

# HYPERSPECTRAL IMAGE CLASSIFICATION FROM MULTISCALE DESCRIPTION WITH CONSTRAINED CONNECTIVITY AND METRIC LEARNING

*Sébastien Lefèvre, Laetitia Chapel, and François Merciol*

Université de Bretagne Sud – IRISA  
Campus de Tohannic, BP 573, 56017 Vannes Cedex, France  
{firstname.lastname}@irisa.fr

## ABSTRACT

Mapping of remote sensing data is usually done through image classification. For hyperspectral images, the classification process often relies only on the spectral signature of each single pixel. Nevertheless, combining spatial and spectral features has been a promising way for accuracy improvement. We address here this problem by computing spectral features from spatially identified regions, sampled from a hierarchical image representation, namely  $\alpha$ -tree, built with prior knowledge. The sampling of the tree nodes (i.e., regions) is based on the paradigm of constrained connectivity and the global range criterion. In this paper, we extend this criterion to hyperspectral data and apply it to our knowledge-based  $\alpha$ -tree. Our results show an improvement of pixelwise classification accuracy over spectral features only.

**Index Terms**— Hyperspectral Image, Metric Learning, Constrained Connectivity, Hierarchical Representation,  $\alpha$ -Tree,  $(\alpha, \omega)$ -Connected Components.

## 1. INTRODUCTION

Thanks to the continuous growth of high and very high spatial resolution sensors, remote sensing images contain great details that can be exploited successfully in various tasks such as object extraction or classification. At the same time, standard pixel-based methods for image understanding are not adapted anymore as they can not deal with the amount of information and with the heterogeneity of the scene. It is then mandatory to rely on image representation of higher level, for instance through regions produced by a segmentation process. To face the challenging issue of building a relevant segmentation from satellite imagery, different multiscale representations have been proposed in the literature and we focus here on the  $\alpha$ -tree model [1], inspired by the well-known principle of single linkage in pattern recognition.

In a previous work [2], we have explored how to build this model from hyperspectral data, and shown the benefit of using metric learning to measure local dissimilarity between neighboring pixels. Nevertheless, the  $\alpha$ -tree is limited by the chaining effect (as any single linkage procedure). Constrained

connectivity has been proposed to solve this issue [3], leading to the definition of  $(\alpha, \omega)$ -connected components. However, the existing computation schemes are not able to deal with high dimension spaces such as hyperspectral images. In this article, we first address this issue and introduce a method for extracting  $(\alpha, \omega)$ -connected components from an hyperspectral image through its  $\alpha$ -tree. We then extract spatial-spectral features by sampling the hierarchical representation at different  $(\alpha, \omega)$  values, which allows to model the spatial context of each pixel. These extracted features are then the inputs of a SVM classifier [4]. We illustrate the interest of our approach by some improvement in pixelwise classification.

This paper is organized as follows. We recall in Sec. 2 the concept of  $\alpha$ -tree, and in particular its extension to hyperspectral images using metric learning. Sec. 3 deals specifically with the drawback of the chaining effect, that is solved with constrained connectivity through a global range constraint. We propose an efficient approximation of this measure for hyperspectral images. The interest of the proposed method is illustrated through experimental results in Sec. 4 before concluding and suggesting directions for future work.

## 2. KNOWLEDGE-BASED $\alpha$ -TREES

### 2.1. $\alpha$ -trees

An  $\alpha$ -tree is a multiscale representation of an image through its  $\alpha$ -connected components (or  $\alpha$ -CCs). While it finds roots in early work in computer vision, it has been revisited only recently by Soille and Ouzounis [5, 1]. This paradigm is very related to the single linkage procedure used in data clustering.

The concept of  $\alpha$ -CC is an extension of the connected component (or CC). We recall that the latter is defined as a set of adjacent pixels that share the same value (either scalar for panchromatic images, or vectorial for multi- or hyperspectral ones). Representing an image by its CCs allows for higher-level analysis (similarly to computer vision techniques relying on superpixels). However, the possibly great number of CCs in an image prevents their practical use. Indeed, adjacent pixels may belong to the same structure but have slightly different values, thus belonging to different CCs. The concept

of  $\alpha$ -CC has been introduced to allow such slight variations, leading to the following definition: an  $\alpha$ -CC is a set of adjacent pixels that share similar values i.e., values with a difference lower or equal to a threshold  $\alpha$ . The  $\alpha$ -CC of a pixel  $p$  will thus contain all pixels  $q$  that can be reached with a path over neighboring pixels  $p_i$  ( $p_1 = p, \dots, p_n = q$ ) from  $p$  to  $q$  such as  $d(p_i, p_{i+1}) \leq \alpha$ .

