A Novel Threshold Detection Technique for the Automatic Construction of Attribute Profiles in Hyperspectral Images

Arundhati Das , Kaushal Bhardwaj , *Member, IEEE*, Swarnajyoti Patra , *Member, IEEE*, and Lorenzo Bruzzone , *Fellow, IEEE*

Abstract—**Attribute profiles are well-acknowledged as one of the most significant techniques to characterize spectral-spatial properties of a hyperspectral image. The spectral-spatial content of an attribute profile is influenced by the threshold values considered during its construction. In this article, we propose a robust method to detect the threshold values automatically by overcoming the limitations of the existing techniques. The proposed method employs a tree structure representing the connected components of the image and evaluates attribute values at each node. Then, a total characteristic function (TCF) is defined that represents these attribute values in a nondecreasing order. The defined TCF is analyzed using a novel technique to detect a few informative thresholds for the construction of a low-dimensional attribute profile representing substantial spectral-spatial information. The proposed threshold detection method is computationally efficient. To assess the effectiveness of the proposed technique experiments are conducted on three real hyperspectral datasets using six different attributes and the results are compared to the recent state-of-the-art method. The results demonstrate that the proposed method has several advantages over the existing state-of-the-art method.**

Index Terms—**Attribute profiles (APs), hyperspectral images (HSIs), mathematical morphology (MM), random forest, spectral-spatial classification, support vector machines (SVMs), threshold detection.**

I. INTRODUCTION

HYPERSPECTRAL images (HSIs) are acquired in hundreds of contiguous channels with a very fine spectral resolution. This allows an accurate class-discrimination among the surface materials in the captured scene. Traditional methods perform the classification of HSI considering only spectral measurements, even if spatially neighboring pixels contain correlated information that could be used to decrease labeling uncertainty. Several approaches exist in the literature to incorporate

spatial information, among which we recall, Markov random field modeling [1], segmentation [2]–[4], intrinsic image decomposition [5], edge preserving filtering [6], Gabor feature extraction [7], [8], sparse representation [9]–[13], and deep convolutional neural network (CNN) [14]–[18]. Another promising and emerging approach is based on mathematical morphology (MM) [19]–[24]. A detailed analysis of the literature on different spectral-spatial techniques employed for HSI classification can be found in [25].

In the MM framework, the morphological filters are able to model different spatial characteristics of a given image [26]. However, they suffer from the limitations related to both the fixed shape of the SE and the inability to filter based on gray-level properties. These limitations are overcome by attribute filters (AFs), which process a given image based on its connected components and provide the flexibility to use any property (attribute) that can be computed for a connected component [27], [28]. The AFs merge the connected components to their background if their attribute values are smaller than a considered threshold value. By considering a sequence of threshold values, a set of filtering results are obtained and concatenated with the original image to form an attribute profile (AP). For an HSI, an extended AP (EAP) is constructed by concatenating the APs generated for each component image obtained by reducing the dimension of the HSI [20].

In the literature to incorporate sufficient spatial information, EAPs are constructed either using a large number of threshold values sampled manually from a wide range [29], [30] or using optimal number of threshold values detected automatically [31]. Only a few methods exist in the literature that detect threshold values automatically for the construction of EAPs [31]–[34]. The first approach in this direction is presented in [31], where a preliminary clustering (or classification) is performed to group the pixels into homogeneous clusters. Then, the considered attribute value is computed for each group of pixels and stored in a vector. Finally, the created vector is again clustered to select representatives of each cluster as the detected threshold values. This approach is sensitive to the preliminary clustering (or classification) results. The approaches presented in [32] and [33] are supervised in nature and concentrate on single attributes namely *standard deviation* and *area*, respectively. They perform statistical analysis on the available training samples to detect threshold values. Another interesting approach has been recently presented

Arundhati Das, Kaushal Bhardwaj, and Swarnajyoti Patra are with the Department of Computer Science and Engineering, Tezpur University, Tezpur 784028, India (e-mail: arundha@tezu.ernet.in; kauscsp@tezu.ernet.in; swpatra @tezu.ernet.in).

Lorenzo Bruzzone is with the Department of Information Engineering and Computer Science, University of Trento I-38123, Trento, Italy (e-mail: lorenzo. bruzzone@ing.unitn.it).

in [34] where the threshold values are detected automatically by exploiting the information contained in the tree representation of the considered image. In this approach, first a large number of threshold values (all possible threshold values) are obtained from the tree representation of the image. Considering each threshold value, a measure of interest is computed and stored in a vector called granulometric characteristic function (GCF). Then, the subset of threshold values which better approximates the GCF is finally selected for constructing the profile. This approach has the drawback to require a high computational time. Moreover, all the mentioned techniques detect some suboptimal thresholds that may construct APs with redundant information.

In this article, we present a novel technique that detects a set of informative threshold values automatically for the construction of APs. To this end, first, the considered image is represented with a tree structure where each node of the tree is associated with a connected component of the image. Then, a total characteristic function (TCF) is defined that stores the attribute values of the nodes in a nondecreasing order. By analyzing the TCF curves, it is possible to observe that irrespectively of the datasets and the attributes considered, initially they have a stable portion, i.e., a large number of nodes in the tree (connected components of the image) have similar attribute values. After that, they have an unstable portion where the attribute values of different nodes have sharp variations. Since the unstable portion on the TCF curve separates some nodes of the tree into two disjoint subsets, one having nodes with lower and another having nodes with higher stable attribute values, selecting appropriate attribute values as thresholds from the unstable portion on the TCF curve results in extracting significant spatial information. We propose a novel technique that automatically selects a set of suitable thresholds from the unstable portion of the TCF curve. The set of thresholds obtained by the proposed technique is used to construct the AP. The proposed approach has the following advantages.

1) It automatically detects a set of informative threshold values based on the tree representation of an image.
2) The initial few threshold values detected by the proposed technique are enough to capture relevant spatial information. As a result, it generates low-dimensional profiles that are informative (less redundant) and mitigate the curse of dimensionality problem.
3) It is computationally efficient.

Experiments are conducted on three real hyperspectral datasets considering six different attributes namely *area*, *diagonal of bounding box*, *standard deviation*, *diameter of equivalent circle*, *area of convex hull*, and *perimeter*. The experimental results validate the aforementioned advantages of the proposed method.

The rest of this article is organized as follows. The basics of APs are briefly recalled in Section II. The proposed method is presented in Section III. Section IV illustrates the datasets used in the experiments. The conducted experiments and the results are presented in Section V. Finally, conclusions are drawn in Section VI.
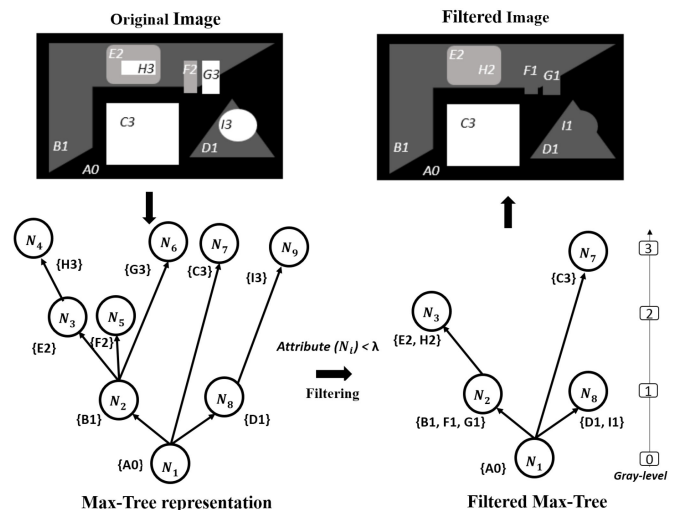

Fig. 1. Attribute filtering of a gray-scale image represented by max-tree structure.

II. BASICS OF APs

In the following subsections, the attribute filtering operations and the construction of APs are recalled.

A. Attribute Filtering

AFs are connected filters that can filter a given image by exploiting different properties of its connected components [27]. The AF operation consists of three phases. The first phase deals with representing a gray-scale image with the help of a tree structure. In this article, max-tree (min-tree) is employed for representing the image. In the next phase, the nodes of the tree structure are filtered based on a condition formulated considering an attribute and a threshold value. In the last phase, the filtered tree is transformed back to a filtered image. The following subsections give details of each phase.

