Fusion of Spectral and Spatial Information for Classification of Hyperspectral Remote-Sensed Imagery by Local Graph

Wenzhi Liao, Member, IEEE, Mauro Dalla Mura, Member, IEEE, Jocelyn Chanussot, Fellow, IEEE, and Aleksandra Pižurica, Member, IEEE

Abstract—Hyperspectral imagery contains a wealth of spectral and spatial information that can improve target detection and recognition performance. Conventional feature extraction methods cannot fully exploit both spectral and spatial information. Data fusion by simply stacking different feature sources together does not take into account the differences between feature sources. In this paper, a local graph-based fusion (LGF) method is proposed to couple dimension reduction and feature fusion of the spectral information (i.e., the spectra in the HS image) and the spatial information [extracted by morphological profiles (MPs)]. In the proposed method, the fusion graph is built on the full data by moving a sliding window from the first pixel to the last one. This yields a clear improvement over a previous approach with fusion graph built on randomly selected samples. Experimental results on real hyperspectral images are very encouraging. Compared to the methods using only single feature and stacking all the features together, the proposed LGF method improves the overall classification accuracy on one of the data sets for more than 20% and 5%, respectively.

Index Terms—Classification, data fusion, graph-based, hyper-spectral image, remote sensing.

I. INTRODUCTION

R ECENT advances in sensors technology have led to an increased availability of hyperspectral data from urban areas at very high spatial and spectral resolutions. Automated image analysis techniques for the high-resolution remote sensing data often make use of mathematical morphology [1],

Manuscript received May 11, 2015; revised September 08, 2015; accepted November 04, 2015. Date of publication November 30, 2015; date of current version February 09, 2016. This work was supported in part by the Fund for Strategic Basic Research supported by the Agency for Innovation by Science and Technology in Flanders (SBO-IWT) project Chameleon: Domain-specific Hyperspectral Imaging Systems for Relevant Industrial Applications and in part by the Fund for Scientific Research in Flanders (FWO) project G037115N "Data fusion for image analysis in remote sensing."

W. Liao and A. Pižurica are with the Department of Telecommunications and Information Processing, Ghent University—iMinds—Image Processing and Interpretation, 9000 Ghent, Belgium (e-mail: Wenzhi.Liao@telin.ugent.be; Aleksandra.Pizurica@telin.ugent.be).

M. Dalla Mura is with the Grenoble Images Parole Signals Automatics Laboratory (GIPSA-Lab), Grenoble Institute of Technology, 38402 Saint Martin d'Hères, France (e-mail: mauro.dalla-mura@gipsa-lab.grenoble-inp.fr).

J. Chanussot is with the Grenoble Images Parole Signals Automatics Laboratory (GIPSA-Lab), Grenoble Institute of Technology, 38402 Saint Martin d'Hères, France, and also with the Faculty of Electrical and Computer Engineering, University of Iceland, Reykjavik 101, Iceland (e-mail: jocelyn.chanussot@gipsa-lab.grenoble-inp.fr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2015.2498664

[2]. Pesaresi and Benediktsson [3] proposed the use of morphological transformations to build an MP. Bellens et al. [4] further explored this approach by using both disk-shaped and linear structuring elements to improve the classification of veryhigh-resolution panchromatic urban imagery. An extension of MPs that can handle efficiently hyperspectral images with high spatial resolution was proposed by Benediktsson et al. [5], where MPs are built on the first few principal components (PCs) extracted from the hyperspectral image. The resulting approach was named extended MP (EMP) and has inspired a number of further developments in the literature. In [6] and [7], the MP was built with partial reconstruction [4], showing an improvement with respect to the EMP in classification of hyperspectral data. In [8], kernel PCs were used to construct the EMP, with a significant improvement in terms of classification accuracy when compared with the conventional EMP built on PCs. In [9], the attribute profiles (APs) [10] were applied to the first PCs extracted from a hyperspectral image, generating an extended AP (EAP). The approach of [11] improved the classification results by constructing the EAP with the independent component analysis.

A limitation of the above approaches is that they rely mainly on geometrical features and do not fully utilize the spectral information in the HS data. In fact, MPs (or their variants) are built by only considering few components extracted from the original HS cube, hence not fully exploiting the spectral information. The information contained in the measured reflectance spectra allows discrimination between different objects based on their material composition. Thus, combining spatial and spectral information can contribute to a more comprehensive interpretation of objects on the ground. For example, spectral signatures cannot differentiate between objects made of the same material (e.g., roofs and roads made with the same asphalt), while they can often be easily distinguished by their geometry. On the other hand, spatial features alone may fail to discriminate between objects that are quite different in nature (e.g., grass field, parking, or a swimming pool), if their shape and size are similar. Many approaches have been developed to fuse the spectral and spatial information for the classification of remote sensing data [13]–[19]. Some of these approaches employ the so-called composite kernel methods [13], [14] or their generalization [16]. Others define spatial information through MPs, and concatenate spectral and spatial features in a stacked architecture for classification [17]-[19]. Recently, Huang et al. [20] proposed a multifeature model to combine

1939-1404 © 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information. multiple spectral and spatial features at both pixel and object levels by a support vector machine (SVM) ensemble, showing an improvement over vector stacking, feature selection, and composite kernels.

While such methods (that simply concatenate several kinds of features together) are appealing due to their simplicity, they may not perform better (or may even perform worse) than using a single feature. Dalla Mura et al. [11] showed examples where the classification accuracies after stacking different morphological attributes even dropped compared to the case of considering a single one. This is because the information contained by different features is not equally represented or measured. The value of different components in the feature vector can be significantly unbalanced. Furthermore, stacking several kinds of feature sources may yield redundant information, making it difficult to select an optimal combination of spectral and spatial features as it was shown by Fauvel et al. in [19]. In addition, the increase in the dimensionality of the stacked features, as well as the limited number of labeled samples may in practice pose the problem of the "curse of dimensionality," consequently increasing the risk to overfit the training data.

To overcome these problems, Debes et al. proposed a graphbased data fusion (GDF) method [21]¹ to couple data fusion and dimension reduction for classification of multisensor imagery. The GDF [21] combined multiple feature sources through a fused graph. The built graph explains the relations between data points in the two data sources and can be seen as a way to model the embedding in the manifold in which the data lie [21]. This approach proved to overcome the conventional approach of stacking different feature sources together in terms of classification accuracy. However, the GDF can cause some problems on storage resources and computational load especially when using large training data sets. This is because finding k nearest neighbors to build a graph is very intensive in both computation and memory consumption. Random sampling was used to speed up the GDF in [21]. However, random sampling can lead to poor representation of the whole area if large areas are not sampled, which will lead to unstable performances. This is even worse if the study area is very large and the number of samples fixed. Moreover, image degradation cannot be avoided during the hyperspectral data acquisition, which will lead to poor performances on finding k nearest neighbors for building a graph on the whole original data or randomly selected samples.

In this paper, we propose a new local graph fusion (LGF) method to overcome the above-mentioned limitations. Specifically, the proposed LGF is used to couple dimensionality reduction and the fusion of spectral (i.e., the original HS image) and spatial features (i.e., the morphological features computed from the HS image). The main contributions of this paper can be summarized as follows.

1) First and foremost, the proposed LGF builds the local fusion graph on the whole data by employing a sliding window. This way we introduce a different approach with regard to GDF [21], where the fusion graph was

¹The GDF method won the "Best Paper Challenge" award of 2013 IEEE Data Fusion Contest.

built globally on randomly selected samples. The local spatial neighborhood information is very important for remote sensing, especially for high-resolution remote sensing imagery. Specifically, many methods [23]-[25], [28] demonstrated notable improvements on the performances of dimensionality reduction, classification, and segmentation, by exploiting the local spatial neighborhood information. In typical remote sensing scenes (especially for high-resolution remote sensing images), pixels in a small spatial neighborhood usually share similar properties (e.g., very similar spectral characteristics). If we build a fusion graph globally, pixels from different objects may become the nearest neighbors of each other, if they share similar spectral characteristics. For example, pixels belonging to a roof of a building may get connected in the graph to pixels of parking lots, because they have very similar spectral characteristics even though they might be not spatially adjacent. Within a small spatial window, the proposed LGF better employs the local spatial neighborhood information to represent objects in the feature space. This way, our approach enables better performances on classification, and better constraints in terms of local connectivity reduce a risk of erroneously selected nearest neighbors even when the spectral characteristics are affected by noise.

