Regional clustering-based spatial preprocessing for hyperspectral unmixing

Xiang Xu[a,b], Jun Li[b,*], Changshan Wu[c], Antonio Plaza[d]

[a] *University of Electronic Science and Technology of China, Zhongshan Institute, Zhongshan 528402, China*
[b] *Guangdong Provincial Key Laboratory of Urbanization and Geo-simulation, Center of Integrated Geographic Information Analysis, School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China*
[c] *Department of Geography, University of Wisconsin-Milwaukee, P.O. Box 413, Milwaukee, WI 53201-0413, United States*
[d] *Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, Escuela Politécnica, University of Extremadura, Cáceres, Spain*

ARTICLE INFO

ABSTRACT

Hyperspectral unmixing is an important technique for remote sensing image exploitation. It aims to decompose a mixed pixel into a collection of spectrally pure components (called *endmembers*), and their corresponding proportions (called *fractional abundances*). In recent years, many studies have revealed that unmixing using spectral information alone does not sufficiently incorporate the spatial information in the remotely sensed hyperspectral image, as the pixels are treated as isolated entities without taking into account the existing local correlation among them. To address this issue, several spatial preprocessing methods have been developed to include spatial information in the spectral unmixing process. In this paper, we present a new spatial preprocessing method which presents several advantages over existing methods. The proposed method is derived from the Simple Linear Iterative Clustering (SLIC) method, which adapts the global search scope of the clustering into local regions. As a result, the spatial correlation and the spectral similarity are intrinsically incorporated at the clustering step, which results in $O(N)$ computational complexity of the clustering procedure with N being the number of pixels in the image. First, a regional clustering is iteratively performed by using spatial and spectral information simultaneously. The obtained result is a set of clustered partitions that exhibit both spectral similarity and spatial correlation. Then, for each partition we select a subset of candidate pixels with high spectral purity. Finally, the obtained candidate pixels are gathered together and fed to a spectral-based endmember extraction method to extract the final endmembers and their corresponding fractional abundances. Our newly developed method naturally integrates the spatial and the spectral information to retain the most relevant endmember candidates. Our experimental results, conducted using both synthetic and real hyperspectral scenes, indicate that the proposed method can obtain accurate unmixing results with less than 0.5% of the number of pixels used by other state-of-the-art methods. This confirms the advantages of integrating spatial and spectral information for hyperspectral unmixing purposes.

1. Introduction

Pixels in a hyperspectral image are often a mixture of different substances (Schaepman et al., 2009). Spectral unmixing (Bioucas-Dias et al., 2012) allows us to model each mixed pixel as a combination of pure materials (*endmembers*), weighted by their corresponding proportions (*fractional abundances*). Endmember extraction is a very important step in the hyperspectral unmixing chain (Plaza et al., 2009). The endmember spectral signatures can be obtained from existing spectral libraries which are acquired from field or laboratory measurements (Somers et al., 2011; Roberts et al., 1993; Herold et al., 2004; Roberts et al., 2004; Okin et al., 2013). Many known spectral libraries are now publicly available, such as the U.S. Geological Survey (USGS) digital

spectral library (available online: http://speclab.cr.usgs.gov/spectral-lib.html), which contains over 1300 mineral spectral signatures. Also, the endmember spectral signatures can be acquired from the image itself (Somers et al., 2011; Dennison and Roberts, 2003). Compared with the former approach, the latter exhibits consistent acquisition and temporal conditions with the image pixels, and thus it can bring more accurate explanatory ability for subsequent unmixing purposes.

Due to the aforementioned reasons, many automatic or semi-automatic techniques have been developed for image endmember extraction, where the endmembers are directly extracted from the image data. Available techniques can be grouped into two main categories: 1) methods that assume the availability of pure signatures in the image, such as the Pixel Purity Index (PPI) (Boardman et al., 1995), Vertex

Component Analysis (VCA) (Nascimento and Biocas-Dias, 2005), Orthogonal Subspace Projection (OSP) (Harsanyi and Chang, 1994), N-FINDR (Winter, 1999) or Iterative Error Analysis (IEA) (Neville et al., 1999), among many others (Bioucas-Dias et al., 2012); and 2) methods that do not assume the presence of pure signatures in the image, such as the Minimum Volume Spectral Analysis (MVSA) (Li et al., 2015), the Simplex Identification via Split Augmented Lagrangian (SISAL), among many others (Bioucas-Dias et al., 2012). In the recent literature, several techniques have been developed to exploit a potentially very large spectral library; the unmixing then amounts to choosing an optimal subset of library endmembers to model each pixel (Iordache et al., 2011). Methods, such as Orthogonal Matching Pursuit (OMP) (Pati et al., 1995), Basis Pursuit (BP) (Chen et al., 2001), Basis Pursuit Denoising (BPDN) (Chen et al., 2001), and Iterative Spectral Mixture Analysis (ISMA) (Rogge et al., 2007b), belong to this category.

The above mentioned techniques used a fixed number of endmember spectra, i.e. one single endmember spectrum per endmember class, which is simple and easy to implement. However, due to environmental, atmospheric and temporal factors, endmember variability commonly exists in hyperspectral image data. Relevant reviews on this topic have been provided in Somers et al. (2011) and Zare and Ho (2014). Compared with the use of a fixed number of endmember spectra, the use of multiple endmembers per class can provide more accurate spectral signature representation and fractional abundance estimation. Numerous techniques and applications have been proposed to consider endmember variability in spectral unmixing, such as Iterative Endmember Selection (IES) (Roth et al., 2012; Schaaf et al., 2011). Of particular importance is the Multiple Endmember Spectral Mixture Analysis (MESMA) techniques (Roberts et al., 1998; Somers and Asner, 2013; Liu and Yang, 2013; Quintano et al., 2013; Fernández-Manso et al., 2012; Delalieux et al., 2012; Thorp et al., 2013; Franke et al., 2009; Powell et al., 2007), which use variable endmember sets to unmix each pixel of the scene.

In recent years, several studies have revealed that hyperspectral unmixing by using spectral information alone does not sufficiently exploit the spatial information in the scene (Shi and Wang, 2015), as the pixels are treated as isolated entities without taking into account the existing local correlation between them. In real hyperspectral images, pure pixels are more likely to be present in spatially homogeneous regions, and the existing spatial correlation among neighboring pixels can be exploited. To address this important issue, several endmember extraction algorithms have been designed with the goal of integrating the spatial and the spectral information. According to the cooperative use of spectral and spatial information, these methods can be divided into two categories: 1) integrated spatial-spectral methods, such as the Automatic Morphological Endmember Extraction (AMEE) (Plaza et al., 2002), Spatial-Spectral Endmember Extraction (SSEE) (Rogge et al., 2007a), Successive Projection Algorithm (SPA) (Zhang et al., 2008), Spatial Purity based Endmember Extraction (SPEE) (Mei et al., 2010), the Hybrid Automatic Endmember Extraction Algorithm (HEEA) (Li and Zhang, 2011), Spatial Adaptive Linear Unmixing Algorithm (SALUA) (Goenaga et al., 2013), the Unsupervised Unmixing based on Multiscale Representation (UUMR) (Torres-Madronero and Velez-Reyes, 2014), or the Image-based Endmember Bundle Extraction Algorithm (Xu et al., 2015), among many others (Bioucas-Dias et al., 2012); and 2) spatial preprocessing methods, which provide an (optional) preprocessing step before the application of a spectral-based endmember extraction algorithm. In this case, the output of the preprocessing is not the final endmember set, but a set of candidate pixels that need to be fed to an existing spectral-based endmember extraction algorithm to obtain the final endmember set. Available methods in this category include the Spatial Preprocessing (SPP) (Zortea and Plaza, 2009), Region-based Spatial Preprocessing (RBSPP) (Martín and Plaza, 2011), Spatial-Spectral Preprocessing (SSPP) (Martín and Plaza, 2012),

Superpixel Endmember Detection Algorithm (SEDA) (Thompson et al., 2010), Spatial Edges and Spectral Extremes based Preprocessing (SE^2PP) (Lopez et al., 2013), a Fast Spatial-Spectral Preprocessing Module (SSPM) (Kowkabi et al., 2016b), Clustering and Over-segmentation-based Preprocessing (COPP) (Kowkabi et al., 2016a), etc.

Compared with integrated spatial-spectral methods, spatial preprocessing methods can be flexibly included with existing spectral-based endmember extraction methods without modifying such methods. Also, since the number of candidate pixels is much smaller than the number of original image pixels, the computational burden is significantly reduced. However, a general issue with available spatial preprocessing methods is that they generally prioritize one of the two sources of information (spatial or spectral) when conducting the preprocessing, which can have an important influence on the final results as some important candidates may be lost in the preprocessing (Martín and Plaza, 2012).

In this paper, we develop a new method for spatial preprocessing for hyperspectral unmixing. The proposed method naturally balances the spatial and the spectral information by means of a regional clustering procedure that is similar to the one performed by the Simple Linear Iterative Clustering (SLIC) (Achanta et al., 2012) method. Compared with conventional global clustering procedures, the proposed method presents two key differences. First, as conventional clustering algorithms need to search the whole image domain to find the clusters, we restricted the search scope to a local neighborhood around each clustering center. Second, we adopted a clustering criterion that integrates spatial and spectral information simultaneously. After the clustering procedure, we obtain a set of clustering partitions that exhibit both spatial correlation and spectral similarity, which are highly desirable properties for spatial preprocessing purposes. Then we select a subset of candidate pixels from each partition by accounting for their spectral purity. Finally, the obtained candidate pixels are gathered together and fed to a spectral-based endmember extraction method to obtain the final endmember set. Our experimental results with synthetic and real hyperspectral scenes indicate that, compared with other available strategies for spatial preprocessing, the newly proposed method is fast and able to consistently provide candidate pixels with higher quality regarding their spatial and spectral information, which represents a significant improvement over other existing methods.

The remainder of this paper is organized as follows. Section 2 provides a review of the endmember extraction methods considered in our experiments. Section 3 describes the newly proposed method in step-by-step fashion. Section 4 performs an extensive validation and quantitative assessment of the proposed method by using both synthetic and real hyperspectral data sets. Finally, Section 5 concludes the paper with some remarks and hints at plausible future research.

2. Related work

The goal of our proposed spatial-spectral preprocessing strategy is to extract spectrally pure candidates from the original image data set, thus reducing the number of candidate endmembers and improving unmixing accuracy simultaneously. Considering that spectrally pure signatures are more likely to appear in spatially homogeneous areas, and that most pure candidates generally exhibit the most singular signature in such homogeneous area, many existing methods adopted certain *homogeneous criteria* to characterize spectral purity. In the AMEE method (Plaza et al., 2002), a Morphological Eccentricity Index (MEI) is assigned to the purest pixel (obtained by the dilation operation) in a spatial kernel, where the MEI is calculated by using the Spectral Angle Distance (SAD) between itself and the most highly mixed pixel (obtained by the erosion operation) in the spatial kernel. In the HEEA method (Li and Zhang, 2011), a joint Spectral Information Divergence and Spectral Angle Mapper (SID-SAM) metric (Du et al., 2004),

combined with a spatial distance weight, is adopted to determine the purity of each candidate endmember in the neighborhood window. In the SSPP method (Martín and Plaza, 2012), the pixel's spectral purity is characterized by using the Root Mean Square Error (RMSE) (Keshava and Mustard, 2002) between two images: the original image and a filtered version obtained by using Multi-Scale Gaussian Filtering (Young and Van Vliet, 1995). The lower the RMSE, the higher the similarity between a pixel and its neighbors. In Xu et al. (2015), a Homogeneous Index (HI) is assigned to each candidate. Such HI is calculated by using the SID criterion between itself and its neighbors, and the maximum SID value is adopted. In the SSPM method (Kowkabi et al., 2016b), the values in cluster labels between the candidate and its neighbors are adopted as the homogeneity criterion, where the cluster label for each pixel is obtained by the *K*-means clustering algorithm. The basic idea of all the above mentioned methods is to perform spectral similarity measurements within a local neighborhood window by using a spectral distance metric, such as the Euclidean Distance (ED), Spectral Correlation Measure (SCM), SAD, and SID. However, there are some pending issues in those methods. For instance, in the AMEE method (Plaza et al., 2002), the properties of the spatial kernel (shape and size) strongly influence the MEI result. Although a progressively increased kernel size can be adopted, this will heavily increase the computational burden. The problem of determining an optimal neighborhood size also exists in other methods (Li and Zhang, 2011; Martín and Plaza, 2012; Xu et al., 2015; Kowkabi et al., 2016b). In addition, the setting of parameters is a difficult issue for all methods.

