Int J Geomath (2016) 7:275–297
DOI 10.1007/s13137-016-0085-0

ORIGINAL PAPER

Spatial-spectral operator theoretic methods for hyperspectral image classification

John J. Benedetto[1] · **Wojciech Czaja**[1] · **Julia Dobrosotskaya**[2] · **Timothy Doster**[1] · **Kevin Duke**[3]

Received: 30 May 2016 / Accepted: 22 June 2016 / Published online: 7 July 2016

Abstract With the emergence of new remote sensing modalities, it becomes increasingly important to find novel algorithms for fusion and integration of different types of data for the purpose of improving performance of applications, such as target/anomaly detection or classification. Many popular techniques that deal with this problem are based on performing multiple classifications and fusing these individual results into one product. In this paper we provide a new approach, focused on creating joint representations of the multi-modal data, which then can be subject to analysis by state of the art classifiers. In the work presented in this paper we consider the problem of spatial-spectral fusion for hyperspectral imagery. Our approach involves machine learning techniques based on analysis of joint data-dependent graphs and the resulting data-dependent fusion operators and their representations.

Keywords Manifold learning · Hyperspectral · Spatial-spectral fusion

Mathematics Subject Classification 42C99

1 Introduction

Hyperspectral imaging (HSI) is among the most significant developments in remote sensing in the recent years. The technology, at its core, recovers radiation reflected

✉ Wojciech Czaja
wojtek@math.umd.edu

1 Department of Mathematics, University of Maryland, College Park, MD, USA

2 Department of Mathematics, Case Western Reserve University, Clevland, OH, USA

3 Department of Mathematics and Statistics, American University, Washington, DC, USA

from the surface of objects across many wavelengths and has a wide range of practical applications, ranging from agriculture, through mineralogy and exploration to applications in the security and defense fields. Farmers are able to utilize hyperspectral images to determine plant stress levels, amount of water being absorbed, and possible insect infestations (Thenkabail et al. 2000). In resource exploration, in particular in mineralogy exploration, hyperspectral images can be quickly obtained for a vast amount of territory and then known spectral signatures corresponding to desirable minerals can be searched for Meer (2004). The defense industry makes wide use of hyperspectral images for target detection and tracking (Manolakis et al. 2003) because, for example, the limitations of the human vision make it difficult to discern modern camouflage from vegetation (Schaum and Stocker 2004).

From a mathematical perspective, a hyperspectral image with hundreds of spectral bands is a far reaching generalization of a standard digital image which has only three spectral bands. Thus, it offers a more complete representation of the light spectrum for viewing and analysis. A regular color image can be interpreted as a collection of three-dimensional spectral vectors, each representing the information for one pixel. Similarly, mathematically, a hyperspectral image, X, can be viewed as a collection of $N \times D$-dimensional spectral vectors, each representing the information for one pixel (Schott 1997). The large number of spectral bands present in hyperspectral images is vital for discerning materials that are spectrally close. This, however, can introduce difficulties for an analyst or algorithm to process. This high dimensionality makes hyperspectral images to be good candidates for the methods of machine learning.

Machine learning is a field at the intersection of mathematics and computer science, which develops algorithms to find representations of training data, which then can be generalized and extended onto unseen data sets. Dimension Reduction represents one family of machine learning algorithms. Dimensionality Reduction algorithms deal with the complexities of large data sets which contain many variables, and they attempt to reduce the dimensionality of the data while preserving its most important characteristics. Such algorithms are becoming more important today because the complexity of sensors have increased, as well as our ability to store massive amounts of data. For example, hyperspectral sensors mentioned above record hundreds of times the amount of information that is found in a typical RGB sensor. With this high number of dimensions being recorded it becomes no longer feasible for analysts to examine the data without the assistance of computer algorithms to reduce the number of dimensions but still keep the intrinsic structure of the data intact.

State of the art methods of analyzing hyperspectral data have primarily focused on studying the spectral information contained in the $N \times D$ data set. However, spatial information is intrinsically included within hyperspectral images. The data is acquired as an $m \times n \times D$ data cube ($N = mn$), with each D-dimensional vector's location in the $m \times n$ array corresponding to its physical location in the scene. Often, spatial information has largely been ignored in hyperspectral image analysis.

Recent work has, however, sought to include this spatial information into a coherent product that improves knowledge (when compared to a single sensor) by means of data fusion. Data fusion, in regards to spatial-spectral fusion, is the process of assembling multiple heterogeneous or homogeneous data sources with the well studied spectral information. Current strategies to integrate spatial and spectral information include

the use of wavelet packets (Benedetto et al. 2010, 2013), modified distance metrics for graph construction (Mohan et al. 2007), combining spectral classification with image segmentation results (Tarabalka et al. 2009), spatially weighted normalized cuts (Gillis and Bowles 2012), and adding spatial information in spectral band grouping (Lee et al. 2011). Other approaches include utilizing iterative information spreading on the data-dependent graphs with SVM to label unknown classes (Camps-Valls et al. 2006, 2007, 2008), using morphological profiles to add spatial information (Fauvel et al. 2008), creating end member sets which are spatially motivated (Plaza et al. 2002), multilevel segmentation (Bruzzone and Carlin 2006) or combining segmentation with pixel classification maps (Tarabalka et al. 2009).

More recently, spatial-spectral fusion techniques have relied on graph-based representations such as, e.g., the graph Laplacian. A comparative study of the spatial and spectral graph representations, together with a range of integration concepts, was presented in Benedetto et al. (2012). In Hou et al. (2013) a pixel distribution flow was constructed, in which a new joint spatial-pixel characteristic distance was developed to improve classification. In Benedetto et al. (2012) an extension to Laplacian Eigenmaps, called Schrödinger Eigenmaps, was proposed as a means of semi-supervised fusion where expert information or spatial information can be utilized during construction of the feature space. This line of investigation was followed by extending the concept of Schroedinger operator through generalizing its potential (Cahill et al. 2014, 2015), or by modifying and replacing the Laplacian operator (Hu et al. 2015; Zhang et al. 2015).

In view of the above, it is within the Dimension Reduction process that we see opportunities, both at the graph level and operator level, to fuse the multiple data sources in a unique way, which shall improve upon the state of the art. At the graph level, we design a distance metric to combine, both, spatial and spectral information, as well as a method to weight the different scales appropriately. At the operator level we propose modifications to the standard Laplacian operator in order to incorporate spatial information in different ways.

Our methods, by design, do not rely on pre- and post-processing of the data, and we use only simple classifiers. By developing algorithms which seek to create a new unified representation of the data, our techniques can be easily combined with different classifiers, majority voting, output smoothing, cloud and vegetation masks; all of which can further improve the numerical outputs. By treating a range of different data sets, we exemplify that our methods can be applied to not just the remote sensing field but to entirely different scientific areas. For example early forms of our ideas were used in the study of gene expressions combined with chromosomal locations (Rajapakse et al. 2012).

The major contributions of the work presented in this paper include: (1) a new approach to fusion via fused representations of heterogeneous data; (2) a family of representation techniques which can be used in different scenarios, based on features of the analyzed data; and (3) an easy combination with state of the art pre- and post-processing techniques.

This paper is organized as follows: In Sect. 2 we will provide a brief summary of Laplacian Eigenmaps as well as the data sources and classification methods and metrics we shall use to justify our results. In Sect. 3 we will motivate the fusion paradigm by

examining separately Spectral Laplacian Eigenmaps and Spatial Laplacian Eigenmaps applied to our scenes. In Sect. 4 we develop our fusion framework, at the graph, operator, and feature space levels and provide numerical justification for their use. Section 5 contains the numerical results. Finally, Sect. 6 includes concluding remarks and future work.