- 2) In addition, with the proposed local fusion graph, our approach reduces computational complexity from $O(N^2)$, which holds for the global fusion graph on the whole data to only $O(NS^2)$, where N denotes the total number of spatial pixels and $S \ll N$ is the size of the sliding window.
- 3) Last but not least, the proposed approach admits a fast implementation by just spatially downsampling the original data, while keeping the performances stable. As shown in the experiments, for the high-resolution remote sensing images, spatial downsampling will not affect much the main spatial structure of the objects (i.e., leading to similar classification performances obtained without subsampling), but can efficiently reduce the computational complexity by a factor equal to the square of the spatial downsampling ratio.

This paper is organized as follows. Section II provides a brief review of morphological features. In Section III, we present the proposed local graph-based fusion (LGF) method. The experimental results on real hyperspectral images are presented and discussed in Section IV. Finally, the conclusion is drawn in Section V.

II. MORPHOLOGICAL FEATURES

Typical morphological features used for characterizing the spatial information of very-high-resolution remote sensing images are generated by applying morphological openings and closings by reconstruction [5] on a grayscale image, using a structural element (SE) of predefined size and shape. An opening acts on objects that are bright compared with their surrounding, while closings act on dark objects. For example, an opening deletes (i.e., the pixels in the object take on the value



Fig. 1. MP built on the first PC of University area. The scale of circular SE varies from 2 to 6, with step size increment of 2.

of their surrounding) bright objects that are smaller than the SE. By increasing the size of the SE and repeating the previous operation, a complete MP is built [5], carrying information about the size and the shape of objects in the image due to its multiscale nature.

An MP is composed of an opening profile (OP) and a closing profile (CP). The OP with M scales at pixel x forms an *M*-dimensional image, and so does the CP. By stacking the OP, the CP, and original image, the MP of pixel x is obtained, leading to a (2M + 1)-dimensional vector. When applying MP to hyperspectral data, PC analysis (PCA) is widely used as a preprocessing step to reduce the dimensionality of the highdimensional original data, as well as to reduce redundancy among the bands [5]. Then, one constructs an MP on each PC independently. An EMP is formed by stacking all the computed MPs. Suppose p PCs are extracted from the original hyperspectral data, then the EMP of a pixel x is a p(2M + 1)-dimensional vector. Fig. 1 shows an MP built on the first PC. The usefulness of using morphological features for classification of remote sensing data on urban areas has been discussed in numerous studies [4]–[11], [15]–[19].

III. PROPOSED METHOD

In this section, we propose a local graph-based fusion (LGF) method.² Suppose $\mathbf{X}^{Spec} = {\{\mathbf{x}_i^{Spec}\}}_{i=1}^N$ and $\mathbf{X}^{Spat} = {\{\mathbf{x}_i^{Spat}\}}_{i=1}^N$ denote the spectral and spatial features after normalization of their values to the same interval (e.g., [0,1]), where $\mathbf{x}_i^{Spec} \in \mathbb{R}^B$, with *B* the number of bands and $\mathbf{x}_i^{Spat} \in \mathbb{R}^D$ (with D = p(2M + 1) being generated by an EMP built on *p* PCs and with *M* filters), and *N* is the total number of spatial pixels in a HS image. Further on, we denote the stacked spectral and spatial features by $\mathbf{X}^{Sta} = {\{\mathbf{x}_i^{Sta}\}}_{i=1}^N = [\mathbf{X}^{Spec}; \mathbf{X}^{Spat}]$, where $\mathbf{x}_i^{Sta} = [\mathbf{x}_i^{Spec}; \mathbf{x}_i^{Spat}] \in \mathbb{R}^{B+D}$.

The goal of this paper is to find a transformation matrix $\mathbf{W} \in \mathbb{R}^{(B+D) \times d}$, which can couple dimensionality reduction (to *d*-dimensions) and feature fusion in a way of

$$\mathbf{z}_i = \mathbf{W}^T \mathbf{x}_i \tag{1}$$

where \mathbf{x}_i is a multivariate variable which can be set to \mathbf{x}_i^{Sta} , and $\{\mathbf{z}_i\}_{i=1}^N$ the fusion features in a lower dimensional feature space with $\mathbf{z}_i \in \mathbb{R}^d$ and $d \leq (B + D)$. The transformation matrix \mathbf{W} should not only fuse different features in a lower dimensional feature space, but also preserve local neighborhood information, hence adapting to the manifold embedded in the high-dimensional feature space. A reasonable way to find

²A MATLAB application that implements the proposed LGF method is available on request.

the transformation matrix \mathbf{W} can be defined as follows (details can be found in [27])

$$\arg\min_{\mathbf{W}\in\mathbb{R}^{(B+D)\times d}} \left(\sum_{i,j=1}^{N} \left\| \mathbf{W}^{T} \mathbf{x}_{i} - \mathbf{W}^{T} \mathbf{x}_{j} \right\|^{2} A_{ij} \right)$$
(2)

where the matrix A represents the edges of an undirected graph $\mathbf{G} = (\mathbf{X}, \mathbf{A})$. The adjacency relation between the graph nodes \mathbf{x}_i and \mathbf{x}_j is expressed through binary edge weights $A_{ij} \in$ $\{0,1\}$. In our case, two data points \mathbf{x}_i and \mathbf{x}_j result in adjacent (connected) graph nodes if they are "close" to each other in terms of some distance. Thus, we have $A_{ij} = 1$ if \mathbf{x}_i and \mathbf{x}_j are "close" and $A_{ij} = 0$ if \mathbf{x}_i and \mathbf{x}_j are "far apart." In particular, \mathbf{x}_i will be "close" to \mathbf{x}_i if it belongs to its k nearest neighbors (kNNs). The kNN is determined first by calculating the distance (here we use the Euclidean distance as it is one of the most simple and popular distance measures) between data point x_i and all the other data points \mathbf{x}_i $(j = 1, ..., N, and i \neq j)$, then sorting the distance and determining nearest neighbors based on the kth minimum distance. Thus, such graph contains the interrelations between data points through the connections between its nodes. The effectiveness of using such graph to fuse multiple feature sources for classification has been discussed in the very recent studies [21], [22].

For calculating the pairwise distance matrix to find kNN, the complexity on storing the data and computational time are of $O(N^2)$ and $O(BN^2)$. In this case, even in conventional remote sensing images, the pairwise distance matrix will exceed the memory capacity of ordinary personal computer. For example, an image of $N = 512 \times 512$ pixels, the size of the distance matrix is $N \times N = (512 \times 512) \times (512 \times 512)$ elements. Therefore, in GDF [21], a small number of samples (e.g., n = 5000) was selected from the whole original data to build the global graph. However, random sampling may not always well represent the full data, especially for the data with large study area, which will lead to unstable performances of the global fusion graph. Moreover, image degradation cannot be avoided during the data acquisition, which will lead to poor performances on finding kNN globally from the randomly selected samples to build global fusion graph.

To overcome the above limitations, we propose an LGF method. The proposed LGF probes an image with an $S \times S$ sliding window, calculate the kNN of the current pixel considering the neighboring samples included by the window, and build the fusion graph within this sliding window. Fig. 2 illustrates an example considering a 7×7 sliding window centered at one pixel \mathbf{x}_i^{Spe} . This way, we reduce the computational complexity of calculating pairwise distance matrix to $O(BNS^2)$ ($S \ll N$), as well as a significant reduction in memory use.



