Local Feature-Based Attribute Profiles for Optical Remote Sensing Image Classification

Minh-Tan Pham, Sébastien Lefèvre, Erchan Aptoula

Local Feature-based Attribute Profiles for Optical Remote Sensing Image Classification

Minh-Tan Pham, *Member, IEEE,* Sébastien Lefèvre, and Erchan Aptoula

***Abstract*—This article introduces an extension of morphological attribute profiles (APs) by extracting their local features. The so-called local feature-based attribute profiles (LFAPs) are expected to provide a better characterization of each APs' filtered pixel (i.e. APs' sample) within its neighborhood, hence better deal with local texture information from the image content. In this work, LFAPs are constructed by extracting some simple first-order statistical features of the local patch around each APs' sample such as mean, standard deviation, range, etc. Then, the final feature vector characterizing each image pixel is formed by combining all local features extracted from APs of that pixel. In addition, since the self-dual attribute profiles (SDAPs) has been proved to outperform the APs in recent years, a similar process will be applied to form the local feature-based SDAPs (LFSDAPs). In order to evaluate the effectiveness of LFAPs and LFSDAPs, supervised classification using both the Random Forest and the Support Vector Machine classifiers is performed on the very high resolution Reykjavik image as well as the hyperspectral Pavia University data. Experimental results show that LFAPs (resp. LFSDAPs) can considerably improve the classification accuracy of the standard APs (resp. SDAPs) and the recently proposed histogram-based APs (HAPs).**

***Index Terms*—Optical remote sensing imagery, attribute profiles (APs), self-dual attribute profiles (SDAPs), local feature-based APs (LFAPs), local feature-based SDAPs (LFSDAPs), supervised classification**

I. Introduction

CLASSIFICATION of optical remote sensing images is one of the most crucial tasks in land use and land cover earth observation. Among a great number of proposed techniques in the literature (see a review in [1]), morphological attribute profiles (APs) [2] have been widely used thanks to their powerful modeling of spatial information from the image content and their efficient implementation via tree structures. In the past few years, a lot of studies have been contributed to exploit and extend the use of APs [3]–[14]. A recent survey of APs and some of their extensions can be found in [15]. In fact, APs provide a multi-level image representation obtained by the sequential application of different filter rules (i.e. attributes) characterizing the size and shape of objects present in the image [2]. By well preserving important spatial properties of regions and objects such as contours, shape, etc., APs become effective to characterize the contextual information of the observed scene, hence relevant for classification task.

M.T. Pham and S. Lefèvre are with the IRISA (Institut de Recherche en Informatique et Systèmes Aléatoires) laboratory - Université de Bretagne Sud, UMR 6074, Vannes 56000, France. Corresponding author: M.T. Pham, e-mail: minh-tan.pham@irisa.fr.

E. Aptoula is with the Institute of Information Technologies - Gebze Technical University, Kocaeli, 41400, Turkey

Considered as an improved version of APs, the self-dual attribute profiles (SDAPs) [7], [8] have been proved to outperform APs in terms of classification accuracy and computational cost. By applying a sequence of self-dual attribute operators based on a tree of shapes (instead of a max-tree and a min-tree employed by the original APs [2]), this technique enables us to simultaneously access and model both dark and bright regions from the image, hence becomes more efficient for modeling the spatial information and reducing the feature dimension.

However, the direct exploitation of APs or SDAPs for classification task may be insufficient for a complete characterization of structural and textural information from the image, especially when regions and objects become more heterogeneous in images acquired by very high resolution (VHR) remote sensing sensors. That is why in [11], the authors proposed the histogram-based attribute profiles (HAPs) for another enhancement of APs. HAPs are built by concatenating the local histograms of attribute filter responses of each pixel. They have been proved to be more efficient and to better deal with local textures in VHR images [11]. However, two limitations of HAPs can be observed involving their very high dimensionality and their high sensitivity to the number of histogram bins (more details will be provided in the rest of the paper). Therefore in this paper, instead of constructing local histograms, our motivation is to exploit certain statistical features to characterize the local neighborhood around each pixel. Similar to HAPs, the proposed local feature-based attribute profiles (LFAPs) can provide a better description of local textures in VHR images than the standard APs. By using some simple first-order local features, LFAPs can overcome the two aforementioned drawbacks of HAPs. Furthermore, the construction of LFAPs is not limited to the use of first-order statistical features, one can take into consideration other kinds of local features to tackle more complex VHR image scenes.

Analogously, we also propose to build the LFSDAPs by extracting and combining some first-order local features from SDAPs. Then, to deal with hyperspectral image data, the extended versions of LFAPs and LFSDAPs (namely ELFAPs and ELFSDAPs, respectively) will be also derived. They are constructed by stacking all the features obtained from some first image components by using the PCA (i.e. principal component analysis) technique, as the principle of generating the extended APs (EAPs) in [3].

The remainder of this paper is organized as follows. Section II reviews some related studies involving the APs, SDAPs and HAPs. The proposed LFAP and LFSDAP techniques together with their extended versions for hyperspectral im-

ages are described in Section III. In Section IV, supervised classification results performed on the VHR Reykjavik image as well as on the hyperspectral Pavia University data yielded by the proposed methods and some reference approaches are evaluated and compared in terms of both classification accuracy and computational cost. Section V finally concludes the paper and discusses some further work.

II. Related work

A. APs and SDAPs

The definition of APs is summarized in Fig. 1(a). Let $X : \mathbb{Z}^2 \rightarrow \mathbb{R}$ be a grayscale image consisting of N pixels and $x_i = X(p_i)$ be the intensity (i.e. gray value) of the i^{th} pixel p_i. The generation of APs on X is achieved by applying a sequence of attribute thickening $\{\phi^{\lambda_\ell}\}_{\ell=1}^{L}$ and attribute thinning $\{\gamma^{\lambda_\ell}\}_{\ell=1}^{L}$ operations as follows:

$$\mathrm{AP}(X) = \Big\{X^{\phi^{\lambda_L}}, X^{\phi^{\lambda_{L-1}}}, \ldots, X^{\phi^{\lambda_1}}, X, X^{\gamma^{\lambda_1}}, \ldots, X^{\gamma^{\lambda_{L-1}}}, X^{\gamma^{\lambda_L}}\Big\}, \tag{1}$$

where $X^{\phi^{\lambda_\ell}}$ is the filtered image obtained by applying the attribute thickening ϕ^{λ} with regard to the threshold λ_ℓ. We denote $x_i^{\phi^{\lambda_\ell}} = X^{\phi^{\lambda_\ell}}(p_i)$ the gray value of the filtered image $X^{\phi^{\lambda_\ell}}$ at pixel position p_i. Similar explanation is made for $X^{\gamma^{\lambda_\ell}}$. As observed, the resulted $\mathrm{AP}(X)$ is a stack of $(2L+1)$ images including the original image, L filtered images from the thickening profiles and the other L from the thinning profiles. For more details about this AP computation, readers are referred to papers [2], [15].

It should be noted that for each pixel p_i $(i = 1, \ldots, N)$ in the definition domain of the image, the following feature vector can be considered as its AP descriptor which has been commonly used for classification task:

$$\chi^{\mathrm{AP}}(p_i) = \Big\{x_i^{\phi^{\lambda_L}}, x_i^{\phi^{\lambda_{L-1}}}, \ldots, x_i^{\phi^{\lambda_1}}, x_i, x_i^{\gamma^{\lambda_1}}, \ldots, x_i^{\gamma^{\lambda_{L-1}}}, x_i^{\gamma^{\lambda_L}}\Big\}. \tag{2}$$

Instead of calculating the APs based on both max-tree and min-tree image representation, the SDAPs were proposed in [7] based on a tree of shapes which possesses a self duality property. This tree structure allows us to simultaneously model dark and bright regions from the image content, hence providing a better simplification with regard to non-dual filtering operators (i.e. attribute thickening or thinning). For more details about attribute filters based on max-tree, min-tree as well as tree of shapes, readers are referred to paper [8]. In short, as illustrated in Fig. 1(b), the SDAPs of a grayscale image X are computed by applying the sequence of self-dual attribute filters $\{\rho^{\lambda_\ell}\}_{\ell=1}^{L}$ as follows:

$$\mathrm{SDAP}(X) = \Big\{X, X^{\rho^{\lambda_1}}, \ldots, X^{\rho^{\lambda_{L-1}}}, X^{\rho^{\lambda_L}}\Big\}. \tag{3}$$

This time, the resulted $\mathrm{SDAP}(X)$ consists of only $(L+1)$ images, thus reduces the dimensionality of the previous $\mathrm{AP}(X)$ by L. Similar to Eq. (2), the SDAP descriptor of any pixel $p_i; i = 1, \ldots, N$ can be extracted:

$$\chi^{\mathrm{SDAP}}(p_i) = \Big\{x_i, x_i^{\rho^{\lambda_1}}, \ldots, x_i^{\rho^{\lambda_{L-1}}}, x_i^{\rho^{\lambda_L}}\Big\}, \tag{4}$$

where $x_i^{\rho^{\lambda_\ell}} = X^{\rho^{\lambda_\ell}}(p_i)$ is the gray value of the filtered image $X^{\rho^{\lambda_\ell}}$ at pixel position p_i.

We note that during the construction of APs or SDAPs, different attributes can be considered to model the spatial and structural properties of regions and objects within the image content. In particularly, there are four attributes which have been commonly used in most AP-based research studies. They are *1) area* (which models the size of regions); *2) moment of inertia* (which helps to discriminate elongated objects from compact ones); *3) standard deviation* (which involves the region's homogeneity); and *4) diagonal of the region's bounding box* (which also models the region's size). In case that more than one attribute is considered, the final APs (resp. SDAPs) are formed by stacking all APs (resp. SDAPs) obtained for each single attribute.

B. HAPs

The histogram-based APs (HAPs) have been recently proposed in [11] and proved to be more efficient than APs for classification of VHR remote sensing images. By modeling the marginal local distributions of attribute filter responses, HAPs can provide a better characterization of textural information within the image content. Fig. 2 outlines the three-step generation of HAPs from a grayscale image X. As observed, the first step is to generate the standard APs (i.e. compute $\mathrm{AP}(X)$ using Eq. (1)). Then, for each sample of each APs' filtered image, the local histogram is estimated from its local patch. In the final step, all local histograms obtained at the same pixel position p_i will be concatenated to form the final HAP descriptor of that pixel as the following equation:

$$\chi^{\mathrm{HAP}}(p_i) = \Big\{h_i^{\phi^{\lambda_L}}, h_i^{\phi^{\lambda_{L-1}}}, \ldots, h_i^{\phi^{\lambda_1}}, h_i, h_i^{\gamma^{\lambda_1}}, \ldots, h_i^{\gamma^{\lambda_{L-1}}}, h_i^{\gamma^{\lambda_L}}\Big\}, \tag{5}$$

where h_i (resp. $h_i^{\phi^{\lambda_\ell}}$, $h_i^{\gamma^{\lambda_\ell}}$ with $\ell = 1, \ldots, L$) is the local histogram estimated from a local patch around p_i from the image X (resp. $X^{\phi^{\lambda_\ell}}$, $X^{\gamma^{\lambda_\ell}}$).