1) Max-Tree (Min-Tree) Representation: A gray-scale image is represented by a tree structure for which several options are available in the literature [28], [35]–[37]. Among those the max-tree (min-tree) is widely accepted choice [20], [22], [28], [30], [38], [39]. Each node of the max-tree (min-tree) represents a connected component of the image. In the max-tree (min-tree), the pixels with minimum (maximum) gray-value are situated in the root and the pixels having maximum (minimum) gray-value are placed in leaf nodes. A gray-scale image (where the component names are followed by their gray-levels) and its corresponding max-tree representation is shown in Fig. 1.

2) Filtering of the Max-Tree (Min-Tree): In the filtering phase, the nodes of the max-tree (min-tree) are filtered by merging them with their parent nodes based on the outcome of a logical condition. The logical condition compares an attribute value evaluated for a node with a predefined threshold value λ. Suppose Attribute(N_i) is an attribute value evaluated on the ith node N_i then during filtering, the node N_i on the tree

will be merged to its parent node if it satisfies the condition $\text{Attribute}(N_i) < \lambda$.

3) Restitution of Filtered Max-Tree (Min-Tree): After filtering of the max-tree (min-tree), the filtered tree is transformed back as a filtered image. Fig. 1 shows the filtered image obtained from the filtered max-tree. For details of restitution process readers are referred to [28].

AFs, as introduced by [27], can be either attribute thinning or its dual attribute thickening. Attribute thinning uses max-tree representation and filters bright objects. Whereas, attribute thickening uses min-tree representation and filters dark objects. When the attribute is increasing, thinning (thickening) is specified as opening (closing).

B. Attribute Profile

Multiple thinning and thickening filters are applied on a gray-scale image with varying threshold values and the resultant set of output images are stacked together with the original image to form an AP. An attribute profile $\text{AP}(I)$ for a gray-scale image I is defined as

$$\text{AP}(I) = \{\phi_{\lambda_t}(I), \phi_{\lambda_{t-1}}(I), \ldots, \phi_{\lambda_1}(I), I, \gamma_{\lambda_1}(I),$$
$$\gamma_{\lambda_2}(I), \ldots, \gamma_{\lambda_t}(I)\}. \tag{1}$$

Here, the AP stores the original image I with its thickening transforms ϕ_{λ_i} and thinning transforms γ_{λ_i} where λ_i is the ith threshold value used during filtering. The threshold value λ_i varies from $i = 1, 2, \ldots, t$ resulting in t thickening and t thinning transforms. Thus, the constructed AP is of size $2t + 1$ and incorporates diverse spatial information from the original image.

C. Extended Attribute Profile

For spectral-spatial analysis of HSIs, an EAP is created. In order to overcome the curse of dimensionality problem, first, the dimension of HSI is reduced using a supervised or unsupervised dimensionality reduction technique. Principal component analysis (PCA) [40] is an unsupervised dimensionality reduction technique widely used for this purpose [19], [21], [22]. To construct an EAP, the APs constructed on the first L principal components (PCs), namely $\text{PC}_1, \text{PC}_2, \ldots, \text{PC}_L$, of the HSI are concatenated together. An EAP for HSI H is formulated as

$$\text{EAP}(H) = \{\text{AP}(\text{PC}_1), \text{AP}(\text{PC}_2), \ldots, \text{AP}(\text{PC}_L)\}. \tag{2}$$

III. PROPOSED METHOD

In this article, we propose a novel computationally efficient method to detect the suitable threshold values automatically for constructing spectral-spatial profiles. In the proposed technique, first the dimension of the HSI is reduced using PCA. Then, for each selected PC, a tree is constructed such that the nodes of the tree represent different connected components of the image. In this work, to represent different connected components of the image, the widely accepted max-tree and min-tree representations are used [20], [22], [28]. Next, for each node of the tree, the attribute value is computed and a TCF is defined to store these
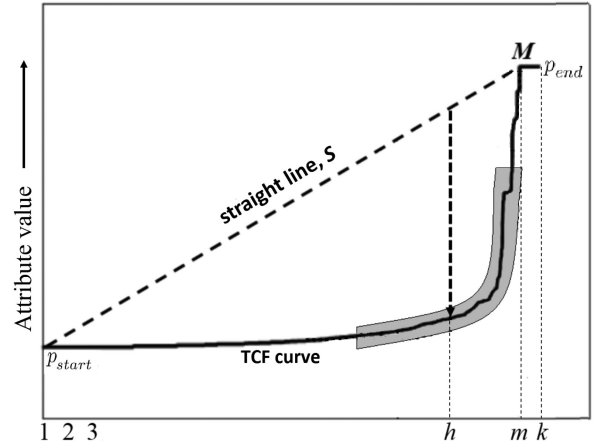

Fig. 2. Analysis of the portion of TCF curve between first index and the index with maximum gradient identified by M for computing $L_{\text{interspace}}$.

attribute values in a nondecreasing order. In detail, for a tree T_I with k nodes constructed corresponding to a component image I in the reduced dimension, the TCF is computed as follows:

$$\text{TCF}_{T_I} = \text{SORT}_{\text{asc}} \left(\{\text{Attribute}(N_i)\}_{i=1}^k\right) \tag{3}$$

where $\text{Attribute}(N_i)$ is the attribute value of the connected component represented by the ith node in the tree and TCF_{T_I} stores the attribute values of all the connected components of the image I in ascending order. Fig. 2 shows a TCF curve corresponding to the first PC of a real hyperspectral dataset considering *diagonal of bounding box* as an attribute. The attribute values are shown versus index $i = 1, 2, \ldots, k$, where k is the total number of nodes in the tree. Note that similar behavior of TCF curves is observed for different attributes as the values are in sorted order. By analyzing the TCF curve shown in Fig. 2, one can see that initially it is almost parallel to the x-axis, i.e., a large number of nodes in the tree (connected components of the image) have similar attribute values. After that, it has an unstable portion (shaded portion) where the attribute values of different nodes vary. Beyond this portion, the curve has higher attribute values that represent important connected components of the image (which need to be preserved). Since the shaded portion on the TCF curve separates some nodes of the tree into two disjoint subsets (one having nodes with lower and the other having nodes with higher attribute values), selecting diverse points as thresholds from the shaded portion on the TCF curve can incorporate significant spatial information. In this work, we propose a novel technique that automatically selects a set of suitable thresholds from the unstable portion of the TCF curve to generate a spectral-spatial profile for HSI. Our goal is to select minimum number of thresholds from the unstable portion of the TCF curve to generate a small number of filtered images that incorporate sufficient spatial information. To achieve this, we cover the TCF curve from its lowest to highest value with a straight line and detect a position among those potential threshold values where a sharp change of attribute value occurs. To this end, the proposed technique first draws a straight line that covers the unstable portion of the TCF curve. To find such line we propose to compute the gradient of all the lines passing through
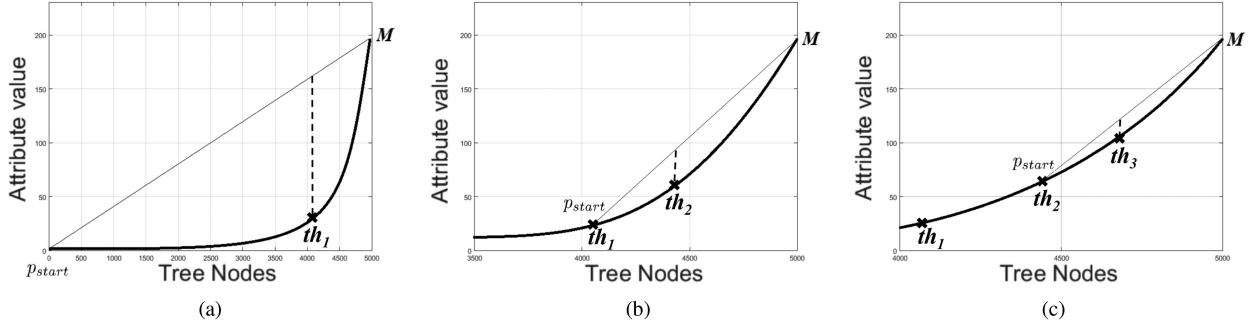
Fig. 3. Example of detection of the first three threshold values considering a sample TCF curve. (a) First detected threshold. (b) Second detected threshold. (c) Third detected threshold.

the starting point to all the other points on the TCF curve; the one which has maximum gradient value is selected. After that, the index on x-axis that corresponds to maximum interspace distance between the TCF curve and the covering straight line is chosen to compute the first suitable threshold value from the unstable portion of the TCF curve. To get the next threshold, we repeat the whole procedure by considering the point on the TCF curve associated with the chosen index as a new starting position.