Fig. 2. Illustration of the 7×7 sliding window centered at pixel \mathbf{x}_i .

Here, we leverage the fact that pixels within a spatial neighborhood are likely to share similar properties.³ If we consider the spectral features (original HS image), we define the "spectral neighbors" within a spatial neighborhood of a pixel with spectrum \mathbf{x}_i^{Spec} as those k pixels that whose values are closest to it in terms of spectral signatures (i.e., nearest neighbors). Let $\mathcal{N}_k^{Spec}(i)$ denote the set of values of the k nearest neighbors of \mathbf{x}_i^{Spec} within the neighborhood, $\mathcal{N}_k^{Spec}(i) = {\{\mathbf{x}_{im}^{Spec}\}_{m=1}^k}$ and $m \neq i$, $|\mathcal{N}_k^{Spec}(i)| = k$ ($|\cdot|$ being the cardinality of the set). Then, the edges in the graph $E_{im}^{Spec} = 1$ for $m \in \mathcal{N}_k^{Spec}(i)$ and $E_{im}^{Spec} = 0$; otherwise, $m \in \{1, \ldots, S^2\}$. Similarly, for the spatial features (i.e., EMP built on the HS image), the k nearest neighbors of the data point with values \mathbf{x}_i^{Spat} are those of the neighborhood that are most similar to it in terms of the spatial characteristics. Analogously, $\mathcal{N}_k^{Spat}(i) = k$, and thus $E_{im}^{Spat} = 1$ if $m \in \mathcal{N}_k^{Spat}(i)$ and $E_{im}^{Spat} = 0$ otherwise. We propose a novel method to construct the fused kNN for the stacked features \mathbf{X}^{Sta} within a spatial window as follows:

$$\mathcal{N}_{k}^{Fus}(i) = \mathcal{N}_{k}^{Spec}(i) \cap \mathcal{N}_{k}^{Spat}(i)$$
(3)

where the operator ' \cap ' denotes the intersection, i.e., the kNN of the stacked vector \mathbf{x}_i^{Sta} : $\mathcal{N}^{Fus}(i) = \{\mathbf{x}_{im}^{Sta}, m \in \mathcal{N}_k^{Spec}(i) \land m \in \mathcal{N}_k^{Spat}(i)\}$. The fused edge \mathbf{E}_i^{Fus} for the stacked data point \mathbf{x}_i^{Sta} must satisfy

$$E_{i,m}^{Fus} = 1, \quad \text{iff} \quad m \in \mathcal{N}_k^{Spec}(i) \land m \in \mathcal{N}_k^{Spat}(i).$$
(4)

For instance, within the 7×7 sliding window centered at pixel \mathbf{x}_i (when the sliding window is close to the image boundary, the symmetric padding is utilized to deal with the margin effect [29], suppose the 6 nearest neighbors of spectral feature point \mathbf{x}_i^{Spec} is $\mathcal{N}_6^{Spec}(i) = {\mathbf{x}_{im}^{Spec} : m \in [2, 6, 11, 15, 23, 36]}$, see Fig. 3. While the 6 nearest neighbors of spatial feature point \mathbf{x}_i^{Spat} is $\mathcal{N}_6^{Spat}(i) = {\mathbf{x}_{im}^{Spat} : m \in [2, 7, 13, 28, 36]}$. Therefore, we can get the kNN of fusion graph $\mathcal{N}_6^{Fus}(i) = {\mathbf{x}_{im}^{Fus} : m \in [2, 36]}$ according to (3). Then, we set their corresponding edges $E_{im}^{Fus} = 1$, for $m \in \mathcal{N}_6^{Fus}(i)$; $E_{im}^{Fus} = 0$ if $m \notin \mathcal{N}_6^{Fus}(i), 1 \le m \le 49$.

This means that the stacked data point \mathbf{x}_i^{Sta} is "close" to \mathbf{x}_{im}^{Sta} only if they have similar spectral and spatial characteristics



Fig. 3. Illustration of local fusion graph building within a 7 × 7 sliding window centered at pixel. The value '1' means connection (i.e., $E_{im} = 1$), while the blank grid means the data point \mathbf{x}_{im} not in the *k*NN of the current pixel \mathbf{x}_i (i.e., $E_{im} = 0$).

within a spatial window. If any individual feature point \mathbf{x}_i^{spec} (or \mathbf{x}_i^{Spat}) is "far apart" from \mathbf{x}_{im}^{Spec} (or \mathbf{x}_{im}^{Spat}), then $E_{im}^{Fus} = 0$. For example, suppose that the data point \mathbf{x}_i^{Sta} belongs to a road and \mathbf{x}_{im}^{Sta} belongs to a flat roof. Since, in practice, roads and roofs are often made with similar materials (e.g., asphalt), the corresponding data points are likely to have similar spectral characteristics $(E_{im}^{Spec} = 1)$, but different spatial information (e.g., shape and size) $(E_{im}^{Spat} = 0)$, so these two data points are not "close" (i.e., $E_{im}^{Fus} = 0$). Similarly, if \mathbf{x}_i^{Sta} and \mathbf{x}_{im}^{Sta} are taken from the grass areas and parking lot, respectively, they will have different spectral characteristics $(E_{im}^{Spec} = 0)$, and even if they might be similar spatially $(E_{im}^{Spat} = 1)$, the resulting $E_{im}^{Fus} = 0$ characterizing these two data points as "far apart." If the fusion graph was globally constructed by using the whole hyperspectral image or randomly selected samples like [21], one may find the kNN of a pixel belonging to a roof (e.g., shopping mall) \mathbf{x}_{i}^{Fus} in pixels belonging to parking lots, because they have very similar spectral and spatial information even though they might be not spatially adjacent. By building a local fusion graph within a spatial window, the proposed LGF overcomes this limitation, and better models the local spatial neighborhood information. In addition, the proposed LGF is robust to image degradation (e.g., noise), which cannot be avoided during the hyperspectral image acquisition (especially when the spectral bands are in correspondence to windows in the electromagnetic spectrum in which the absorption of the atmosphere is high). The spectra of the same land cover type might exhibit a high variability. This is due to different factors such as the intrinsic variability of the reflectance, differences in illumination, and image artifacts. However, typically, the spectra of pixels belonging to the same object are correlated even if they might differ to those of objects of the same thematic class but located in other parts of the image for the above-mentioned reasons. Thus, by looking for the kNN within a spatial neighborhood can enforce to establish among pixels relations that are meaningful (in terms of representation of the objects). In a similar fashion, the approaches in [28] and [30] showed better denoising results and efficient target detection by considering a local neighborhood.

Then, we can rearrange the edge of each stacked data point \mathbf{E}_i^{Fus} into a sparse matrix \mathbf{A}^{Fus} by using

$$A_{ij}^{Fus} = \begin{cases} E_{ij}^{Fus}, & \text{if } j \in \mathcal{N}^{Fus}(i), j \in [1, \dots, N] \\ 0, & \text{otherwise}, \end{cases}$$
(5)

³This assumption is particularly valid when dealing with images of very-high spatial resolutions.

The matrix $\mathbf{A}^{Fus} \in \mathbb{R}^{N \times N}$ represents the adjacency relation of all data points (e.g., full edge) built on the stacking features (i.e., $\mathbf{G}^{Fus} = (\mathbf{X}^{Sta}, \mathbf{A}^{Fus})$). When using the same constraint in [21] for avoiding degeneracy

$$\mathbf{W}^{T}\left(\mathbf{X}^{Sta}\right)\mathbf{D}^{Fus}\left(\mathbf{X}^{Sta}\right)^{T}\mathbf{W} = \mathbf{I}$$
(6)

where \mathbf{D}^{Fus} is a diagonal matrix with $D_{ii}^{Fus} = \sum_{j=1}^{n} A_{ij}^{Fus}$ and I the identity matrix. We can obtain the transformation matrix $\mathbf{W} = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_r)$ which is made up of r eigenvectors associated with the least r eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_r$ of the following generalized eigenvalue problem:

$$\left(\mathbf{X}^{Sta}\right)\mathbf{L}^{Fus}\left(\mathbf{X}^{Sta}\right)^{T}\mathbf{w} = \lambda\left(\mathbf{X}^{Sta}\right)\mathbf{D}^{Fus}\left(\mathbf{X}^{Sta}\right)^{T}\mathbf{w}$$
(7)

where $\mathbf{L}^{Fus} = \mathbf{D}^{Fus} - \mathbf{A}^{Fus}$ is the fusion Laplacian matrix.