It should be noted that in addition to the attribute and threshold values needed for AP generation, the HAP method requires two more parameters which are dedicated to the computation of local histograms. They are the size of local patch (i.e. neighborhood) around each APs' sample (denoted by w) and the number of histogram bins (denoted by nb). The choice of w and nb involves a significant influence on the performance of HAPs in terms of discrimination capacity as well as computational cost. In fact, one can observe two drawbacks of HAPs including their very high dimensionality (which is $(2L+1) \times nb$ and leads to the problem of memory requirement and computation time) and their high sensitivity to the parameter nb. Readers are invited to the experimental study of the related paper [11] for more details. We will get back to discuss the HAPs' behavior when evaluating and

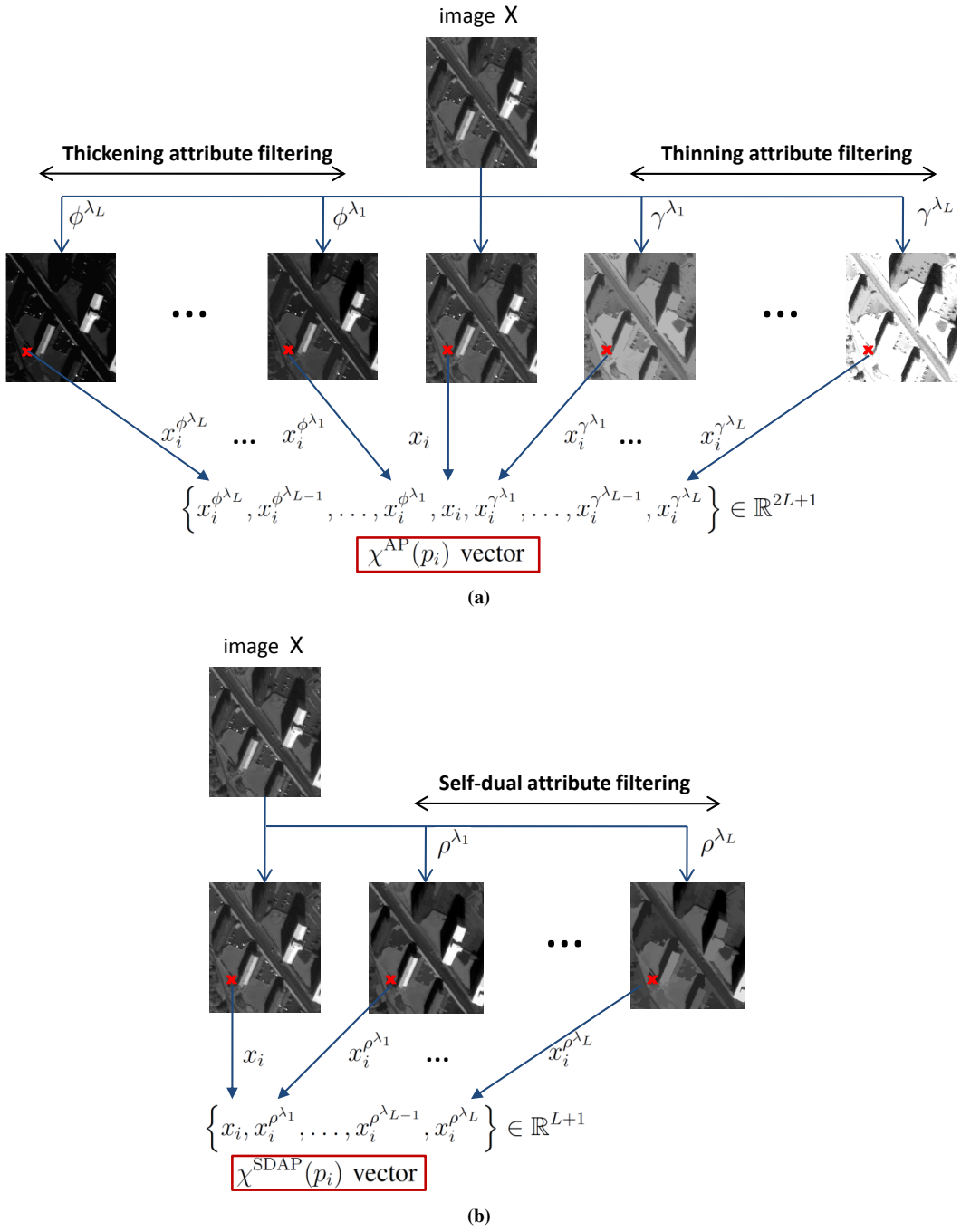


Fig. 1. Generation of **(a)** attribute profiles (APs) and **(b)** self-dual attribute profiles (SDAPs) from a grayscale image.

comparing their performance to the proposed strategy during our experimental study in Section IV.

Last but not least, although there has not been a definitive formulation of the histogram-based SDAPs (denoted by HSDAPs) from the literature yet, they can be simply deduced by applying the above three-step algorithm to the SDAPs. Therefore, similar to Eq. (5), HSDAP descriptor can be defined for each pixel $p_i; i = 1, \ldots, N$:

$$\chi^{\mathrm{HSDAP}}(p_i) = \left\{h_i, h_i^{\rho^{\lambda_1}}, \ldots, h_i^{\rho^{\lambda_L-1}}, h_i^{\rho^{\lambda_L}}\right\}. \quad (6)$$

III. Proposed Methodology

Due to the increase in spatial resolution, the appearance of geometrical and textural information in VHR remote sensing images becomes more and more significant. Hence, classification tasks should take more into account textural features when dealing with VHR image data. It should be noted that textures are generally not derived from a single image pixel, but from a local neighborhood (i.e. local patch) around it. That is the reason why the direct application of AP or SDAP feature vector in Eq. (2) or Eq. (4) for classification task may be not sufficient to account for textures in VHR images. To this end, replacing each pixel sample from the standard APs

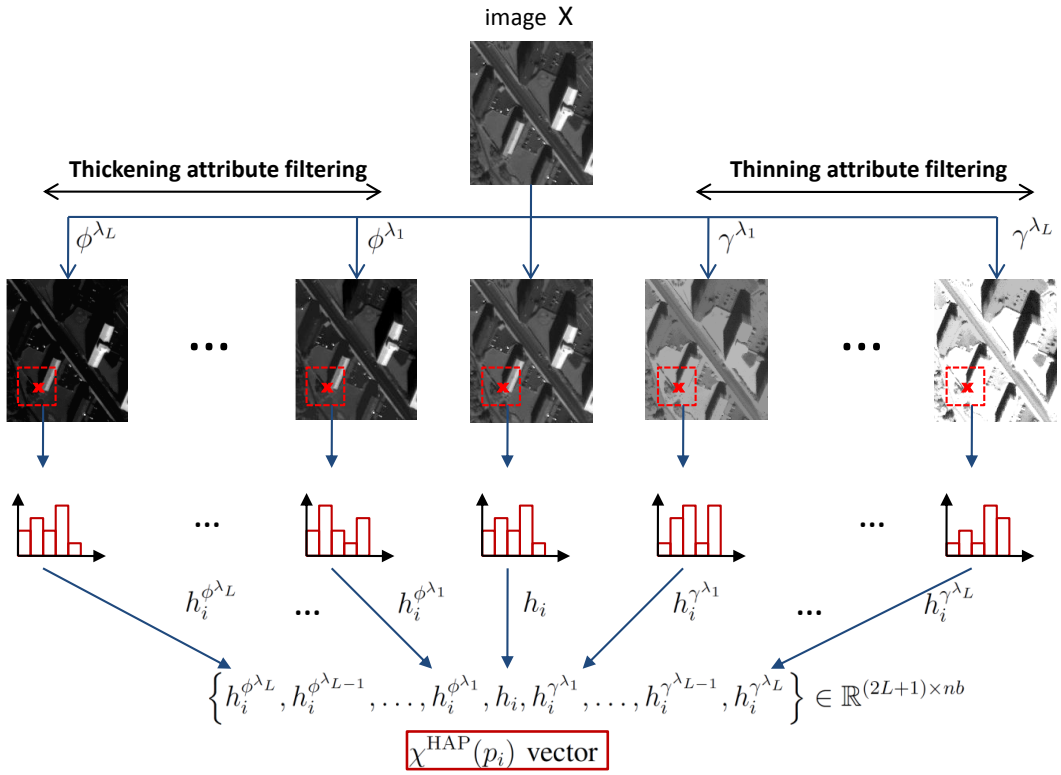


Fig. 2. Generation of histogram-based attribute profiles (HAPs) from a grayscale image.

or SDAPs with the statistical features extracted from its local neighborhood appears to be a good strategy to resolve this issue, which constitutes our motivation for this work.

In the rest of this section, we first describe the construction of the proposed LFAPs. Then, a similar principle is applied to form the LFSDAPs. Section III-C finally extends the proposed methods for hyperspectral images.

A. LFAPs

Local feature extraction has been one of the main approaches for texture analysis for decades [16]–[19]. First-order, second-order or higher order statistical features extracted from the local patch can describe the roughness, regularity, homogeneity, contrast, etc. of the related texture [16]. An important remark is that AP images mostly consist of homogeneous (i.e. flat) regions. Hence, only first-order local features such as mean, range, variance or standard deviation can be sufficient to model texture information. That is why we recommend to exploit only first-order features to construct LFAPs. Particularly in this work, we propose to use the mean (μ) and the range (r) thanks to their good representation of smooth textures, their high performance and fast computation during our experimentation.