In greater detail, let p_{start} be the starting point $(i_{\text{start}}, \text{TCF}(i_{\text{start}}))$ and p_{end} be the end point $(k, \text{TCF}(k))$ of the TCF curve, as shown in Fig. 2. For detecting the first threshold, the starting index on x-axis of the TCF curve is set to $i_{\text{start}} = 1$. Our technique first computes the gradient of the lines passing through the starting point p_{start} and a point $(i, \text{TCF}(i))$ for $i = i_{\text{start}} + 1, i_{\text{start}} + 2, \ldots, k$. The gradient of a line passing through p_{start} and ith point on the TCF curve (i.e., $(i, \text{TCF}(i))$) is computed as

$$\text{Gradient}(i) = \frac{\text{TCF}(i) - \text{TCF}(i_{\text{start}})}{i - i_{\text{start}}}. \quad (4)$$

After computing the gradients, if a line passing through $p_{\text{start}}(i_{\text{start}}, \text{TCF}(i_{\text{start}}))$ and $M(m, \text{TCF}(m))$ provides maximum gradient at m, i.e.,

$$m = \arg \max_{i = i_{\text{start}}+1, \ldots, k} \{\text{Gradient}(i)\} \quad (5)$$

then a straight line S is drawn between p_{start} and M to cover the unstable portion of the TCF curve. Now, to select the first suitable threshold from the unstable portion of the TCF curve, the interspace distance ($L_{\text{interspace}}$) between S and TCF is computed at each index as follows:

$$L_{\text{interspace}}(i) = S(i) - \text{TCF}(i), \quad i = i_{\text{start}}, \ldots, m \quad (6)$$

and the index h on the x-axis within the range $[i_{\text{start}}, m]$ that provides maximum interspace distance is chosen i.e.,

$$h = \arg \max_{i = i_{\text{start}}, \ldots, m} \{L_{\text{interspace}}(i)\}. \quad (7)$$

Finally, the attribute value at index h, i.e., $\text{TCF}(h)$, is selected as the first suitable threshold. For selecting the next threshold, the same procedure is repeated considering $i_{\text{start}} = h$ so that $(h, \text{TCF}(h))$ becomes the new p_{start}. Fig. 3 demonstrates the detection of first three threshold values in a synthetic TCF curve. The detection of the first threshold th_1 is shown in

---

Algorithm 1: Proposed Threshold Detection Technique.

Input: Tree T_I,
  Attribute A (*area*, *perimeter*, *etc.*),
  Required number of thresholds C.
Output: Vector (V_{th}) consisting of C threshold values.
1: Compute the attribute value at each node of the tree T_I and define the TCF_{T_I} using (3).
2: $V_{th} \leftarrow \phi$.
3: $i_{\text{start}} \leftarrow 1$ //starting index.
4: $j \leftarrow 1$ //loop counter.
5: **while** $j <= C$ **do**
6:    Find the point M (i.e.$(m, TCF(m))$) on the TCF_{T_I} that has maximum gradient from p_{start}, i.e., $(i_{\text{start}}, TCF_{T_I}(i_{\text{start}}))$, using (4).
7:    Draw a straight line S connecting p_{start} and M.
8:    Considering S and TCF_{T_I}, find the index h of TCF_{T_I} having maximum $L_{\text{interspace}}$ using (7).
9:    $th_j \leftarrow TCF_{T_I}(h)$ //detected threshold.
10:   $V_{th} \leftarrow \{V_{th}, th_j\}$.
11:   $i_{\text{start}} \leftarrow h$.
12:   $j \leftarrow j + 1$.
13: **end while**
14: Return V_{th}.

Fig. 3(a) where a straight line is drawn between p_{start} and M and the position corresponding to maximum $L_{\text{interspace}}$ is chosen as threshold th_1. For detecting the second threshold th_2, the point th_1 on TCF curve is treated as new p_{start}, as shown in Fig. 3(b). Similarly, for the third detected threshold th_3, the same procedure is repeated considering the point th_2 as new p_{start}. The steps of the proposed method are summarized in Algorithm 1.

Fig. 4 shows the overall architecture for the construction of EAP using the proposed threshold detection technique. One can see from the figure that for each image component in reduced HSI space a separate set of threshold values are detected automatically corresponding to a tree representation (max-tree and min-tree). The APs are constructed for all component images. The constructed APs are then concatenated together to form an EAP for the HSI using (2). Such constructed EAP is compact in size and effective for spectral-spatial classification of HSI.
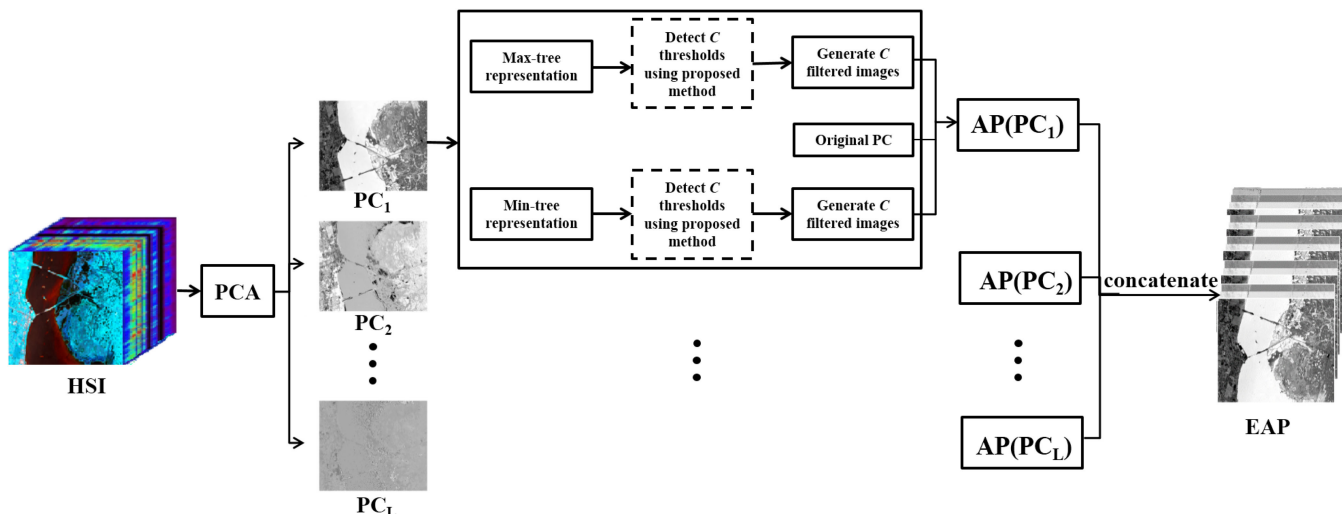
Fig. 4.   Overall architecture to construct EAP for HSI using the proposed automatic threshold detection technique.

A. Complexity Analysis

The proposed method analyzes the attribute values computed on the nodes of the tree to detect effective threshold values. Let us assume that, the image has k connected components. Accordingly, the constructed tree will possess k nodes. For constructing TCF, we need to sort the attribute values from all the k nodes. Thus, the time complexity for constructing the TCF is $\mathcal{O}(k \log k)$. Once we have the TCF, the gradient and the $L_{\text{interspace}}$ can be computed in $\mathcal{O}(k)$ each for detecting a threshold, resulting in a time complexity of $2 \times \mathcal{O}(k)$. The time complexity for detecting C threshold values is $C \times 2 \times \mathcal{O}(k)$. Therefore, the total time complexity required to detect C effective thresholds for a given image using the proposed method is $\mathcal{O}(k \log k) + 2C \times \mathcal{O}(k)$, which is equivalent to $\mathcal{O}(k \log k)$. It is worth noting that the number of connected components in an image is far less than the number of pixels in it. This shows that the proposed method is efficient in detection of threshold values.

IV. DESCRIPTION OF DATASETS

The effectiveness of the proposed technique are assessed by a set of experiments carried out on the following three real hyperspectral datasets.

ROSIS University of Pavia, Italy: The first dataset[1] is an HSI of size 610×340 pixels with 1.3 m spatial resolution. This scene was acquired by the airborne sensor ROSIS (Reflective Optics System Imaging Spectrometer) over the University of Pavia situated in the North Italian urban area of Pavia. In this dataset, there are 103 spectral bands available for use after discarding 12 bands due to irregularity in signals during transmission. Fig. 5 shows a three-band false color composite of the University of Pavia along with its ground truth having nine thematic classes.

