Hyperspectral Data Classification Using Extended Extinction Profiles

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

Abstract—This paper proposes a new approach for the spectral-spatial classification of hyperspectral images, which is based on a novel extrema-oriented connected filtering technique, here entitled as *extended extinction profiles* (EEPs). The proposed approach progressively simplifies the first informative features extracted from hyperspectral data considering different attributes. Then, the classification approach is applied on two well-known hyperspectral data sets: Pavia University and Indian Pines, and compared with one of the most powerful filtering approaches in the literature, extended attribute profiles (EAPs). Results indicate that the proposed approach is able to efficiently extract spatial information for the classification of hyperspectral images automatically and swiftly. In addition, an array-based node-oriented max-tree representation was carried out to efficiently implement the proposed approach.

Index Terms—Extended multi-extinction profile, Hyperspectral data classification, support vector machines, random forests.

I. INTRODUCTION

The classification of hyperspectral images using both spectral and spatial information has become a vibrant topic of research recently. In order to efficiently extract spatial information, there is an extensive number of research works in the literature, based either on crisp or adaptive neighborhood systems. Among these approaches, the ones that are based on mathematical morphology have obtained great attention [1].

In [2], the concept of morphological transformations was considered to form the so-called morphological profiles (MPs). Then, in [3], the concept of MPs was successfully generalized to hyperspectral data leading to the extended morphological profiles (EMPs). Since then, EMPs and their modifications have been enormously used to extract existing spatial information from hyperspectral data. Although MPs and their modifications can produce accurate classification maps, their concepts suffer from a few shortcomings, such as: (i) the shape of SEs is fixed and (ii) structuring elements (SEs) cannot characterize information related to the gray-level characteristics of the regions.

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

J. A. Benediktsson is with the Faculty of Electrical and Computer Engineering, University of Iceland, 107 Reykjavik, Iceland.

Roberto Souza, Letícia Rittner and Roberto Lotufo are with the School of Electrical and Computer Engineering - UNICAMP, Brazil.

To overcome the above-mentioned shortcomings, morphological attribute profiles (APs) were introduced in [4]. The AP is the generalization of the MP, which provides a multilevel characterization of an image using the sequential application of morphological attribute filters (AFs). Although the AP has been recently introduced, there is a considerable number of contributions based on that. As discussed in [1, 4], APs are more flexible than MPs since APs process images, based on 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 on different characteristics, such as spatial relations to other connected components. In [5, 6], automatic frameworks have been proposed for the classification of hyperspectral data, which are able to accurately classify hyperspectral images in an acceptable CPU processing time. A comprehensive survey on APs and their capabilities for the classification of remote sensing data can be found in [1].

In [7], Ghamisi *et al.* proposed the concept of extinction profiles (EPs), based on extinction filters (EFs), to further improve the classification accuracies of the APs. EPs are extrema-oriented connected filters, which are automatic by nature and in this context, they address the main shortcoming of the conventional APs (i.e., the manual setting of threshold values). Unlike AFs, EFs preserve the height of the extrema kept. In some experiments conducted on benchmark gray scale images and panchromatic remote sensing data [7, 8], the capability of EFs and EPs have been demonstrated through experiments, which confirm that they are better alternatives than AFs in terms of simplification for recognition and obtained classification accuracies.

In this paper, the concept of EP is generalized for the classification of hyperspectral data sets, entitled as *extended extinction profile* (EEPs). EEPs simultaneously simplify the input image by discarding unimportant spatial details and preserves the geometrical characteristics of the other regions from the first informative features extracted by a feature extraction approach [e.g., Independent Component Analysis (ICA) or Principal Component Analysis (PCA)] on a hyperspectral data set. In addition, the proposed approach is fully automatic in the sense that it can adjust the filtering parameters, based on the number of extrema. The output of this step provides a few informative features, which can be fed to a classification approach, e.g., random forest (RF). RF and support vector machines (SVMs) are well-established classifiers in the hyperspectral community since they can handle high dimensional data with a limited number of training
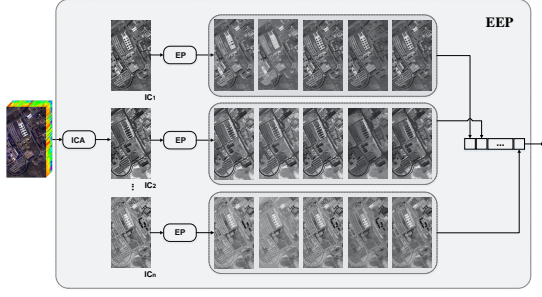
Fig. 1. EEP general architecture with ICA. The EPs are applied to the most informative independent components. Finally, EPs are concatenated to generate the EEP.

