Fusion of Hyperspectral and LiDAR Data Using Sparse and Low-Rank Component Analysis

Behnood Rasti, *Member, IEEE,* Pedram Ghamisi, *Member, IEEE,* Javier Plaza, *Senior Member, IEEE,* and Antonio Plaza, *Fellow, IEEE,*

***Abstract*—The availability of diverse data captured over the same region makes it possible to develop multisensor data fusion techniques to further improve the discrimination ability of classifiers. In this paper, a new sparse and low-rank technique is proposed for the fusion of hyperspectral and LiDAR-derived features. The proposed fusion technique consists of two main steps. First, extinction profiles are used to extract spatial and elevation information from hyperspectral and LiDAR data, respectively. Then, the sparse and low-rank technique is utilized to estimate the low-rank fused features from the extracted ones which are eventually used to produce a final classification map. The proposed approach is evaluated over an urban data set captured over Houston, USA, and a rural one captured over Trento, Italy. Experimental results confirm that the proposed fusion technique outperforms the other techniques used in the experiments based on the classification accuracies obtained by random forest (RF) and support vector machine (SVM) classifiers. Moreover, the proposed approach can effectively classify joint LiDAR and hyperspectral data in an ill-posed situation when only a limited number of training samples are available.**

***Index Terms*—Feature fusion; hyperspectral; LiDAR; sparse and low-rank component analysis; extinction profiles.**

I. Introduction

Nowadays, diverse remote sensors are available, thus allowing to obtain complementary information from different sources for materials on the surface of the Earth. Such information can vary from spectral information obtained by passive sensors [e.g., multispectral and hyperspectral images (HSIs)], to height and shape information acquired by Light Detection And Ranging (LiDAR) sensors, as well as texture information to amplitude and phase by Synthetic Aperture Radar (SAR). The availability of data coming from these multiple sources now allows researchers worldwide to integrate such diverse information to improve object detection ability and classification performance. Regardless of the great amount of knowledge available in such data sets, automatic interpretation of remote sensed data still remains a challenge [1].

B. Rasti is with Keilir Institute of Technology (KIT), Graenasbraut 910, 235 Reykjanesbaer, Iceland (behnood@Keilir.net) and the Department of Electrical and Computer Engineering, University of Iceland, Hjardarhagi 6, 107 Reykjavik, Iceland (behnood.rasti@hi.is).

P. Ghamisi is with German Aerospace Center (DLR), Remote Sensing Technology Institute (IMF) and Technische Universität München (TUM), Signal Processing in Earth Observation, Munich, Germany (corresponding author, e-mail: pedram.ghamisi@dlr.de).

J. Plaza and A. Plaza are with the Hyperspectral Computing Laboratory, Escuela Politecnica de Caceres, University of Extremadura (e-mail: jplaza@unex.es, aplaza@unex.es).

Manuscript received April 19, 2005; revised January 11, 2007.

HSIs are capable of defining the phenomenology and the spectral characteristics of different objects over a detailed spectral signature. LiDAR data, instead, can be used to characterize the elevation and object height information of the scene. These two datasets have been intensively investigated for different tasks [2].

Urban scenes are usually highly complex and challenging. It is generally optimistic to assume that a single sensor can provide enough information for classification and feature extraction [3]. To this end, HSIs may not be able to precisely differentiate objects composed of the same material (i.e., objects with the same spectral characteristics). For instance, roofs and roads, which are made by the same material, exhibit the same spectral characteristics [4]. Therefore, it is difficult to differentiate such categories in the feature space. On the other hand, the use of LiDAR elevation data alone cannot discriminate objects with the same elevation but made of different materials (e.g., roofs with the same elevation built by concrete or asphalt). Furthermore, the individual use of LiDAR data for complex areas, e.g., where many classes are located close to each other, is very limited compared to optical data, due to the lack of spectral information provided by this type of sensors [5], [6].

To take advantage of information provided by different sensors, multisensor data fusion can be taken into account. The joint use of HSI and LiDAR has been investigated in several applications, such as shadow, height, and gap related masking techniques [7]–[9], above-ground biomass estimates [10], micro-climate modelling [11], quantifying riparian habitat structure [12], and fuel type mapping [13]. Moreover, the joint use of LiDAR and HSI has led to higher discrimination power in the feature space compared to the individual use of each source [1], [14]–[18]. For instance, in [17], the joint investigation of HSI and LiDAR was taken into account for the classification of complex forested areas using only pure spectral information classified by either support vector machines (SVMs) or Gaussian maximum likelihood. In [19], deep convolutional neural network was developed to fuse features extracted from HSI and LiDAR to precisely classify land-cover classes. In [18], graph-based feature fusion and morphological profiles were used to fuse LiDAR-derived features and HSI in a fewer dimensional space. In [20], HSI and LiDAR have been considered to deal with individual tree classification, which could also be extended to areas of shadow caused by the illumination of tree crowns with sunlight. That approach was applied to a complex forested area to classify the scene into 16 classes of tree species. In [21], spectral, spatial, and elevation

features extracted from HSI and LiDAR were fused via orthogonal total variation component analysis, which is able to estimate the fused features in a lower dimensional space and also promote piece-wise smoothness while maintaining the spatial structures. The aforementioned works indicate that LiDAR and HSI can complement each other effectively and that integrating these two sources of information in an appropriate manner, one can make the most of the advantages of the two, while addressing the shortcomings of each of them. However, the automatic integration of multiple types of data is not a trivial task [16]. In addition, spatial information, which plays a key role for the classification of HSI, in particular the ones which are of high spatial resolution, has been neglected in most of the conventional approaches.

In 2016, the concept of extinction profiles (EPs) was proposed [22] and later in [23], the concept of EPs has been generalized to extract spatial and contextual information from HSI. In contrast with attribute profiles (APs) [24]–[26], EPs preserve the height of the extrema kept [22], which leads to a higher simplification for recognition capability. This advantage leads to higher classification accuracy for EPs compared to the results obtained by APs. More importantly, EPs are automatic in nature, and also independent from the kind of the attribute being used (e.g. area, volume, etc). However, the initialization of threshold values used in APs is difficult and time consuming.

Although the joint use of HSI and LiDAR information can potentially improve classification accuracies, the automatic fusion of such data is not straightforward [16]. Moreover, the simple concatenation of extracted features obtained by different sensors might increase the so-called *curse of dimensionality*, while the number of training samples is limited [25], [27]–[31]. To address this issue, different feature reduction approaches can be taken into account [25]. This encourages to develop an effective and efficient fusion approach to perform both dimensionality reduction and feature fusion simultaneously [18].

Due to the spectral redundancy in HSI, it has been shown that low-rank HSI modeling provides great advantages in HSI analyses, such as denoising [32], unmixing [33], and feature extraction [34]. In [34], HSI low-rank modeling was incorporated with sparse regression called sparse and low-rank component analysis. A sparse and low-rank component analysis using wavelet bases given in [35] where shown that HSI can be restored based on a few sparse components.

Hyperspectral and LiDAR data provide valuable information of a scene such as height, spatial, and spectral characteristics. On the one hand, extracting different kinds of information and features is crucial for the classification task. On the other hand, it could decrease the classification accuracy due to the Hughes phenomenon. As a result, in this contribution we seek for a remedy to the aforementioned dilemma. Therefore, instead of stacking the extracted features, here, we estimate the fused features assumed to live in a lower dimension space. More specifically, in this paper, a new fusion technique for HSI and LiDAR data is proposed based on the use of sparse and low-rank component analysis applied on the EPs. First, EPs are used to extract spatial and elevation information from HSI and LiDAR, respectively. At the next stage, the HSI as well as the extracted EPs from HSI and LiDAR are fused based on a sparse and low-rank technique, where the spectral redundancy of the features are captured by the low-rank property, while the sparsity property helps to capture spatial redundancy of the features. The sparsity property promotes the spatial smoothness on the fused features which leads to a region-wise homogeneous classification map and the low-rank property of the fusion technique avoid the Hughes phenomenon [27] and therefore, both improve the classification accuracy. The proposed fusion technique, is evaluated based on classification accuracies obtained by applying both RF and SVM classifiers on the fused features. The performance of the proposed fusion approach is also evaluated in a situation when there is only a limited number of training samples are available.

The rest of the paper is organized as follows: After giving a short description for the notations used in the paper, Section II describes the proposed fusion technique. The experiments are described in Section III. Finally, Section IV concludes the paper with some remarks.

A. Notation

In this paper, the number of bands and pixels in each band of the HSI are denoted by p and n, respectively. Matrices are denoted by bold and capital letters, column vectors by bold letters, the element placed in the ith row and jth column of matrix $\mathbf{X}$ by x_{ij} and the ith column by $\mathbf{x}_{(i)}$. The identity matrix of size $p \times p$ is denoted by $\mathbf{I}_p$. $\hat{\mathbf{X}}$ stands for the estimate of the variable $\mathbf{X}$, and $\mathbf{X}^m$ denotes the estimate of the variable $\mathbf{X}$ at mth iteration. The Frobenius norm and the Kronecker product are denoted by $\|.\|_F$ and $\otimes$, respectively. The matrix vectorization operator is shown by vec.

II. Methodology

As can be seen from Algorithm 1, the proposed feature fusion approach is composed of two main phases. In the first phase, EPs are used to extract spatial and elevation features from HSI and LiDAR, respectively. The second phase fuses spectral, spatial, and elevation features using sparse and low-rank component analysis. Below, we elaborate on these two phases.

