super-pixel are then used as an input to any available endmember extraction methods. Different from spatial windows with a fixed size and shape that are commonly used in other spatial methods, super-pixels are adaptive spatial neighborhoods that are suitable for analyzing natural features with irregular
shape. Other benefits gained by this method include the reduction of the susceptibility to noise and outliers as well as the improvement of computational efficiency for subsequent endmember extraction. However, the accuracy of extracted endmembers largely depends on the effectiveness of image segmentation, for which optimal parameters are often hard to determine since they vary as the images or the following endmember extraction methods change (Thompson et al., 2010).

A similar method was proposed by Martin and Plaza (2011) that addressed a region-based spatial preprocessing (RBSPP). This method comes with two steps. First, after partitioning a scene into regions using unsupervised clustering techniques, an initial set of spectrally pure regions is detected by applying OSP to the mean spectra of these regions. Second, only pixels within the initial pure regions are considered for another round of endmember extraction. This method suffers from the same problems as the super-pixel endmember extraction (Martin & Plaza, 2011).

3.2. Incorporating spatial information in selection of endmember combinations

In conventional spectral unmixing techniques, a single fixed set of endmembers is often used to decompose each pixel in the whole scene. In spite of its simplicity, such methods suffer from two main drawbacks when processing real remotely sensed images. First, the spectral signatures of endmembers are assumed to be unique and constant over the entire study site. However, variations in endmember spectra are not uncommon due to changing environmental factors (e.g., illumination conditions) and physical differences in broadly defined ground substances (e.g., vegetation, soil). This has been reported as the endmember variability problem (Somers et al., 2011; Zare & Ho, 2014). Second, the number of endmembers for unmixing pixels remains unchanged throughout the whole image. Although a large number of underlying components might exist across the entire scene, the actual number of present endmembers contained in a single pixel is much smaller. Errors will be accrued if too many endmembers are used to unmix a pixel (Roberts et al., 1998; Sabol et al., 1992).

In order to address these two problems, a number of spectral-only unmixing methods have been developed to allow the number and type of endmembers as well as their corresponding spectral signatures to vary on a per-pixel basis (Asner & Heidbrecht, 2002; Bateson, Asner, & Weisman, 2000; Dennison & Roberts, 2003; Maselli, 1998; Petrou & Foschi, 1999; Roberts et al., 1998; Rogge, Rivard, Zhang, & Feng, 2006; Silvan-Cardenas & Wang, 2010b; Song, 2005). An optimization procedure of selecting endmember combinations for each pixel is often involved in these methods. For example, Roberts et al. (1998) used a maximal covering criterion to select the smallest subset of endmembers that could model the largest area. Dennison and Roberts (2003) employed endmember average root mean square error (EAR) to select endmembers from a spectral library for use in mixture models. However, two challenges are still present. 1) Due to the non-orthogonality of endmember signatures (Van der Meer & De Jong, 2000), it is likely that different combinations of endmembers lead to the same mixed spectra, making it difficult to determine the optimal combination; 2) the computational complexity becomes an issue as the number of candidate endmember combinations increases. Given the fact that pixels in close spatial proximity share similar endmember combinations, the incorporation of spatial information may provide a solution to the aforementioned challenges in the selection of endmember combinations, e.g., reducing the number of possible mixtures, alleviating spectral confusion, and improving computational efficiency.

Based on the extent to which endmember combinations can vary, the spatial spectral unmixing techniques reviewed in this section are grouped as per-pixel methods and per-field methods. Per-pixel methods allow each pixel to have a specific set of endmembers, whereas per-field methods select a distinct endmember combination for each homogenous field partitioned from the remotely sensed image.

3.2.1. Per-pixel methods

One of the earliest attempts for a per-pixel strategy that incorporates spatial information in the selection of endmember combinations is the neighborhood-oriented iterative unmixing procedure (Roessner et al., 2001). Given an initial set of endmembers as training data, this method starts with the search for spectrally pure pixels using supervised classification with rigorous criteria. The pure pixels present in a predefined spatial neighborhood of a nonclassified pixel are then identified and their corresponding ground cover classes are very likely to be present in the nonclassified pixel. All combinations of the endmembers belonging to these classes are tested for unmixing the nonclassified pixel and the most likely endmember combination is selected. In the case that only one pure pixel is identified in the neighborhood, all combinations involving the endmember associated with that pure pixel are checked. If no pure pixels exist in the neighborhood, the nonclassified pixel will not be processed at the current stage. During the next iteration, all classes that are contained in a nonclassified pixel at a preceding iteration have to be considered in the selection of endmember combinations for unmixing pixels. This iterative process is repeated until all pixels are unmixed. Compared to the popular Multiple Endmember Spectral Mixture Analysis (MESMA) that exhaustively tests all combinations of endmembers in the initial set (Roberts et al., 1998), the number of possible endmember combinations are significantly reduced, and thus, spectral confusion of similar endmember combinations is mitigated.

An alternative to selecting endmember combinations with the help of neighboring pixels was presented by King and Younan (2006). Based on the initial attempt made by Luo, King, and Younan (2002), the authors introduced an endmember weighting strategy to determine a particular endmember subset for each pixel. This method is also an iterative unmixing procedure starting with image classification. In order to unmix a nonclassified pixel, different weights are assigned to the endmembers via the information of its neighboring pixels. More specifically, an endmember will receive a large weight if it is present in pure pixels in the neighborhood, whereas a small weight will be assigned to that endmember if it is contained in mixed pixels as one component. Spectral similarity between the nonclassified pixel and each endmember is also utilized to determine the weights. The endmembers receiving the largest weights make up the endmember subset to unmix the nonclassified pixel. Although both of the iterative methods can achieve an improvement of the unmixing accuracy, potential drawbacks might include the sufficiency of spectrally pure pixels required to start the iteration and the sensitivity to the unmixing results of neighboring pixels (King & Younan, 2006; Luo et al., 2002; Roessner et al., 2001).

More recently, Deng and Wu (2013) developed a Spatially Adaptive Spectral Mixture Analysis (SASMA) method to synthesize endmember spectra on a per-pixel basis. In SASMA, candidate endmember pixels and mixed pixels are first classified by using a classification tree method. To decompose a mixed pixel, particular endmember spectra are synthesized as a weighted average of the candidate endmember pixels lying within its spatial neighborhood. Similar to Zortea and Plaza (2009), a distance-based weighting strategy is employed. The synthesized endmember spectra rather than the original ones are then input to a linear mixture model to estimate the abundances. It has been reported that SASMA has a comparable performances with MESMA in terms of the accuracy of abundance estimation. Additionally, the computational burden of SASMA is much less thanks to the synthesis of spatially adaptive endmember spectra. However, the limitation of SASMA lies in the empirical setting of the size of spatial neighborhood. Moreover, the comparison between the synthetic and original endmember spectra need to be further investigated.