samples. The application of the above-mentioned filtering approaches decreases the nonlinearity of input images by simplifying unnecessary information. In this context, with respect to the studies reported in [1, 9], RF usually leads to higher classification accuracy than kernel SVMs in a faster way for the classification of APs. RF is composed of a collection of tree-like classifiers. Ideally, an RF classifier should be an independent and identically distributed randomization of weak learners. RF conducts a lot of individual decision trees, all of which are trained (grown) in order to tackle the same problem. A sample is assigned to the most frequent labels as determined by the individual trees [1]. Finally, the obtained results are compared with EAP in terms of classification accuracies on two hyperspectral data sets; Indian Pines and Pavia University. Results confirm the capability of EEPs for the classification of hyperspectral data.

The rest of the paper is organized as follows: Section II is devoted to the EEP methodology. Section III presents experimental results on two well-known data sets. Section IV wraps up the paper by providing the main concluding remarks.

II. METHODOLOGY

Fig. 1 shows the general idea of the proposed EEP. As can be seen, first, a feature extraction approach, such as ICA, is applied to the input data in order to avoid producing redundant features using EPs and address the so-called *curse of dimensionality*. Then, EP is applied to the most informative ICs/PCs to simultaneously extract spatial features. The produced features are concatenated into a stacked vector and classified by a RF. More details are given in the following subsections.

III. EXTINCTION FILTER

EFs are connected filters, which preserve the relevant image extrema. Relevance is measured by the concept of extinction values defined by Vachier [10]. EFs can be formally defined as follows: Let $Max(f) = \{M_1, M_2, ..., M_N\}$ denote the regional maxima of the image f. Each regional maxima M_i has an extinction value ϵ_i corresponding to the increasing attribute being analyzed. The extinction filter of f preserves the n maxima with highest extinction values, $EF^n(f)$, is given as follows:

$$EF^n(f) = R_g^\delta(f), \tag{1}$$

where $R_g^\delta(f)$ is the reconstruction by dilation [11] of the mask image g from marker image f. The mask image g is given by:

$$g = \max_{i=1}^{n}\{M_i'\}, \tag{2}$$

where $\max$ is the pixel-wise 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 using the max-tree structure [8]. The n maxima (max-tree leaves) with the highest extinction values for the attribute being analyzed are chosen. All max-tree nodes that are not in the paths from these leaves to the root are pruned.

EFs are idempotent. Unlike the usual contrast and size filters, such as the *hmax* [12] and the *area-open* [13] filters, the heights of the extrema kept in the image are completely preserved.

IV. EXTINCTION PROFILE AND ITS EXTENSION TO HYPERSPECTRAL DATA

In order to extract detailed information from an input hyperspectral image, instead of applying one filtering step, a sequence of filtering steps with progressively higher threshold values can be taken into consideration. In this way, the EPs can be produced. An EP is composed of a sequence of thinning and thickening transformations defined with a sequence of progressively stricter criteria. An EP for the input gray scale image, f, can be defined as follows:

$$EP(f) = \{\underbrace{\phi^{P_{\lambda_L}}(f), \phi^{P_{\lambda_{L-1}}}(f), \ldots, \phi^{P_{\lambda_1}}(f)}_{\text{thickening profile}}, \tag{3}$$
$$f, \underbrace{\gamma^{P_{\lambda_1}}(f), \ldots, \gamma^{P_{\lambda_{L-1}}}(f), \gamma^{P_{\lambda_L}}(f)}_{\text{thinning profile}}\},$$

with $P_\lambda : \{P_{\lambda_i}\}$ $(i = 1, \ldots, L)$ a set of L ordered predicates (i.e., $P_{\lambda_i} \subseteq P_{\lambda_k}$, $i \leq k$). For EPs, the number of extrema can be considered as the predicates.

