Vector attribute profiles for hyperspectral image classification

Erchan Aptoula, Mauro Dalla Mura, *Member, IEEE,* and Sébastien Lefèvre

***Abstract*—Morphological attribute profiles are among the most prominent spectral-spatial pixel description methods. They are efficient, effective and highly customizable multi-scale tools based on hierarchical representations of a scalar input image. Their application to multivariate images in general, and hyperspectral images in particular, has been so far conducted using the *marginal strategy*, i.e. by processing each image band (eventually obtained through a dimension reduction technique) independently. In this paper, we investigate the alternative *vector strategy*, which consists in processing the available image bands simultaneously. The vector strategy is based on a vector ordering relation that leads to the computation of a single max- and min-tree per hyperspectral dataset, from which attribute profiles can then be computed as usual. We explore known vector ordering relations for constructing such max-trees and subsequently *vector attribute profiles*, and introduce a combination of marginal and vector strategies. We provide an experimental comparison of these approaches in the context of hyperspectral classification with common datasets, where the proposed approach outperforms the widely used marginal strategy.**

***Index Terms*—Morphological attribute profiles, multivariate morphology, vector ordering, hyperspectral images.**

I. Introduction

THE advent of satellite sensors with very high spatial and spectral resolutions has had a profound effect on the remote sensing community by providing us with datasets richer in information than ever before. Although this unprecedented level of detail has paved the way for several new applications, it has also scaled the problem of classification to a more challenging level, as images are now more heterogeneous and contain highly complex spatial structures and inter-pixel relations, thus making effective spectral-spatial content description a necessity.

Given its rigorous non-linear mathematical foundation and shape-oriented spatial tools, it is thus no surprise that mathematical morphology has become largely popular with remote sensing applications. It is employed in a rich variety of tasks such as building detection [1], urban area detection [2], forest mapping [3], road network extraction [4], content-based retrieval [5], rubble detection [6], segmentation [7]–[10], change detection [11] and spectral unmixing [12]. In particular, the largest part of the morphology oriented efforts in the remote sensing context, have focused on the topic of spatial-spectral pixel description and classification; with the most widely applied and researched tool being morphological profiles (MP) [13]–[16].

Indeed MP have been one of the most prominent spectral-spatial pixel description methods of the last decade, with numerous variations, the latest of which have been attribute profiles (AP) [17]. APs are a more efficient generalization of MPs, acting directly on connected components instead of pixels, thus enabling object based image analysis. APs have effectively replaced MPs, by removing the structuring element shape restriction, and by further enabling the description of image components through arbitrary parametric features, thus leading to more flexible, complete and accurate content representations. For a very recent survey on AP based spectral-spatial classification the reader is referred to [18].

An AP is indeed a multi-scale, efficient and effective tool constructed from hierarchical representations of its input image [17]. It is based on a min- and a max-tree, also referred to as component trees [19]. The inclusion tree (or tree of shapes) [20], has been also considered, leading to a self dual AP [21]. Yet, when it comes to dealing with multivalued images, as with all morphological tools, complications inevitably arise. Specifically, component trees rely on a total order among the image pixel values, a concept which is absent among multivalued (vector) pixel values. So far, in order to deal with this issue, all instances of APs with multivalued data have been employing the same technique as MPs, the so-called *marginal strategy*; i.e. band-wise independent processing [16]. It basically consists in constructing one tree per band using the usual scalar order among pixel values, from which then an AP is computed, followed finally by their concatenation in order to form the pixel feature vector [22].

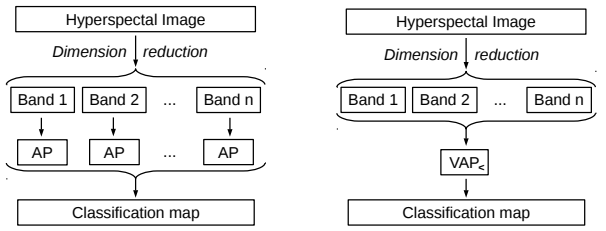
Although this strategy is feasible when dealing with multispectral images (i.e. up to about ten spectral channels), it is impractical with hyperspectral images (i.e. with hundreds of spectral channels). In this case, the conventionally used approach relies on a dimension reduction step in which the initial number of bands is reduced to few according to some criterion (e.g. maximizing the variance of the data or the discriminability of the classes) [23]. The tree representation can then be built analogously on the retained components.

However, it is well known from the context of color morphological analysis that there exists an alternative, called *vector processing strategy*, that can handle all available image bands simultaneously [24]. In this context, it manifests itself through the use of a vector ordering relation, that enables one to compute a single max-tree (and not multiple as in the case of marginal strategy) from a hyperspectral dataset. Such trees are straightforward to construct [25], by working on the

E. Aptoula is with the Department of Computer Engineering, Okan University, Istanbul, 34959 Turkey e-mail: erchan.aptoula@okan.edu.tr.

M. Dalla Mura is with the GIPSA-Laboratory, Departement Image and Signal, Grenoble-INP, Saint Martin d'Heres Cedex F-38402, France.

S. Lefèvre is with the Université de Bretagne-Sud, UMR 6074, IRISA, Vannes 56000, France.



(a) Marginal strategy (b) Vector strategy

Fig. 1. Outline of the marginal and vector strategies for attribute profile computation from hyperspectral images.

corresponding rank image of the input multivalued data. The main question rather is which vector ordering relation among the tens of options is most appropriate for computing AP and subsequently classifying hyperspectral remote sensing images, and furthermore, whether there is a superior alternative to the widely used marginal strategy.

This paper focuses on both of the aforementioned questions. Specifically, our contributions are first the investigation of AP defined using a vector strategy, then the definition of a new approach combining the vector and marginal strategies, and an experimental comparison of AP constructed using the aforementioned strategies, with two common real hyperspectral remote sensing datasets and one multispectral, where the proposed method is shown to outperform the marginal strategy that is currently used exclusively in the state-of-the-art.

In the sequel of this paper, following an overview of related work on morphological pixel description of hyperspectral images (Section II), we elaborate on vector AP and introduce a new combined vector-marginal strategy for profile construction (Section III). Next, we present and discuss the results of our classification experiments (Section IV), while Section V is devoted to concluding remarks.

II. Related work

Soon after the initial introduction more than a decade ago of morphological profiles [13], as a tool designed to process grayscale panchromatic data, they were extended and have been widely used on multivalued (and hyperspectral in particular) images.

One of the initial attempts by Dell'Acqua et al. [26] oriented towards the extension of MP to hyperspectral images, consisted in simply reducing the input image into grayscale. Specifically, the hundreds of available channels were reduced to a single one, containing however the majority of variational information by means of a principal component analysis (PCA). Since such a radical dimension reduction led to an inevitable and significant loss of information, Benediktsson et al. [16], proposed later on extended morphological profiles. In particular, they suggested the use of more than one principal components (PC). In more detail, the pixel descriptions computed independently (or marginally) from each PC, were then concatenated to form the final pixel feature vector (Fig. 1a). In the years that followed, a lot of effort has been put in optimizing the dimension reduction step, with several approaches having been investigated; for a detailed analysis of this aspect the reader is referred to [23].

Although the widely used marginal approach, outlined in Fig. 1a, has the important advantage of being straightforward, since it continues to rely on existing grayscale operators, it does have significant drawbacks as well. In particular, the independent processing of each band multiplies the computational load, ignores any eventual inter-band correlation related information, and in the case of MP, the filters can output vectors not included in the input image. That is why, a second strategy has emerged.

The alternative of marginal processing, which does not suffer from the aforementioned disadvantages, consists in processing all bands simultaneously by means of the so-called vector processing strategy (Fig. 1b), which has been extensively studied in the context of color morphology [24]. The core idea is surprisingly simple; as mathematical morphology is based on complete lattice theory, one can replace the underlying ordering relation (e.g. the marginal ordering relation in the case of marginal processing), with any algebraically valid alternative, and still obtain theoretically correct results with various properties. For an in-depth account of multivariate morphology the reader is referred to [27].

The first attempt of this type in hyperspectral images was made by Plaza et al. [14] who investigated an approach based on cumulative spectral angular distances. Li and Hu [28] on the other hand, experimented with using an ordering relation based on the Euclidean norm of the vector pixel values. Moreover, Aptoula and Lefèvre [24] provided a comparative study of various ordering relations in the context of multispectral image classification using MP, where it was shown that although alternative ordering relations lead to pixel descriptions with diverse class-specific performances, the marginal approach still outperforms overall. Velasco and Angulo [29], have also worked on this topic and proposed a supervised ordering relation, constructed through machine learning methods from user provided training pixel sets, though with no applications in pixel description or classification.

More recently, Courty et al. [30], proposed a method based on the PerTurbo classification algorithm, that leads to class-specific ordering relations, while in [31] an ordering relation exploiting the end-member vectors of a given hyperspectral image has been introduced; although it has been shown to lead to classification performances that can catch up with the marginal approach, it still requires a precise computation of end-member vectors.

As far as attribute profiles are concerned, so far all reported instances of their use with hyperspectral images, have been based on the marginal strategy [18], [32].

III. Vector attribute profiles

In this section, we will first recall the principles behind AP, and then explain how to construct a single max-tree from a multivariate image and finally introduce a new approach combining the advantages of marginal and vector strategies.

For a detailed account of attribute filters and profiles the reader is referred to [17], [33].

A. Background

APs are multiscale image description tools, constructed similarly to MPs, by applying successively a morphological filter along a variety of settings. Specifically, APs rely on morphological attribute filters (AFs) [33]; which belong to the class of connected morphological filters. As such, AF and AP revolve around the core concept of connectivity, and deal directly with connected components (CCs) instead of pixels.