Thus, we define the LFAP feature vector for each pixel p_i as follows:

$$\chi^{\mathrm{LFAP}}(p_i) = \left\{\chi_\mu^{\mathrm{LFAP}}(p_i), \chi_r^{\mathrm{LFAP}}(p_i)\right\} \tag{7}$$

in which

$$\chi_\mu^{\mathrm{LFAP}}(p_i) = \left\{\mu_i^{\phi^{\lambda_L}}, \mu_i^{\phi^{\lambda_L-1}}, \ldots, \mu_i^{\phi^{\lambda_1}}, \mu_i, \mu_i^{\gamma^{\lambda_1}}, \ldots, \mu_i^{\gamma^{\lambda_L-1}}, \mu_i^{\gamma^{\lambda_L}}\right\}, \tag{8}$$

$$\chi_r^{\mathrm{LFAP}}(p_i) = \left\{r_i^{\phi^{\lambda_L}}, r_i^{\phi^{\lambda_L-1}}, \ldots, r_i^{\phi^{\lambda_1}}, r_i, r_i^{\gamma^{\lambda_1}}, \ldots, r_i^{\gamma^{\lambda_L-1}}, r_i^{\gamma^{\lambda_L}}\right\}, \tag{9}$$

where μ_i and r_i (resp. $\mu_i^{\phi^{\lambda_\ell}}$ and $r_i^{\phi^{\lambda_\ell}}$) are the mean and range values extracted from the local patch $\mathcal{N}(p_i)$ of size $w \times w$ from the image X (resp. $X^{\phi^{\lambda_\ell}}$):

$$\mu_i = \frac{1}{w^2} \sum_{p_j \in \mathcal{N}(p_i)} x_j,$$

$$\mu_i^{\phi^{\lambda_\ell}} = \frac{1}{w^2} \sum_{p_j \in \mathcal{N}(p_i)} x_j^{\phi^{\lambda_\ell}},$$

$$r_i = \max_{p_j \in \mathcal{N}(p_i)}\{x_j\} - \min_{p_j \in \mathcal{N}(p_i)}\{x_j\},$$

$$r_i^{\phi^{\lambda_\ell}} = \max_{p_j \in \mathcal{N}(p_i)}\{x_j^{\phi^{\lambda_\ell}}\} - \min_{p_j \in \mathcal{N}(p_i)}\{x_j^{\phi^{\lambda_\ell}}\}.$$

Similar calculations are adopted to extract the mean and range values from every other APs' filtered image (i.e. every $X^{\phi^{\lambda_\ell}}, X^{\gamma^{\lambda_\ell}}$; $\ell = 1, \ldots, L$). The final dimension of a LFAP feature vector is $2 \times (2L + 1)$.

From Fig. 3, one can observe that the construction of LFAPs also consists of three steps (similar to that of HAPs). After generating APs, the selected local features are extracted from

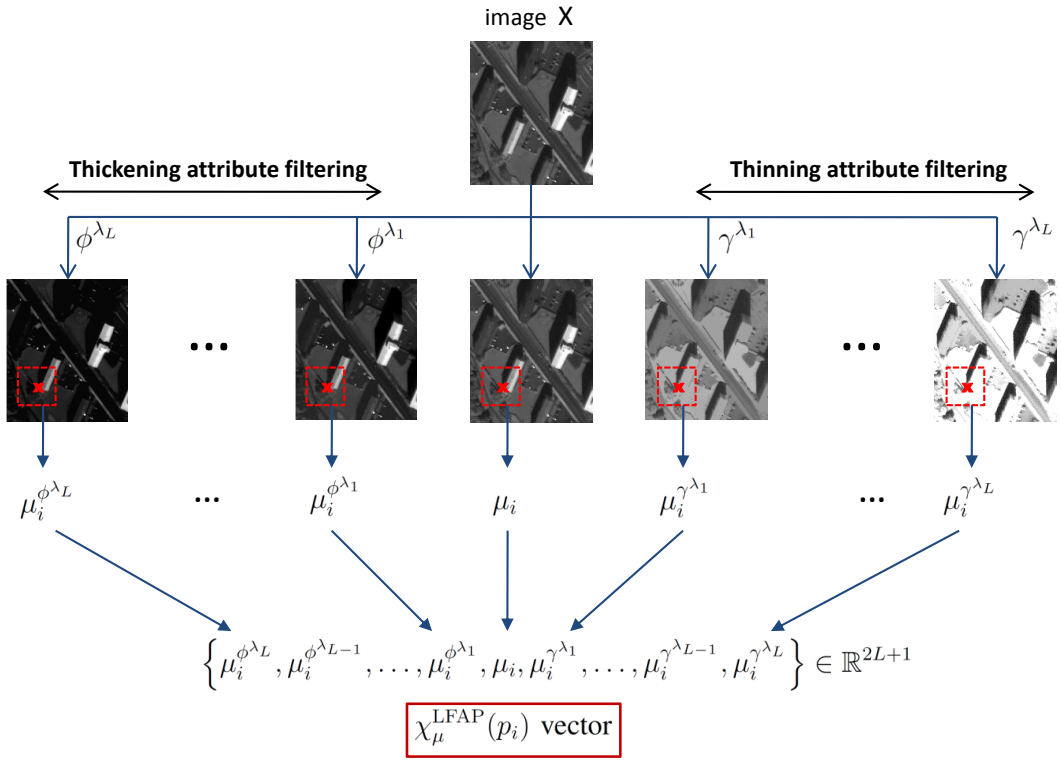


Fig. 3. Generation of local feature-based attribute profiles (LFAPs) from a grayscale image. In this example, the extracted local feature is the mean (μ) of local patch.

the local patch of each APs' sample. They are then combined to form the final LFAP feature vector. Note that in the figure, only the mean feature μ is extracted for an illustration. As defined in Eq. (7), our proposed LFAPs in this paper are built by combining the mean and the range features. However, other kinds of local features such as the Haralick features from the grey level cooccurrence matrix (GLCM) technique [20], the Gabor filter [21], wavelet transform [22], [23], pointwise features [24]–[27], local binary pattern (LBP) [28], morphological descriptors [29], [30], or covariance descriptors [31], [32], etc. can be also exploited to deal with more complex textures from VHR image data.

Another remark is that the HAPs (Section II-B) can be considered as a specific case of LFAPs since the local histogram is in fact one tool to model and characterize the statistical distribution of the pixel's local neighborhood. Thus, one can write:

$$\begin{aligned}\chi^{\mathrm{HAP}}(p_i) &= \chi_h^{\mathrm{LFAP}}(p_i) \\ &= \left\{h_i^{\phi^{\lambda_L}}, h_i^{\phi^{\lambda_L-1}}, \ldots, h_i^{\phi^{\lambda_1}}, h_i, \right. \\ &\qquad \left. h_i^{\gamma^{\lambda_1}}, \ldots, h_i^{\gamma^{\lambda_L-1}}, h_i^{\gamma^{\lambda_L}}\right\}. \end{aligned} \tag{10}$$

We remind that h_i (resp. $h_i^{\phi^{\lambda_\ell}}$, $h_i^{\gamma^{\lambda_\ell}}$ with $\ell = 1, \ldots, L$) is the local histogram estimated from the local patch $\mathcal{N}(p_i)$ from X (resp. $X^{\phi^{\lambda_\ell}}$, $X^{\gamma^{\lambda_\ell}}$).

An advantage of the proposed LFAPs compared to HAPs is that their construction does not require the parameter nb (i.e. number of histogram bins). We have mentioned the negative impact of this parameter to the performance of HAPs which causes their high dimensionality and high sensitivity to nb in Section II-B. In our work, the extraction of the mean and range features to form the LFAP descriptor in Eq. (7) still requires the patch size w. This parameter does exist in most of the local descriptors in the literature. It represents the level of exploiting the information from the neighborhood environment around each pixel. We will study the sensitivity of the proposed method to this parameter in Section IV-C.

B. LFSDAPs

As previously mentioned in Section II-A, the efficiency of SDAPs applied to remote sensing image classification compared to the original APs has been confirmed thanks to its construction via self-dual filtering operators based on a tree of shapes [7], [8]. It is motivating to apply the proposed approach to SDAPs. Hence, by extracting the mean and range features from the local patch around each SDAP's sample, the LFSDAP descriptor of the pixel p_i can be similarly defined:

$$\chi^{\mathrm{LFSDAP}}(p_i) = \left\{\chi_\mu^{\mathrm{LFSDAP}}(p_i), \chi_r^{\mathrm{LFSDAP}}(p_i)\right\}, \tag{11}$$

where

$$\chi_\mu^{\mathrm{LFSDAP}}(p_i) = \left\{\mu_i, \mu_i^{\rho^{\lambda_1}}, \ldots, \mu_i^{\rho^{\lambda_L-1}}, \mu_i^{\rho^{\lambda_L}}\right\}, \tag{12}$$

$$\chi_r^{\mathrm{LFSDAP}}(p_i) = \left\{r_i, r_i^{\rho^{\lambda_1}}, \ldots, r_i^{\rho^{\lambda_L-1}}, r_i^{\rho^{\lambda_L}}\right\}. \tag{13}$$

As a reminder, $\{\rho^{\lambda_\ell}\}_{\ell=1}^{L}$ is the sequence of self-dual attribute operators considered to generate the SDAPs as in Eq. (3). The dimension of LFSDAP feature vector equals to $2 \times (L+1)$, which is twice of the SDAP dimension.

C. Extended versions

The extended versions of APs (EAPs) and SDAPs (ESDAPs) were proposed in [3] and [9], respectively, for classification of multi-channel remote sensing data, in particular hyperspectral images. In general, a feature extraction or feature selection method is first applied to reduce the dimensionality and remove redundant information from the image. For example, in the original work of EAPs [3], the authors proposed to first apply the PCA and then compute APs from each of the first few components. Then, the final EAPs were constructed by stacking all the obtained APs as follows:

$$\mathrm{EAP} = \{\mathrm{AP}(PC_1), \mathrm{AP}(PC_2), \ldots, \mathrm{AP}(PC_K)\}, \quad (14)$$

where K is the number of first PCA images preserved for EAP construction.

We note that other techniques can be exploited to replace PCA for the feature reduction task such as the kernel PCA (KPCA) [33], independent component analysis (ICA) [34], discriminant analysis feature extraction (DAFE) [35], non-parametric weighted feature extraction (NWFE) [9], etc. A survey was conducted in [15]. Also in [14], a vector strategy based on vector-ordering relation was proposed to adapt the APs for hyperspectral data. Here in our work, the PCA is selected as the work of EAPs [3], but any adaptation or improvement using the above mentioned techniques can be undoubtedly applied. To this end, we define the extended versions of LFAPs as well as LFSDAPs by the two following equations:

$$\mathrm{ELFAP} = \{\mathrm{LFAP}(PC_1), \mathrm{LFAP}(PC_2), \ldots, \mathrm{LFAP}(PC_K)\}, \quad (15)$$

$$\mathrm{ELFSDAP} = \{\mathrm{LFSDAP}(PC_1), \mathrm{LFSDAP}(PC_2), \ldots, \mathrm{LFSDAP}(PC_K)\}, \quad (16)$$

where K is again the number of preserved principal components (PCs).

IV. Experimental Study

This section describes our experimental study to evaluate the performance of the proposed methods. Supervised classification has been carried out on both VHR panchromatic and hyperspectral image data in order to confirm the effectiveness of LFAPs and LFSDAPs as well as their extended versions. We first introduce the two data sets used in our experiments. Next, the experimental setup is described in details. We then provide the classification results yielded by the proposed algorithms compared to some reference methods. Both qualitative and quantitative assessments will be delivered in terms of classification accuracy as well as computational cost. Finally, the experimental study is completed by analyzing the parameter sensitivity of the proposed approaches.