For better examining the spatially homogeneous regions in hyperspectral images, unsupervised clustering methods have been adopted in existing spatial-spectral unmixing models. Specifically, RBSPP (Martín and Plaza, 2011), SSPP (Martín and Plaza, 2012) and SSPM (Kowkabi et al., 2016b) apply unsupervised clustering algorithms [such as ISODATA (Ball and Hall, 1965), *K*-means, or the Hierarchical Segmentation (HSEG) algorithm (Tilton, 2003)] to segment the original or transformed image [e.g., by using the Principal Component Transform (PCT)] into a set of spectral clusters, each made up of one or more spatially connected regions. Compared with the SSEE (Rogge et al., 2007a), SPEE (Mei et al., 2010), HEEA (Li and Zhang, 2011), LLC (Canham et al., 2011) and the method in Xu et al. (2015), which all segment the original image into non-overlapping sub-blocks with a fixed size and shape, the partitions obtained by clustering algorithms generally exhibit better smoothness and coherence. However, in most of the aforementioned clustering algorithms, only spectral information is considered during the clustering process, and the size of segmented partitions is subject to diversity and inconsistencies. The SSPP method (Martín and Plaza, 2012), which includes the spatial information, however fuses the spatial and spectral information in a separate way. Therefore, how to naturally combine the spatial information during the clustering stage is a main difficulty for SSPP. In addition, other relevant issues such as the computational complexity of clustering algorithms, the determination of the number of clusters in advance, and the computational efficiency need further discussion.

In recent years, many superpixel-based segmentation methods, such as Graph-based Algorithms (Felzenszwalb and Huttenlocher, 2004), Turbopixel method (Levinshtein et al., 2009), and SLIC (Achanta et al., 2012), have been proposed. In essence, any image segmentation method can generate superpixels, and the aforementioned methods show faster speed and more accurate segmentation results than some traditional methods such as Mean-Shift-based (Vedaldi and Soatto, 2008) and Watershed-based approaches (Vincent and Soille, 1991). Superpixel methods have also been widely used in the hyperspectral imaging community; relevant work can be found in Thompson et al. (2010) and Saranathan and Parente (2016). In hyperspectral images, a superpixel represents a homogeneous region which contains a number of pixels that exhibit spatial continuity and spectral similarity. When

superpixels are used to represent spatial correlation information in hyperspectral images, compared with the fixed size and shape of window-based methods, they can represent adaptive spatial neighborhoods as expected in natural scenes, which generally exhibit arbitrary morphologies or sharp boundaries, thus reducing the sensitivity to noise and outliers and significantly improving the computational efficiency for subsequent processing tasks, such as endmember extraction and spectral unmixing. In Thompson et al. (2010) and Saranathan and Parente (2016), an efficient graph-based image segmentation algorithm (Felzenszwalb and Huttenlocher, 2004), which performs an agglomerative clustering of pixels as nodes on a graph such that each superpixel represents the minimum spanning tree of the constituent pixels, is adopted. This method adjusts well to natural image boundaries, but produces superpixels with very irregular sizes and shapes (Achanta et al., 2012). In Massoudifar et al. (2014), the well known Ultrametric Contour Map (UCM) algorithm is extended to hyperspectral images, conducting comparative experiments on the original hyperspectral image, the transformed image by PCT, and a monochromatic image, respectively. The results indicate that superpixel estimation on the transformed image is generally more efficient than other methods. The most important aspect of superpixel segmentation methods is how to determine the number of superpixels. In order to obtain a good superpixel representation, over-segmentation is generally adopted. In Saranathan and Parente (2016), a segment uniformity criterion is proposed to control the segmentation scale, which adopts a threshold to limit maximum variability inside one segment. The threshold is computed by a statistical analysis of the within-class and between-class spectral divergences of several endmember classes.

In Saranathan and Parente (2016), Thompson et al. (2010), and Massoudifar et al. (2014), superpixel segmentation methods have been shown to succeed when applied to hyperspectral imagery, and presented competitive results and significant improvements in the subsequent processing steps (particularly from a computational standpoint). However, the processing time of these superpixel segmentation methods increases as the size of the images become larger. As a result, finer segmentations are likely to result in significant computation times and sensitivity to noise. Based on this observation, in this work we have developed a new spatial preprocessing approach for hyperspectral unmixing. Compared with the aforementioned superpixel segmentation methods, the proposed approach is faster and requires less memory space. In addition, its computational complexity is not affected by the image size. A detailed description of our proposed method is given in the following section.

3. Regional clustering-based spatial preprocessing

In this section we describe a new method for spatial preprocessing which consists of two main steps. The first one is a clustering procedure which divides the original image into a set of homogeneous partitions. As opposed to the RBSPP (Martín and Plaza, 2011) and the SSPP (Martín and Plaza, 2012), we introduce a new efficient clustering strategy which naturally integrates the spatial and the spectral information contained in the data. In a second step, we extract candidate pixels from each partition. After these two steps, the candidate pixels are gathered together and fed to a spectral-based method to obtain the final endmembers and their corresponding abundances. An illustrative flowchart of the method is shown in Fig. 1. A detailed description of each step is given below.

3.1. Regional clustering

The proposed clustering strategy is inspired by the SLIC algorithm. The core idea of SLIC is to constrain the search scope from the whole image to a local region around the centroid. Compared with the SLIC
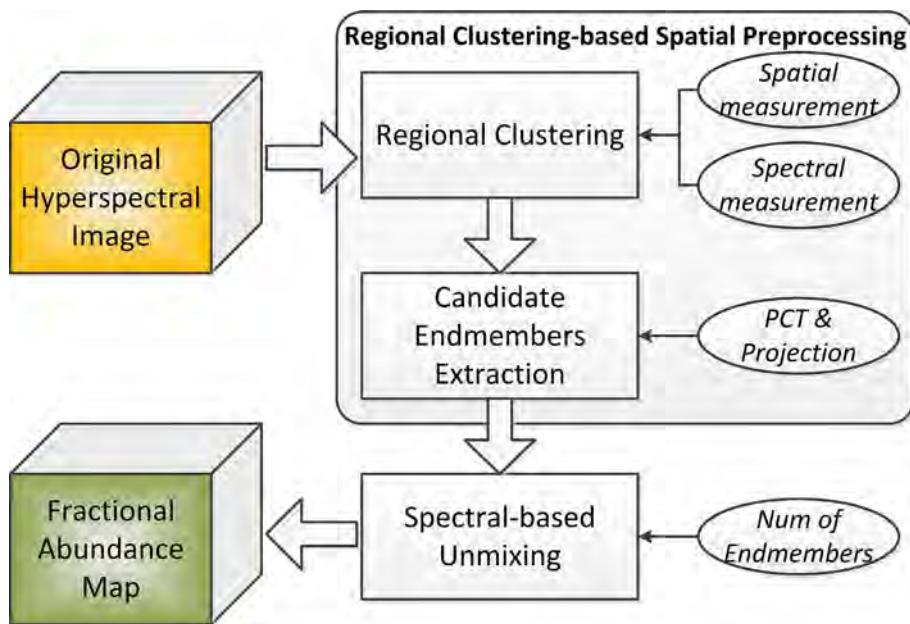
algorithm, our method introduces several differences:

- First and foremost, we adopt a moderate number of partitions based on the spatial resolution of the hyperspectral image. In the SLIC algorithm, the number of partitions is a predefined parameter. Generally, in order to obtain the lowest under-segmentation error, the SLIC algorithm tends to perform an over-segmentation, thus the number of partitions is set to a very large number. In our method, we only need to divide the original image into a series of homogeneous regions with high spectral similarity and spatial correlation. As a result, a very detailed segmentation is not required. According to the definition of Spatial Neighborhood Purity Index (SNPI) in the SPEE method (Mei et al., 2010), if SNPI of the generated cluster partitions is greater than a predefined threshold, the clustering resolution does not need further subdivision. This principle is adopted by our method in order to decide the number of partitions.
- Second, we use a different clustering criterion which is specifically designed for hyperspectral images (Shi and Malik, 2000; Gillis and Bowles, 2012). In the SLIC algorithm, a weighted 5-dimensional vector is adopted, which includes a 3-dimensional color vector and a 2-dimensional spatial vector. This 5-dimensional vector is modified in our work to work with hyperspectral data. In our strategy, the color distance is substituted by a spectral distance such as the SID-SAM.
- Most importantly, our newly developed strategy reduces the number of unsuccessfully clustered pixels and gives a special treatment to these pixels. In the SLIC algorithm, some pixels cannot be successfully clustered because of the over-segmentation. To address this issue, the SLIC performs a rough postprocessing to force those unsuccessfully clustered pixels into nearby superpixels. We emphasize that the proposed method does not apply the final superpixel removal step proposed in SLIC to keep outliers distinct. In our proposed method, the unsuccessfully clustered pixels are very few because of the moderate segmentation scale. Also, in hyperspectral images these pixels are often outliers or heterogeneous pixels. As a result, we add these pixels into the set of candidate pixels for subsequent endmember identification purposes.

Next, we will give a detailed description of our proposed regional clustering strategy.

In the initialization step, we first choose the initial cluster centers.

Like the SLIC algorithm, we also adopt an equidistant sampling process. For example, if the spatial size of the hyperspectral image is 100×100 and the number of initial clusters is 100, we divide the original image into 100 sub-blocks with the same size of 10×10 pixels and the initial cluster centers are selected from pixels closest to the center position of each sub-block, e.g. anyone from the coordinates {(5,5),(5,6),(6,5),(6,6)}, as illustrated in Fig. 2 (a). Then, in order to restrict the clustering search scope to a local neighborhood around each cluster center, we set the neighborhood as a rectangle around each cluster center with width m and height n. For each cluster center we only search for the surrounding $m \times n$ pixels, hence significantly reducing the computational complexity as compared to conventional clustering algorithms (that perform the search over the entire image) [see Fig. 2 (b)]. At this point, it is important to emphasize that how to choose the neighborhood size is very important for the algorithm. The larger the values of m and n, the greater the computational complexity of the clustering process. However, if the values of m and n are too small, many suitable neighbors will not be assigned to the adjacent clusters. In our experiments, we have empirically tested that a value that represents two times the size of the initial image sub-block is generally required for m and n, as shown in Fig. 2 (c). For values of m and n greater than two times, we did not find obvious improvements. Although this choice cannot provide a precise clustering for all pixels, from a local point of view the generated cluster partitions exhibit good homogeneity which meet the requirements for the subsequent extraction of candidate pixels.