More formally, given a grayscale image $f : E \to \mathbb{Z}, E \subseteq \mathbb{Z}^2$, its upper-level sets are defined as $\{f \geq t\}$ with $t \in \mathbb{Z}$ (resp. lower-level sets $\{f \leq t\}$), i.e. the set of images obtained by thresholding an image at all possible values of their pixels. The connected components ($CC \subseteq E$) composing the upper or lower level sets are referred to as peak components. AFs are applied to these peak components, using a predefined logical predicate T_κ^α consisting of comparing the attribute α computed on CC against a threshold κ; e.g. T_{300}^{area}: "is the area of CC larger than 300 pixels?". Depending on the binary outcome of $T_{300}^{area}(CC)$, the connected component is either preserved or removed from the image. An AF's output is computed by processing all the connected components present in the input image, thus evaluating the underlying predicate for all of them.

Subsequently, an AP can be straightforwardly constructed using a sequence of AFs (often attribute thinnings and thickenings [33]), that are applied to the input image using a set of ordered logical predicates. More precisely, given a predicate T and a collection of L thresholds $\{\kappa_i\}_{1 \leq i \leq L}$ let γ^{κ_i} and ϕ^{κ_i} denote respectively the attribute thinnings and thickenings employing them. In which case the AP of a grayscale image f would become:

$$AP(f) = \{\phi^{\kappa_L}(f), \phi^{\kappa_{L-1}}(f) \ldots, \phi^{\kappa_1}(f), f, \gamma^{\kappa_1}(f), \ldots, \gamma^{\kappa_{L-1}}(f), \gamma^{\kappa_L}(f)\} \quad (1)$$

Thus a pixel p of an image f can be characterized using the values it obtains across this sequential filtering process.

As far as the extension of AP to a hyperspectral image $\mathbf{f} : E \to \mathbb{Z}^r, r > 1$ is concerned, the widely (and exclusively) encountered marginal strategy consists in first reducing the number of spectral dimensions (from r to n, $n << r$) through a variety of methods [18], and then in computing independently the AP of each resulting image band, that are finally concatenated in order to form the so-called extended attribute profile (EAP):

$$EAP(\mathbf{f}) = \{AP(band_1), AP(band_2), \ldots, AP(band_n)\} \quad (2)$$

Here $band_i$ refers to an image component after dimension reduction, but it might equivalently denote an actual spectral band of the input image if no reduction were performed, still leading to an EAP constructed with the marginal approach.

Although AFs have been part of the morphological toolbox for almost two decades, the reasons for their only recent popularity have been mostly computational, in addition to their flexibility. Efficient implementations of AFs have become possible thanks to a tree based image representation [34].

In particular, according to that representation, the connected components of a given grayscale image are represented in a max-tree (alternatively min-tree) structure through their inclusion relations. Thus a grayscale image can be fully represented as a rooted tree of peak components. Its advantage stems from the implementation possibility of filtering in the form of node or branch removal from the tree representing the input image. This becomes especially interesting for the computation of AP since each tree needs to be computed only once and then multiple filtering outputs can be derived easily from it.

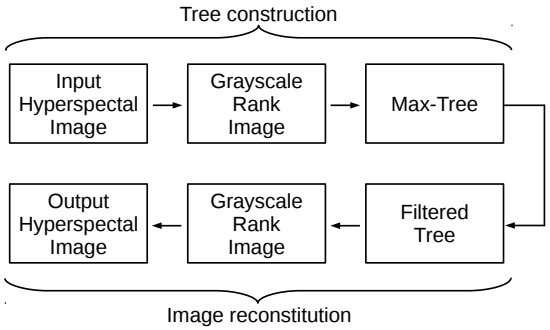


Fig. 2. Processing flow of the vector strategy for vector attribute filtering.

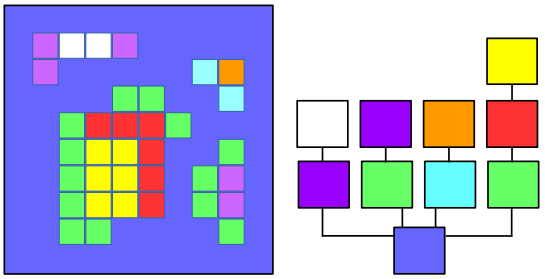


Fig. 3. A max-tree example constructed using the lexicographical ordering relation in the RGB color space (blue $<_L$ green $<_L$ cyan $<_L$ red $<_L$ magenta $<_L$ orange $<_L$ yellow $<_L$ white).

B. Vector strategy

The vector strategy (Fig. 1b) involves employing all available image components simultaneously. In detail, attribute filters, as well as the AP and EAP that derive from them, operate on connected components, in other words on sets of pixels coordinates. Consequently, they are fully independent from whether those coordinates within the CC represent scalar or vector pixel values. Thus, the only practical requirement for defining AP based on a vector strategy is being able to compute the tree representation of a multivariate image. In the following we will refer to a max-tree, but the reasoning clearly applies on the other hierarchical representations for which an underlying vector ordering between pixel values is required as well.

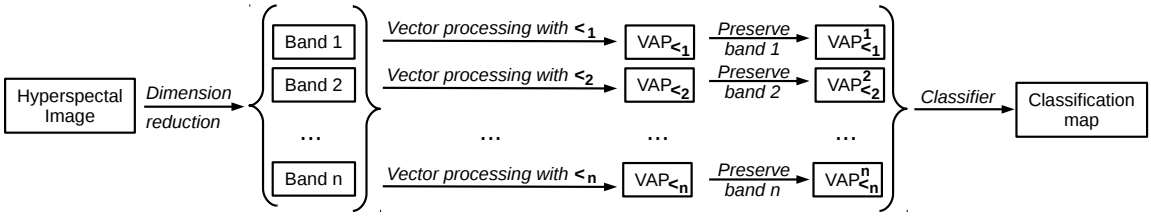


Fig. 4. Outline of the proposed strategy for attribute profile computation from hyperspectral images.

The hierarchy of the levels of a max-tree, as presented in [34], is determined by the natural order of scalar pixel values (0 to 255 in the case of 8-bit grayscale images). Hence, the root of the tree denotes the entire image coordinate space, with the lowest possible pixel value, while the leaves of the tree represent the connected components with the highest pixel values. Of course, as there is no universally accepted vector ordering method, the computation of a max-tree from multivariate images becomes evidently problematic. If however we equip with such a total ordering relation, e.g. lexicographical ordering:

$$\begin{aligned}\forall\, \mathbf{v}, \mathbf{v}' \in \mathbb{Z}^n,\ \ \mathbf{v} <_L \mathbf{v}' \Leftrightarrow \exists\, i \in \{1, \ldots, n\}\,,\\ (\forall\, j < i,\ v_j = v'_j)\ \text{ and }\ (v_i < v'_i)\end{aligned} \tag{3}$$

or Euclidean norm ($\|\cdot\|$) based ordering:

$$\begin{aligned}\forall\, \mathbf{v}, \mathbf{v}' \in \mathbb{Z}^n,\ \ \mathbf{v} <_E \mathbf{v}' \Leftrightarrow \|\mathbf{v}\| < \|\mathbf{v}'\| \ \text{ or }\\ (\|\mathbf{v}\| = \|\mathbf{v}'\| \ \text{ and } \mathbf{v} <_L \mathbf{v}')\end{aligned} \tag{4}$$

in that case one can trivially transform any multivariate image $\mathbf{f} : E \to \mathbb{Z}^n$ into a grayscale image $g : E \to \mathbb{Z}$ where $\forall p \in E, g(p) = \mathrm{rank}_\prec(\mathbf{f}(p))$, with $\mathrm{rank}_\prec$ denoting the integer rank of a vector according to the ordering relation $\prec$. This approach has been already successfully applied to computing multivariate quasi-flat zones [35]. Consequently, one can straightforwardly then apply the usual max-tree construction algorithm, hence leading to a tree containing the connected components not just of a single band, but of the entire multivariate input image (Fig. 2). After attribute filtering, the filtered multivariate image can be trivially and efficiently reconstituted using a reverse rank lookup. A visual example of using the lexicographical ordering for constructing a color max tree in the RGB color space is provided in Fig. 3.

Thus, given a multivariate image $\mathbf{f} : E \to \mathbb{Z}^n, n > 1$ and a total ordering relation $\prec$ such as Eq. (3) or (4), one can apply vector attribute thinnings and thickenings on it ($\boldsymbol{\gamma}^{\kappa_i}$ and $\boldsymbol{\phi}^{\kappa_i}$), with the only exception being that the input and output trees represent multivariate images. Hence, vector attribute profiles ($VAP_\prec$) can be straightforwardly defined as:

$$\begin{aligned}VAP_\prec(\mathbf{f}) = \{\boldsymbol{\phi}^{\kappa_L}(\mathbf{f}), \boldsymbol{\phi}^{\kappa_{L-1}}(\mathbf{f}) \ldots, \boldsymbol{\phi}^{\kappa_1}(\mathbf{f}), \mathbf{f},\\ \boldsymbol{\gamma}^{\kappa_1}(\mathbf{f}), \ldots, \boldsymbol{\gamma}^{\kappa_{L-1}}(\mathbf{f}), \boldsymbol{\gamma}^{\kappa_L}(\mathbf{f})\}\end{aligned} \tag{5}$$

i.e. a sequence of $2L + 1$ multivariate images, each one with the same dimension as $\mathbf{f}$.

C. Proposed strategy

Although the debate between marginal and vector strategies has been long going in the field of color morphology [24], quantitatively speaking, the vector processing strategy has seldom outperformed the marginal approach in an actual image description task [24], [36].