The size of the sliding window has significant influence on the preservation of local spatial neighborhood information (e.g., texture). On one hand, when the window size is too small, the neighborhood contains too few samples for properly modeling the local spatial information and $\mathcal{N}_{k}^{Fus}(i)$ is composed of almost all data points within the window. On the other hand, if the window is too large, then the local spatial information might not be retrieved. In the case limit in which the neighborhood is the whole image, the proposed LGF equals to GDF (thus, GDF can be considered a special case of LGF). In our experiments, we set the sliding window with a fixed intermediate size, and change k nearest neighbors to obtain a satisfying result. By building a local fusion graph within a sliding window, we not only reduce memory cost and computational complexity, but also increase the preservation of local spatial neighborhood information. The algorithmic procedure of the proposed method which uses LGF to couple data fusion and dimension reduction of spectral and spatial features for classification is formally stated in Algorithm 1.

When dealing with high-resolution hyperspectral data, we can fast implement the proposed LGF by spatially downsampling the original HS image and EMP to the same ratio. The main spatial structure of the objects in a high-resolution remote sensing image will be preserved after spatially downsampling within a certain value of downsampling ratio. This way, the proposed LGF will keep stable on the classification performances while reducing the computational complexity. The computational complexity can be reduced by a factor equal to the square of the spatial downsampling ratio. For example, if we downsample original HS image by a factor of R (e.g., R = 4) along both spatial directions, the total number of spatial pixels can be reduced to N/R^2 ; thus, the computational complexity is reduced to $O(BNS^2/R^2)$.

IV. EXPERIMENTAL RESULTS

A. Hyperspectral Image Data Sets

Experiments were run on two data sets, namely the 'Indian Pine' and 'University Area'. The first data set was captured by airborne visible/infrared imaging spectrometer (AVIRIS) over northwestern Indiana in June 1992, with 220 spectral bands **Algorithm 1.** Proposed LGF of spectral and spatial information for classification

- 1: Build the Extended Morphological Profile (EMP) on the first p PCs (usually with cumulative variance near to 99%) of the original hyperspectral data sets. Actually, the EMP are defined th We used a symmetric padding to avoid the margin effect, when the sliding window closed to the margin of image same way as [5]. An MP consist of the original image (one of the PC features) and M openings with SE of increasing size (all applied on the original image) and M closings with the same SE. Then, an EMP is obtained with $d = p \times (2M + 1)$ dimension;
- 2: Find k nearest neighbors for each pixel within a $S \times S$ sliding window. For example, a spectral data point \mathbf{x}_{i}^{Spec} with its kNN index can be denoted by $\mathcal{N}_{k}^{Spec}(i) = {\mathbf{x}_{im}^{Spec}, 1 \le m \le S^2}$. Whereas a spatial data point \mathbf{x}_{i}^{Spat} with its kNN index included by the sliding window can be denoted by $\mathcal{N}_{k}^{Spat}(i)$;
- 3: Construct the local fusion graph for stacked features by using equations (3), (4) and (5);
- 4: Compute the eigenvectors for the generalized eigenvalue problem in (7). The projection matrix $\mathbf{W} = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_r)$ is made up by the *r* eigenvectors associat We used a symmetric padding to avoid the margin effect, when the sliding window closed to the margin of imageed with the least *r* eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_r$;
- 5: Obtain the fused features by projecting the high dimensional staked features $(\mathbf{x}_i^{Sta} \in \mathbb{R}^{B+D})$ into a lower dimensional subspace $(\mathbf{z}_i \in \mathbb{R}^d)$ with equation (1);
- 6: Use the fused features **Z** in the lower dimensional subspace as an input to do classification.

in the wavelength range $0.4-2.5 \,\mu\text{m}$ and low spatial resolution of 20 m by pixel. The calibrated data are available online (along with detailed ground-truth information).⁴ The second data set was acquired over an urban areas in the city of Pavia, Italy. The data were collected by the reflective optics system imaging spectrometer (ROSIS) sensor, with 115 spectral bands in the wavelength range $0.43-0.86 \,\mu\text{m}$ and very fine spatial resolution of 1.3 m by pixel.

Indian pines: The whole scene $(145 \times 145 \text{ pixels})$ contains 16 classes, ranging in size 20–2468 pixels. We keep all 220 bands (including some noisy bands) to see the effect of noise on the classification. In our experiments, classes with less than 30 labeled pixels were removed, resulting thus in 14 classes with available labeled samples in Table II. Note that the color in the cell denotes different classes in the classification maps (Fig. 4).

University area: The image composed of 610×340 pixels was collected over the University of Pavia, Italy, and contains 103 spectral channels after removal of noisy bands. This data set includes nine land cover/use classes, see Table III. Note that the color in the cell denotes different classes in the classification maps (Fig. 5). Available training and testing sets are

⁴[Online]. Available: http://cobweb.ecn.purdue.edu/~biehl/

given in Table III (# number of training samples /# number of test samples).

B. Experimental Setup

Prior to applying the MPs to hyperspectral images, PCA was first applied to the original hyperspectral data set, and the first few PCs (the first 4 PCs for *Indian pine* and the first 3 for *University area*) were selected (representing 99% of the cumulative variance) to construct the EMP. A circular SE ranging from 1 to 10 with step size increment of 1 was used. Ten openings and closings were computed for each PC, resulting in an EMP of dimension 84 for *Indian pine* and 63 for *University area*.

We used an SVM [31] classifier, as it performs well even with a limited number of training samples, limiting the Huges phenomenon. The SVM classifier with radial basis function (RBF) kernels in MATLAB SVM toolbox, LIBSVM [32], is applied in our experiments. SVM with RBF kernels has two parameters: 1) the penalty factor C and 2) the RBF kernel widths γ . We apply a grid-search on C and γ using fivefold cross-validation to find the best C within the given set $\{10^{-1}, 10^{0}, 10^{1}, 10^{2}, 10^{3}\}$ and the best γ within the given set $\{10^{-3}, 10^{-2}, 10^{-1}, 10^{0}, 10^{1}\}$. We compare our proposed LGF with the schemes of 1) using original HS image (Raw_{HSI}); 2) using EMP computed on the first three PCs of the original HS image (EMP $_{HSI}$); 3) stacking all feature sources together, i.e., \mathbf{X}^{Sta} (STA); 4) stacking all the features extracted by PCA from each individual feature source (PCA); 5) stacking all the features extracted by NWFE [33] from each individual feature source (NWFE), similar as [19]; and 6) the GDF [21] with its extension to fuse two feature sources. The classification results are quantitatively evaluated by measuring the overall accuracy (OA), the average accuracy (AA), and the kappa coefficient (κ). The experiments were carried out on 64-b, 3.40 GHz Intel i7-4930K (1 core) CPU computer with 64 GB memory. The consumed time reported in our experiments includes both feature fusion and parameters optimization for SVM classification.

C. Results on Indian Pines Data Set

In this experiment, the whole data of *Indian pines* was used to construct the fusion graph in both GDF [21] and the proposed LGF. We set the size of sliding window to 15×15 , and kNN to 30. Twenty samples per class were randomly selected from the labeled data set to train SVM classifiers; all results were evaluated against the remaining labeled samples in the ground truth. After repeating the selection of training samples and classification process five times, we report the mean classification results and their standard deviation in Table I. In order to compare the accuracy for each class and final classification maps, we show the best results (in term of OA) of each method in Table II and Fig. 4. From the tables and figure, we have the following findings.

