Extinction Profiles for the Classification of Remote Sensing Data

Pedram Ghamisi, *Member, IEEE*, Roberto Souza, *Student Member, IEEE*, Jon Atli Benediktsson, *Fellow, IEEE*, Xiao Xiang Zhu, *Senior Member, IEEE*, Letícia Rittner, *Member, IEEE*, and Roberto A. Lotufo, *Member, IEEE*

Abstract—With respect to recent advances in remote sensing technologies, the spatial resolution of airborne and spaceborne sensors is getting finer, which enables us to precisely analyze even small objects on the Earth. This fact has made the research area of developing efficient approaches to extract spatial and contextual information highly active. Among the existing approaches, morphological profile and attribute profile (AP) have gained great attention due to their ability to classify remote sensing data. This paper proposes a novel approach that makes it possible to precisely extract spatial and contextual information from remote sensing images. The proposed approach is based on extinction filters, which are used here for the first time in the remote sensing community. Then, the approach is carried out on two well-known high-resolution panchromatic data sets captured over Rome, Italy, and Reykjavik, Iceland. In order to prove the capabilities of the proposed approach, the obtained results are compared with the results from one of the strongest approaches in the literature, i.e., APs, using different points of view such as classification accuracies, simplification rate, and complexity analysis. Results indicate that the proposed approach can significantly outperform its alternative in terms of classification accuracies. In addition, based on our implementation, profiles can be generated in a very short processing time. It should be noted that the proposed approach is fully automatic.

Index Terms—Attribute profile (AP), extinction profile (EP), image classification, random forest (RF), remote sensing data.

I. Introduction

UNDOUBTEDLY, supervised classification plays a fundamental role in the analysis of remote sensing images, and many applications, such as crop monitoring, forest applications, urban development, mapping and tracking, and risk management, can be handled by an efficient classifier [1].

A large amount of data with different specifications (e.g., spectral, spatial, and temporal resolutions) is progressively being made available for different applications. It is now well known that the extraction of spatial information can greatly improve the classification of remote sensing images, particularly for those with high spatial resolution [1].

Conventional spectral classifiers exploit the input image as a stack of spectral measurements without considering their spatial arrangement. To further improve the performance of conventional classification approaches, one can feed spatial information to the classification system. The spatial information can provide extra information related to the shape and size of different objects, which is useful to reduce the salt-and-pepper appearance of labeled pixels and the labeling uncertainty that exits when only spectral information is taken into account [1].

Two types of approaches are commonly used for spatial information extraction: a crisp neighborhood system and an adaptive neighborhood system. While the former mostly considers spatial and contextual dependence within a predefined neighborhood system, the latter is more flexible and is not bounded within a given neighborhood system. As an example for the first group, one can consider Markov random field (MRF)-based approaches. For instance, in [2], a fully automatic spatial–spatial classifier was proposed, based on the integration of *hidden MRF* and Support Vector Machine (SVM). In [3], the result of the probabilistic SVM was regularized by an MRF. Although the spectral–spatial classifiers based on a set of crisp neighbors can provide acceptable results in terms of classification accuracies, their concept suffers from a few shortcomings, such as the following: 1) a crisp neighborhood system may not contain enough samples to characterize the specification of the object, and consequently, this downgrades the effectiveness of the classification step (in particular, when the input data set is of high resolution and the neighboring pixels are highly correlated [4]; and 2) a larger neighborhood system leads to intractable computational problems [4].

To address the shortcomings of the crisp neighborhood system, one can consider spectral–spatial classification approaches based on adaptive neighborhood approaches. Among those approaches, there are a considerable number of contributions in the literature on the use of segmentation approaches such as [5]–[7]. Another possible set of adaptive approaches used to extract spatial information is based on the concept of morphological profiles (MPs). MPs are composed of a number of features constructed by applying a set of openings and closings by reconstruction with a structuring element (SE) of an increasing size. Pesaresi and Benediktsson [8] considered

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

2 IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING

using morphological transformations to build an MP. There are a considerable number of works in the literature in which MPs have been taken into consideration. For example, in [9], an MP was used to assess the damage caused by the earthquake in Bam on a Quickbird panchromatic image. In [10], an automatic hierarchical segmentation approach was proposed based on the analysis of the derivative of the MP (DMP). The MP was also investigated in [11] by extracting a fuzzy measure of the characteristic scale and contrast of each structure in the image. The computed measures were compared with the possibility distribution predefined for each thematic class, generating a value of membership degree for each class used for classification. In [12], a binary optimization approach was proposed, based on fractional-order Darwinian particle swarm optimization [13], in order to select the most informative features produced by MP.

Based on the aforementioned examples, it can be seen that multiscale processing based on MPs has proven to be effective in extracting informative spatial features from images [1], [14]. However, MPs have a few limitations, for example, the shape of SEs is fixed, and SEs are only able to extract information related to the size of existing objects and are unable to characterize information related to the gray-level characteristics of the regions.

To overcome these shortcomings, the morphological attribute profile (AP) was introduced in [15] as a generalization of the MP, which provides a multilevel characterization of an image by using the sequential application of morphological attribute filters (AFs). AFs analyze an input image by considering only its connected components based on a connectivity rule. Although the AP has been recently introduced, there are a significant number of contributions based on it for the analysis of different types of data. For example, AFs were taken into account in modeling the structural information of the scene to increase the efficiency of the classification and building extraction in [15] and [16], respectively. Those approaches proved to be efficient methods of modeling structural information in very high resolution images. As discussed in [1], [14], and [15], APs are a more flexible tool than MPs, since APs can process images based on many different types of attributes. In fact, the attributes can be of any type. For example, they can be purely geometric, or related to the spectral values of the pixels, or based on different characteristics, such as spatial relations to other connected components. In [17] and [18], APs have been investigated for the integration of optical and LiDAR data. In [19] and [20], an automatic method was proposed for the classification of hyperspectral data by considering only two attributes (area and standard deviation) that are able to accurately classify hyperspectral images in an acceptable CPU processing time. In addition, the information provided by such approaches (i.e., MPs and APs) can be extremely redundant and demands an efficient form of representation in order to avoid increasing the dimensionality of the problem, which can downgrade the subsequent classifier performance due to the curse of dimensionality. To address such an issue, in [21], a sparse classification using both spectral and spatial information was investigated based on extended multivariate AP (EMAP). On the other hand, in [22], the performance of different feature extraction approaches, including linear, nonlinear, and manifold

approaches, has been investigated to generate base images to construct EMAPs. In a few works such as [1], [14], [22], and [23], the efficiency of different classifiers has been investigated for the classification of features produced by EMAP. For instance, a few classifiers, such as kernel SVMs, can handle high-dimensional data due to their lower sensitivity to the imbalance between dimensionality and the number of training samples. However, the construction of the aforementioned filtering approaches leads to a very sparse feature space. In this context, kernel SVMs may complicate the problem and lead to overfitting of the classification model. A comprehensive survey on APs and their capabilities for the classification of remote sensing data can be found in [1] and [14].

Extinction filters (EFs) are extrema-oriented connected filters, which, unlike AFs, preserve the height of the extrema kept. Souza *et al.* [24] showed through experiments that EFs are better than AFs with respect to simplification for recognition. EFs are based on the concept of extinction values proposed by Vachier [25]. They are a generalization of the dynamics concept [26], which is equivalent to the height extinction value, to any increasing attribute. Informally speaking, extinction values measure the persistence of the extrema in the image. The main contribution of this paper is to propose a fast, accurate, and automatic approach for the extraction of spatial and contextual information from remote sensing data. In more detail, our main contributions are listed in the following with respect to their importance.

1) Above all, this paper proposes a novel approach, i.e., the extinction profile (EP), for information extraction and classification. The proposed approach is based on EFs. This approach simultaneously discards unimportant spatial details and preserves the geometrical characteristics of the other regions.

2) The most common attributes in the literature are area, standard deviation, diagonal of the bounding box, and moment of inertia. In this paper, we have taken into account a few other attributes such as volume and height and investigate their usefulness for the classification of remote sensing data.

3) In addition, the proposed approach is extrema oriented, in contrast to APs, which are threshold oriented, making it less sensitive to image resolution. In other words, the proposed approach is data set distribution independent. This contribution makes the proposed approach fully automatic for all different attributes. It should be noted that, although there are a few recent research works trying to make the APs automatic [19], [20], [27], its conventional concept is highly dependent on the threshold values of different attributes. These values need to be set manually, and they can be changed from one data set to another one. However, the proposed approach can, by nature, solve this issue.

4) The derivative of EP (DEP) is proposed by differentiating the features produced by EP.

5) The proposed approach is applied to two well-known panchromatic data sets. We believe that APs are among the strongest approaches in the literature with regard

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

GHAMISI *et al.*: EXTINCTION PROFILES FOR THE CLASSIFICATION OF REMOTE SENSING DATA

3

TABLE I
LIST OF NOTATIONS AND ACRONYMS

Notations and Acronyms	Definition	Notations and Acronyms	Definition	Notations and Acronyms	Definition
f and g	Input images	RF	Random forest	MP	Morphological profile
ψ	A transformation	AP	Attribute profile	DMP	Derivative of MP
$\mathbf{x}$	Pixel position	$x_{i,min}, x_{i,max} y_{i,min}, y_{i,max}$	Bounding box coordinates	μ_h	height attribute
μ_v	Volume attribute	μ_{STD}	Standard deviation attribute	μ_a	Area attribute
$\mu_{gray-level}$	Mean gray-level	EP	Extinction Profile	ϕ	Thickening
MAP	Multi-AP	EF	Extinction filter	γ	Thinning
MEP	Multi-EP	s	Number of thresholds	C_i	Connected component i
Π_ϕ	Thickening profile	Δ_ϕ	Derivative of thickening profile	SVM	Support vector machine
Π_γ	Thinning profile	Δ_γ	Derivative of thinning profile	MRF	Markov random field

to extracting and modeling spatial information. To this end, we only compared the proposed approach with APs from different perspectives, such as classification accuracies, analysis of complexities, simplification rate, and recognition power. Results indicate that EP is a powerful approach with respect to simplification for recognition, since it preserves more regions and correspondences found by affine region detectors. In addition, the proposed method can significantly outperform APs in terms of classification accuracies, and it is fully automatic.

It is important to note that the concept proposed in this paper is being used for the first time in the remote sensing community. In summary, in this paper, we tried to prove the capability of the proposed approach through several experiments. The approach is capable of providing very high classification accuracy swiftly and automatically. In this context, the proposed approach outperforms one of the most well-known approaches in the literature, i.e., AP, in terms of classification accuracy and automation capability within the same CPU processing time.