A. Phase I: Extinction Profiles (EPs)

Ghamisi *et al.* [22] proposed EPs using a set of extinction filters, which are connected and able to maintain relevant image extrema. Relevance here is defined with respect to the concept of extinction value, proposed by Vachier [36]. The extinction value of a regional extremum (minimum or maximum) of any increasing attribute is the maximal size of the attribute filter [37], such that this extremum still exists after filtering [36].

The definition of the extinction value for a regional maximum given by Vachier [36] is as follows: Let M be a regional maximum of a gray scale image $\mathbf{X}$, and $\Psi = (\psi_\lambda)_\lambda$ represents a family of decreasing connected anti-extensive transformations. The extinction value corresponding to M with

Algorithm 1: SLRCA Fusion

Input:
HSI = Hyperspectral data,
L = LiDAR data,
r: Number of fused features,
λ: Regularization tuning parameter,
ϵ: Tolerance values.
Output:
$\hat{\mathbf{F}}_{fused}$: Fused features estimated.

Phase I :
$\mathbf{H}$ = *normalized*(**HSI**),
$\mathbf{F}_{\text{HSI}} = \textit{normalized}(\text{EP}(\mathbf{HSI}))$,
$\mathbf{F}_{\text{LiDAR}} = \textit{normalized}(\text{EP}(\mathbf{L}))$,
$\mathbf{F} = [\mathbf{F}_{\text{HSI}}, \mathbf{H}, \mathbf{F}_{\text{LiDAR}}]$,

Phase II :
Initialization;$\mathbf{V}^0$,
while $\left|J^{(k+1)} - J^{(k)}\right| \leq \epsilon$ **do**
  W-step :
  $\mathbf{S} = \mathbf{D}^T\mathbf{F}\mathbf{V}^m$,
  $\mathbf{W}^{m+1} = \text{Soft}(\mathbf{S}, \lambda)$,
  V-step :
  ${\mathbf{W}^{m+1}}^T\mathbf{D}^T\mathbf{F} = \mathbf{Q}\mathbf{\Sigma}\mathbf{G}^T$,
  $\mathbf{V}^{m+1} = \mathbf{Q}\mathbf{G}^T$,
end
$\hat{\mathbf{F}}_{fused} = \mathbf{D}\hat{\mathbf{W}}$.

respect to Ψ denoted by $\varepsilon_\Psi(M)$ is the maximal λ value, such that M is still a regional maxima of $\psi_\lambda(\mathbf{X})$. This definition can be shown as follows:

$$\varepsilon_\Psi(M) = sup\{\lambda \geq 0 | \forall \mu \leq \lambda, M \subset Max(\psi_\mu(\mathbf{X}))\}, \quad (1)$$

where $Max(\psi_\mu(\mathbf{X}))$ is a set containing all the regional maxima of $\psi_\mu(\mathbf{X})$. In the same manner, extinction values of regional minima can be defined. Extinction values can be efficiently computed on a max-tree structure [38].

Extinction filters (EFs) are connected filters, which preserve the relevant extrema of the gray scale image $\mathbf{X}$. This filtering approach are defined as follows: Let $Max(\mathbf{X}) = \{\mathbf{M}_1, \mathbf{M}_2, ..., \mathbf{M}_N\}$ be the regional maxima of the gray scale image $\mathbf{X}$. For the input gray scale image $\mathbf{X}$, the EF preserves the n maxima with the highest extinction values, $\text{EF}^n(\mathbf{X}) = \text{R}^\delta_{\mathbf{X}}(\mathbf{G})$, where $\text{R}^\delta_{\mathbf{X}}(\mathbf{G})$ denotes the reconstruction by dilation [39] of the mask image $\mathbf{X}$ from marker image $\mathbf{G}$. The marker image can be obtained through $\mathbf{G} = \max_{i=1}^{n}\{\boldsymbol{M}'_i\}$, where $\max$ is the pixel-wise maximum operation. $\boldsymbol{M}'_1$ is the maximum with the highest extinction value, followed by $\boldsymbol{M}'_2$ with the second highest extinction value, and so on. For detailed description, please see [22].

EFs for increasing attributes (e.g., area, height, volume, and diagonal of the bounding box) can be efficiently implemented using the max-tree structure [40]. The use of EFs for non-increasing attributes (e.g., standard deviation), however, demands the construction of the second tree which takes the image to the space of shapes [41] allowing the creation of a novel class of connected operators from the leveling family and more complex morphological analysis, such as the computation of extinction values for non-increasing attributes. For more information please see [22].

EPs are constructed by applying several extinction filters, *i.e.*, a sequence of thinning and thickening transformations, with progressively higher threshold values to extract spatial and contextual information of the input data. The EP for the input gray scale image, $\mathbf{X}$, is obtained by (2):

$$\begin{aligned}\text{EP}(\mathbf{X}) = \{&\underbrace{\phi^{\text{P}_{\lambda_s}}(\mathbf{X}), \phi^{\text{P}_{\lambda_{s-1}}}(\mathbf{X}), \ldots, \phi^{\text{P}_{\lambda_1}}(\mathbf{X})}_{\text{thickening profile}}, \mathbf{X}, \\ &\underbrace{\gamma^{\text{P}_{\lambda_1}}(\mathbf{X}), \ldots, \gamma^{\text{P}_{\lambda_{s-1}}}(\mathbf{X}), \gamma^{\text{P}_{\lambda_s}}(\mathbf{X})}_{\text{thinning profile}}\},\end{aligned} \quad (2)$$

where $\text{P}_\lambda : \{\text{P}_{\lambda_i}\}$ $(i = 1, \ldots, s)$ is a set of s ordered predicates (i.e., $\text{P}_{\lambda_i} \subseteq \text{P}_{\lambda_k}$, $i \leq k$). It should be noted that the number of extrema is considered as the predicates. ϕ and γ are thickening and thinning transformations, respectively.

EPs can be of any type. In this context, multi-EPs (MEPs) concatenates several types of EPs (e.g., area, height, volume, diagonal of bounding box, and standard deviation) as a single stacked vector, which is defined as follows:

$$\text{MEP}(\mathbf{X}) = \{\text{MEP}_{\mathbf{a_1}}(\mathbf{X}), \text{MEP}_{\mathbf{a_2}}(\mathbf{X}), ..., \text{MEP}_{\mathbf{a_w}}(\mathbf{X})\}, \quad (3)$$

where a_k, $k = \{1, ..., w\}$ represents different types of extinction attributes. Since different extinction attributes provide complementary spatial and contextual information, the MEP has a greater ability in extracting spatial information than a single EP.

The EP (and its extension MEP) described above was introduced to extract spatial and contextual information from gray scale images. In order to make the EP applicable for HSI, one can extract a few informative features from the whole dimensionality using an approach such as independent component analysis (ICA). Then the extracted features are considered as base images to produce EPs [23]. In this way, an extended extinction profile (EEP) can be obtained, which is a generalization of the EPs. More precisely, EEP, first, reduces the dimensionality of the data from $E \subseteq \mathbf{Z}^n$ to $E' \subseteq \mathbf{Z}^m$ $(m \leq n)$ with a generic transformation $\Psi : E \to E'$ (i.e., ICA). Second, the EP is performed on the most informative features $\mathbf{Q}_i$ $(i = 1, \ldots, m)$, which can defined as:

$$\text{EEP}(\mathbf{Q}) = \{\text{EP}(\mathbf{Q}_1), \text{EP}(\mathbf{Q}_2), \ldots, \text{EP}(\mathbf{Q}_m)\}. \quad (4)$$

In order to effectively exploit spatial and contextual information from HSIs, different extinction attribute filters a_k, $k = \{1, ..., w\}$ can be applied to the first informative features of ICA [i.e., $\mathbf{Q}_i$ $(i = 1, \ldots, m)$]. In this manner, EMEP can be constructed.

$$\text{EMEP}(\mathbf{Q}) = \{\text{MEP}(\mathbf{Q}_1), \text{MEP}(\mathbf{Q}_2), \ldots, \text{MEP}(\mathbf{Q}_m)\}. \quad (5)$$

Fig. 1 illustrates a general work flow of the EMEP. For detailed information, please see [22], [23]. It is important to note that the EMEP and EP approximately require the same computational time since the most time consuming part is on the construction of the max-tree and min-tree, which are computed only once for each gray scale image [22], [23].

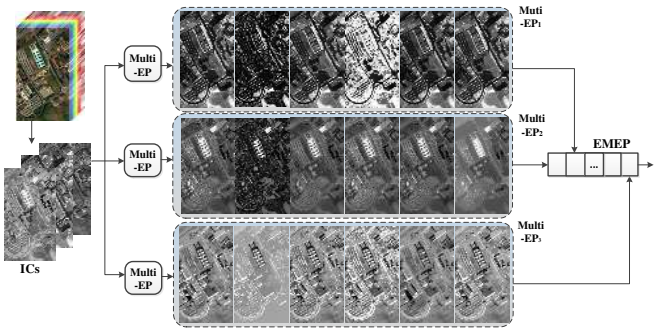


Fig. 1: A simple work flow of the EMEP. First, we only preserve the first three independent components extracted from the whole hyperspectral data cube. Then, each independent component is treated as one base image to produce multi-EPs including five extinction attributes (i.e., area, volume, standard deviation, diagonal of the bounding box, and height). Finally, all Muli-EPs extracted from three different ICs are concatenated to construct the EMEP.

