

Attribute Profiles from Partitioning Trees

Petra Bosilj^{1,4,*}, Bharath Bhushan Damodaran¹, Erchan Aptoula², Mauro Dalla Mura³, and Sébastien Lefèvre¹

¹ Univ. Bretagne Sud - IRISA, Vannes, France; firstname.lastname@irisa.fr

² Institute of Information Technologies, Gebze Technical University, Turkey; eaptoula@gtu.edu.tr

³ GIPSA Laboratory, Department Image and Signal, Grenoble-INP, Saint Martin d'Herès Cedex, France; mauro.dalla-mura@gipsa-lab.grenoble-inp.fr

⁴ University of Lincoln, Lincoln, UK; pbosilj@lincoln.ac.uk

* The contributions made by author were done during the post-doc position at ¹.

Abstract. Morphological attribute profiles are among the most prominent spatial-spectral pixel description tools. They can be calculated efficiently from tree based representations of an image. Although widely and successfully used with various inclusion trees (i.e., component trees and tree of shape), in this paper, we investigate their implementation through partitioning trees, and specifically α - and (ω) -trees. Our preliminary findings show that they are capable of comparable results to the state-of-the-art, while possessing additional properties rendering them suitable for the analysis of multivariate images.

Keywords: attribute profiles, partitioning trees, α -tree, (ω) -tree, hyperspectral images

1 Introduction

Mathematical Morphology has offered effective ways to perform spatial analysis in many application domains of digital image processing. In the context of Earth Observation through satellite or airborne remote sensing, morphological tools have been popular the last decades, especially due to their intrinsic ability to model spatial information within a multiscale framework. Following (Differential) Morphological Profiles (DMP) in the early 2000s [1], Attribute Profiles (AP) [2] and more recently Self-Dual Attribute Profile (SDAP) [3] are recognized as an efficient solution to provide multilevel spatial-spectral description of image pixels. Such a description can then be used as a pixelwise feature in a subsequent image interpretation task such as land cover classification.

Profiles such as AP or SDAP rely on a tree-based representation of an image (min- and max-tree for AP, Tree of Shapes for SDAP). All these trees are indeed inclusion trees, which rely on an ordering relation of the image pixels. When dealing with multivariate images such as multi- or hyperspectral images, defining such an ordering is not straightforward and neither is the computation of AP/SDAP from multivariate images [4].

Conversely, partitioning trees have received much less attention in the context of multiscale characterizations of images. However, such trees are appealing in this context since they do not require the definition of a vectorial ordering, but only the selection of an appropriate distance measure. In this paper, we present some preliminary work aiming to demonstrate the relevance of building attribute profiles from partitioning trees.

The paper is organized as follows. In Sec. 2, we recall the principles of attribute profiles and the main tree models they rely on. We discuss the different steps to compute attribute profiles from partitioning trees in Sec. 3, before providing experimental results in Sec. 4. Finally, Sec. 5 concludes the paper and gives future research directions.

2 Related Work

In this section, we recall the principles of attribute profiles, attribute filtering and the different inclusion and partitioning trees.

2.1 Attribute Profiles

APs are multiscale image description tools, constructed similarly to MPs, by successively applying a morphological operator with progressively increasing filter parameters (leading to a sequence of increasingly coarser filters). Specifically, APs rely on morphological attribute filters (AFs); 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 $f : E \rightarrow \mathbb{Z}, E \subseteq \mathbb{Z}^2$, its upper-level sets are defined as $\mathcal{L}^t = \{f \geq t\}$ with $t \in \mathbb{Z}$ (resp. lower-level sets $\mathcal{L}_t = \{f \leq t\}$), i.e. the set of images obtained by thresholding an image at all possible values of their pixels. The connected components ($CC \subseteq E$), typically based on 4 or 8-connectivity, composing the upper level sets $\mathcal{L}^t = \{\mathcal{L}^{t,i}\}$ or lower level sets $\mathcal{L}_t = \{\mathcal{L}_{t,i}\}$ are referred to as peak components. AFs are applied to these peak components, using a predefined logical predicate T_κ^α , in general 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 output of the AF 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), 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 this case the AP of

a grayscale image f would become:

$$AP(f) = \{\phi^{\kappa L}(f), \phi^{\kappa L-1}(f) \dots, \phi^{\kappa 1}(f), f, \gamma^{\kappa 1}(f), \dots, \gamma^{\kappa L-1}(f), \gamma^{\kappa L}(f)\}. \quad (1)$$