TABLE I
SUMMARY OF THE QUALITIES OF THE MARGINAL, VECTOR AND PROPOSED APPROACHES

Marginal strategy
+ one max-tree per band; band specific description - oblivious to what happens in adjacent bands - data is processed n times
Vector strategy
+ takes into account information from adjacent bands + data is processed once - one max-tree per dataset; i.e. limited data representation
Proposed strategy (combine the two)
+ one max-tree per band, computed using a vector ordering + each tree takes into account adjacent bands w.r.t. arbitrary weights - data is processed n times

This has motivated us to explore a third alternative combining the best qualities of each (Table I). To explain, given a hyperspectral image of n bands (and after an eventual dimension reduction step), we propose to process the data n times, just as in the case of the marginal approach, but at each iteration we process not a single band, but the entire hyperspectral dataset, using a distinct vector ordering $\prec_i$ (Fig. 4).

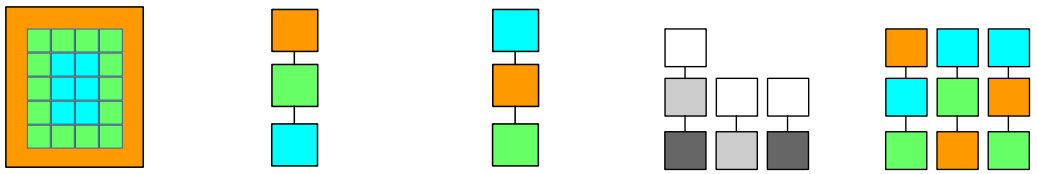
Formally, let $b_i : \mathbb{Z}^n \to \mathbb{Z}$ be the mapping returning the i-th grayscale band of a multivariate image $\mathbf{f} : E \to \mathbb{Z}^n$ and start by defining a set $\{\prec_i\}_{1 \le i \le n}$ of n distinct weighted norm based Euclidean ordering relations:

$$\begin{aligned}\forall\, \mathbf{v}, \mathbf{v}' \in \mathbb{Z}^n,\ \ \mathbf{v} \prec_i \mathbf{v}' \Leftrightarrow \|\mathbf{v}\|_i < \|\mathbf{v}'\|_i \ \text{ or }\\ (\|\mathbf{v}\|_i = \|\mathbf{v}'\|_i \ \text{ and } \mathbf{v} <_L \mathbf{v}')\end{aligned} \tag{6}$$

where:

$$\|\mathbf{v}\|_i = \sqrt{\sum_{j=1}^{n} v_j w_{ij}} \tag{7}$$

with $w_{ij} \in W = [0, 1]^{n \times n}$ a set of weights. Then, given a multivariate image $\mathbf{f}$, we compute for each i-th band, a vector attribute profile $VAP_{\prec_i}$ using its corresponding ordering relation $\prec_i$, by first applying Eq. (5), and then extracting from



(a) Color image (b) Lexicographical (c) Euclidean norm (d) Marginal (e) Weighted Euclidean

Fig. 5. From left to right, a color image with three distinct colors: $c_1 = [255, 150, 100]$ orange, $c_2 = [150, 255, 100]$ green, $c_3 = [100, 255, 255]$ cyan, and the resulting max-trees (rooted at the bottom), for various ordering relations: lexicographical ($<_L$), Euclidean Norm ($<_E$), Marginal and Weighted euclidean with the weight matrix W_2 of Eq. (10).

each resulting image its i-th band:

$$VAP^i_{\prec_i}(\mathbf{f}) = \{b_i(h) \,|\, \forall\, h \in VAP_{\prec_i}(\mathbf{f})\} \tag{8}$$

and once we repeat this for every band, we obtain the final weighted vector attribute profile:

$$VAP_W(\mathbf{f}) = \left\{VAP^1_{\prec_1}(\mathbf{f}), VAP^2_{\prec_2}(\mathbf{f}), \ldots, VAP^n_{\prec_n}(\mathbf{f})\right\} \tag{9}$$

Thus, VAP_W will have the same number of grayscale components as EAP and $VAP_\prec$ (i.e. $(2L+1) \times n$).

In order to better illustrate the interest of our approach, let us consider the simple scenario of a multivariate image with only three bands and the following three distinct weight sets:

$$W_1 = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} \quad W_2 = \begin{vmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \end{vmatrix} \quad W_3 = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} \tag{10}$$

Basically, the j-th column of the i-th row of the weight matrix, represents the relative contribution of the j-th band when processing the i-th band. This gives us a very high level of flexibility for fine tuning the behavior of our approach.

To explain, in the case of using equal weights (W_1), $\prec_1$, $\prec_2$ and $\prec_3$ are identical to the Euclidean norm based ordering $<_E$, thus VAP_{W_1} becomes equal to $VAP_{<_E}$. In other words, all bands contribute equally to the resulting profile. On the other extreme, we have the imbalanced W_3, where for each band's profile computation we employ only that band while other bands do not contribute at all, thus rendering VAP_{W_3} identical to the marginal approach (EAP). In-between the two, there is a more balanced weight matrix W_2, where the processing focuses primarily on one band but without ignoring the others. For a more visual illustration, Fig. 5 shows a RGB color image with three distinct colors, along with the resulting max-trees for various ordering relations.

D. Weight selection

Obviously, the selection of the weight matrix is of prime importance. As far as the weight matrix of Eq. (9) is concerned, it can modify the behavior of VAP_W from acting exactly as the marginal strategy all the way to acting identically to $VAP_{<_E}$. The general idea, when selecting the relative weights of each band (i.e. every row of W in Eq. (7)) is to set them according to the contribution potential of that band, in other words proportionally to the amount of useful information in that band; an excessively weighted noisy band is bound to corrupt results, while an under-weighted useful band would lead to suboptimal scores.

Of course an automatic and unsupervised weight computation system would be the ideal solution, however the design of such a method is beyond the scope of this paper. Nevertheless we would still like to present some directions that we consider as promising in this regard.

- Supervised weights: Similarly to the work of Velasco and Angulo [29], a straightforward approach for setting the weight matrix could be to employ a small set of training pixels, labeled by an expert. Their labels could be used for evaluating various weight matrices in order to determine the optimal one, in terms of classification performance. This could be achieved by starting with a predetermined sub-optimal weight matrix, and then feeding it to any standard optimization solution, such as genetic algorithms, in order to determine the optimal one.
- Eigenvalue based weights: an alternative unsupervised solution could be to exploit the eigenvalues of the bands of the image under study, after applying a PCA. The weight of each band could be set proportionally to its eigenvalue. Although fully unsupervised, this solution would be of interest only in cases where a dimension reduction step occurs.
- Statistically set weights: a further solution for setting the weights independently of any dimension reduction stage, could be to exploit the statistical properties of the bands under study. Since we are interested in the amount of useful information that they contain, one could measure their variance, or better yet their entropy, and then set the weight of each band proportionally to them.
- Morphologically set weights: another approach in this context, could be to quantify the usefulness of every image band not by measuring a global property such as entropy, but more specifically by measuring the amount of spatial structures and irregularities in the spatial domain of the image. This property can be quantified for instance through the application of tophat (residue of morphological opening) and bothat (residue of morphological closing) operators. Consequently, the weights could be set proportionally to the amount of image volume corre-

sponding to the respective tophat and bothat applications of that band. In fact, with an intelligently chosen structuring element, one could even measure the presence or not of particular objects of interest.

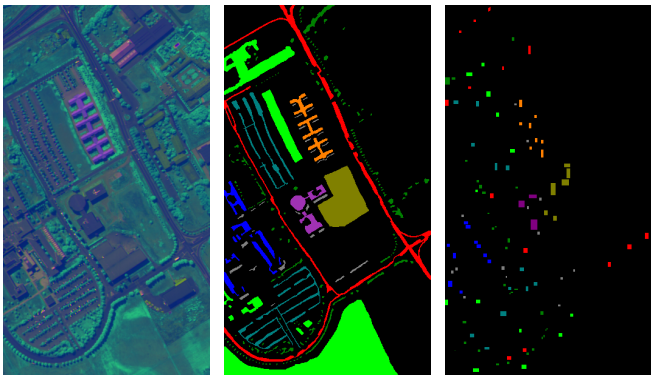
Let us now study the classification performance of EAP, VAP and VAP_W with two common hyperspectral datasets and one multispectral dataset, to see whether we can finally find a suitable alternative to the widely and exclusively used marginal strategy for attribute profile computation.

IV. Experiments

The main goal of the experiments presented in this section is to test the hypothesis that a combination of marginal and vector strategies for AP computation from hyperspectral images, as defined in Eq. (9), can lead to superior classification results w.r.t. their marginal counterpart. The tested approaches are EAP (Eq. (2)), $VAP_{<_L}$ & $VAP_{<_E}$ (Eq. (5)) and the proposed strategy VAP_W (Eq. (9)). Our experiments include two hyperspectral and one multispectral datasets, two geometric attributes (one increasing and one not-increasing), and study the effect of the number of image components as well as the effect of various attribute threshold sets.

A. Setup

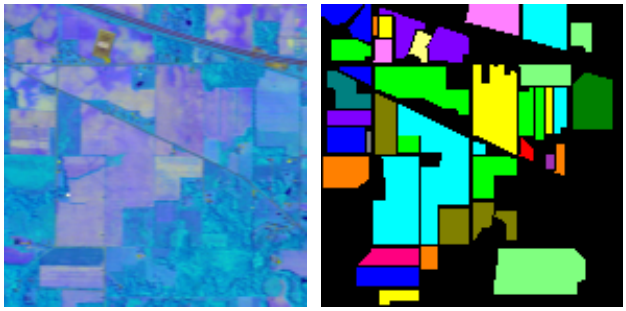
The experiments have been conducted using three datasets, of which two are fairly common; namely Pavia University and Indian Pines. The first is an urban area of size 340×610 pixels and 9 thematic classes, and has been acquired using the ROSIS-03 sensor with a 1.3m spatial resolution over the city of Pavia, Italy. The ROSIS-03 sensor has 115 data channels with a spectral coverage ranging from 0.43 to 0.86μm. After the elimination of 12 noisy bands, 103 bands have been left for processing (Fig. 6).