CASI University of Houston, USA: The second dataset[2] is an HSI of size 349×1905 pixels with 2.5 m spatial resolution. The

[1]Available at: http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes
[2]Available at:http://hyperspectral.ee.uh.edu/?page_id=459
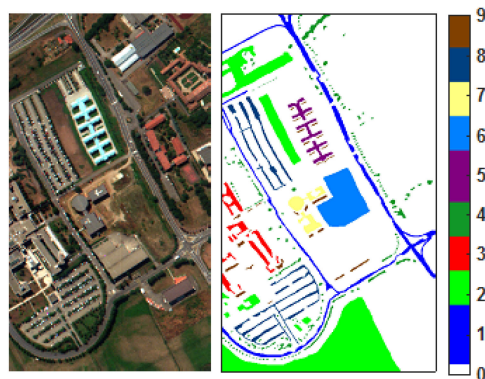
Fig. 5.   Three-band false-color composite of the University of Pavia image along with ground truth and color legend.
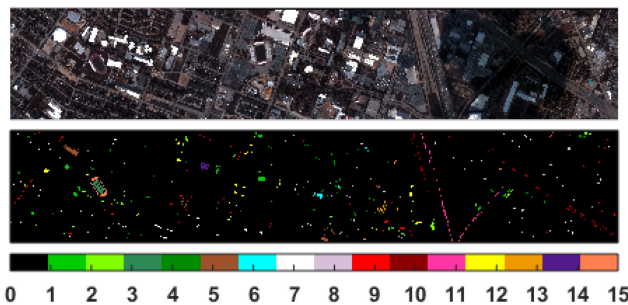
Fig. 6.   Three-band false-color composite of the University of Houston image along with ground truth and color legend.

CASI sensor acquired the data over the campus of University of Houston, Texas, United States, and its neighboring urban area with a spectral coverage ranging between 380 and 1050 nm. The acquired dataset consists of 144 spectral bands after preprocessing. Fig. 6 shows a three-band false color composite of the University of Houston data along with its ground truth having fifteen thematic classes.

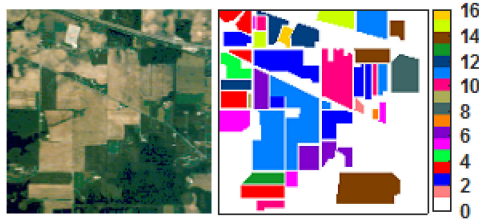Fig. 7. Three-band false-color composite of the Indian Pines image along with ground truth and color legend.
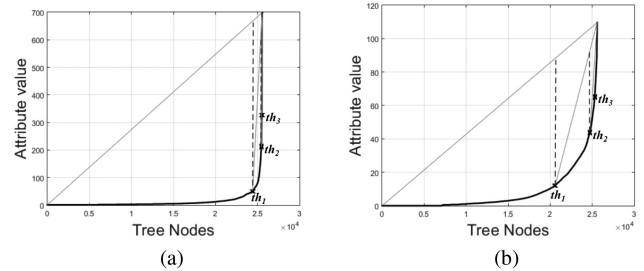

Fig. 8. First three threshold values detected by the proposed method on the TCF curves constructed for the first PC of the University of Pavia dataset considering attributes (a) Diagonal of bounding box and (b) Standard deviation.

AVIRIS Indian Pines, USA: The third dataset[1] is an HSI of 145 × 145 pixels with 20 m spatial resolution. It was acquired by the AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) sensor over the agricultural land of Indian Pines, Indiana, USA. The total number of spectral bands of the acquired scene was initially 220 out of which 200 spectral bands are available for use after preprocessing. The coverage of the sensor ranges between 400 and 2500 nm. Fig. 7 shows a three-band false color composite of the Indian Pines dataset and its related ground truth categorized into sixteen thematic classes.

V. EXPERIMENTAL RESULTS

A. Experimental Set-Up

The effectiveness of the proposed method is assessed by the experimental results obtained on the abovementioned hyperspectral datasets. The different spectral-spatial profiles are constructed based on six different attributes including *area* (A), *diagonal of bounding box* (DBB), *standard deviation* (SD), *diameter of equivalent circle* (DEC), *area of convex hull* (ACH), and *perimeter* (PER). Among these, A, DBB, DEC, and ACH are of increasing nature, whereas SD and PER are of nonincreasing nature. PCA is used for reducing the spectral dimension of HSI and spectral-spatial profiles are constructed on the first 5 PCs, which retain most of the cumulative variance in the original HSI data. The profiles constructed by the proposed method ($EAP_{proposed}$) are compared to those constructed using automatically detected threshold values by the state-of the-art method [34]. The profiles constructed by the state-of-the-art method are based on the measures *number of regions* (EAP_{NR}), *number of pixels* (EAP_{NP}), and *sum of gray-values* (EAP_{SG}). The dimension of the profiles constructed by the proposed and the state-of-the-art method are kept the same for a fair comparison. The profile sizes are 15, 25, and 35 (considering the first one, first two, and first three automatically detected thresholds, respectively). During profile construction one can adopt min, max, direct, or subtractive filtering rule. In our experiment, we have adopted the subtractive filtering as it is more robust than the other filtering rules [20], [21], [41]. The $EAP_{proposed}$ is also compared to a large spectral-spatial profile (EAP_{manual}) generated based on the following nine manually selected threshold values (the set of threshold values showing best result in trial and error) that yields a profile of size 95.

   1) A = [50 250 450 650 850 1050 1250 1450 1650].
   2) DBB = [9 16 25 36 49 64 81 100 121].
   3) SD = [1 2.7 4.1 5.5 6.9 8.3 9.7 11.1 12.5].

   4) DEC = [5 15 25 35 45 55 65 75 85].
   5) ACH = [100 200 300 600 900 1200 1500 2000 2500].
   6) PER = [30 180 330 480 630 780 930 1080 1230].

All the experiments related to implementation of the proposed technique as well as the state-of-the-art method are carried out using 64-b MATLAB (R2015a) running on a workstation with CPU Intel(R) Xeon(R) 3.60 GHz and 16 GB RAM. For all the considered datasets, the effectiveness of the proposed method is evaluated by the two sets of experiments. The first set of experiments assess the stability of the proposed technique and compares it with a state-of-the-art technique by using randomly selected training and test sets. The second set of experiments analyzes the effectiveness of the proposed technique by comparing it with several state-of-the-art spectral-spatial classification techniques by using standard/fixed training and test sets.

B. Results on the Randomly Selected Training and Test Sets

In this experiment, for all the three datasets, the proposed method is compared to the state-of-the-art method using a one-against-all support vector machine (SVM) classifier trained with 30% of the labeled data (randomly selected from each class) and considering the rest 70% for testing. The experiments are run ten times to show the stability of the proposed method and classification results are reported taking average of the quality indices, namely the overall accuracy ($\overline{OA}$), the related standard deviation ($\overline{OA}\_std$), and the kappa coefficient (kappa). The SVM classifier is implemented by using the LIBSVM library [42]. The radial basis function (RBF) kernel is adopted for the classifier. The SVM parameters $\{\sigma, C\}$ (i.e., the spread of the RBF kernel and the regularization parameter) are obtained by applying grid search with five-fold cross validation.

Figs. 8–10 show the three thresholds automatically detected by the proposed method on the basis of the TCF curves for the University of Pavia, the University of Houston, and the Indian Pines datasets, respectively. The TCF curves in the figures are constructed for the max-tree corresponding to the first PC for all the datasets considering the increasing DBB attribute and the nonincreasing SD attribute. One can observe from these figures that the first detected threshold is at a position in the beginning of the unstable portion of the TCF curves. This allows the proposed method to incorporate significant amount of spatial information in its first filtering operation by removing most of the outliers present on the image. Moreover, from these figures one can see

TABLE I
AVERAGE OVERALL ACCURACY ($\overline{OA}$), RELATED STANDARD DEVIATION (OA_STD), AND KAPPA COEFFICIENT (KAPPA) FOR PROFILES CONSTRUCTED BY THE STATE-OF-THE-ART AND THE PROPOSED METHODS CONSIDERING SIX DIFFERENT ATTRIBUTES (UNIVERSITY OF PAVIA DATA SET)