The rest of this paper is organized as follows. Section II introduces the methodology of this paper. Section III is devoted to experimental results. The main concluding remarks are provided in Section IV.

II. METHODOLOGY

The classification framework is composed of two steps: spatial information extraction using EPs and a classification step. For the classification step, a random forest (RF) classifier is taken into account since it can efficiently handle the high redundancy existing in features produced by EPs (as well as APs). RF is a classifier composed of a collection of tree-like classifiers. Ideally, an RF classifier should be an independent and identically distributed randomization of weak learners. RF uses a lot of individual decision trees, all of which are trained (grown) to tackle the same problem. A sample is labeled as the most frequently occurring of the classes as determined by the individual trees [1]. The following subsections elaborate on the proposed EPs and some primary backgrounds.

A. Notations and Acronyms

To make this paper easier to follow, Table I presents the main notations and acronyms.

B. Fundamental Properties

Here, a few preliminary concepts of mathematical morphology are recalled. To do so, let us consider f as a discrete grayscale image.

- *Idempotence*: A transformation ψ is idempotent if the output of the transformation is independent of the number of times it is applied to the image, which can be mathematically shown as $\psi(\psi(f)) = \psi(f)$.
- *Increasingness*: A transformation is known as increasing if and only if it keeps the ordering relation between images, which can be mathematically shown as $f \leq g \iff \psi(f) \leq \psi(g) \forall f, g$. The notation $f \leq g$ means that $f(\mathbf{x}) \leq g(\mathbf{x})$ for each pixel $\mathbf{x}$ in the definition domain of the images. A few examples of increasing criteria might include increasing attributes such as area, volume, and size of the bounding box. In contrast, nonincreasing attributes, such as scale-invariant measures (e.g., gray-level homogeneity, shape descriptors, and region orientation), lead to nonincreasing criteria.
- *Extensivity and antiextensivity*: A transformation ψ is extensive if, for each pixel, the transformation output is greater than or equal to the original image, which can be mathematically shown as $f \leq \psi(f)$. The correspondent property is antiextensive if $f \geq \psi(f)$ for all the pixels in the image is satisfied.
- *Absorption property*: The absorption property is fulfilled when two transformations, which are defined by the parameters i, j, are applied to the image, and the following relation is verified: $\psi_i \psi_j = \psi_j \psi_i = \psi_{\max(i,j)}$.

Another preliminary concept is the so-called *connected component*. In a binary image, a connected component is defined as a set of connected pixels. For grayscale images, we have the concept of *flat zone*, which is defined as a set of connected isointensity pixels. Two pixels are connected with respect to a connectivity rule. The 4- and 8-connected are considered common connectivity rules, in which a pixel is said to be adjacent to four or eight of its neighboring pixels, respectively. The connectivity can be extended by more general criteria defining a connectivity class [28].

C. Max-Trees

In order to have efficient implementations of AFs and EFs, one can take advantage of a tree representation of the input data,

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

4   IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING
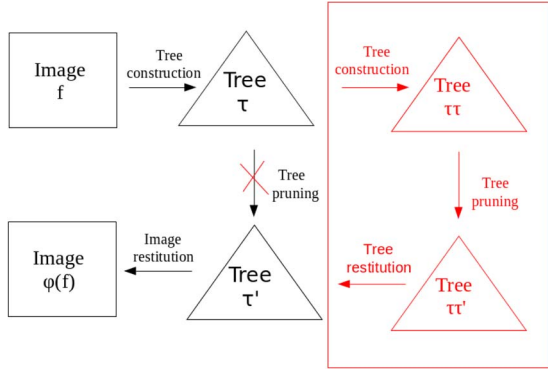

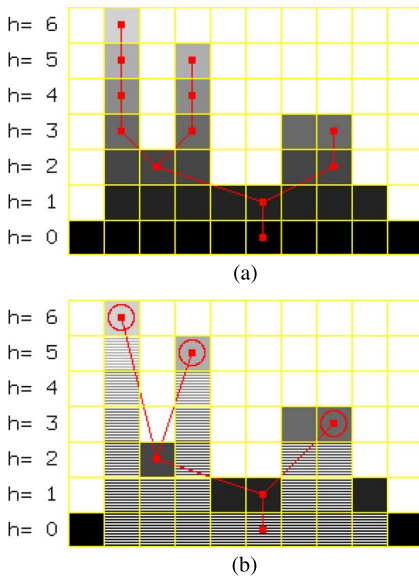Fig. 1. Classical max-tree filtering (black path) and space of shapes filtering (black + red paths).

Fig. 2. (a) Component tree and (b) max-tree corresponding to the 1-D image with intensity values of $f = \{0, 6, 2, 5, 1, 1, 3, 3, 1, 0\}$.

which was introduced in [29]. This representation has received much interest due to the fact that it increases the efficiency of filtering by dividing the transformation process into three steps: 1) tree creation; 2) filtering; and 3) image restitution (see the black path in Fig. 1) [15], [19].

The component tree [30] represents the input image through the hierarchical relationship of its connected components. Each node of the component tree stores all the pixels of the connected component it represents. The max-tree can be considered as a compact structure for the component tree representation. The only differences are that a connected component whose area remains unchanged for a sequence of threshold values is stored in a single node (i.e., composite node) and each node stores only the pixels of the connected component that are visible in the image. In order to express this concept in a simple way, Fig. 2 illustrates the component tree and the max-tree corresponding to the 1-D image $f = \{0, 6, 2, 5, 1, 1, 3, 3, 1, 0\}$. In the figure, the max-tree composite nodes are shown with double circles. The leaves of the component tree and the max-tree correspond to regional maxima. It should be noted that the

min-tree, which is the dual of the max-tree, can be computed using the max-tree algorithm, with the negated image as input.

D. Max-Tree Node Attributes

One of the main advantages of the max-tree data structure is that one can compute diverse attributes of the max-tree nodes that can be used for filtering and object recognition purposes. The height of a max-tree node is a contrast attribute, the area is a size attribute, and the volume is a combination of contrast and size. The diagonal of the bounding box enclosing the node-connected component is a shape and size attribute. These are all increasing attributes which can be computed through the following equations, respectively:

$$\mu_h(C_i) = \max_{\forall k \in \text{descendants}(i)} \{h_k - h_i\} \tag{1}$$

$$\mu_a(C_i) = \sum_{\forall \mathbf{x} \in C_i} 1 \tag{2}$$

$$\mu_v(C_i) = \mu_a(C_i) + \sum_{k \in \text{descendants}(i)} \mu_a(C_k) \times \text{nlevels}(k) \tag{3}$$

$$\mu_{\text{diag}}(C_i) = \sqrt{(x_{i,\max} - x_{i,\min})^2 + (y_{i,\max} - y_{i,\min})^2} \tag{4}$$

where i is a max-tree node, C_i is its corresponding connected component, h_i is the gray level of the node i, nlevels represents the number of sequential threshold values in which the component stayed the same, and descendants(i) is a set containing all the descendants of node i. $x_{i,\max}$, $x_{i,\min}$, $y_{i,\max}$, and $y_{i,\min}$ are the coordinates of the node bounding box.

The standard deviation of a max-tree node is also a contrast attribute, but unlike the height attribute, it is a nonincreasing attribute. It is given by the following equation:

$$\mu_{\text{std}}(C_i) = \sqrt{\frac{1}{\mu_a} \sum_{\forall \mathbf{x} \in C_i} (f(\mathbf{x}) - \mu_{\text{gray-level}}(i))^2} \tag{5}$$

where $\mu_{\text{gray-level}}(i)$ is the average gray-level intensity of node i, which is given by

$$\mu_{\text{gray-level}}(C_i) = \frac{1}{\mu_a} \sum_{\forall \mathbf{x} \in C_i} f(\mathbf{x}). \tag{6}$$

E. Extinction Values

Extinction values are a measure of persistence of regional extrema. They are a powerful tool with which to measure the persistence of an attribute and are useful to discern relevant from irrelevant extrema, usually noise. Loosely speaking, the extinction value of a regional extremum (minimum or maximum) of any increasing attribute is the maximal size of the AF [31], such that this extremum still exists after filtering [25].

The formal definition of the extinction value of a regional maximum given by Vachier [25] is the following: consider M a regional maximum of an image f, whereas $\Psi = (\psi_\lambda)_\lambda$ is a family of decreasing connected antiextensive transformations.

The extinction value corresponding to M with respect to Ψ, which is denoted by $\varepsilon_\Psi(M)$, is the maximal λ value, such that M is still a regional maxima of $\psi_\lambda(f)$. This definition can be expressed through the following equation:

$$\varepsilon_\Psi(M) = \sup\{\lambda \geq 0 | \forall \mu \leq \lambda, M \subset \mathrm{Max}(\psi_\mu(f))\} \quad (7)$$

where $\mathrm{Max}(\psi_\mu(f))$ is the set containing all the regional maxima of $\psi_\mu(f)$. Extinction values of regional minima can be defined in a similar way. Extinction values can be efficiently computed on the max-tree structure [32].

F. EFs

An EF is a connected filter that preserves the relevant extrema of the image. This filter can be defined as follows: let $\mathrm{Max}(f) = \{M_1, M_2, \ldots, M_N\}$ be the set of regional maxima of the image f. Each regional maximum M_i has a corresponding extinction value ϵ_i, which is defined by Vachier [25]. The EF of f set to preserve the n maxima with the highest extinction values is given by

$$\mathrm{EF}^n(f) = R_g^\delta(f) \quad (8)$$

where $R_g^\delta(f)$ is the reconstruction by dilation [33] of the mask image g from marker image f. The mask image g is given by

$$g = \max_{i=1}^{n}\{M_i'\} \quad (9)$$

where $\max$ is the pixelwise maximum operation. M_1' is the maximum with the highest extinction value, M_2' has the second highest extinction value, and so on.

EFs can be efficiently implemented in the max-tree structure [24]. The n maxima (max-tree leaves) with the highest extinction values with respect to the attribute being analyzed are chosen. The nodes in the paths from these leaves to the root are marked as to be kept. All other nodes are pruned from the max-tree. Due to the fact that the contraction of max-tree nodes is a connected filter, the EF is a connected filter. Moreover, it is also idempotent. Fig. 3 illustrates a simple workflow of the EF. Let $n = 3$ and nodes 8, 15, and 16 (the blue nodes) in Fig. 3(a) be the leaves with the highest extinction values with respect to the attribute being analyzed. The nodes in the paths from these leaves to the root are presented in red. In Fig. 3(b), the remaining nodes are pruned. The resulting tree is illustrated in Fig. 3(c).