(a) Pavia University (b) Ground Truth (c) Training Set

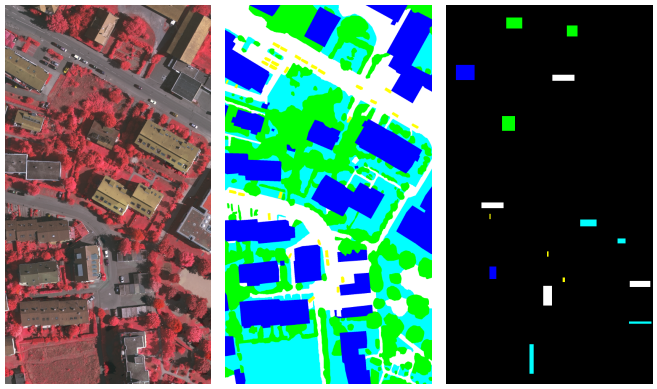
Fig. 6. The Pavia University dataset (false colors) and its corresponding ground truth; its thematic classes (training set size/ground truth size) are: Asphalt (548/6631), Trees (524/3064), Bitumen (375/1330), Meadows (540/18649), Metal sheets (265/1345), Shadows (231/947), Gravel (392/2099), Bare soil (532/5029) and Self-blocking bricks (514/3682).

The second dataset has been captured over the Indian Pines test site in north-western Indiana, USA, through the AVIRIS sensor. This sensor has 224 data channels with a spectral coverage ranging from 0.43 to 2.5μm. The scene depicts primarily agricultural fields of regular geometry, with 16 thematic classes, has a spatial resolution of 20m and is of size 145×145 pixels. Plus, 24 spectral bands affected by atmospheric absorption have been discarded, thus resulting in a dataset of 200 bands (Fig. 7).



(a) Indian Pines (b) Ground Truth

Fig. 7. The Indian Pines dataset and its corresponding ground truth; its thematic classes (ground truth size) are: Alfalfa (46), Corn-notill (1428), Corn-mintill (830), Corn (237), Grass-pasture (483), Grass-trees (730), Grass-pasture-mowed (28), Hay-windrowed (478), Oats (20), Soybean-notill (972), Soybean-mintill (2455), Soybean-clean (593), Wheat (205), Woods (1265), Buildings-Grass-Trees-Drives (386) and Stone-Steel-Towers (93).



(a) Vaihingen (b) Ground Truth (c) Training Set

Fig. 8. The Vaihingen dataset and its corresponding ground truth; its thematic classes (training set size/ground truth size) are: Impervious Surfaces (3858/113k), Building (2554/134k), Low vegetation (2259/116k), Tree (3356/141k), Car (170/3127).

Both hyperspectral datasets have had their spectral dimension reduced by means of a PCA as per [22]. Four PCs representing approximately 99% of the total image variance have been preserved for both. As a note, we are fully aware of the plethora of alternative dimension reduction methods used in this context [37], [38]; the reason we have selected PCA is because of its widespread use, and because our goal is to measure the relative differences between the tested strategies, and not to achieve absolute performance maximization.

As to the third multispectral dataset, it has been captured over the small village of Vaihingen, Germany and contains mostly detached buildings [39] (Area 34). It possesses a very high spatial resolution of 23cm per pixel, and has a size of

543×1000 pixels. The Vaihingen dataset, possesses 3 spectral channels: near infrared, red and green with 5 thematic classes (Fig. 8).

Classification has been conducted using a Random Forest classifier composed of 100 trees. The number of variables involved in the training of the classifier was set to the square root of the feature vector length, as suggested by [40]. Classification performance has been measured by means of the kappa statistic. For the Pavia University dataset, training has been realized using the standard training set widely used in the state-of-the-art [22]. As to Indian Pines, randomly selected 1% pixels per class of the ground truth have been used for training, along with 30 Monte Carlo runs to minimize training set selection bias. As a side-note, larger training sets (5%, 10% and 15%) have also been explored, however their effect of the relative performances of the methods under study have been negligible. For the Vaihingen dataset, we have used the training set shown in Fig. 8c.

We have selected two geometrical attributes for constructing the attribute profiles, the area (an increasing attribute, i.e. if it is verified for a connected component, then it will also be verified by all the regions brighter or darker, depending on the transformation) and the first moment invariant of Hu (a non-increasing attribute, also known as moment of inertia) [41]. Moreover, the subtractive rule has been used for dealing with the non-increasing attribute as in [17], [42]. For area thresholds ($\lambda_{a,Pav}, \lambda_{a,Pin}$), we have computed the automatic settings according to [43] and for the moment of inertia ($\lambda_{m,Pav}, \lambda_{m,Pin}$) we have employed the manual settings used in [22], [37], [44]:

$$\begin{aligned}\lambda_{a,Pav} = \{&770, 1538, 2307, 3076, 3846, 4615, 5384\\ &6153, 6923, 7692, 8461, 9230, 10000, 10769\}\\ \lambda_{a,Pin} = \{&50, 100, 150, 200, 250, 300, 350, 400, 450,\\ &500, 550, 600, 650, 700\}\\ \lambda_{m,Pav} = \lambda_{m,Pin} = \{&0.2, 0.3, 0.4, 0.5\}\end{aligned}$$

However, it is well known that attribute threshold selections have a very significant effect on the performance of AP [17], and the aforementioned settings, regardless of whether they are automatic or manual, have been empirically determined by their respective authors with always the marginal strategy in mind. Which is why, for the sake of fairness, we employ multiple sets ($\Lambda_{a,Pav}, \Lambda_{a,Pin}, \Lambda_{m,Pav}, \Lambda_{m,Pin}$) of thresholds for both attributes, by simply scaling them with various multipliers μ:

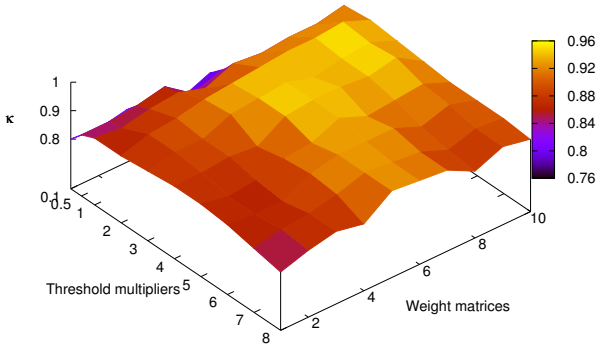
$$\begin{aligned}\Lambda_{a,Pav} &= \{\mu \cdot \lambda_{a,Pav} \,|\, \forall\, \mu \in \{0.1, 0.5, 1, 2, 3, 4, 5, 6, 7, 8\}\}\\ \Lambda_{a,Pin} &= \{\mu \cdot \lambda_{a,Pin} \,|\, \forall\, \mu \in \{0.1, 0.5, 1, 2, 3, 4, 5, 6, 7, 8\}\}\\ \Lambda_{m,Pav} &= \{\mu \cdot \lambda_{m,Pav} \,|\, \forall\, \mu \in \{0.1 \cdot k \,|\, k \in \{1, \ldots, 20\}\}\}\\ \Lambda_{m,Pin} &= \{\mu \cdot \lambda_{m,Pin} \,|\, \forall\, \mu \in \{0.1 \cdot k \,|\, k \in \{1, \ldots, 20\}\}\}\end{aligned}$$

B. Weight selection

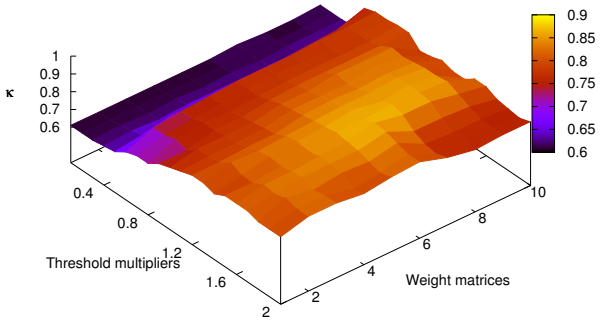
As the design of an automatic weight selection solution is beyond the scope of this paper, at this stage of our work we have proceeded with an empirical selection. Since our goal is to combine marginal and vector strategies we started with a pool of weight matrices ranging between the two extremes (W_1: Euclidean norm based ordering and W_{10}: Marginal ordering):

$$\begin{aligned}\{&W_1 = 0.1 \cdot J_4, W_2 = 0.1 \cdot J_4 + 0.1 \cdot I_4,\\ &\ldots, W_9 = 0.1 \cdot J_4 + 0.8 \cdot I_4, W_{10} = I_4\}\end{aligned} \tag{11}$$

where J and I denote respectively the matrix of ones and the identity matrix. Fig. 9 shows the classification scores obtained with the Pavia University dataset for both attributes. Based on these results, we have selected to use the weight matrix $W_6 = 0.1 \cdot J_4 + 0.5 \cdot I_4$ for both attributes and all datasets.