Since the LiDAR-derived DSM contains only one component, its corresponding elevation information can be extracted using equation (3) where $\mathbf{X}$ denotes the LiDAR-derived DSM feature.

EPs extract spatial and contextual information considering the number of extrema, which leads not only to better results in terms of classification accuracies compared to APs but also decreases the burden of setting threshold values, which was a burden for conventional APs [22].

For the sake of simplicity, in the rest of the paper, **HSI** denotes the input hyperspectral data (i.e., spectral information), $\mathbf{EP}_{\text{HSI}}$ represents EMEP($\mathbf{Q}$) (i.e., spatial information), and $\mathbf{EP}_{\text{LiDAR}}$ demonstrates MEP($\mathbf{X}$) (i.e., elevation information).

B. Phase II: Feature Fusion Using Sparse and Low-Rank Component Analysis

To fuse spectral (**HSI**), spatial ($\mathbf{EP}_{\text{HSI}}$), and elevation ($\mathbf{EP}_{\text{LiDAR}}$) features, one needs to normalize the number of dimensionalities to put the same weight on each type of the features and reduce the computational cost and noise throughout the feature space [42]. To do so, Kernel principal component analysis (KPCA) [43] was used as an effective tool to reduce the dimensionality of each type of features independently, since it can represent a higher-order complex and nonlinear distribution in a fewer number of dimensions to address Hughes phenomenon [27] and high computational cost. The normalized dimension of **HSI**, $\mathbf{EP}_{\text{HSI}}$, and $\mathbf{EP}_{\text{LiDAR}}$ is automatically set to the smallest dimension of the above-mentioned features. For example, for the Houston data, this value is set to 71 [1].

Let $\mathbf{H}$ be the normalized matrix contained the input spectral information (band i is located in column i). $\mathbf{F}_{\text{HSI}}$ represents the normalized spatial features produced by EPs on the first three independent components (*i*th feature in its *i*th column). $\mathbf{F}_{\text{LiDAR}}$ is the normalized elevation features obtained by EPs on the LiDAR derived digital surface model (DSM) (*i*th feature in its *i*th column).

Extracted features from HSI and LiDAR are highly redundant. In order to reduce the features redundancy, we propose a low-rank model for the fused features. In other words, the extracted features from LiDAR and HSI can be represented in a space of lower dimension. Note that, this redundancy of features can affect the classification results due to the Hughes phenomenon [27] and also fused features are expected to have a lower dimension. Hence, we suggest to use the following low-rank model

$$\mathbf{F} = \mathbf{D}\mathbf{W}\mathbf{V}^T + \mathbf{N}, \tag{6}$$

where $\mathbf{F} = [\mathbf{F}_{\text{HSI}}, \mathbf{H}, \mathbf{F}_{\text{LiDAR}}] = \left[\mathbf{f}_{(i)}\right]$ is an $n \times p$ matrix containing the ith vectorized feature in its ith column, $\mathbf{V}$ is an unknown subspace (low-rank) basis ($p \times r$), $\mathbf{D}$ is an $n \times n$ orthogonal 2D wavelet transform matrix, $\mathbf{W} = \left[\mathbf{w}_{(i)}\right]$ is an $n \times r$ matrix containing the unknown 2D wavelet coefficients for the i-th component in its i-th columns, and $\mathbf{N} = \left[\mathbf{n}_{(i)}\right]$ is an $n \times p$ matrix containing the vectorized noise and error at band i in its ith column. The purpose of the 2D wavelet transformation $\mathbf{D}$ is to capture the spatial correlations in the model and the purpose of $\mathbf{V}$ is to capture the low-rank structure of the features. Note that r is the number of fused features ($1 \leq r \leq p$), and p is the total number of extracted features including spectral features. It is worth mentioning that model (6) preserves spatial information of the features (see Appendix A).

In order to estimate the wavelet coefficients $\mathbf{W}$ and the basis matrix $\mathbf{V}$ in (6), we use a sparse and low-rank component analysis given in [35], which is based on solving the following non-convex ℓ_1 penalized least squares problem

$$(\hat{\mathbf{W}}, \hat{\mathbf{V}}) = \arg\min_{\mathbf{W},\mathbf{V}} \; J(\mathbf{W}, \mathbf{V}) = \arg\min_{\mathbf{W},\mathbf{V}} \; \frac{1}{2}\left\|\mathbf{F} - \mathbf{D}\mathbf{W}\mathbf{V}^T\right\|_F^2 + \lambda \sum_{i=1}^{r} \left\|\mathbf{w}_{(i)}\right\|_1 \quad \text{s.t.} \quad \mathbf{V}^T\mathbf{V} = \mathbf{I}_r, \tag{7}$$

The estimated fused features are given by $\hat{\mathbf{F}}_{fused} = \mathbf{D}\hat{\mathbf{W}}$.

C. Estimation

A cyclic descent (CD)-type algorithm given in [35], [44] is used to solve (7) called sparse and low-rank component analysis (SLRCA), which solves the non-convex problem (7) w.r.t. one variable at a time while the other variable is assumed to be fixed. Therefore, SLRCA consists of the following two steps:

1) W-step: Given a fixed V the optimization problem (7) can be rewritten as

$$\arg\min_{\mathbf{W}} \; \frac{1}{2}\left\|\mathbf{F} - \mathbf{D}\mathbf{W}\mathbf{V}^T\right\|_F^2 + \lambda \sum_{i} \left\|\mathbf{w}_{(i)}\right\|_1. \tag{8}$$

Since $\mathbf{D}$ and $\mathbf{V}$ are orthogonal matrices, it can be shown [35] that the minimization problem (8) is equivalent to

$$\arg\min_{\mathbf{W}} \; \frac{1}{2}\left\|\mathbf{S} - \mathbf{W}\right\|_F^2 + \lambda \sum_{i} \left\|\mathbf{w}_{(i)}\right\|_1, \tag{9}$$

where $\mathbf{S} = \mathbf{D}^T\mathbf{F}\mathbf{V}$. It can be shown ([35]) that the solution to this minimization problem is given by

$$\hat{w}_{ji} = \max\left(0, |s_{ji}| - \lambda\right) \frac{s_{ji}}{|s_{ji}|}. \tag{10}$$

Function (10) is called soft-thresholding and often is written as

$$\hat{\mathbf{W}} = \text{soft}\left(\mathbf{S}, \lambda\right). \tag{11}$$

Note that soft funstion in (11) is applied element-wise on the matrix $\mathbf{S}$.

2) V-step: Given a fixed $\mathbf{W}$, the optimization problem (7) turns to a Reduced-Rank Procrustes problem [45] as

$$\arg\min_{\mathbf{V}} \left\|\mathbf{F} - \mathbf{D}\mathbf{W}\mathbf{V}^T\right\|_F^2 \quad \text{s.t.} \quad \mathbf{V}^T\mathbf{V} = \mathbf{I}_r,$$

which has a solution given by $\hat{\mathbf{V}} = \mathbf{Q}\mathbf{G}^T$ where $\mathbf{Q}$ and $\mathbf{G}$ are computed using Singular Value Decomposition (SVD) of $\mathbf{M} = \mathbf{W}^T\mathbf{D}^T\mathbf{F} = \mathbf{Q}\mathbf{\Sigma}\mathbf{G}^T$. A description of the method is given in Algorithm 1.

III. Experimental Results

A. Data Description

1) Houston Data: The data are composed of a hyperspectral image and a LiDAR-derived digital surface model (DSM). This dataset was distributed for the 2013 GRSS data fusion contest. The hyperspectral data were acquired by the Compact Airborne Spectrographic Imager (CASI) over the University of Houston campus and the neighboring urban area on June 23, 2012. The LiDAR data were acquired on June 22, 2012. The datasets were collected by the NSF-funded Center for Airborne Laser Mapping (NCALM). The size of the data is 349 $\times$ 1905 with the spatial resolution of 2.5*m*. The hyperspectral dataset consists of 144 spectral bands ranging 0.38-1.05μm. The 15 classes of interests are: Grass Healthy, Grass Stressed, Grass Synthetic, Tree, Soil, Water, Residential, Commercial, Road, Highway, Railway, Parking Lot 1, Parking Lot 2, Tennis Court and Running Track. The "Parking Lot 1" includes parking garages at the ground level and also in elevated areas, while "Parking Lot 2" corresponded to parked vehicles. Fig. 2 shows a false color composite representation of the hyperspectral data and the corresponding training and test samples. Table I gives information about the number of training and test samples for different classes of interests.

It should be noted that we have used the standard sets of training and test samples for the above-mentioned datasets to make the results fully comparable with the available literature.

Cloud shadows in the hyperspectral data were detected using thresholding of illumination distributions calculated by the spectra. Relatively small structures in the thresholded illumination map were removed based on the assumption that cloud shadows are larger than structures on the ground.[1]

[1]The enhanced dataset was provided by Prof. Naoto Yokoya from Technical University of Munich (TUM).