Along the direction of MESMA, a hierarchical method was developed to apply spatial constrains to it (Franke, Roberts, Halligan, & Menz, 2009). Different levels of classification complexity were introduced, in
which the simplest level contained two classes, namely, impervious and paved, while the most detailed level consisted of twenty classes. Since higher classification accuracies are achieved at lower levels of complexity, results from one level can be used as spatial constraints for the selection of endmember combinations at a higher level. As such, a hierarchical MESMA with spatially adjusted endmember sets was constructed. It can also effectively reduce the number of possible endmember combinations.

Maselli (2001) introduced the definition of spatially variable spectral endmembers to derive endmember spectra via multivariate regression. Given the fractions of the ground cover classes for certain pixels as training data, the fractions and the spectral values of these pixels serve as independent and dependent variables, respectively. Endmember spectra are then derived by extrapolating the constructed regression model to 100% cover fraction for the corresponding class. In order to incorporate endmember variability in the regression, regression statistics including mean vectors and variance/covariance matrices are locally calibrated to allow them to vary on a per-pixel basis. The spectral values of the training pixels are weighted in accordance with their distances to the examined pixel and the range of the spatial variability of the endmembers derived by the semivariogram analysis. The resulting endmember spectra thus vary pixel by pixel and the intra-class variability of endmembers are properly accounted for (Maselli, 2001).

3.2.2. Per-field methods

When a large number of materials exist throughout a complex scene or the intra-class variability of endmembers hinders the inter-class separability, the use of a single large and fixed set of endmembers in spectral unmixing is unreliable in general, and even impracticable when examining the remotely sensed images with a limited number of wavebands (e.g., Landsat TM images). Several per-field spatial unmixing methods tackle this problem by partitioning a large image into a number of separated fields and assigning each field a smaller set of endmembers that are more spectrally distinct (Canham et al., 2011; Garcia-Haro, Sommer, & Kemper, 2005; Rand & Keenan, 2001; Shoshany & Svoray, 2002; Zare et al., 2013). Similar as their methodological schemes, these methods differ in the techniques adopted to partition the image and in the determination of endmembers for each partitioned region (Table 2).

Rand and Keenan (2001) developed an unsupervised Gibbs-based algorithm that utilizes the properties of Markov random fields (MRFs) to partition a hyperspectral image into homogeneous regions. In their case studies, the partitions resulting from the Gibbs-based algorithm were further grouped into two fields that represented natural and cultural land cover types according to the authors’ scene knowledge. Endmember spectra were manually extracted from the image, and based on the ground-truth information, two sets of representative endmembers for both land cover types were assigned to the two regions, respectively. The experiments showed a significant reduction of confusion between asphalt and (slightly shaded) vegetation.

A similar method was applied to study Mediterranean vegetation (Shoshany & Svoray, 2002). The authors employed the unsupervised ISODATA algorithm to obtain image partitions and then grouped them into two regions corresponding to Mediterranean-type and arid-type environments. Endmembers specific to each region, including vegetation, bare soil and rock, were selected from the pixels lying within the respective region based on the ground-truth information, and thus, spectral variability of endmembers, especially vegetation variations in different environments, was effectively accounted for.

A Variable Multiple Endmember Spectral Mixture Analysis (VMESMA) was developed by Garcia-Haro et al. (2005) to incorporate a hierarchical image segmentation into MESMA. As an iterative feedback process, VMESMA begins its unmixing procedure with an initial endmember set. The unmixing results (e.g., RMSEs) and a priori knowledge from auxiliary data are used to partition the image into regions with user-specified attributes. For the region associated with low unmixing accuracy, additional endmembers are manually updated to determine the endmember set which is only valid within this region. The region-dependent endmember set is then employed in MESMA and the resulting information is used in the subsequent region partition. This process is repeated until the satisfying accuracy is achieved.

More recently, a spatially adaptive method for linear spectral unmixing has been developed (Canham et al., 2011). It uses local EM selection, local abundance map generation, and global EM clustering and is termed the LLG method. A set of square-shaped, non-overlapping sub-images with the same size, called tiles, by the authors, is obtained by conducting image partition, the same process employed by Rogge et al. (2007) and Li and Zhang (2011). The number of endmembers needed to accurately unmix each tile is automatically estimated by using the Gram matrix method (Messinger, Ziemann, Schlamm, & Basener, 2010), and then, the Max-D algorithm (Bajorski, 2004) is applied to identify the endmember spectra from the pixels within each tile. Linear spectral unmixing then operates on the tiles using their corresponding endmembers. Finally, a spectral clustering process is applied to group all of the local endmembers into several common ones across the whole image. In spite of a potential drawback of the user-defined tile size, this method effectively reduces the errors by taking into account the spatial variation of endmember spectra. Another advantage is its modularity as the algorithms used in this method are interchangeable with other functionally similar ones.

Zare et al. (2013) proposed a Piecewise Convex Multiple-Model Endmember Detection (PCOMMEND) algorithm to constrain groups of pixels (not necessarily spatial neighbors) to have same endmember combinations. PCOMMEND conducts an iterative fuzzy clustering process simultaneously with spectral unmixing to partition an image into multiple regions where each region contains a distinct set of endmembers that defines a simplex. All pixels in the image are thus represented by a union of all of the simplices. Parameters such as the number of endmembers per set and the number of sets of endmembers can be autonomously estimated by a cluster validity index (Anderson & Zare, 2012). In addition, neighboring pixels can be encouraged to have same set of endmembers by adding spatial constraints on fuzzy membership values (Zare, Bchir, Frigui, & Gader, 2010).