The complexity and number of  $\alpha$ -CCs are directly related to  $\alpha$ . It allows one to build a hierarchical representation of an image, and to perform subsequent multiscale analysis (e.g., in an object-oriented strategy). This representation is called an  $\alpha$ -tree. Each level of the tree is indexed by an  $\alpha$  value, and its nodes are the corresponding  $\alpha$ -CCs. A leaf in the tree is a 0-CC i.e., a standard CC in the image. Increasing  $\alpha$  leads to the connection of  $\alpha$ -CCs, resulting in the creation of higher nodes in the tree, until the root that contains the whole image.

## 2.2. Extension to hyperspectral data

Constructing an  $\alpha$ -tree requires to define a similarity (or dissimilarity) measure between values of any pair of adjacent pixels. While such a dissimilarity can be simply written as a difference of grey levels in the panchromatic case, there is no universal solution for multidimensional spaces. Among the most usual choices, we can cite  $L_q$  norms (with  $q = 1, 2$  or even  $\infty$ ) as well as SAM (Spectral Angle Mapper).

However, we have shown in a recent study [2] that these norms can be advantageously replaced by some other metrics relying on prior knowledge provided with the data. This option is particularly relevant in the case of hyperspectral satellite images, where classes or labels are often provided by an expert with some training samples (pixels).

More precisely, we have explored the use of the Euclidean norm using a Mahalanobis metric  $M$  learned from data i.e.,

$$d(\mathbf{a}, \mathbf{b}) = \left( (\mathbf{a} - \mathbf{b})^T M (\mathbf{a} - \mathbf{b}) \right)^{0.5}$$

Learning the  $M = WW^T$  metric is driven by data labels, and more particularly by a set  $\mathcal{S}$  of *must-link* constraints and a set  $\mathcal{D}$  of *cannot-link* constraints:

$$\begin{aligned} \mathcal{S} &= \{(\mathbf{a}, \mathbf{b}) \mid \mathbf{a} \text{ and } \mathbf{b} \text{ belong to the same class}\} \\ \mathcal{D} &= \{(\mathbf{a}, \mathbf{b}) \mid \mathbf{a} \text{ and } \mathbf{b} \text{ belong to different classes}\} \end{aligned}$$

Matrix  $W$  is then built such as the sum of squared distances of points from  $\mathcal{S}$  is minimal, and the sum of points from  $\mathcal{D}$  is maximal. In other words, we have the following optimization problem:

$$W^* = \arg \max_{W^T W = I} \frac{\text{trace}(W^T \hat{S}_{\mathcal{S}} W)}{\text{trace}(W^T \hat{S}_{\mathcal{D}} W)},$$

where  $\hat{S}_{\mathcal{S}}$  and  $\hat{S}_{\mathcal{D}}$  are covariance matrices of sets  $\mathcal{S}$  and  $\mathcal{D}$ , solved here using [6]. The final Mahalanobis metric is then built such as  $M = W^*(W^*)^T$ .

In this paper, we build upon this previous work to propose a multiscale description scheme of hyperspectral images. More precisely, we address the major drawback of the  $\alpha$ -tree, namely the chaining effect, and solve it with the paradigm of constrained connectivity proposed by Soille [3].

## 3. CONSTRAINED CONNECTIVITY

The chaining effect is a well-known issue for the single-linkage paradigm. Since the  $\alpha$ -tree is built on this paradigm, it is naturally affected by this problem. To explain, let us consider a series of neighboring pixels with values  $u_1, \dots, u_n$ . If  $d(u_i, u_{i+1}) \leq \alpha$  for all pairs  $(i, i+1)$ , these pixels will belong to the same  $\alpha$ -CC. This remark holds even for a small value of  $\alpha$ , values of neighboring pixels being then always very similar. However, the difference between any pair of pixels within the  $\alpha$ -CC (e.g.,  $d(u_1, u_n)$ ) might be much larger than  $\alpha$ . In this case, pixels with very distinct values are gathered in a unique  $\alpha$ -CC assumed to be of low complexity (low  $\alpha$  value). Such a situation is frequently observed on transition regions between objects of different spectral signatures, due to the discretization of edges in digital images [7].