An example of a 41×68 region of interest of a satellite image and the result of an area-open and an area EF are displayed in Fig. 4. The fact that EFs preserve the height of the maxima can be noted in the max-tree representation and the filtered images. The tops of the buildings are much brighter in the area EF result than in the area-open result.

EFs are closely related to dynamic openings [34], and unlike the usual contrast and size filters, the heights of the remaining extrema in the image are completely preserved. Souza *et al.* showed that EFs are a more efficient alternative than AFs with respect to simplification for recognition [24] since they are able to preserve more regions and correspondences found by affine region detectors. Another advantage of EFs over AFs is that
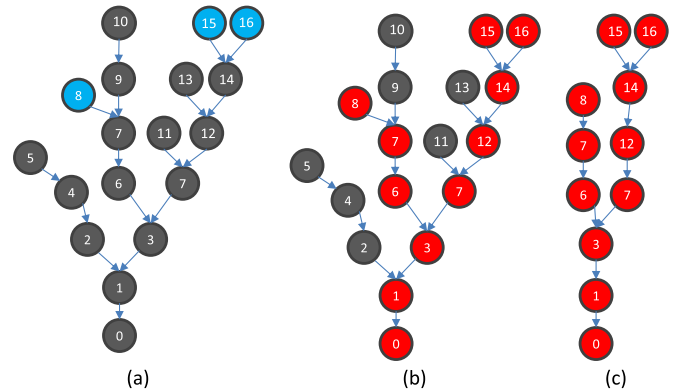

Fig. 3. (a) Original max-tree; the blue nodes represent the three nodes with the highest extinction values. (b) Nodes in the path from the three leaves with the highest extinction values to the root are presented in red. (c) Result of the pruning of the nodes is not marked in red.
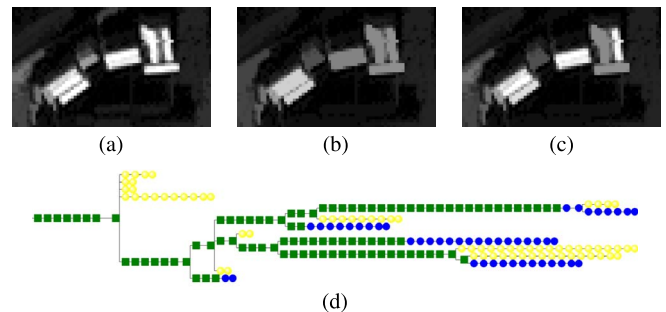

Fig. 4. (a) Original image (all max-tree nodes), (b) area open (green rectangles) with $t = 58$, (c) area EF (green rectangles + blue circles) with $n = 5$, and (d) max-tree representation.

EFs' parameters are easier to set, independently of the kind of attribute being used (e.g., area and volume), since they are based on the number of extrema. The thresholds used by AFs vary greatly according to the attribute being used and the data set being analyzed and are therefore more difficult to set. In other words, the main shortcoming of conventional APs, which are related to the initialization of the threshold values, is being addressed by EPs.

G. EFs With Nonincreasing Attributes

Until recently, it was not possible to compute extinction values for nonincreasing attributes. In 2012, Xu *et al.* [35] proposed an approach called space of shapes, which basically consists of computing a max-tree from another tree representation (e.g., max-tree and min-tree). During the second tree construction, the attribute used should be the attribute you want to become increasing in the new representation. The methodology described corresponds to the black path plus the red path in Fig. 1. The second max-tree construction is much faster than the first, since the complexity of the second max-tree construction depends on the number of nodes in the initial max-tree, whereas the first max-tree construction depends on the number of image pixels, which, in general, is much higher than the number of max-tree nodes [36].
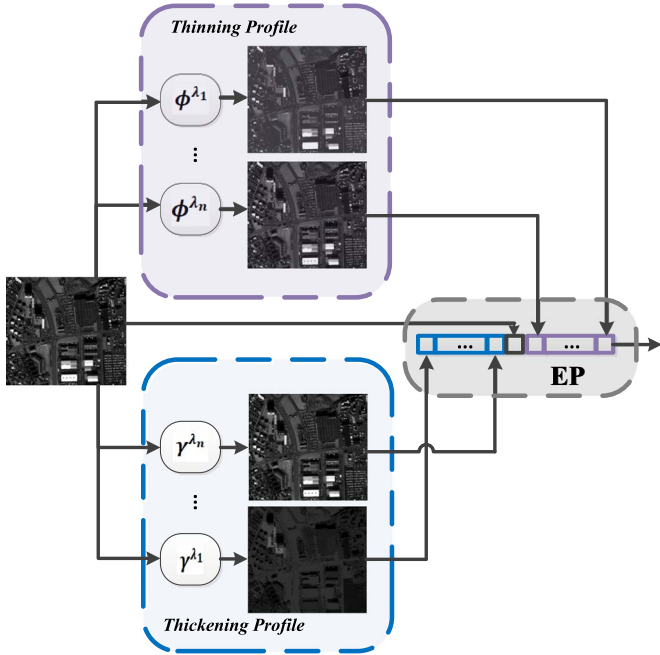
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

6

IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING

Fig. 5. Simple representation of EP.

H. EPs and DEPs

EPs are built by applying a sequence of thinning and thickening transformations to a grayscale image. In other words, EPs carry out a multilevel decomposition of the input image based on EFs. Similar to MPs and APs, the EP can be described as a concatenation of a thickening EP, i.e., Π_{ϕ^λ}, and a thinning EP, i.e., Π_{γ^λ}, computed with a generic ordered criterion λ (also called threshold or criteria), which can be mathematically given by

$$
\text{EP}(f) = \left\{ \underbrace{\phi^{\lambda_1}(f), \dots, \phi^{\lambda_{n-1}}(f), \phi^{\lambda_n}(f),}_{\text{Thickening Profile } (\Pi_\phi)} \right.
$$
$$
\left. f, \underbrace{\gamma^{\lambda_n}(f), \gamma^{\lambda_{n-1}}(f), \dots, f, \gamma^{\lambda_1}(f)}_{\text{Thinning Profile } (\Pi_\gamma)} \right\} \quad (10)
$$

and/or

$$
\text{EP}(f) = \begin{cases} \Pi_{\phi^{\lambda_n}}, & n = (n - i + 1), \quad \forall i \in [1, n]; \\ \Pi_{\gamma^{\lambda_n}}, & n = (i - n), \quad \forall i \in [n+1, 2n] \end{cases}. \quad (11)
$$

where n is the number of thresholds (i.e., criteria). The set of ordered thresholds $\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$ for λ_i, $\lambda_j \in \lambda$ and $j \geq i$; the relation $\lambda_i \leq \lambda_j$ holds for thickening, and $\lambda_i \geq \lambda_j$ holds for thinning.[1] The equation also depicts that, by defining n threshold values λ (here, the number of extrema), one will come up with $2n + 1$ features, including n features produced

[1] Please note that, for the proposed method, the higher value of extrema can provide more detail. This contrasts with the conventional thresholding approach applied on APs, in which the higher value of the threshold causes more smoothness. In other words, for the proposed approach, the feature produced by the higher number of extrema is placed closer to the input image in the profile.

by thickening, n features produced by thinning, and the input grayscale image.

The thickening profile is considered in reverse order in which the highly filtered out image is placed first and the original image last. The input grayscale image f is also placed in the profile, since it can be considered as the level zero of both the thickening and thinning profiles. An EP can be mathematically given as in (11), where λ_i represents the number of extrema (maxima or minima) preserved by EFs, and $\lambda_n > \lambda_{n-1} > \cdots > \lambda_1$. Fig. 5 illustrates the general idea of the EP. As shown, the profile is the concatenation of a set of features produced by thinning and thickening and the input grayscale data set. Fig. 6 is provided for a better understanding of the concept of a profile produced by EFs. As shown, the EP is composed of a number of features produced by an area attribute with extrema values of 8192, 2048, 512, 128, 32, 8, and 2. The higher number of extrema contains more detail, whereas the smaller number smooths out the input data to a great extent. In other words, the EP simplifies the input data by excluding unnecessary information, based on the attribute, as the number of extrema decreases.

In contrast to MPs that are only able to model the size and structure of different objects, EFs are more flexible and can be of any type. To this extent, a multi-EP (MEP) concatenates different EPs (e.g., area, height, volume, diagonal of the bounding box, and standard deviation) into a single stacked vector, which can be mathematically given by

$$
\text{MEP} = \{\text{EP}_{a_1}, \text{EP}_{a_2}, \dots, \text{EP}_{a_m}\} \quad (12)
$$

where a_k, $k = \{1, \dots, m\}$, represents different types of attributes. Since different extinction attributes can extract complementary spatial information, the MEP can extract more spatial information than a single EP while, at the same time, the computational cost of producing these features is almost the same since the max-tree and the min-tree are computed only once for each grayscale image (except for the standard deviation extinction attribute) and are filtered with different attributes at different levels.

As an example, Fig. 7 shows a profile that considers different attributes (i.e., from top to bottom: thinning profile of area, thickening profile of area, thinning profile of volume, thickening profile of volume, thinning profile of height and thickening profile of height, thinning profile of diagonal of the bounding box, thickening profile of diagonal of the bounding box, and thinning profile of standard deviation and thickening profile of standard deviation). The number of extrema is set as 512, 128, 64, 16, 4, and 1. In order to have a complete profile, the input feature should be also concatenated, along with all the features, into a stacked vector. As can be seen, the higher number of extrema extracts more detail from the input data, whereas the lower number of extrema considerably simplifies the input data. By considering different attributes, a diverse set of features can be obtained, which are useful in terms of classification accuracy. For example, the features produced by the standard deviation attribute are different from the ones produced by other attributes. In other words, different features emphasize different objects of interest while eliminating other objects.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