3.3. Incorporating spatial information in abundance estimation

The primary objective of spectral unmixing is to estimate the fractional abundances of endmembers present in each mixed pixel in the scene (Keshava & Mustard, 2002). The spectrum of a pixel is often modeled as a linear combination of endmember spectra, and the

Table 2

<table>
<thead>
<tr>
<th>Method</th>
<th>Image partitioning</th>
<th>Endmember determination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shoshany and Svoray (2002)</td>
<td>ISODATA</td>
<td>Manual determination based on the ground-truth information</td>
</tr>
<tr>
<td>Garcia-Haro et al. (2005)</td>
<td>Unmixing results and a priori knowledge</td>
<td>Manual determination based on unmixing results</td>
</tr>
<tr>
<td>Canham et al. (2011)</td>
<td>Evenly partitioned, local “tiles”</td>
<td>Automated determination using the Gram matrix algorithm and the Max-D algorithm</td>
</tr>
<tr>
<td>Zare et al. (2013)</td>
<td>Iterative fuzzy clustering</td>
<td>Automated determination in a simultaneous fashion with fuzzy membership values of clusters</td>
</tr>
</tbody>
</table>
abundances can be derived using the least squares method to minimize the mean square error between the actual spectrum and the reconstructed spectrum (Settle & Drake, 1993; Shimabukuro & Smith, 1991). Although mathematically straightforward and broadly applied in spectral unmixing, this procedure addresses the abundance estimation problem in a pixel-by-pixel manner and the abundances of each pixel are derived independently of its neighboring pixels. Nonetheless, due to the spatial arrangement of the pixels in an image, their corresponding abundances often show a pronounced spatial dependence. In addition, the mean square errors should be randomly distributed across the image if each pixel is precisely modeled by the endmember spectra. Recent studies have shown that the incorporation of spatial information in abundance estimation can improve the accuracy of spectral unmixing and also make the obtained abundances more consistent, which match the spatial characteristics of the ground surface.

3.3.1. Spatial characteristics of residuals

Although the least squares method can produce the optimal estimates of abundances for a single pixel, as mentioned earlier, the residuals of all pixels in the image might portray spatial autocorrelation that the conventional inversion methods hardly account for. Defined on the root-mean-square errors (RMSEs) for each pixel, spatial criteria, in addition to other statistical criteria, were employed for an iterative spectral unmixing (ISU) method to address this problem (Van der Meer, 1999). The first spatial criterion is to minimize spatial structure in the RMSEs. An effective measure is the ratio between the sill and the nugget derived from the variogram of the RMSEs. Second, anisotropic variograms can be used to minimize the directional spatial structure. Third, the local variance of RMSEs in a spatial neighborhood is also expected to be as small as possible. In the proposed method, the initial endmembers were obtained using PPI. Next, the RMSEs resulting from spectral unmixing were used to re-position endmembers to improve the results in terms of the aforementioned criteria. This process was repeated iteratively until the criteria were satisfied or no significant improvements were observed. Although this iterative spectral unmixing method takes into account spatial autocorrelation of unmixing errors, the spatial criteria used in this method serve as the termination conditions of the iterative process and are not included in a specific objective function to be optimized.

3.3.2. Spatial characteristics of abundances

In contrast to unmixing residuals, fractional abundances of each endmember are expected to be spatially correlated. The spatial characteristics of endmember abundances have been utilized to improve the accuracy of abundance estimation. In the spectral and spatial complexity blind source separation (SSCBSS) method (Jia & Qian, 2007), the conventional temporal complexity (Stone, 2001) was extended to the spatial complexity which properly expresses local spatial dependence of abundances. The essence of the spatial complexity is the spatial predictability. Specifically, if the abundance in a pixel can be accurately predicted by the weighted average of the abundances in its neighboring pixels, low complexity or high predictability is observed. The objective function was then constructed to maximize the spatial predictability of the abundances. In addition, the temporal complexity of endmember spectra was incorporated in the objective function to achieve the spectral smoothness. The maximal values of the objective function were obtained using a gradient ascent algorithm. It is worth noting that the SSCBSS is a blind source separation algorithm which does not assume endmember spectra are known in advance. The optimal endmember spectra are jointly estimated with their corresponding abundances. This merit particularly fits the remotely sensed images in which pure pixels are insufficient, if not absent. Nevertheless, the method makes use of the unconstrained linear mixture model, and thus, may produce negative values which are physically meaningless.

A different strategy of incorporating spatial dependence in abundance estimation is to add a spatially defined regularizer to the classic spectral-only optimization procedure (Castrodad et al., 2011; Iordache et al., 2012; Jia & Qian, 2009; Liu, Xia, Wang, & Zhang, 2011; Song, Jiang, & Rui, 2010). Song et al. (2010) employed a spatial statistic, Moran’s I (Moran, 1950), to describe spatial autocorrelation of fractional abundances. Among the mixed pixels falling into the same class label, the abundances were assumed similar. Global Moran’s I was used to characterize the spatial consistency of abundances for each endmember. Additionally, local Moran’s I was used to represent the spatial autocorrelation of abundances between a pixel and its neighbors. These indices were combined with the average of the mean square errors to form an objective function which aimed to minimize unmixing errors while maximizing the spatial autocorrelation of abundances. Subject to the nonnegativity and full additivity constraints, a genetic algorithm, optimization procedure that mimics the process of natural selection, was applied to estimate the abundances of endmembers. While the resulting abundance images are more homogenous than the ones derived from the least squares solution in a pixel-by-pixel context, this method is highly computationally intensive due to the slow convergence of the genetic algorithm.

Along this direction, Castrodad et al. (2011) presented the spatial spectral coherence regularization that was imposed to allow abundance estimation for a pixel to be influenced by its spectrally similar and spatially adjacent neighbors. In their proposed dictionary modeling with spatial coherence (DMS) method, the abundances of a pixel were compared with the weighted average abundances of its neighboring pixels. The weights were dependent on spectral similarity between the central pixel and its neighbors, which is different from a weighting strategy based on inverse distances. More specifically, the neighboring pixels whose spectrum was highly correlated with the central pixel’s spectrum received a large weight, while no weight was assigned if the two pixels were orthogonal. By minimizing the discrepancy between a pixel’s abundances and the weighted average abundances of its spatial neighbors, pixels with spatial proximity and spectral similarity were impelled to have similar abundance vectors.

Another choice of the regularizer is the total variation utilized in the sparse unmixing via variable splitting augmented Lagrangian and total variation (SunSAL-TV) algorithm (Iordache et al., 2012). The total variation is essentially the cumulative sum of the 1-norm distances between the abundance vectors of the neighboring pixels. The incorporation of the total variation regularization in the optimization promotes piecewise constant fractional abundances between neighboring pixels. Other constraints on the spatial autocorrelation of abundances have been seen in nonnegative matrix factorization for unmixing hyperspectral data (Jia & Qian, 2009; Liu et al., 2011).