In the main loop step, for each cluster center the algorithm adopts a distance criterion to cluster each pixel in the local neighborhood. In order to integrate the spectral and the spatial information simultaneously, we considered several distance metrics in our experiments, including the Euclidean Distance (ED), SAD, SID, and SID-SAM (Du et al., 2004). After different experiments, we choose the SID-SAM as our spectral distance measurement as it integrates the variability and similarity of pixels in the image. For the spatial distance measurement, we use the spatial Euclidean Distance (ED) to account for the distance to the considered cluster center. The clustering criterion is the weighted sum of the spectral and spatial measurements. If we denote by $x_i = (x_{i1}, x_{i2}, \cdots, x_{il})$ a given pixel with spatial coordinate (ix, iy) in the hyperspectral image and by $x_c = (x_{c1}, x_{c2}, \cdots, x_{cl})$ the corresponding cluster center with spatial coordinate (cx, cy), the clustering criterion $D_{i,c}$ can be simply defined by the following distance:
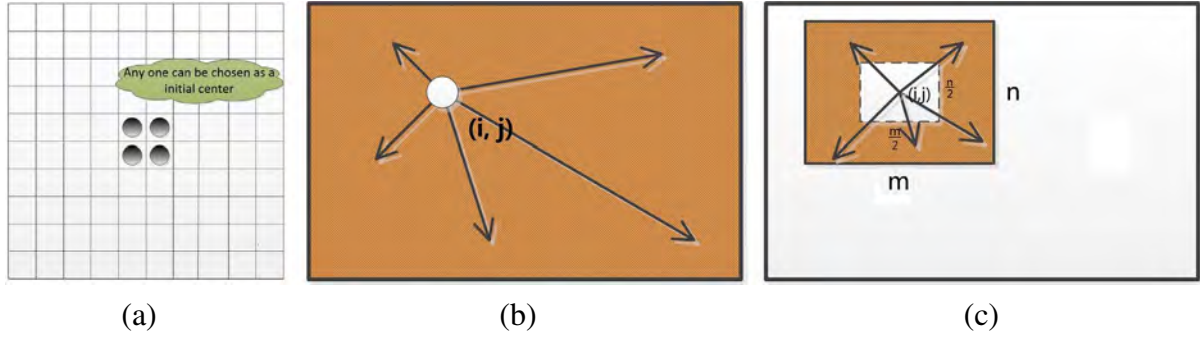
Fig. 2. An illustration of the proposed regional clustering procedure. (a) Choice of the initial cluster center. (b) Global search strategy implemented by conventional clustering algorithms. (c) Local search strategy implemented by our regional clustering procedure.
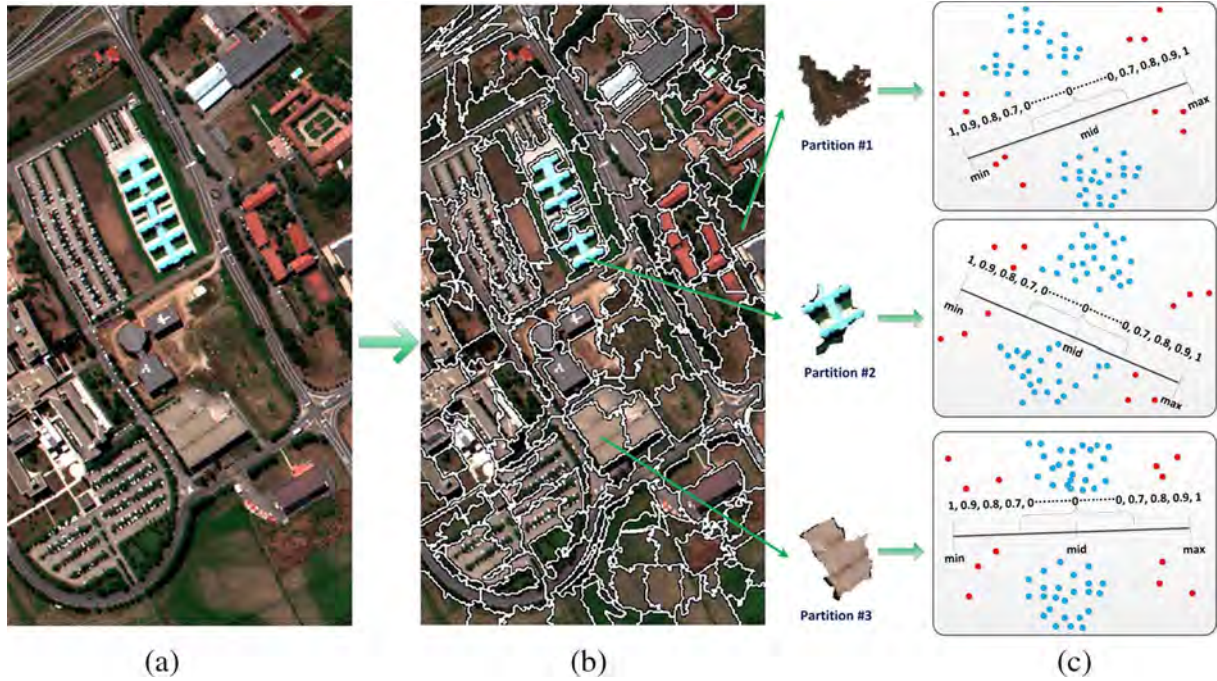

Fig. 3. (a) False color composition of the ROSIS Pavia University scene. (b) Results of the proposed regional clustering strategy (the image is divided into 183 homogeneous partitions). (c) Illustrative example of the extraction of candidate endmembers using three different partitions. Here, the min-max axis is the first principal component vector of the partition, **min** and **max** denote two extreme projection positions, the pixels in red have the weight value ≥ 0.7, and the pixels in blue have the weight value < 0.7. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$$D_{i,c} = (1-\lambda) \times D_{spectral} + \lambda \times D_{spatial},$$
$$D_{spectral} = SID(x_i, x_c) \times \tan(SAM(x_i, x_c)),$$
$$D_{spatial} = \sqrt{(ix-cx)^2 + (iy-cy)^2} \;/\; r, \tag{1}$$

where $D_{spectral}$ denotes the spectral distance defined by the SID-SAM measurement, $SID(x_i, x_c)$ denotes the SID measurement and $SAM(x_i, x_c)$ denotes the SAM measurement as defined in Eqs. (2) and (3), respectively, where $p_{ik} = x_{ik} / \sum_{k=1}^{l} x_{ik}$, and $p_{ck} = x_{ck} / \sum_{k=1}^{l} x_{ck}$. $D_{spatial}$ denotes the spatial Euclidean Distance with normalization, r is the length of the diagonal for the search neighborhood which is used to normalize the spatial ED, λ is a weighting factor used to balance the spectral and spatial terms.

In our proposed regional clustering strategy, as the search scope is restricted to a local neighborhood around each cluster center, the spatial correlation is locally constrained. Therefore, the spectral information plays a more important role than the spatial information. The balance parameter λ is then suggested to be less than 0.5. Taking into account the

higher the spatial resolution, the greater the impact of the spatial information, we empirically set a small λ for images with coarse spatial resolutions, and a large λ for images with fine spatial resolutions.

$$
\begin{aligned}
SID(x_i, x_c) &= D(x_i \| x_c) + D(x_c \| x_i) \\
&= \sum_{k=1}^{l} p_{ik} \log\left(\frac{p_{ik}}{p_{ck}}\right) + \sum_{k=1}^{l} p_{ck} \log\left(\frac{p_{ck}}{p_{ik}}\right) \\
&= \sum_{k=1}^{l} (p_{ik} - p_{ck}) \log\left(\frac{p_{ik}}{p_{ck}}\right)
\end{aligned}
\tag{2}
$$

$$SAM(x_i, x_c) = \cos^{-1}\left(\frac{<x_i, x_c>}{\|x_i\| \bullet \|x_c\|}\right) \tag{3}$$

After each iteration, the new cluster centers are updated for the next iteration. Here, we also give a detailed pseudocode illustrating the adopted regional clustering strategy in Algorithm 1.

Algorithm 1. Pseudocode of the proposed regional clustering strategy

Input: Hyperspectral image data set, $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_{np}) \in \mathbb{R}^{l \times np}$ with l bands and np pixels;
   Initial number of clusters, $cnum$;
   Weight value for the distance metric, λ;
   Number of iterations, $iter$;
Output: Cluster label for each pixel, **labels** $= (lb_1, lb_2, ..., lb_{np})$;

Step 1: Initialization.
   Define equidistant sampling clustering centers as c_k, save index of clustering centers as
   c_idx_k and spectral signatures of clustering centers as $\mathbf{c\_spec}_k$, where $k \in [1, cnum]$;
   Set the width and height of the initial sub-block as w, h;
   Set the size of the local search neighborhood to $m = 2 * w, n = 2 * h$;
   Set the initial cluster label for each pixel, $labels(i) = -1, i \in [1, np]$;
   Set the initial distance for each pixel, $distance(i) = +\infty, i \in [1, np]$;
Step 2: Main Loop.
   for $it = 1$ to $iter$ **do**
      for each cluster center c_k **do**
         for each pixel j in a $m \times n$ neighborhood region around c_k **do**
            Compute the spectral distance between pixel j and c_k, $D_{spectral}$;
            Compute the spatial distance between pixel j and c_k, $D_{spatial}$;
            $D_{j,c_k} = (1 - \lambda)D_{spectral} + \lambda D_{spatial}$;
            if $D_{j,c_k} < distance(j)$ **then**
               $distance(j) = D_{j,c_k}$;
               $label(j) = k$;
            end if
         end for
      end for
      Update new cluster centers;
   end for

Fig. 3 illustrates the performance of the proposed regional clustering strategy. Here we use the well-known Pavia University hyperspectral data set, collected by the ROSIS sensor over the University of Pavia, Italy. The image size in pixels is 610×340, with very high spatial resolution of 1.3 m per pixel. The number of spectral channels in the acquired image is 103 (with spectral range from 0.43 to 0.86 μm). After applying the Minimum Noise Fraction (MNF) (Green et al., 1988) transformation, we obtained a reduced image with 20 bands. Fig. 3 (a) shows a false color composite of this image, while Fig. 3 (b) shows the results of the proposed regional clustering with 183 partitions.

3.2. Extraction of candidate endmembers

According to the description given in the previous subsection, the output from the clustering step is a set of partitions and a few unsuccessfully clustered pixels. In this step, we need to select a subset of pixels that will be used for endmember identification purposes. These candidate pixels should contain the most spectrally pure signatures in the image. To achieve this goal, two strategies can be adopted. One is to filter the candidate pixels as implemented by the SPEE algorithm (Mei et al., 2010), which uses the SNPI to select the most spectrally pure partitions (i.e., those with a SNPI that is above a predefined threshold). Then, candidate pixels can be selected from these filtered partitions. Another strategy is to choose spectrally pure pixels on all available partitions, since these partitions represent spatially homogeneous areas in the image. Our proposed method adopts this second strategy. As suggested in Martín and Plaza (2012), for each partition we perform a Principal Component Analysis (PCA) and choose the first q principal components as the projection axis. Then, we project all pixels on these principal components. For each projection axis, we assign a weight of 1 to the pixels with extreme projection positions, and mark them as *max* and *min* projection values; then, the weight of the other pixels is calculated as $w_i = (proj_i - min)/(max - min)$, where $proj_i$ is the projection value of pixel i. If $w_i \geq 0.7$, the final weight is w_i. Otherwise if $w_i \leq 0.3$, the final weight is $1 - w_i$. Further, we assign a weight of 0 to those pixels which meet the condition $w_i \in (0.3, 0.7)$. Finally, the sum of all the weights for each pixel is tallied as its spectral purity index. By defining a percentage of pixels that will be selected per partition, we

obtain the final candidate pixels from each partition according to their spectral purity index. Fig. 3 (c) illustrates the performance of the extraction of candidate endmembers step using an example based on three different partitions. Here, the *min-max* projection axis is the first principal component vector in each partition, the red points denote the pixels with the weights in the interval [0.7,1], and the blue points denote the pixels with the weights in the interval [0,0.7].

3.3. Extraction of final endmembers

Finally, endmember identification and abundance estimation algorithms can be applied to the set of retained candidate pixels after the procedure described in the previous subsection to obtain the final endmembers and their corresponding abundances. Compared with other spatial preprocessing techniques such as SPP (Zortea and Plaza, 2009), RBSPP (Martín and Plaza, 2011) and SSPP (Martín and Plaza, 2012), the proposed method offers two significant improvements: 1) the computational complexity is much lower, mainly because we adopt a regional clustering strategy that restricts the search scope to a small neighborhood window around each cluster center, thus greatly accelerate the clustering speed; and 2) the proposed method is highly adaptive by means of several parameters that can be configured depending on the characteristics of the hyperspectral image, including the number of partitions, percentage of candidate pixels, spectral and spatial distance criteria, etc. In the following section we perform extensive experiments using both synthetic and real hyperspectral images to validate the proposed method. We also provide a quantitative assessment of our method with state-of-the-art endmember extraction algorithms.