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 multivariate image $\mathbf{f} : E \rightarrow \mathbb{Z}^r, r > 1$ is concerned, the widely (and mostly exclusively) encountered marginal strategy consists in first reducing the number of spectral dimensions (from r to $n, n \ll r$) through a variety of methods, 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) [5]:

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

Here $band_i$ refers to an image component after dimensionality 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 marginally.

Since their initial introduction by Dalla Mura et al. [2], APs have been intensively studied in various aspects, such as combining them with alternative dimension reduction methods [6], implementation of multi-dimensional attributes [7], design of techniques for automatic threshold production [8], multivariate tree representations [9], histogram based AP representations [10], as well as their combination with deep learning [11]. For a recent survey on the topic the reader is referred to [12].

2.2 Trees

The various hierarchical image representations used in contemporary mathematical morphology to manipulate the connected components of images can be divided into inclusion and partitioning hierarchies. On one hand, the *inclusion* hierarchies comprise partial partitions of the image with nested supports and their components are formed by creating, inflating and merging image blocks [13] (examples can be see in Fig. 1). On the other hand, the hierarchies from the class of *partitioning* trees comprise full partitions of the image. The leaves of the hierarchy form the finest partition and are iteratively merged until a single root node is formed (examples given in Fig. 2). However, inclusion trees require a total order between the pixel values, which limits their power to process multivariate data. In contrast, partitioning trees usually require only a dissimilarity metric between neighboring pixels (i.e. a total ordering on the image edges instead of the image pixels). They are additionally also not extrema-oriented, capturing information about objects at intermediate gray levels.

The seminal hierarchies **min** and the **max-trees** [14], belong to the class of inclusion trees. They are dual hierarchies structuring the inclusion relations between the peak components of the lower (resp. upper) level sets of the image, thus well suited for representing bright (resp. dark) structures in the image. The peak components of lower level sets $\mathcal{L}_{t,i}$ (resp. upper level sets $\mathcal{L}^{t,i}$) are nested

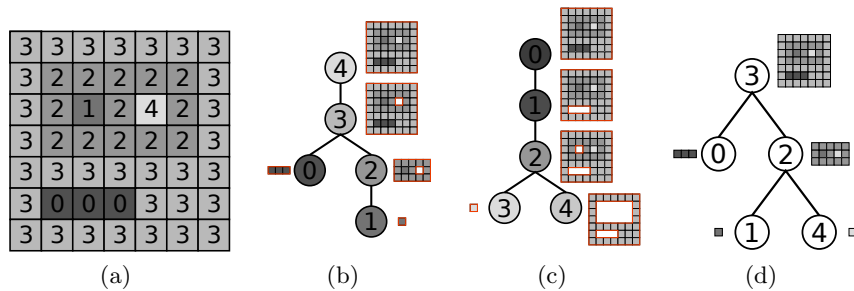


Fig. 1. The three different inclusion trees of a toy image (a). The min-tree is displayed in (b), while its dual max-tree is shown in (c). The self-dual tree of shapes is shown in (d). The (gray) levels of the nodes are displayed in the nodes, and the corresponding regions are shown beside the nodes.

for increasing (resp. decreasing) values of t and form the nodes of the min-tree (resp. max-tree), thus fully representing a grayscale image. The difference between a min and a max-tree hierarchy for the same image given in Fig. 1(a) can be seen in Figs. 1(b) and 1(c).