Different from the above regularization methods, Eches et al. (2011) exploited a Markov random field for spectral unmixing. In the proposed method, spatial dependence of fractional abundances was accounted for by assuming the abundances of the pixels that fell into a homogeneous region shared the same first and second order statistical moments, i.e., means and variances. An implicit image classification was jointly conducted with abundance estimation, and the spatial dependence of the hidden class labels assigned for each pixel in the image were assumed to follow a Markov random field. The nonnegativity and full additivity constraints in the linear mixture model were satisfied by using logistic coefficients to reparameterize the abundances. A hierarchical Bayesian framework was introduced to derive the joint posterior distribution of the abundances, the means and variances of the abundances, and the class labels. Due to the intrinsic complexity of the distribution, a Markov chain Monte Carlo (MCMC) algorithm was utilized to generate samples asymptotically distributed according to this joint posterior. The fractional abundances and the hidden labels for each pixel were then jointly estimated from the obtained samples. In spite of its expensive computational cost, this method provides the opportunity of determining confidence intervals for the estimated abundances by using the samples generated from the joint posterior distribution (Eches et al., 2011).
4. Experiments

In this section, we illustrate the performance of spatial spectral unmixing in endmember extraction, selection of endmember combinations and abundance estimation using two synthetic hyperspectral data sets. The goal is to demonstrate the advantages of spatial spectral unmixing over conventional spectral-only unmixing techniques. The reason for using synthetic images in our experiments is that all information of the synthetic data including endmember composition and fractional abundances are known in advance. Therefore, the performance of different methods can efficiently and precisely be examined. Fig. 2 shows a flowchart of our experiment-based evaluation of the representative spectral-only and spatial spectral unmixing methods. All of the steps in the flowchart are elaborated in the remainder of this section. Section 4.1 describes the endmember set and the two synthetic images used in the experiments. Section 4.2 introduces two different metrics employed to assess the accuracy of endmember extraction and abundance estimation. The results of the experiments of endmember extraction, selection of endmember combinations and abundance estimation are shown in Sections 4.3, 4.4 and 4.5, respectively.

4.1. Data description

Five ground materials including alunite, dumotierite, halloysite, kaolinite and sphene were selected from the U.S. Geological Survey (USGS) digital spectral library (Clark et al., 2007) to carry out our experiments. The spectral signatures of the endmembers, available in a reflectance unit, contain 425 bands distributed in the wavelength range between 380 nm and 2500 nm. Fig. 3 plots the spectral reflectance curves of these five materials. Two synthetic images with distinct spatial patterns were generated as follows:

1) Synthetic image 1 (IMG1): this synthetic image was generated using the Hyperspectral Imager Synthesis toolbox for MATLAB (available at http://www.ehu.es/ccwintco/index.php/Hyperspectral_Imagery_Synthesis_tools_for_MATLAB). It provides different options of creating synthetic images with spatial patterns and allows users to define spatial dimensions of the images as well as the number of endmembers utilized. In particular, IMG1, with a size of $100 \times 100$ pixels and 425 spectral bands, was generated according to the Spheric Gaussian Field with the five endmembers. The ground-truth fractional abundance maps for IMG1 are shown in Fig. 4. The abundances are subject to the nonnegativity and full additivity constraints and depict spatial homogeneity in the images. IMG1 has been employed to conduct experiments for endmember extraction and abundance estimation.

2) Synthetic image 2 (IMG2): this synthetic image, with a size of $100 \times 100$ pixels, was generated according to the procedure presented by Mei et al. (2010). The whole image was first partitioned into $5 \times 5$ square-shaped, non-overlapping sub-images with the same size of $20 \times 20$ pixels. Five materials were assigned to the five sub-images in each row so that the pixels in the same sub-image comprised a homogeneous region. In order to simulate global endmember variability, five spectral signatures of each material were created by adding 20 dB Gaussian noises to the original reflectance curve of the corresponding material and each was assigned to one of the five rows. In addition, local endmember variability was simulated by adding 30 dB Gaussian noise to the pixels in each sub-image. Finally, a $15 \times 15$ local window slid across the entire image on a per-pixel basis and the spectra of the center pixel was replaced by the average reflectance values of the pixels lying within the window. As such, mixed pixels were generated if the pixels inside the window were occupied by different materials. The nonnegativity and full additivity conditions of the abundances...
were inherently guaranteed by the simulation procedure. Fig. 5 shows the ground-truth fractional abundance maps for IMG2 which has been employed to conduct experiments for selection of endmember combinations.

4.2. Accuracy assessment

In order to compare the performance of different spectral unmixing methods, two accuracy assessment metrics have been adopted. The first metric is the average SAD between extracted endmembers and their corresponding ground-truth reflectance signatures. Since the set of ground-truth endmembers are known in advance, the type of material that each extracted endmember belongs to can be obtained by identifying its most similar spectral signature across the ground-truth endmember set. SAD has been used to examine spectral similarity of reflectance spectra. The second metric is the average RMSE between the estimated and ground-truth fractional abundances. The availability of the ground-truth abundances associated with synthetic data make it feasible to calculate the RMSE between the estimates and the truth for each endmember. The average RMSE for all endmembers indicates the overall accuracy of abundance estimation.

4.3. Experiment 1: endmember extraction

In this subsection, we conduct a group of comparative experiments of endmember extraction algorithms including PPI, N-FINDR, SSEE and SPP using the synthetic image IMG1. In order to test the algorithms’ susceptibility to noise, white Gaussian noise was added to a certain percentage of the pixels in the image. Two levels of the signal-to-noise ratio (SNR), i.e., 30 dB and 20 dB, were considered in noise simulation, and the percentage of noise-contaminated pixels was set from 10% to 100% in increments of 10%. As a result, 20 synthetic scenes with average SNR values ranging from 40.0 dB to 20.0 dB were generated for the experiments.

First, PPI and SSEE were adopted for the experiments. In our implementation of PPI, 10,000 random vectors were generated for projection and the Otsu’s method (Otsu, 1979) was used to perform thresholding. The final endmembers were obtained by applying k-means clustering to the set of candidates to achieve full automation. In order to make a comparison with PPI, SSEE was applied to the same image. We partitioned the image into 4 square sub-images to generate compiled vectors. Three window sizes of 3 × 3, 5 × 5 and 7 × 7 were used for the spatial averaging process with 100 iterations. The Otsu’s thresholding and k-means clustering methods were also adopted to automate SSEE. The additional experiments were carried out using N-FINDR and SPP. N-FINDR does not require any input parameters other than the number of endmembers to be extracted. SPP, as a preprocessing module, was combined with N-FINDR (referred to as SPP + N-FINDR hereafter) in the experiments. Three window sizes of 3 × 3, 5 × 5 and 7 × 7 were tested for sensitivity analysis.