In [7], it was shown that EPs are a more efficient alternative than APs with respect to simplification for recognition, since it is able to preserve more regions and correspondences found by affine region detectors. In addition, the parameters of EPs can be simply set in contrast to the parameters of APs since they are independent of the kind of attribute being used (e.g. area, volume,...), and are only based on the number of extrema [7]. On the other hand, the thresholds used by APs vary greatly according to the attribute being used as well as the data set being analyzed. Therefore, the thresholds are more difficult to be set. In other words, the main shortcoming of APs, which is related to the initialization of the threshold values, is being addressed by EPs [7].

In order to generalize the concept of EP and produce extended EP (EEP), which was originally proposed for gray scale images, to hyperspectral data, one possible way is based on performing a feature reduction approach like PCA or ICA and applying EPs on the most informative features [1]. In more detail, this approach is based on the reduction of the dimensionality of the image from $E \subseteq \mathbf{Z}^n$ to $E' \subseteq \mathbf{Z}^m$ ($m \leq n$) with a generic transformation $\Psi : E \to E'$

applied to an input image $\mathbf{f} = \{f_i\}_{i=1}^{n}$ (i.e., $\mathbf{g} = \Psi(\mathbf{f})$) where $\mathbf{g} = \{g_i\}_{i=1}^{m}$. Then, EP can be performed on the most informative features $\mathbf{g}_i$ $(i = 1, \ldots, m)$ of the transformed image, which can be mathematically given as follows:

$$EEP(\mathbf{g}) = \{EP(g_1), EP(g_2), \ldots, EP(g_m)\}. \quad (4)$$

In contrast with 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, extended multi-EP (EMEP) concatenates different EEPs (e.g., area, height, volume, diagonal of bounding box, and standard deviation on different extracted features) into a single stacked vector, which can be mathematically defined as follows:

$$EMEP(\mathbf{g}) = \left\{ EEP_{\mathcal{A}_1}(\mathbf{g}), EEP'_{\mathcal{A}_2}(\mathbf{g}), \ldots, EEP'_{\mathcal{A}_k}(\mathbf{g}) \right\}, \quad (5)$$

where $EEP_{\mathcal{A}_i}$ is an EEP built with a set of predicates evaluating the attribute $\mathcal{A}_i$ and $EEP' = EEP \backslash \{g_i\}_{i=1,\ldots,m}$ in order to avoid redundancy since the original components $\{g_i\}$ are present in each EEP. Due to the fact that different extinction attributes can extract complementary spatial information, the EMEP has a greater capability in extracting spatial information than a single EP. In addition, the computational cost of the EMEP and EEP are almost the same since the max-tree and min-tree construction, which are the most time consuming part of producing profiles, are done only once for each gray scale image (except for the standard deviation extinction attribute).

A. EEP Computational Complexity Analysis

The computational complexity of EEP is m times the complexity of computing EP, where m is the number of informative features kept by either ICA or PCA. The most time consuming part is on the max-tree and min-tree constructions, which require to compute the thickening and thinning profiles. Their complexity for a generic floating point structure is $O(NlogN)$, where N is the number of image pixels. A complete analysis of the max-tree construction complexity for different data types and different implementations is given in [14].

In our implementation, we use the array-based node-oriented max-tree representation proposed in [15]. This representation is very flexible, and for some attributes, like height, it reduces their computational complexity from $O(N)$ to $O(M)$, where M is the number of max-tree nodes. Also, the structure is suitable for parallel processing of the max-tree. Table I summarizes the usual complexities of each stage in the EEP computation using sequential algorithms and their respective number of occurrences throughout the profile computation. For a more complete discussion on the complexities of the max-tree construction, attributes computation and filtering, see [15].