(a) Area



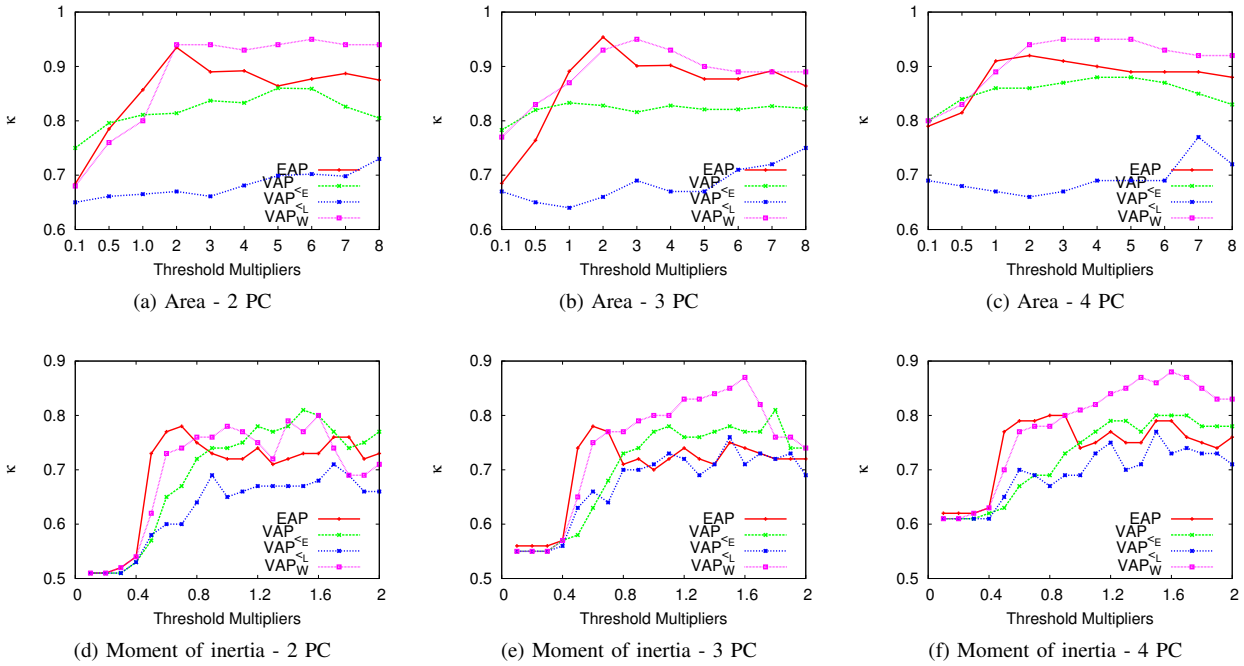
(b) Moment of inertia

Fig. 9. Plots illustrating the effect of various weight matrices on the performance of both attributes using all 4 PC of the Pavia University dataset.

C. Classification

The classification scores (in terms of the kappa statistic) that have been obtained for the Pavia University dataset are shown in Fig. 10, the classification maps corresponding to the various studied strategies are provided in Fig. 14 and its class based performances are at Table II.

As far as the increasing attribute, area, is concerned, the overall performances are affected moderately by the various



(a) Area - 2 PC (b) Area - 3 PC (c) Area - 4 PC

(d) Moment of inertia - 2 PC (e) Moment of inertia - 3 PC (f) Moment of inertia - 4 PC

Fig. 10. Classifications scores (κ statistic) for the Pavia University dataset for the area and moment of inertia attributes using a variety of principal components and threshold settings.

TABLE II
CLASSIFICATION ACCURACIES (%) FOR THE PAVIA UNIVERSITY DATASET WITH 4 PCS AND OPTIMAL SETTINGS OF EACH METHOD.

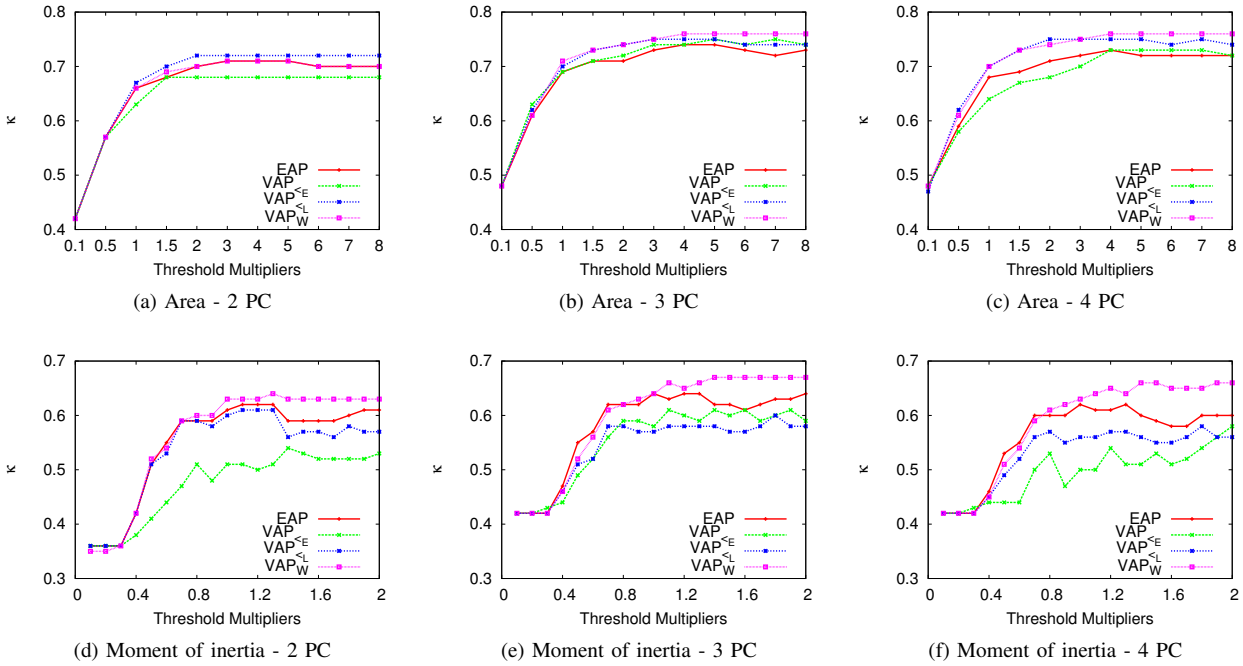
		Area				Moment			
Class	Set Size	EAP	$VAP_{<_L}$	$VAP_{<_E}$	VAP_W	EAP	$VAP_{<_L}$	$VAP_{<_E}$	VAP_W
Asphalt	6631	**95.4**	91.9	92.1	95.1	92.7	92.8	90.6	**94.1**
Meadows	18649	94.0	77.8	96.5	**97.1**	78.2	74.0	87.3	**94.1**
Gravel	2099	**90.2**	78.8	54.8	79.4	**80.3**	61.7	45.7	76.7
Trees	3064	**98.4**	95.1	95.3	97.6	**99.5**	98.4	98.9	98.8
Painted metal sheets	947	99.2	**99.9**	**99.9**	**99.9**	99.3	**99.6**	99.0	99.0
Bare Soil	3682	84.9	63.7	74.9	**97.1**	**65.8**	62.6	63.4	64.6
Bitumen	1345	**99.9**	95.6	97.4	**99.9**	**99.8**	91.6	97.8	99.6
Self-Blocking Bricks	1330	**99.6**	91.7	90.7	99.1	**98.0**	91.7	85.4	97.2
Shadows	5029	**98.5**	98.4	87.8	95.6	**99.6**	**99.6**	97.5	92.7
Average Accuracy		95.57	88.1	87.71	**95.64**	90.36	85.78	85.07	**90.76**
Overall Accuracy		94.2	82.5	90.6	**96.3**	84.1	80.1	84.6	**90.7**
κ		0.92	0.77	0.88	**0.95**	0.8	0.75	0.8	**0.88**

thresholds sets. However, the relative performances are consistent across 2, 3 and 4 principal components. The vector approaches ($VAP_{<_E}$ and $VAP_{<_L}$) are inferior to the marginal strategy, while the lexicographical approach is consistently the poorest among them. Considering the fact that the lexicographical ordering takes almost exclusively into account the first dimension of vectors, this practically means that it almost ignores all PC beyond the first, thus taking a serious performance hit. Besides, these results confirm the similar MP based experiments in [24]. The overall best approach with area is VAP_W which is systematically superior to the marginal approach.

In terms of class based performances, the vector approaches ($VAP_{<_E}$ and $VAP_{<_L}$) exhibit drastically distinct performances both w.r.t. each other and w.r.t. VAP_W and EAP. The latter two are mostly similar, with the principal exceptions being the classes Bare Soil and Gravel, where they outperform each other respectively by a significant margin. These results are particularly motivating for exploring the collaboration of distinct ordering relations.

On the other hand, with the non increasing attribute, overall performance drops, as area apparently is a more effective attribute with this dataset. Moreover, the performances vary more erratically across the various attribute thresholds, especially for fewer PC, thus underlining the effect of the selected attribute. All the same, VAP_W clearly outperforms the marginal strategy by multiple percentile points, especially for 4 principal components and higher scaling factors. If we take a further and closer look to the involved classes, one can observe additionally, that for their optimal settings, the relative performances of the methods under study have remained relatively similar w.r.t. the area attribute.

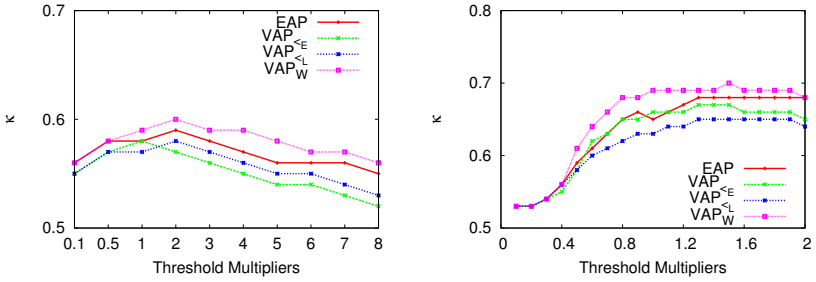
The corresponding results of the Indian Pines dataset are provided in Fig. 11, its classification maps are shown in Fig. 15



(a) Area - 2 PC (b) Area - 3 PC (c) Area - 4 PC

(d) Moment of inertia - 2 PC (e) Moment of inertia - 3 PC (f) Moment of inertia - 4 PC

Fig. 11. Classifications scores (κ statistic) for the Indian Pines dataset for the area and moment of inertia attributes using a variety of principal components and threshold settings.