4. Experimental results

This section is organized as follows. Section 4.1 describes data sets, including synthetic and real hyperspectral data. Section 4.2 presents a comparative assessment of the segmentation performance using the proposed regional clustering strategy and graph-based segmentation methods. Section 4.3 presents a quantitative assessment of our spatial preprocessing methods using synthetic and real hyperspectral data. Real hyperspectral image data is collected by the Hyperspectral Digital Imagery Collection Experiment (HYDICE). Section 4.4 presents a comparative assessment of the adaptivity of the method to spectral variability in the scene by analyzing the results obtained by our method with regards to those provided by other state-of-the-art algorithms such as Sparse Unmixing by Variable Splitting and Augmented Lagrangian (SUnSAL) introduced in Iordache et al. (2011) and its collaborative version (CLSUnSAL) (Iordache et al., 2014). Section 4.5 analyzes the sensitivity of our proposed method to different parameter settings, including the number of partitions, clustering criteria, search scope, etc. Finally, Section 4.6 analyzes the computational complexity of our spatial preprocessing methods. All our experiments have been conducted by using Matlab R2013a in a desktop PC equipped with an Intel Core i7 CPU (at 3.6 GHz) and 32 GB of RAM.

4.1. Hyperspectral data sets

4.1.1. Synthetic data

The synthetic hyperspectral image used in the study is generated by using fractals, which are often adopted to simulate distinct spatial patterns found in nature. First, a set of endmember signatures are randomly chosen from the USGS spectral library. Then, linear mixtures of different endmembers are constructed to form the initial fractal image. This initial fractal image is further divided into a number of clusters by using the conventional *K*-means clustering method, where the number of clusters is larger than the number of endmember signatures used in the image. In our experiments, the number of endmembers is fixed 9 (Fig. 4 shows the reflectance curves of the USGS
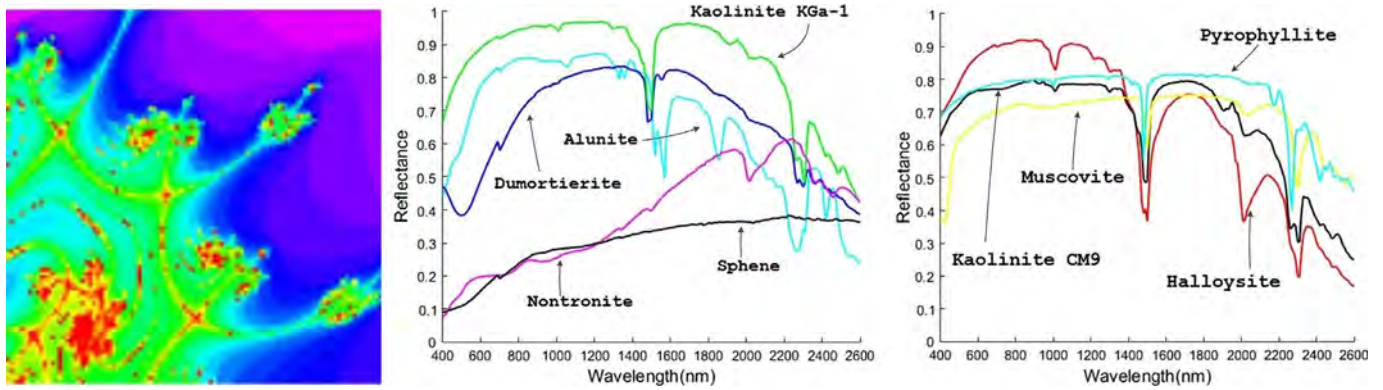
Fig. 4. Fractal 1 synthetic hyperspectral image and the corresponding USGS library signatures used to construct the Fractal 1 image.

signatures used in the Fractal 1 synthetic image). Then, a Gaussian filter is applied to generate the abundance proportions of pixels in each cluster partition, so that the pixels that are closer to the border of partition are highly mixed, while the pixels located at the center of the partition are spectrally pure in nature. The fractional abundances of each pixel meet the abundance non-negativity (ANC) and sum-to-one (ASC) constraints, thus ensuring that all pixels strictly adhere to a fully constrained linear mixture model. Finally, zero-mean Gaussian noise is added to the synthetic image in different signal-to-noise ratios (SNRs), from 30:1 to 90:1, to simulate contributions from ambient and instrumental sources, following the procedure described in Harsanyi and Chang (1994).

4.1.2. HYDICE data

The real hyperspectral image used in experiments is an urban scene, which was acquired by HYDICE in October 1995. This urban scene is located at Copperas Cove, TX, US, and is available online from the following website: http://www.tec.army.mil/Hypercube. The portion used in our experiments comprises a total of 307 × 307 pixels and 210 spectral bands, with spectral resolution of 10 nm, covering the wavelength range from 0.4 to 2.5 μm. The spatial resolution is $2 \times 2 \, m^2$ per pixel. Considering the dense water vapor and atmospheric effects, bands 1–4, 76, 87, 101–111, 136–153 and 198–210 were removed, leaving a total of 162 bands to be used in the experiments. Fig. 5 (a) shows a false color composition of the image.

This region contains a mixture of man-made objects and forestry. The top of the image contains a highway that crosses the region from left to right, a shopping mall along the highway, and a parking lot in front of the mall. Some roads, houses, grass, and trees are scattered in various areas of the scene. In addition, due to the low solar altitude at the time of data collection, trees and houses cast long shadows on the ground. Here we choose four classes of endmembers: asphalt, roof, grass, and tree. However, some phenomena should be clarified. For the roof target, there are mainly two classes. One is bright, while the other one resembles the shadows on the ground. For the road target, most of them, including the highway and the parking lot, are asphalt paved, while a few others are made by concrete. Trees and grass also exhibit some differences. In order to assess the methods quantitatively, we extract the reference endmember spectra manually from the scene. Using the asphalt class as an example, we first pick out 20 pixels in the region defined by spatial coordinates [75,226] to [82,234] as the asphalt spectrum [see Fig. 5 (a)]. As mentioned before, this region comprises a parking lot, which belongs to the asphalt class. These asphalt spectra are averaged to obtain a reference spectral signature (Jia and Qian, 2009). It should be noted that the SAD between this reference spectral signature and the asphalt spectrum in the spectral library available in ENVI is less than 0.05, which indicates that the obtained reference signature is reliable. The same has been done for other classes. Fig. 5 (b) shows the obtained reference spectral signatures. As we can see in this figure, the signatures of grass and trees exhibit very similar spectral shape, which makes their separation difficult. In addition, using the obtained reference signatures as endmembers, the fractional abundances are estimated by the Fully Constrained Least Squares (FCLS) method. Fig. 6 shows the extracted fractional abundance maps of the four reference endmember classes. In Fig. 6, the grayscale values (from black to white) represents abundance values (from 0 to 1).

4.2. Experiments with the segmentation strategy

In this set of experiments, we performed a comparison of the segmentation performance for the hyperspectral images. The chosen
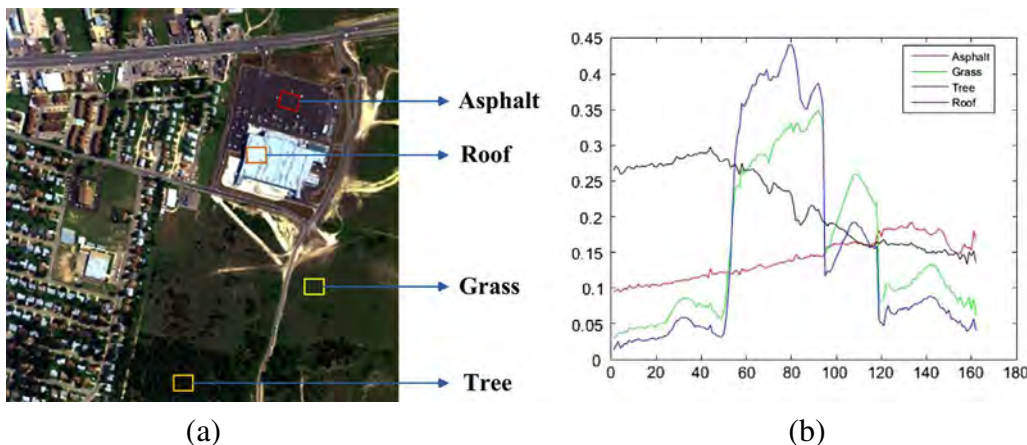

(a)

(b)

Fig. 5. (a) Urban HYDICE scene collected at Copperas Cove, TX, US. Four endmembers are considered in our experiments. For the reference endmember spectra, we take 20 pixels in each rectangle block, and calculate the average spectrum of these pixels in order to obtain a reference spectral signature. (b) The obtained endmember spectral signatures considered in experiments.
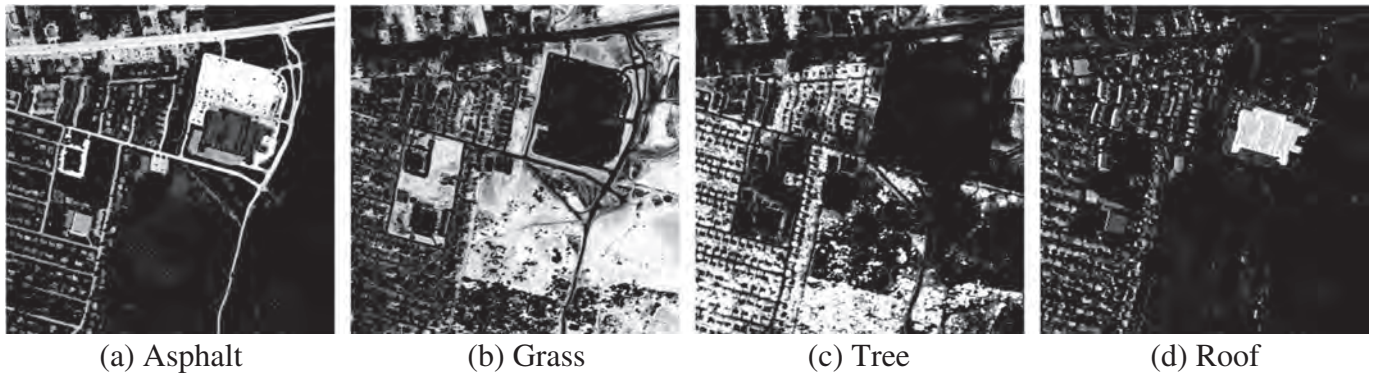
|(a) Asphalt|(b) Grass|(c) Tree|(d) Roof|

Fig. 6. The fractional abundance maps of the four reference spectral signatures in the urban HYDICE image, which is estimated by the FCLS method. In these images, pure black denotes that the abundance of an endmember class in this pixel is 0, while pure white denotes an abundance of 1.
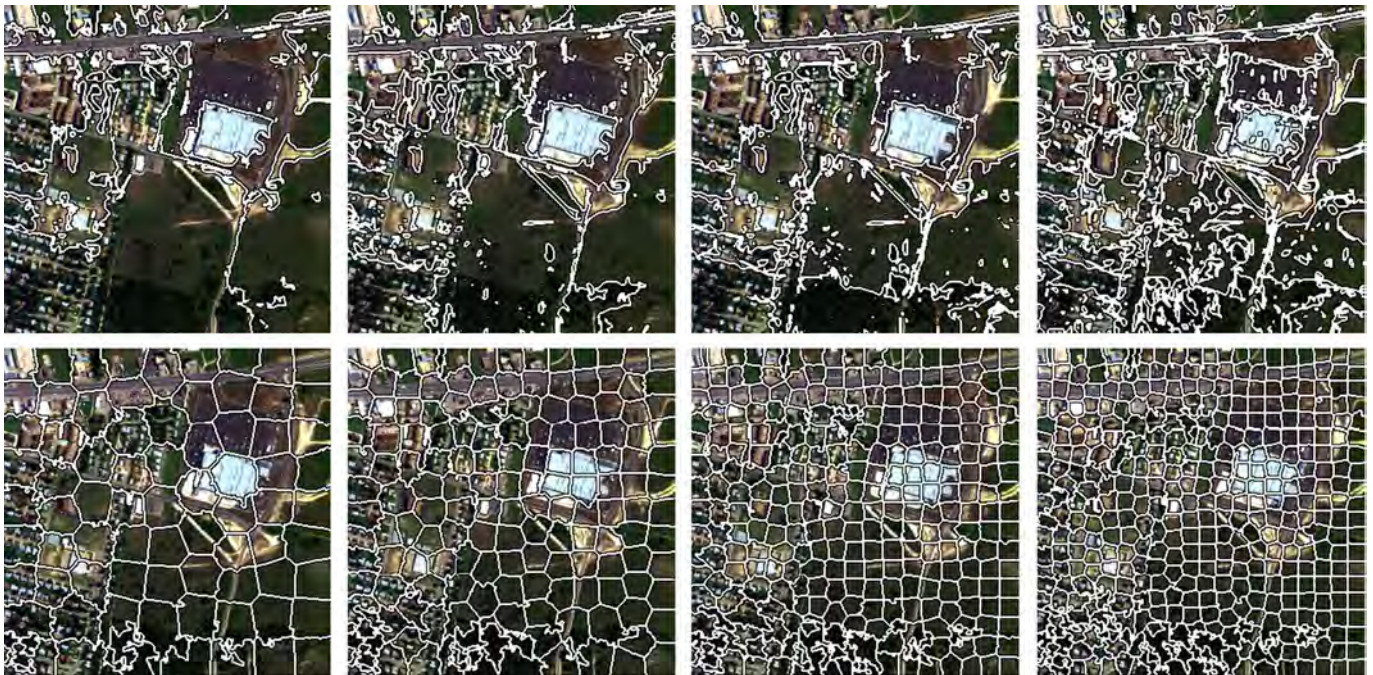