GHAMISI *et al.*: EXTINCTION PROFILES FOR THE CLASSIFICATION OF REMOTE SENSING DATA
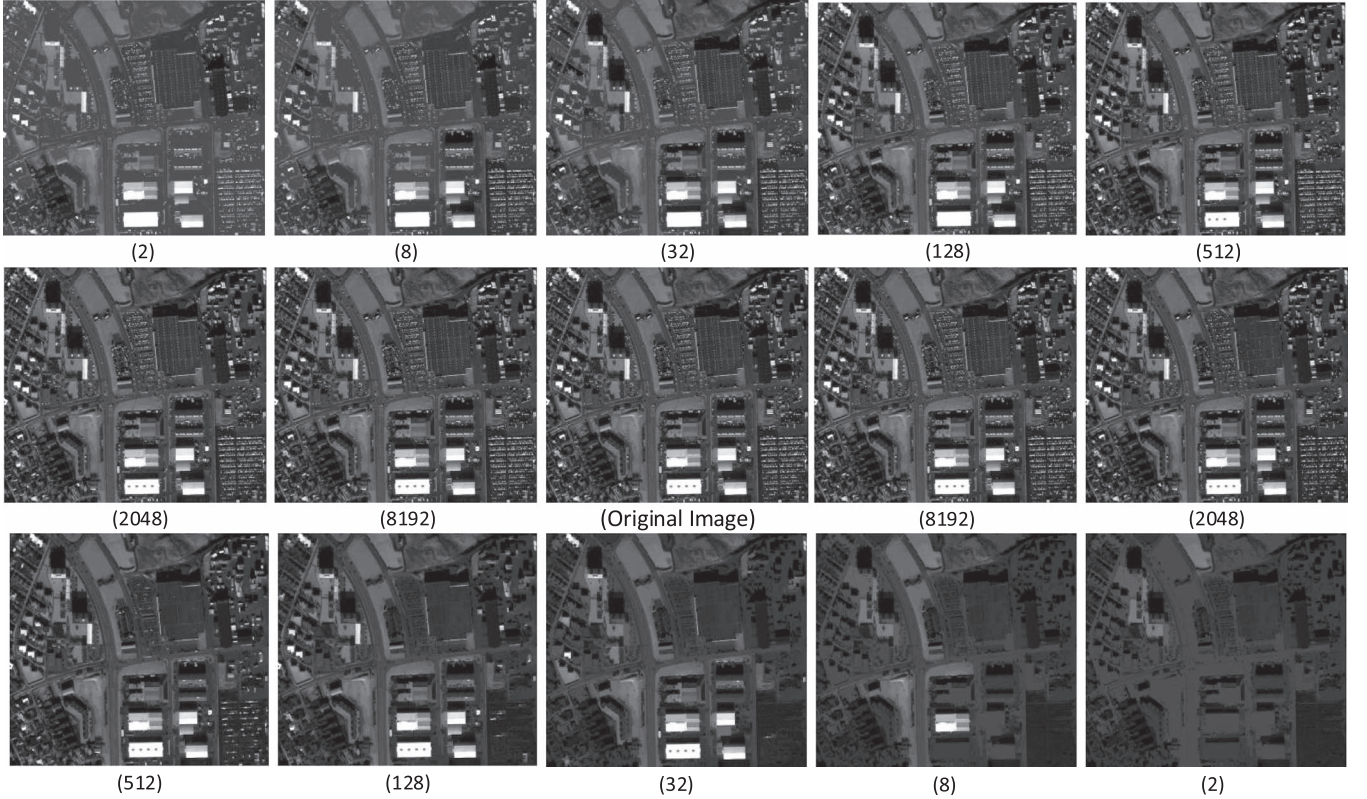7

Fig. 6. EP made by considering an area attribute. The number of extrema is set to be 8192, 2048, 512, 128, 32, 8, and 2. The first seven features are produced by thickening, the eighth one is the original image, and the rest are produced by thinning. The higher number of extrema contains more detail, whereas the smaller number smooths out the input data to a great extent.

MPs are, by nature, able to fulfill the absorption property, since they are a sequential composition of openings and closings that consider a structural element of increasing size. However, this characteristic might not be always verified by EP. If we consider attributes that cannot fulfill the increasing property, instead of opening and closing, one might come up with thinning (or thickening). Since attribute thinning (or thickening) is not always increasing, the absorption law might not be satisfied in the profile. This can produce a profile in which the elements are not sequential with respect to the attribute. For instance, regions eliminated at a certain level of the profile might appear again in subsequent levels associated with more relaxed criteria [15]. This effect is not as desirable in applications for which the order of the features with respect to the attribute is important since it can be an issue when a derivative of the profile needs to be computed. If the absorption property is fulfilled by filtered images, one can obtain a progressively increased simplification of the image as the filter values get stricter, which makes the computation of the derivative of the profile well defined. EPs are able to fulfill this property when the considered criterion is increasing. Otherwise, a constraint on the criterion has to be taken into account. The set of criteria $\lambda = \{\lambda_1, \ldots, \lambda_n\}$ considered in the profile must be formally ordered, leading to $\gamma^{\lambda_i} \supseteq \gamma^{\lambda_j}$ and $\phi^{\lambda_i} \subseteq \phi^{\lambda_j}$ for $i \leq j$. In EP, since the higher number of extrema leads to more detail, the higher value of extrema should be located closer to the original image.

The DEP is composed of the residues of two subsequent filtering operations for two adjacent levels existing in the profile. The obtained map is generated by associating each pixel with

the level in which the maximum of the DEP (evaluated at the given pixel) occurs. Since the DEP is the derivative of the EP, it has a number of levels that are one less than the number of levels in the EP. Inspired by the DMP and the derivative of AP (DAP), the DEP can be mathematically presented as

$$\mathrm{DEP}(f) = \begin{cases} \Delta_{\phi^{\lambda_k}}, & k = (n - i + 1), \quad \forall i \in [1, n]; \\ \Delta_{\gamma^{\lambda_k}}, & k = (i - n), \quad \forall i \in [n+1, 2n] \end{cases}. \quad (13)$$

In the equation, Δ_ϕ and Δ_γ are the derivative of the thickening profile and the derivative of the thinning profile, respectively. Fig. 8(a)–(d) shows the DEP by considering 1, 9, and 81 extrema for both thickening and thinning profiles applied on the Rome data set (see Section III-A) for area, diagonal of the bounding box, height, and standard deviation, respectively. In this paper, DEP has been introduced as a new modification of EP. Analogous to DMP and DAP, this concept can be used for different applications such as image segmentation [10] and classification [37]. However, here, we intentionally put the emphasis on EP, which is the main contribution of this paper.

I. Analysis of Computational Complexity

The MP always demands two complete image transformations: one performed by a closing and the other by an opening for each level of the profile. However, in order to produce an EP or an AP, it is only necessary to represent the input image once as max-tree for the thinning and once as min-tree for the
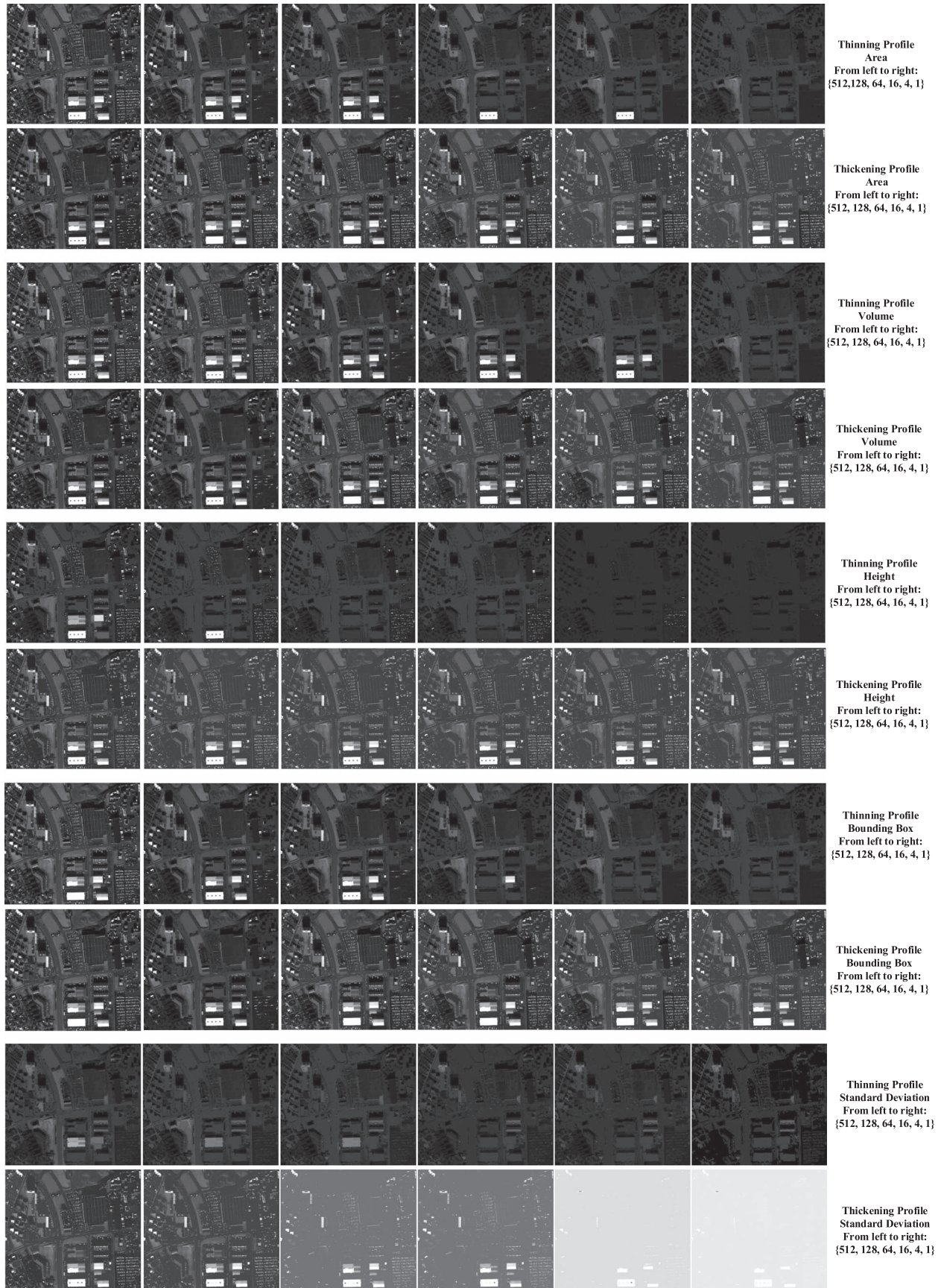
Fig. 7.   EP made by considering different attributes. (Top to bottom) Thinning and thickening profiles of area, thinning and thickening profiles of volume, thinning and thickening profiles of height, thinning and thickening profiles of diagonal of the bounding box, and thinning and thickening profiles of standard deviation. The number of extrema is set to be 512, 128, 64, 16, 4, and 1.