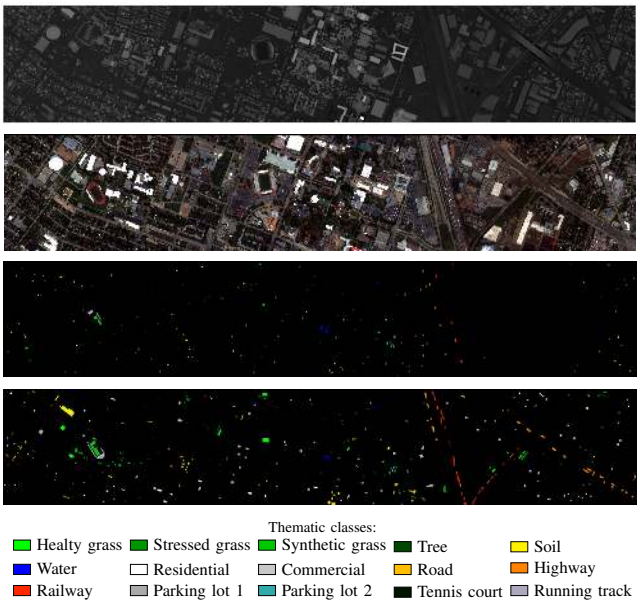


Fig. 2: Houston - From top to bottom: LiDAR-derived rasterized data set, a color composite representation of the HSI using bands 64, 43, and 22 as R, G, and B, respectively; Training samples; Test samples; and legend of different classes.

TABLE I: Houston - Number of Training and Test Samples.

Class		Number of Samples	
No	Name	Training	Test
1	Grass Healthy	198	1053
2	Grass Stressed	190	1064
3	Grass Synthetic	192	505
4	Tree	188	1056
5	Soil	186	1056
6	Water	182	143
7	Residential	196	1072
8	Commercial	191	1053
9	Road	193	1059
10	Highway	191	1036
11	Railway	181	1054
12	Parking Lot 1	192	1041
13	Parking Lot 2	184	285
14	Tennis Court	181	247
15	Running Track	187	473
Total		2,832	12,197

2) Trento Data: The second dataset was captured over a rural area in the south of the city of Trento, Italy. The size of the dataset is of 600 by 166 pixels. The LiDAR DSM data were acquired by the Optech ALTM 3100EA sensor and the hyperspectral data captured by the AISA Eagle sensor, all with the spatial resolution of 1*m*. The hyperspectral data consist of 63 bands ranging from 402.89 to 989.09*nm*, where the spectral resolution is 9.2*nm*. The spatial resolution of this dataset is 1*m*. For this dataset, six classes of interests were extracted, including Building, Woods, Apple trees, Roads, Vineyard, and Ground. Fig. 3 shows a false color composite representation of the hyperspectral data and the corresponding training and test samples. Table II gives information about the number of training and test samples for different classes of interests.

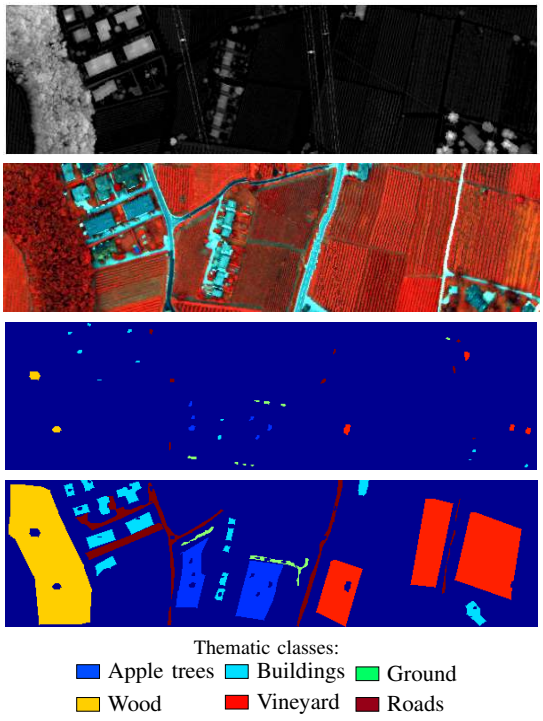


Fig. 3: Trento - From top to bottom: LiDAR-derived rasterized data set, a color composite representation of the HSI using bands 40, 20, and 10, as R, G, and B, respectively; Training samples; Test samples; and legend of different classes.

TABLE II: Trento - Number of Training and Test Samples.

Class		Number of Samples	
No	Name	Training	Test
1	Apple trees	129	3905
2	Buildings	125	2778
3	Ground	105	374
4	Wood	154	8969
5	Vineyard	184	10317
6	Roads	122	3252
Total		819	29595

B. Algorithm Setup

For the EPs, one only needs to define the number of desired levels (s) as the whole process is automatic. In this context, in order to generate the EP for area, volume, and diagonal of the bounding box, the threshold values used to generate the profile are automatically given by $\lfloor 3^j \rfloor$, where $j = 0, 1, ..., s-1$. The size of the EPs is $2s+1$, since the original image should also be included in the profile. The profiles have been computed using the 4-connected connectivity rule. Here, s is set to seven, as suggested in [23].

SLRCA is initialized as suggested in [35]. The tuning parameter λ indicates the level of smoothness. In the experiments, λ is set to one percent of the intensity range of the extracted features. A fast wavelet toolbox provided in [46] was used for the implementation of Wavelet transforms. Daubechies wavelet with 2 coefficients and five decomposition levels is used in all the experiments.

In terms of the SVM, a radial basis function (RBF) kernel is used. The optimal hyperplane parameters C (parameter that controls the amount of penalty during the SVM optimization) and γ (spread of the RBF kernel) have been traced in the range of $C = 10^{-2}, 10^{-1}, ..., 10^4$ and $\gamma = 10^{-3}, 10^{-2}, ..., 10^4$ using five-fold cross validation.

For the RF, the number of trees is set to 300. The number of the prediction variable is set approximately to the square root of the number of input bands.

For the sake of simplicity, the following names are used in the experimental part: **LiDAR** and **HSI** show the classification accuracies of the LiDAR-derived DSM and HSI, respectively. $\mathbf{EP}_{\text{LiDAR}}$, and $\mathbf{EP}_{\text{HSI}}$ show the classification accuracies of EPs applied to LiDAR, and HSI. $\mathbf{EP}_{\text{LiDAR+HSI}}$ refers to the classification accuracies of EPs applied to the stack of LiDAR and HSI.

C. Classification Experiments

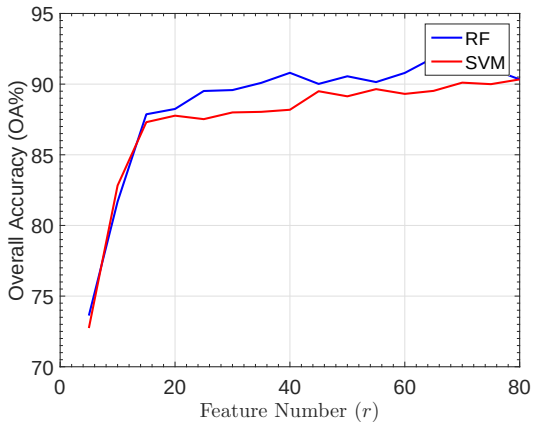
1) Classification accuracies w.r.t. the number of features: Parameter r gives the number of fused features estimated by the proposed algorithm. Therefore, it is of interest to see the performance of the fusion algorithm w.r.t. the number of features selected. To do so, the standard training and test samples given in Tables I and II are used in this section. Fig. 4 (a) and (b) demonstrate the OA in percentage w.r.t. r for Houston and Trento datasets. The simulations are given for $5 \leq r \leq 80$ for every 5 features increment and for both RF and SVM classifiers. As it can be seen in Fig. 4 (a), for Houston, RF gives OA over 90 percent for $r \geq 35$ and OA for SVM gradually increases by r and goes over 90 percent when using more than 70 features.

In the case of the Trento dataset, Fig. 4 (b), for $r \geq 15$ the OA is around 99 percent (the minimum OA is 98.95% for $r = 15$ and the maximum OA is 99.37 for $r = 20$) and for SVM is around 98% (the minimum OA is 97.78% for $r = 40$ and the maximum OA is 98.35% for $r = 15$).

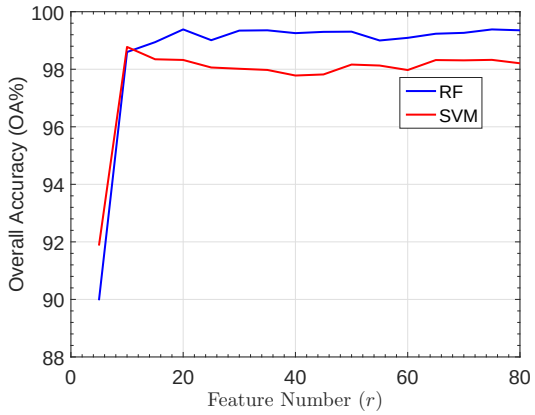
Note that RF classifier considerably outperforms SVM based on OA when using higher features ($r > 10$) while for fewer features ($r = 5$ and $r = 10$) SVM performs better especially in the case of Trento dataset.

The simulations in this section show that, the SLRCA fusion technique by fusing the EP-derived features applied to LiDAR and HSI is able to provide high classification accuracies by using fewer fused features. However, for the proposed method, we suggest the number of fused features to be automatically set to the minimum number of the features in **HSI**, $\mathbf{EP}_{\text{HSI}}$, and $\mathbf{EP}_{\text{LiDAR}}$. In this context, for the Houston data, this value is set to 71 (i.e., the minimum value among 144 features of **HSI**, 213 features of $\mathbf{EP}_{\text{HSI}}$, and 71 features of $\mathbf{EP}_{\text{LiDAR}}$), while for the Trento data, this value is set to 63 (i.e., the minimum value among 63 features of **HSI**, 213 features of $\mathbf{EP}_{\text{HSI}}$, and 71 features of $\mathbf{EP}_{\text{LiDAR}}$).