Fig. 7. Segmentation results for the urban HYDICE scene. The first row shows the results obtained by the graph-based segmentation approach (Felzenszwalb and Huttenlocher, 2004). The images in this row contain 107, 198, 287 and 424 partitions, respectively. The second row shows the results obtained by the regional clustering approach. The images in this row contain 77, 159, 285 and 398 partitions, respectively.

approaches include our regional clustering strategy and the graph-based segmentation developed by Felzenszwalb and Huttenlocher (2004), which is a highly efficient segmentation algorithm with $O(n_p \log(n_p))$ complexity, with n_p being the number of pixels in the image. This method has been used in the past to analyze hyperspectral images (Thompson et al., 2010; Saranathan and Parente, 2016). Fig. 7 illustrates the segmentation results obtained for the urban HYDICE scene with different levels of detail. As it can be seen in this figure, the partitions from the graph-based approach show obvious diversity in terms of both size and shape, with many partitions containing less than 5 pixels, while other partitions containing a large number of pixels. Moreover, the number of partitions is hard to establish in advance by this approach. Relatively speaking, the regional clustering strategy can generate more regular partitions, and the number of partitions is easy to control. Most importantly, the obtained partitions naturally present local spatial correlation, which can benefit the subsequent candidate endmember extraction in such local areas.

Fig. 7 shows a visual comparison of the segmentation results. Here we also provide a quantitative assessment by using some clustering validity indices, aiming at measuring both the compactness and the

separation of clusters in the partitioning, where high compactness indicates that pixels in a cluster are close to each other, while high separation indicates that clusters are widely separated. In Legány et al. (2006) and Li et al. (2016), several clustering validity indices are introduced and compared to each other. Here, we choose the *SD* Validity Index (Halkidi et al., 2000) to assess the segmentation performance, which combines both compactness and separation. At the same time, we adopted two 'SNPI' (Mei et al., 2010) metrics (including 'AvgPI' and 'SV DPI') to focus on the measurement of the spatial correlation within each partition. Next, we will first give the definition of these indices, and then we will analyze the segmentation performance by using these indices.

- The *SD* validity index is defined as $SD = \alpha \cdot Scatt + Dis$, where *Scatt* is the average scattering of clusters, defined by

$$Scatt = \frac{\frac{1}{c}\sum_{i=1}^{c}\|\sigma(\nu_i)\|}{\|\sigma(x)\|},\qquad(4)$$

where $\sigma(x)$ is the variance of the entire data set, and $\sigma(\nu_i)$ is the

variance of the ith cluster, and c is the number of clusters. Dis is the total separation of clusters, defined by

$$Dis = \frac{\max\limits_{i,j=1\ldots c}\left(\left\|\nu_j - \nu_i\right\|\right)}{\min\limits_{i,j=1\ldots c}\left(\left\|\nu_j - \nu_i\right\|\right)} \sum_{i=1}^{c}\left(\sum_{j=1, j\neq i}^{c}\left\|\nu_j - \nu_i\right\|\right)^{-1}.$$
(5)

Finally, α is a weighting factor that is equal to the Dis value in case that we have a maximum number of clusters. A lower SD index means better cluster configuration.

- Two types of 'SNPI' metrics are used. One is called '$AvgPI$', which is formulated as $AvgPI = \sum_{i=1}^{k} dist(x_i, \bar{x})/k$, where $\bar{x}$ is the average value of pixels in a partition, k is the number of pixels in this partition, and $dist(x_i, \bar{x})$ is the distance of x_i to $\bar{x}$. Here we adopted the correlation coefficient as the distance measurement, which is defined as $dist(x_1, x_2) = x_1^T x_2/(\sqrt{x_1^T x_1}\sqrt{x_2^T x_2})$. The range of the correlation coefficient is $[-1,1]$. The greater the correlation coefficient is, the more similar/relevant the two variables are. The '$AvgPI$' quantitatively measures the average variance of all pixels in a partition to their average value, which is also in the range $[-1,1]$. A larger '$AvgPI$' denotes better spatial homogeneity. The other metric is called '$SV DPI$', which is formulated as $SVDPI = S_{11}/(\sum_{i=1}^{k} S_{ii})$, where S_{11} is the largest singular value computed by the Singular Value Decomposition (SVD) and S_{ii} is the ith singular value. The '$SV DPI$' represents the domination of the first feature component. A larger '$SV DPI$' denotes that the partition exhibits better spatial purity.

Table 1 illustrates the results of the aforementioned validity indices with different segmentation levels for the urban HYDICE image. As we can see in the table, if we only consider the compactness (such as $Scatt$ index value in Table 1), we obtain better compactness for small segmentation scales. However, the separability increases when the segmentation scale is moderate, which can be seen from the Dis index value. By considering both the compactness and separability we can observe that, when the segmentation scale is 100, the value of SD index is smaller, which means better cluster configuration. As a result, we suggest to choose a moderate segmentation scale in the experiments. This is also reflected by the minimum value of the '$AvgPI$' and '$SV DPI$' metrics, where the moderate segmentation scale of 100 resulted in better spatial purity than the other cases (with a greater minimum value in '$AvgPI$' and '$SV DPI$').

4.3. Comparative experiments to evaluate RCSPP

4.3.1. Experiments with synthetic data

In our first experiment with the synthetic images, we used the SAD and the Root Mean Square Error (RMSE) to assess the performance of different endmember extraction methods. Specifically, the SAD is used to compare the spectral similarity between the extracted endmembers and their corresponding reference signatures from the USGS library (the lower the SAD, the more similar the extracted endmembers). The RMSE is used to compare the original image and the reconstructed one, which

is generated under the linear mixing model by using the extracted endmembers and their corresponding fractional abundances. In this work, the fractional abundances are estimated by the FCLS method. Table 2 shows the SAD (in radians) and the RMSE calculated by different spectral-based endmember extraction methods, combined with our spatial preprocessing methods. For the proposed Regional Clustering-Based Spatial Preprocessing (hereinafter denoted as RCSPP), we set the number of partitions to 25, the weighting value of $\lambda = 0.1$ and the percentage of candidate pixels to 20% of the total number of pixels in each partition. The parameters involved in the other methods used for comparison were carefully optimized, so that the best performance results are reported in each case. The number of endmembers to be extracted was set in all cases to 9, according to our previous knowledge on the simulation process in which 9 USGS signatures were used to simulate each scene. Nevertheless, the number of endmembers estimated by the Hyperspectral Signal Identification with Minimum Error (HySime) (Bioucas-Dias and Nascimento, 2008) is also 9, while the Virtual Dimensionality (VD) (Chang and Du, 2004) algorithm provided an estimation of 8 endmembers. We would like to emphasize that the synthetic data sets are generated based on the linear mixing model, with the ultimate goal of assessing the proposed method in a fully controlled environment. For the linear model, many of the existing methods can obtain very good results, such as VCA and MVSA. Here we only choose OSP and N-FINDR (as highly representative algorithms) for evaluation purposes. From Table 2 we can conclude that, as a preprocessing for the OSP and N-FINDR, the proposed RCSPP can provide comparable results to other preprocessing algorithms, especially in low SNR circumstances, with the advantage that a much smaller number of pixels needs to be used in the process. As a result, we can conclude that the combination of spatial preprocessing with spectral-based methods generally exhibits good results in the task of endmember extraction.

4.3.2. Experiments with the urban HYDICE data set

In this experiment, we use the urban HYDICE data set to assess the performance of the proposed RCSPP method in comparison with other algorithms. For the RCSPP, we obtained 77 partitions after segmentation, with $\lambda = 0.4$, and the percentage of candidate pixels was reduced to only 0.5% of the total number of pixels in each partition. The parameters in the other methods used for comparison were carefully optimized. For the number of endmembers, both VD (Chang and Du, 2004) and HySime (Bioucas-Dias and Nascimento, 2008) obtained the same estimation of 9 endmembers. Considering that the estimated number of endmembers is 9 and the number of reference endmembers is 4, we used the Earth Movers Distance (EMD) (Zare and Anderson, 2013) to simultaneously compare the estimated endmember set and the corresponding fractional abundances with the available reference endmember set. EMD is a flexible distance metric between two sets of distributions (Rubner et al., 2000). In Zare and Anderson (2013), the EMD is adopted as a method to compare two endmember sets and the corresponding fractional abundances. For a single pixel, let us assume that $h_1 = \{E, p\}$ and $h_2 = \{B, q\}$, where $E \equiv [e_1, e_2, \ldots, e_M]$ is the first endmember set with e_i denoting the spectral signature of the ith endmember in this set, and p denoting the corresponding abundance vector. $B \equiv [b_1, b_2, \ldots, b_N]$ is the second endmember set, and q is the

Table 1
Assessment of the SD and 'SNPI' indices with different segmentation levels for the urban HYDICE image. The first row tabulates the segmentation levels, where we indicate the preset segmentation scale and (in the parentheses) the actual segmentation scale. For the two 'SNPI' indices we also give the minimum and maximum values reached in each segmentation level.

Indices	100 (77)	200 (159)	300 (285)	400 (398)
$Scatt$	0.7942	0.7424	0.6814	0.6425
Dis	0.0741	0.1534	0.17	0.354
SD	**0.3552**	0.4162	0.4112	0.5814
$AvgPI$	[0.9417,0.9982]	[0.9292,0.9994]	[0.9346,0.9995]	[0.9356,0.9996]
$SV DPI$	[0.6008,0.8264]	[0.5986,0.8511]	[0.6002,0.8654]	[0.5994,0.8667]

The entry value in bold presents the best segmentation level.

Table 2
SAD (in radians) and RMSE after reconstructing the Fractal 1 synthetic hyperspectral image. SAD is measured between the USGS mineral spectra and the corresponding endmember pixels obtained by different endmember extraction methods. RMSE is measured between the original and the reconstructed version of the Fractal 1 image using different endmember extraction algorithms and their corresponding fractional abundances, estimated by the FCLS method.