The **Tree of Shapes** (ToS) [15] was introduced in order to unify a representation of bright and dark image structures, which are treated equivalently to form a hierarchy based on their contrast with the background. The *shapes* which compose the ToS are formed by filling the holes in the peak components of the image (used to form the min and the max-tree), and do not intersect and are either nested or disjoint [16, 17]. This makes the hierarchy an inclusion tree as well, which is additionally self-dual and contrast invariant, while remaining a full representation of the image. It can be seen on Fig. 1(d).

The α -tree is a partitioning tree and is formed based on the local range of its components. It comprises α -connected components [18, 19], with α corresponding to their local range. Flat zones form the 0-CCs (for $\alpha = 0$) and correspond to the connected components of maximal size containing pixels at the same graylevel ($\mathcal{F}_t = \{f = t\}$). For $\alpha > 0$, an α -CC is defined as the CC of the maximal size such that only the neighboring pixels with gray level difference less or equal to α are considered connected. The α -CCs are nested for increasing values of α , forming a complete, self-dual hierarchy which can represent dark and bright but also regions at intermediate image levels. An example of this hierarchy can be seen on Fig. 2(b) for the image in Fig. 2(a). Due to the locality of the metric used, gray level variations within regions can be much higher than α , which is called the chaining effect [19] and is most prominent when the region gray levels gradually increase and decrease (e.g. $\alpha = 2$ in Fig. 2(b)).

To deal with this, different constrained hierarchies [19–21] are constructed from the α -tree, most notably the (ω)-**tree** [18], constraining the α -CCs by their global range. The (ω)-CCs are defined as the largest α -CCs in the image with the global range (i.e. the maximal dissimilarity between any two pixels belonging to that component) less or equal to some ω . Some of the α -CCs are removed

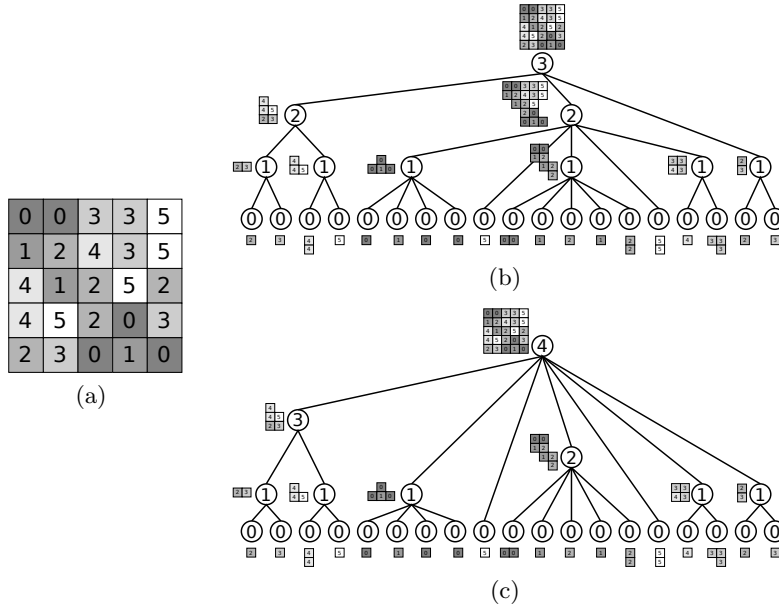


Fig. 2. The two different partitioning trees of a toy image (a). The α -tree is displayed in (b), while the constrained hierarchy (ω)-tree is shown in (c). The α (resp. (ω)) levels of the nodes are displayed in the nodes and indicated by their height, while the regions are displayed besides the nodes. It can be seen how the (ω)-tree contains a subset of the nodes of the α -tree, but arranged through a larger span of levels.

from the hierarchy, but the maximal global range and thus the maximal level in the tree will typically be higher in the (ω)-tree. The hierarchy remains self-dual, complete and capable of capturing regions of low, intermediate and high gray levels, but global range provides better grouping per level than just a local measure. An example of this can be seen in Fig. 2(c).