		EAP_{Manual}	EAP_{NR}			EAP_{NP}			EAP_{SG}			$EAP_{proposed}$		
Attributes		95	15	25	35	15	25	35	15	25	35	15	25	35
A	$\overline{OA}$	99.13	87.01	89.55	92.54	93.13	99.32	99.33	90.02	98.85	99.40	**98.00**	**99.74**	**99.77**
	kappa	0.988	0.824	0.860	0.901	0.908	0.991	0.991	0.866	0.985	0.992	**0.976**	**0.997**	**0.997**
	$\overline{OA}$_std	0.055	**0.072**	0.121	0.112	0.090	0.054	0.051	0.132	0.075	0.036	0.117	**0.038**	**0.032**
DBB	$\overline{OA}$	99.19	86.81	88.88	91.42	89.26	96.72	99.55	88.67	95.02	99.31	**95.64**	**99.18**	**99.72**
	kappa	0.989	0.822	0.850	0.885	0.856	0.956	0.994	0.847	0.934	0.991	**0.940**	**0.989**	**0.996**
	$\overline{OA}$_std	0.042	**0.090**	0.104	0.088	0.131	0.094	0.033	0.202	0.109	**0.028**	0.095	**0.048**	0.037
SD	$\overline{OA}$	98.04	86.88	88.98	91.64	89.29	97.29	**99.57**	88.67	96.06	99.36	**97.84**	**98.03**	98.05
	kappa	0.973	0.822	0.852	0.888	0.856	0.964	**0.994**	0.847	0.948	0.991	**0.971**	**0.973**	0.974
	$\overline{OA}$_std	0.061	0.166	0.157	0.111	**0.077**	0.082	**0.040**	0.089	**0.060**	0.056	0.083	0.061	0.064
DEC	$\overline{OA}$	99.76	86.99	89.41	92.34	93.02	99.34	99.34	90.22	98.81	99.41	**96.22**	**99.62**	**99.79**
	kappa	0.996	0.824	0.858	0.898	0.907	0.991	0.991	0.869	0.984	0.992	**0.949**	**0.996**	**0.997**
	$\overline{OA}$_std	0.036	0.093	0.099	0.133	0.081	0.040	0.056	0.115	**0.054**	0.057	**0.073**	0.055	**0.035**
ACH	$\overline{OA}$	99.10	86.85	88.89	91.49	90.47	98.44	99.57	89.95	97.71	99.32	**97.25**	**99.68**	**99.78**
	kappa	0.988	0.822	0.850	0.886	0.872	0.979	0.994	0.865	0.970	0.991	**0.964**	**0.995**	**0.997**
	$\overline{OA}$_std	0.048	0.137	0.148	0.116	0.214	0.086	0.045	0.169	0.092	0.058	**0.089**	**0.036**	**0.023**
PER	$\overline{OA}$	99.39	87.26	89.39	92.73	89.22	93.78	98.93	88.75	93.29	98.59	**97.42**	**99.70**	**99.65**
	kappa	0.992	0.828	0.857	0.903	0.855	0.917	0.986	0.848	0.911	0.981	**0.965**	**0.997**	**0.996**
	$\overline{OA}$_std	0.057	0.134	0.154	0.063	0.136	0.109	0.039	0.134	0.071	0.050	**0.063**	**0.036**	**0.030**

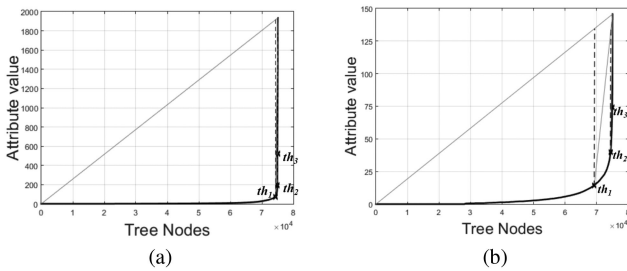The best results among the proposed and the state-of-the-art methods are highlighted in boldface.

Fig. 9.   First three threshold values detected by the proposed method on the TCF curves constructed for the first PC of the University of Houston data set considering attributes (a) Diagonal of bounding box and (b) Standard deviation.
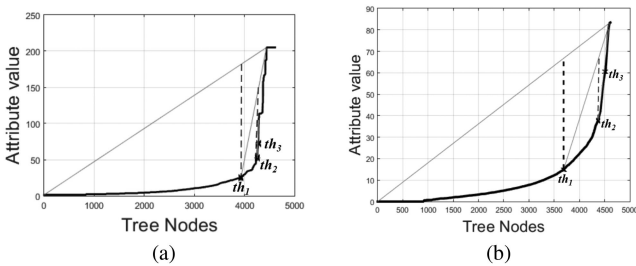

Fig. 10.   First three threshold values detected by the proposed method on the TCF curves constructed for the first PC of the Indian Pines dataset considering attributes (a) Diagonal of bounding box and (b) Standard deviation.

that the three thresholds detected by the proposed technique are diverse from each other and cover the unstable portion of the TCF curves. Thus, the proposed technique is able to filter the less significant objects by detecting few thresholds. Although the figures show the TCF curves for only two attributes, it is worth noting that similar behaviors are observed for other attributes as well. The effectiveness of the proposed technique is evident from the classification results obtained for the University of Pavia, the University of Houston, and the Indian Pines datasets,

as shown in Tables I–III, respectively. The tables present the classification results obtained for EAP_{manual}, EAP_{NR}, EAP_{NP}, EAP_{SG}, and $EAP_{proposed}$ considering the attributes A, DBB, SD, DEC, ACH, and PER. One can observe from these tables that while comparing the accuracies of the profiles with 15 features, the $EAP_{proposed}$ is outperforming the profiles constructed by the state-of-the-art method for all the attributes. In addition, the $EAP_{proposed}$ with just 15 features has similar (and sometimes better) accuracies than the EAP_{NR}, EAP_{NP}, and EAP_{SG} with 25 features. Thus, the proposed technique is robust enough to construct a low-dimensional profile that can contribute to mitigate the curse of dimensionality problem. Fig. 11 shows the filtered images obtained considering the first PC of the University of Pavia for the attributes area, standard deviation, and area of convex hull employing the first threshold automatically detected by the state-of-the-art and the proposed method. It is evident from the filtered images that the first threshold detected by the proposed method can preserve more objects of interest than the one detected by the state-of-the-art method. This again confirms that the first automatically detected threshold value by the proposed method incorporates more discriminative spatial information than that of the state-of-the-art method. Similarly, for the profiles with feature size 25 and 35, the $EAP_{proposed}$ outperforms all the three respective profiles EAP_{NR}, EAP_{NP}, and EAP_{SG}. Moreover, the results obtained by the $EAP_{proposed}$ with a small number of features are also similar or better than those obtained by the high-dimensional EAP_{manual}. This confirms the robustness of the proposed method.

C. Results on the Standard/Fixed Training and Test Sets

In this experiment, we compare the proposed technique with several recent spectral-spatial classification techniques reported in [25]. For these experiments, the same training and test sets as used in [25] are considered. These standard training and test sets are generated by the remote sensing community for

TABLE II
AVERAGE OVERALL ACCURACY ($\overline{OA}$), RELATED STANDARD DEVIATION ($\overline{OA}$_STD), AND KAPPA COEFFICIENT (KAPPA) FOR PROFILES CONSTRUCTED BY THE STATE-OF-THE-ART AND THE PROPOSED METHODS CONSIDERING SIX DIFFERENT ATTRIBUTES (UNIVERSITY OF HOUSTON DATA SET)