A. Data Description

1) Reykjavik data set: The first data set is an image of size 976×640 pixels acquired by the IKONOS Earth imaging satellite in Reykjavik, Iceland. The original image consists of a VHR panchromatic image (1-m resolution) and a four-band multispectral image at lower resolution (4 m). A panshapenring process using the undecimated discrete wavelet transform method [36] was then applied in order to generate the 1-m high resolution multispectral product. In our experiments, the panchromatic (PAN) and the pansharpened multispectral (MS) images (both have 1-m resolution) were exploited. They are shown in Fig. 4(a) together with the related ground truth consisting of six thematic classes.

2) Pavia University data set: The second data set is the hyperspectral image acquired by the ROSIS airborne sensor with 1.3-m spatial resolution over the region of Pavia University, Italy. The image consists of 610×340 pixels with 103 spectral bands (from 0.43 to 0.86 μm). From the image scene, nine thematic classes were identified including trees, asphalt, bitumen, gravel, metal sheets, shadows, meadows, self-blocking bricks and bare soil. The false-color image (made by combining the bands 31, 56 and 102) and the dedicated ground truth are shown in Fig. 4(b).

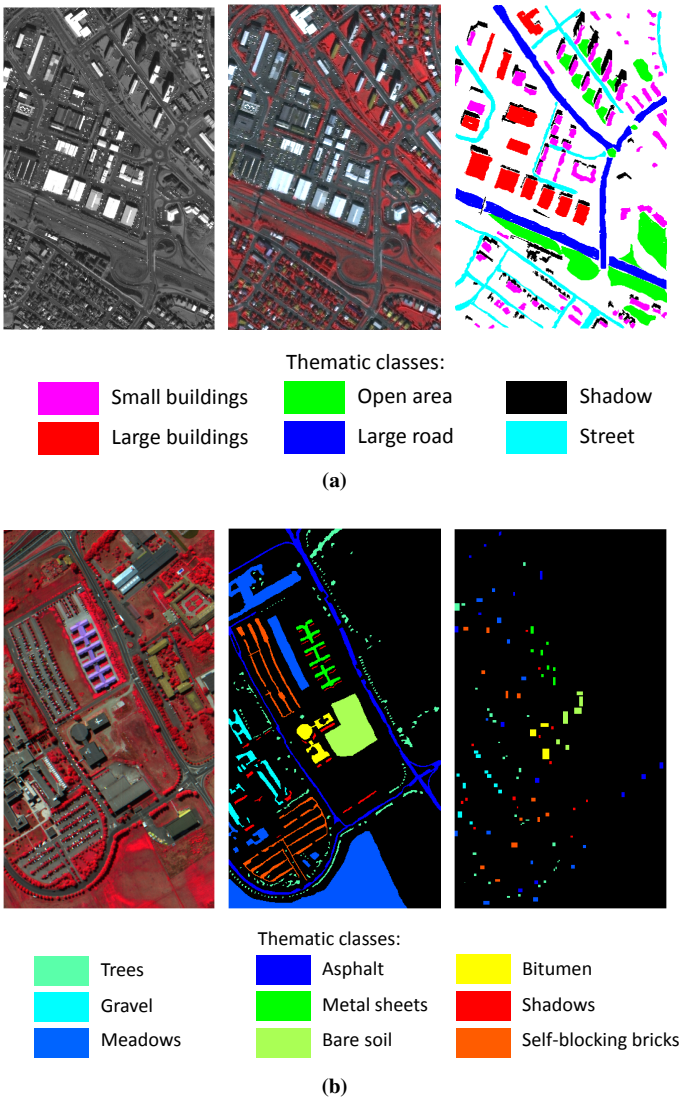


Fig. 4. Two data sets used in our experimental study. **(a)** The Reykjavik data (From left to right: panchromatic, false-color image made by bands 1-2-4 and ground truth including six thematic classes); **(b)** The hyperspectral Pavia University data (From left to right: false-color image made by bands 31-56-102, ground truth including nine thematic classes and training set).

B. Experimental Setup

For the Rekjavik data set, supervised classification results obtained by exploiting the proposed LFAP and LFSDAP features will be evaluated and compared to those yielded by using the PAN (only panchromatic band), the standard AP, SDAP, HAP and HSDAP features. The attributes and their threshold values used for generating all AP-based features are reported in Table I (second column). This setting was adopted from the recent work in paper [8] to provide an equivalent comparison. In details, three attributes including *area*, *standard deviation* and *moment of inertia* were considered and ten threshold values were set for each one. For the classification stage, Random Forest (RF) classifier [37] was employed by setting the number of trees to 200 and the number of training variables to the square root of the feature length. We randomly selected 1% of samples from each class for training and the rest for testing. Then, another experiment scenario was conducted by using 10% of training samples. All experiments were run 10 times in order to report the mean and standard deviation of classification accuracy, in terms of overall accuracy (OA), average accuracy (AA) and kappa coefficient.

For the hyperspectral Pavia University data set, we first performed the PCA on the image and the first four PCs (involving more than 99% of the total variance) were preserved for our experiments. As mentioned, this data set serves for evaluating the extended version of the proposed methods. Hence, classification results yielded by the ELFAP and ELFSDAP techniques are compared to those produced by the EAP, ESDAP as well as the extended versions of HAP and HSDAP (i.e. EHAP and EHSDAP, respectively). The attributes and their threshold values are reported in the third column of Table I. Here, four attributes including the previous three plus the *diagonal of bounding box* were considered. Four thresholds were set for each one, similar to the parameter setting from the original EAP paper [3]. Then, both RF [37] and SVM [38] were employed for supervised classification stage. As previously, the number of trees was set to 200 for the RF. For SVM, we exploited the LIBSVM [39] for implementation. We note that as recommended by the authors in [11], the histogram intersection (HI) kernel [40] was adopted for the two histogram-based approaches (i.e. EHAP and EHSDAP) while the radial basis function (RBF) kernel [41] was used for the other methods. The SVM regularization parameter C and the γ parameter of RBF kernel were estimated by a grid search method using 5-fold cross validation. We set $C = 2^{\alpha}$ with $\alpha = \{-5, -4, \ldots, 10\}$ and $\gamma = 2^{\beta}$ with $\beta = \{-10, -9, \ldots, 5\}$. The OA, AA and kappa coefficient were again reported for the evaluation and comparison of classification performance.

For both data sets, the local patch size (w) used for extracting the mean and range features (for LFAP and LFSDAP techniques) as well as for extracting the histogram signature (for HAP and HSDAP approaches) was tested from 3 to 11 with a step of 2. We also varied the number of bins (nb) from 5 to 9 for the two histogram-based reference methods in order to better investigate their performance, due to their high sensitivity to this parameter. All experiments were implemented using MATLAB on a standard personal computer with 3.4GHz CPU and 16GB RAM.

C. Results

1) Performance in terms of classification accuracy:

a) For the Reykjavik data set: Table II reports the classification results obtained by the proposed LFAPs and LFSDAPs ($w = 7$) compared to the reference methods for both two experimental scenarios (i.e. using 1% and 10% training samples). From the upper part of the table, only the PAN image was exploited to generate different feature descriptors. Moreover, in the table's lower part, we also provide the results obtained by using the PAN image to construct the descriptors and then adding the spectral information from the MS image (4 bands) into each feature vector (denoted by PAN+MS, AP+MS, HAP+MS, LFAP+MS, SDAP+MS, HSDAP+MS and LFSDAP+MS). That is why their dimensions (second column) were all increased by 4. We note that the vector length of LFAP (resp. LFSDAP) is twice of that of AP (resp. SDAP) since two local features (mean and range) were extracted to form these descriptors. Meanwhile, the dimensions of HAP and HSDAP are much higher (multiplied by a factor of nb).

From the table, we observe that the proposed LFAP technique has produced better classification performance (in terms of OA, AA and kappa) than the standard AP and the HAP (with 3 different values of nb). Analogously, LFSDAP has outperformed SDAP and HSDAP. LFSDAP has been also more efficient than LFAP, which shows a similar behavior to their standard SDAP and AP versions. In case of using 10% training samples, the best classification result on the PAN image in terms of OA was achieved by LFSDAP with 98.24%, better than 97.18% yielded by LFAP. The standard AP and SDAP have produced an OA of 91.68% and 92.53%. Therefore, LFAP and LFSDAP have improved 5.5% and 5.7% from AP and SDAP, respectively. The results of HAP and HSDAP are quite sensitive to the number of bins. HAP has attained its highest OA equal to 96.13% with $nb = 7$, while the best OA of 95.01% has been recorded by HSDAP with $nb = 9$. HAP has also outperformed the original AP (as proved in [11]). However, when using $nb = 5$, HSDAP has obtained a considerably inferior OA compared to SDAP (i.e. 85.34% compared to 92.53%). This remark again emphasizes the significant dependence of the histogram-based approaches to the number of bins, which causes one of their disadvantages compared to our local feature-based approaches. For a qualitative comparison, Fig. 5 shows the related classification maps. We observe that the thematic maps yielded by the proposed LFAP and LFSDAP [Fig. 5(d)-(h)] are more accurate than the original AP and SDAP [Fig. 5(b)-(f)] and quite close to the reference ground truth [Fig. 5(a)]. These results are smoother since the neighborhood information (from the local patch of each filtered pixel) was taken into account. The two histogram-based approaches also yielded smooth classification maps but still involved more noisy points within objects, which reduced their classification accuracy. Hence, the mean and range features used by LFAP and LFSDAP methods seem to provide better characterization of homogeneous textures within

TABLE I
ATTRIBUTES AND THRESHOLD VALUES CONSIDERED IN THE EXPERIMENTS.

Attribute	Reykjavik	Pavia University
Area	25, 100, 500, 1000, 5000, 10000, 20000, 50000, 100000, 150000	100, 500, 1000, 5000
Standard deviation	2.5, 5, 7.5, 10, 15, 20, 25, 30, 35, 40	20, 30, 40, 50
Moment of inertia	0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5 0.55, 0.6, 0.65	0.2, 0.3, 0.4, 0.5
Length of the diagonal	-	10, 25, 50, 100

TABLE II
CLASSIFICATION ACCURACY OF THE REYKJAVIK DATA SET OBTAINED BY DIFFERENT METHODS USING RF CLASSIFIER WITH 200 TREES.