3 Proposed Method

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 [14].

The max-tree (alternatively min-tree) used to represent the connected components of a given grayscale image is a basis for AP calculation. The advantage of these representations stems from the implementation possibility of attribute 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. Following their calculation using min and max-trees,

the AP have also been calculated on the Tree of Shapes (ToS) [22]. Extending the idea of using different trees to calculate the AP, we examine the α and (ω)-trees as a basis for profile calculation.

Filtering a partitioning tree is not as straightforward as filtering an inclusion tree, since all the children of a single parent need to be processed in the same manner. To construct APs from these trees, we consider an equivalent definition of APs through pixels. For some threshold κ_i , a pixel p would be characterized by the value it obtains in γ^{κ_i} and ϕ^{κ_i} . These values can also be interpreted as the levels of the nodes closest to a leaf in a max-tree (respectively min-tree) containing the pixel p , and satisfying the logical predicate $T_{\kappa_i}^\alpha$ based on the threshold κ_i employed by the thinning and thickening. Following this definition, when using an α -tree, a pixel value at level i of an AP is also determined based on the lowest node in the tree containing that pixel and satisfying the logical predicate $T_{\kappa_i}^\alpha$. Similarly to AP calculation for inclusion trees, this process also allows us to re-use the hierarchy in the calculation of subsequent AP components, and is elaborated in the remainder of the section.

For the sake of fair comparison with state-of-the-art, we follow here a marginal strategy as used in EAP and ESDAP. But let us recall that partitioning trees offer a greater flexibility than inclusion trees to deal with multivariate data such as hyperspectral images [23]. Exploring various dissimilarity metrics and constrained connectivity criteria is however left for future work.

3.1 Filtering a partitioning tree

Filtering a partitioning tree is more restrictive than inclusion trees, since removing certain nodes would invalidate the hierarchy (e.g. child nodes of any node need to be either all preserved or all removed, i.e. the same operation needs to be applied to all the “sibling” nodes). Therefore, the AP is calculated following the interpretation that the lowest node containing a pixel and satisfying the logical predicate $T_\alpha^{\kappa_i}$ determines the value of that pixel in the i^{th} component of the AP. It is possible that a region $R \subseteq E$ satisfying the logical predicate T_α^κ has both child nodes which satisfy and which do not satisfy T_α^κ . Thus, the region R will determine the value of a subset of pixels belonging to it while the values of the other pixels will be determined based on the values of the child and descendant regions. From the implementation point of view, as all the “sibling” regions need to be processed equally, so the child regions not satisfying the logical predicate T_α^κ are “collapsed” into their parents by assigning them the same node representation (cf. Sec. 3.2) as their parents, while their children are filtered from the tree. Similarly to AP calculation of inclusion trees, this enables us reusing the filtered hierarchy from one component of the AP to the next.

3.2 Node representation

Once a node in the hierarchy is selected to represent a pixel at a certain level of the AP, the pixel needs to be assigned to a value in the profile based on the selected node.

With the inclusion trees, such as the min and max-trees used in the original AP calculation, as well as their extension to the ToS, the choice of representation is straightforward. This is because the AP calculation on these hierarchies corresponds precisely to performing an AF repeatedly on the same image. As a consequence, a node which is selected for representing one of its pixels will represent all the pixels of the corresponding region. To obtain the pixel values, an image is reconstructed from the filtered hierarchy. The reconstruction is done by assigning to every pixel the level node in the hierarchy closest to the leaf which contains that pixel.

On the other hand, no universal solution exists for node representation with partitioning trees. While the problem of representation already exists during the filtering step, in the case of AP calculation the issue is even more complex due to the fact that a single tree node might not be chosen to represent all of the pixels belonging to it. We explore in this paper three different node representation strategies.