1) The results confirm that it is better sometimes to use single feature source than simply stacking many of them for

 TABLE I

 AVERAGE CLASSIFICATION ACCURACIES ON Indian Pines USING SVM

	Features	OA	AA	κ	Std
Raw _{HSI}	220	55.37	66.32	50.32	1.91
EMP_{HSI}	84	82.39	88.29	80.11	1.90
STA	304	71.76	81.13	68.30	1.88
PCA	14 (4.10)	77.83	85.49	75.07	1.01
NWFE	42 (12,30)	87.59	91.63	85.93	1.48
GDF	40	83.59	88.70	81.49	1.87
Proposed LGF	40	93.05	94.83	92.08	0.89

The number in brackets is the number of features from the spectral feature and spatial feature, respectively.

classification. Compared to the situation with single spatial features (EMP_{HSI}), the OA of simply stacking original spectral and spatial features (STA) decreases more than 10 percentage points, while increasing the dimensionality. Our proposed LGF produced the best results, with OA improvements of 10.66–37.68 percentage points over only using the single spectral/spatial feature source, with improvements of 5.46–21.29 percentage points over stacking both the spectral and the spatial features by STA, PCA, and NWFE, and with 9.46 percentage points improvement over the GDF [21].

2) From the class-specific accuracies, the EMP_{HSI} approach performed much better for most classes than the Raw_{HSI} approach, especially for some classes (e.g., 'Corn-min' and 'Soybean-clean,' and 'Bldgs-Grass-Trees-Drives'), with more than 40 percentage points improvement in accuracy. By simply stacking original feature sources or stacking features extracted by PCA representing more than 99% of the cumulative variance, the accuracies on classes 'Corn' and 'Soybean-mintill' drop almost by 20 percentage points compared to the EMP_{HSI} approach. By stacking spectral and spatial features extracted by NWFE representing more than 90% of the cumulative variance, better accuracies were produced in classes 'Corn-notil' and 'Soybean-clean' (with OA improvements of 13.39-37.66 and 10.42-50.63 percentage points over Raw_{HSI} and EMP_{HSI}), but the performance dropped significantly on classes 'Corn' compared to EMP_{HSI} . By building fusion graph on the whole data, the GDF approach performed much better than both Raw_{HSI} and EMP_{HSI} on classes 'grasspasture' and 'grass-tress' (with OA improvements of 8.86–13.69 and 8.97–18.95 percentage points), but much worse on class 'Corn' and 'Soybean-clean' compared to only using single spatial feature source. Our proposed LGF demonstrated better performance on almost all the classes than the methods that use using single features (Raw_{HSI} and EMP_{HSI}), stacked features (i.e., STA, PCA, and NWFE), and the GDF, and produced much better results on classes 'Corn-notil' and 'Corn-min.' For class 'Corn-notil,' the proposed LGF approach had more than 25, 10, and 25 percentage points improvement compared to the approaches using single features, stacked features, and GDF, respectively. For the class 'Corn,' both the approaches of stacked features and GDF performed better than only using spectral feature, but worse than only using the spatial feature, while our

	Raw _{HSI}	EMP _{HSI}	STA	PCA	NWFE	GDF	Proposed LGF
Number of features	220	84	304	14(4,10)	42(12,30)	40	40
OA (%)	57.48	84.34	73.93	79.10	88.68	86.53	94.01
AA (%)	65.91	89.41	82.21	86.00	91.55	89.93	95.17
κ	0.522	0.823	0.708	0.765	0.871	0.848	0.932
Consumed time (s)	13.91	4.49	16.59	3.41	6.55	63.03	58.85
(#54) Alfalfa	88.89	100	98.15	96.30	98.15	98.15	98.15
(#1434) Corn-notil	38.70	62.97	56.00	70.29	76.36	64.71	89.75
(#834) Corn-min	46.76	86.93	79.91	76.98	84.29	83.81	96.64
(#234) Corn	54.27	94.02	79.91	78.63	79.06	87.61	97.01
(#497) Grass-Pasture	82.29	87.12	89.34	91.15	93.96	95.98	95.57
(#747) Grass-Trees	78.71	88.49	91.97	95.72	96.92	97.46	95.45
(#489) Hay-windrowed	70.14	99.39	99.39	98.98	99.80	99.59	98.16
(#968) Soybean-Notill	46.28	82.64	66.63	82.75	79.24	88.02	89.67
(#2468) Soybean-mintill	46.19	84.89	65.03	65.68	87.84	85.41	93.23
(#614) Soybean-clean	41.69	82.90	69.22	69.38	93.32	73.94	85.99
(#212) Wheat	95.75	99.53	97.64	99.53	99.53	99.53	99.53
(#1294) Woods	89.88	88.72	84.00	86.01	97.22	99.23	99.77
(#380) Bldgs-Grass-Trees-Drives	44.21	96.32	89.47	94.74	96.05	93.95	97.63
(#95) Stone-steel-towers	98.95	97.89	95.79	97.89	100	91.58	95.79

TABLE II RESULT FOR Indian Pines WITH BEST CLASSIFICATION ACCURACY OVER 5 RUNS

Twenty training samples per class with SVM classifier were used.



(a)



(d)



Fig. 4. Classification maps for Indian pines. (a) RGB composition with 14 classes labeled and highlighted in the image, and thematic map using (b) original HS image, (c) EMP, (d) stacked features by original HS image and EMP, (e) PCA, (f) NWFE, (g) GDF [21], and (h) proposed LGF.

approach produced best result, even 3 percentage points higher than only using the spatial feature.

3) From the classification maps, we can see visually that combining spatial information will produce smoother results. In particular, the proposed method leads to smoother classification maps than the other methods, with more stable performances (lower standard deviation in Table I). The processing time of the proposed method is faster than GDF, but slower than other schemes. This is because of limited training samples. With larger number

of training samples, the consumed time of extracting features by NWFE and using the features with higher dimension for classification will increase.

The hyperspectral remote sensing data contain a wealth of spectral and spatial information. Only using single spectral/spatial feature is not enough for a reliable classification. The approaches of PCA and NWFE are similar to the STA approach in terms of a stacked architecture, all these three approaches first applied feature extraction on original HS data and EMP, then concatenated the extracted feature vectors from

	Raw _{HSI}	EMP _{HSI}	STA	PCA	NWFE	GDF	Proposed LGF
Number of features	103	63	166	42(2,40)	40(8,32)	36	28
OA (%)	79.75	84.61	85.04	93.29	93.15	94.16	98.29
AA (%)	88.26	91.25	91.93	92.51	93.92	94.68	98.66
κ	0.747	0.802	0.809	91.05	0.909	0.923	0.977
Consumed time (s)	65.05	27.28	75.42	25.67	876.78	22.69	20.46
(#548 / #6641) Asphalt	84.17	95.04	93.14	98.45	97.74	93.89	98.45
(#540 / #18649) Meadows	67.42	76.79	75.82	95.79	94.71	93.77	97.95
(#392 / #2099) Gravel	73.70	82.18	80.51	77.99	82.28	77.94	98.67
(#524 / #3064) Trees	94.78	96.77	97.10	84.86	97.94	93.24	93.77
(#265 / #1345) Metal Sheets	99.63	99.93	99.93	99.85	99.93	99.78	99.78
(#532 / #5029) Soil	92.30	71.72	82.04	80.25	73.61	95.67	99.88
(#375 / #1330) Bitumen	91.20	99.77	99.77	99.92	99.85	99.25	99.92
(#514 / #3682) Bricks	91.47	99.27	99.29	99.40	99.27	99.32	99.51
(#231 / #947) Shadows	99.68	99.79	99.79	96.09	100	99.26	100

 TABLE III

 CLASSIFICATION ACCURACY FOR University Area WITH SVM CLASSIFIER

We built the local fusion graph of the proposed LGF on both the downsampled original HS image and EMP of factor 5.

both the original HS data and the EMP into one stacked vector. The differences are that each individual feature has different characteristics, e.g., the features extracted by PCA represent most of the cumulative variance in the data, while the features extracted by NWFE respect the class discriminant. In this case, the supervised NWFE performs better than unsupervised PCA. The class 'Corn' is not classified well by fusing features in a stacked architecture. The approaches of PCA, NWFE, and STA produced lower accuracies than only using single spatial feature, indicating that the spatial information contained in the original EMP was not well exploited in such a stacked architecture. The performances of GDF is not better than NWFE even with fusion graph built on all data point of the original feature sources. This is not surprising because the original HS image contains noise [see Fig. 4(b)], thus may affect the kNNsearching in fusion graph building. By building local fusion graph with a spatial window, the proposed LGF method better preserves local spatial information, and more robust to image noise.