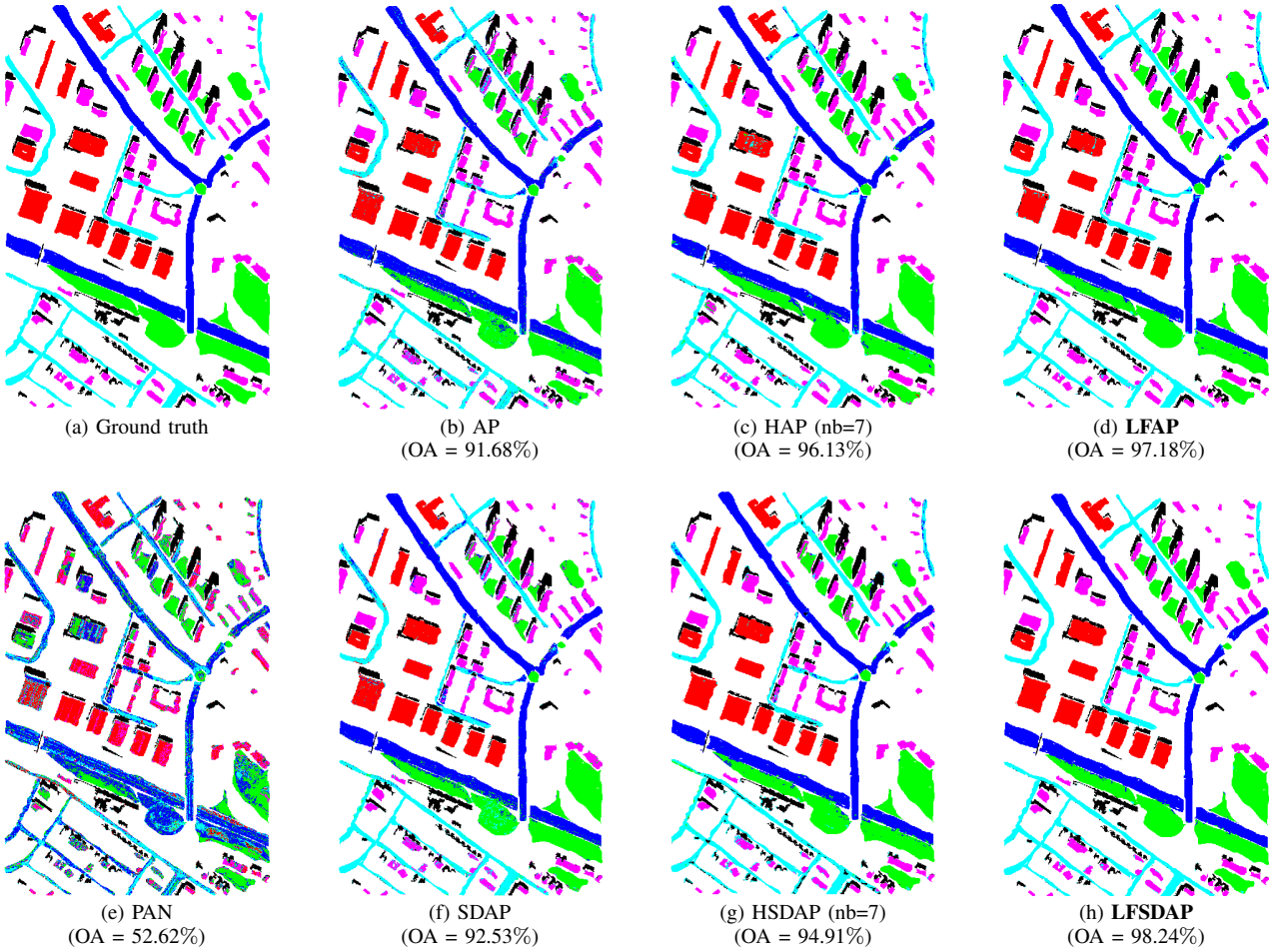
Method	Dimension	1% training samples			10% training samples		
		OA (%)	AA (%)	*Kappa*100*	OA (%)	AA (%)	*Kappa*100*
Using only panchromatic (PAN) image							
PAN	1	50.93 ± 0.44	52.61 ± 0.48	40.67 ± 0.54	52.62 ± 0.13	54.34 ± 0.12	42.74 ± 0.14
AP	63	84.81 ± 0.51	85.23 ± 0.48	81.65 ± 0.62	91.68 ± 0.12	90.84 ± 0.11	89.96 ± 0.15
HAP(nb=5)	315	84.91 ± 0.42	85.16 ± 0.43	81.77 ± 0.51	92.93 ± 0.07	93.06 ± 0.07	91.46 ± 0.09
HAP(nb=7)	441	88.25 ± 0.30	88.59 ± 0.32	85.81 ± 0.36	96.13 ± 0.07	96.28 ± 0.07	95.33 ± 0.08
HAP(nb=9)	567	87.40 ± 0.31	87.71 ± 0.31	84.79 ± 0.38	95.83 ± 0.10	95.92 ± 0.10	94.96 ± 0.12
LFAP	126	**89.44 ± 0.44**	**89.66 ± 0.44**	**87.26 ± 0.53**	**97.18 ± 0.13**	**97.19 ± 0.13**	**96.60 ± 0.16**
SDAP	33	88.12 ± 0.32	88.30 ± 0.34	85.65 ± 0.39	92.53 ± 0.12	92.63 ± 0.12	90.98 ± 0.14
HSDAP(nb=5)	165	80.29 ± 0.31	79.71 ± 0.31	76.22 ± 0.38	85.34 ± 0.11	84.59 ± 0.12	82.30 ± 0.14
HSDAP(nb=7)	231	89.96 ± 0.47	90.00 ± 0.46	87.88 ± 0.57	94.91 ± 0.13	94.84 ± 0.17	93.85 ± 0.16
HSDAP(nb=9)	297	88.22 ± 0.47	88.22 ± 0.45	85.78 ± 0.57	95.01 ± 0.13	94.94 ± 0.12	93.98 ± 0.15
LFSDAP	66	**92.68 ± 0.34**	**92.78 ± 0.32**	**91.16 ± 0.41**	**98.24 ± 0.09**	**99.21 ± 0.09**	**97.87 ± 0.11**
Using panchromatic (PAN) + multispectral (MS) images							
PAN+MS	1+4	68.98 ± 0.30	69.07 ± 0.29	62.53 ± 0.36	75.31 ± 0.09	75.50 ± 0.09	70.18 ± 0.11
AP+MS	63+4	88.10 ± 0.27	88.15 ± 0.31	85.63 ± 0.33	94.41 ± 0.09	$94.39 \pm 0.0.08$	93.25 ± 0.11
HAP(nb=5)+MS	315+4	88.24 ± 0.39	88.37 ± 0.44	85.80 ± 0.48	95.01 ± 0.13	95.06 ± 0.13	93.98 ± 0.16
HAP(nb=7)+MS	441+4	90.38 ± 0.53	90.54 ± 0.54	88.39 ± 0.65	96.92 ± 0.09	96.99 ± 0.09	96.28 ± 0.11
HAP(nb=9)+MS	567+4	89.67 ± 0.35	89.84 ± 0.34	87.53 ± 0.42	96.45 ± 0.16	96.49 ± 0.16	95.71 ± 0.19
LFAP+MS	126+4	**90.98 ± 0.53**	**91.13 ± 0.49**	**89.12 ± 0.64**	**97.36 ± 0.09**	**97.36 ± 0.07**	**96.82 ± 0.09**
SDAP+MS	33+4	90.26 ± 0.33	90.27 ± 0.33	88.24 ± 0.40	95.30 ± 0.06	95.26 ± 0.06	94.33 ± 0.07
HSDAP(nb=5)+MS	165+4	88.49 ± 0.40	88.25 ± 0.44	86.11 ± 0.48	93.85 ± 0.13	93.63 ± 0.13	92.58 ± 0.16
HSDAP(nb=7)+MS	231+4	92.18 ± 0.29	92.18 ± 0.30	90.57 ± 0.35	96.99 ± 0.05	96.96 ± 0.05	96.37 ± 0.06
HSDAP(nb=9)+MS	297+4	91.99 ± 0.33	91.98 ± 0.33	90.33 ± 0.40	97.09 ± 0.06	97.05 ± 0.06	96.48 ± 0.08
LFSDAP+MS	66+4	**93.53 ± 0.23**	**93.59 ± 0.22**	**92.19 ± 0.28**	**98.35 ± 0.04**	**98.33 ± 0.04**	**98.00 ± 0.05**

AP filtered images than the local histogram descriptors (which in fact consist of many zeros [11]). Furthermore, we will analyze later the significant benefit of the proposed descriptors compared to those histogram-based approaches in terms of computational time.

Similar remarks can be observed from the table in case of using 1% training samples, which provides no doubt inferior classification accuracy compared to the 10% training scenario. Then, by adding the complementary spectral information from the MS image, all methods have slightly improved their performance but not very significantly. For example, LFAP+MS and LFSDAP+MS have achieved their OA of 97.36% and 98.35%, which in fact provided an enhancement of 0.18% and 0.11% from the previous case using only the PAN image, respectively. In general, the behavior of all descriptors remains quite consistent when switching from 1% to 10% training samples, with or without using the MS image. That is, the local feature-based methods significantly improve the classification accuracy from the standard AP and SDAP; slightly enhance the performance from the histogram-based approaches but consisting of a fixed and fewer number of features. Further discussions about the time consumption will be provided in Section IV-C2.

b) For the Pavia University data set: The classification results obtained by the proposed ELFAP and ELFSDAP ($w = 7$) compared to reference methods including the 4 PCs (exploiting directly the first 4 PCA bands), the EAP, EHAP, ESDAP and EHSDAP are shown in Table III. The results yielded by RF and SVM classifiers can be observed from the middle and the right side of the table, respectively. Also, we enrich the experimental study by providing the performance of all descriptors in case of using *1)* only the *area attribute*; *2)* only the *moment of inertia attribute*; and *3) all four attributes*. We remind that when using SVM classifier, the HI kernel was exploited for EHAP and EHSDAP approaches, as recommended by the authors in [11], while the RBF kernel was used for the others. In fact, our experiments also showed that the HI kernel are more relevant and can produce better performance for these two histogram-based techniques than the RBF kernel.

As observed from the table, the classification results yielded by the proposed methods are more accurate compared to reference approaches. Again, ELFAP and ELFSDAP have performed a great improvement from their standard versions



(a) Ground truth
(b) AP (OA = 91.68%)
(c) HAP (nb=7) (OA = 96.13%)
(d) **LFAP** (OA = 97.18%)
(e) PAN (OA = 52.62%)
(f) SDAP (OA = 92.53%)
(g) HSDAP (nb=7) (OA = 94.91%)
(h) **LFSDAP** (OA = 98.24%)

Fig. 5. Classification maps of the Reykjavik image obtained by different methods using a RF classifier with 200 trees and 10% training samples.