Algorithm	SNR = 30:1		SNR = 50:1		SNR = 70:1		SNR = 90:1	
	SAD	RMSE	SAD	RMSE	SAD	RMSE	SAD	RMSE
OSP	0.038	0.233	0.014	0.0252	0.0134	0.006	0.0136	0.0048
N-FINDR	0.0346	0.228	0.0135	0.0245	0.013	0.0048	0.0131	0.0036
RCSPP + OSP	0.0367	0.231	0.0142	0.0262	0.0133	0.0065	0.0132	0.0052
RCSPP + N-FINDR	0.0351	0.23	0.0137	0.0266	0.0131	0.0073	0.0132	0.0076

corresponding abundance vector. Our goal is to find an optimal flow $F = [f_{ij}]$ (with f_{ij} denoting the flow between p_i and q_j) that minimizes the objective function

$$WORK(E, B, p, q, f) = \sum_{i=1}^{M} \sum_{j=1}^{N} d_{ij} f_{ij}, \qquad (6)$$

with respect to the following constraints:

$$f_{ij} \geqslant 0 \qquad 1 \leqslant i \leqslant M, 1 \leqslant j \leqslant N,$$

$$\sum_{j=1}^{N} f_{ij} \leqslant p_i \qquad 1 \leqslant i \leqslant M,$$

$$\sum_{i=1}^{M} f_{ij} \leqslant q_j \qquad 1 \leqslant j \leqslant N,$$

$$\sum_{i=1}^{M} \sum_{j=1}^{N} f_{ij} = \min\left(\sum_{i=1}^{M} p_i, \sum_{j=1}^{N} q_j\right). \qquad (7)$$

In Eq. (6), d_{ij} denotes the ground distance between endmember e_i and endmember b_j. Here we adopt SAD as the distance measurement. Once the optimal f^* is found, the EMD is calculated as

$$EMD(E, B, p, q, f^*) = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} d_{ij} f_{ij}^*}{\sum_{i=1}^{M} \sum_{j=1}^{N} f_{ij}^*}. \qquad (8)$$

In order to measure the dissimilarity of all pixels in the hyperspectral data set, we adopted an appropriate summation of the EMD values over all pixels, as defined in Zare and Anderson (2013):

$$EMD_{tot}(E, B, P, Q, F^*) = \sum_{k=1}^{N_x} EMD(E, B, p_k, q_k, f_k^*), \qquad (9)$$

where N_x is the total number of pixels in the hyperspectral data set.

Table 3 presents the total EMD (EMD_{tot}) of all pixels for any two different methods. In Table 3, 'Ref' denotes the reference endmember set and its corresponding fractional abundances. As shown by the results reported in Table 3, relative to the reference endmember set, the RCSPP based methods (RCSPP + OSP, RCSPP + N-FINDR) have smaller EMD_{tot}, indicating that the RCSPP based methods can find more spectrally similar endmembers (with regards to the reference spectral signatures) than the spectral-based methods.

4.4. Analysis of the spectral variability

Due to variable illumination, atmospheric interferers and other

Table 3
EMD_{tot} values ($\times 10^4$) between the reference spectral signatures and the endmembers obtained by different methods for the urban HYDICE scene, which is calculated by Eq. (9). The four reference spectral signatures are obtained as described in Section 4.1.2.

Algorithm	OSP	RCSPP + OSP	N-FINDR	RCSPP + N-FINDR
	(M = 9)	(M = 9)	(M = 9)	(M = 9)
Ref (M = 4)	4.391	3.1571	4.251	**2.8782**

The entry value in bold presents the best segmentation level.

environmental conditions, spectral variability inevitably exists in real hyperspectral scenes. In order to evaluate the quality of the proposed RCSPP in the presence of this phenomenon, we conduct a set of experiments to analyze its sensitivity to spectral variability. In our experiments, considering that the knowledge about the different materials in the scene is often not available in advance, we consider approaches able to autonomously estimate the endmember sets and the corresponding fractional abundances simultaneously. Here, we choose SUnSAL and its collaborative version (CLSUnSAL) as our reference unmixing approaches. These two sparse unmixing approaches can be applied using large spectral libraries. In our RCSPP algorithm, the candidate endmember set (obtained in Step 2 of the algorithm) can be used to generate this kind of spectral library. At the same time, in order to quantitatively evaluate the performance of RCSPP, we compared the results with a previously developed Automated Endmember Bundles (AEB) method in Somers et al. (2012). For the AEB method, we randomly choose several sub-images from the original image, and then use a spectral-based endmember extraction method to acquire an endmember set in each sub-image. Then, all endmembers are gathered together and fed to SUnSAL or CLSUnSAL for unmixing purposes.

Fig. 8 shows the error maps obtained using the AEB method and our proposed RCSPP. For the AEB method, we randomly choose 20 subblocks with a size that is 10% the size of the original image, ending up with 180 candidate endmembers (each sub-block consists of 9 endmembers). For the RCSPP, as mentioned before the number of partitions is 77, and the percentage of candidates selected in each partition is 0.5%. Thus, for the RCSPP a total of 513 candidate endmembers are used for constructing the endmember bundles. From Fig. 8, we can observe that the errors obtained by the proposed RCSPP are significantly lower than those provided by the AEB method, and the reconstructions achieved by the RCSPP exhibit much less error than those obtained by the AEB. This is considered as a good indicator that the spectral bundles estimated by the proposed RCSPP are better than those provided by the AEB.

Finally, Fig. 9 shows the bundles obtained by the different methods for the urban HYDICE scene. Here, we simply used the *K*-means clustering method with ED distance and set the number of clusters to 9 to construct bundles for four classes: 'Asphalt', 'Grass', 'Tree', 'Roof', in order to provide a visual assessment of the bundles constructed by the two considered methods. It can be observed that man-made objects such as 'Asphalt' and 'Roof' exhibit higher spectral variability than the natural endmembers. Table 4 tabulates the SAD of the obtained bundles. As can be seen in Table 4, for the proposed approach the variation of the two natural classes ('Grass' and 'Tree') is much smaller than that of the two man-made classes ('Asphalt' and 'Roof'). Similar observations can be made for the AEB method, revealing that our proposed regional clustering strategy also has the ability to perform similarly with regards to a method that is specifically designed to model spectral variability, such as the AEB.

4.5. Sensitivity to input parameters

To evaluate the sensitivity of the proposed RCSPP method to input parameters, we performed a series of tests using the Fractal 1 synthetic
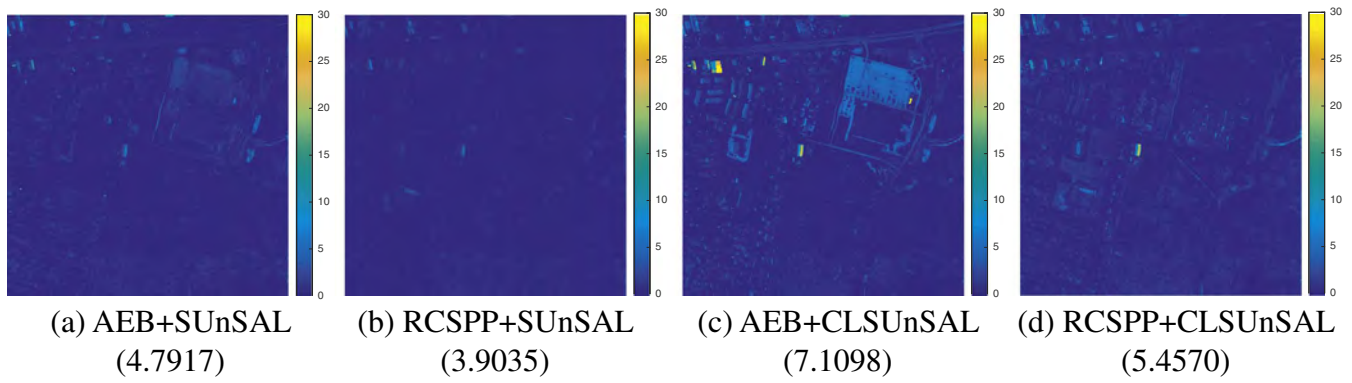
(a) AEB+SUnSAL   (b) RCSPP+SUnSAL   (c) AEB+CLSUnSAL   (d) RCSPP+CLSUnSAL
    (4.7917)   (3.9035)   (7.1098)   (5.4570)

Fig. 8. Error maps obtained after reconstructing the urban HYDICE scene using the AEB in Somers et al. (2012) and the proposed RCSPP. For the AEB method, 20 sub-blocks with a size that is 10% of the size of the original image are randomly chosen, ending up with 180 candidate endmembers (each sub-block consists of 9 endmembers). For the RCSPP, 77 partitions are derived and the percentage of candidates selected in each partition is 0.5%. Thus, a total of 513 candidate endmembers is used in experiments.

scene and the urban HYDICE scene with different numbers of partitions, clustering criteria, and search scope. When one parameter is tested, the other parameters are fixed and set to the same values reported in Section 4.3.1 and Section 4.3.2. The adopted algorithm is the RCSPP + OSP combination, mainly because the OSP algorithm can produce more stable endmember results after several executions than other tested methods. Table 5 shows the obtained SAD and RMSE values for different parameter settings using the Fractal 1 synthetic scene. First of all, we can observe that the SAD and RMSE results are very similar when the number of partitions is greater than 25. These results suggest that our method is not sensitive to the number of partitions. Considering that the larger the number of partitions, the heavier the computational burden, we suggest to choose a relatively low number of partitions in practice. Regarding the impact of different clustering criteria, including the ED, SAD, SID, and SID-SAM (Du et al., 2004), we have empirically found that the proposed method is not very sensitive to these metrics. As a result, we suggest to choose one of the aforementioned distance metrics in practice. Finally, We also analyze the impact of the neighborhood size in the search strategy implemented by the considered RCSPP, which is set to two, three, and four times the size of the initial image partition [see Table 5]. As we can observe, for the synthetic Fractal 1 scene, the SAD and RMSE results are very similar in different scenarios. However, we emphasize that the computational burden becomes heavier when the size of the spatial neighborhood

Table 4
Average(± standard deviation), minimum and maximum SAD (in radians) of the bundles obtained by the AEB in Somers et al. (2012) and the proposed RCSPP.

		AEB	RCSPP
Asphalt	Average(± stdev)	0.2202 ± 0.0951	0.162 ± 0.082
	Min	0	0.0204
	Max	0.4288	0.4148
Grass	Average(± stdev)	0.218 ± 0.1264	0.103 ± 0.0464
	Min	0	0.0179
	Max	0.5085	0.2236
Tree	Average(± stdev)	0.1584 ± 0.0991	0.1006 ± 0.053
	Min	0	0.0108
	Max	0.4287	0.3092
Roof	Average(± stdev)	0.1764 ± 0.071	0.1982 ± 0.1773
	Min	0	0.0289
	Max	0.0161	0.6760

becomes larger. As a result, we recommend a neighborhood size that is two times of the initial partition size in practice.

Table 6 shows the EMD_{tot} between the obtained unmixing results and the reference data with different parameter settings on the urban HYDICE scene. We can see from Table 6 that, for numbers of partitions and neighborhood sizes that are similar to those in the pervious experiment, the proposed RCSPP is not sensitive to the parameter settings.
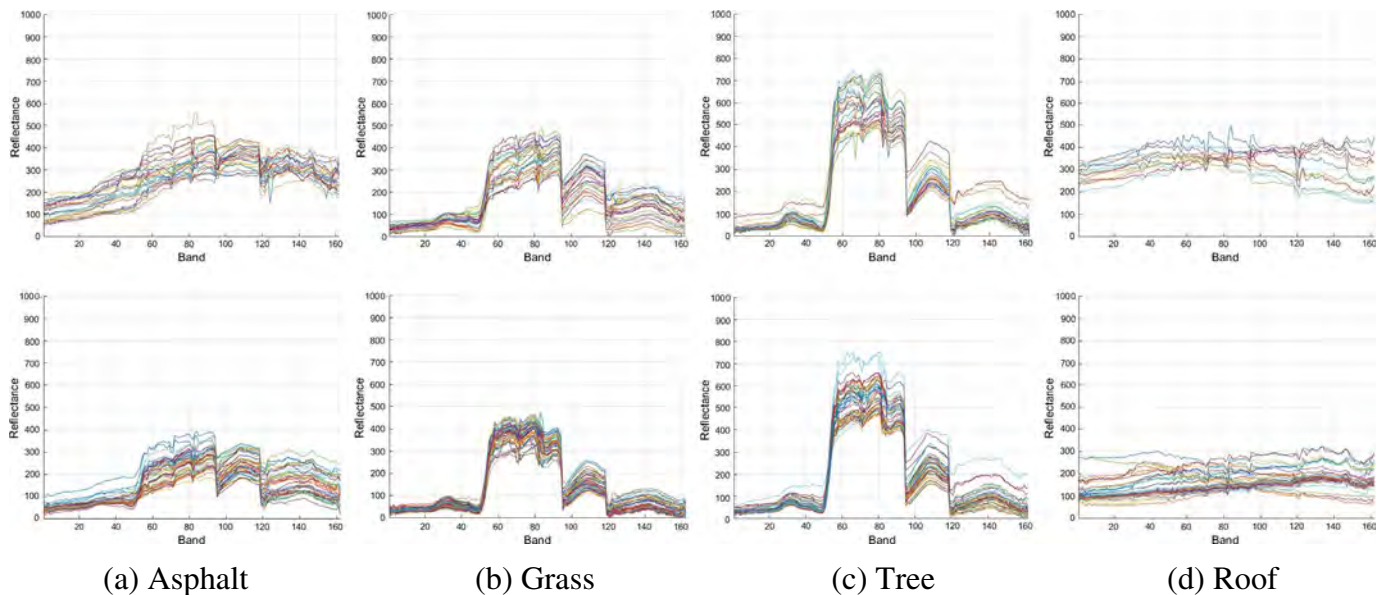

(a) Asphalt   (b) Grass   (c) Tree   (d) Roof

Fig. 9. Endmember bundles constructed for the urban HYDICE scene using the AEB in Somers et al. (2012) (first row) and the proposed RCSPP (second row).