V. EXPERIMENTAL RESULTS

A. Data Set Descriptions

Two widely used hyperspectral data, the AVIRIS Indian Pines and the ROSIS Pavia University, were used in the experiments. It should be noted that the standard set of training

TABLE I
COMPLEXITY ANALYSIS OF THE EEP. THE PARAMETER "s" REFERS TO THE NUMBER OF THRESHOLD VALUES IN THE PROFILE. THE PARAMETER m REPRESENTS THE NUMBER OF INFORMATIVE FEATURES KEPT AFTER PERFORMING A FEATURE REDUCTION APPROACH.

Operation	Complexity	# ocurrences
Max-tree construction	$O(NlogN)$	$2m$
Attribute computation	$O(N)$	$2m$
Extinction values computation	$O(M)$	$2m$
Filtering	$O(N)$	$2ms$

and test samples for the data sets have been used to make the results fully comparable with the literature. For detailed information about the data sets and their corresponding test and training samples, please see [1, 5].

B. Algorithm Setup and Discussion

A 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. In the experiments, **Raw** refers to a situation where the RF is applied to the input hyperspectral data sets. In this paper, a, v, h, bb, and std are area, volume, height, diagonal of the bounding box, and standard deviation attributes [7], respectively. In order to examine the ability of the proposed approach, we defined two scenarios.

In order to generate EEPs for the first scenario, the values of n used to generate the profile for different attributes are automatically given by $\lfloor \alpha^j \rfloor$ $\quad j = 0, 1, \ldots, s-1$. The total EP size is $2s+1$, since the original image is also included in the profile. The term above was determined experimentally. For instance, if an image is filtered using an EF set to preserve 1 extrema and another filter set to preserve 2 extrema, the difference between these two images will be higher than applying an EF set to preserve 1000 extrema and the other set to preserve 1001 extrema. That occurs because extrema with highest extinction values are the ones that contain most of the image structural information [8]. In summary, our reason to define the equation for selecting the values of n was to generate a profile with more images with few extrema, where most of the changes occur and has more information for the classifier to learn, but also keep some images with a higher number of extrema, where less changes occur, nonetheless they contain important information for the classifier to learn. We recommend to use an α between 2 and 5. These values were derived experimentally through a grid search through different values of α. In the experiments here, we used $\alpha = 3$, which obtained the best results in our experiments, and set $s = 7$. The profiles were computed considering the 4-connected connectivity rule.

For the second scenario, the proposed EEPs are compared with EAPs in terms of classification accuracies, while the parameters were set in favor of the EAPs. To do so, we only selected the attributes that are available in both our implementation and the ones published in [16]. In this context, the threshold values for EAPs are (i) $\lambda_a = \{100, 500, 1000, 5000\}$, (ii) $\lambda_{bb} = \{10, 25, 50, 100\}$, and (iii) $\lambda_{std} = \{20, 30, 40, 50\}$. For
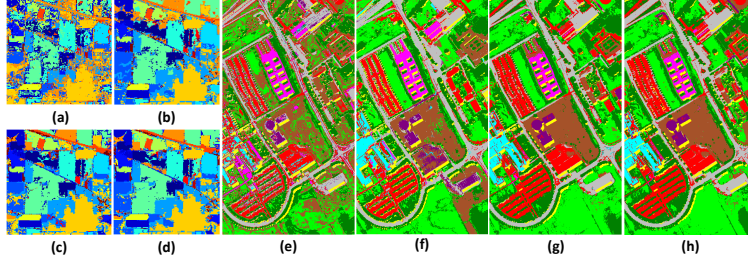
Fig. 2. Classification maps obtained on Indian Pines using (a) **Raw** (b) $\mathbf{EEP}_{all}$ using ICA (c) $\mathbf{EEP}_{bb}$ using PCA (d) $\mathbf{EEP}_{all}$ using PCA. Classification maps obtained on Pavia University using (e) **Raw** (f) $\mathbf{EEP}_a$ using PCA (g) $\mathbf{EEP}_a$ using ICA (h) $\mathbf{EEP}_{all}$ using ICA.