Firstly, we attempted **min and max** representation motivated by the fact that the original APs are a concatenation between a series of attribute thinnings and thickenings, that is two complementary series. In order to construct an anti-extensive (resp. extensive) series, the nodes are represented with the minimal (resp. maximal) pixel value they contain. While we obtain two complementary series to form an AP from, both of these series are based on representing the same nodes and simply changing their representation, thus introducing redundancy to the profile.

In order to avoid this redundancy, we further experimented with the **average** representation, where the results for each threshold value are used in the profile only once while the nodes are represented by the average pixel value present in them. This was also inspired by the typical choice of node representation when the partitioning trees are used for image simplification.

Finally, we considered a strategy based on the **level** representation of the nodes. This representation is inspired by considering that the values assigned to the pixels when calculating an AP from inclusion trees correspond to their assigned level in the tree. With our choices of partitioning trees, this means we are assigning the α (resp. ω) levels to the nodes from the α -tree (resp. ω -tree) hierarchy. This is motivated by the fact that their levels are actually what characterizes the nodes of the partitioning trees dictating their structure. Furthermore, an advantage of this approach is that it will always assign a single scalar value as a node representation, which becomes significant when dealing with multivariate data (i.e. when the values of the pixels belonging to a single α -CC or ω -CC are not scalar but rather come from multiple image channels).

4 Experiments

The main goal of the experiments presented here is to compare the proposed tree based image representations against established alternatives, as a basis for com-

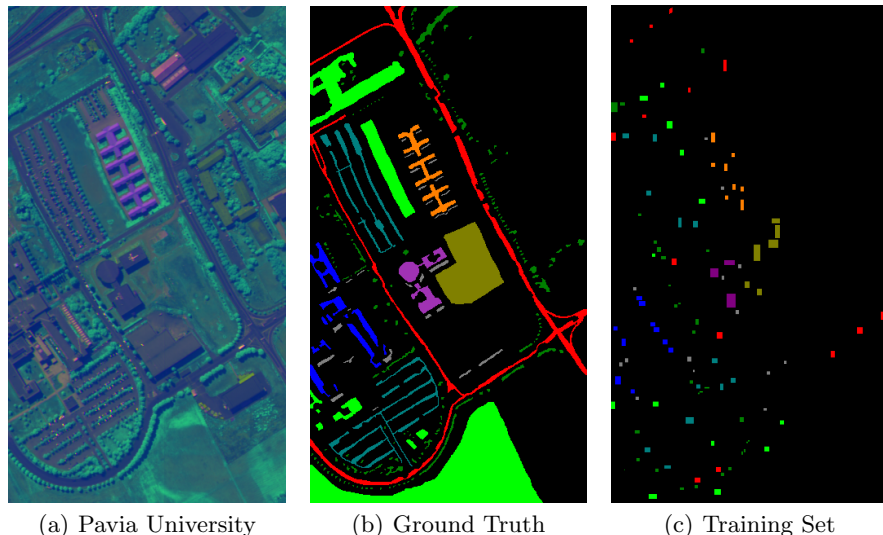


Fig. 3. The Pavia University dataset (false colors) and its corresponding ground truth; its thematic classes (training set size/ground truth size) are: ■ Asphalt (548/6631), ■ Trees (524/3064), ■ Bitumen (375/1330), ■ Meadows (540/18649), ■ Metal sheets (265/1345), ■ Shadows (231/947), ■ Gravel (392/2099), ■ Bare soil (532/5029) and ■ Self-blocking bricks (514/3682).

puting attribute profiles with the end goal of pixel classification. More precisely, we compare α and (ω) -trees against max/min trees and the tree of shapes.

The dataset under consideration is an urban area of size 340×610 pixels and 9 thematic classes, 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 $0.86\mu\text{m}$. After the elimination of 12 noisy bands, 103 bands have been left for processing (Fig. 3).

Classification has been realized 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 [24]. Classification performance has been measured by means of the κ statistic. For the Pavia University dataset, the learning step has been carried out using the standard training set widely used in the literature [2].