Tables 3 and 4 show average SADs (in degrees) between the ground-truth endmember signatures and their corresponding image endmembers produced by PPI and SSEE with the respective SNRs of 30 dB and 20 dB used for noise simulation. As shown by the tables, SSEE generally provided better results (lower SAD values) across all noise-contaminated scenes. PPI failed to identify all five endmembers when the average SNR value was lower than 30 dB, which might be attributable to its nature of susceptibility to noise. Conversely, SSEE,
thanks to nonrandom selection of projection vectors from sub-images and the iterative spatial averaging process, produced endmembers more similar to the ground truth. Most of the minimum SAD values were obtained using a window size of 3 × 3. When larger windows were employed, spectrally similar but spatially independent endmembers might be averaged, resulting in a negative impact on the accuracy. The results produced by N-FINDR and SPP + N-FINDR are shown in Tables 5 and 6. It is clear from both tables that SPP + N-FINDR outperformed N-FINDR in most cases. As the average SNR value decreased, the endmembers extracted by N-FINDR depicted increasing discrepancy from the ground-truth signatures, indicating that N-FINDR was highly sensitive to noise. On the contrary, SPP + N-FINDR consistently identified more accurate endmembers regardless of average SNR levels. It is worth noting that the results of SPP + N-FINDR were not very sensitive to the window size, as the SAD values with different window sizes were generally lower than the ones without using preprocessing. N-FINDR produced results that were close or equal to the ones obtained by SPP + N-FINDR when large percent of pixels were contaminated by noise. In this scenario, it was difficult for SPP to find homogenous regions from the image, therefore, the positive impact of SPP became trivial.

4.4. Experiment 2: selection of endmember combinations

Our experiment on the selection of endmember combinations was conducted on the synthetic image IMG2. Three unmixing methods were compared: 1) fully constrained linear spectral unmixing (FCLSU) (Silvan-Cardenas & Wang, 2010a), a spectral-only unmixing method, 2) SASMA, a per-pixel spatial spectral unmixing method, and 3) LLG, a per-field spatial spectral unmixing method. FCLSU accounts for both full additivity and nonnegativity constraints and does not require any parameters. In our implementation, the endmembers input into FCLSU were extracted by the N-FINDR algorithm. For SASMA, a supervised classification tree method was applied to IMG2 to classify the image into pure pixels (in the five material types) and mixed pixels. Three window sizes of 21 × 21, 25 × 25 and 31 × 31 were used to produce spatially adaptive endmember spectra on a per-pixel basis. The fractional abundances of the five materials were estimated by FCLSU. For LLG, Table 3

<table>
<thead>
<tr>
<th>Percentage</th>
<th>Average SNR (dB)</th>
<th>PPI</th>
<th>SSEE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Window size</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Window size</td>
<td>3</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>10%</td>
<td>40.0</td>
<td>2.5122</td>
<td>0.7483</td>
</tr>
<tr>
<td>20%</td>
<td>37.0</td>
<td>2.1026</td>
<td>0.7060</td>
</tr>
<tr>
<td>30%</td>
<td>35.2</td>
<td>1.6269</td>
<td>0.6915</td>
</tr>
<tr>
<td>40%</td>
<td>34.0</td>
<td>1.8038</td>
<td>0.6816</td>
</tr>
<tr>
<td>50%</td>
<td>33.0</td>
<td>1.5860</td>
<td>0.6371</td>
</tr>
<tr>
<td>60%</td>
<td>32.2</td>
<td>1.5567</td>
<td>0.8312</td>
</tr>
<tr>
<td>70%</td>
<td>31.6</td>
<td>1.5183</td>
<td>0.9440</td>
</tr>
<tr>
<td>80%</td>
<td>31.0</td>
<td>1.5612</td>
<td>0.9248</td>
</tr>
<tr>
<td>90%</td>
<td>30.5</td>
<td>1.6517</td>
<td>1.1642</td>
</tr>
<tr>
<td>100%</td>
<td>30.0</td>
<td>1.4803</td>
<td>1.2984</td>
</tr>
</tbody>
</table>

Table 4

<table>
<thead>
<tr>
<th>Percentage</th>
<th>Average SNR (dB)</th>
<th>PPI</th>
<th>SSEE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Window size</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Window size</td>
<td>3</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>10%</td>
<td>29.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20%</td>
<td>27.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30%</td>
<td>25.2</td>
<td>2.9686</td>
<td></td>
</tr>
<tr>
<td>40%</td>
<td>24.0</td>
<td>3.1162</td>
<td></td>
</tr>
<tr>
<td>50%</td>
<td>23.0</td>
<td>3.1928</td>
<td></td>
</tr>
<tr>
<td>60%</td>
<td>22.2</td>
<td>2.22</td>
<td></td>
</tr>
<tr>
<td>70%</td>
<td>21.6</td>
<td>2.3099</td>
<td></td>
</tr>
<tr>
<td>80%</td>
<td>21.0</td>
<td>2.01</td>
<td></td>
</tr>
<tr>
<td>90%</td>
<td>20.5</td>
<td>2.5427</td>
<td></td>
</tr>
<tr>
<td>100%</td>
<td>20.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 5. Ground-truth fractional abundances of the endmembers in the synthetic image 2 (IMG2): (a) abundances of alunite; (b) abundances of dumortierite; (c) abundances of halloysite; (d) abundances of kaolinite; and (e) abundances of sphene.

Table 3

Average SADs (in degrees) between the ground-truth endmember signatures and their corresponding image endmembers produced by PPI and SSEE with the SNR of 30 dB used for noise simulation. The minimum values of SADs for each noise-contaminated scene are outlined in bold typeface.

Table 4

Average SADs (in degrees) between the ground-truth endmember signatures and their corresponding image endmembers produced by PPI and SSEE with the SNR of 20 dB used for noise simulation. The minimum values of SADs for each noise-contaminated scene are outlined in bold typeface.

---

* The percentage of pixels in the image that are contaminated by noise.

* The average SNR value accounting for all pixels in the image.

* The results are not shown due to incompleteness of extracted endmember set.
we partitioned IMG2 into four sub-images with the same size of 50 × 50 pixels. In each sub-image, N-FINDR was adopted to extract local endmembers that were then input into FCLSU to produce fractional abundances. Spatial neighborhoods with small sizes were not suitable, since sufficient pixels in a sub-image were needed to extract all types of endmembers.