2) Classification Accuracies w.r.t. the Number of Training Samples: In this subsection, the performance of the proposed fusion technique is investigated in terms of the number of training samples selected for the classification task based on the OA. Here, the number of features selected for the Houston dataset is 71 and for the Trento dataset is 63 as explained

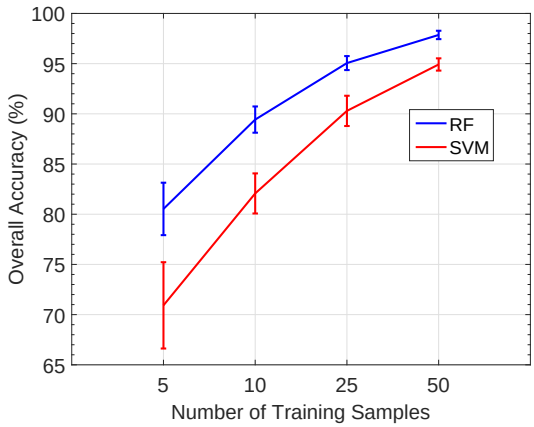


(a) Houston

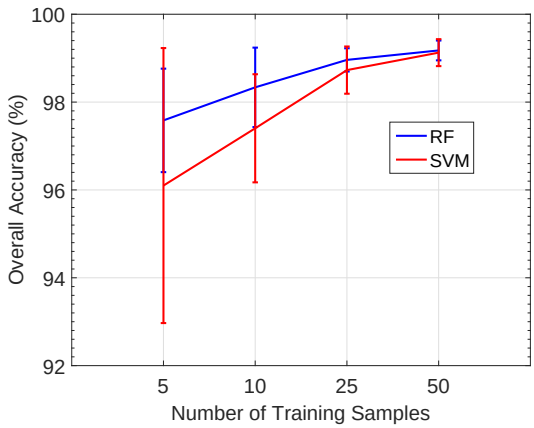


(b) Trento

Fig. 4: Performance of OA w.r.t. feature number r obtained by applying RF and SVM classifiers on the fused features estimated by SLRCA from (a) the University of Houston dataset (b) Trento dataset.



(a) Houston



(b) Trento

Fig. 5: Performance of OA w.r.t. the number of training samples obtained by applying RF and SVM classifiers on the fused features estimated by SLRCA from (a) the University of Houston dataset (b) Trento dataset.

in Subsection III-C1. Fig. 5 (a) and (b) depicts the OA's obtained by RF and SVM using 5, 10, 25, and 50 samples per class for the classification of Houston and Trento dataset, respectively. Here, the results shown are the mean values of 20 times selecting the training samples randomly and the standard deviations were shown by the error bars. As can be seen from the figures, the trends of the graphs are similar for both Houston and Trento datasets. The outcome of this experiment can be summarized as follows:

- SLRCA fusion gives accurate classification results also in the case of having low training samples. For example, in the case of Houston and using RF, as can be seen from Fig. 5 (a), the OA is 89.43% by using only 10 samples per class and it goes over 95% by using 25 samples. Note that, by using 50 training samples per class, SLRCA fusion provides a very high accurate classification results where OA = 97.86%. Also, SVM gives OA = 90.30% and 94.92% using 25 and 50 samples, respectively.
- For both datasets and in all the cases, RF outperforms SVM. This confirms that RF is a better option as a classifier for the proposed technique.
- Comparing the error bars for RF and SVM, confirms the robustness of RF for the classification tasks. This fact is clearer in the case of selecting 5 samples for the classification task. Specifically, in the case of Trento and using 5 training samples per class, the SVM standard deviation is more than double of the RF one.
- Classification accuracies obtained are higher in all cases for Trento dataset compared to Houston, which is expected since dataset was captured over a rural area and therefore, there exists less structure and detail in the scene (and consequently fewer class of interests). For Trento, by using RF, OA starts from 97.58% at 5 samples and goes up to 99.18% at 50 samples and by using SVM, OA starts from 96.10% at 5 samples and goes up to 99.13% at 50 samples.

3) Estimated Fused Features: The fused features estimated by applying the SLRCA fusion are shown in Fig. 6 for the Houston dataset. From top to bottom, the number of features are 1, 3, 6, 13, 20, 30, and 60. As can be seen from Fig. 6, features with lower numbers are more informative visually compared to the ones with higher numbers which is due to the tendency of SLRCA to find a low-rank representation for the high dimensional features. This also can be seen in Fig

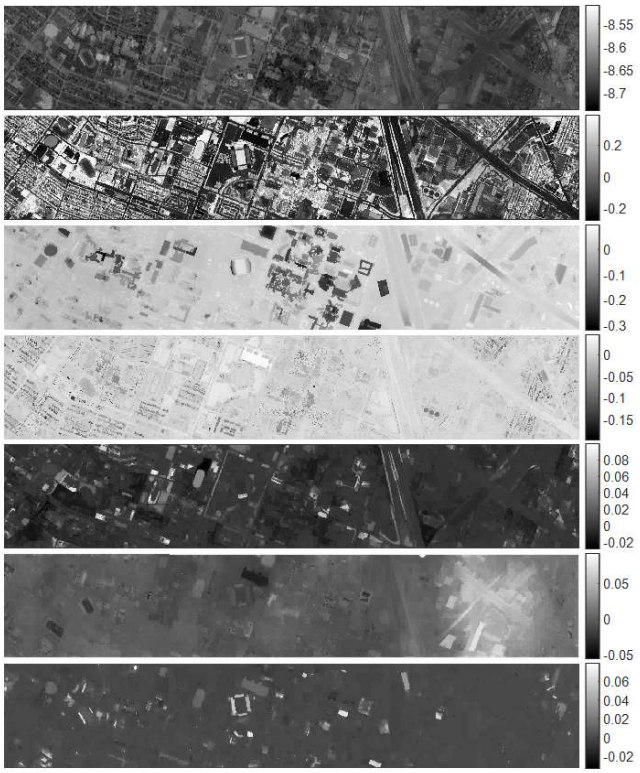


Fig. 6: Fused features estimated by applying SLRCA on EPs obtained from HSI and LiDAR of Houston dataset. From top to bottom number of features are; 1, 3, 6, 13, 20, 30, and 60.

4, where OAs reach to quiet high values using only a few features.

4) SLRCA Fusion Compared With Other Techniques: Here, the proposed fusion technique is compared with other techniques based on the classification accuracies. The classification results are given in Table III for the Houston dataset and Table IV for the Trento dataset. The results are compared based on class accuracies, OA, AA, and kappa coefficient (κ). The number of features used for the classification task in each case were given in the brackets.

As can be seen from Table III, the consideration of the spatial information extracted by the EP can considerably improve classification accuracies compared to the situations where the SVM and RF have directly been applied to the input datasets. For example, EPs on the LiDAR data (with 71 features) substantially improve the OA of **LiDAR** by almost 42% for the RF classifier. In the case of the HSI, due to the rich spectral information, the consideration of the EPs slightly improves the OA by almost 1% and 3%, using SVM and RF, respectively. **LiDAR+HSI** outperform the individual use of each data, which confirms that HSI and LiDAR provide complement information to distinguish different classes of interest. In contrast, the integration of EPs for HSI and LiDAR decreases the OA by almost 1.5% which confirms the drawbacks of high dimensionality on the classification task which can be more severe for the higher number of classes. It can also be seen that, by integrating HSI with EPs for HSI and LiDAR, the accuracies obtained by SVM have not been changed while the ones obtained by RF have been slightly improved. The SLRCA fusion method clearly captures the redundant information existing in the HSI and LiDAR profiles and leads to the accuracy of over 91% which is the best classification accuracy among all the approaches considered in this paper for SVM. Note that the number of features used in the case of SLRCA fusion is 25% of the integration of the profiles.

A similar trend can be seen in the case of using RF, with a difference that the high dimensionality of the EPs does not affect the performance of the RF classifier and as can be seen the integration of HSI and LiDAR profiles improves OA by over 6%. However, SLRCA fusion improves the OA by more than 4% using 75% less features.

As can be seen from Table IV, in the case of the Trento dataset, the use of the EP can considerably improve classification accuracies due to the fact that the EP can effectively extract spatial and contextual information. In addition, for the RF classifier, the 63 fused features obtained by applying OTVCA improve the classification accuracies compared with the integrated profiles (284 features) and for SVM the classification accuracies obtained are slightly less than the integrated profiles.

Overall, Table III and IV show that the SLRCA fusion improves the classification accuracies using fewer number of features for both rural and urban datasets. Moreover, from the tables, it can be seen that the RF provides higher OA, AA, and Kappa coefficients than the SVM.

The classification maps obtained by applying RF and SVM on **HSI**, **LiDAR+HSI**, and the fused features using SLRCA are shown in Fig. 7 and Fig. 8 for Houston and Trento datasets, respectively. It can be seen that the proposed fusion technique provides classification maps having homogeneous regions while preserving the structures, which is greatly of the interest specifically in the case of urban datasets. This is because of using spatial filtering obtained by the sparsity penalty on the sparse wavelet coefficients.

D. Comparison with other Techniques in the Literature

In this subsection, the proposed approach is compared with the state-of-the-art based on the classification accuracies. In the case of Trento dataset, the proposed technique outperforms the ones published in [16], [47] in terms of classification accuracies for both SVM and RF. This improvement might be due to the use of EPs instead of APs in the proposed approach.