2 Methodology

There are two main branches of dimensionality reduction: linear and non-linear. Linear methods, which include Principal Component Analysis (PCA) (Smith et al. 1985) and Multi-Dimensional Scaling (MDS) (Torgerson 1952), have been used widely in the past because of their obvious computational advantage over non-linear methods. However, making the assumption that hyperspectral data lies on a linear manifold is problematic (Prasad and Bruce 2008). There are many popular non-linear Dimension Reduction techniques, such as, Laplacian Eigenmaps (LE) (Belkin and Niyogi 2003), Schroedinger Eigenmaps (SE) (Czaja and Ehler 2013), Diffusion Maps (DM) (Coifman et al. 2005), Local Linear Embedding (LLE) (Roweis and Saul 2000), Local Tangent Space Alignment (LTSA) (Zhang and Zha 2002) and Hessian LLE (HLLE) (Donoho and Grimes 2003) [(all can be described as Output Normalized Algorithms (Goldberg et al. 2008)] which attempt to overcome this problem. These non-linear methods have been used to study hyperspectral data in the past, particularly (Kim and Finkel 2003; Coifman and Hirn 2014; Fauvel et al. 2009; Halevy 2011; Ma et al. 2010), thus we see them as a strong starting point for our spatial-spectral fusion work.

Our techniques are built upon a foundation of diffusion operators on data-dependent manifolds. The primary example of such operators is the Laplace-Beltrami operator. Given a point cloud of data which lies on an unknown manifold, the eigenfunctions of the Laplace-Beltrami operator applied to the data, describe the optimal smooth mappings which preserve the structure of the manifold in lower dimensions. To compute such embeddings, a discrete approximation (utilizing a graph representation of the data structure) of the Laplace-Beltrami operator must be derived. The non-linear Dimension Reduction algorithms mentioned above are such approximations. For this study, we will concentrate on the Laplacian Eigenmaps approach, but the principles described in the following pages can be easily extended to the other non-linear Dimension Reduction algorithms mentioned. For the sake of completeness, we provide a brief overview of the Laplacian Eigenmaps algorithm below. We shall conclude this section with a description of the data sets we will use to test our methods and with the details of the classification algorithm used to compute our results.

The ideas presented in this paper were motivated by a preliminary study in Benedetto et al. (2012). Among our major new contributions we consider strengthening the concept of the data representation via fusion. This is achieved by means of deemphasizing the importance of classification in the integration process. In this regard we use 1NN—one of the simplest classification techniques—to show that the high quality of classification results is a consequence of the quality of data representation, rather than the quality of the classifier itself. Furthermore, we develop the methodology for parameter selection based specifically on the input data, instead of relying entirely

on broad parameter swipes. Finally, our demonstrated classification accuracies are independent of numerical acceleration techniques, and are applied to a larger range of HSI data sets.

2.1 Laplacian Eigenmaps

Let $X = \{x_1, \ldots, x_N\} \subset \mathbb{R}^D$. We assume that the points of X lie on a d-dimensional manifold ($d \leq D$) that reflects the local geometry of the data set. The goal is to find a mapping

$$y : \mathbb{R}^D \to \mathbb{R}^n, \ d \leq n \leq D$$

that emphasizes the local geometric structure of X. The Laplacian Eigenmaps algorithm (Belkin and Niyogi 2003, 2008), pioneered by Belkin and Niyogi, is divided into three steps:

1. Construct a k-nearest neighbor (kNN) graph.
   An edge is drawn between x_i and x_j, if x_i is among the k nearest neighbors of x_j with respect to the Euclidean metric. This information is stored in an adjacency matrix G, by letting $G_{i,j} = 1$. We assume that our graph is undirected, which is equivalent to taking $G = \max\{G, G^T\}$.
2. Define a graph Laplacian.
   The edges of the graph are assigned weights and this information is stored in the weight matrix, W. Let D be the diagonal matrix given by $D_{i,i} = \sum_j W_{i,j}$. The graph Laplacian is given by

   $$L = D - W.$$

   A common choice of weights, motivated by the heat kernel, is the *diffusion weight matrix*:

   $$W_{i,j} = \begin{cases} e^{-\frac{\|x_i - x_j\|_2^2}{2\sigma^2}} & \text{if } G_{i,j} = 1 \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

   where σ is the heat kernel parameter which controls the spread of information in the graph. A large σ will essentially treat all neighbors of equal importance while a small σ will enforce the closeness relation between only the closest of neighbors.
3. Find the mapping $y = \{y_1, \ldots, y_N\}^T$, $y_i \in \mathbb{R}^n$ by solving

   $$\arg\min_{y^T D y = I} \frac{1}{2} \sum_{i,j} \|y_i - y_j\|_2^2 W_{i,j}. \tag{2}$$

   Note that (2) is equivalent to solving

   $$\arg\min_{y^T D y = I} \operatorname{trace}(y^T L y).$$

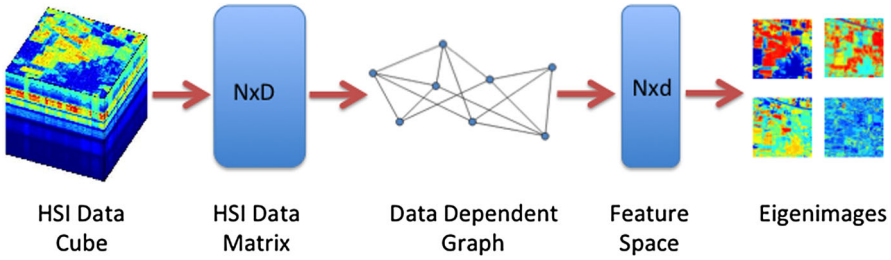


Fig. 1 Diagram depicting the Laplacian Eigenmaps algorithm stages. Given a $m \times n \times D$ hyperspectral data cube we reshape this to form a $N \times D$ data matrix. A data dependent graph is then learned using kNN. A cost function is minimized to produce a lower dimensional feature space representation of the data in the form of a $N \times d$ matrix. Each column of this matrix can be reshaped to form an eigenimage

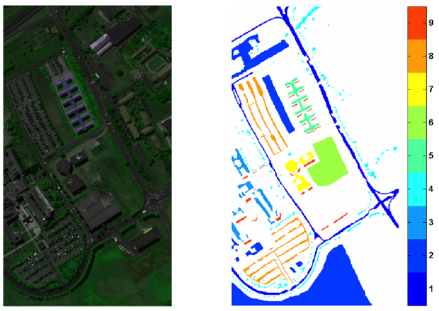
Letting $z = D^{1/2}y$ in (2), it follows that the minimization problem is equivalent to finding first n solutions to the generalized eigenproblem $Lv = \lambda Dv$, sorted in increasing order of λ.
It is clear that $v_0 = (1, 1, \ldots, 1)^T$ is a solution for $\lambda = 0$. If G is connected, this solution is unique for $\lambda = 0$. Hence, the problem can be refined further by assuming that $v_0 \in \ker(y^T)$, i.e. we only look for eigenvectors corresponding to nonzero eigenvalues.

A diagram of the Laplacian Eigenmaps algorithm applied to a hyperspectral data cube can be seen in Fig. 1.

2.2 Data

For our experiments, we take advantage of the hyperspectral data sets known as Pavia University, Pavia Centre, and Indian Pines. The Pavia University scene is a 610×340 pixel image that contains 103 spectral bands with approximately 1.3 meter resolution. The Pavia Centre scene is a 1096×492 pixel image (we have chosen to crop out the right side of the image where little ground truth is available) that contains 102 spectral bands with approximately 1.3 meter resolution. Both images were acquired in a flyover of Pavia, Italy using a ROSIS sensor. There is a collection of 42,776 ground truth pixels for the Pavia University scene and 103,539 ground truth pixels for the Pavia Centre scene, covering a total of 9 classes for each data set. Both the ground truth sets and the images, with bad spectral bands removed, are readily available on the web. We chose to work with these data sets primarily because of the spatial diversity found in its ground truth. Many data sets have ground truth whose classes are clearly separated spatially. In Pavia University and Pavia Centre, however, the classes are more interspersed, and we felt that this would provide a sufficient challenge to our hypothesis that including spatial information improves classification. Figs. 2 and 3 contain a 3 color composite image of Pavia University and Pavia Centre, respectively, as well as ground truth masks and tables.