D. Results on University Area Data Set

In order to make fair comparisons, for the approaches of PCA and NWFE in all our experiments, we use the best combination of the extracted spectral and the extracted spatial features for the classification. We search the best combination of the spectral and the spatial dimensions using the crossvalidation according to the OA, with both the spectral dimension and the spatial dimension ranging from 2 to 40 (with step size increment of 2). The best combination is obtained when OA reaches the maximum; Fig. 6(a) and (b) shows the results. 5000 samples were randomly selected to build the global fusion graph in GDF, similar as we did in [21]. Fig. 6(c) shows the performance with different number of nearest neighbors and dimensions (of fused features). For the proposed LGF, we first downsampled both original HS image and EMP of factor 5 on both spatial directions to speed up the processing time, and set the size of sliding window to 15×15 . Table IV reports

the accuracies and consumed time as the downsampled size changes. The classification results using the best combination are shown in Table III and Fig. 5.

From the tables and figure, we can make the following remarks.

- 1) The results confirm that the integration of spectral and spatial features can improve the classification performance on HS images. Compared to the situation with single spectral or spatial feature, the OA of stacking spectral and spatial features has 8.68-13.54 and 8.54-13.4 percentage points improvements for PCA and NWFE, respectively. The improvements of simply stacking original spectral and spatial features (STA) over only using the single spectral/spatial features is not significant, while increasing both the dimensionality and computational time. Our proposed LGF produced better results, with OA improvements of 13.68–18.54 percentage points over only using the single spectral/spatial feature, with OA improvements of 5–13.25 percentage points over stacking both the spectral and the spatial features by PCA, NWFE, and STA, and with 4.13 percentage points improvement over GDF. As far as we know, these accuracies are higher than all the previously reported in the literature for this data set with SVM classifier and without postprocessing [15], [16].
- 2) From the class-specific accuracies, the EMP_{HSI} approach performed much better for most classes than the Raw_{HSI} approach, especially for the classes 'asphalt' and 'meadows,' with more than 10 percentage points improvement in accuracy. However, the EMP_{HSI} approach produced much worse accuracy in class 'soil,' dropping by 20 percentage points compared to the Raw_{HSI} approach. By stacking spectral and spatial features extracted by PCA/NWFE, better accuracies were produced in class 'meadows' (with improvements of 19–28.37 and 17.92–27.29 percentage points, respectively, over Raw_{HSI} and EMP_{HSI}), but the performance dropped significantly on classes 'soil' compared to



Fig. 5. Classification maps produced by the described schemes. (a) Test samples and thematic maps using (b) original HS data, (c) EMP of HS data, (d) PCA, (e) NWFE, (f) stacked features \mathbf{X}^{Sta} , (g) GDF [21], and (h) proposed LGF.



Fig. 6. Surface of the classification accuracies as a function of (a) number of extracted spectral features and spatial features by PCA; (b) number of extracted spectral features and spatial features by NWFE; (c) number of extracted features and K nearest neighbors in GDF [21]; and (d) number of extracted features and K nearest neighbors in the proposed LGF.

Raw_{HSI}. By building the fusion graph on randomly selected samples, the GDF approach consumes less time and performed much better than both Raw_{HSI} and EMP_{HSI} on classes 'meadows' and 'soil' (with OA improvements of 16.98–26.35 and 3.37-23.95 percentage points), but worse on class 'gravel' compared to only using the spatial features. Our proposed LGF demonstrated better performance on almost all the classes than the methods that use using single features (Raw_{HSI})

and EMP_{*HSI*}), stacked features (i.e., PCA, NWFE, and STA), and the GDF, and produced much better results on classes 'gravel' and 'soil.' For class 'gravel,' the proposed LGF approach had improvements of 16.49-24.97, 16.39-20.68, and 20.73 percentage points compared to the approaches using single features, stacked features, and GDF, respectively.

3) Using only single spectral/spatial feature is not enough to get very accurate result, see Fig. 6(a) and (b). In order



Fig. 7. Surface of the classification accuracies as a function of the size of sliding window and the number of K nearest neighbors in the proposed LGF.

to get the best OA, PCA needs 2 spectral features and 40 spatial features, and NWFE needs 8 spectral features and 32 spatial features. In GDF and our proposed LGF, the selection of dimension becomes much easier, see the surface plots of the classification accuracies as a function of dimension and the number of nearest neighbors in Fig. 6(c) and (d). The dimension of extracted features varies from 2 to 40 (with step size increment of 2) and the number of nearest neighbors k varies from 10 to 50 (with step size of 2). This latter analysis leads us to state that the proposed is not sensitive to the parameter k, as the dimension increases to 36. The size of the sliding window has significant influence on the preservation of local spatial neighborhood information. When the window size is too small, the neighborhood contains too few samples for properly modeling the local spatial information. Fig. 7 shows that small size of sliding window with bigger kleads to poor performances of classification. If the window is too large, then the local spatial information might not be retrieved, resulting decreased classification accuracies, see Fig. 7. With sliding window sizes varying from 13 to 21 and nearest neighbors from 20 to 40, we can get satisfying results for University area.

When stacking different features extracted by methods like PCA and NWFE, it is not easy to select the optimal combination of the spectral and the spatial dimensions, as was also discussed by Fauvel *et al.* in [19]. These optimal combinations of spectral and spatial dimensions are different for different data sets. Even for the same data set, when the training sample size changes, the combination of spectral and spatial dimensions will change.

Many approaches selected the optimal combination of spectral and the spatial dimensions according to the cumulative variance [19]. However, these approaches do not always work well. For example, in PCA, the number of PCs which represent more than 99% of the cumulative variance depends on the statistical distribution of the data. The extracted PCs which represent 99% of the cumulative variance may not contain enough information of the data, resulting in a worse performance. When

TABLE IV University Area: THE ACCURACIES AND CONSUMED TIME AS THE DOWNSAMPLED SIZE OF ORIGINAL FEATURE SOURCES INCREASES

	1×1	2×2	3×3	4×4	5×5
Consumed time (s)	1551.5	136.7	32.1	14.3	8.5
OA (%)	97.11	98.23	97.56	97.63	98.29
AA (%)	97.04	98.58	98.12	98.11	98.66
κ	0.962	0.977	0.968	0.969	0.977

 2×2 means we downsample both the original HS image and EMP of a factor 2 on both spatial directions.

the data contain non-Gaussian noise, the number of PCs needed to reach 99% of the cumulative variance is higher, which may contain redundant information. Although some algorithms (e.g., cross-validation) can be used to find the best combination of dimensions, it increases the processing time. In our experiments, the elapsed time of searching the best combination is 7.99 and 7.63 h for PCA and NWFE, respectively.