(a)

(b)

(c)

(d)

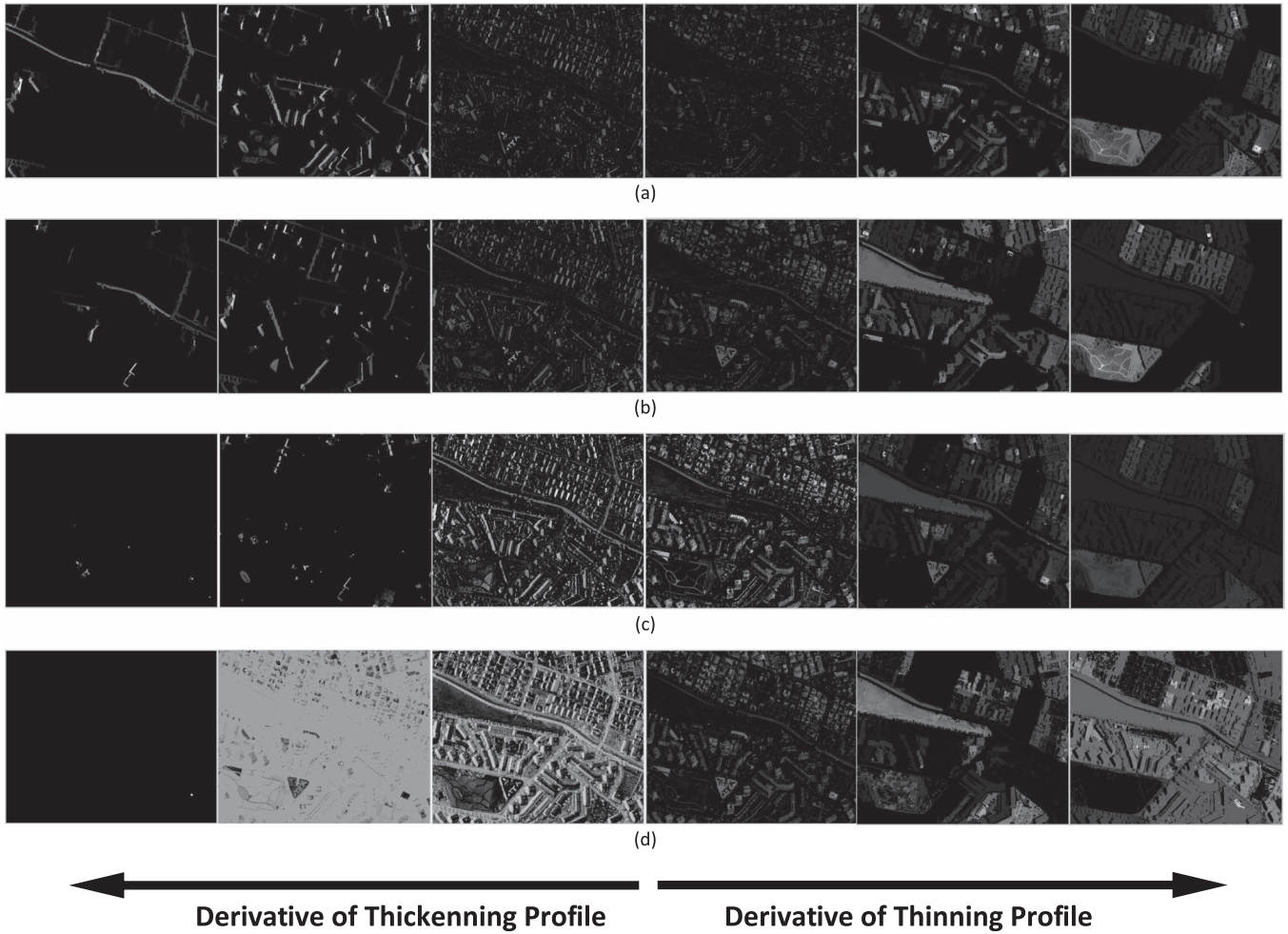**← Derivative of Thickenning Profile**   **Derivative of Thinning Profile →**

Fig. 8. DEP made by considering levels 1, 9, and 81 extrema for both thickening and thinning profiles: (a) area, (b) diagonal of the bounding box, (c) height, and (d) standard deviation.

TABLE II
COMPLEXITY ANALYSIS OF THE DIFFERENT STAGES IN AP AND EP COMPUTATION. THE PARAMETER "S"
REFERS TO THE NUMBER OF THRESHOLD VALUES IN THE PROFILE

Operation	Complexity	# ocurrences AP	# ocurrences EP
Max-tree construction	$O(N log N)$	2	2
Attribute computation	$O(N)$	2	2
2nd max-tree construction	$O(M log M)$	0	2 (Only for non-increasing attributes)
Extinction values computation	O(M)	0	2
Filtering	O(N)	2s	2s

thickening for the entire profile, unless it is a nonincreasing attribute. In this case, it is necessary to build two max-trees and two min-trees to compute the EP. The set of filtering is obtained by sequential pruning of the same trees with different values of criteria (for EP, the number of extrema). This greatly reduces the burden of the analysis with respect to MPs, since the most demanding phase of the filtering, which is the construction of a tree [38], is done only once.

Computing EPs and APs with the same size and for the same attribute has similar processing times. Both profiles will have to build the max-tree and the min-tree once, compute the attribute being analyzed, and perform the same number of filtering steps. The main difference is that EFs have to compute the extinction values for the attribute, but this can be done simultaneously with the number of nodes [32]. Therefore, it does not add much to the processing time. For nonincreasing attributes, EPs also require a second max-tree (min-tree) construction, which APs do not need, but this second max-tree (min-tree) can be computed much faster, as discussed before, since its complexity is proportional to the number of nodes M of the first tree instead of the number of pixels N in the original image, and usually, $M \ll N$ [36]. Another factor that influences the computational time of building EPs and APs is the data type of the images: the max-tree construction algorithm has different behaviors for integer images with low quantization, high quantization, and floating-point images. Table II summarizes the complexity of each stage in the computation of APs and EPs for floating-point images using sequential algorithms, considering the most usual attributes. For a complete survey on max-tree construction, filtering, and attributes computation algorithms, please see [38]–[40].
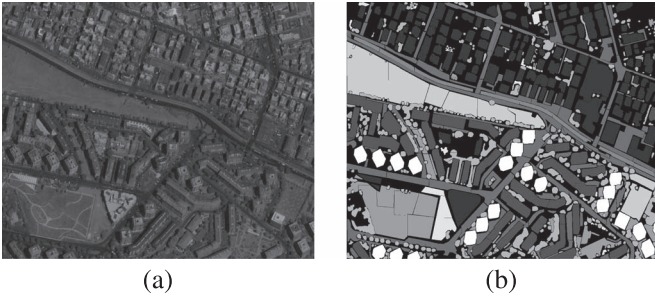
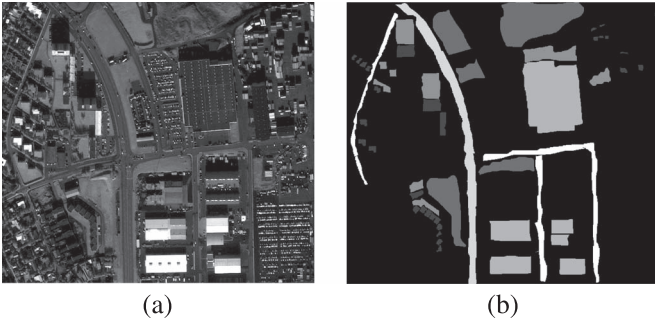Fig. 9. (a) Rome satellite image and (b) its reference data.

Fig. 10. (a) Reykjavik satellite image and (b) its reference data.

TABLE III
REYKJAVIK: THE NUMBER OF TRAINING AND
TEST SAMPLES FOR DIFFERENT CLASSES

Class	Training	Test
Residential	1863	4350
Soil	6068	22076
Shadow	2619	7991
Commercial	5599	24169
Highway	2489	9562
Road	4103	7837

III. EXPERIMENTAL RESULTS

A. Data Sets Descriptions

1) Rome Data Set: The first data set (see Fig. 9) was captured over an urban area of Rome, Italy, which was acquired by the QuickBird satellite. It consists of a high-resolution (0.6 m) panchromatic image. It has nine classes, namely, buildings (181 255 samples), blocks (109 816 samples), roads (163 529 samples), light train (16 060), vegetation (69 617 samples), trees (90 880 samples), bare soils (81 271 samples), soil (15 057 samples), and towers (47 916 samples). The classification step for the Rome data set is repeated ten times, randomly selecting 10% of the samples for training and the rest as test samples to avoid any bias induced by random sampling of the training and test samples.

2) Reykjavik Data Set: The second data set (see Fig. 10) is a high-resolution panchromatic image of Reykjavik, Iceland, acquired by the Ikonos satellite. This data set consists of six classes, namely, residential, commercial, shadow, soil, highway, and road. The Reykjavik data set comes with the training and test sets already split. Therefore, RF has been applied only once to different images. The number of training and test samples for each class is listed in Table III.

B. Algorithm Setup

An RF classifier with 200 trees is used to classify the images. In order to compare classification accuracies of different approaches, overall accuracy (OA), average accuracy (AA), and kappa coefficient (K) have been taken into account. The classification procedure was evaluated using the AP and EP features.

In [41], in order to generate the APs using the area attribute, ten thresholds {25, 100, 500, 1000, 5000, 10 000, 20 000, 50 000, 100 000, 150 000} were taken into account. This is our main motivation to set the number of thresholds for all APs and EPs to 10. The AP thresholds used for the other thresholds were chosen based on the maximum value of each attribute, disregarding extreme values such as the root node, which usually has a much higher attribute than the other nodes. We divided this maximum value into ten equidistant parts. The difficulty in choosing the thresholds for APs is one of the major improvements that EPs address. In order to generate the EP, the values of n used to generate the profile for different attributes are automatically given by the following equation:

$$\lfloor \alpha^j \rfloor, \qquad j = 0, 1, \ldots, s - 1. \tag{14}$$

The total EP size is $2s + 1$, since the original image is also included in the profile. The preceding equation was determined experimentally. The larger α is, the larger are the differences between consecutive images in the profile. The smaller α is, the fewer extrema there will be, where most of the image information is usually present [24]. Our recommendation is to use an α between 2 and 5. In our experiments in this paper, we used $\alpha = 2$ and set $s = 10$, so that both the AP and the EP have the same length, making it a fair comparison. The profiles were computed considering the 4-connected connectivity rule.