The Indian Pines scene is a 145×145 pixel image that contains 224 spectral bands with approximately 20 meter resolution; we discard 4 of the bands due to noise and water. It was acquired in a flyover of northwest Indiana using an AVIRIS spectrometer.

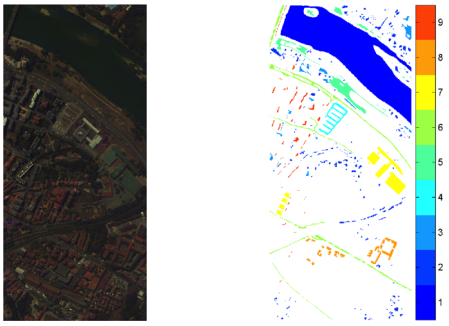


(a) (b)

#	Class	Samples
1	Asphalt	6631
2	Meadows	18649
3	Gravel	2099
4	Trees	3064
5	Painted metal sheets	1345
6	Bare Soil	5029
7	Bitumen	1330
8	Self-Blocking Bricks	3682
9	Shadows	947

(c)

Fig. 2 **a** Three color composite image of Pavia University. **b** Ground truth for Pavia University. **c** Ground truth descriptions and sample count for Pavia University



(a) (b)

#	Class	Samples
1	Water	65278
2	Trees	6508
3	Asphalt	2905
4	Self-Blocking Bricks	2140
5	Bitumen	6549
6	Tiles	7585
7	Shadows	7287
8	Meadows	3122
9	Bare Soil	2165

(c)

Fig. 3 **a** Three color composite image of Pavia Centre. **b** Ground truth for Pavia Centre. **c** Ground truth descriptions and sample count for Pavia Centre

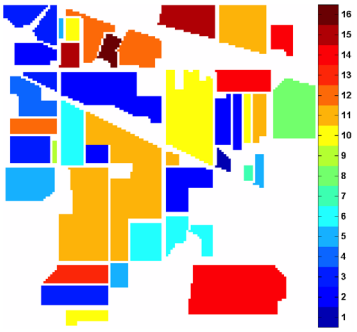
There is a collection of 10, 249 ground truth pixels associated with the data set coving a total of 16 classes. We choose to use this image, seen in Fig. 4 rendered as a 3-color composite with ground truth mask and table, because of its wide use in the community.

2.3 Classification and validation metrics

All of the results presented in the next section were computed using $k = 20$ nearest neighbors and $d = 25$ embedding dimensions for Pavia University and Pavia Centre, and $k = 20$ and $d = 50$ for Indian Pines. The intrinsic dimensions were chosen to reflect what was seen in previous literature and to approximate the material complexity of the scenes. The number of nearest neighbors was chosen so as to result in a connected graph that was still sparse as to aid in the minimization phase.

With embeddings in hand, we use a vector angle 1 nearest neighbor classifier with 1 % for Pavia University and Pavia Centre and 10 % for Indian Pines training data from each class, the remaining number of pixels, N_v, comprise the validation set. We took an average of ten runs to produce the confusion matrices (C) and classification

(a)



(b)

#	Class	Samples
1	Alfalfa	46
2	Corn-notill	1428
3	Corn-mintill	830
4	Corn	237
5	Grass-pasture	483
6	Grass-trees	730
7	Grass-pasture-mowed	28
8	Hay-windrowed	478
9	Oats	20
10	Soybean-notill	972
11	Soybean-mintill	2455
12	Soybean-clean	593
13	Wheat	205
14	Woods	1265
15	Buildings-Grass-Trees-Drives	386
16	Stone-Steel-Towers	93

(c)

Fig. 4 **a** The color composite image of Indian Pines. **b** Ground truth for Indian Pines. **c** Ground truth descriptions and sample count for Indian Pines

Table 1 SVM classification results with parameters

Image name	C	γ	OA	AA
Pavia University	1024	0.5	90.6500	88.1568
Pavia Centre	1024	0.125	93.4879	82.3235
Indian Pines	512	0.0625	82.6021	82.5542

accuracies presented in the following pages. In the following tables Overall Accuracy (OA) is the ratio of correctly labeled validation pixels to the total number of validation pixels, Average Accuracy (AA) is a weighted accuracy measurement where weights are added so each class has the same influence, and Cohen's Kappa, κ, is defined by $\kappa = \frac{N_v \sum_i (C_{i,i})^2 - \Omega}{N_v^2 - \Omega}$, where $\Omega = \sum_i C_{i,} . C_{.,i}$. The class images and eigenimages come from a single run. We note that we could potentially improve the classification accuracy by applying more complicated classification schemes such as Support Vector Machines (SVM) (Cortes and Vapnik 1995) or Artificial Neural Networks (Bishop 1995); or clustering and sieving the results.

For a comparison to the state of the art of classification we used the LIBSVM (Chang et al. 2001) multiclass SVM classifier with a five-fold cross validation method to optimize the parameters C and γ; results of which can be found in Table 1.

3 Spectral Laplacian Eigenmaps, spatial Laplacian Eigenmaps, and their integration

As a baseline for our analysis, we first look at Spectral Laplacian Eigenmaps and Spatial Laplacian Eigenmaps. We define Spectral Laplacian Eigenmaps to be the standard Laplacian Eigenmaps algorithm where the L_2 distance between two pixels spectra is used to determine the nearest neighbors and the weights of the kernel matrix. Spatial Laplacian Eigenmaps also uses the standard Laplacian Eigenmaps algorithm but uses the L_2 distance between pixel locations (where the origin is selected to be the upper left corner of the image) to select nearest neighbors as well as the weights in the kernel matrix. Throughout the rest of this paper, we choose $\sigma = 0.8$, $k = 20$, and $d = 25, 25, 50$ for Pavia University, Pavia Centre, and Indian Pines respectively. Parameter σ was selected to maximize the classification score over a set, k was selected to so as to provide a fully connected graph but be minimal over a set, and d was chosen based on the material complexity in the scene and from past literature. The selection of these parameters is motivated by the results of our prior work (Benedetto et al. 2012), where we studied the problem of appropriate parameter selection for joint spatial-spectral representations.

3.1 Spectral Laplacian Eigenmaps

In Figs. 5, 6, and 7 we can see several of the eigenimages that are produced. Eigenmaps correspond to the eigenvectors produced by the Laplacian Eigenmaps algorithm, they are reshaped to form a matrix corresponding to the original dimensions of the hyperspectral Image and rescaled linearly to produce a desired colormap. As with previous research, the eigenimages that are presented here work well as a feature extraction technique. These eigenmaps can be interpreted as highlighting similar features one-by-one in the various eigenimages. In Fig. 8 and Table 2 we can see the classification results for Spectral Laplacian Eigenmaps.