Table 7 shows the individual and average RMSEs between the ground-truth fractional abundances and their corresponding abundances estimated by FCLSU, SASMA and LLG. Among these three methods, FCLSU produced the lowest accuracy for each material, as it failed to account for endmember variability present in the image. LLG generated four spectral signatures for each endmember and used a set of local endmembers to unmix pixels in each sub-image. As such, the RMSE values between the estimated and the ground-truth abundances were reduced. The best results were obtained by SASMA. It allowed the spectral signatures of the five endmembers to vary on a per-pixel basis, and each pixel in the image was assigned a unique set of endmembers that were very similar to the endmembers used in the underlying linear mixing process. The window size of 31 × 31 provided the most accurate estimation of abundances, which was attributed to a substantial number of pure pixels considered in synthesis of endmember spectra.

The abundance maps estimated for the five materials in IMG2 are shown in Fig. 6. The maps of SASMA were produced with a window size of 31 × 31. The visual delineation coincides with the quantitative results reported in Table 7. FCLSU produced accurate abundances in the regions where the image endmembers were extracted. However, the overall quality of the abundance maps was lowered, as variable endmembers were not taken into account in other regions. By conducting endmember extraction and abundance estimation in local sub-images, LLG had better capacity of mitigating endmember variability than FCLSU. But it should be noted that a tiling artifact was observed between adjacent sub-images due to clustering of the locally identified endmembers into global endmember groups. The abundance maps produced by SASMA were the most similar to the ground truth, which revealed its strength of characterizing endmember variability by the incorporation of spatial information.

4.5. Experiment 3: abundance estimation

Our experiment of abundance estimation evaluated the performance of FCLSU and FCLSU-TV with the total variation (TV) regularization (referred to as FCLSU-TV hereafter). The synthetic image IMG1 was used to carry out the experiment. White Gaussian noise with the SNR values of 25 dB, 20 dB, 15 dB and 10 dB was added to all pixels in the image. In order to eliminate the impact of endmember extraction algorithms, the five ground-truth endmember spectra were used to estimate fractional abundances. No parameters were specified for FCLSU, while different values of the TV regularization parameter ranging from 0.001 to 0.3 were evaluated for FCLSU-TV. In all the tests of FCLSU-TV, a 4-neighborhood system with a cross shape was considered.

The average RMSEs between the ground-truth fractional abundances and their corresponding abundances estimated by FCLSU and FCLSU-TV are shown in Table 8. From the table, it can be seen that FCLSU-TV produced more accurate abundances in all considered SNR levels. When the SNR was high, the improvement obtained by the inclusion of the TV regularizer with regard to the standard FCLSU was not significant. However, FCLSU-TV outperformed FCLSU more apparently as the SNR decreased. It is worth noting that the TV regularization parameter controlled the influence of the TV regularizer and promoted spatial consistency of estimated abundances. Therefore, it is not surprising to see that bigger parameters were needed to achieve high accuracy as the SNR got lower. Fig. 7 shows the abundance maps obtained by FCLSU and FCLSU-TV for the five endmembers in IMG1 with the SNR value of 10 dB. Visual inspection confirms the results reported in Table 8. The abundances estimated by FCLSU-TV became more homogeneous in nature. The spatial distribution of the materials was approximated with higher accuracy by the incorporation of the TV regularizer.

5. Considerations

The integration of spatial and spectral information has been identified as a new trend by many researchers devoted to remotely sensed data analysis (Plaza et al., 2009a). The use of spatial spectral unmixing has drawn increasing attention in the remote sensing research community and has been shown to be successful at improving the performances of endmember extraction, selection of endmember combinations and abundance estimation. Many benefits can be gained by integrating spatial
information with spectral information, as illustrated in our comparative experiments. First, extracted endmembers that account for spatial context of candidate pixels may represent the majority of pixels more consistently and be less susceptible to outlier pixels (which tend to be incorrectly extracted by spectral-only endmember extraction methods). With regard to selection of endmember combinations, the incorporation of spatial information balances the tradeoff between the insufficiency of a single fixed set of endmembers and the complexity of multiple variable endmembers so that the spectral confusion and the computational burden caused by numerous endmember combinations are reduced. Furthermore, abundance estimation benefits from incorporating spatial information since the estimated abundance fractions become more...
spatially consistent. Spatial dependence of abundances is hardly characterized by spectral-only unmixing methods. Along with these advantages, spatial spectral unmixing is also confronted with some issues that are not fully resolved and yet need to be carefully considered in method design.

5.1. Spatial dependence

The property of spatial dependence that delineates the coincidence between the similarity of the ground objects and their locational adjacency is the theoretical foundation of spatial spectral unmixing. Understanding spatial dependence present in remotely sensed images is likely the most important condition for developing spatial spectral unmixing methods. Nonetheless, spatial dependence has rarely been investigated to determine whether it is present in the remotely sensed data before it is utilized to analyze the scene. A critical assumption of many methods described above is that spatial dependence between neighboring pixels so strongly exists that the examined image is comprised of a series of homogenous regions. The efficacy of these methods will be significantly reduced if the assumption does not hold (Zortea & Plaza, 2009). It is even possible that the incorporation of spatial information in the analysis of a highly heterogeneous scene will worsen the results obtained by a conventional spectral-only method. Therefore, it is important to assess
spatial dependence of the image beforehand in order to justify the necessity of the incorporation of spatial information. This assessment cannot be accomplished based on visual inspection alone. Quantitative analysis has to be introduced to fulfill this need. In this regard, a suite of spatial statistics, such as Moran’s I (Moran, 1950), Geary’s c (Geary, 1954), and Getis’ G (Getis & Ord, 1992), can derive the degree to which an observed spatial pattern deviates from complete spatial randomness (Cressie, 1993) and might be potential resources for assessing spatial dependence of the remotely sensed data (Emerson, Lam, & Quattrochi, 2005; Su et al., 2008; Wulder & Boots, 1998).

Another consideration is where spatial dependence is present in the remotely sensed data. An image might contain certain regions that depict strong spatial dependence between neighboring pixels, while in other regions weak or no spatial dependence is observed. However, most of the methods described above treat each spatial neighborhood evenly when analyzing the scene. Little effort has been made in the analysis of varying degrees of spatial dependence in different regions. Ideally, the extent to which spatial information is used to analyze a pixel should match the degree of spatial dependence inherent in this pixel and its spatial neighbors. Spectral-only methods would be preferred in case of no significant spatial dependence present in some particular spatial neighborhoods. One of our ongoing studies is focused on a hybrid framework that combines spatial spectral and spectral-only methods to analyze complex scenes where both homogenous and heterogeneous regions coexist.