Table 5
Experimental results obtained by the proposed method using different parameter settings with the Fractal 1 scene, including the number of partitions, the clustering criteria and the size of the spatial neighborhood used to implement the search strategy. Here, the SAD (in radians) is measured between the USGS library signatures and the extracted endmember spectra by the RCSPP + OSP algorithm. RMSE is measured between the original and the reconstructed version of the Fractal 1 image using the RCSPP + OSP algorithm and their corresponding fractional abundances, estimated by the FCLS method. When one parameter is tested, the other parameters are fixed and set to the same values reported in Section 4.3.1.

Parameter type	Parameter setting	SNR = 30:1		SNR = 50:1		SNR = 70:1		SNR = 90:1	
		SAD	RMSE	SAD	RMSE	SAD	RMSE	SAD	RMSE
Number of partitions	16	0.044	0.247	0.0202	0.0444	0.0206	0.0031	0.0206	0.0031
	25	0.0367	0.231	0.0142	0.0262	0.0133	0.0065	0.0132	0.0052
	50	0.04	0.237	0.0169	0.0270	0.0157	0.0090	0.0159	0.0077
	100	0.0407	0.238	0.0136	0.0258	0.0131	0.0074	0.0131	0.0064
	400	0.037	0.233	0.0143	0.0250	0.0129	0.0050	0.0131	0.0036
Cluster criteria	ED	0.04	0.235	0.0155	0.0257	0.0147	0.0063	0.0147	0.0051
	SAD	0.0391	0.239	0.0157	0.0284	0.0143	0.01	0.0145	0.0089
	SID	0.0375	0.238	0.015	0.0268	0.0136	0.0079	0.0136	0.0069
	SID-SAM	0.0367	0.231	0.0142	0.0262	0.0133	0.0065	0.0132	0.0052
Times of initial partition	2	0.0367	0.231	0.0142	0.0262	0.0133	0.0065	0.0132	0.0052
	3	0.0367	0.231	0.0142	0.0262	0.0133	0.0065	0.0132	0.0052
	4	0.0367	0.231	0.0142	0.0262	0.0133	0.0065	0.0132	0.0052

Table 6
EMD_{tot} values ($\times 10^4$) between the obtained unmixing results and the reference data, obtained by considering different parameter settings with the urban HYDICE scene. We consider different number of partitions, clustering criteria and sizes of the spatial neighborhood used to implement the search strategy. The obtained values are calculated by using Eq. (9). Here, the four reference spectral signatures are acquired using the procedure described in Section 4.1.2. When one parameter is tested, the other parameters are fixed and set to the same values reported in Section 4.3.2.

Para. type	Para. setting	OSP	N-FINDR	RCSPP + OSP	RCSPP + N-FINDR
Number of partitions	100	4.391	4.251	3.1571	2.8782
	200			4.2079	2.9693
	300			4.8628	2.8377
	400			4.1731	4.5604
Cluster criteria	ED	4.391	4.251	4.9004	2.4896
	SAD			3.6485	2.5699
	SID			3.7310	3.3660
	SID-SAM			3.1571	2.8782
Times of initial partition	2	4.391	4.251	3.1571	2.8782
	3			4.3534	2.2611
	4			2.8160	3.2273

Therefore, in order to reduce computational complexity, we suggest to use a moderate number of partitions and neighborhood size. With respect to the clustering criteria, it can be seen that the use of SID-SAM obtained better results. This is expected, as it simultaneously considers the variability and similarity of endmembers. Therefore, we suggest to choose SID-SAM for the implementation of the adopted clustering criteria.

4.6. Analysis on the computational complexity

In this section, we performed an analysis on the computational complexity of our proposed RCSPP method. Our method includes two main steps: 1) regional clustering, and 2) extraction of candidate endmembers. After obtaining the candidate endmembers, we can use them as an image endmember library directly, or feed them to a spectral-based endmember extraction method (such as OSP and N-FINDR.) for further refinement. For the regional clustering step, the search scope is localized in a neighborhood, so the complexity is $O(N)$ with N being the number of pixels in the image. This means the complexity of regional clustering step is linear with the number of pixels, irrespective of the number of partitions. For the extraction of candidate endmembers, the algorithm performs a loop on all partitions. In each partition, the main computational complexity is the PCT, which has an upper bound of $O(n_i^3)$ complexity with n_i being the number of pixels in the ith partition. If we assume that the number of partitions is c, and $n_i \approx N/c$, the computational complexity of the endmember candidates extraction step

Table 7
Processing times (in seconds) measured for different methods when processing the urban HYDICE data. Here, the Preprocessing time includes the regional clustering step and the candidate endmembers extraction step. The number of partitions is 77, and the number of candidate endmembers is 513.

Algorithm	Preprocessing	Endmember extraction	Total
OSP	–	105.391	105.391
N-FINDR	–	86.172	86.172
RCSPP + OSP	9.250	8.594	17.844
RCSPP + N-FINDR	9.250	8.312	17.562

is $O(c \times (n_i^3)) \approx O(N^3/c^2)$. Here, we also present the computational times (in the considered computing environment) measured for different methods when processing the urban HYDICE scene. The obtained results are reported on Table 7. As shown by this table, the RCSPP (in combination with endmember extraction methods) significantly reduces the processing time as compared with the endmember extraction methods alone. This is because the number of candidate pixels to be considered by the spectral-based endmember extraction methods is significantly reduced (in our experiments, only 0.5% of the original image pixels are selected). This indicates that the newly proposed RCSPP offers a relevant strategy to improve the endmember extraction process from the viewpoint of both the quality of the extracted endmembers (achieved by the integration of spatial information into the process) and the computational performance (achieved by intelligently selecting suitable candidate pixels).

5. Conclusions and future research

In this paper, we have developed a new spatial preprocessing method for remotely sensed hyperspectral unmixing. The proposed technique is derived from the SLIC method and naturally integrates the spatial and the spectral information. It comprises two main steps. First, a regional clustering process is applied to obtain partitions with high spectral similarity and spatial correlation. Second, candidate pixels are intelligently extracted from each partition. The proposed method can be run as an independent module, and the output candidate pixels are fed to an existing spectral-based endmember extraction method. Our experimental results, conducted using both synthetic and real hyperspectral images, indicate that the proposed method exhibits the potential to improve the accuracy of spectral-based endmember identification methods, regardless of the presence of pure signatures in the original image data. Moreover, since the set of candidate endmembers is much smaller than the set of pixels in the original image, the proposed method exhibits stable and fast results as compared with other traditional spectral-based endmember extraction methods.

As with any new approach, there are some unresolved issues that may present challenges over time. In this sense, future research includes a more detailed investigation of the fact that combined spectral-based endmember extraction techniques generally result in fixed endmember spectra, and it is difficult to represent endmember variability in real hyperspectral scenes with these techniques. How to combine these methods with multiple endmember-based methods such as MESMA (Roberts et al., 1998) or other endmember bundles methods (Bateson et al., 2000; Xu et al., 2015) (and the effect of this kind of combination) also deserves further research. Another issue is related to the application of RCSPP on urban scenes, in which the determination of endmember classes is very subjective and perhaps not accurate in some cases. For instance, different buildings' roofs are made up of different materials, which may result in different endmember features. Different varieties of trees may also result in different endmember features. How to represent such classes more accurately requires further investigation. Moreover, in our RCSPP method, the number of partitions is determined by a rough estimation that depends on the spatial resolution of the hyperspectral image. In turn, the investigation of whether there is an optimal partition size and how to obtain this value need further research. In addition, our regional clustering strategy may produce some unsuccessfully clustered pixels, and these pixels represent outliers or heterogeneous pixels. Whether these pixels are suitable to be used as candidate endmembers or not also need further analysis. A more detailed investigation on the impact of the segmentation results (from coarse to fine), and the weighting value of λ in Eq. (1) is also a topic deserving attention in future research. Finally, in future developments we will also provide additional experiments with other real hyperspectral data sets, and attempt to apply our method for spatial preprocessing in the context of hyperspectral image classification problems.

Acknowledgments

References

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Susstrunk, S., 2012. SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Trans. Pattern Anal. Mach. Intell. 34 (11), 2274–2282.

Ball, G.H., Hall, D.J., 1965. ISODATA, a novel method of data analysis and pattern classification. In: Tech. Rep. DTIC Document.

Bateson, C.A., Asner, G.P., Wessman, C.A., 2000. Endmember bundles: a new approach to incorporating endmember variability into spectral mixture analysis. IEEE Trans. Geosci. Remote Sens. 38 (2), 1083–1094.

Bioucas-Dias, J.M., Nascimento, J.M., 2008. Hyperspectral subspace identification. IEEE Trans. Geosci. Remote Sens. 46 (8), 2435–2445.

Bioucas-Dias, J.M., Plaza, A., Dobigeon, N., Parente, M., Du, Q., Gader, P., Chanussot, J., 2012. Hyperspectral unmixing overview: geometrical, statistical and sparse regression-based approaches. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 5 (2), 354–379.

Boardman, J.W., Kruse, F.A., Green, R.O., 1995. Mapping target signatures via partial unmixing of AVIRIS data. In: Summaries, Fifth JPL Airborne Earth Science Workshop. Jet Propulsion Laboratory, Pasadena, CA, pp. 23–26.

Canham, K., Schlamm, A., Ziemann, A., Basener, B., Messinger, D., 2011. Spatially adaptive hyperspectral unmixing. IEEE Trans. Geosci. Remote Sens. 49 (11), 4248–4262.

Chang, C.-I., Du, Q., 2004. Estimation of number of spectrally distinct signal sources in hyperspectral imagery. IEEE Trans. Geosci. Remote Sens. 42 (3), 608–619.

Chen, Scott Shaobing, Donoho, David L., Saunders, Michael A., 2001. Atomic decomposition by basis pursuit. SIAM Rev. 43 (1), 33–61.

Delalieux, S., Somers, B., Haest, B., Spanhove, T., Borre, J.V., Mücher, C., 2012. Heathland conservation status mapping through integration of hyperspectral mixture analysis and decision tree classifiers. Remote Sens. Environ. 126, 222–231.

Dennison, P.E., Roberts, D.A., 2003. Endmember selection for multiple endmember spectral mixture analysis using endmember average RMSE. Remote Sens. Environ. 87 (2–3), 123–135.

Du, Y., Chang, C.-I., Ren, H., Chang, C.-C., Jensen, J.O., et al., 2004. New hyperspectral discrimination measure for spectral characterization. Opt. Eng. 43 (8), 1777–1786.

Felzenszwalb, P.F., Huttenlocher, D.P., 2004. Efficient graph-based image segmentation. Int. J. Comput. Vis. 59 (59), 167–181.

Fernández-Manso, A., Quintano, C., Roberts, D., 2012. Evaluation of potential of multiple endmember spectral mixture analysis (MESMA) for surface coal mining affected area mapping in different world forest ecosystems. Remote Sens. Environ. 127, 181–193.

Franke, J., Roberts, D.A., Halligan, K., Menz, G., 2009. Hierarchical multiple endmember spectral mixture analysis (MESMA) of hyperspectral imagery for urban environments. Remote Sens. Environ. 113 (8), 1712–1723.