In this paper we are considering two of the most popular attributes: area and moment of inertia. For area thresholds (λ_a), we have computed the automatic settings according to [25] and for the moment of inertia (λ_m) we have employed the manual settings used in [2]: $\lambda_a = \{770, 1538, 2307, 3076, 3846, 4615, 5384, 6153, 6923, 7692, 8461, 9230, 10000, 10769\}$ and $\lambda_m = \{0.2, 0.3, 0.4, 0.5\}$. However, it is well known that attribute threshold selections have a very significant effect on the performance of AP [2], and the aforementioned settings, re-

Table 1. The classification performance (κ statistic) of α - and (ω) -tree based AP for different node representations and for both attributes as well as their combination.

	α -tree			(ω) -tree		
	area	moment	comb.	area	moment	comb.
max/min	0.9327	0.8521	0.9080	0.9264	0.8546	0.9429
average	0.9240	0.8468	0.9379	0.8876	0.8449	0.9302
level	0.8662	0.8477	0.8898	0.9482	0.8489	0.9479

ardless of whether they are automatic or manual, have been empirically determined by their respective authors with always the max/min strategy in mind. Which is why, for the sake of fairness, we employ multiple sets (Λ_a, Λ_m) of thresholds for both attributes, by simply scaling them with various multipliers $\mu = \{0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 3, 4, 5, 6, 7, 8\}$.

In order to compare the effectiveness of different node representations for partitioning trees, we compare the performance of different profiles at the threshold value $\mu = 1$. These results can be seen in Tab.4. It can be remarked that even though the max/min node representation effectively uses the same regions twice, producing the descriptor of twice the size and containing redundancy, it is still more effective than the average representation which attempts to represent each region only once by its average gray level. The level representation performs much better with the (ω) -tree than for the α -tree, which can be explained by its closer correspondence to region complexity due to global range constraints. The level representation also has the advantage of being half the size of the max/min representation (60 vs. 116 for the area attribute and 20 vs. 36 for the moments). For this reason, and as this representation corresponds most closely to the pixel value interpretation for the original AP, we chose the level representation as our preferred one. The additional advantage of this representation is its extendability to multivariate data: while the descriptor size for the min/max and the average representations will be multiplied by the number of bands in case of multivariate data representations, the length does not change when using the level representation.

We have also tried combining the moments and the area attribute, but have noticed that for the peak values of the area AP, addition of the moments AP does not improve performance. For this reason, the remainder of the experiments were carried out separately for the area and for the moments AP.

As far as their performance across a range of threshold values is concerned, in the case of the area attribute, the performances of all four tested trees are relatively similar and consistent, underlining the reliability of the attribute under consideration (Fig. 4(a)). All the same, one can still observe that the (ω) -tree systematically outperforms the α -tree. Even though (ω) -tree leads to the best scores for lower thresholds, it is eventually the tree of shapes, that achieves the overall best classification performance level with the area attribute.