(i.e. EAP and ESDAP). Compared to the histogram-based approaches (i.e. EHAP and EHSDAP), they have also been more efficient with a slight enhancement when using the RF or a more considerable enhancement when using the SVM. The highest accuracy was adopted by ELFSDAP with its OA, AA and kappa coefficient respectively equal to 95.13%, 94.57% and 93.46% yielded by RF; and 98.01%, 97.35% and 97.37% yielded by SVM, in case of considering only the *area* attribute. In fact, during our experiments, the *area* attribute provided the superior performance compared to the three other attributes. We can observe this behavior from the table by comparing the results obtained by the *area* against those obtained by the *moment of inertia*. This observation also explains why, when combining all four attributes, the performance of all descriptors has been slightly decreased. For example, OA of ELFSDAP reached 94.27% (with RF) and 97.79% (with SVM) when considering four attributes, thus 0.86% and 0.22%, respectively, lower than the case using only the *area*. That is to say, for this data set, using only the *area* attribute can provide more interesting and competitive performance than combining all the four, not only in terms of classification accuracy, but also in terms of feature dimensionality.

For qualitative analysis of classification results, Fig. 6 illustrates the classified images yielded by different descriptors in case of considering all four attributes and using RF classifier. Compared to the reference ground truth in Fig. 6(a), the result yielded by EAP [Fig. 6(b)] (OA= 89.08%) still consists of many noisy points. Other results, except the 4 PCs, are smoother thanks to the use of neighborhood information extracted from local patch (here $w = 7$). Despite its pretty high OA of 93.45%, EHSDAP [Fig. 6(g)] seems to over-smooth the self-blocking bricks (orange color) which may cause the loss of some small structures or details. Meanwhile, the results provided by the two proposed methods, i.e. ELFAP and ELFSDAP [Fig. 6(d)-(h)], are very promising by better smoothing homogeneous regions and preserving small structures. In terms of OA, an increase of 3.05% and 1.88% has been achieved compared to the standard EAP and ESDAP, respectively. In case of using SVM, this improvement is even more significant (i.e. 3.26% and 5.57%). Consequently, all experiments on this Pavia University data set confirm the effectiveness of the proposed LFAP and LFSDAP techniques in term of classification performance when extended to hyperspectral data. In the following subsection, we provide a detailed comparison in terms of computational time.

2) Performance in terms of calculation time: Tables IV and V provide the computational time required by each method to produce the classification results shown in the previous subsection. Here, we separate the time necessary for the feature extraction stage from the time required by the classification

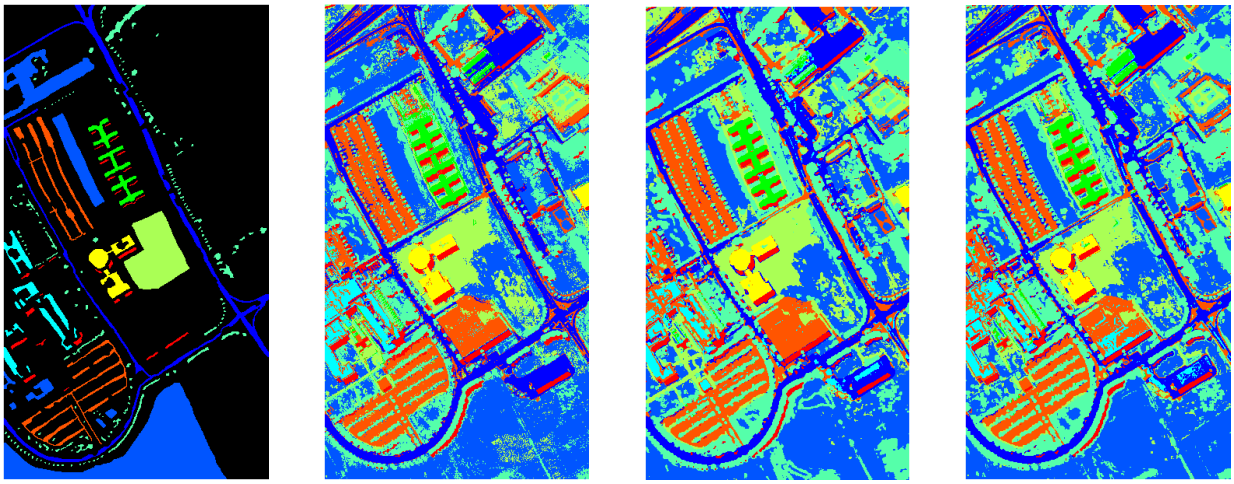
TABLE III
CLASSIFICATION ACCURACY OF THE PAVIA UNIVERSITY DATA SET OBTAINED BY DIFFERENT METHODS USING RF AND SVM CLASSIFIERS.

Method	**Dimension**	**RF classifier**			**SVM classifier**		
		OA (%)	AA (%)	*Kappa*100*	OA (%)	AA (%)	*Kappa*100*
4 PCs	4	70.70	80.63	68.89	75.79	81.33	69.01
Area attribute							
EAP	36	90.09	91.27	86.85	92.20	94.58	89.84
EHAP(nb=5)	180	90.13	89.01	86.89	89.57	89.32	86.11
EHAP(nb=7)	252	91.32	89.08	88.44	91.01	89.79	88.15
EHAP(nb=9)	324	89.84	90.94	86.66	90.76	90.89	87.56
ELFAP	72	**92.15**	**91.66**	**89.56**	**96.28**	**95.57**	**95.09**
ESDAP	20	92.45	91.59	89.92	89.69	94.18	86.80
EHSDAP(nb=5)	100	95.08	94.30	93.41	94.99	95.05	93.35
EHSDAP(nb=7)	140	93.70	92.09	91.54	93.68	93.13	91.53
EHSDAP(nb=9)	180	94.53	93.71	93.02	96.12	95.85	94.81
ELFSDAP	40	**95.13**	**94.57**	**93.46**	**98.01**	**97.35**	**97.37**
Moment of Inertia attribute							
EAP	36	80.52	89.48	75.25	92.69	93.51	90.28
EHAP(nb=5)	180	89.87	89.39	86.45	91.75	91.55	88.98
EHAP(nb=7)	252	90.42	91.02	87.23	91.68	92.57	89.02
EHAP(nb=9)	324	89.99	90.06	86.64	90.53	91.23	87.47
ELFAP	72	**91.72**	**91.82**	**88.94**	**93.15**	**93.21**	**90.84**
ESDAP	20	86.01	90.35	81.90	91.75	92.49	89.15
EHSDAP(nb=5)	100	85.19	84.50	80.82	83.32	81.96	78.28
EHSDAP(nb=7)	140	88.76	89.76	85.37	88.47	90.41	85.03
EHSDAP(nb=9)	180	88.29	89.62	84.92	87.02	90.71	83.40
ELFSDAP	40	**90.14**	**88.23**	**86.82**	**94.92**	**93.62**	**93.25**
All 4 attributes							
EAP	36×4	89.08	91.34	85.57	91.45	92.98	88.76
EHAP(nb=5)	180×4	91.77	90.90	89.03	92.79	92.45	90.35
EHAP(nb=7)	252×4	91.78	91.51	89.04	91.57	92.87	89.03
EHAP(nb=9)	324×4	90.87	91.10	87.86	90.55	91.65	87.47
ELFAP	72×4	**92.13**	**91.54**	**89.49**	**94.71**	**94.09**	**91.61**
ESDAP	20×4	92.39	92.12	89.96	92.22	95.74	89.96
EHSDAP(nb=5)	100×4	93.12	93.10	91.49	94.91	95.18	93.42
EHSDAP(nb=7)	140×4	93.98	92.11	91.95	93.91	93.50	91.88
EHSDAP(nb=9)	180×4	93.45	92.75	91.26	94.97	95.89	93.09
ELFSDAP	40×4	**94.27**	**93.64**	**92.31**	**97.79**	**96.76**	**97.07**

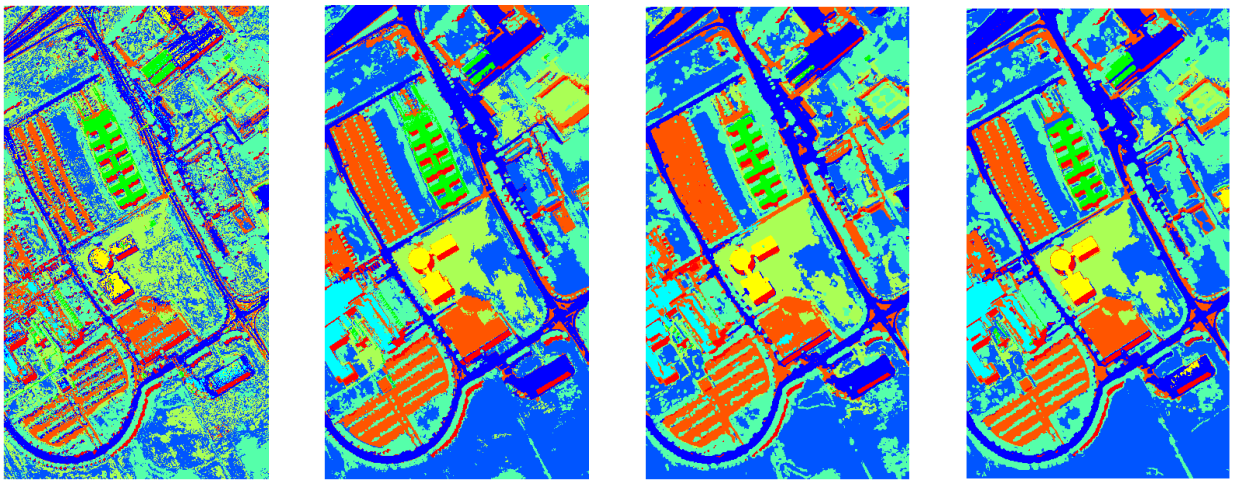
TABLE IV
COMPARISON OF FEATURE DIMENSION AND CALCULATION TIME OF DIFFERENT METHODS. EXPERIMENTS WERE CONDUCTED ON THE PANCHROMATIC REYKJAVIK IMAGE USING RF CLASSIFIER WITH 200 TREES.