Gillis, D.B., Bowles, J.H., 2012. Hyperspectral image segmentation using spatial-spectral graphs. In: SPIE Defense, Security, and Sensing. vol. 8390 pp. 83901Q-1–83901Q-11.

Goenaga, M.A., Torres-Madronero, M.C., Velez-Reyes, M., Van Bloem, S.J., Chinea, J.D., 2013. Unmixing analysis of a time series of hyperion images over the Guánica dry forest in Puerto Rico. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 6 (2), 329–338.

Green, A.A., Berman, M., Switzer, P., Craig, M.D., 1988. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. IEEE Trans. Geosci. Remote Sens. 26 (1), 65–74.

Halkidi, M., Vazirgiannis, M., Batistakis, Y., 2000. Quality Scheme Assessment in the Clustering Process. Springer Berlin Heidelberg.

Harsanyi, J.C., Chang, C.-I., 1994. Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach. IEEE Trans. Geosci. Remote Sens. 32 (4), 779–785.

Herold, M., Roberts, D.A., Gardner, M.E., Dennison, P.E., 2004. Spectrometry for urban area remote sensing-development and analysis of a spectral library from 350 to 2400 nm. Remote Sens. Environ. 91 (3–4), 304–319.

Iordache, M.-D., Bioucas-Dias, J.M., Plaza, A., 2011. Sparse unmixing of hyperspectral data. IEEE Trans. Geosci. Remote Sens. 49 (6), 2014–2039.

Iordache, M.D., Bioucas-Dias, J.M., Plaza, A., 2014. Collaborative sparse regression for hyperspectral unmixing. IEEE Trans. Geosci. Remote Sens. 1 (1), 341–354.

Jia, S., Qian, Y., 2009. Constrained nonnegative matrix factorization for hyperspectral unmixing. IEEE Trans. Geosci. Remote Sens. 47 (1), 161–173.

Keshava, N., Mustard, J.F., 2002. Spectral unmixing. IEEE Signal Process. Mag. 19 (1), 44–57.

Kowkabi, F., Ghassemian, H., Keshavarz, A., 2016a. Enhancing hyperspectral endmember extraction using clustering and oversegmentation-based preprocessing. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9 (6), 2400–2413.

Kowkabi, F., Ghassemian, H., Keshavarz, A., 2016b. A fast spatial-spectral preprocessing module for hyperspectral endmember extraction. IEEE Geosci. Remote Sens. Lett. 13 (6), 782–786.

Legány, C., Juhász, S., Babos, A., 2006. Cluster validity measurement techniques. In: Proceedings of the 5th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases, pp. 388–393.

Levinshtein, A., Stere, A., Kutulakos, K.N., Fleet, D.J., Dickinson, S.J., Siddiqi, K., 2009. TurboPixels: fast superpixels using geometric flows. IEEE Trans. Pattern Anal. Mach. Intell. 31 (12), 2290–2297.

Li, H., Zhang, L., 2011. A hybrid automatic endmember extraction algorithm based on a local window. IEEE Trans. Geosci. Remote Sens. 49 (11), 4223–4238.

Li, H., Zhang, S., Ding, X., Zhang, C., Dale, P., 2016. Performance evaluation of cluster validity indices (CVIs) on multi/hyperspectral remote sensing datasets. Remote Sens. 8 (4), 295.

Li, J., Agathos, A., Zaharie, D., Bioucas-Dias, J.M., Plaza, A., Li, X., 2015. Minimum volume simplex analysis: a fast algorithm for linear hyperspectral unmixing. IEEE Trans. Geosci. Remote Sens. 53 (9), 5067–5082.

Liu, T., Yang, X., 2013. Mapping vegetation in an urban area with stratified classification and multiple endmember spectral mixture analysis. Remote Sens. Environ. 133, 251–264.

Lopez, S., Moure, J.F., Plaza, A., Callico, G.M., Lopez, J.F., Sarmiento, R., 2013. A new preprocessing technique for fast hyperspectral endmember extraction. IEEE Geosci. Remote Sens. Lett. 10 (5), 1070–1074.

Martín, G., Plaza, A., 2011. Region-based spatial preprocessing for endmember extraction and spectral unmixing. IEEE Geosci. Remote Sens. Lett. 8 (4), 745–749.

Martín, G., Plaza, A., 2012. Spatial-spectral preprocessing prior to endmember identification and unmixing of remotely sensed hyperspectral data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 5 (2), 380–395.

Massoudifar, P., Rangarajan, A., Gader, P., 2014. Superpixel estimation for hyperspectral imagery. In: IEEE Comput. Vis. Pattern Recognit. Workshops, pp. 287–292.

Mei, S., He, M., Wang, Z., Feng, D., 2010. Spatial purity based endmember extraction for spectral mixture analysis. IEEE Trans. Geosci. Remote Sens. 48 (9), 3434–3445.

Nascimento, J.M., Bioucas-Dias, J.M., 2005. Vertex component analysis: a fast algorithm to unmix hyperspectral data. IEEE Trans. Geosci. Remote Sens. 43 (4), 898–910.

Neville, R., Staenz, K., Szeredi, T., Lefebvre, J., Hauff, P., 1999. Automatic endmember extraction from hyperspectral data for mineral exploration. In: Proc. 21st Can. Symp. Remote Sens, pp. 21–24.

Okin, G.S., Clarke, K.D., Lewis, M.M., 2013. Comparison of methods for estimation of absolute vegetation and soil fractional cover using MODIS normalized BRDF-adjusted reflectance data. Remote Sens. Environ. 130 (130), 266–279.

Pati, Y.C., Rezaiifar, R., Krishnaprasad, P.S., 1995. Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. In: Conference on Signals, pp. 1–3.

Plaza, A., Benediktsson, J.A., Boardman, J.W., Brazile, J., Bruzzone, L., Camps-Valls, G., Chanussot, J., Fauvel, M., Gamba, P., Gualtieri, A., et al., 2009. Recent advances in techniques for hyperspectral image processing. Remote Sens. Environ. 113, S110–S122.

Plaza, A., Martinez, P., Pérez, R., Plaza, J., 2002. Spatial/spectral endmember extraction by multidimensional morphological operations. IEEE Trans. Geosci. Remote Sens. 40 (9), 2025–2041.

Powell, R.L., Roberts, D.A., Dennison, P.E., Hess, L.L., 2007. Sub-pixel mapping of urban land cover using multiple endmember spectral mixture analysis: Manaus, Brazil. Remote Sens. Environ. 106 (2), 253–267.

Quintano, C., Fernández-Manso, A., Roberts, D.A., 2013. Multiple endmember spectral mixture analysis (MESMA) to map burn severity levels from Landsat images in Mediterranean countries. Remote Sens. Environ. 136, 76–88.

Roberts, D.A., Gardner, M., Church, R., Ustin, S., Scheer, G., Green, R.O., 1998. Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models. Remote Sens. Environ. 65 (3), 267–279.

Roberts, D.A., Smith, M.O., Adams, J.B., 1993. Green vegetation, nonphotosynthetic vegetation, and soils in AVIRIS data. Remote Sens. Environ. 44, 255–269.

Roberts, D.A., Ustin, S.L., Ogunjemiyo, S., Greenberg, J., Dobrowski, S.Z., Chen, J., Hinckley, T.M., 2004. Spectral and structural measures of northwest forest vegetation at leaf to landscape scales. Ecosystems 7 (5), 545–562.

Rogge, D., Rivard, B., Zhang, J., Sanchez, A., Harris, J., Feng, J., 2007a. Integration of spatial-spectral information for the improved extraction of endmembers. Remote Sens. Environ. 110 (3), 287–303.

Rogge, D.M., Rivard, B., Zhang, J., Feng, J., 2007b. Iterative spectral unmixing for optimizing per-pixel endmember sets. IEEE Trans. Geosci. Remote Sens. 44 (12), 3725–3736.

Roth, K.L., Dennison, P.E., Roberts, D.A., 2012. Comparing endmember selection techniques for accurate mapping of plant species and land cover using imaging spectrometer data. Remote Sens. Environ. 127 (4), 139–152.

Rubner, Y., Tomasi, C., Guibas, L.J., 2000. The earth mover's distance as a metric for image retrieval. Int. J. Comput. Vis. 40 (2), 99–121.

Saranathan, A.M., Parente, M., 2016. Uniformity-based superpixel segmentation of hyperspectral images. IEEE Trans. Geosci. Remote Sens. 54 (3), 1419–1430.

Schaaf, A.N., Dennison, P.E., Fryer, G.K., Roth, K.L., Roberts, D.A., 2011. Mapping plant functional types at multiple spatial resolutions using imaging spectrometer data. GIsci. Remote Sens. 48 (3), 324–344.

Schaepman, M.E., Ustin, S.L., Plaza, A.J., Painter, T.H., Verrelst, J., Liang, S., 2009. Earth system science related imaging spectroscopy-an assessment. Remote Sens. Environ. 113, S123–S137.

Shi, C., Wang, L., 2015. Incorporating spatial information in spectral unmixing: a review. Remote Sens. Environ. 149, 70–87.

Shi, J., Malik, J., 2000. Normalized cuts and image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 22 (8), 888–905.

Somers, B., Asner, G.P., 2013. Multi-temporal hyperspectral mixture analysis and feature selection for invasive species mapping in rainforests. Remote Sens. Environ. 136 (5), 14–27.

Somers, B., Asner, G.P., Tits, L., Coppin, P., 2011. Endmember variability in spectral mixture analysis: a review. Remote Sens. Environ. 115 (7), 1603–1616.

Somers, B., Zortea, M., Plaza, A., Asner, G.P., 2012. Automated extraction of image-based endmember bundles for improved spectral unmixing. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 5 (2), 396–408.

Thompson, D.R., Mandrake, L., Gilmore, M.S., Castano, R., 2010. Superpixel endmember detection. IEEE Trans. Geosci. Remote Sens. 48 (11), 4023–4033.

Thorp, K., French, A., Rango, A., 2013. Effect of image spatial and spectral characteristics on mapping semi-arid rangeland vegetation using multiple endmember spectral mixture analysis (MESMA). Remote Sens. Environ. 132, 120–130.

Tilton, J.C., 2003. Analysis of hierarchically related image segmentations. In: Advances in Techniques for Analysis of Remotely Sensed Data, 2003 IEEE Workshop on. IEEE, pp. 60–69.

Torres-Madronero, M.C., Velez-Reyes, M., 2014. Integrating spatial information in unsupervised unmixing of hyperspectral imagery using multiscale representation. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7 (6), 1985–1993.

Vedaldi, A., Soatto, S., 2008. Quick Shift and Kernel Methods for Mode Seeking. Springer Berlin Heidelberg, pp. 705–718.

Vincent, L., Soille, P., 1991. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. IEEE Trans. Pattern Anal. Mach. Intell. 13 (6), 583–598.

Winter, M.E., 1999. N-FINDR: an algorithm for fast autonomous spectral endmember determination in hyperspectral data. In: SPIE's International Symposium on Optical Science, Engineering, and Instrumentation. International Society for Optics and Photonics, pp. 266–275.

Xu, M., Zhang, L., Du, B., 2015. An image-based endmember bundle extraction algorithm using both spatial and spectral information. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 8 (6), 2607–2617.

Young, I.T., Van Vliet, L.J., 1995. Recursive implementation of the Gaussian filter. Signal Process. 44 (2), 139–151.

Zare, A., Anderson, D.T., 2013. Earth movers distance-based simultaneous comparison of hyperspectral endmembers and proportions. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7 (6), 1910–1921.

Zare, A., Ho, K.C., 2014. Endmember variability in hyperspectral analysis: addressing spectral variability during spectral unmixing. IEEE Signal Process. Mag. 31 (1), 95–104.

Zhang, J., Rivard, B., Rogge, D., 2008. The successive projection algorithm (SPA), an algorithm with a spatial constraint for the automatic search of endmembers in hyperspectral data. Sensors 8 (2), 1321–1342.

Zortea, M., Plaza, A., 2009. Spatial preprocessing for endmember extraction. IEEE Trans. Geosci. Remote Sens. 47 (8), 2679–2693.