In the case of Houston dataset, compared to the 2013 Fusion Contest[2], SLRCA fusion gives competitive classification accuracies. Note that, the classification techniques participated in the contest had been specifically developed for the Houston data and they include several overheads, preprocessing and postprocessing approaches for further classification improvements. In this paper, we have tried to propose a scheme, which is also applicable to other data sets composing of co-registered HSI and LiDAR by preserving the generalization capability of the proposed approach, while achieving the

[2]http://www.grss-ieee.org/community/technical-committees/data-fusion/2013-ieee-grss-data-fusion-classification-contest-results/

TABLE III: Houston - Classification accuracies obtained by different approaches using RF and SVM. The metrics AA an OA are reported in percentage. Kappa coefficient is of no units. The best result is shown in bold. The number of features are written in parentheses.

	LiDAR (1)		**HSI** (144)		$\mathbf{EP}_{\text{LiDAR}}(71)$		$\mathbf{EP}_{\text{HSI}}(213)$		**LiDAR+HSI** (145)		$\mathbf{EP}_{\text{LiDAR+HSI}}(284)$		$\mathbf{EP}_{\text{LiDAR+HSI}}$ + **HSI**(414)		SLRCA fusion (71)	
	SVM	RF	SVM	RF	SVM	RF	SVM	RF	SVM	RF	SVM	RF	SVM	RF	SVM	RF
OA	28.82	31.83	84.69	77.47	67.2	73.42	85.82	80.36	88.43	80.50	86.87	86.98	86.43	88.14	91.21	**91.30**
AA	36.31	37.43	86.34	80.34	70.00	75.97	83.08	83.47	89.64	82.84	88.78	88.54	88.46	89.52	90.09	**91.95**
K	0.2422	0.2677	0.8340	0.7563	0.6440	0.7120	0.8168	0.7876	0.8745	0.7887	0.8577	0.8592	0.8531	0.8717	0.8924	**0.9056**
1	11.68	13.49	**83.48**	83.38	57.36	74.26	79.39	77.49	83.00	**83.57**	79.39	78.06	79.58	81.96	80.06	81.58
2	0.00	16.26	96.43	98.40	40.79	61.75	78.85	78.48	97.84	98.12	80.36	84.96	81.39	97.93	94.74	**99.44**
3	87.13	56.63	99.80	98.02	98.61	97.23	**100.00**	**100.00**	99.80	98.42	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	98.61
4	51.80	44.03	98.77	97.54	92.33	58.14	87.78	82.77	**98.96**	97.82	95.83	95.45	96.12	98.86	96.88	96.12
5	12.12	58.05	98.11	96.40	83.43	82.10	99.81	97.73	98.48	96.40	**99.81**	98.77	98.11	94.41	99.05	99.72
6	78.32	58.04	95.10	97.20	78.32	83.22	95.80	95.80	99.30	95.80	95.80	95.80	95.80	95.80	95.80	**98.60**
7	56.90	39.09	89.09	82.09	55.22	77.33	85.17	73.23	87.41	84.61	80.41	73.41	74.63	74.25	85.73	**90.39**
8	13.11	29.53	45.87	40.65	29.06	68.28	65.15	59.92	70.94	57.74	90.41	85.28	88.89	88.60	86.42	**95.73**
9	14.92	13.60	82.53	69.78	67.33	59.40	89.90	83.00	86.69	70.35	89.80	93.96	80.17	86.12	86.02	**98.21**
10	8.30	11.29	83.20	57.63	61.39	66.89	51.54	64.09	82.53	56.95	56.66	67.08	**72.68**	67.08	66.99	63.42
11	72.68	40.42	83.87	76.09	99.72	99.91	87.76	84.72	89.94	79.98	90.70	90.89	88.61	91.46	**98.29**	90.70
12	0.00	9.99	70.99	49.38	63.11	64.75	84.34	78.10	78.19	59.65	89.91	88.57	88.57	88.38	**96.35**	91.07
13	12.28	15.09	70.53	61.40	49.12	58.60	**84.56**	77.89	72.98	65.26	84.56	76.14	84.21	77.89	81.75	76.49
14	97.57	80.16	100.00	99.60	**100.00**	**100.00**	**100.00**	99.60	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
15	27.91	75.90	97.46	97.67	74.21	87.74	97.25	99.37	98.52	97.89	98.10	**99.79**	98.10	**100.00**	**100.00**	99.15

TABLE IV: Trento - Classification accuracies obtained by different approaches using RF and SVM. The metrics AA an OA are reported in percentage. Kappa coefficient is of no units. The best result is shown in bold. The number of features are written in parentheses.

	LiDAR (1)		**HSI** (63)		$\mathbf{EP}_{\text{LiDAR}}(71)$		$\mathbf{EP}_{\text{HSI}}(213)$		**LiDAR+HSI** (64)		$\mathbf{EP}_{\text{LiDAR+HSI}}(284)$		$\mathbf{EP}_{\text{LiDAR+HSI}}$ + **HSI**(347)		SLRCA fusion (63)	
	SVM	RF	SVM	RF	SVM	RF	SVM	RF	SVM	RF	SVM	RF	SVM	RF	SVM	RF
OA	63.30	46.70	84.55	84.92	81.27	85.17	96.28	95.90	85.06	90.61	98.68	98.39	98.77	9898	98.13	**99.27**
AA	46.14	43.31	85.14	85.01	82.50	84.43	93.89	93.53	88.70	89.17	97.79	97.06	97.08	9765	97.45	**98.55**
K	0.5039	0.3350	0.7965	0.8004	0.7637	0.8099	0.9505	0.9453	0.8068	0.8566	0.9824	0.9785	0.9835	0.9863	0.9751	**0.9902**
1	37.10	42.50	88.40	86.20	98.56	96.06	99.93	97.82	91.27	86.09	99.95	97.62	99.62	**100.00**	**100.00**	99.87
2	41.40	51.30	82.60	85.90	96.21	98.42	97.97	94.25	95.83	93.87	97.24	96.80	96.51	97.48	97.37	**98.74**
3	0.00	34.20	97.60	96.80	70.15	72.03	97.08	94.99	91.65	**97.91**	96.45	94.36	92.25	94.92	94.92	97.33
4	67.40	52.60	96.90	95.70	98.64	99.45	99.81	99.22	98.48	97.05	99.57	99.97	99.79	99.99	99.99	**100.00**
5	87.60	46.50	77.10	80.10	58.39	69.89	99.57	98.76	66.82	82.76	99.26	99.10	99.61	**99.84**	96.90	99.67
6	79.90	32.40	67.90	65.00	73.06	70.79	69.03	76.15	88.19	86.01	94.27	94.55	94.72	93.64	95.54	**95.68**

Fig. 7: Classification maps for Houston data: (a) the outputs of RF on **HSI**, (b) the output of SVM on **HSI**, (c) the output of RF on **LiDAR+HSI**, (d) the outputs of SVM on **LiDAR+HSI**, (e) the output of the proposed method using RF, and (f) the output of the proposed method using SVM.

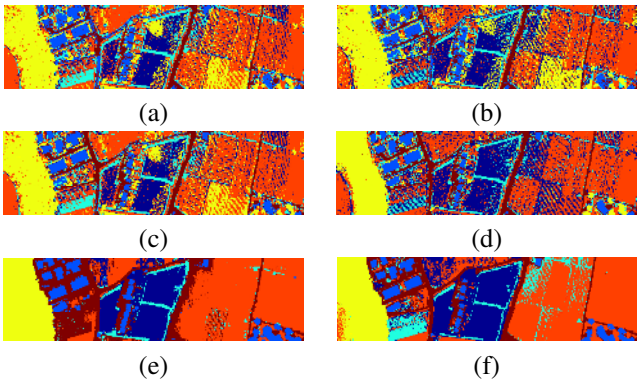


Fig. 8: Classification maps for Houston data: (a) the outputs of RF on **HSI**, (b) the output of SVM on **HSI**, (c) the output of RF on **LiDAR+HSI**, (d) the outputs of SVM on **LiDAR+HSI**, (e) the output of the proposed method using RF, and (f) the output of the proposed method using SVM.

highest classification accuracy on one specific data set is not expected.

Also, SLRCA fusion outperforms the fusion technique given in [2] for the Houston dataset, when no postprocessing is applied. Even when Markov random field (MRF) is used for postprocessing in [2] the classification results obtained by SRLCA fusion are still competitive. However, the considerable improvements obtained in [2] by using the postprocessing step has encouraged us to examine the effect of the hidden MRF proposed in [48] to further improve the classification accuracy of the proposed approach in future.

IV. CONCLUSIONS

In this paper, a new technique for the fusion of hyperspectral and LiDAR data was proposed called SLRCA fusion. The fusion methodology consists of two main phases. At phase I, spatial and elevation information from hyperspectral and LiDAR datasets are extracted using extinction profiles. At phase II, the SLRCA fusion utilizes a sparse and low-rank component analysis to fuse extracted features. The resulting fused features are of lower dimension than the profiles.

For the experiments, Houston (urban) and Trento (rural) datasets have been used. Both RF and SVM classifiers were used to perform the classification task. It has been shown that applying extinction profiles considerably improves the classification accuracies due to effectively extract spatial and contextual information. The integration of LiDAR and HSI profiles increases the dimensionality, and as shown in the experiments, might not successfully improve the classification accuracies due to the Hughes phenomenon. Therefore, SLRCA fusion was used to decrease the dimensionality by fusing the LiDAR and HSI profiles while preserving extracted information.