		EAP_{Manual}	EAP_{NR}			EAP_{NP}			EAP_{SG}			$EAP_{proposed}$		
	Attributes	95	15	25	35	15	25	35	15	25	35	15	25	35
A	$\overline{OA}$	96.28	86.58	88.73	91.54	91.37	96.92	97.57	90.33	95.35	96.83	**94.55**	**97.88**	**98.01**
	kappa	0.960	0.855	0.878	0.909	0.907	0.967	0.974	0.895	0.950	0.966	**0.940**	**0.978**	**0.978**
	$\overline{OA}$_std	0.087	0.296	0.340	0.188	0.207	0.187	0.186	**0.203**	0.270	0.283	0.341	**0.131**	**0.177**
DBB	$\overline{OA}$	95.90	86.41	88.56	90.93	89.43	93.68	97.36	88.23	91.49	96.46	**93.18**	**96.17**	**97.96**
	kappa	0.956	0.853	0.876	0.902	0.886	0.932	0.971	0.873	0.908	0.962	**0.927**	**0.960**	**0.978**
	$\overline{OA}$_std	0.172	0.381	0.181	0.260	0.302	0.339	**0.158**	0.184	0.211	0.247	**0.178**	**0.131**	0.159
SD	$\overline{OA}$	95.55	86.41	91.29	94.89	91.99	**95.06**	**95.85**	87.41	93.26	95.53	**93.30**	93.27	93.25
	kappa	0.948	0.853	0.906	0.946	0.913	**0.946**	**0.955**	0.864	0.927	0.952	**0.931**	0.924	0.928
	$\overline{OA}$_std	0.251	0.381	0.212	0.395	**0.227**	0.227	0.267	0.257	**0.158**	**0.205**	0.266	0.277	0.224
DEC	$\overline{OA}$	97.27	86.66	88.49	91.28	91.34	**96.96**	97.60	90.08	95.32	96.88	**93.33**	96.84	**98.07**
	kappa	0.967	0.856	0.875	0.906	0.906	**0.967**	0.974	0.893	0.949	0.966	**0.931**	0.964	**0.978**
	$\overline{OA}$_std	0.257	0.412	0.312	0.293	0.304	0.199	0.221	0.334	**0.128**	0.163	**0.219**	0.242	**0.115**
ACH	$\overline{OA}$	96.11	86.74	89.10	91.27	90.46	95.37	97.77	88.39	91.66	96.46	**94.65**	**97.15**	**97.88**
	kappa	0.956	0.857	0.882	0.906	0.897	0.950	0.976	0.874	0.910	0.962	**0.945**	**0.971**	**0.977**
	$\overline{OA}$_std	0.223	0.404	0.274	0.303	0.217	0.261	0.240	**0.126**	0.213	**0.186**	0.193	**0.179**	0.198
PER	$\overline{OA}$	96.38	86.84	88.80	91.34	89.98	93.63	**96.74**	89.24	93.21	96.62	**94.13**	**95.40**	96.25
	kappa	0.962	0.858	0.879	0.906	0.892	0.931	**0.965**	0.884	0.926	0.963	**0.936**	**0.949**	0.962
	$\overline{OA}$_std	0.225	0.158	**0.230**	0.423	**0.141**	0.250	0.171	0.368	0.247	**0.168**	0.161	0.252	0.216

The best results among the proposed and the state-of-the-art methods are highlighted in boldface.

TABLE III
AVERAGE OVERALL ACCURACY ($\overline{OA}$), RELATED STANDARD DEVIATION ($\overline{OA}$_STD), AND KAPPA COEFFICIENT (KAPPA) FOR PROFILES CONSTRUCTED BY THE STATE-OF-THE-ART AND THE PROPOSED METHODS CONSIDERING SIX DIFFERENT ATTRIBUTES (INDIAN PINES DATASET)

		EAP_{Manual}	EAP_{NR}			EAP_{NP}			EAP_{SG}			$EAP_{proposed}$		
	Attributes	95	15	25	35	15	25	35	15	25	35	15	25	35
A	$\overline{OA}$	96.36	80.73	83.26	88.20	85.88	**96.07**	**96.09**	85.27	95.88	96.05	**94.51**	95.48	95.80
	kappa	0.960	0.779	0.809	0.865	0.839	**0.955**	**0.955**	0.832	0.953	0.955	**0.942**	0.950	0.953
	$\overline{OA}$_std	0.195	0.309	0.417	0.312	0.306	0.334	0.175	0.295	0.216	0.246	0.233	0.243	0.230
DBB	$\overline{OA}$	96.86	80.80	82.20	87.19	82.73	93.25	**96.24**	81.78	89.20	95.92	**92.42**	**95.86**	96.12
	kappa	0.964	0.780	0.796	0.854	0.802	0.923	**0.957**	0.792	0.877	0.953	**0.912**	**0.953**	**0.957**
	$\overline{OA}$_std	0.247	0.242	0.292	0.442	0.304	0.212	0.232	0.418	0.381	0.185	0.192	0.323	0.251
SD	$\overline{OA}$	92.91	80.76	86.28	93.15	92.58	**94.96**	94.55	89.32	94.10	94.35	92.70	94.66	**95.12**
	kappa	0.920	0.780	0.843	0.922	0.915	**0.942**	0.938	0.878	0.933	0.935	0.904	0.947	**0.945**
	$\overline{OA}$_std	0.284	0.353	0.415	0.294	0.297	0.302	0.370	0.429	0.320	0.200	0.504	0.454	0.246
DEC	$\overline{OA}$	96.54	80.86	83.36	87.96	85.87	**96.15**	**96.19**	84.71	95.85	95.93	**94.19**	95.73	96.07
	kappa	0.960	0.781	0.810	0.862	0.838	**0.956**	**0.957**	0.825	0.953	0.954	**0.934**	0.948	0.960
	$\overline{OA}$_std	0.292	0.376	0.404	0.435	0.439	0.219	0.205	0.345	0.172	0.154	0.263	0.336	0.289
ACH	$\overline{OA}$	96.74	80.73	82.17	86.72	83.45	95.35	**96.30**	82.15	93.73	96.14	**94.39**	**95.49**	95.61
	kappa	0.963	0.779	0.796	0.848	0.811	0.947	**0.958**	0.796	0.928	0.956	**0.938**	**0.946**	0.947
	$\overline{OA}$_std	0.192	0.580	0.538	0.481	0.126	0.279	0.225	0.467	0.282	0.268	0.439	0.323	0.355
PER	$\overline{OA}$	96.30	80.99	85.25	92.39	83.13	89.48	**96.29**	82.19	88.21	95.85	**94.59**	95.39	95.21
	kappa	0.958	0.782	0.831	0.913	0.807	0.880	**0.958**	0.796	0.865	0.953	**0.939**	**0.942**	0.942
	$\overline{OA}$_std	0.201	0.378	0.426	0.236	0.295	0.404	0.164	0.372	0.368	0.158	0.176	0.353	0.306

The best results among the proposed and the state-of-the-art methods are highlighted in boldface.

spectral-spatial classification of the abovementioned datasets. Note that the number of training samples in these datasets is much smaller than the one used in the previous experiments. For a fair comparison, the profiles constructed by the proposed technique are classified using the same classifier as applied in [25] (i.e., random forest classifier with 200 trees). The classification results are reported in terms of overall accuracy (OA), average accuracy (AA), and kappa coefficient (kappa).

The different spectral-spatial techniques considered in these experiments include profiles constructed based on binary partition tree with $\alpha = 0.5$ (BPT $\alpha = 0.5$) [2], methods based on sparse representation (i.e., unmixing with sparse representation classifier (USRC) [9], fixed region-based model (JSRC)

[10], extended multi-AP used with sparse representation classifier (EMAP+SRC) [11], superpixel-based sparse representation classifier (SBSDM) [12], and shape adaptive sparse(SAS) [13]), and methods based on convolution neural networks (i.e., CNN [16] and PCA-CNN [15]). Table IV reports the results obtained by the EAP_proposed having 35 features considering six different attributes and those of the aforementioned techniques for University of Pavia, University of Houston, and Indian Pines datasets, respectively. One can observe from the table that in most of the cases the proposed technique considering separately each of six different attributes is able to produce better results than the other spectral-spatial classification techniques in the literature for all the three datasets. Therefore, we can conclude
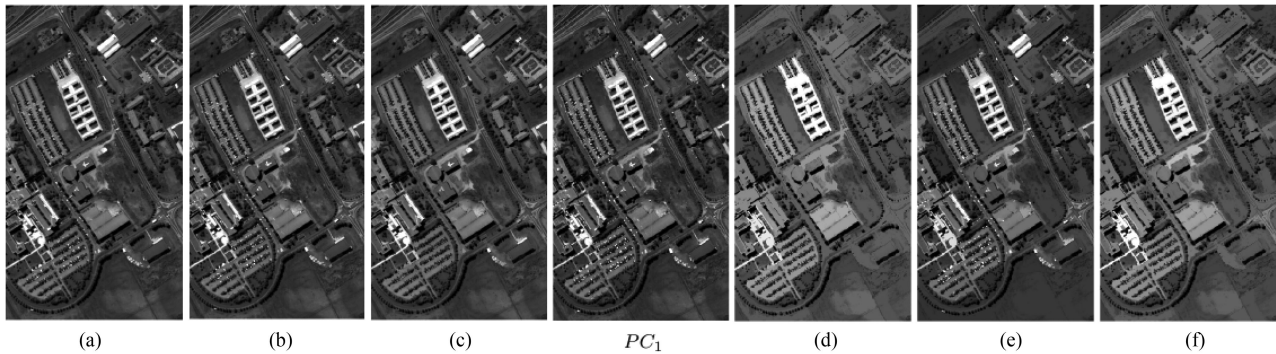
Fig. 11.   Filtered images obtained using the automatically detected first threshold value for first PC of the University of Pavia dataset by the state-of-the-art method (considering attributes (a) area, (b) standard deviation, (c) area of convex hull) and by the proposed method (considering attributes (d) area, (e) standard deviation, (f) area of convex hull).