5.2. Spatial neighborhood

The most significant distinction between the spatial spectral methods and the spectral-only methods is the definition and description of a spatial neighborhood. A well-defined spatial neighborhood system should capture spatial dependence between pixels and also avoid exaggerating it. With the purpose of an overview of the spatial neighborhood systems used by the spatial spectral unmixing methods reviewed in this paper, Table 9 summarizes their main characteristics such as the shape, the size, the motion capacity of the neighborhood, whether weights are assigned to the elements in the neighborhood, and the similarity measures used. It should be noted that the detailed characteristics of each method included in the table are filled in according to the corresponding published paper. However, these methods are not limited to the particular characteristics specified in the papers. Other functionally similar ones can also be used without changing the overall methods.

From Table 9, it can be seen that the majority of the methods employed a spatial neighborhood system with a regular shape (e.g., square or cross) and a predefined window size. Such a spatial window often slides through the pixels of interest in the image to account for their spatial context. It is worth noting that the selection of the window size is largely dependent on the scale of spatial heterogeneity of the ground surface. In practice, a small spatial kernel (4- or 8-neighborhood system) is often used in order to prevent the proposed algorithm from introducing errors as a consequence of the exaggeration of spatial dependence. The sensitivity to the size of spatial neighborhood can be reduced by an iterative process with an increasing window size, which was applied in AMEE (Plaza et al., 2002). In contrast with the regularly shaped, moving spatial neighborhood, other methods utilize a spatially adaptive neighborhood system to perform spectral unmixing (see Table 9 for details). Resulting from image segmentation, these neighborhoods depict varying sizes and shapes, and thus, seem better suited to address real scenarios with irregular morphologies. However, the errors of image segmentation might be propagated to the stages of spectral unmixing. Also, the generated spatial neighborhoods become static, losing the motion capacity presented by moving windows.

As shown in Table 9, plain kernels without weights often comprise a spatial neighborhood system. However, the exploitation of different weighting strategies has also been seen in spatial spectral unmixing. A popular method of assigning weights is based on spatial distance (Deng & Wu, 2013; Jia & Qian, 2007; Li & Zhang, 2011; Song et al., 2010; Zortea & Plaza, 2009). Specifically, a neighboring pixel receives a higher weight if it is closer to the central pixel in the spatial window. Another criterion of the weighting strategy is spectral similarity (Castrodad et al., 2011). According to this criterion, the neighboring pixel that is spectrally similar to the central pixel is assigned a heavy weight. It has been reported that assigning each pixel in the spatial window a particular weight is beneficial to the reduction of the sensitivity to the window size (Zortea & Plaza, 2009), as shown by SPP + N-FINDR in our experiments (see Tables 5 and 6).

Table 9 lists a collection of similarity measures that are often used to examine the homogeneity of spectral signatures or abundance vectors.

<table>
<thead>
<tr>
<th>References</th>
<th>Shape</th>
<th>Size</th>
<th>Motion</th>
<th>Weighting</th>
<th>Similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endmember extraction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plaza et al. (2002)</td>
<td>Square</td>
<td>Varying</td>
<td>Moving</td>
<td>No</td>
<td>SAD</td>
</tr>
<tr>
<td>Mei et al. (2010)</td>
<td>Cross</td>
<td>Fixed</td>
<td>Moving</td>
<td>No</td>
<td>SAD</td>
</tr>
<tr>
<td>Rogge et al. (2007)</td>
<td>Square</td>
<td>Fixed</td>
<td>Static</td>
<td>Moving</td>
<td>No</td>
</tr>
<tr>
<td>Zhang et al. (2008)</td>
<td>Square</td>
<td>Fixed</td>
<td>Static</td>
<td>No</td>
<td>SAD</td>
</tr>
<tr>
<td>Li and Zhang (2011)</td>
<td>Square</td>
<td>Fixed</td>
<td>Static</td>
<td>Yes</td>
<td>SAD</td>
</tr>
<tr>
<td>Zortea and Plaza (2009)</td>
<td>Square</td>
<td>Fixed</td>
<td>Moving</td>
<td>Yes</td>
<td>SAD</td>
</tr>
<tr>
<td>Thompson et al. (2010)</td>
<td>Adaptive</td>
<td>Varying</td>
<td>Static</td>
<td>No</td>
<td>SAD</td>
</tr>
<tr>
<td>Selection of endmember combinations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Martin and Plaza (2011)</td>
<td>Adaptive</td>
<td>Varying</td>
<td>Static</td>
<td>No</td>
<td>EuD</td>
</tr>
<tr>
<td>Roessner et al. (2001)</td>
<td>Square</td>
<td>Fixed</td>
<td>Moving</td>
<td>No</td>
<td>N/A</td>
</tr>
<tr>
<td>King and Younan (2006)</td>
<td>Square</td>
<td>Fixed</td>
<td>Moving</td>
<td>Yes</td>
<td>N/A</td>
</tr>
<tr>
<td>Deng and Wu (2013)</td>
<td>Square</td>
<td>Fixed</td>
<td>Moving</td>
<td>Yes</td>
<td>N/A</td>
</tr>
<tr>
<td>Shoshany and Svoray (2002)</td>
<td>Adaptive</td>
<td>Varying</td>
<td>Static</td>
<td>No</td>
<td>EuD</td>
</tr>
<tr>
<td>Garcia-Haro et al. (2005)</td>
<td>Adaptive</td>
<td>Varying</td>
<td>Static</td>
<td>No</td>
<td>N/A</td>
</tr>
<tr>
<td>Canham et al. (2011)</td>
<td>Square</td>
<td>Fixed</td>
<td>Static</td>
<td>No</td>
<td>SAD</td>
</tr>
<tr>
<td>Abundance estimation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Van der Meer (1999)</td>
<td>Square</td>
<td>Fixed</td>
<td>Moving</td>
<td>No</td>
<td>N/A</td>
</tr>
<tr>
<td>Jia and Qian (2007)</td>
<td>Square</td>
<td>Fixed</td>
<td>Moving</td>
<td>Yes</td>
<td>EuD</td>
</tr>
<tr>
<td>Song et al. (2010)</td>
<td>Square</td>
<td>Fixed</td>
<td>Moving</td>
<td>Yes</td>
<td>N/A</td>
</tr>
<tr>
<td>Castrodad et al. (2011)</td>
<td>Square</td>
<td>Fixed</td>
<td>Moving</td>
<td>Yes</td>
<td>EuD</td>
</tr>
<tr>
<td>Iordache et al. (2012)</td>
<td>Cross</td>
<td>Fixed</td>
<td>Moving</td>
<td>No</td>
<td>1 norm</td>
</tr>
<tr>
<td>Eches et al. (2011)</td>
<td>Cross</td>
<td>Fixed</td>
<td>Moving</td>
<td>No</td>
<td>N/A</td>
</tr>
</tbody>
</table>
in a spatial neighborhood. The SAD is a widely used measure of spectral similarity due to its insensitivity to illumination. The similarity between abundance vectors is often measured by the Euclidean distance. Other similarity measures such as correlation coefficient, SID–SAD and 1-norm distance have been utilized in the recently proposed methods (Iordache et al., 2012; Li & Zhang, 2011; Mei et al., 2010). However, the comparison between these similarity measures need to be further investigated (van der Meer, 2006). Potential questions include, for example, which similarity measure is the most effective in discriminating spectral signatures of endmember candidates, how sensitive the similarity measures are to spectral mixing and spectral noise, and whether (and how) the similarity between abundance vectors need to be addressed differently from spectral similarity.