The experimental results revealed that applying SLRCA fusion improves the classification accuracies compared to the integrated profiles by using low dimensional fused features. In other words, SLRCA fusion is able to capture the redundancy of the features while improves the classification accuracies. This has been shown based on the classification accuracies obtained by using both RF and SVM classifiers for both rural and urban datasets. Furthermore, RF has demonstrated to be well-suited for the classification of the features obtained by the proposed approach.

In addition, it has been shown that the low-dimensional fused features obtained by applying SLRCA fusion provides accurate classification results in the case of selecting few training samples. Last, the experiments confirm that SLRCA fusion is well- competitive with other fusion techniques in the literature which also exploit pre and post processing technique.

V. ACKNOWLEDGMENT

The authors would like to thank Prof. Lorezo Bruzzone of the University of Trento for providing the Trento data set. In addition, the authors would like to express their appreciation to the National Center for Airborne Laser Mapping (NCALM) for providing the Houston data set. The shadow-removed hyperspectral data is provided by Prof. Naoto Yokoya.

APPENDIX A
TWO DIMENSIONAL WAVELET TRANSFORM

The 2D wavelet transform (shown by matrix $\mathbf{D}$) we use is separable in the sense that first 1D wavelet transform is applied on the rows of the image and then on the columns (separable bases). Assuming 2D image $\mathbf{X}$, applying 2D wavelet transform (WT_{2D}) can be written as

$$WT_{2D}(\mathbf{X}) = \mathbf{D}_{1D}\mathbf{X}\mathbf{D}_{1D}^{T}. \tag{12}$$

where $\mathbf{D}_{1D}$ is a matrix contains the 1D wavelet bases in its columns. If we vectorize this expression we get

$$vec(\mathbf{D}_{1D}\mathbf{X}\mathbf{D}_{1D}^{T}) = \mathbf{D}\mathbf{x}, \tag{13}$$

where $\mathbf{D} = \mathbf{D}_{1D} \otimes \mathbf{D}_{1D}$ and $\mathbf{x} = vec(\mathbf{X})$ [49]. We can go from (12) to (13) and vice versa and therefore no spatial information is lost.

REFERENCES

[1] C. Debes, A. Merentitis, R. Heremans, J. Hahn, N. Frangiadakis, T. van Kasteren, W. Liao, R. Bellens, A. Pizurica, S. Gautama, W. Philips, S. Prasad, Q. Du, and F. Pacifici, "Hyperspectral and LiDAR data fusion: Outcome of the 2013 GRSS data fusion contest," *IEEE Jour. Selec. Top. App. Earth Obs. Remote Sens.*, vol. 7, no. 6, pp. 2405–2418, June 2014.

[2] M. Khodadadzadeh, J. Li, S. Prasad, and A. Plaza, "Fusion of hyperspectral and LiDAR remote sensing data using multiple feature learning," *IEEE Jour. Sel. Top. App. Earth Obs. Remote Sens.*, vol. 8, no. 6, pp. 2971–2983, 2015.

[3] P. Gamba, F. Dell'Acqua, and B. V. Dasarathy, "Urban remote sensing using multiple data sets: Past, present, and future," *IEEE Trans. Geos. Remote Sens.*, vol. 6, p. 319326, 2005.

[4] J. Plaza, A. Plaza, P. Gamba, and G. Trianni, "Efficient multi-band texture analysis for remotely sensed data interpretation in urban areas," in *IEEE Urban Remote Sensing Joint Event*, April 2007.

[5] Q. Chen, "Airborne lidar data processing and information extraction," *Photo. Eng. & Remote Sens.*, vol. 73, no. 2, pp. 109 – 112, 2007.

[6] B. Höfle, M. Hollaus, and J. Hagenauer, "Urban vegetation detection using radiometrically calibrated small-footprint full-waveform airborne LiDAR data," *ISPRS Jour. Photo. Remote Sens.*, vol. 67, pp. 134–147, 2012.

[7] G. P. Asner, D. E. Knapp, T. Kennedy-Bowdoin, M. O. Jones, R. E. Martin, and J. Boardman, "Invasive species detection in hawaiian rainforests using airborne imaging spectroscopy and LiDAR," *Remote Sens. Env.*, vol. 112, pp. 1942–1955, 2008.

[8] G. A. Blackburn, "Remote sensing of forest pigments using airborne imaging spectrometer and LiDAR imagery," *Remote Sens. Env.*, vol. 82, p. 311321, 2002.

[9] M. Voss and R. Sugumaran, "Seasonal effect on tree species classification in an urban environment using hyperspectral data, LiDAR, and an object-oriented approach," *Sensors*, vol. 8, pp. 3020–3036, 2008.

[10] R. M. Lucas, A. C. Lee, and P. J. Bunting, "Retrieving forest biomass through integration of CASI and LiDAR data," *Int. Jour. Remote Sens.*, vol. 29, pp. 1553–1577, 2008.

[11] U. Heiden, W. Heldens, S. Roessner, K. Segl, T. Esch, and A. Mueller, "Urban structure type characterization using hyperspectral remote sensing and height information," *Landsc. Urban Plann.*, vol. 105, no. 6, pp. 361–375, June 2012.

[12] R. K. Hall, R. L. Watkins, D. T. Heggem, K. B. Jones, P. R. Kaufmann, and S. B. Moore, "Quantifying structural physical habitat attributes using LiDAR and hyperspectral imagery," *Env. Mon. Asses.*, vol. 159, pp. 63–83, 2009.

[13] B. Koetz, F. Morsdorf, S. Linder, T. Curt, and B. Allgower, "Multi-source land cover classification for forest fire management based on imaging spectrometry and LiDAR data." *For. Ecol. Management*, vol. 256, pp. 263–271, 2008.

[14] J. T. Mundt, D. R. Streutker, and N. F. Glenn, "Mapping sagebrush distribution using fusion of hyperspectral and LiDAR classifications," *Phot. Eng. Remote Sens.*, vol. 72, no. 1, p. 4754, Sept 2006.

[15] R. Sugumaran and M. Voss, "Object-oriented classification of LiDAR fused hyperspectral imagery for tree species identification in an urban environment," *in Proc. Urban Remote Sens. Joint Event*, p. 16, April 2007.

[16] P. Ghamisi, J. A. Benediktsson, and S. Phinn, "Land-cover classification using both hyperspectral and LiDAR data," *Int. Jour. Image Data Fus.*, vol. 6, no. 3, pp. 189–215, 2015.

[17] M. Dalponte, L. Bruzzone, and D. Gianelle, "Fusion of hyperspectral and lidar remote sensing data for classification of complex forest areas," *IEEE Trans. Geos. Remote Sens.*, vol. 46, no. 5, pp. 1416–1427, 2008.

[18] W. Liao, R. Bellens, A. Pizurica, S. Gautama, and W. Philips, "Generalized graph-based fusion of hyperspectral and LiDAR data using morphological features," *IEEE Geos. Remote Sens. Let.*, vol. 12, no. 3, pp. 552–556, 2015.

[19] P. Ghamisi, B. Hfle, and X. X. Zhu, "Hyperspectral and lidar data fusion using extinction profiles and deep convolutional neural network," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. in press, no. 99, pp. 1–14, 2016.

[20] T. Matsuki, N. Yokoya, and A. Iwasaki, "Hyperspectral tree species classification of japanese complex mixed forest with the aid of lidar data," *IEEE Jour. Sel. Topi. App. Earth Obs. Remote Sens.*, vol. 8, no. 5, pp. 2177–2187, May 2015.

[21] B. Rasti, P. Ghamisi, and R. Gloaguen, "Hyperspectral and lidar fusion using extinction profiles and total variation component analysis," *IEEE Trans. Geos. Remote Sens.*, vol. 55, no. 7, pp. 3997–4007, July 2017.

[22] P. Ghamisi, R. Souza, J. A. Beneiktsson, X. X. Zhu, L. Rittner, and R. Lotufo, "Extinction profiles for the classification of remote sensing data," *IEEE Trans. Geos. Remote Sens.*, vol. 54, no. 10, pp. 5631–5645, 2016.

[23] P. Ghamisi, R. Souza, J. A. Benediktsson, L. Rittner, R. Lotufo, and X. X. Zhu, "Hyperspectral data classification using extended extinction profiles," *IEEE Geos. Remote Sens. Let.*, vol. 13, no. 11, pp. 1641–1645, 2016.

[24] P. Ghamisi, M. Dalla Mura, and J. A. Benediktsson, "A survey on spectral–spatial classification techniques based on attribute profiles," *IEEE Trans. Geos. Remote Sens.*, vol. 53, no. 5, pp. 2335–2353, 2015.

[25] J. A. Benediktsson and P. Ghamisi, *Spectral-Spatial Classification of Hyperspectral Remote Sensing Images*. Artech House Publishers, INC, Boston, USA, 2015.

[26] M. Dalla Mura, J. A. Benediktsson, B. Waske, and L. Bruzzone, "Morphological Attribute Profiles for the Analysis of Very High Resolution Images," *IEEE Trans. Geos. Remote Sens.*, vol. 48, pp. 3747–3762, 2010.

[27] G. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Trans. Inf. Theory*, vol. IT, no. 14, pp. 55 – 63, 1968.

[28] P. Ghamisi, "Spectral and spatial classification of hyperspectral data," Ph.D. dissertation, University of Iceland, 2015.

[29] P. Ghamisi and J. A. Benediktsson, "Feature selection based on hybridization of genetic algorithm and particle swarm optimization," *IEEE Geos. Remote Sens. Let.*, vol. 12, no. 2, pp. 309 – 313, 2015.