(a) Area (b) Moment of inertia

Fig. 12. Classifications scores (κ statistic) for the Vaihingen dataset using the area and moment of inertia attributes along a variety of thresholds.

and its class based performances are at Table III.

Judging from the scores obtained with area, one can immediately observe that the performances vary much less than before, and are much closer. However, they diversify as the number of components increases, and with 4 principal components we can see that the lexicographical method performs way better w.r.t. the previous dataset, while VAP_W is overall the best with a very small margin. The marginal approach on the other hand performs poorer. One possible explanation could be the fact that the first PC of the Pines dataset contains a significantly higher amount of information w.r.t. to the first PC of Pavia University, thus placing lexicographical ordering at an advantage. Hence, a dataset specific selection of weights would probably improve even further the performance of VAP_W.

Once more, the non-increasing moment attribute is more sensitive to threshold selections, with $VAP_{<_E}$ performing the poorest, probably because the 2nd, 3rd and 4th principal components lack valuable information so the pixel signatures become corrupted as the Euclidean norm takes the entire vectors into account. Thanks to its balanced design, VAP_W once more is the best performing approach among them, while both vector methods are clearly inferior to the marginal strategy. As far as class based performances are concerned, the Pines dataset exhibits a high level of inter-class and inter-method variance. In classes such as Grass-Pasture-Mowed, Oats and Stone-Steel-Towers, the best performing method can even triple the performance of the worst one.

The classification scores that have been obtained for the Vaihingen dataset are shown in Fig. 12, the classification maps corresponding to the various studied strategies are provided in Fig. 16 and its class based performances are at Table IV.

In terms of the area attribute, there is an obvious correlation between all studied approaches across attribute thresholds, similarly to the Pines dataset and on the contrary of the Pavia University dataset. The overall performances are very close, and conversely to the other two datasets, they are globally

TABLE III
CLASSIFICATION ACCURACIES (%) AND THEIR DEVIATIONS FOR THE INDIAN PINES DATASET WITH 4 PCS AND OPTIMAL SETTINGS OF EACH METHOD.

		Area				Moment			
Class	Set Size	EAP	$VAP_{<_L}$	$VAP_{<_E}$	VAP_W	EAP	$VAP_{<_L}$	$VAP_{<_E}$	VAP_W
Alfalfa	46	78.3 ± 1.6	82.6 ± 1.7	58.7 ± 1.8	**97.8** ± 1.9	69.39 ± 1.5	79.14 ± 1.2	44.12 ± 1.3	**93.68** ± 1.9
Corn-notill	1428	70.6 ± 1.7	65.9 ± 1.4	66.8 ± 1.3	**75.1** ± 1.6	54.89 ± 1.4	48.87 ± 1.2	65.4 ± 1.5	**68.22** ± 1.5
Corn-mintill	830	76 ± 1.5	**87.1** ± 1.2	61.4 ± 1.6	55.9 ± 1.5	36.67 ± 1.2	30.36 ± 1.5	**51.62** ± 1.4	42.32 ± 1.3
Corn	237	25.3 ± 1.8	22.4 ± 1.8	**26.2** ± 1.4	23.2 ± 1.6	14.77 ± 1.1	12.96 ± 0.8	**36.29** ± 1.1	25.36 ± 1.2
Grass-past.	483	62.9 ± 1.5	72.9 ± 1.8	66.9 ± 1.4	**78.9** ± 1.8	54.04 ± 1.7	58.57 ± 1.1	50.64 ± 1.4	**65.48** ± 1.6
Grass-trees	730	88.6 ± 1.3	**98.9** ± 1.6	82.2 ± 1.3	87.1 ± 1.6	86.88 ± 1.5	**87.75** ± 1.8	72.36 ± 1.8	87.64 ± 2.0
Grass-past.-mow.	28	**100** ± 1.2	96.4 ± 1.4	71.4 ± 0.1	96.4 ± 0.2	58.92 ± 1.7	57.86 ± 1.4	22.15 ± 0.9	**66.79** ± 1.7
Hay-windrowed	478	88.1 ± 1.4	96.4 ± 1.7	**99.8** ± 1.6	**99.8** ± 1.5	**92.67** ± 1.8	83.41 ± 1.7	83.07 ± 1.8	91.91 ± 1.9
Oats	20	**60** ± 1.9	25 ± 0.3	30 ± 1.6	30 ± 0.4	22.5 ± 1.0	17 ± 0.7	14 ± 0.8	**27.5** ± 1.4
Soybean-notill	972	68.9 ± 1.4	47.6 ± 1.7	**69.4** ± 1.4	61.7 ± 1.8	**54.37** ± 1.3	42.55 ± 1.3	38.93 ± 1.2	48.44 ± 1.4
Soybean-mintill	2455	87.2 ± 1.8	**94.3** ± 1.5	92.8 ± 1.7	93.7 ± 1.4	82.38 ± 1.6	82.45 ± 1.7	80.29 ± 1.9	**87.26** ± 1.7
Soybean-clean	593	36.1 ± 1.7	67.1 ± 1.4	54.1 ± 1.9	**76.9** ± 1.3	37.07 ± 1.3	28.46 ± 1.1	31.76 ± 1.2	**47.41** ± 1.4
Wheat	205	**98.5** ± 1.6	96.6 ± 1.4	73.2 ± 1.2	**98.5** ± 1.2	**92.11** ± 1.6	83.33 ± 1.9	61.71 ± 1.8	57.61 ± 1.5
Woods	1265	89.2 ± 1.3	82.1 ± 1.1	**89.9** ± 1.7	89.7 ± 1.7	93.56 ± 1.4	**95.72** ± 1.8	85.26 ± 1.9	95.45 ± 1.9
Build.-Gr.-Tr.-Dr.	386	68.7 ± 1.5	**75.1** ± 1.3	66.8 ± 1.8	74.4 ± 1.5	**47.92** ± 1.1	30.61 ± 1.2	35.82 ± 1.1	44.11 ± 1.4
St.-St.-Tow.	93	16.1 ± 1.9	48.4 ± 0.5	52.7 ± 1.8	**51.6** ± 1.6	37.1 ± 1	22.81 ± 1.1	24.84 ± 1.1	**41.94** ± 1.4
Average Accuracy		69.65 ± 1.3	72.42 ± 1.2	66.39 ± 1.5	**74.42** ± 1.4	58.02 ± 0.9	53.87 ± 1.4	49.89 ± 1.3	**61.95** ± 1.6
Overall Accuracy		75.9 ± 1.2	78.7 ± 1.2	76.6 ± 1.7	**79.9** ± 1.6	67.04 ± 1.1	63.29 ± 1.4	63.94 ± 1.5	**71** ± 1.6
κ		0.72 ± 0.01	0.75 ± 0.01	0.73 ± 0.01	**0.77** ± 0.01	0.62 ± 0.01	0.57 ± 0.01	0.58 ± 0.01	**0.66** ± 0.01

TABLE IV
CLASSIFICATION ACCURACIES (%) FOR THE VAIHINGEN DATASET WITH OPTIMAL SETTINGS OF EACH METHOD.

		Area				Moment			
Class	Set Size	EAP	$VAP_{<_L}$	$VAP_{<_E}$	VAP_W	EAP	$VAP_{<_L}$	$VAP_{<_E}$	VAP_W
Impervious Surfaces	113549	64.4	64.1	63.8	**65.2**	**87.6**	84.5	75.0	83
Building	133719	64.9	64.7	64.6	**65.8**	77.9	75.5	81	**83**
Low vegetation	116526	35.5	33.6	33.1	**36.7**	30.3	34.3	**45.7**	19.2
Tree	141474	57.4	57.1	57	**57.9**	**70.4**	69.2	62.4	69.5
Car	3127	60.1	60	60.1	**60.3**	70.2	69.5	**80.5**	75
Average Accuracy		56.5	55.9	55.72	**57.2**	67.28	66.6	**68.82**	65.94
Overall Accuracy		61.7	61.3	61.2	**62.2**	75.9	74.1	74.9	**77.3**
κ		0.59	0.58	0.58	**0.6**	0.68	0.65	0.67	**0.7**

worse than the moment of inertia. This is not unexpected, since the dataset in question is of very high spatial resolution and as such, elongation (as provided by the moment of inertia) can be a more discriminatory property of objects than their area. Still, the proposed strategy (VAP_W), achieved systematically better scores, though with a very small margin. Furthermore, the vector strategies once again perform worse than their alternatives, VAP_W and EAP. In terms of class based performances, all four tested approaches have achieved similar scores per class, with VAP_W outscoring its alternatives with all classes.

As to the moment attribute, this time overall performances are higher, and more diversified across attribute thresholds. Even so, the proposed approach still outperforms the marginal strategy, while the vector strategies follow them closely. As far as class specific classification scores are concerned, low vegetation proves to be particularly challenging to recognize, while the tested approaches exhibit a more diversified performance across classes.

D. Computational cost

In order to better evaluate the studied approaches, we present in Fig. 13 their execution times (in seconds). We measured using a standard laptop (3GHz, 8GB memory) the time required for a full profile computation (tree construction, filtering and tree reconstitution) for every strategy using the area attribute and all three datasets along with optimal threshold settings for each.