3.2 Spatial Laplacian Eigenmaps

In order to motivate how we can improve the results from the previous section, we present the result of Laplacian Eigenmaps applied to X_s, the $N \times 2$ array of purely

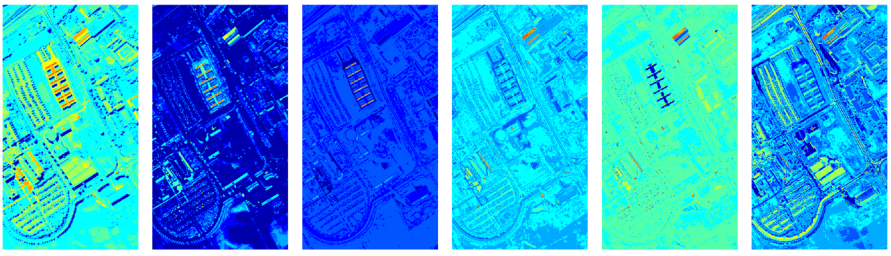


Fig. 5 Several eigenimages from Pavia University scene

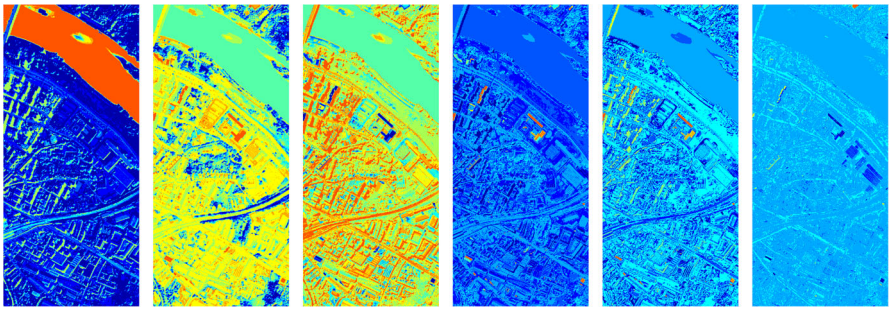


Fig. 6 Several eigenimages from Pavia Centre scene

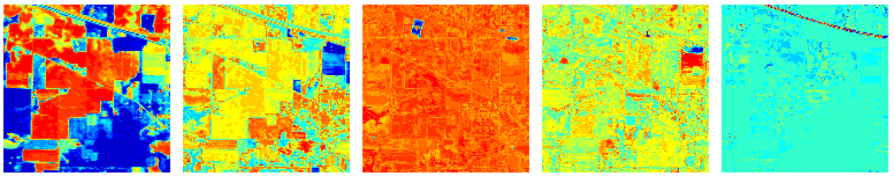


Fig. 7 Several eigenimages from Indian Pines scene

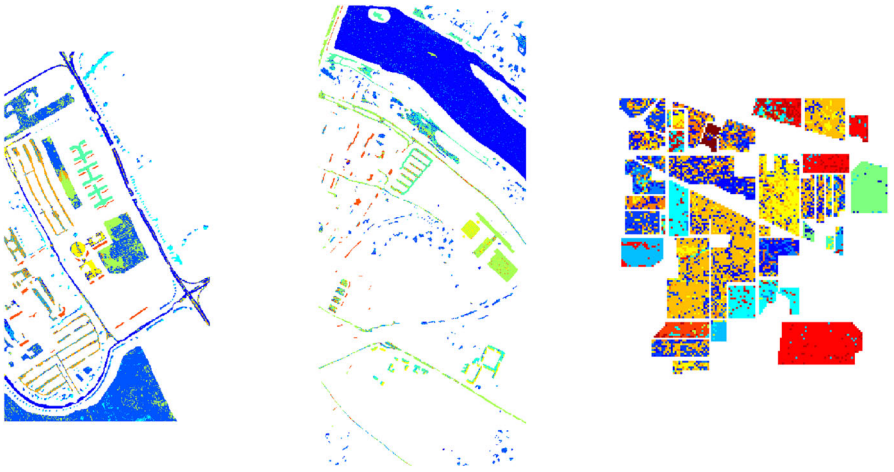
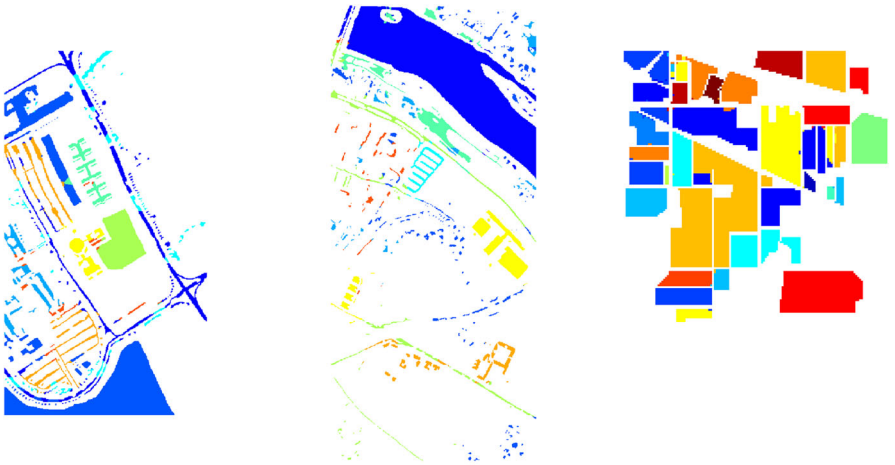


Fig. 8 Class map for **a** Pavia University scene, **b** Pavia Centre scene, and **c** Indian Pines scene, using Spectral Laplacian Eigenmaps

spatial information. That is, $X_s(i)$, the ith row of the array, contains the spatial coordinates for the ith pixel. The Spatial Laplacian Eigenmaps results provide a means of measuring to what degree the ground truth classes are spatially interspersed. In Fig. 9, as well as in Table 2, we can see the classification results for Spatial Laplacian Eigenmaps. We note here that these results in most cases significantly outperform the Spectral Laplacian Eigenmaps classifications.

Table 2 Spatial and spectral Laplacian Eigenmaps

	Spectral	Spatial
Pavia University		
OA	0.7326	0.8994
AA	0.7215	0.8207
κ	0.6440	0.8593
Pavia Centre		
OA	0.9568	0.9497
AA	0.8807	0.8728
κ	0.9258	0.9129
Indian Pipes		
OA	0.6041	0.9795
AA	0.5552	0.9583
κ	0.5491	0.9767



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

Fig. 9 Class map for **a** Pavia University scene, **b** Pavia Centre scene and **c** Indian Pines scene, using Spatial Laplacian Eigenmaps

Laplacian Eigenmaps analysis of spatial information is not commonly used, due to the artificial nature of the resulting eigenimages. However, we note here that it has been already established in literature (Gillis and Bowles 2012) that the spatial structure is crucial for many applications, including clustering, classification, and segmentation of high dimensional data. Table 2 provides further evidence to that effect.

3.3 Comments

As we see neither approach is perfect. The Spectral Laplacian Eigenmaps is able to make primarily global connections between pixels, whereas the Spatial Laplacian

Eigenmaps algorithm is simply creating local clusters around the ground truth. Even though the resulting classification results for the spatial approach are impressive, their class maps are practically meaningless. A fusion of these two ideas, which we will discuss in the following section, is required to improve upon these individual results.

4 Spatial-spectral fusion via Laplacian Eigenmaps

Our goal is to find meaningful ways of including spatial information, which is acquired with no additional cost, with the spectral data in Laplacian Eigenmaps framework. To accomplish this, we studied the effect of introducing spatial information in the first two steps of the Laplacian Eigenmaps algorithm, separately and simultaneously (Benedetto et al. 2012; Benedetto et al. 2012; Duke 2012; Doster 2014).

4.1 Feature space fusion for benchmark comparison

Laplacian Eigenmaps feature space provides us with a first natural environment for data fusion. In this setting we can combine eigenvectors obtained by our two methods, Spectral and Spatial Laplacian Eigenmaps, by stacking the resulting eigenvectors and then applying a classifier. In Table 3 we present the classification results obtained by combining different ratios of the Spatial Laplacian Eigenmaps eigenvectors with the Spectral Laplacian Eigenmaps eigenvectors. To provide a better comparison with our other methods, we force the total the number of stacked features to be consistent with the chosen intrinsic dimensions (25 for the Pavia scenes and 50 for the Indian Pines scene). The columns indicate the ratio of spatial eigenvectors to spectral eigenvectors chosen for this experiment.