The performances of the proposed LGF are less sensitive to the values of the free parameters. We keep the parameters (the number of nearest neighbors and the number of extracted features) the same for feature sources with different downsampling ratios, see Table IV. We get very similar classification results for *University area*, with processing time dropping from 1551.5 to 8.5 s. Downsampling might cause a reduction in the intraclass heterogeneity (i.e., objects belonging to the same class will be more spectrally similar). If the training samples are taken far from the objects' edges, they will likely correspond to areas of a unique thematic class (i.e., they do not correspond to mixed pixels) leading to a simpler classification problem.

V. CONCLUSION

In this paper, we present a novel method for LGF of spectral and spatial information. The morphological features, which are used to characterize the spatial information, are first generated on the first few PCs of HS image. Then, we build a local fusion graph within a sliding window where only the feature points that have similar both spectral and spatial characteristics are connected. Finally, we solve the problem of data fusion by projecting all the features onto a linear subspace, in which local neighborhood data points (i.e., with similar spectral and spatial characteristics) in the high-dimensional feature space are kept on local neighborhood in the low-dimensional subspace as well. The proposed LGF technique effectively employs the local spatial information of different feature sources within a spatial window. This allows to obtain better performances in classification in particular with respect to the GDF approach which is global. In addition, with respect to this latter, we reduce both memory cost and computational complexity for graph building and increase robustness to image noise thanks to considering a small sliding window.

The experiments confirmed expected improvements of such an approach over both stacking different feature sources together and building full fusion graph. The feature stacking approaches experience serious problems in selecting the best combination of the spectral dimension and the spatial dimension, and can be affected by the redundancy in the stacked data. On the other hand, feature extraction on all the features together does not take into account the properties of different feature sources, while building a full fusion graph does not take into account of the local spatial information and may require more computer resources. The proposed LGF overcomes these problems and makes full advantage of both feature sources through the local fusion graph. Classification results on two real HS data show the efficiency of the proposed LGF. Recently, some approaches show great improvements in the classification of remote sensing images by using APs [35] and by combining postprocessing [34], which will be exploited in our future work.

ACKNOWLEDGMENT

The authors would like to thank Prof. P. Gamba from the University of Pavia, Italy, for kindly providing the *University area* data, Prof. Landgrebe for providing the AVIRIS *Indian Pines*, and Prof. Kuo for providing NWFE codes.

REFERENCE

- [1] P. Soille, *Morphological Image Analysis, Principles and Applications*, 2nd ed. Berlin, Germany: Springer-Verlag, 2003.
- [2] P. Soille and M. Pesaresi, "Advances in mathematical morphology applied to geoscience and remote sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 9, pp. 2042–2055, Sep. 2002.
- [3] M. Pesaresi and J. A. Benediktsson, "A new approach for the morphological segmentation of high-resolution satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 2, pp. 309–320, Feb. 2001.
- [4] R. Bellens, S. Gautama, L. Martinez-Fonte, W. Philips, J. C.-W. Chan, and F. Canters, "Improved classification of VHR images of urban areas using directional morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 10, pp. 2803–2812, Oct. 2008.
- [5] J. A. Benediktsson, J. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 480–491, Mar. 2005.
- [6] W. Liao, R. Bellens, A. Pižurica, W. Philips, and Y. Pi, "Classification of hyperspectral data over urban areas using directional morphological profiles and semi-supervised feature extraction," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 4, pp. 1177–1190, Aug. 2012.
- [7] W. Liao, R. Bellens, A. Pižurica, W. Philips, and Y. Pi, "Classification of hyperspectral data over urban areas based on extended morphological profile with partial reconstruction," in *Proc. Adv. Concepts Intell. Vis. Syst. (ACIVS'12)*, 2012, pp. 278–289.
- [8] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "Kernel principal component analysis for the classification of hyperspectral remote-sensing data over urban areas," *EURASIP J. Adv. Signal Process.*, vol. 2009, p. 14, Feb. 2009.
- [9] M. Dalla Mura, J. A. Benediktsson, B. Waske, and L. Bruzzone, "Extended profiles with morphological attribute filters for the analysis of hyperspectral data," *Int. J. Remote Sens.*, vol. 31, no. 22, pp. 5975–5991, Nov. 2010.
- [10] M. Dalla Mura, J. A. Benediktsson, B. Waske, and L. Bruzzone, "Morphological attribute profiles for the analysis of very high resolution images," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3747–3762, Oct. 2010.
- [11] M. Dalla Mura, A. Villa, J. A. Benediktsson, J. Chanussot, and L. Bruzzone, "Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 3, pp. 541–545, May 2011.
- [12] J. Crespo, J. Serra, and R. Shafer, "Theoretical aspects of morphological filters by reconstruction," *Signal Process.*, vol. 47, no. 2, pp. 201–225, Nov. 1995.
- [13] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. Vila-Frances, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 93–97, Jan. 2006.

- [14] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "A spatial-spectral kernel-based approach for the classification of remote-sensing images," *Pattern Recognit.*, vol. 45, no. 1, pp. 381–392, 2012.
- [15] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Advances in spectral-spatial classification of hyperspectral images," *Proc. IEEE*, vol. 101, no. 3, pp. 652–675, Mar. 2013.
- [16] J. Li, P. R. Marpu, A. Plaza, J. M. Bioucas-Dias, and J. A. Benediktsson, "Generalized composite kernel framework for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 9, pp. 4816– 4829, Sep. 2013.
- [17] J. A. Palmason, J. A. Benediktsson, J. R. Sveinsson, and J. Chanussot, "Fusion of morphological and spectral information for classification of hyperspectral urban remote sensing data," in *Proc. Int. Geosci. Remote Sens. Symp. (IGARSS'06*, Jul. 2006, pp. 2506–2509.
- [18] M. Fauvel, J. Chanussot, J. A. Benediktsson, and J. R. Sveinsson, "Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles," in *Proc. Int. Geosci. Remote Sens. Symp.* (*IGARSS'07*), 2007, pp. 4834–4837.
- [19] M. Fauvel, J. A. Benediktsson, J. Chanussot, and J. R. Sveinsson, "Spectral and spatial classification of hyperspectral data using svms and morphological profile," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 11, pp. 3804–3814, Nov. 2008.
- [20] X. Huang and L. Zhang, "An SVM ensemble approach combining spectral, structural, and semantic features for the classification of highresolution remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 257–272, Jan. 2013.
- [21] C. Debes *et al.*, "Hyperspectral and LiDAR data fusion: Outcome of the 2013 GRSS data fusion contest," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2405–2418, Jun. 2014.
- [22] W. Liao, R. Bellens, A. Pizurica, S. Gautama, and W. Philips, "Generalized graph-based fusion of hyperspectral and LiDAR data using morphological features," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 3, pp. 552–556, Mar. 2015.
- [23] Y. Tarabalka, J. A. Benediktsson, and J. Chanussot, "Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2973–2987, Aug. 2009.
- [24] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Spectral-spatial hyperspectral image segmentation using subspace multinomial logistic regression and Markov random field," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 809–823, Feb. 2012.
- [25] G. Camps-Valls, N. Shervashidze, and K. M. Borgwardt, "Spatio-spectral remote sensing image classification with graph Kernels," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 741–745, Oct. 2010.
- [26] B. Scholkopf, A. J. Smola, and K. R. Muller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Comput.*, vol. 10, pp. 1299–1319, 1998.
- [27] M. Belkin and P. Niyogi, "Laplacia eigenmaps and spectral techniques for embedding and clustering," in *Advances in Neural Information Processing Systems 14*, Cambridge, MA, USA: MIT Press, 2002, pp. 585–591.
- [28] G. Chen and S. E. Qian, "Dimensionality reduction of hyperspectral imagery using improved locally linear embedding," J. Appl. Remote Sens., vol. 1, pp. 1–10, Mar. 2007.
- [29] M. D. Jimenez and N. Prelcic, "Linear boundary extensions for einite length signals and paraunitary two-channel filterbanks," *IEEE Trans. Signal Process*, vol. 52, no. 11, pp. 3213–3226, Nov. 2004.
- [30] G. Chen, T. D. Bui, and A. Krzyzak, "Image denoising with neighbour dependency and customized wavelet and threshold," *Pattern Recognit.*, vol. 38, no. 1, pp. 115–124, 2005.
- [31] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," in *Data Mining and Knowledge Discovery*, vol. 2, 1998, pp. 121–167.
- [32] C. C. Chang and C. J. Lin. (2001). LIBSVM: A Library for Support Vector Machines [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm
- [33] B. C. Kuo and D. A. Landgrebe, "Nonparametric weighted feature extraction for classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 5, pp. 1096–1105, May 2004.
- [34] X. Huang, Q. Lu, L. Zhang, and A. Plaza, "New postprocessing methods for remote sensing image classification: A systematic study," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 11, pp. 7140–7159, Nov. 2014.
- [35] M. Pedergnana, P. Reddy Marpu, M. Dalla Mura, J. A. Benediktsson, and L. Bruzzone, "Classification of remote sensing optical and LiDAR data using extended attribute profiles," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 7, pp. 856–865, Nov. 2012.