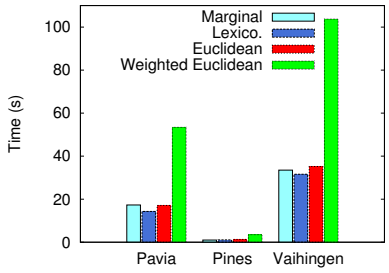
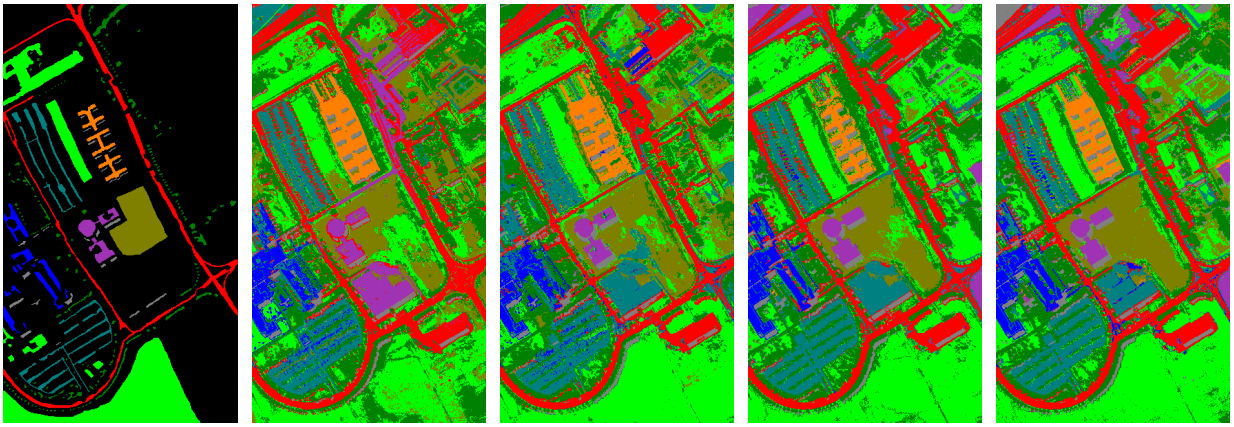


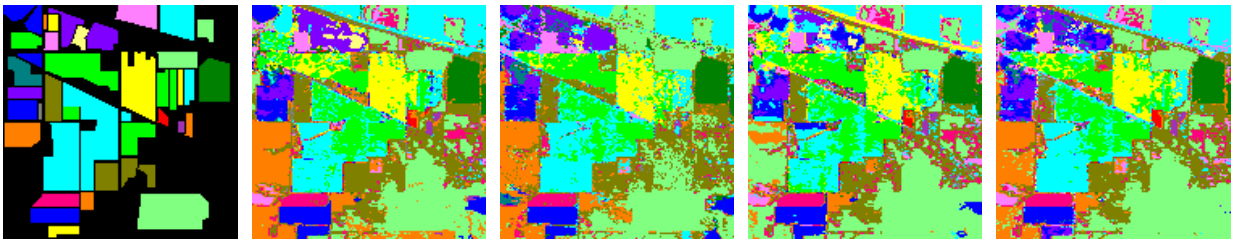
Fig. 13. The computational cost of the studied approaches across all three datasets using their optimal settings along with the area attribute.

As expected, the vector approaches ($VAP_{<_E}$ and $VAP_{<_L}$) are faster than the marginal approach. However, since the marginal max/min trees are developed across only 256 intensity levels, while purely vector strategies operate on 256^4 possible spectral values, their difference is relatively negligible. Naturally the slowest among them is $VAP_{<_W}$, since it needs to process the data both with a vector strategy and with multiple passes as in the case of the marginal approach. Let it be noted however, that our implementations are preliminary and sub-optimal single threaded prototypes.



(a) Ground Truth (b) Lexicographical (c) Euclidean Norm (d) Marginal (e) Weighted Euclidean

Fig. 14. The classification maps of the Pavia University dataset using the area attribute for each AP computation strategy, along with their optimal settings.



(a) Ground Truth (b) Lexicographical (c) Euclidean Norm (d) Marginal (e) Weighted Euclidean

Fig. 15. The classification maps of the Indian Pines dataset using the area attribute for each AP computation strategy, along with their optimal settings.

V. Conclusion

The focus of this paper has been on attribute profiles, a powerful morphological spectral-spatial pixel description tool, and especially on how to improve their strategy of computation from hyperspectral images. To this end, we have transferred from color morphology some of the accumulated know-how on multivariate image processing. Specifically, we have first formulated an effective alternative vector strategy to the widely and exclusively used marginal processing method, which is capable of processing all image bands simultaneously.

We have further introduced a combination of both marginal and vector strategies as a more balanced approach, integrating the best aspects of the two available methods. In particular, the approach that has been developed can be configured through its weight matrix to a behavior ranging from being identical to the widely used marginal strategy (hence providing band specific descriptions) all the way to the other extreme, of being a purely distance based vector approach (thus taking into account information from adjacent bands). As a result, the approach that has been developed can be configured to deal with a very large variety of dataset properties.

Moreover we have provided an experimental comparison of the studied approaches with three datasets, two attributes, multiple number of bands, and various threshold levels. Our experiments have shown that despite using the same relatively simple weight matrices, that have been determined empirically, the classification performances of the proposed method have been clearly superior to the marginal approach. These results have motivated us greatly for pursuing this method further, especially in order to design an automatic solution for dataset specific weight matrix determination, for which multiple directions have been presented.

Acknowledgments

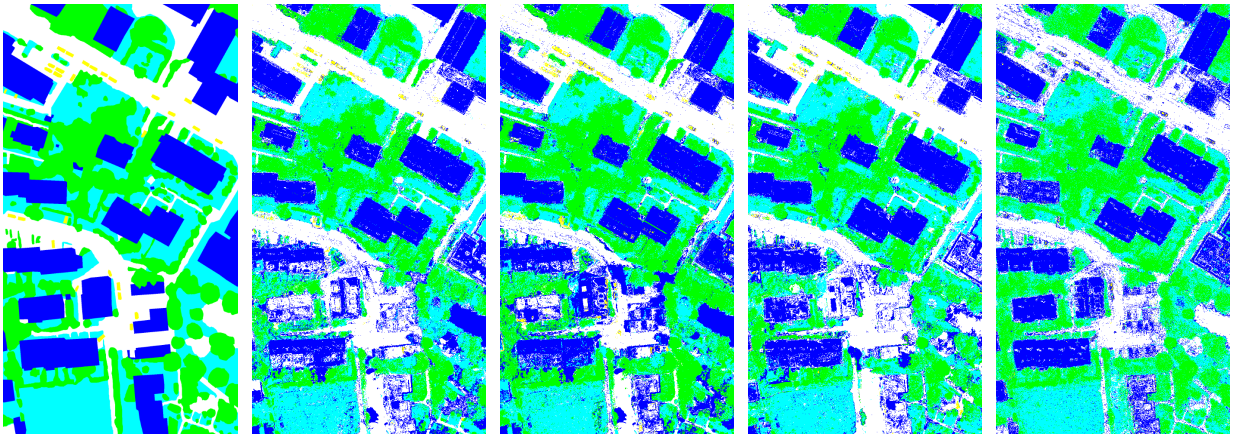
This work was supported by the Turkish TUBITAK Career Grant 112E210 and by the French Agence Nationale de la Recherche (ANR) under reference ANR-13-JS02-0005-01 (Asterix project). The authors would like to thank Prof. P. Gamba and Prof. D. Langrebe for making available to the community the Pavia University and Indian Pines hyperspectral datasets respectively. Moreover, the Vaihingen dataset has ben kindly provided by the International Society for Photogrammetry and Remote Sensing (ISPRS).

References

[1] K. Stankov and D.-C. He, "Building detection in very high spatial resolution multispectral images using the hit-or-miss transform," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 1, pp. 86–90, January 2013.

[2] L. Gueguen and M. Pesaresi, "Multi scale harris corner detector based on differential morphological decomposition," *Pattern Recognition Letters*, vol. 32, no. 14, pp. 1714–1719, October 2011.

[3] X. Huang, L. Zhang, and L. Wang, "Evaluation of morphological texture features for mangrove forest mapping and species discrimination using multispectral IKONOS imagery," *IEEE Geoscience and Remote Sensing Letters*, vol. 6, no. 3, pp. 393–397, July 2009.



(a) Ground Truth (b) Lexicographical (c) Euclidean Norm (d) Marginal (e) Weighted Euclidean

Fig. 16. The classification maps of the Vaihingen dataset using the area attribute for each AP computation strategy, along with their optimal settings.

[4] S. Valero, J. Chanussot, J. A. Benediktsson, H. Talbot, and B. Waske, "Advanced directional mathematical morphology for the detection of the road network in very high resolution remote sensing images," *Pattern Recognition Letters*, vol. 31, no. 10, pp. 1120–1127, July 2007.

[5] E. Aptoula, "Remote sensing image retrieval with global morphological texture descriptors," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 5, pp. 3023–3034, 2014.

[6] G. Ouzounis, P. Soille, and M. Pesaresi, "Rubble detection from VHR aerial imagery data using differential morphological profiles," in *Proceedings of the 34th International Symposium for Remote Sensing of the Environment*, Sydney, Australia, April 2011.

[7] G. Noyel, J. Angulo, and D. Jeulin, "A new spatio-spectral morphological segmentation for multispectral remote sensing images," *International Journal of Remote Sensing*, vol. 31, no. 22, pp. 5895–5920, 2010.

[8] K. Bernard, Y. Tarabalka, J. Angulo, J. Chanussot, and J. A. Benediktsson, "Spectral-spatial classification of hyperspectral data based on a stochastic minimum spanning forest approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 21, no. 4, pp. 2008–2021, 2012.

[9] E. Aptoula, S. Lefèvre, and C. Collet, "Mathematical morphology applied to the segmentation and classification of galaxies in multispectral images," in *Proceedings of the European Signal Processing Conference*, Florence, Italy, 2006.