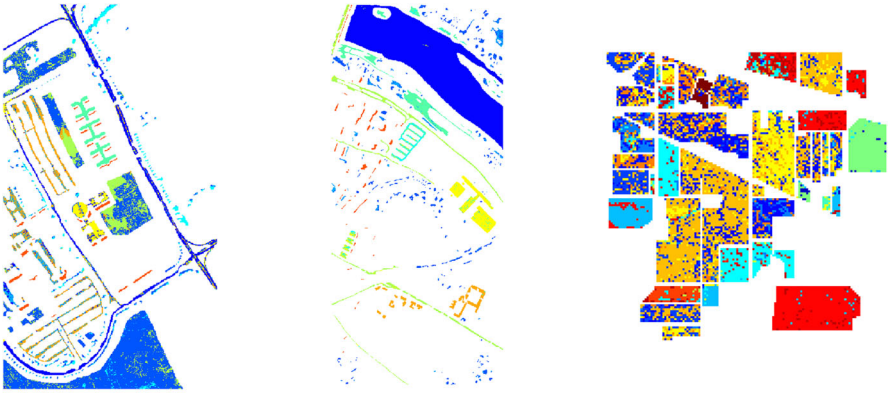
The outcome of this analysis is very promising, as the results in Table 3 show that after appropriate Dimension Reduction, choosing a balanced selection of spatial and spectral eigenvectors allows to significantly increase the classification rates as compared to purely spectral approach. The reason for this improvement and for the fact that such an arbitrage works is the difference between spatial and spectral confusion maps, which can be exploited via data integration. As promising as this method is, however, it has certain drawbacks. The optimal ratios range from 0.4 to 0.92, across the analyzed data, with differences that cannot be easily explained by the data structure alone. Moreover, some of the class maps show strong preference towards spatial classification. Therefore, notwithstanding the superior classification rates and the novelty of the approach which combines the spatial and spectral Laplacian features of HSI, in the reminder of this paper we introduce a novel methodology based on data-dependent representations.

4.2 Mixing spatial/spectral graphs and weights

Motivated by the observations of the previous section, we seek to explore a new avenue for data fusion by combining spatial and spectral information in the graph construction phase of the Laplacian Eigenmaps algorithm. Let G_f be the kNN-graph computed

Table 3 Feature space fusion overall classification percentage with varying percentage of principle spatial feature vectors included

	[$100p\%$] Primary spatial and [$100(1-p)\%$] primary spectral vectors							
	0	0.08	0.16	0.4	0.6	0.84	0.92	1.00
Pavia University								
OA	0.7326	0.8991	0.9224	0.9599	0.9748	0.9822	0.9644	0.8944
AA	0.7215	0.8739	0.9108	0.9506	0.9684	0.9763	0.9509	0.8207
κ	0.6440	0.8652	0.8970	0.9465	0.9664	0.9764	0.9525	0.8593
Pavia Centre								
OA	0.9568	0.9706	0.9758	0.9817	0.9815	0.9787	0.9679	0.9497
AA	0.8807	0.9186	0.9339	0.9521	0.9562	0.9495	0.9160	0.8728
κ	0.9258	0.9495	0.9584	0.9685	0.9682	0.9633	0.9446	0.9129
Indian Pines								
OA	0.6041	0.7878	0.8477	0.9301	0.9646	0.9817	0.9844	0.9795
AA	0.5552	0.7161	0.7927	0.9235	0.9584	0.9792	0.9788	0.9583
κ	0.5491	0.7577	0.8261	0.9204	0.9596	0.9791	0.9822	0.9767



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

Fig. 10 Class map for **a** Pavia University scene, **b** Pavia Centre and **c** Indian Pines Scene using spectral neighborhood construction with spatial kernel

using the spectral metric; W_f is the weight matrix constructed for a given kNN-graph using weights determined by (1). Let G_s be the kNN-graph constructed using the *spatial* metric:

$$\|x_i - x_j\|_s = \|s(i) - s(j)\|_2;$$

W_s is the weight matrix obtained by replacing the spectral weights with their spatial analogue. In this spatial fusion attempt a global structure is first defined by the initial construction of G_f using spectral information. Next, by replacing the weights in the kernel matrix with spatial distances, a priority in the embedding will be given to preserving distances between spatially close pixels. However, since the connections in G_f were defined spectrally, materials with similar spectra shall still be linked in the embedding. We refer to Fig. 10 and Table 5 for the outcomes of this approach.

4.3 Fusion operators

Having already explored the problem of data fusion in the graph construction phase of the Laplacian Eigenmaps, we now move to looking at how to properly perform the data fusion in the graph Laplacian stage of the Laplacian Eigenmaps algorithm. In this section, we fix $G = G_f$. We choose to fix the construction of G based on spectral information, in order to preserve a global perspective to the embedding; otherwise a very local and patchy embedding results. Spatial information is introduced by modifying the operator L in three different ways. Let L_{ff} be the operator using spectral information for the graph construction and graph weights, let L_{fs} be the operator using spectral information for the graph and construction and spatial information for the graph weights. We have the following:

- $L_1(\sigma, \eta) = L_{ff}(\sigma) \cdot L_{fs}(\eta)$, where $\cdot$ denotes entry-wise multiplication,
- $L_2(\sigma, \eta) = L_{ff}(\sigma) + L_{fs}(\eta)$,

– $L_3(\sigma, \eta) = G \cdot (L_{ff}(\sigma) \times L_{fs}(\eta))$, i.e., the usual matrix product of the two Laplacians, overlaid with the sparsity structure determined by the adjacency matrix.

The parameters σ and η are the heat kernel parameters from (1) for the spectral and spatial constructions, respectively. In this fusion operator approach to adding spatial information, three separate methodologies are considered based on how the fused diffusion weight matrix is defined. $L_1(\sigma, \eta)$ is perhaps the simplest and is akin to defining a new distance metric which weights spectral and spatial information differently for the diffusion weights matrix. This new metric promotes sparsity of the resulting new fused graphs, by disconnecting the vertices which are not connected by an edge in both spatial and spectral representations. This type of representation is useful in images with spatially distributed classes of objects with high spectral correlations (e.g., Pavia Centre). $L_2(\sigma, \eta)$, by adding the two diffusion weight matrices separately allows for pixels that are, both, close spectrally and spatially to dramatically dominate the embedding while allowing pixels which are only close in one regard to have approximately equal importance in the embedding. The role of L_2 is similar to that of L_1, but without completely disconnecting the nodes, which are far apart in one of the representations. $L_3(\sigma, \eta)$ seeks to take the fusion deeper by considering weights derived from common neighborhoods amongst pixels. In addition, by further requiring that the original sparsity structure carried by $G = G_f$ be preserved, L_3 keeps the computational complexity of the problem at hand on par with other fusion operator approaches. This technique shows its strength for datasets with highly spatially correlated classes, such as, e.g., Indian Pines.

Across all three images we found that a value 0.2 for η produced the best results, σ was set to 0.8. These parameters where chosen based on the results of our prior work (Benedetto et al. 2012). We refer to Figs. 11, 12 and 13 and Table 6 for results using this methodology.

We emphasize here that all three representations, L_1, L_2, and L_3, play very similar roles from the perspective of Representation Theory, with differences addressing certain specific data structures. Further in this section we shall show that a much more complete picture will be obtained when merging operator fusion with graph integration. This will not only improve the results, but also decrease variability of the outputs.