TABLE IV
CLASSIFICATION RESULTS OBTAINED FOR PROPOSED METHOD AND DIFFERENT SPECTRAL-SPATIAL METHODS IN [25] BASED ON STANDARD TRAINING
AND TEST SETS

Datasets	Accu-racies	Proposed spectral-spatial technique						Spectral-spatial techniques in literature						
		Area	DBB	SD	DEC	ACH	PER	BPT $\alpha = 0.5$	USRC	JSRC	EMAP +SRC	SBSDM	SAS	CNN	PCA -CNN
University of Pavia	OA	93.20	91.17	87.77	**93.28**	92.98	91.59	85.74	75.05	71.78	69.49	72.00	68.95	87.01	88.93
	AA	93.43	90.64	89.60	**93.96**	93.08	92.45	**93.96**	80.88	79.3	74.2	76.63	71.48	87.15	86.37
	kappa	0.908	0.879	0.838	**0.909**	0.905	0.885	0.818	0.682	0.638	0.618	0.639	0.602	0.831	0.854
University of Houston	OA	84.29	83.24	78.64	83.95	**84.81**	84.48	83.78	70.49	76.35	71.44	75.66	75.72	82.75	83.22
	AA	86.43	85.78	81.75	86.34	**87.02**	86.77	85.34	78.35	74.66	76.56	78.26	78.08	84.04	85.61
	kappa	0.830	0.818	0.770	0.826	**0.835**	0.832	0.824	0.680	0.745	0.691	0.737	0.738	0.806	0.816
Indian Pines	OA	91.40	92.29	90.77	**92.38**	91.42	91.21	84.36	75.03	88.24	80.43	89.9	90.61	91.53	91.99
	AA	93.93	94.44	93.63	94.49	93.80	94.11	91.61	83.34	91.23	85.43	94.1	94.3	**95.24**	94.98
	kappa	0.901	0.911	0.894	**0.913**	0.901	0.899	0.822	0.717	0.866	0.778	0.885	0.893	0.901	0.906

The best results among the proposed and the state-of-the-art methods are highlighted in boldface.

that the proposed method is promising for spectral-spatial classification of HSI.

D. Analysis of Computational Time

To assess the efficiency of the proposed method the computational time (in seconds) required by both the proposed and the recent state-of-the-art method are compared as shown in Table V. The time required for constructing EAP_{NR}, EAP_{NP}, EAP_{SG}, and $EAP_{proposed}$ are referred as T_{NR}, T_{NP}, T_{SG}, and $T_{proposed}$, respectively. From the table, one can observe that for all the datasets and all the attributes, the $T_{proposed}$ is at least 15 times faster than those of T_{NR}, T_{NP}, and T_{SG}. This confirms that the proposed method is more efficient than the state-of-the-art method for constructing spectral-spatial profiles by detecting threshold values automatically. Note that, the methods initially exploit the tree representation and store the evaluated attribute values for further processing. The state-of-the-art method creates GCF after computing a measure of interest considering each attribute value as a candidate threshold, whereas the TCF stores the attribute values as it is. Furthermore, the state-of-the-art method employs regression to approximate the GCF curves. The time required for regression is proportional to the size of the GCF. In contrast, the proposed technique requires only two scans of the TCF curve to detect each threshold. For these reasons, the proposed method is significantly faster than the state-of-the-art method.

TABLE V
COMPUTATIONAL TIME (IN SECONDS) REQUIRED FOR CONSTRUCTING
SPECTRAL-SPATIAL PROFILE WITH 35 FEATURES BY THE STATE-OF-THE-ART
AND THE PROPOSED METHODS

Datasets	Attribute	T_{NR}	T_{NP}	T_{SG}	$T_{proposed}$
University of Pavia	Area	3979	4148	4156	135
	DBB	3636	2686	3646	137
	SD	795807	802349	816577	153
	DEC	4100	4143	4164	136
	ACH	5601	5587	5444	139
	PER	2508	2490	2488	132
University of Houston	Area	8066	8258	8253	289
	DBB	7525	6181	7632	295
	SD	1097565	1007517	1087365	387
	DEC	8388	8559	8499	290
	ACH	10703	10793	10740	299
	PER	4240	4321	4316	269
Indian Pines	Area	874	869	854	22
	DBB	384	377	393	22
	SD	30859	31349	31442	23
	DEC	852	842	856	22
	ACH	716	722	718	23
	PER	519	521	528	21

VI. CONCLUSION

APs are a rich source of spectral-spatial information for classification of HSIs. For an effective characterization of the spectral and spatial information present in the APs, it is important to identify optimal threshold values. In the literature, the threshold values have mostly been chosen manually. Recently

few works have addressed the issue of detecting the threshold values automatically but have limitations. The most recent state-of-the-art method overcomes these limitations but is computationally demanding and detects threshold values that may incorporate redundant information. This article has presented a novel technique for automatically detecting the threshold values to construct informative (less redundant) APs in considerably less amount of time. To this end, a tree representation is exploited where each node represents a connected component of the image and a TCF is created by considering attribute values at each node of the tree. The TCF curve is analyzed using a novel criterion to detect the threshold values automatically. The first detected threshold is able to characterize a significant portion of the image and, thus, captures remarkable contextual information on filtering. Using the detected threshold values for each component image in reduced representation of HSI, separate APs are constructed and concatenated to form an EAP, which is highly informative, class discriminative, and compact in size.

To demonstrate the effectiveness of the proposed method experiments have been conducted on three real hyperspectral datasets considering six different attributes. From the classification results it is observed that the proposed method outperforms the state-of-the-art method. The computational times required by the methods are also compared to confirm that the proposed method is significantly faster than the state-of-the-art one. Moreover, the proposed method is also compared with different spectral-spatial classification methods in the literature to show its potentiality. The proposed method can be summarized to have the following advantages:

1) It automatically detects a set of informative threshold values based on the tree representation of an image.
2) The initial few threshold values detected by the proposed technique are enough to capture sufficient spatial information. As a result, it generates low-dimensional profiles, which are informative (less redundant) and capable to mitigate the curse of dimensionality problem; and
3) It is computationally efficient.

In this article, the proposed method has been tested using a subtractive filtering rule on the considered images. Even if we do not anticipate any difference in the performance of our technique, we plan as a future work to study the performance of the proposed method over different tree representations.

REFERENCES

[1] Y. Tarabalka, M. Fauvel, J. Chanussot, and J. A. Benediktsson, "SVM-and MRF-based method for accurate classification of hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 736–740, Oct. 2010.

[2] S. Valero, P. Salembier, and J. Chanussot, "Hyperspectral image representation and processing with binary partition trees," *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1430–1443, Apr. 2012.

[3] P. Ghamisi, M. S. Couceiro, M. Fauvel, and J. A. Benediktsson, "Integration of segmentation techniques for classification of hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 342–346, Jan. 2014.

[4] S. Jia, X. Deng, J. Zhu, M. Xu, J. Zhou, and X. Jia, "Collaborative representation-based multiscale superpixel fusion for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 7770–7784, Oct. 2019.

[5] X. Kang, S. Li, L. Fang, and J. A. Benediktsson, "Intrinsic image decomposition for feature extraction of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 2241–2253, Apr. 2015.

[6] X. Kang, S. Li, and J. A. Benediktsson, "Spectral–spatial hyperspectral image classification with edge-preserving filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2666–2677, May 2014.

[7] Z. Zhu, S. Jia, S. He, Y. Sun, Z. Ji, and L. Shen, "Three-dimensional Gabor feature extraction for hyperspectral imagery classification using a memetic framework," *Inf. Sci.*, vol. 298, pp. 274–287, 2015.

[8] S. Jia, Z. Lin, B. Deng, J. Zhu, and Q. Li, "Cascade superpixel regularized gabor feature fusion for hyperspectral image classification," *IEEE Trans. Neural Netw. Learn. Syst.*, Jul. 2019, doi: 10.1109/TNNLS.2019.2921564.

[9] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2014–2039, Jun. 2011.

[10] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Hyperspectral image classification using dictionary-based sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3973–3985, Oct. 2011.

[11] B. Song *et al.*, "Remotely sensed image classification using sparse representations of morphological attribute profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 5122–5136, Aug. 2013.