[10] Y. Tarabalka, J. Chanussot, and J. A. Benediktsson, "Segmentation and classification of hyperspectral images using watershed transformation," *Pattern Recognition*, vol. 43, no. 7, pp. 2367–2379, July 2010.

[11] X. Huang, X. Guan, J. A. Benediktsson, and L. Zhang, "Multiple morphological profiles from multicomponent-base images for hyperspectral image classification," *Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 12, pp. 4653–4669, December 2014.

[12] S. Bernabé, P. R. Marpu, A. Plaza, and J. A. Benediktsson, "Spectral unmixing of multispectral satellite images with dimensionality expansion using morphological profiles," in *SPIE Proceedings on Satellite Data Compression, Communications, and Processing*, vol. 8514, San Diego, USA, 2012.

[13] M. Pesaresi and J. Benediktsson, "A new approach for the morphological segmentation of high resolution satellite imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 2, pp. 309–320, February 2001.

[14] A. Plaza, P. Martinez, R. Pérez, and J. Plaza, "A new approach to mixed pixel classification of hyperspectral imagery based on extended morphological profiles," *Pattern Recognition*, vol. 37, no. 6, pp. 1097–1116, June 2004.

[15] M. Fauvel, J. A. Benediktsson, J. Chanussot, and J. R. Sveinsson, "Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 11, pp. 3804–3814, November 2008.

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

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

[18] P. Ghamisi, M. Dalla Mura, and J. A. Benediktsson, "A survey on spectral-spatial classification techniques based on attribute profiles," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 5, pp. 2335–2353, May 2015.

[19] L. Najman and M. Couprie, "Building the component tree in quasi-linear time," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3531–3539, November 2006.

[20] P. Monasse and F. Guichard, "Fast computation of a contrast-invariant image representation," *IEEE Transactions on Image Processing*, vol. 9, no. 5, pp. 860–872, May 2000.

[21] M. Dalla Mura, J. A. Benediktsson, and L. Bruzzone, "Self-dual attribute profiles for the analysis of remote sensing images," in *Mathematical Morphology and Its Applications to Image and Signal Processing*, P. Soille, G. K. Ouzounis, and M. Pesaresi, Eds. Verbania-Intra, Italy: Springer-Verlag, 2011, pp. 320–330.

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

[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," *International Journal of Image and Data Fusion*, vol. 3, no. 3, pp. 269–298, 2012.

[24] E. Aptoula and S. Lefèvre, "A comparative study on multivariate mathematical morphology," *Pattern Recognition*, vol. 40, no. 11, pp. 2914–2929, November 2007.

[25] C. Kurtz, B. Naegel, and N. Passat, "Connected filtering based on multivalued component-trees," *IEEE Transactions on Image Processing*, vol. 23, no. 12, pp. 5152–5164, December 2014.

[26] F. Dell'Acqua, P. Gamba, A. Ferrari, J. A. Palmason, J. A. Benediktsson, and K. Arnason, "Exploiting spectral and spatial information in hyperspectral urban data with high resolution," *IEEE Geosci. and Remote Sensing Letters*, vol. 1, no. 4, pp. 322–326, October 2004.

[27] J. Serra, "Anamorphoses and function lattices," in *Mathematical Morphology in Image Processing"*, E. R. Dougherty, Ed. New York: Marcel Dekker, 1993, ch. 13, pp. 483–523.

[28] P. Li and H. Hu, "Segmentation of high-resolution multispectral image based on extended morphological profiles," in *Proceedings of the International Geoscience and Remote Sensing Symposium*, Barcelona, Spain, July 2007, pp. 1481–1484.

[29] S. Velasco-Forero and J. Angulo, "Supervised ordering in $\mathbb{R}^n$: Application to morphological processing of hyperspectral images," *IEEE Transactions on Image Processing*, vol. 20, no. 11, pp. 3301–3308, 2011.

[30] N. Courty, E. Aptoula, and S. Lefevre, "A classwise supervised ordering approach for morphology based hyperspectral image classification," in *Proceedings of the International Conference on Pattern Recognition*, Tsukuba, Japan, November 2012, pp. 1997–2000.

[31] E. Aptoula, N. Courty, and S. Lefèvre, "An end-member based ordering relation for the morphological description of hyperspectral images," in *Proceedings of the IEEE International Conference on Image Processing*, Paris, France, October 2014, pp. 5097–5101.
[32] E. Aptoula, "Hyperspectral image classification with multi-dimensional attribute profiles," *IEEE Geoscience and Remote Sensing Letters*, 2015, In press.
[33] E. J. Breen and R. Jones, "Attribute openings, thinnings and granulometries," *Computer Vision and Image Understanding*, vol. 64, no. 3, pp. 377–389, November 1996.
[34] P. Salembier, A. Oliveras, and L. Garrido, "Anti-extensive connected operators for image and sequence processing," *IEEE Transactions on Image Processing*, vol. 7, no. 4, pp. 555–570, April 1998.
[35] E. Aptoula, J. Weber, and S. Lefèvre, "Vectorial quasi-flat zones for color image simplification," in *Proceedings of the International Symposium on Mathematical Morphology*, Uppsala, Sweden, April 2013.
[36] E. Aptoula, "The impact of multivariate quasi-flat zones on the morphological description of hyperspectral images," *International Journal of Remote Sensing*, vol. 35, no. 10, pp. 3482–3498, May 2014.
[37] 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, May 2011.
[38] S. Bernabé, P. R. Marpu, A. Plaza, M. Dalla Mura, and J. A. Benediktsson, "Spectral-spatial classification of multispectral images using kernel feature space representation," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 1, pp. 288–292, January 2014.
[39] M. Cramer, "The DGPF test on digital aerial camera evaluation - overview and test design," *Photogrammetrie - Fernerkundung - Geoinformation*, vol. 2, pp. 73–82, 2010.
[40] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
[41] M. Hu, "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory*, vol. 8, no. 2, pp. 179–187, 1962.
[42] E. R. Urbach, J. B. T. M. Roerdink, and M. H. F. Wilkinson, "Connected shape-size pattern spectra for rotation and scale-invariant classification of gray-scale images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 2, pp. 272–285, February 2007.
[43] P. Ghamisi, J. A. Benediktsson, and J. R. Sveinsson, "Automatic spectral-spatial classification framework based on attribute profiles and supervised feature extraction," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 9, pp. 5771–5782, September 2014.
[44] G. Cavallaro, N. Falco, M. Dalla Mura, L. Bruzzone, and J. A. Benediktsson, "Automatic threshold selection for profiles of attribute filters based on granulometric characteristic functions," in *Mathematical Morphology and Its Applications to Signal and Image Processing*, J. Benediktsson, J. Chanussot, L. Najman, and H. Talbot, Eds., 2015, vol. 9082, pp. 169–181.

Erchan Aptoula received the B.Sc. degree in computer engineering from Galatasaray University, Istanbul, Turkey in 2004 and the M.Sc. and Ph.D. degrees in computer science from Strasbourg University, Strasbourg, France in 2005 and 2008 respectively. He is currently an Associate Professor at the computer engineering department of Okan University at Istanbul, Turkey, working on morphological image analysis for color and hyper-spectral images, and content-based image description.

Mauro Dalla Mura Mauro Dalla Mura (S'08 M'11) received the laurea (B.E.) and laurea specialistica (M.E.) degrees in Telecommunication Engineering from the University of Trento, Italy, in 2005 and 2007, respectively. He obtained in 2011 a joint Ph.D. degree in Information and Communication Technologies (Telecommunications Area) from the University of Trento, Italy and in Electrical and Computer Engineering from the University of Iceland, Iceland. In 2011 he was a Research fellow at Fondazione Bruno Kessler, Trento, Italy, conducting research on computer vision. He is currently an Assistant Professor at Grenoble Institute of Technology (Grenoble INP), France. He is conducting his research at the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA-Lab). His main research activities are in the fields of remote sensing, image processing and pattern recognition. In particular, his interests include mathematical morphology, classification and multivariate data analysis. Dr. Dalla Mura was the recipient of the IEEE GRSS Second Prize in the Student Paper Competition of the 2011 IEEE IGARSS 2011 and co-recipient of the Best Paper Award of the International Journal of Image and Data Fusion for the year 2012-2013 and the Symposium Paper Award for IEEE IGARSS 2014. He is a member of the Geoscience and Remote Sensing Society (GRSS) and IEEE GRSS Data Fusion Technical Committee (DFTC) and Secretary of the IEEE GRSS French Chapter (2013-2016). He was a lecturer at the RSSS12 - Remote Sensing Summer School 2012 (organized by the IEEE GRSS), Munich, Germany.

Sébastien Lefèvre Sbastien Lefvre received his M.Sc. and Eng. degrees in Computer Engineering from the University of Technology of Compigne in 1999, and his Ph.D. in Computer Science from the University of Tours in 2002. In 2009, he earned his French HDR degree (Habilitation to Supervise Doctoral Studies) in Computer Science from the University of Strasbourg. From 2003 to 2010, he was an Associate Professor in the Department of Computer Sciences and the Image Sciences, Computer Sciences and Remote Sensing Laboratory (LSIIT), University of Strasbourg - CNRS. In 2009-2010, he was an INRIA invited scientist within the TEXMEX team of IRISA/INRIA Rennes. In 2010, he joined the University of Bretagne-Sud as a Full Professor in Computer Science, in the Institute of Technology of Vannes and the Institute for research in computer science and random systems (IRISA). Within IRISA, he is leading the OBELIX team dedicated to image analysis and machine learning for remote sensing and earth observation (www.irisa.fr/obelix). He has coauthored more than 100 papers in image analysis and pattern recognition. His current research interests are mainly related to multivariate mathematical morphology, hierarchical models and machine learning applied to remote sensing of environment.