Method	**Feature extraction**		**RF Classification**			
			1% training sample		10% training sample	
	Dimension	Time	Training	Testing	Training	Testing
AP	63	8.4s	2.7s	2.8s	41.0s	3.3s
HAP(nb=5)	315	3m 10.6s + 8.4s	12.1s	5.5s	3m 30.4s	6.6s
HAP(nb=7)	441	3m 14.3s + 8.4s	17.1s	5.9s	5m 01.2s	7.6s
HAP(nb=9)	567	3m 20.2s + 8.4s	22.0s	6.6s	6m 26.9s	8.3s
LFAP	126	48.7s + 8.4s	5.0s	3.4s	1m 24.7s	4.1s
SDAP	33	7.2s	1.5s	2.3s	20.7s	2.7s
HSDAP(nb=5)	165	1m 35.9s + 7.2s	6.2s	4.1s	1m 43.6s	4.6s
HSDAP(nb=7)	231	1m 37.3s + 7.2s	8.8s	4.8s	2m 33.1s	5.6s
HSDAP(nb=9)	297	1m 40.1s + 7.2s	12.2s	5.8s	3m 24.4s	6.5s
LFSDAP	66	35.3s + 7.2s	2.8s	2.5s	44.1s	2.9s

stage (which includes in fact the training and testing phases). From the two tables, the first observation is that all SDAP-based techniques require less computational time than AP-based techniques, for both feature extraction and classification stages. This behavior can be observed from several literature studies which performed a comparative study of APs and SDAPs such as [7]–[9]. In fact, by only constructing the tree of shapes instead of both min-tree and max-tree, the extraction time of SDAPs is lower than APs (7.2s compared to 8.4s in Table IV, or 10.6s compared to 17.7s in Table V). We note that these time results were obtained with the Matlab code provided by the authors of [8]. Then, since the SDAP feature dimension is lower, their corresponding classification stage is again less costly (for both training and testing phases; both RF and SVM). That is why between our two proposed methods, LFSDAP is completely more efficient than LFAP in terms



(a) Ground truth
(b) EAP (OA = 89.08%)
(c) EHAP (nb=9) (OA = 90.87%)
(d) **ELFAP** (OA = 92.13%)



(e) 4 PCs (OA = 70.70%)
(f) ESDAP (OA = 92.39%)
(g) EHSDAP (nb=9) (OA = 93.45%)
(h) **ELFSDAP** (OA = 94.27%)

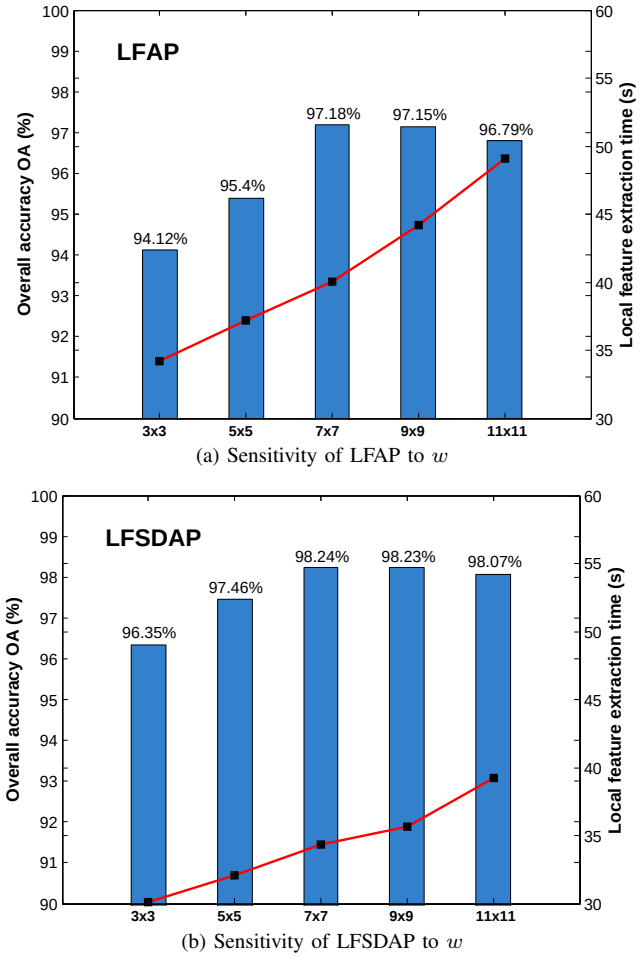
Fig. 6. Classification results of the Pavia University data obtained by different methods using a RF classifier with 200 trees.

TABLE V
COMPARISON OF FEATURE DIMENSION AND CALCULATION TIME OF DIFFERENT METHODS. EXPERIMENTS WERE CONDUCTED ON THE PAVIA UNIVERSITY DATA USING BOTH THE RF AND SVM CLASSIFIERS.

Method	**Feature extraction**		**RF Classification**		**SVM Classification**	
	Dimension	Time	Training	Testing	Training	Testing
EAP	36×4	17.7s	8.4s	3.3s	0.5s	21.3s
EHAP(nb=5)	180×4	2m 13.1s + 17.7s	46.1s	6.4s	2m 34.8s	27m 24.6s
EHAP(nb=7)	252×4	4m 01.2s + 17.7s	1m 06.5s	6.7s	2m 54.7s	33m 13.4s
EHAP(nb=9)	324×4	4m 07.1s + 17.7s	1m 27.2s	6.9s	3m 21.7s	38m 10.3s
ELFAP	72×4	1m 35.6s + 17.7s	18.3s	3.9s	0.7s	43.9s
ESDAP	20×4	10.6s	4.3s	2.3s	0.3s	11.4s
EHSDAP(nb=5)	100×4	1m 14.0s + 10.6s	21.6s	4.2s	1m 28.5s	15m 47.0s
EHSDAP(nb=7)	140×4	1m 15.1s + 10.6s	30.2s	4.7s	2m 01.3s	20m 21.9s
EHSDAP(nb=9)	180×4	1m 16.8s + 10.6s	44.3s	5.2s	2m 08.4s	22m 45.2s
ELFSDAP	40×4	11.9s + 10.6s	9.2s	3.0s	0.5s	28.0s

of computational cost, as well as in terms of classification accuracy as previously remarked.

The second observation is that, the calculation time of LFAP is higher than the standard AP but much lower than HAP. Similarly, LFSDAP requires more time than SDAP but less than HSDAP. And the same remarks can be given for their extended versions (Table V). This behavior can be easily explained by the fact that extracting the mean and range features from a pixel patch requires less time that constructing a local histogram (with a predefined number of bins). Then, the computational cost of classification stage significantly depends on the feature dimension of the exploited descriptor. From



(a) Sensitivity of LFAP to w

(b) Sensitivity of LFSDAP to w

Fig. 7. Sensitivity of the proposed methods to the patch size w in terms of overall accuracy (%) and local feature extraction time (s). Experiments were performed on the panchromatic Reykjavik image using the RF classifier with 10 runs, 200 trees and 10% training samples.

now, one can understand why the use of HAP or HSDAP may be limited due to their high dimensionality. In particular when applying the HI kernel within the SVM, a great amount of time needs to be considered for those histogram-based approaches. For example from Table V, the testing time of SVM algorithm for EHAP with nb equal to 5, 7 and 9 is more than 27, 33 and 38 minutes, respectively. Similarly, more than 15, 20 and 22 minutes of SVM testing is required by EHSDAP. By exploiting the RBF kernel, ELFAP and ELFSDAP only require 43.9s and 28s, respectively, for SVM testing phase. We also note that classification stage using RF is less costly but always confirms the behavior: AP<LFAP<HAP (resp. SDAP<LFSDAP<HSDAP). To this end, although the computation time required by LFAP (resp. LFSDAP) is higher than AP (resp. SDAP), the amount of time difference is not very significant compared to the HAP (resp. HSDAP). However, by providing competitively superior classification results, the proposed methods can be considered as an effective improvement of the standard AP or SDAP, especially more promising than the histogram-based approaches.

3) Sensitivity to parameter: This subsection finally aims at studying the sensitivity of the proposed methods to the patch size w used for extracting local features. Fig. 7 shows the performance of LFAP and LFSDAP obtained by varying w from 3 to 11. Experiments were conducted on the PAN Reykjavik image using RF classifier with 10 runs, 200 trees and 10% training samples. Here, both the classification accuracy (in terms of OA) and the feature extraction time are investigated. From the figure, we remark that LFSDAP generally provides higher OA and requires less calculation time than LFAP. However, they all perform a similar behavior when w varies. First, a quite stable performance is provided by LFSDAP with an OA from 96.35% to 98.24%, and by LFAP with an OA varying from 94.12% to 97.18%. The variation of OA involves 2 stages. When w increases from 3 to 7, OA is enhanced to reach the highest value (98.24% for LFSDAP and 97.18% for LFAP). Then, when w continues to increase from 7 to 11, OA starts to be reduced. Here is our explanation. As previously mentioned, the parameter w represents how much information from the neighborhood environment will be taken into consideration to characterize textural features at each pixel position. Within the first stage, increasing the patch size can provide more useful information to enhance the texture discrimination capacity of the two methods. Hence, OA is improved. If we continue to increase w, although more neighborhood information is exploited, we may lose the notion of local features and signal stationarity, which normally causes the over-smoothing issue. The classification accuracy is therefore decreased. In terms of feature extraction time, the more the patch size increases, the higher the computation time is required for extracting local features. However, since only simple first-order features (mean and range) are exploited, the amount of increasing time is not significant. This issue again makes the proposed strategy very effective and competitive in terms of both classification accuracy and computational cost. In conclusion, Fig. 7 shows that the proposed techniques are not very sensitive to w. A stable classification performance can be adopted by setting $w \in \{3, \ldots, 11\}$.

V. Conclusion and Perspectives

We have presented novel extensions to attribute profiles: local feature-based attribute profiles and local-feature based self-dual attribute profiles with the end of improving their pixel description capacity. Instead of using the pixels' values across progressive filtering, our approach exploits relatively simple statistical properties of local pixel neighborhoods. More specifically, we have explored the potential of two simple first-order features: the mean and the range, that have both proved to be of practical interest.

We have conducted an experimental study encompassing two widely used data sets, and compared their classification performance against known approaches. Both LFAPs and LFSDAPs have systematically outperformed their counterparts. Moreover, they have accomplished this through both shorter feature vector lengths (w.r.t. HAPs) and lower computational cost.

Besides not having to deal with the histogram bin parameter of HAPs, the presented approaches can be combined with any arbitrary local texture descriptor, hence possessing a high level

of application and dataset specific customization potential. All the same, the strategies under consideration still depend on the patch size and the particular choice of attributes; for the threshold selection of which unsupervised methods have been reported.

Future work will focus on the investigation of alternative and computationally efficient local texture descriptors, such as local binary patterns and morphological descriptors, that are expected to further improve performance. We also intend to explore the extraction of LFAPs from alternative hierarchical tree representations such as partitioning trees (alpha trees, omega trees, etc.) as well as vector strategies during the construction of their extended versions. Last but not least, a more complete study on spatial-spectral incorporation of extra spatial processing by LFAPs (and LFSDAPs) with nonlinear feature extraction techniques such as KPCA, DAFE, or NWFE for multispectral and hyperspectral image classification will be quite prospective.