C. Results and Discussion for EPs

1) Rome: Tables IV and V present classification accuracies (OA, AA, and kappa coefficient) of the Rome data set by considering AP and EP composed of a single attribute with parameters defined in Section III-B. As can be seen, both approaches considerably outperform the result of the RF on the original input data (i.e., the Raw data). In this manner, the extinction attributes EP_a, EP_h, EP_v, EP_{bb}, and EP_{std} improve the classification accuracy of *Raw* by almost 43%, 38%, 42%, 38%, and 31%, respectively. The results confirm that the spatial information extracted by EPs can significantly improve the classification accuracies of the situation when spatial information is discarded by the classification system. The proposed approach has also outperformed APs for all single attributes, namely, EP_a, EP_h, EP_v, EP_{bb}, and EP_{std}, by almost 1%, 4%, 46%, 5%, and 7% in terms of OA, respectively. The main reasons for the superior performance of EPs over APs are as follows: the EPs' ability to preserve more relevant regions suitable for classification and the EFs' ability to preserve the height of the extrema, which is useful for differentiating among different classes of interest. In addition, it is important to note that EPs are not as dependent on the threshold values as APs. In other words, EPs are initialized automatically with respect to the

TABLE IV
ROME: CLASSIFICATION RESULTS OF THE ROME DATA SET CONSIDERING AP AND EP WITH PARAMETERS DEFINED IN SECTION III-B. IN THIS TABLE, a, v, AND h REPRESENT THE AREA, VOLUME, AND HEIGHT ATTRIBUTES, RESPECTIVELY. METRICS AA AND OA ARE REPORTED IN PERCENTAGES. KAPPA IS A COEFFICIENT WITH CHANGES IN THE RANGE OF 0 AND 1

Metrics	Raw	AP_a	EP_a	AP_h	EP_h	AP_v	EP_v
AA	26.62 ± 0.16	81.95 ± 0.11	82.86 ± 0.14	72.87 ± 0.29	78.41 ± 0.08	18.64 ± 0.00	82.83 ± 0.11
OA	41.29 ± 0.07	82.78 ± 0.06	83.49 ± 0.05	75.62 ± 0.18	79.87 ± 0.08	37.18 ± 0.02	83.52 ± 0.06
Kappa	0.2790 ± 0.0014	0.7949 ± 0.0007	0.8036 ± 0.0006	0.7065 ± 0.0024	0.7603 ± 0.0009	0.1898 ± 0.0001	0.8041 ± 0.0007

TABLE V
ROME: CLASSIFICATION RESULTS OF THE ROME DATA SET CONSIDERING AP AND EP WITH PARAMETERS DEFINED IN SECTION III-B. IN THIS TABLE, bb AND std REPRESENT THE DIAGONAL OF THE BOUNDING BOX AND STANDARD DEVIATION ATTRIBUTES, RESPECTIVELY. METRICS AA AND OA ARE REPORTED IN PERCENTAGES. KAPPA IS A COEFFICIENT WITH CHANGES IN THE RANGE OF 0 AND 1

Metrics	AP_{bb}	EP_{bb}	AP_{std}	EP_{std}
AA	67.52 ± 0.41	77.39 ± 0.16	56.91 ± 0.30	69.92 ± 0.18
OA	74.22 ± 0.07	79.97 ± 0.08	64.69 ± 0.23	72.07 ± 0.08
Kappa	0.6920 ± 0.0008	0.7615 ± 0.001	0.5734 ± 0.0036	0.6653 ± 0.0009

TABLE VI
ROME: CLASSIFICATION RESULTS OF THE ROME DATA SET CONSIDERING MAP AND MEP WITH PARAMETERS DEFINED IN SECTION III-B. IN THIS TABLE, a, v, h, AND bb REPRESENT THE AREA, VOLUME, HEIGHT, AND DIAGONAL OF THE BOUNDING BOX ATTRIBUTES, RESPECTIVELY. METRICS AA AND OA ARE REPORTED IN PERCENTAGES. KAPPA IS A COEFFICIENT WITH CHANGES IN THE RANGE OF 0 AND 1

Metrics	$AP_{a,bb}$	$EP_{a,bb}$	$AP_{a,v,h}$	$EP_{a,v,h}$	$AP_{a,v,h,bb}$	$EP_{a,v,h,bb}$
AA	83.37 ± 0.14	84.80 ± 0.09	83.27 ± 0.14	84.26 ± 016	84.74 ± 0.07	83.60 ± 0.14
OA	83.95 ± 0.05	85.57 ± 0.05	83.89 ± 0.05	84.91 ± 0.05	85.44 ± 0.04	84.15 ± 0.04
Kappa	0.8092 ± 0.0006	0.8284 ± 0.07	0.8085 ± 0.0006	0.8203 ± 0.0008	0.8269 ± 0.0004	0.8117 ± 0.0005

TABLE VII
ROME: CLASSIFICATION RESULTS OF THE ROME DATA SET CONSIDERING MAP AND MEP WITH PARAMETERS DEFINED IN SECTION III-B. IN THIS TABLE, a, v, h, bb, AND std REPRESENT THE AREA, VOLUME, HEIGHT, DIAGONAL OF THE BOUNDING BOX, AND STANDARD DEVIATION ATTRIBUTES, RESPECTIVELY. METRICS AA AND OA ARE REPORTED IN PERCENTAGES. KAPPA IS A COEFFICIENT WITH CHANGES IN THE RANGE OF 0 AND 1

Metrics	$AP_{a,std}$	$EP_{a,std}$	$AP_{a,v,h,bb,std}$	$EP_{a,v,h,bb,std}$
AA	84.85 ± 0.14	83.28 ± 0.12	85.13 ± 0.10	83.75 ± 0.15
OA	85.72 ± 0.06	83.87 ± 0.04	85.90 ± 0.05	84.30 ± 0.05
Kappa	0.8301 ± 0.0007	0.8082 ± 0.0005	0.8323 ± 0.0006	0.8135 ± 0.0006

TABLE VIII
REYKJAVIK: CLASSIFICATION RESULTS OF THE REYKJAVIK DATA SET CONSIDERING AP AND EP WITH PARAMETERS DEFINED IN SECTION III-B. IN THIS TABLE, a, v, h, bb, AND std REPRESENT THE AREA, VOLUME, HEIGHT, DIAGONAL OF THE BOUNDING BOX, AND STANDARD DEVIATION ATTRIBUTES, RESPECTIVELY. METRICS AA AND OA ARE REPORTED IN PERCENTAGES. KAPPA IS A COEFFICIENT WITH CHANGES IN THE RANGE OF 0 AND 1

Metrics	Raw	AP_a	EP_a	AP_h	EP_h	AP_v	EP_v	AP_{bb}	EP_{bb}	AP_{std}	EP_{std}
AA	52.87	73.4	76.93	53.47	60.98	55.78	73.17	68.77	75.44	69.51	60.91
OA	62.45	78.4	83.63	66.81	67.65	66.58	77.93	76.77	79.82	77.21	71.2
Kappa	0.5123	0.7263	0.7898	0.5589	0.5887	0.5615	0.7195	0.7041	0.7435	0.7052	0.6211

number of extrema. Therefore, by considering a set of numbers, one can obtain high classification accuracies. However, this is not applicable to conventional APs. For APs, one needs to have detailed information regarding the type, sensors, and spatial resolution of the data in order to specify meaningful thresholds. These threshold values can be varied for different attributes and different data sets. However, this is not an issue for EPs.

In terms of multiple attributes, in some cases such as $EP_{a,v,h,bb}$ and $EP_{a,v,h,bb,\text{std}}$, Multi AP (MAP) provides better results in terms of classification accuracies than MEP (see Tables VI and VII). The reason why MAP outperforms MEP in those cases might be that different EPs may produce a few similar features. In this context, redundant features may reduce the quality of the classification. However, due to our way of thresholding for APs, we were able to initialize the threshold values with respect to the maximum value of each attribute divided in a few equidistant parts. As a result, with a high probability, MAP is composed of different features that are not that similar due to the fact that the maximum value of each attribute is different. It is also important to note that the attributes *volume* and *height* have been considered here for the first time along with APs, which show their efficiency for extracting complementary information for MAP.

2) Reykjavik: Tables VIII and IX give information on the classification accuracies obtained by different APs, EPs, MAPs, and MEPs, respectively. Similar to the Rome data set, both approaches exponentially boost *Raw* in terms of classification accuracies. For instance, the extinction attributes EP_a, EP_h, EP_v, EP_{bb}, and EP_{std} improve the classification accuracy of *Raw* by approximately 21%, 4%, 15%, 17%, and 9%, respectively.

TABLE IX
REYKJAVIK: CLASSIFICATION RESULTS OF THE REYKJAVIK DATA SET CONSIDERING MAP AND MEP WITH PARAMETERS DEFINED IN SECTION III-B.
IN THIS TABLE, a, v, h, bb, AND std REPRESENT THE AREA, VOLUME, HEIGHT, DIAGONAL OF THE BOUNDING BOX,
AND STANDARD DEVIATION ATTRIBUTES, RESPECTIVELY. METRICS AA AND OA ARE REPORTED
IN PERCENTAGES. KAPPA IS A COEFFICIENT WITH CHANGES IN THE RANGE OF 0 AND 1

Metrics	$AP_{a,bb}$	$EP_{a,bb}$	$AP_{a,v,h}$	$EP_{a,v,h}$	$AP_{a,v,h,bb}$	$EP_{a,v,h,bb}$	$AP_{a,std}$	$EP_{a,std}$	$AP_{a,v,h,bb,std}$	$EP_{a,v,h,bb,std}$
AA	76.83	78.87	74.57	78.74	76.61	79.36	74.74	77.71	76.76	79.09
OA	80.7	84.96	79.24	85.18	80.6	85.7	80.48	84.93	80.67	85.58
Kappa	0.7551	0.8069	0.737	0.8093	0.754	0.816	0.7516	0.8061	0.7548	0.8145

With reference to the tables, EPs and MEPs provide better results in terms of classification accuracies than APs and MAPs. This way, EP outperforms AP for attributes a, h, v, and bb by almost 5%, 1%, 11%, and 3% in terms of OA for attributes area and volume, respectively. For the Reykjavik data set, AP_{std} improves the OA of EP_{std} by 6%. The main reason is because of the way in which the thresholds have been set for AP_h, which proves that the results of APs are highly dependent on the initialization step. In terms of multiple attributes, in all cases, namely, $\{a, bb\}$, $\{a, v, h\}$, $\{a, v, h, bb\}$, $\{a, stdd\}$, and $\{a, v, h, bb, std\}$, MEP outperforms its alternative in terms of OA by 4%, 6%, 5%, 4%, and 5%, respectively.

D. Structural Similarity, Image Simplification, and Affine Region Detection Analysis

Here, we analyze the structural similarity, image simplification, and affine region detection rates for the EP and the AP. In order to evaluate the simplification performance, we computed the flat-zones simplification rate given by the number of flat zones in the filtered image divided by the number of flat zones in the original image. This metric ranges between zero and one, and the closer to zero this metric is, the higher the image simplification becomes. The images' similarity was evaluated through the structural similarity (SSIM) index [42] between the original image and the filtered image. This metric also ranges between zero and one, but the higher the metric, the higher the structural similarity of the images.