[12] L. Fang, S. Li, X. Kang, and J. A. Benediktsson, "Spectral–spatial classification of hyperspectral images with a superpixel-based discriminative sparse model," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 8, pp. 4186–4201, Aug. 2015.

[13] W. Fu, S. Li, L. Fang, X. Kang, and J. A. Benediktsson, "Hyperspectral image classification via shape-adaptive joint sparse representation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 2, pp. 556–567, Feb. 2016.

[14] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep learning-based classification of hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2094–2107, Jun. 2014.

[15] J. Yue, W. Zhao, S. Mao, and H. Liu, "Spectral–spatial classification of hyperspectral images using deep convolutional neural networks," *Remote Sens. Lett.*, vol. 6, no. 6, pp. 468–477, 2015.

[16] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, "Deep feature extraction and classification of hyperspectral images based on convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6232–6251, Oct. 2016.

[17] S. Hao, W. Wang, Y. Ye, E. Li, and L. Bruzzone, "A deep network architecture for super-resolution-aided hyperspectral image classification with classwise loss," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4650–4663, Aug. 2018.

[18] S. Hao, W. Wang, Y. Ye, T. Nie, and L. Bruzzone, "Two-stream deep architecture for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 2349–2361, Apr. 2018.

[19] J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 480–491, Mar. 2005.

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

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

[22] P. Ghamisi, M. Dalla 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, May 2015.

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

[24] K. Bhardwaj, S. Patra, and L. Bruzzone, "Threshold-free attribute profile for classification of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 7731–7742, Oct. 2019.

[25] P. Ghamisi *et al.*, "New frontiers in spectral-spatial hyperspectral image classification: The latest advances based on mathematical morphology, Markov random fields, segmentation, sparse representation, and deep learning," *IEEE Geosci. Remote Sens. Mag.*, vol. 6, no. 3, pp. 10–43, Sep. 2018.

[26] J. A. Benediktsson, M. Pesaresi, and K. Amason, "Classification and feature extraction for remote sensing images from urban areas based on morphological transformations," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 9, pp. 1940–1949, Sep. 2003.

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

[28] P. Salembier, A. Oliveras, and L. Garrido, "Antiextensive connected operators for image and sequence processing," *IEEE Trans. Image Process.*, vol. 7, no. 4, pp. 555–570, Apr. 1998.

[29] M. Pedergnana, P. R. Marpu, M. Dalla Mura, J. A. Benediktsson, and L. Bruzzone, "A novel technique for optimal feature selection in attribute profiles based on genetic algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 6, pp. 3514–3528, Jun. 2013.

[30] K. Bhardwaj and S. Patra, "An unsupervised technique for optimal feature selection in attribute profiles for spectral-spatial classification of hyperspectral images," *ISPRS J. Photogrammetry Remote Sens.*, vol. 138, pp. 139–150, 2018.

[31] Z. Mahmood, G. Thoonen, and P. Scheunders, "Automatic threshold selection for morphological attribute profiles," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2012, pp. 4946–4949.

[32] P. R. Marpu, M. Pedergnana, M. Dalla Mura, J. A. Benediktsson, and L. Bruzzone, "Automatic generation of standard deviation attribute profiles for spectral–spatial classification of remote sensing data," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 2, pp. 293–297, Mar. 2013.

[33] 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, Sep. 2014.

[34] G. Cavallaro, N. Falco, M. Dalla Mura, and J. A. Benediktsson, "Automatic attribute profiles," *IEEE Trans. Image Process.*, vol. 26, no. 4, pp. 1859–1872, Apr. 2017.

[35] E. Carlinet and T. Géraud, "MToS: A tree of shapes for multivariate images," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5330–5342, Dec. 2015.

[36] M. Dalla Mura, J. A. Benediktsson, and L. Bruzzone, "Self-dual attribute profiles for the analysis of remote sensing images," in *Proc. Int. Symp. Math. Morphol. Its Appl. Signal Image Process.*, 2011, pp. 320–330.

[37] P. Salembier and L. Garrido, "Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval," *IEEE Trans. Image Process.*, vol. 9, no. 4, pp. 561–576, Apr. 2000.

[38] J. Xia, M. Dalla Mura, J. Chanussot, P. Du, and X. He, "Random subspace ensembles for hyperspectral image classification with extended morphological attribute profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 9, pp. 4768–4786, Sep. 2015.

[39] 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, Apr. 2016.

[40] I. Jolliffe, *Principal Component Analysis*. Hoboken, NJ, USA: Wiley, 2002.

[41] E. R. Urbach, J. B. Roerdink, and M. H. Wilkinson, "Connected shape-size pattern spectra for rotation and scale-invariant classification of gray-scale images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 2, pp. 272–285, Feb. 2007.

[42] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, 2011, Art. no. 27.

Arundhati Das received the B.E. and M.Tech. degrees in computer science and engineering from GIMT, Guwahati, India and NIT Agartala, Tripura, India, in 2013 and 2016, respectively. She is currently working toward the Ph.D. degree in engineering with the Department of Computer Science and Engineering, Tezpur University, Tezpur, India.

Her research interests include machine learning, mathematical morphology, and remote sensing image analysis.

Kaushal Bhardwaj (Member, IEEE) received the M.Tech degree in 2013 in information technology from Tezpur University, Tezpur, India, where he is currently working toward the Ph.D. degree in engineering with the Department of Computer Science and Engineering.

His research interests include machine learning, mathematical morphology, and remote sensing image analysis.

Mr. Bhardwaj is a Reviewer for journals, including the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE TRANSACTIONS ON IMAGE PROCESSING, and *the International Journal of Remote Sensing*.

Swarnajyoti Patra (Member, IEEE) received the B.Sc. and the M.C.A. degrees in computer science from Vidyasagar University, Midnapur, India, in 1999 and 2003, respectively, and the Ph.D. degree in engineering from Jadavpur University, Kolkata, India, in 2009.

He is currently an Assistant Professor with the Department of Computer Science and Engineering, Tezpur University, Tezpur, India. His research interests include feature extraction, feature selection, mathematical morphology, pattern recognition, and remote sensing image analysis.

Lorenzo Bruzzone (Fellow, IEEE) received the Laurea (M.S.) degree (*summa cum laude*) in electronic engineering and the Ph.D. degree in telecommunications from the University of Genoa, Genoa, Italy, in 1993 and 1998, respectively.

He is currently a Full Professor of telecommunications with the University of Trento, Trento, Italy, where he teaches remote sensing, radar, and digital communications. He is the Founder and the Director of the Remote Sensing Laboratory, Department of Information Engineering and Computer Science, University of Trento. His research interests include remote sensing, radar and SAR, signal processing, machine learning, and pattern recognition. He promotes and supervises research on these topics within the frameworks of many national and international projects.

Dr. Bruzzone was ranked the First Place in the Student Prize Paper Competition of the 1998 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Seattle, July 1998. Since 2009, has been a member of the Administrative Committee of the IEEE Geoscience and Remote Sensing Society (GRSS), where he has been the Vice President for Professional Activities since 2019. Since that, he was a recipient of many international and national honors and awards, including the recent IEEE GRSS 2015 Outstanding Service Award, the 2017 IEEE IGARSS Symposium Prize Paper Award, and the 2018 IEEE IGARSS Symposium Prize Paper Award. Since 2003, he has been the Chair of the SPIE Conference on Image and Signal Processing for Remote Sensing. He is the Principal Investigator of many research projects. Among the others, he is currently the Principal Investigator of the Radar for icy Moon exploration (RIME) instrument in the framework of the JUpiter ICy moons Explorer (JUICE) mission of the European Space Agency (ESA) and the Science Lead for the High Resolution Land Cover Project in the framework of the Climate Change Initiative of ESA. His papers are highly cited, as proven from the total number of citations (more than 32 000) and the value of the h-index (84) (source: Google Scholar). He was invited as keynote speaker in more than 32 international conferences and workshops. He was a Guest Co-Editor of many Special Issues of international journals. He is the Co-Founder of the IEEE International Workshop on the Analysis of Multi-Temporal Remote-Sensing Images (MultiTemp) series and is currently a member of the Permanent Steering Committee of this series of workshops. He is the Founder of the IEEE Geoscience and Remote Sensing Magazine for which he has been the Editor-in-Chief from 2013 to 2017. He is currently an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. He has been a Distinguished Speaker of the IEEE Geoscience and Remote Sensing Society from 2012 to 2016.