VI. Acknowledgement

This work was supported by the French Agence Nationale de la Recherche (ANR) under the reference ANR-13-JS02-0005-01 (Asterix project) and the Région Bretagne grant. It was also supported by the BAGEP Award of the Science Academy and the Tubitak grant 115E857.

The authors would like to thank Prof. Jon Atli Benediktsson and Prof. Paolo Gamba for making available the Reykjavik image and the hyperspectral Pavia University data. Also, we are grateful to Dr. Mauro Dalla Mura and Dr. Gabriele Cavallaro for providing the Matlab codes of APs and SDAPs.

References

[1] L. Bruzzone and B. Demir, "A review of modern approaches to classification of remote sensing data," in *Land Use and Land Cover Mapping in Europe*. Springer, 2014, pp. 127–143.

[2] M. D. Mura, J. A. Benediktsson, B. Waske, and L. Bruzzone, "Morphological attribute profiles for the analysis of very high resolution images," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3747–3762, 2010.

[3] M. D. Mura, J. A. Benediktsson, B. Waske, and L. Bruzzone, "Extended profiles with morphological attribute filters for the analysis of hyperspectral data," *Int. J. Remote Sens.*, vol. 31, no. 22, pp. 5975–5991, 2010.

[4] B. Song, J. Li, M. Dalla Mura, P. Li, A. Plaza, J. M. Bioucas-Dias, J. A. Benediktsson, and J. Chanussot, "Remotely sensed image classification using sparse representations of morphological attribute profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 5122–5136, 2014.

[5] P. Ghamisi, J. A. Benediktsson, and J. R. Sveinsson, "Automatic spectral–spatial classification framework based on attribute profiles and supervised feature extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 9, pp. 5771–5782, 2014.

[6] X. Huang, X. Han, L. Zhang, J. Gong, W. Liao, and J. A. Benediktsson, "Generalized differential morphological profiles for remote sensing image classification," *IEEE J. Sel. Topics Appl. Earth Observat. Remote Sens.*, vol. 9, no. 4, pp. 1736–1751, 2016.

[7] M. D. Mura, J. A. Benediktsson, and L. Bruzzone, "Self-dual attribute profiles for the analysis of remote sensing images," in *Int. Symp. Math. Morpho. Appl. Sig. Image Proc.* Springer, 2011, pp. 320–330.

[8] G. Cavallaro, M. D. Mura, J. A. Benediktsson, and A. Plaza, "Remote sensing image classification using attribute filters defined over the tree of shapes," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 7, pp. 3899–3911, 2016.

[9] G. Cavallaro, M. D. Mura, J. A. Benediktsson, and L. Bruzzone, "Extended self-dual attribute profiles for the classification of hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 8, pp. 1690–1694, 2015.

[10] E. Aptoula, "Hyperspectral image classification with multidimensional attribute profiles," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 10, pp. 2031–2035, 2015.

[11] B. Demir and L. Bruzzone, "Histogram-based attribute profiles for classification of very high resolution remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 2096–2107, 2016.

[12] E. Aptoula, M. C. Ozdemir, and B. Yanikoglu, "Deep learning with attribute profiles for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 12, pp. 1970–1974, 2016.

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

[14] E. Aptoula, M. D. Mura, and S. Lefèvre, "Vector attribute profiles for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3208–3220, 2016.

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

[16] A. Materka, M. Strzelecki, *et al.*, "Texture analysis methods–a review," *Technical university of lodz, institute of electronics, COST B11 report, Brussels*, pp. 9–11, 1998.

[17] D. ping Tian *et al.*, "A review on image feature extraction and representation techniques," *Int. J. Multimedia Ubiquitous Engin.*, vol. 8, no. 4, pp. 385–396, 2013.

[18] F. Tomita and S. Tsuji, *Computer analysis of visual textures*. Springer Science & Business Media, 2013, vol. 102.

[19] J. Tang, S. Alelyani, and H. Liu, "Feature selection for classification: A review," *Data Classification: Algorithms and Applications*, p. 37, 2014.

[20] R. M. Haralick, K. Shanmugam, *et al.*, "Textural features for image classification," *IEEE Trans. Sys. Man Cybern.*, vol. 3, no. 6, pp. 610–621, 1973.

[21] A. K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using gabor filters," *Pattern Recognition*, vol. 24, no. 12, pp. 1167–1186, 1991.

[22] S. K. Meher, B. U. Shankar, and A. Ghosh, "Wavelet-feature-based classifiers for multispectral remote-sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1881–1886, 2007.

[23] L. Ruiz, A. Fdez-Sarría, and J. Recio, "Texture feature extraction for classification of remote sensing data using wavelet decomposition: a comparative study," in *20th ISPRS Congress*, vol. 35, no. part B, 2004, pp. 1109–1114.

[24] M.-T. Pham, G. Mercier, and J. Michel, "Pointwise graph-based local texture characterization for very high resolution multispectral image classification," *IEEE J. Sel. Topics Appl. Earth Observat. Remote Sens.*, vol. 8, no. 5, pp. 1962–1973, 2015.

[25] M.-T. Pham, G. Mercier, and J. Michel, "PW-COG: An effective texture descriptor for VHR satellite imagery using a pointwise approach on covariance matrix of oriented gradients," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3345–3359, 2016.

[26] M.-T. Pham, G. Mercier, and J. Michel, "Textural features from wavelets on graphs for very high resolution panchromatic Pléiades image classification," *Revue française de photogrammétrie et de télédétection*, no. 208, pp. 131–136, 2014.

[27] M.-T. Pham, G. Mercier, O. Regniers, and J. Michel, "Texture retrieval from VHR optical remote sensed images using the local extrema descriptor with application to vineyard parcel detection," *Remote Sensing*, vol. 8, no. 5, p. 368, 2016.

[28] Z. Guo, L. Zhang, and D. Zhang, "A completed modeling of local binary pattern operator for texture classification," *IEEE Trans. Image Processing*, vol. 19, no. 6, pp. 1657–1663, 2010.

[29] E. Aptoula, "Comparative study of moment based parameterization for morphological texture description," *J. Visual Comm. Image Represent.*, vol. 23, no. 8, pp. 1213–1224, 2012.

[30] E. Aptoula and S. Lefèvre, "Morphological texture description of grey-scale and color images," *Advances in imaging and electron physics*, vol. 169, p. 1, 2011.

[31] O. Tuzel, F. Porikli, and P. Meer, "Region covariance: A fast descriptor for detection and classification," in *European Conf. Comp. Vis.* Springer, 2006, pp. 589–600.

[32] M.-T. Pham, G. Mercier, and J. Michel, "Covariance-based texture description from weighted coherency matrix and gradient tensors for polarimetric SAR image classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*. IEEE, 2015, pp. 2469–2472.

[33] S. Bernabe, P. R. Marpu, A. Plaza, M. Dalla Mura, and J. A. Benediktsson, "Spectral–spatial classification of multispectral images using kernel feature space representation," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 288–292, 2014.

[34] M. Dalla Mura, A. Villa, J. A. Benediktsson, J. Chanussot, and L. Bruzzone, "Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 3, pp. 542–546, 2011.
[35] P. R. Marpu, M. Pedergnana, M. D. Mura, S. Peeters, J. A. Benediktsson, and L. Bruzzone, "Classification of hyperspectral data using extended attribute profiles based on supervised and unsupervised feature extraction techniques," *Int. J. Image Data Fusion*, vol. 3, no. 3, pp. 269–298, 2012.
[36] F. Palsson, J. R. Sveinsson, J. A. Benediktsson, and H. Aanaes, "Classification of pansharpened urban satellite images," *IEEE J. Sel. Topics Appl. Earth Observat. Remote Sens.*, vol. 5, no. 1, pp. 281–297, 2012.
[37] A. Liaw and M. Wiener, "Classification and regression by randomForest," *R news*, vol. 2, no. 3, pp. 18–22, 2002.
[38] J. A. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural processing letters*, vol. 9, no. 3, pp. 293–300, 1999.
[39] C. C. Chang and C. J. Lin, "LIBSVM: a library for support vector machines," *ACM Trans. Intell. Sys. Tech. (TIST)*, vol. 2, no. 3, p. 27, 2011.
[40] S. Maji, A. C. Berg, and J. Malik, "Classification using intersection kernel support vector machines is efficient," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. IEEE, 2008, pp. 1–8.
[41] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1351–1362, 2005.

Minh-Tan Pham (S'13-M'17) received the M.Eng. and M.Res. degrees in electronics and telecommunications from the Institute Mines-Telecom, Telecom Bretagne, France in 2013. He obtained his Ph.D. in Information and Image processing from Telecom Bretagne in collaboration with the French Space Agency (CNES) in 2016. He is now a post-doctoral researcher at the OBELIX team, IRISA laboratory, Vannes, France. His research interests include image processing, computer vision and machine learning applied to remote sensing imagery with the current focus on texture analysis, mathematical morphology, hierarchical representation and deep networks for feature extraction, object detection and classification of remote sensing data. He actually serves as reviewer of IEEE Transactions on Geosciences and Remote Sensing, IEEE Geosciences and Remote Sensing Letters and several MDPI journals including Remote Sensing, Sensors and Electronics.

Sébastien Lefèvre graduated from University of Technology of Compiègne (M.Sc. and Eng. degrees, 1999), University of Tours (PhD, 2002) and University of Strasbourg (HDR, 2009). From 2003 to 2010, he was an Associate Professor in the Department of Computer Sciences and the Image Sciences, Computer Sciences and Remote Sensing Laboratory (LSIIT), University of Strasbourg-CNRS. In 2009-2010, he was an INRIA invited scientist within the TEXMEX team of IRISA/INRIA Rennes. In 2010, he joined the University Bretagne Sud as a Full Professor in Computer Science, in the Institute of Technology of Vannes and the Institute for Research in Computer Science and Random Systems (IRISA). Within IRISA, he is leading the OBELIX team dedicated to image analysis and machine learning for remote sensing and earth observation (www.irisa.fr/obelix). He has coauthored more than 110 papers in image analysis and pattern recognition. His current research interests are mainly related to hierarchical image analysis and deep learning applied to remote sensing of environment. He was co-chair of GEOBIA 2016 and is co-chair of JURSE 2019.

Erchan Aptoula received the B.Sc. degree in computer engineering from Galatasaray University, Turkey in 2004 and the M.Sc. and Ph.D. degrees in computer science from Strasbourg University, France in 2005 and 2008 respectively. He is currently an associate professor at the Institute of Information Technologies of Gebze Technical University at Turkey, working on mathematical morphology, hyper-spectral image analysis as well as on content-based image description and retrieval.