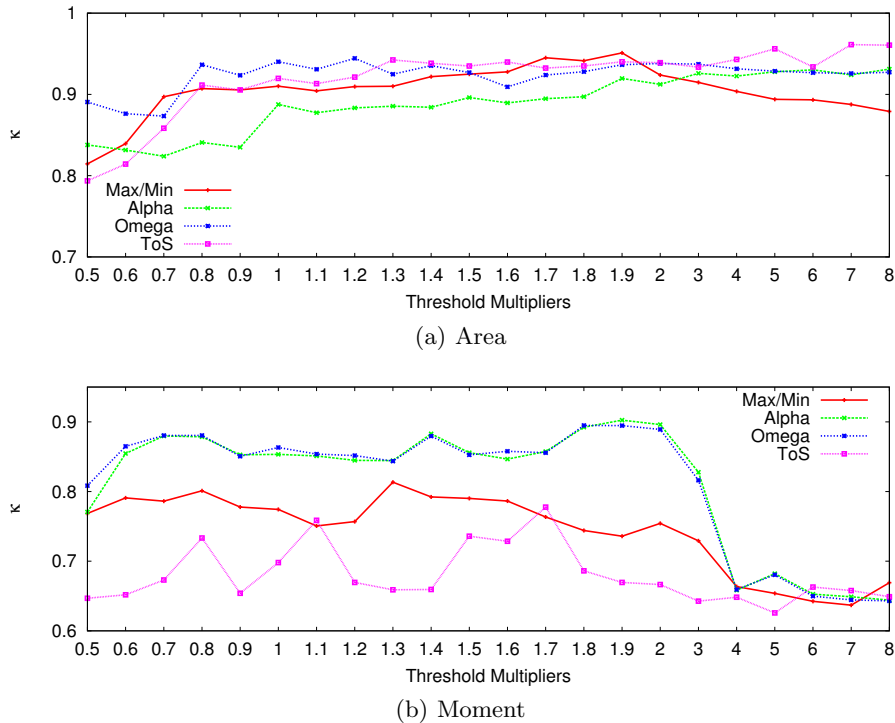


Fig. 4. Classifications scores (κ statistic) for the Pavia University dataset for the area and moment of inertia attributes using a variety of threshold settings.

As to the moment of inertia attribute (Fig. 4(b)), this time the relative scores vary at a greater level with respect to area. In more detail, the AP based on inclusion trees are clearly outperformed by both partition tree based profiles, while the previously best performing tree of shapes is now performing the worst, and with a higher level of variance across thresholds. Interestingly, both (ω)-tree and (α)-tree lead to very similar performances, both clearly surpassing their inclusion based counterparts.

5 Conclusion

Attribute profiles have been one of the most successful morphological tools in remote sensing recently. They allow for multiscale image description through successive attribute filterings, and can be efficiently computed using inclusion trees such as min- and max-tree (AP) and tree of shapes (SDAP). In this paper, we propose to compute these profiles using another class of tree representations, namely partitioning trees.

We consider in this first attempt α -tree and (ω)-tree, applied with a marginal strategy on each band of an hyperspectral image (similarly to the standard computation of extended AP). We discuss the specific challenges raised when computing attribute profiles from partitioning trees, in particular the way to represent tree nodes and build the AP feature vectors. Our preliminary results show the relevance of such an approach, and calls for further exploration.

Indeed, while we have used here a marginal strategy to ensure a fair comparison with existing AP definitions, using partitioning trees allows building a single tree from a whole dataset (either the original image or its first principal components). Future work will thus aim to evaluate different local dissimilarity metrics as well as constrained connectivity criteria (e.g. global range) defined in a multivariate space.

Acknowledgments

This work was supported by the French Agence Nationale de la Recherche (ANR) under reference ANR-13-JS02-0005-01 (Asterix project), by the BAGEP Award of the Science Academy and by the Turkish TUBITAK Grant 115E857.

References

1. Benediktsson, J., Pesaresi, M., Arnason, K.: Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. *IEEE Transactions on Geoscience and Remote Sensing* **41**(9) (2003) 1940–1949
2. Dalla Mura, M., Benediktsson, J.A., Waske, B., Bruzzone, L.: Morphological attribute profiles for the analysis of very high resolution images. *IEEE Transactions on Geoscience and Remote Sensing* **48**(10) (2010) 3747–3762
3. Dalla Mura, M., Benediktsson, J., Bruzzone, L.: Self-dual attribute profiles for the analysis of remote sensing images. In: *Mathematical Morphology and Its Applications to Image and Signal Processing (ISMM)*. (2011) 320–330
4. Carlinet, E., Geraud, T.: Mtos: A tree of shapes for multivariate images. *IEEE Transactions on Image Processing* **24**(12) (2015) 5330–5342
5. Dalla Mura, M., Atli Benediktsson, J., Waske, B., Bruzzone, L.: Extended profiles with morphological attribute filters for the analysis of hyperspectral data. *International Journal of Remote Sensing* **31**(22) (2010) 5975–5991
6. Dalla Mura, M., Villa, A., Benediktsson, J.A., Chanussot, J., Bruzzone, L.: Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis. *IEEE Geoscience and Remote Sensing Letters* **8**(3) (2011) 542–546
7. Aptoula, E.: Hyperspectral image classification with multi-dimensional attribute profiles. *IEEE Geoscience and Remote Sensing Letters* **12**(10) (2015) 2031–2035
8. Marpu, P., Pedernana, M., Dalla Mura, M., Benediktsson, J.A., Bruzzone, L.: Automatic generation of standard deviation attribute profiles for spectral spatial classification of remote sensing data. *IEEE Geoscience and Remote Sensing Letters* **10**(2) (2013) 293–297