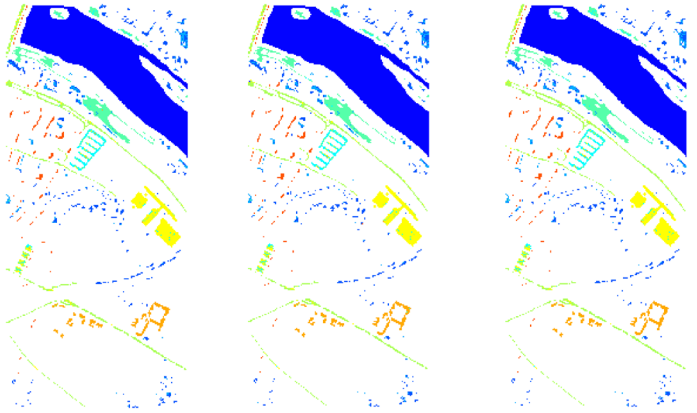
4.4 Fusion metrics

We close this section with the final approach to integrating spatial and spectral information. We propose to adjust the norm used to determine the kNN-graph and the weight matrix, basing it on distances computed in both modalities. In this manner we are able to claim that by carefully adjusting the norms used to construct the graph we can inject important spatial information into the optimization phase and thus favor solutions (eigenimages) that are spatially and spectral motivated. For $\gamma > 0$, we define the fusion metric as:

$$\|x_i - x_j\|_\gamma = \sqrt{\|x_i - x_j\|_f^2 + \gamma \|x_i - x_j\|_s^2}.$$

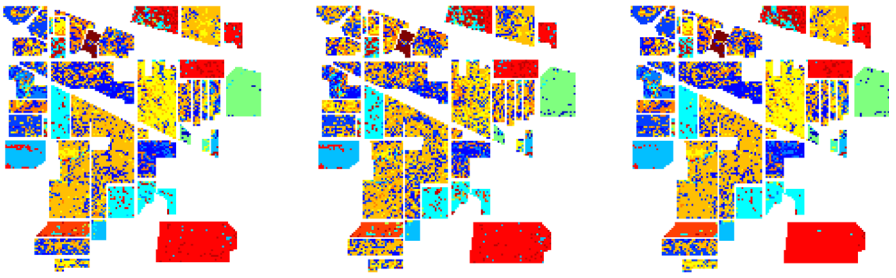
(a) L_1 **(b)** L_2 **(c)** L_3

Fig. 11 Pavia University class maps using spectral neighborhood construction and fusion operators kernel matrix construction



(a) L_1 **(b)** L_2 **(c)** L_3

Fig. 12 Pavia Centre class maps using spectral neighborhood construction and fusion operators kernel matrix construction



(a) L_1 **(b)** L_2 **(c)** L_3

Fig. 13 Indian Pines class maps using spectral neighborhood construction and fusion operators kernel matrix construction

Table 4 Parameter γ values found using (3) for Pavia University, Pavia Centre, and Indian Pines

	γ
Pavia University	44.2011
Pavia Centre	43.0527
Indian Pines	894.8153

The parameter γ allows the user to control how much spatial information should be included in the analysis. Our goal was to choose γ so that the contribution of spatial information to $\|\cdot\|_\gamma$ reflected the relative importance of spatial information to spectral information inherent in the image. This would guarantee that two pixels in the image would be considered close with respect to $\|\cdot\|_\gamma$ if and only if they are close spectrally and "close enough" spatially. Motivated by this idea, let γ be determined as follows.

Let N_i be the set of indices for the spectral k-nearest neighbors of the pixel x_i. Define γ_i to be the ratio of the spectral and spatial spread of this neighborhood:

$$\gamma_i = \frac{\sum_{j\in N_i} \|x_i - x_j\|_f^2}{\sum_{j\in N_i} \|x_i - x_j\|_s^2}.$$

Define γ to be the global average of these local weights:

$$\gamma = \frac{1}{N}\sum_{i=1}^{N} \gamma_i. \tag{3}$$

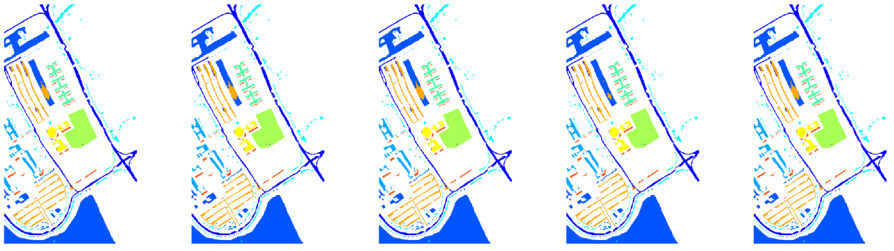
In Table 4 we show the γ-values we calculated for the images of study, we note that these calculated values match closely with the values found using a grid search optimization method in Benedetto et al. (2012).

This fusion metric graph construction can also be combined with the fusion operators to form fusion metric-fusion operator based spectral-spatial fusion; we denote this as $\Gamma + L_1$, $\Gamma + L_2$, and $\Gamma + L_3$ where Γ is the graph constructed using the fusion metric to determine neighborhoods and weights. If we use the fusion metric derived weights with the spectral kNN graph construction we denote this fusion as $G_\gamma + L_f$ and if we use the fusion metric to construct both the kNN and weight the graph we denote this as L_γ. See Figs. 14, 15, and 16, as well as Table 7 for the results of the analysis using this fusion metric.

We note here that similar attempts to inject spatial information into Laplacian Eigenmaps outputs have been shown in the literature, see, e.g., Hou et al. (2013). It is evident, however, from our work and from comparison of results of Hou et al. (2013) to the results we present in this paper, that the modification of the graph construction alone is insufficient to obtain the best results in spatial-spectral fusion.

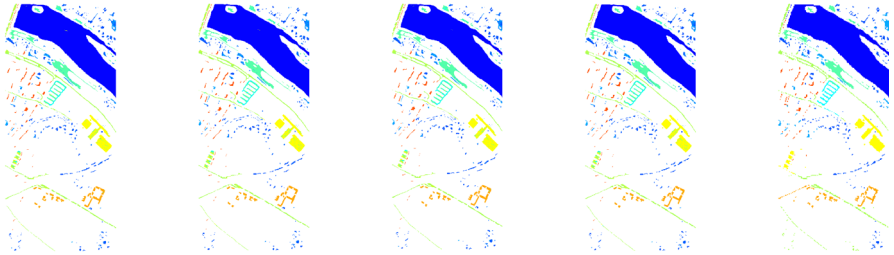
5 Results

We start this section by recalling from the previous section that we consider fusion at two levels: the graph level and the operator level. At the graph level we developed a



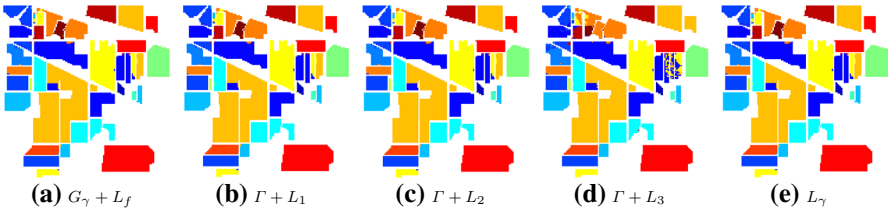
(a) $G_\gamma + L_f$ **(b)** $\Gamma + L_1$ **(c)** $\Gamma + L_2$ **(d)** $\Gamma + L_3$ **(e)** L_γ

Fig. 14 Pavia University class maps graph-operator level joint fusion using fusion metrics



(a) $G_\gamma + L_f$ **(b)** $\Gamma + L_1$ **(c)** $\Gamma + L_2$ **(d)** $\Gamma + L_3$ **(e)** L_γ

Fig. 15 Pavia Centre class maps using graph-operator level joint fusion using fusion metrics



(a) $G_\gamma + L_f$ **(b)** $\Gamma + L_1$ **(c)** $\Gamma + L_2$ **(d)** $\Gamma + L_3$ **(e)** L_γ

Fig. 16 Indian Pines class maps using graph-operator level joint fusion using fusion metrics

mixed spatial and spectral weights method to create a fused graph representation of the data. At the operator level we introduced three fused operators, L_1, L_2, and L_3. We also developed the fusion metric $dist_\gamma(x_i, x_j) = \|x_i - x_j\|_\gamma$, which was used to choose neighborhoods which spatially and spectrally balanced chosen neighboring nodes. Finally, we presented a mixture of the graph and operator based fusion approaches with optimized weights based on fusion metrics. This led us to consider a range of fusion techniques, denoted by $G_\gamma + L_f$, L_γ, and $\Gamma + L_t$, where $t = 1, 2, 3$.