5.3. Sensitivity to the specification of parameters

Spatial spectral unmixing methods often have multiple parameters to be specified before they are carried out. Nevertheless, the sensitivity of these methods to the specified parameters has seldom been discussed. One of the most common parameters is the window size (if the spatial neighborhood system is of a regular shape). Plaza et al. (2002), Zortea and Plaza (2009) and Canham et al. (2011) conducted an analysis of the effect that the window size had on their methods. However, many studies either exploited a predefined spatial neighborhood size (Eches et al., 2011; Iordache et al., 2012; King & Younan, 2006; Roessner et al., 2001; Zhang et al., 2008) or determined the size based on the authors’ knowledge about the study area (Deng & Wu, 2013).

In our experiments, different window sizes have been used for SSEE, SPP and SASMA. The results show that the optimal window size is dependent on both data and methods. As of today, the determination of the optimal window size has not been fully resolved. However, a general suggestion is to use a small window size for images with a high degree of spatial heterogeneity whereas a large size can be adopted to analyze a spatially homogeneous scene. In addition, it may be possible to employ geostatistical approaches (e.g., variogram models) to identify the range of the underlying spatial autocorrelation in remotely sensed images and help determine the appropriate spatial neighborhood size (Curran, 1988; Curran & Atkinson, 1998; Van der Meer, 2012).

With regard to other parameters specified for the spatial spectral unmixing methods, an empirical optimization procedure is often employed to achieve maximum performance of the methods (Iordache et al., 2012; Thompson et al., 2010). In our experiments with FCLSU-TV, a wide range of values for the TV regularization parameter has been evaluated. Although an optimal parameter can be determined through such an exhaustive search, it will undoubtedly introduce extra computational burden. How to specify parameters that are adaptive to both the remotely sensed data and the adopted spatial spectral unmixing methods still remains an open question.

5.4. Comparative analysis

Although research on the incorporation of spatial information in spectral unmixing began about a decade ago, comparative studies of spatial spectral methods and spectral-only methods are still insufficient. An example has been shown by Plaza et al. (2004) who compared a number of standard endmember extraction methods and concluded that AMEE which combined spatial and spectral information outperformed several spectral-only methods. Few studies have been conducted to compare spatial spectral and spectral-only methods for the selection of endmember combinations or abundance estimation. In this paper, we select five spatial spectral methods (including SSEE, SPP, LLG, SASMA, and FCLSU-TV) and three spectral-only methods (including PPI, N-FINDR, and FCLSU) that represent the three aspects of the spectral unmixing technique, and carry out a series of experiments using two synthetic hyperspectral images. The results show that the spatial spectral methods can improve performances of endmember determination and abundance estimation. However, there are over twenty spatial spectral methods reviewed in this paper (Table 1) and the number will be rising. In addition, a large collection of the spectral-only methods has been presented to the remote sensing community (Bioucas-Dias et al., 2012; Quintino et al., 2012). Therefore, more comparative analyses between spectral-only and spatial spectral methods are needed to fully evaluate the benefits (and potential drawbacks) of incorporating spatial information in spectral unmixing.

5.5. Integrated framework

To date, spatial information has been incorporated in the three aspects of spectral unmixing, namely, endmember extraction, selection of endmember combinations and abundance estimation. In each method discussed above, only a single aspect was considered in the incorporation of spatial information. The integration of different aspects is rarely seen in spatial spectral unmixing. In this regard, optimization of integrated objective functions of spectral unmixing may provide a potential framework in which multiple tasks including spatial consistency of endmember spectra or fractional abundances can be taken on simultaneously. For example, Iordache et al. (2012) included the total variation regularization to the sparse unmixing formulation in order to find the optimal endmember combinations from a large spectral library and achieve spatial consistency of estimated abundances. Zare et al. (2013) estimated endmember spectra and fractional abundances through a piecewise convex representation of hyperspectral data and determined multiple sets of endmembers that accounted for endmember variability in different regions in the image. A spatial constraint on membership values to different endmember sets could also be integrated in the objective function (Zare et al., 2010). Although spatial information is still involved in a single aspect of spectral unmixing, these methods have shed light on an integrated framework in which multiple objectives involving both spatial and spectral information can be included in a single function to be optimized. The resultant objective functions may be complex and the optimal solutions may not be easy to derive through classic linear or quadratic programming routines. Nonetheless, the alternating direction method of multipliers (ADMM) has the capacity of decomposing a difficult optimization problem into a series of simpler ones (Boyd, Parikh, Chu, Peleato, & Eckstein, 2011; Eckstein & Bertsekas, 1992), and has shown its power in spectral unmixing (Bioucas-Dias & Figueiredo, 2010). Our current research has been exploring the possibility of an integrated framework that incorporates spatial information in multiple aspects of spectral unmixing with the ADMM optimization procedure.

6. Summary

Spectral unmixing provides a way of quantitatively analyzing subpixel components for remotely sensed images. However, spatial dependence among pixels is neglected in many spectral-only unmixing methods. This paper focused on summarizing the different methods that incorporate spatial information, inherent to remotely sensed imagery, in order to improve the performance of spectral unmixing. The available methods in the literature were reviewed in the three aspects of spectral unmixing: 1) endmember extraction, 2) selection of endmember combinations, and 3) abundance estimation. An experiment-based comparison between spectral-only and spatial spectral unmixing methods was also conducted. In addition, we pointed out five important issues that need to be carefully considered in method design of spatial spectral unmixing: 1) spatial dependence, 2) spatial neighborhood, 3) sensitivity to the specification of parameters, 4) comparative analysis, and 5) integrated framework. With this review, we hope more spectral unmixing methods can benefit from the integration of both spectral and spatial information in the future.