EEPs, s is simply set to 4. Since in [16], only Pavia data were considered, we compare EAPs and EEPs only for Pavia University using ICA.

C. Discussion

As can be seen in Tables II and III, the features produced by EEPs can significantly outperform the classification accuracies of **Raw**. For example, for the Pavia data, $\mathbf{EEP}_{all}$ obtained by ICA improves **Raw** by more than 24% in terms of overall accuracy. In the same way, for the Indian Pines, $\mathbf{EEP}_{all}$ obtained by PCA improves **Raw** by almost 23% in terms of overall accuracy.

The classification accuracies of EAPs obtained by ICA and PCA are published in [16]. The highest overall accuracies reported in [16] for Pavia University using PCA and ICA are 90.00% and 94.47%, respectively. However, in this paper, 91.28% and 95.97% are obtained for Pavia University using PCA and ICA, respectively. In [17], a few results on the use of Kernel PCA (KPCA) as the feature extraction approach are reported. The best result reported in [17] for Pavia University using the same set of training samples as ours is 94.09%, while in this paper the corresponding overall accuracy is 95.97%.

For Indian Pines, the methodology using PCA (92.92%) also improves the classification accuracies of EAPs using Discriminant Analysis Feature Extraction (DAFE) [5] with the overall accuracy of (91.13%) and Decision Boundary Feature Extraction (DBFE) with the overall accuracy of (89.56%) [5].

In the proposed approach, the higher number of extrema extracts more detail from the input data while the lower number of extrema considerably simplifies the input data. By considering different attributes, a diverse set of features can be obtained that are efficient for classification. Fig. 2 shows a few classification maps obtained by the proposed approach on Indian Pines and Pavia University.

Table IV shows the obtained classification accuracies for the second scenario. As can be seen, EEPs can significantly improve EAPs in terms of classification accuracy. The main reason is that EEPs preserve more regions and correspondences than EAPs, which is useful from the stand point of classification accuracy.

Fig. 3 displays processing times for different EPs computed on one component of the Pavia University data set. The image is of 610×340 pixels. The processing times were measured in a 2.7GHz 12-core Intel Xeon E5 with 64GB of RAM running OSX 10.9.5. The final processing time for EEP is proportional
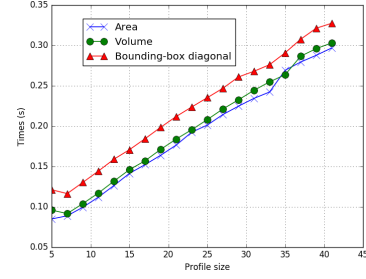

Fig. 3. Processing time for different EP attributes for one of the components of the hyperspectral data.

to the number of hyperspectral components kept for sequential implementations.

VI. CONCLUSION

In this letter, a novel approach was proposed for the classification of hyperspectral data, based on EFs, as extended extinction profiles (EEPs). An EEP is composed of a sequence of thinning and thickening transformations applied to the first informative features extracted by a feature exaction approach. The proposed approach is efficiently able to carry out a multi-level decomposition of the input image based on EFs. Then, the proposed approach was applied to two well-known hyperspectral data sets; Pavia University and Indian Pines and compared with extended attribute profiles (EAPs). Results confirm that the proposed approach can precisely classify hyperspectral data sets within a short period of time, and can improve EAPs in terms of classification accuracies. In addition, the proposed approach is automatic and there is no need to initialize the threshold values *a priori*, which is considered as an important advantage over traditional APs whose threshold values need to be predefined by users. Furthermore, the proposed approach is fast.

VII. ACKNOWLEDGMENT

REFERENCES

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

TABLE II
PAVIA: CLASSIFICATION ACCURACIES OBTAINED BY RF ON DIFFERENT EXTINCTION ATTRIBUTES ON THE FIRST THREE FEATURES PRODUCED BY PCA AND ICA. AA AND OA ARE IN PERCENTAGE.