Tables 5, 6, and 7 contain the numerical results for our methods. In Figs. 10, 11, 12, 13, 14, 15 and 16 we show the corresponding class maps. Table 7 contains our main results, the mixture of graph and operator fusion methods. Tables 5 and 6 serve as motivation for our final results.

To capture the essence of the results presented in this Section we refer again to Table 7, which provides the overall and average classification rates for the three examples of hyperspectral imagery we have chosen to analyze. Comparing to the most recent

Table 5 Mixing spectral neighborhoods with spatial weights

	Mixed
Pavia University	
OA	0.7334
AA	0.7224
κ	0.6450
Pavia Centre	
OA	0.9558
AA	0.8764
κ	0.9240
Indian Pines	
OA	0.6049
AA	0.5594
κ	0.5502

Table 6 Fusion operators

	L_1	L_2	L_3
Pavia University			
OA	0.7392	0.7390	0.7392
AA	0.7253	0.7228	0.7259
κ	0.6482	0.6525	0.6528
Pavia Centre			
OA	0.9584	0.9566	0.9580
AA	0.8840	0.8788	0.8820
κ	0.9285	0.9524	0.9279
Indian Pines			
OA	0.6130	0.5262	0.6473
AA	0.5656	0.4723	0.6182
κ	0.5593	0.4587	0.5983

results in the field of spatial-spectral integration for hyperspectral imagery, we clearly show that our results are in line with those presented recently in literature, or better, see e.g., in Fauvel et al. (2013). Moreover, we showed that across the range of various methods presented in this section, the standard deviation for classification results is negligible, and as such, all methods perform equally well. This is due to the fact that varying the graph constructions with a fixed method for computing the Laplacian, or varying the construction of the Laplacian while retaining a fixed graph, leads to similar outcomes. At the same time, we emphasize that we have introduced a completely novel approach to spatial-spectral fusion, which possibly can be combined with the state of the art classification techniques such as those in Tilton et al. (2012), Bernard et al. (2012) and Tarabalka et al. (2012) to provide additional improvements.

Table 7 Graph-operator level joint fusion using fusion metrics

	$G_\gamma + L_f$	$\Gamma + L_1$	$\Gamma + L_2$	$\Gamma + L_3$	L_γ
Pavia University					
OA	0.9715	0.9699	0.9703	**0.9815**	0.9714
AA	0.9747	0.9739	0.9739	**0.9781**	0.9746
κ	0.9623	0.9603	0.9607	**0.9754**	0.9622
Pavia Centre					
OA	0.9557	0.9581	0.9568	0.9577	**0.9728**
AA	0.8762	0.8826	0.8794	0.8811	**0.9291**
κ	0.9239	0.9280	0.9258	0.9273	**0.9532**
Indian Pines					
OA	0.9881	**0.9887**	0.9882	0.8795	0.9881
AA	0.9852	**0.9856**	0.9853	0.7889	0.9852
κ	0.9864	**0.9872**	0.9865	0.8625	0.9864

Bold represents the top method for each given image and metric pair

6 Conclusions

In this paper we have analyzed data integration and fusion techniques for taking advantage of, both, spatial and spectral information in hyperspectral imagery. These techniques are based on outcomes of Laplacian-type operators on data dependent graphs. As expected, this approach yields significant improvements when compared to classification based entirely on either Spectral or Spatial Laplacian Eigenmaps. As is evident from the outcome, the best results are observed when we utilize spatial and spectral information at all stages: in constructing the data adjacency matrix, and in building the new joint spatial-spectral Laplacian on such a modified graph.

We note here that numerical values for classification results obtained through selecting a mix of eigenvectors from purely spatial and purely spectral Laplacian Eigenmaps, are not far away from our fused classification results. At the same time there are qualitative differences between the class maps obtained through these two methods, with the advantage clearly on the side of the graph-based fusion.

We would like to stress that classification is just a simple example of an application of the operator fusion we have introduced. Unlike other methods that concentrate on improving classification methodologies, we make use of the simplest classifier possible, 1NN, and put our effort into creating more universal technique based on the quality of our data representation. Aside from classification we are actively developing our techniques for target detection and feature extraction.

We are already in the process of exploring further fusion possibilities, particularly LIDAR and HSI fusion, cf., Cloninger et al. (2013), utilizing our graph construction operator methodologies.

Acknowledgments The authors would like to thank Professor Landgrebe (Purdue University, USA) for providing the Indian Pines data and Professor Paolo Gamba (Pavia University, Italy) for providing the Pavia University and Pavia Centre data. The work presented in this paper was supported in part by NSF

through grant CBET 0854223, by DTRA through Grant HDTRA 1-13-1-0015, and by ARO through Grant W911NF1610008.

References

Belkin, M., Niyogi, P.: Laplacian eigenmaps for dimensionality reduction and data representation. Neural Comput. **1396**, 1373–1396 (2003)

Belkin, M., Niyogi, P.: Convergence of Laplacian eigenmaps. Statistics **19**, 1–31 (2008)

Benedetto, J.J., Czaja, W., Dobrosotskaya, J., Doster, T., Duke, K., Gillis, D.: Integration of heterogeneous data for classification in hyperspectral satellite imagery. In: Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVIII, Proc. SPIE, vol. 839078, pp. 1–12. International Society for Optics and Photonics (2012)

Benedetto, J.J., Czaja, W., Dobrosotskaya, J., Doster, T., Duke, K., Gillis, D.: Semi-supervised learning of heterogeneous data in remote sensing imagery. In: Independent Component Analyses, Compressive Sampling, Wavelets, Neural Net, Biosystems, and Nanoengineering X, Proc. SPIE, vol. 840104, pp. 1–12. International Society for Optics and Photonics (2012)

Benedetto, J.J., Czaja, W., Ehler, M.: Wavelet packets for time frequency analysis of multispectral data. Int. J. Geomath. **4**(2), 137–154 (2013)

Benedetto, J.J., Czaja, W., Ehler, M., Flake, C., Hirn, M.: Wavelet packets for multi- and hyper-spectral imagery. Proc. SPIE **753508**, 1–11 (2010)

Bernard, K., Tarabalka, Y., Angulo, J., Chanussot, J., Benediktsson, J.: Spectral-spatial classification of hyperspectral data based on a stochastic minimum spanning forest approach. IEEE Trans. Image Process. **21**(4), 2008–2021 (2012)

Bishop, C.: Neural Networks for Pattern Recognition. Clarendon Press, Oxford (1995)

Bruzzone, L., Carlin, L.: A multilevel context-based system for classification of very high spatial resolution images. IEEE Trans. Geosci. Remote Sens. **44**(9), 2587–2600 (2006)

Cahill, N.D., Chew, S.E., Wenger, P.S.: Spatial-spectral dimensionality reduction of hyperspectral imagery with partial knowledge of class labels. In: Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXI, Proc. SPIE. International Society for Optics and Photonics (2015)

Cahill, N.D., Messinger, D.W., Czaja, W.: Schroedinger eigenmaps with nondiagonal potentials for spatial-spectral clustering of hyperspectral imagery. In: Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XX, Proc. SPIE, pp. 1–13. International Society for Optics and Photonics (2014)

Camps-Valls, G., Bandos Marsheva, T.V., Zhou, D.: Semi-supervised graph-based hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. **45**(10), 3044–3054 (2007)