[30] P. Ghamisi, M. S. Couceiro, and J. A. Benediktsson, "A novel feature selection approach based on FODPSO and SVM," *IEEE Trans. Geos. Remote Sens.*, vol. 53, no. 5, pp. 2935–2947, 2015.

[31] Y. Bazi and F. Melgani, "Toward an optimal svm classification system for hyperspectral remote sensing images," *IEEE Trans. Geos. Remote Sens.*, vol. 44, no. 11, pp. 3374–3385, 2006.

[32] B. Rasti, J. R. Sveinsson, M. O. Ulfarsson, and J. A. Benediktsson, "Hyperspectral image denoising using a new linear model and sparse regularization," in *2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS*, July 2013, pp. 457–460.

[33] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 354–379, April 2012.

[34] B. Rasti, "Sparse Hyperspectral Image Modeling and Restoration," Ph.D. dissertation, University of Iceland, Dec. 2014.

[35] B. Rasti, J. Sveinsson, and M. Ulfarsson, "Wavelet-based sparse reduced-rank regression for hyperspectral image restoration," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 10, pp. 6688–6698, Oct 2014.

[36] C. Vachier, "Extinction value: a new measurement of persistence," in *IEEE Workshop on Nonlinear Signal and Image Processing*, vol. I, 1995, pp. 254–257.

[37] E. J. Breen and R. Jones, "Attribute openings, thinnings, and granulometries." *Comp. Vis. Image Und.*, vol. 64, no. 3, pp. 377–389, 1996.

[38] A. Silva and R. Lotufo, "New extinction values from efficient construction and analysis of extended attribute component tree," in *SIBGRAPI'08*, 2008, pp. 204–211.

[39] P. Soille, *Morphological Image Analysis: Principles and Applications*, 2nd ed. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2003.

[40] R. Souza, L. Rittner, R. Machado, and R. Lotufo, "A comparison between extinction filters and attribute filters," in *ISMM'15*, 2015, pp. 63–74.

[41] Y. Xu, T. Géraud, and L. Najman, "Morphological filtering in shape spaces: Applications using tree-based image representations," *CoRR*, vol. abs/1204.4758, 2012.

[42] W. Liao, R. Bellens, S. Gautama, and W. Philips, "Feature fusion of hyperspectral and LiDAR data for classification of remote sensing data from urban area," in *EARSeL Special Interest Group on Land Use and Land Cover, 5th Workshop*, S. V. D. Linden, T. Kuemmerle, and K. Janson, Eds., 2014, pp. 34–34.

[43] B. Schölkopf, A. Smola, and K. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Comp.*, vol. 10, no. 5, pp. 1299–1319, 1998.

[44] B. Rasti and K. S. Gudmundsson, "Sparse and low-rank feature extraction for the classification of target's tracking capability," *Proc. SPIE*, vol. 9970, pp. 99 701U–99 701U–7, 2016. [Online]. Available: http://dx.doi.org/10.1117/12.2240282

[45] H. Zou, T. Hastie, and R. Tibshirani, "Sparse principal component analysis," *Journal of Computational and Graphical Statistics*, vol. 15, p. 2006, 2004.

[46] B. Rasti, "Wavelab Fast," May 2016. [Online]. Available: https://www.researchgate.net/publication/303445667_Wavelab_fast

[47] M. Pedergnana, P. R. Marpu, M. Dalla Mura, J. A. Benediktsson, and L. Bruzzone, "Classification of remote sensing optical and lidar data using extended attribute profiles," *IEEE Jour. Sel. Top. Signal Proc.*, vol. 6, no. 7, pp. 856–865, 2012.

[48] P. Ghamisi, J. A. Benediktsson, and M. O. Ulfarsson, "Spectral-spatial classification of hyperspectral images based on hidden Markov random fields," *IEEE Trans. Geos. Remote Sens.*, vol. 52, no. 5, pp. 2565–2574, 2014.

[49] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus With Applications in Statistics and Econometrics*, 3rd ed. John Wiley & Sons Ltd, Jan. 2007.

Behnood Rasti (S'12, M'14) received the B.Sc. and M.Sc. degrees in electrical and electronics engineering from the University of Guilan, Rasht, Iran, in 2006 and 2009, respectively, and the Ph.D. degree in electrical and computer engineering from the Department of Electrical and Computer Engineering, University of Iceland, Reykjavik, Iceland, in 2014.

From 2015 to 2016, he was a Post-Doctoral Researcher with the University of Iceland. In 2016, he joined the Keilir Institute of Technology (KIT), Reykjanesbaer, Iceland, where he is currently an Instructor. His research interests include signal and image processing, hyperspectral image analysis, remote sensing data fusion, biomedical engineering, biomedical data analysis, control system, and robotics.

Pedram Ghamisi (S'12, M'15) graduated with a B.Sc. in civil (survey) engineering from the Tehran South Campus of Azad University. He obtained an M.Sc. degree with first class honors in remote sensing at K.N.Toosi University of Technology in 2012. In 2013/2014, he spent seven months at the school of Geography, Planning and Environmental Management, the University of Queensland, Australia. He received a Ph.D. in electrical and computer engineering at the University of Iceland, Reykjavik in 2015 and subsequently worked as a postdoctoral research fellow at the University of Iceland. In 2015, Dr. Ghamisi won the prestigious Alexander von Humboldt Fellowship and started his work as a postdoctoral research fellow at Technical University of Munich (TUM) and Heidelberg University, Germany from October 2015. He has also been working as a researcher at German Aerospace Center (DLR), Remote Sensing Technology Institute (IMF), Germany since October 2015. His research interests are in remote sensing and image analysis, with a special focus on spectral and spatial techniques for hyperspectral image classification, multisensor data fusion, machine learning, and deep learning.

In the 2010-11 academic year, he received the Best Researcher Award for M.Sc. students in K. N. Toosi University of Technology. In July, 2013, he presented at the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Melbourne, and was awarded the IEEE Mikio Takagi Prize for winning the conference Student Paper Competition against almost 70 people. In 2016, he was selected as a talented international researcher by Iran's National Elites Foundation. In 2017, he won the Data Fusion Contest 2017 organized by the Image Analysis and Data Fusion Technical Committee (IADF) of the Geoscience and Remote Sensing Society (IEEE-GRSS). His model was the most accurate among more than 800 submissions. Additional information: http://pedram-ghamisi.com/

Javier Plaza (M09-SM15) is a member of the Hyperspectral Computing Laboratory at the Department of Technology of Computers and Communications, University of Extremadura, where he received the M.Sc. degree in 2004 and the PhD degree in 2008, both in Computer Engineering. He was the recipient of the Outstanding Ph.D. Dissertation Award at the University of Extremadura in 2008. His main research interests comprise hyperspectral data processing and parallel computing of remote sensing data. He has authored more than 140 publications, including 45 JCR journal papers, 10 book chapters, and 90 peer-reviewed conference proceeding papers. He has guest edited 2 special issues on hyperspectral remote sensing for different journals. He is an Associate Editor for IEEE Geoscience and Remote Sensing Letters and an Associate Editor of the IEEE Remote Sensing Code Library. He is a recipient of the Best Column Award of the IEEE Signal Processing Magazine in 2015 and the most highly cited paper (2005-2010) in the Journal of Parallel and Distributed Computing. He received best paper awards at the IEEE International Conference on Space Technology and the IEEE Symposium on Signal Processing and Information Technology. http://www.umbc.edu/rssipl/people/jplaza

Antonio Plaza (M05-SM07-F15) is the Head of the Hyperspectral Computing Laboratory at the Department of Technology of Computers and Communications, University of Extremadura, where he received the M.Sc. degree in 1999 and the PhD degree in 2002, both in Computer Engineering. His main research interests comprise hyperspectral data processing and parallel computing of remote sensing data. He has authored more than 600 publications, including 197 JCR journal papers (over 140 in IEEE journals), 23 book chapters, and 285 peer-reviewed conference proceeding papers. He has guest edited 10 special issues on hyperspectral remote sensing for different journals. Prof. Plaza is a Fellow of IEEE for contributions to hyperspectral data processing and parallel computing of Earth observation data. He is a recipient of the recognition of Best Reviewers of the IEEE Geoscience and Remote Sensing Letters (in 2009) and a recipient of the recognition of Best Reviewers of the IEEE Transactions on Geoscience and Remote Sensing (in 2010), for which he served as Associate Editor in 2007-2012. He is also an Associate Editor for IEEE Access, and was a member of the Editorial Board of the IEEE Geoscience and Remote Sensing Newsletter (2011-2012) and the IEEE Geoscience and Remote Sensing Magazine (2013). He was also a member of the steering committee of the IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS). He is a recipient of the Best Column Award of the IEEE Signal Processing Magazine in 2015, the 2013 Best Paper Award of the JSTARS journal, and the most highly cited paper (2005-2010) in the Journal of Parallel and Distributed Computing. He received best paper awards at the IEEE International Conference on Space Technology and the IEEE Symposium on Signal Processing and Information Technology. He served as the Director of Education Activities for the IEEE Geoscience and Remote Sensing Society (GRSS) in 2011-2012, and as President of the Spanish Chapter of IEEE GRSS in 2012-2016. He has reviewed more than 500 manuscripts for over 50 different journals. He is currently serving as the Editor-in-Chief of the IEEE Transactions on Geoscience and Remote Sensing journal. Additional information: http://www.umbc.edu/rssipl/people/aplaza