	Raw	ICA						PCA					
		EEP_a	EEP_v	EEP_h	EEP_{bb}	EEP_{std}	EEP_{all}	EEP_a	EEP_v	EEP_h	EEP_{bb}	EEP_{std}	EEP_{all}
AA	82.25	96.38	95.35	91.62	95.02	94.31	**97.09**	93.03	93.48	84.49	91.06	86.25	92.74
OA	71.64	95.21	94.48	90.4	90.06	94.72	**95.97**	90.79	92.79	80.8	87.31	83.45	91.28
K	0.6512	0.9375	0.928	0.8732	0.8727	0.9308	**0.9474**	0.8807	0.9056	0.7526	0.837	0.7865	0.8866

TABLE III
INDIAN PINES: CLASSIFICATION ACCURACIES OBTAINED BY RF ON DIFFERENT EXTINCTION ATTRIBUTES ON THE FIRST THREE FEATURES PRODUCED BY PCA AND ICA. AA AND OA ARE IN PERCENTAGE.

	Raw	ICA						PCA					
		EEP_a	EEP_v	EEP_h	EEP_{bb}	EEP_{std}	EEP_{all}	EEP_a	EEP_v	EEP_h	EEP_{bb}	EEP_{std}	EEP_{all}
AA	76.71	92.47	92.49	91.53	92.49	87.7	92.16	94.2	94.66	93.64	95.31	94.14	**95.49**
OA	69.97	89.57	89.3	89.18	90.32	89.14	90.59	90.83	91.11	89.59	92.45	89.74	**92.92**
K	0.6600	0.8808	0.878	0.8764	0.8894	0.8759	0.8900	0.8951	0.8984	0.8812	0.9136	0.8828	**0.9147**

TABLE IV
PAVIA UNIVERSITY: COMPARISON BETWEEN EAPs AND EEPs.

	EEP_{bb}	EAP_{bb}	EEP_a	EAP_a	EEP_{std}	EAP_{std}	EEP_{all}	EAP_{all}
AA	96.53	88.37	96.71	91.29	92.94	91.61	96.68	93.59
OA	94.74	82.19	96.23	85.25	94.3	92.07	96.13	92.16
K	0.9316	0.7711	0.9506	0.81	0.9254	0.8962	0.9497	0.8974

[2] 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, 2001.

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

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

[5] 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 Jour. Sel. Top. App. Earth Obs. Remote Sens.*, vol. 7, no. 6, pp. 2147–2160, June 2014.

[6] P. Ghamisi, J. A. Benediktsson, and J. R. Sveinsson, "Automatic spectral-spatial classification framework based on attribute profiles and supervised feature extraction," *IEEE Trans. Geoscience and Remote Sens.*, vol. 52, no. 5, pp. 342–346, 2014.

[7] P. Ghamisi, R. Souza, J. A. Beneiktsson, X. X. Zhu, L. Rittner, and R. Lotufo, "Extinction profiles for the classification of remote sensing data," *IEEE Trans. Geos. Remote Sens.*, In press.

[8] R. Souza, L. Rittner, R. Machado, and R. Lotufo, "A comparison between extinction filters and attribute filters," in *ISMM'15, Reykjavik, Iceland, May 27-29, 2015. Proceedings*, 2015, pp. 63–74.

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

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

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

[12] L. Vincent, "Morphological area openings and closings for grey-scale images," in *Shape in Picture*, Y.-L. O, A. Toet, D. Foster, H. Heijmans, and P. Meer, Eds. Springer Berlin Heidelberg, 1994, vol. 126, pp. 197–208.

[13] P. Salembier and A. Oliveras, "Practical extensions of connected operators," in *Mathematical Morphology and its Applications to Image and Signal Processing*, ser. Computational Imaging and Vision. Springer US, 1996, vol. 5, pp. 97–110.

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

[15] R. Souza, L. Rittner, R. Lotufo, and R. Machado, "An array-based node-oriented max-tree representation," in *ICIP'15*, Sept 2015, pp. 3620–3624.

[16] M. Dalla Mura, A. Villa, J. A. Benediktsson, J. Chanussot, and L. Bruzzone, "Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 3, pp. 542–546, 2011.

[17] 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 Geos. Remote Sens. Lett.*, vol. 11, no. 1, pp. 288–292, 2014.