The affine region detection analysis was performed using the maximally stable extremal region (MSER) [43] affine detector. We chose this detector because it is the one that achieves the best results in the survey by Mikolajczyk *et al.* [44], and like EFs and AFs, it can be efficiently computed from the max-tree structure.

The flat-zones simplification rates for each image in the area profile are illustrated in Fig. 11. Both EP and AP have a similar behavior, as expected, but AP simplifies the image more. The SSIM indexes are shown in Fig. 12. EP has higher structural similarity than most corresponding AP images in the Rome and Reykjavik profiles. These results are due the fact that EFs preserve the height of the extrema they keep, therefore simplifying the image less, but preserving more image details than AFs.

The number of MSERs detected is illustrated in Fig. 13. The MSERs are computed from the corresponding thickening and thinning profiles; therefore, the horizontal axis has only $n + 1$ points. Index 0 corresponds to the number of regions found in the original image, index 1 corresponds to the number of
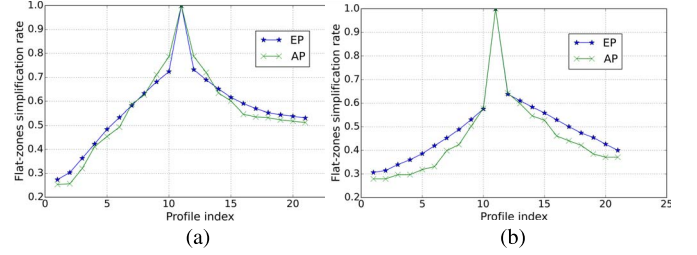

Fig. 11. Flat-zones simplification profiles for the area EP and AP. (a) Rome. (b) Reykjavik.
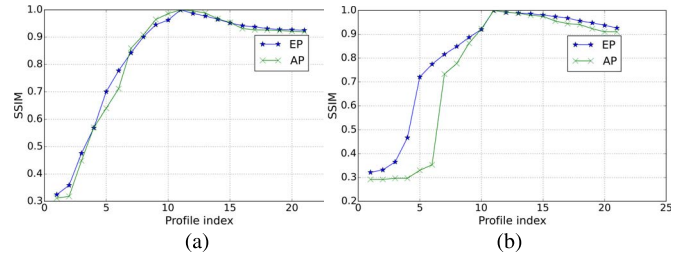

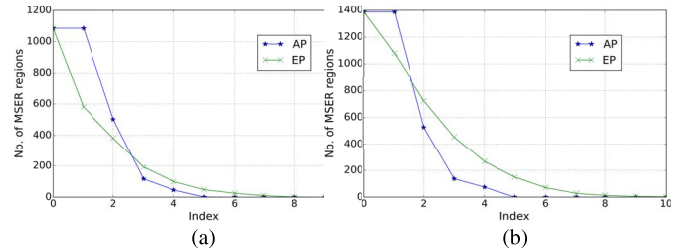Fig. 12. SSIM profiles for the area EP and AP. (a) Rome. (b) Reykjavik.

Fig. 13. Number of MSERs detected. (a) Rome. (b) Reykjavik.

regions found in the first filtering step, and so on. The plot shows that the number of MSERs found in the EP decreases more slowly than in the AP. Many images in the AP find 0 MSER, which is undesirable, since this indicates that the relevant information in the image was completely filtered out. We believe that this is one of the reasons why the EP outperforms the AP, and it is indirectly related to the problem of setting the profile thresholds, which is much harder for APs.

E. Quick Note on Processing Time

Our max-tree implementation is a didactic one that computes many more attributes than necessary to compute the profiles. The max-tree structure and the filtering algorithm used are the
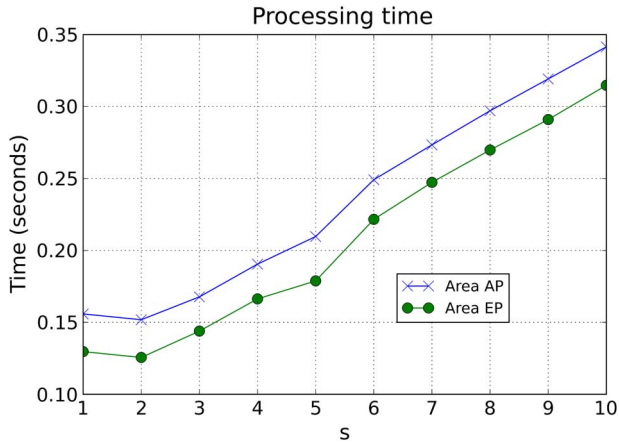
Fig. 14. Area EP and area AP processing times for different profile sizes (s).

ones proposed by Souza *et al.* [36]. In order to illustrate the processing times to extract APs and EPs, in Fig. 14, we display the processing time required to compute the area EP and the area AP with different sizes s using the Reykjavik data set. We can see that, even without an optimized code, computation of the profiles occurs very fast, taking less than half a second for this data set. The processing times were measured on a 4-core virtual machine running in an Intel Xeon X5675 server with a clock of 3.06 GHz and 12 MB of cache memory. Our experience and the results reported by Souza *et al.* [24] indicate that EPs and APs of the same size require similar processing times.

IV. CONCLUSION

In this paper, a novel approach has been proposed for the analysis of remote sensing data based on EFs. The proposed approach is an EP composed of a sequence of thinning and thickening transformations applied to a grayscale image. The proposed approach is efficiently able to carry out a multilevel decomposition of the input image based on EFs. In addition, we have adopted a few new attributes, such as volume and height, for the first time in the remote sensing community. Furthermore, based on the proposed EP, the derivative of the EP was proposed in this paper. Then, the proposed approach was performed on two well-known panchromatic data sets: the Rome and Reykjavik data sets. The obtained results have been compared with one of the strongest approaches in the literature, i.e., APs, from different points of view, including classification accuracies, the complexity analysis, and the simplification and recognition rate. With respect to the experiments, the following promising points can be obtained: 1) EPs can significantly outperform APs in terms of classification accuracies due to their ability to preserve more regions and correspondences and to preserve the height of the extrema; 2) this method works naturally with the number of extrema, which seems to provide better results in terms of classification accuracies and decreases the burden of setting threshold values, which was a shortcoming for conventional APs; 3) the EP can be generated very fast and applied in only a few seconds; and, finally, 4) an EP is a more efficient tool than an AP in terms of simplification for

TABLE X
PERFORMANCE EVALUATION OF APs AND EPs (AS WELL AS MAPs AND MEPs) BASED ON DIFFERENT METRICS, NAMELY, CLASSIFICATION ACCURACIES, AUTOMATION, COMPLEXITY, AND SIMPLIFICATION FOR RECOGNITION. THE HIGHER THE NUMBER OF BULLETS, THE BETTER THE PERFORMANCE OF THE INDIVIDUAL

Techniques	AP	EP	MAP	MEP
Accuracy	●●	●●●	●●	●●●
Automation	●	●●●	●	●●●
Complexity	●●●	●●●	●●●	●●●
Simplification for Recognition	●●	●●●	●●	●●●

recognition. Furthermore, the proposed approach is extrema oriented, which makes the proposed approach fully automatic and data set distribution independent. In other words, there is no need to have any *a priori* information about the resolution, type of the sensor, or regions; and the proposed approach can automatically extract informative features for classification.

Table X summarizes the comparison between EPs and APs from different points of view. The higher the number of bullets, the better the performance of the individual.

In future, we intend to investigate the use of EP for other types of remote sensing data and evaluate the efficiency of different classifiers for the classification of features produced by EP. In addition, one might consider the effectiveness of different classifiers on APs and EPs and compare them from different perspectives, such as the number of input features and training samples. For example, in some papers, such as [45], SVM could lead to better classification accuracy than RF, whereas RF provided better results than SVM in terms of classification accuracy in other papers, such as [23] and [27].

REFERENCES

[1] J. A. Benediktsson and P. Ghamisi, *Spectral-Spatial Classification of Hyperspectral Remote Sensing Images*.   Boston, MA, USA: Artech House, 2015.
[2] P. Ghamisi, J. A. Benediktsson, and M. O. Ulfarsson, "Spectral–spatial classification of hyperspectral images based on hidden Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2565–2574, May 2014.
[3] 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.
[4] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Advances in spectral–spatial classification of hyperspectral images," *Proc. IEEE*, vol. 101, no. 3, pp. 652–675, Mar. 2013.
[5] P. Ghamisi, M. S. Couceiro, F. M. Martins, and J. A. Benediktsson, "Multilevel image segmentation approach for remote sensing images based on fractional-order Darwinian particle swarm optimization," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2382–2394, May 2014.
[6] P. Ghamisi, M. 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.

[7] J. Li, H. Zhang, and L. Zhang, "Supervised segmentation of very high resolution images by the use of extended morphological attribute profiles and a sparse transform," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 8, pp. 1409–1413, Aug. 2014.

[8] M. Pesaresi and J. A. Benediktsson, "A new approach for the morphological segmentation of high-resolution satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 2, pp. 309–320, Feb. 2001.

[9] M. Chini, N. Pierdicca, and W. Emery, "Exploiting SAR and VHR optical images to quantify damage caused by the 2003 Bam earthquake," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 1, pp. 145–152, Jan. 2009.

[10] H. Akcay and S. Aksoy, "Automatic detection of geospatial objects using multiple hierarchical segmentations," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 7, pp. 2097–2111, Jul. 2008.

[11] J. Chanussot, J. A. Benediktsson, and M. Fauvel, "Classification of remote sensing images from urban areas using a fuzzy possibilistic model," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 40–44, Jan. 2006.

[12] P. Ghamisi, M. S. Couceiro, and J. A. Benediktsson, "Classification of hyperspectral images with binary fractional order Darwinian PSO and random forests," *Proc. SPIE*, vol. 8892, pp. 1–8, 2013.

[13] P. Ghamisi, M. S. Couceiro, J. A. Benediktsson, and N. M. F. Ferreira, "An efficient method for segmentation of images based on fractional calculus and natural selection," *Expert Syst. Appl.*, vol. 39, no. 16, pp. 12 407–12 417, 2012.

[14] 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.

[15] 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.

[16] M. Dalla Mura, J. A. Benediktsson, and L. Bruzzone, "Modeling structural information for building extraction with morphological attribute filters," in *Proc. SPIE Eur. Remote Sens.*, Berlin, Germany, 2009, pp. 747703-1–747703-9.

[17] P. Ghamisi, J. A. Benediktsson, and S. Phinn, "Fusion of hyperspectral and lidar data in classification of urban areas," in *Proc. IGARSS*, Jul. 2014, pp. 181–184.

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

[19] 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. 5, pp. 342–346, Sep. 2014.