Wenzhi Liao received the B.S. degree in mathematics from Hainan Normal University, HaiKou, China, in 2006, the Ph.D. degree in engineering from the South China University of Technology, Guangzhou, China, in 2012, and the Ph.D. degree in computer science engineering from Ghent University, Ghent, Belgium, in 2012.

Since 2012, he has been working as a Postdoc with Ghent University. His research interests include pattern recognition, remote sensing, and image processing, mathematical morphology, multitask feature

learning, multisensor data fusion, and hyperspectral image restoration.

Dr. Liao is a Member of the Geoscience and Remote Sensing Society (GRSS) and IEEE GRSS Data Fusion Technical Committee (DFTC). He was the recipient of the "Best Paper Challenge" Awards on both 2013 IEEE GRSS Data Fusion Contest and 2014 IEEE GRSS Data Fusion Contest.

Mauro Dalla Mura (S'08–M'11) received the B.E. (Laurea) and M.E. (laurea specialistica) degrees in telecommunication engineering from the University of Trento, Trento, Italy, in 2005 and 2007, respectively, and the joint Ph.D. degree in information and communication technologies (telecommunications area) from the University of Trento and in electrical and computer engineering from the University of Iceland, Reykjavik, Iceland, in 2011.

In 2011, he was a Research fellow with Fondazione Bruno Kessler, Trento, Italy. He is currently an

Assistant Professor with Grenoble Institute of Technology (Grenoble INP), Grenoble, France. He is conducting his research at the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA-Lab). His research interests include remote sensing, image processing, pattern recognition, mathematical morphology, classification, and multivariate data analysis.

Dr. Mura is a Reviewer of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, the IEEE JOURNAL OF SELECTED TOPICS IN EARTH OBSERVATIONS AND REMOTE SENSING, the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, Pattern Recognition Letters, ISPRS Journal of Photogrammetry and Remote Sensing, PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING (PE&RS). He is a Member of the Geoscience and Remote Sensing Society (GRSS) and IEEE GRSS Data Fusion Technical Committee (DFTC), and Secretary of the IEEE GRSS French Chapter (2013-2016). He was a lecturer at the RSSS12-Remote Sensing Summer School 2012 (organized by the IEEE GRSS), Munich, Germany. He was the recipient of the IEEE GRSS Second Prize in the Student Paper Competition of the 2011 IEEE IGARSS 2011 and co-recipient of the Best Paper Award of the International Journal of Image and Data Fusion for the year 2012-2013 and the Symposium Paper Award for IEEE IGARSS 2014.



Jocelyn Chanussot (M'04–SM'04–F'12) received the M.Sc. degree in electrical engineering from the Grenoble Institute of Technology (Grenoble INP), Grenoble, France, in 1995, and the Ph.D. degree from Savoie University, Annecy, France, in 1998.

In 1999, he was with the Geography Imagery Perception Laboratory for the Delegation Generale de l'Armement (DGA—French National Defense Department), France. Since 1999, he has been with Grenoble INP, where he was an Assistant Professor from 1999 to 2005, an Associate Professor from 2005

to 2007, and is currently a Professor of Signal and Image Processing. He is conducting his research at the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA-Lab). His research interests include image analysis, multicomponent image processing, nonlinear filtering, and data fusion in remote sensing. He has been a Visiting Scholar at Stanford University, Stanford, CT, USA, KTH, Stockholm, Sweden, and NUS, Singapore. Since 2013, he has been an Adjunct Professor with the University of Iceland, Reykjavik, Iceland. He has supervised Ph.D. students from 10 different countries (Brazil, China, Egypt, France, Italy, Montenegro, Pakistan, Portugal, Ukrain, and Spain).

Dr. Chanussot was a Member of the IEEE Geoscience and Remote Sensing Society AdCom (2009–2010), in charge of membership development. He was the General Chair of the first IEEE GRSS Workshop on Hyperspectral Image and Signal Processing, Evolution in Remote sensing (WHISPERS). He was the Chair (2009-2011) and Cochair of the GRS Data Fusion Technical Committee (2005-2008). He was a Member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society (2006-2008) and the Program Chair of the IEEE International Workshop on Machine Learning for Signal Processing (2009). He was an Associate Editor for the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS (2005-2007) and for Pattern Recognition (2006–2008). Since 2007, he is an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. Since 2011, he is the Editor-in-Chief of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING. In 2013, he was a Guest Editor for the Proceedings of the IEEE and in 2014 a Guest Editor for the IEEE SIGNAL PROCESSING MAGAZINE. He is a member of the Institut Universitaire de France (2012-2017). He is the Founding President of IEEE Geoscience and Remote Sensing French Chapter (2007-2010), which received the 2010 IEEE GRS-S Chapter Excellence Award. He was the co-recipient of the NORSIG 2006 Best Student Paper Award, the IEEE GRSS 2011 Symposium Best Paper Award, the IEEE GRSS 2012 Transactions Prize Paper Award, and the IEEE GRSS 2013 Highest Impact Paper Award.



Aleksandra Pižurica received the Diploma degree in electrical engineering from the University of Novi Sad, Novi Sad, Serbia, the M.Sc. degree in telecommunications from the University of Belgrade, Belgrade, Serbia, and the Ph.D. degree in engineering from Ghent University, Ghent, Belgium, in 1994, 1997, and 2002, respectively.

She is a Professor in statistical image modeling at Ghent University. She was a Postdoctoral Fellow with the Fund for Scientific Research in Flanders— FWO (2005 to 2011) and was elected as a Principal

Investigator at the Research Department Multimedia Technology, iMinds (since 2009). In 2011, she has founded Statistical Image Modeling Laboratory, Ghent University. Her research interests include multiresolution statistical image modeling, graphical models (Markov random field models for spatial context), image and video restoration (denoising, deblurring, super-resolution, and inpainting) using Bayesian inference, image reconstruction from undersampled measurements and feature extraction from multimodal image data.

Dr. Pižurica currently serves as an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING and was the Lead Guest Editor for EURASIP Journal on Advances in Signal Processing (for the Special Issue on Advanced Statistical Tools for Enhanced Quality Digital Imaging with Realistic Capture Models). She has also served as a program committee Member, Area Chair, and/or conference Co-Chair of several world-renowned workshops and conferences including International Traveling Workshop for Interacting Sparse Models and Technology (iTWIST), IEICE Information and Communication Technology Forum, European Signal Processing Conference (EUSIPCO), the IEEE International Conference on Image Processing (ICIP), and the IEEE International Conference on Acoustic, Speech, and Signal Processing (ICASSP). She is a member of ACM. She was the recipient of the best paper award of the IEEE GRSS (Geoscience and Remote Sensing Society) Data Fusion Contest (twice), in 2013 and 2014. She was also the recipient of the Scientific Prize "de Boelpaepe" from 2013 to 2014, awarded by the Académie Royale des Sciences, des Lettres et des Beaux-arts de Belgique.