9. Aptoula, E., Dalla Mura, M., Lefèvre, S.: Vector attribute profiles for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing* **54**(6) (2016) 3208–3220
10. Demir, B., Bruzzone, L.: Histogram-based attribute profiles for classification of very high resolution remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing* **54**(4) (2016) 2096–2107
11. Aptoula, E., Ozdemir, M.C., Yanikoglu, B.: Deep learning with attribute profiles for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters* **13**(12) (2016) 1970–1974
12. Ghamisi, P., Dalla Mura, M., Benediktsson, J.A.: A survey on spectral-spatial classification techniques based on attribute profiles. *IEEE Transactions on Geoscience and Remote Sensing* **53**(5) (2015) 2335–2353
13. Ronse, C.: Ordering partial partitions for image segmentation and filtering: merging, creating and inflating blocks. *Journal of Mathematical Imaging and Vision* **49**(2) (2014) 202–233
14. Salembier, P., Oliveras, A., Garrido, L.: Anti-extensive connected operators for image and sequence processing. *IEEE Transactions on Image Processing* **7**(4) (1998) 555–570
15. Monasse, P., Guichard, F.: Scale-space from a level lines tree. *Journal of Visual Communication and Image Representation* **11**(2) (2000) 224–236
16. Song, Y., Zhang, A.: Monotonic tree. In: *Discrete Geometry for Computer Imagery*, Springer (2002) 114–123
17. Ballester, C., Caselles, V., Monasse, P.: The tree of shapes of an image. *ESAIM: Control, Optimisation and Calculus of Variations* **9** (2003) 1–18
18. Soille, P.: On genuine connectivity relations based on logical predicates. In: *International Conference on Image Analysis and Processing*. (2007) 487–492
19. Soille, P.: Constrained connectivity for hierarchical image partitioning and simplification. *IEEE transactions on pattern analysis and machine intelligence* **30**(7) (2008) 1132–1145
20. Soille, P.: Preventing chaining through transitions while favouring it within homogeneous regions. In: *International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing*, Springer (2011) 96–107
21. Soille, P., Najman, L.: On morphological hierarchical representations for image processing and spatial data clustering. In: *Applications of Discrete Geometry and Mathematical Morphology*. Springer (2012) 43–67
22. Cavallaro, G., Dalla Mura, M., Benediktsson, J., Plaza, A.: Remote sensing image classification using attribute filters defined over the tree of shapes. *IEEE Transactions on Geoscience and Remote Sensing* **54**(7) (2016) 3899–3911
23. Lefèvre, S., Chapel, L., Merciol, F.: Hyperspectral image classification from multiscale description with constrained connectivity and metric learning. In: *IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*. (2014)
24. Breiman, L.: Random forests. *Machine Learning* **45**(1) (2001) 5–32
25. Ghamisi, P., Benediktsson, J.A., Sveinsson, J.R.: Automatic spectral–spatial classification framework based on attribute profiles and supervised feature extraction. *IEEE Transactions on Geoscience and Remote Sensing* **52**(9) (2014) 5771–5782