[20] P. Ghamisi, J. A. Benediktsson, G. Cavallaro, and A. Plaza, "Automatic framework for spectral–spatial classification based on supervised feature extraction and morphological attribute profiles," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2147–2160, Jun. 2014.

[21] 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. 2014.

[22] X. Huang *et al.*, "Multiple morphological profiles from multicomponent base images for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 12, pp. 4653–4669, Dec. 2014.

[23] P. R. Marpu, M. Pedergnana, M. Dalla 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 Fus.*, vol. 3, no. 3, pp. 269–298, 2012.

[24] R. Souza, L. Rittner, R. Machado, and R. Lotufo, "A comparison between extinction filters and attribute filters," in *Proc. ISMM*, Reykjavik, Iceland, May 27–29, 2015, pp. 63–74.

[25] C. Vachier, "Extinction value: A new measurement of persistence," in *Proc. IEEE Work. Nonlinear Sig. Image*, 1995, vol. 1, pp. 254–257.

[26] M. Grimaud, "New measure of contrast: The dynamics," in *Proc. SPIE*, 1992, vol. 1769, pp. 292–305.

[27] P. 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.

[28] J. Serra, "Connectivity on complete lattices," *J. Math. Imag. Vis.*, vol. 9, no. 3, pp. 231–251, 1998.

[29] 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.

[30] R. Jones, "Connected filtering and segmentation using component trees," *Comp. Vis. Image Und.*, vol. 75, pp. 215–228, Apr. 1999.

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

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

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

[34] P. Salembier and M. Wilkinson, "Connected operators," *IEEE Signal Process. Mag.*, vol. 26, no. 6, pp. 136–157, Nov. 2009.

[35] Y. Xu, T. Géraud, and L. Najman, "Morphological filtering in shape spaces: Applications using tree-based image representations," in *Proc. 21st Int. Conf. Pattern Recog.*, Tsukuba, Japan, Nov. 2012, pp. 485–488. [Online]. Available: https://hal.archives-ouvertes.fr/hal-00714847

[36] R. Souza, L. Rittner, R. Lotufo, and R. Machado, "An array-based node-oriented max-tree representation," in *Proc. ICIP*, Sep. 2015, pp. 3620–3624.

[37] J. A. Benediktsson, M. Pesaresi, and K. Arnason, "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.

[38] M. H. F. Wilkinson, H. Gao, W. H. Hesselink, J.-E. Jonker, and A. Meijster, "Concurrent computation of attribute filters on shared memory parallel machines," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 10, pp. 1800–1813, Oct. 2008.

[39] E. Carlinet and T. Geraud, "A comparative review of component tree computation algorithms," *IEEE Trans. Image Process*, vol. 23, no. 9, pp. 3885–3895, Sep. 2014.

[40] Y. Xu, E. Carlinet, T. Géraud, and L. Najman, "Efficient computation of attributes and saliency maps on tree-based image representations," in *Proc. 12th ISMM Appl. Signal Image Process.*, Reykjavik, Iceland, 2015, pp. 693–704. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-18720-4_58

[41] G. Cavallaro, M. D. Mura, J. A. Benediktsson, and L. Bruzzone, "A comparison of self-dual attribute profiles based on different filter rules for classification," in *Proc. IGARSS*, 2014, pp. 1265–1268.

[42] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[43] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," in *Proc. Brit. Mach. Vis. Conf.*, 2002, pp. 384–393.

[44] K. Mikolajczyk *et al.*, "A comparison of affine region detectors," *Int. J. Comp. Vis.*, vol. 65, p. 2005, 2005.

[45] S. Bernabe, P. R. Marpu, A. Plaza, M. D. 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, Jan. 2014.

Pedram Ghamisi (S'12–M'15) received the B.Sc. degree in civil (survey) engineering from Islamic Azad University, South Tehran, Tehran, Iran, the M.Sc. degree (with first-class honors) in remote sensing from K. N. Toosi University of Technology, Tehran, in 2012, and the Ph.D. degree in electrical and computer engineering from the University of Iceland, Reykjavik, Iceland, in 2015.

He has been a Postdoctoral Research Fellow with the University of Iceland. In 2015, he won the prestigious Alexander von Humboldt Fellowship and has started his work as a Postdoctoral Research Fellow in the Signal Processing in Earth Observation at Technische Universität München (TUM), Munich, Germany, since October, 2015. Since October 2015, he has been also a Researcher with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Weßling, Germany. His research interests are in remote sensing and image analysis, with a special focus on spectral and spatial techniques for hyperspectral image classification and the integration of LiDAR and hyperspectral data for land cover assessment.

Dr. Ghamisi was a recipient of the Best Researcher Award for M.Sc. students in K. N. Toosi University of Technology in the academic year 2010–2011. He was also a recipient of the IEEE Mikio Takagi Prize for winning the Student Paper Competition at the 2013 IEEE International Geoscience and Remote Sensing Symposium held in Melbourne, in July 2013, at the conference of almost 70 people.

Roberto Souza (S'10) received the Bachelor's degree (top of his class) in electrical engineering from the Federal University of Pará (UFPA), Pará, Brazil, and the M.Sc. degree in electrical engineering from the University of Campinas (UNICAMP), São Paulo, Brazil. He is currently working toward the Ph.D. degree at UNICAMP, doing a six-month exchange at the Calgary Image Processing and Analysis Centre, University of Calgary, Calgary, Canada.

As an undergraduate student, he also studied for six months as an exchange student at the Institut Polytechnique de Grenoble, Grenoble, France. As a graduate student, he had a three-month internship at the Medical Image Processing Group, University of Pennsylvania, Philadelphia, PA, USA. His research focus is mathematical morphology theory, algorithms, and applications, with the focus of the applications being medical image processing and developing semiautomated segmentation techniques. His other applications of interest are content based image retrieval, digital forensics, and image quality metrics.

Jon Atli Benediktsson (S'84–M'90–SM'99–F'04) received the Cand.Sci. degree from the University of Iceland, Reykjavik, Iceland, in 1984 and the M.S.E.E. and Ph.D. degrees from Purdue University, West Lafayette, IN, USA, in 1987 and 1990, respectively, all in electrical engineering.

From 2009 to 2015, he was the Pro Rector of Science and Academic Affairs and a Professor of electrical and computer engineering with the University of Iceland, where he became the President/Rector on July 1, 2015. He is a Cofounder of the biomedical start-up company Oxymap. His research interests are in remote sensing, biomedical analysis of signals, pattern recognition, image processing, and signal processing. He has published extensively in those fields.

Prof. Benediktsson is a Fellow of International Society for Optics and Photonics (SPIE) and a member of the 2014 IEEE Fellow Committee, the Association of Chartered Engineers in Iceland (VFI), and Societas Scinetiarum Islandica. He was the 2011–2012 President of the IEEE Geoscience and Remote Sensing Society (GRSS) and has been on the GRSS Administrative Committee since 2000. He was the Editor-in-Chief of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGRS) from 2003 to 2008 and has served as an Associate Editor of TGRS since 1999, the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS since 2003, and IEEE ACCESS since 2013. He is on the Editorial Board of the PROCEEDINGS OF THE IEEE and the International Editorial Board of the *International Journal of Image and Data Fusion* and was the Chairman of the Steering Committee of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (J-STARS) in 2007–2010. He was a recipient of the Stevan J. Kristof Award from Purdue University in 1991 as an outstanding graduate student in remote sensing. He was also the recipient of the Icelandic Research Council's Outstanding Young Researcher Award in 1997, the IEEE Third Millennium Medal in 2000, the yearly research award from the Engineering Research Institute of the University of Iceland in 2006, the Outstanding Service Award from the IEEE GRSS in 2007, and the IEEE/VFI Electrical Engineer of the Year Award in 2013. He was a corecipient of the University of Iceland's Technology Innovation Award in 2004, the 2012 IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING Paper Award, the IEEE GRSS Highest Impact Paper Award in 2013, and the International Journal of Image and Data Fusion Best Paper Award in 2014. He is also a member of Tau Beta Pi.

Xiao Xiang Zhu (S'10–M'12–SM'14) received the Bachelor's degree in space engineering from the National University of Defense Technology, Changsha, China, in 2006 and the M.Sc. degree, the Dr.-Ing. degree, and the "Habilitation" degree in the field of signal processing from the Technical University of Munich (TUM), Munich, Germany, in 2008, 2011, and 2013, respectively.

Since 2011, she has been a Scientist with the Remote Sensing Technology Institute, German Aerospace Center (DLR), Weßling, Germany, where she is the Head of the Team Signal Analysis. Since 2013, she has been also a Helmholtz Young Investigator Group Leader and appointed as a TUM Junior Fellow. In 2015, she was appointed as the Professor for Signal Processing in Earth Observation at TUM. She was a Guest Scientist or a Visiting Professor at the Italian National Research Council (CNR-IREA), Naples, Italy; Fudan University, Shanghai, China; The University of Tokyo, Tokyo, Japan; and the University of California, Los Angeles, CA, USA, in 2009, 2014, 2015, and 2016, respectively. Her main research interests are advanced interferometric SAR techniques such as high-dimensional tomographic SAR imaging and SqueeSAR; computer vision in remote sensing, including object reconstruction and multidimensional data visualization; and big data analysis in remote sensing and modern signal processing, including innovative algorithms such as sparse reconstruction, nonlocal means filter, robust estimation, and deep learning, with applications in the field of remote sensing such as multispectral/hyperspectral image analysis.

Dr. Zhu is an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.

Letícia Rittner (M'05) received the Bachelor's degree in electrical engineering and the M.Sc. and Ph.D. degrees from the University of Campinas, São Paulo, Brazil, in 1994, 2004, and 2009, respectively, with a six-month internship at the Montreal Neurological Institute, McGill University, Montréal, Canada.

She was a Postdoctoral Fellow with the School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. She is currently an Assistant Professor with the School of Electrical and Computer Engineering, University of Campinas.

Roberto A. Lotufo (M'90) received the B.S. degree from Instituto Tecnológico de Aeronáutica, Brazil, in 1978 and the Ph.D. degree in electrical engineering from the University of Bristol, Bristol, U.K., in 1990.

Since 1981, he has been with the University of Campinas (UNICAMP), São Paulo, Brazil, where he is currently a Full Professor in the School of Electrical and Computer Engineering. His principal research interests are in the areas of image processing and analysis, pattern recognition, and machine learning. Prof. Lotufo has authored or coauthored over 150 refereed international journal and full conference papers.

Prof. Lotufo was a recipient of the Innovation Personality in 2008 and the Zeferino Vaz Academic Recognition in 2011 from the University of Campinas.