Camps-Valls, G., Gómez-Chova, L., Muñoz-Marí, J., Rojo-Álvarez, J.L., Martínez-Ramón, M.: Kernel-based framework for multitemporal and multisource remote sensing data classification and change detection. IEEE Trans. Geosci. Remote Sens. **46**(6), 1822–1835 (2008)

Camps-Valls, G., Gomez-Chova, L., Munoz-Mari, J., Vila-Frances, J., Calpe-Maravilla, J.: Composite kernels for hyperspectral image classification. IEEE Geosci. Remote Sensing Lett. **3**(1), 93–97 (2006)

Chang, C.C., Lin, C.J.: Libsvm: a library for support vector machines. http://www.csie.ntu.edu.tw/~cjlin/libsvm/ (2001). Accessed June 2013

Cloninger, A., Czaja, W., Doster, T.: Operator analysis and diffusion based embeddings for heterogeneous data fusion. In: 3rd Place 2013 IEEE GRSS Data Fusion Contest: Fusion of Hyperspectral and LiDAR Data (2013)

Coifman, R., Lafon, S., Lee, A., Maggioni, M., Nadler, B., Warner, F., Zucker, S.: Geometric diffusions as a tool for harmonic analysis and structure definition of data: diffusion maps. Proc. Natl. Acad. Sci. USA **102**(21), 7426–7431 (2005)

Coifman, R.R., Hirn, M.J.: Diffusion maps for changing data. Appl. Comput. Harmonic Anal. **36**(1), 79–107 (2014)

Cortes, C., Vapnik, V.: Support-vector networks. Mach. Learn. **20**(3), 273–297 (1995)

Czaja, W., Ehler, M.: Schröedinger eigenmaps for the analysis of bio-medical data. IEEE Trans. Pattern Anal. Mach. Intell. **35**(5), 1274–1280 (2013)

Donoho, D., Grimes, C.: Hessian eigenmaps: locally linear embedding techniques for high-dimensional data. Proc. Natl. Acad. Sci. **100**(10), 5591–5596 (2003)

Doster, T.: Harmonic analysis inspired data fusion for applications in remote sensing. Ph.D. thesis, University of Maryland, College Park (2014)

Duke, K.: A study of the relationship between spectrum and geometry through fourier frames and Laplacian eigenmaps. Ph.D. thesis, University of Maryland, College Park (2012)

Fauvel, M., Benediktsson, J., Chanussot, J., Sveinsson, J.: Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles. IEEE Trans. Geosci. Remote Sens. **46**(11), 3804–3814 (2008)

Fauvel, M., Chanussot, J., Benediktsson, J.: Kernel principal component analysis for the classification of hyperspectral remote sensing data over urban areas. EURASIP J. Adv. Signal Process **2009**, 1–14 (2009)

Fauvel, M., Tarabalka, Y., Benediktsson, J., Chanussot, J., Tilton, J.: Advances in spectral-spatial classification of hyperspectral images. Proc. IEEE **101**(3), 652–675 (2013)

Gillis, D., Bowles, J.: Hyperspectral image segmentation using spatial-spectral graphs. In: SPIE Defense, Security, and Sensing, vol. 83901Q, pp. 1–11. International Society for Optics and Photonics (2012)

Goldberg, Y., Zakai, A., Kushnir, D., Ritov, Y.: Manifold learning: the price of normalization. J. Mach. Learn. Res. **9**, 1909–1939 (2008)

Halevy, A.: Extensions of Laplacian eigenmaps for manifold learning. Ph.D Thesis, University of Maryland, College Park (2011)

Hou, B., Zhang, X., Ye, Q., Zheng, Y.: A novel method for hyperspectral image classification based on Laplacian eigenmap pixels distribution-flow. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. **6**(3), 1602–1618 (2013)

Hu, Y., Saber, E., Monteiro, S.T., Cahill, N.D., Messinger, D.W.: Classification of hyperspectral images based on conditional random fields. In: Image Processing: Machine Vision Applications VIII, Proc. SPIE. International Society for Optics and Photonics (2015)

Kim, D., Finkel, L.: Hyperspectral image processing using locally linear embedding. In: Neural Engineering, 2003. Conference Proceedings. First International IEEE EMBS Conference on, pp. 316–319 (2003)

Lee, M., Bruce, L., Prasad, S.: Concurrent spatial-spectral band grouping: Providing a spatial context for spectral dimensionality reduction. In: 3rd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), pp. 1–4 (2011)

Ma, L., Crawford, M., Tian, J.: Generalised supervised local tangent space alignment for hyperspectral image classification. Electron. Lett. **46**(7), 497–498 (2010)

Manolakis, D., Marden, D., Shaw, G.: Hyperspectral image processing for automatic target detection applications. Lincoln Lab. J. **14**(1), 79–116 (2003)

Mohan, A., Sapiro, G., Bosch, E.: Spatially coherent nonlinear dimensionality reduction and segmentation of hyperspectral images. IEEE Geosci. Remote Sens. Lett. **4**, 206–210 (2007)

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

Prasad, S., Bruce, L.: Limitations of principal components analysis for hyperspectral target recognition. IEEE Geosci. Remote Sens. Lett. **5**, 625–629 (2008)

Rajapakse, V., Czaja, W., Pommier, Y., Reinhold, W., Varma, S.: Predicting expression-related features of chromosomal domain organization with network-structured analysis of gene expression and chromosomal location. In: Proceedings of the ACM Conference on Bioinformatics, Computational Biology and Biomedicine, pp. 226–233 (2012)

Roweis, S., Saul, L.: Nonlinear dimensionality reduction by locally linear embedding. Science **290**(5500), 2323–2326 (2000)

Schaum, A., Stocker, A.: Hyperspectral change detection and supervised matched filtering based on covariance equalization. In: Proceedings of SPIE, vol. 5425, pp. 77–90 (2004)

Schott, J.: Remote Sensing: The Image Chain Approach. Oxford University Press, New York (1997)

Smith, M., Johnson, P., Adams, J.: Quantitative determination of mineral types and abundances from reflectance spectra using principal components analysis. In: Lunar and Planetary ScienceConference Proceedings, vol. 15, pp. 797–804 (1985)

Tarabalka, Y., Benediktsson, J., Chanussot, J.: Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques. IEEE Trans. Geosci. Remote Sens. **47**, 2973–2987 (2009)

Tarabalka, Y., Benediktsson, J., Chanussot, J.: Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques. IEEE Trans. Geosci. Remote Sens. **47**(8), 2973–2987 (2009)

Tarabalka, Y., Tilton, J., Benediktsson, J., Chanussot, J.: A marker-based approach for the automated selection of a single segmentation from a hierarchical set of image segmentations. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. **5**(1), 262–272 (2012)

Thenkabail, P., Smith, R., De Pauw, E.: Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. Remote Sens. Environ. **71**(2), 158–182 (2000)

Tilton, J., Tarabalka, Y., Montesano, P., Gofman, E.: Best merge region growing with integrated region object classification. IEEE Trans. Geosci. Remote Sens. **50**, 4454–4467 (2012)

Torgerson, W.S.: Multidimensional scaling: I. theory and method. Psychometrika **17**(4), 401–419 (1952)

Van Der Meer, F.: Analysis of spectral absorption features in hyperspectral imagery. Int. J. Appl. Earth Obs. Geoinform. **5**(1), 55–68 (2004)

Zhang, X., Chew, S.E., Xu, Z., Cahill, N.D.: Slic superpixels for efficient graph-based dimensionality reduction of hyperspectral imagery. In: IAlgorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXI, Proc. SPIE. International Society for Optics and Photonics (2015)

Zhang, Z., Zha, H.: Principal manifolds and nonlinear dimension reduction via local tangent space alignment. SIAM J. Sci. Comput. **26**, 313–338 (2002)