### 3.1. Global range and $(\alpha, \omega)$ -connected components

To counter this chaining effect, Soille has introduced the paradigm of constrained connectivity [3]. It consists in imposing a set of constraints to be fulfilled by the  $\alpha$ -CCs. The most representative example is to apply a threshold  $\omega$  of the global range of the  $\alpha$ -CC i.e.,  $d(p, q) \leq \omega$  for all pairs  $(p, q)$  of pixel values belonging to this  $\alpha$ -CC. Such a criterion leads to the definition of the  $(\alpha, \omega)$ -CCs. Computing this global range is efficiently achieved in the grayscale case by storing the minimum and maximum values of each  $\alpha$ -CC, and updating these extrema when merging two  $\alpha$ -CCs in the  $\alpha$ -tree construction process.

However, defining constrained connectivity on multivariate data (e.g., hyperspectral images) remains very challenging: i) it is not possible anymore to compute the global range of a given  $\alpha$ -CC from its minimum and maximum; ii) there is no universal ordering for vectorial data, and thus no unique way to determine minimum and maximum of a set of vectors (i.e., spectral signatures of pixels contained in the  $\alpha$ -CC). The global range can only be computed using an exhaustive search, thus requiring to compute distances between any pair of pixels in the  $\alpha$ -CC. Furthermore, it has to be done each time two  $\alpha$ -CCs are merged in the tree. Considering the higher bound of the tree size (i.e. number of nodes) is twice the image size, it leads to an unfordable  $O(n^3)$  complexity.

### 3.2. Global range for multivariate data

Only very few attempts have been made to address this issue. The first solution has been proposed by Soille in [3, 8]

in the context of remote sensing. A marginal strategy was suggested, where both  $\alpha$  and  $\omega$  criteria shall be fulfilled in all spectral bands to build the  $(\alpha, \omega)$ -CCs. This presents several drawbacks: i) each band is processed independently of the others, ignoring the possible correlation among bands; ii) the same (local or global) range is imposed on all bands of the image, while the behavior of spectral signatures of different materials might vary very much in the observation spectrum.

Conversely, a purely vectorial solution was recently proposed [9]. Assuming a total ordering can be imposed on pixel values, each pixel is described by its rank within the pixel value space. The range of the multivariate set is then defined as the range of the ranks within the set. While this approach exhibits appealing properties, it still requires the definition of a total ordering. To do so, the standard lexicographical ordering or one of its variants needs to prioritize the image channels. While this is possible for color images [10], there is no adequate solution for hyperspectral data yet.

### 3.3. Efficient overestimation for hyperspectral images

Considering the computation of the global range cannot be achieved efficiently on multivariate data such as hyperspectral images, we suggest here to rely on an overestimation of it. This approximation is relevant as long as the overestimated global range is still an increasing criterion w.r.t.  $\alpha$ .

Representing an  $\alpha$ -CC by the set of multivariate data it contains, the simplest scheme to overestimate its global range is to compute the diagonal of the set. To do so, we can approximate the set by its bounding box. We only need then to store the minimum and maximum values within each band, which could be very easily updated when merging two  $\alpha$ -CCs.

However, the overestimation error will increase exponentially with the number of dimensions. So we rather rely on the diameter of the bounding sphere. It is known that this technique leads to much lower estimation error than the bounding box one, when the number of dimensions is high (e.g., in the case of hyperspectral images). Similarly to the case of the bounding box, we can rely on efficient schemes to compute bounding spheres. We only need to store the center of the sphere (i.e., the center of gravity of the set of pixels contained in the  $\alpha$ -CC), as well as its radius (i.e., the distance between the center and the furthest element within the  $\alpha$ -CC, which might not correspond to an actual observed pixel). When merging two  $\alpha$ -CCs, updating these attributes is trivial: the new center of gravity is a linear combination of the centers of gravity of the two  $\alpha$ -CCs to be merged (we also store the  $\alpha$ -CCs size); its furthest element is the extreme point at the surface of one of the two spheres from the merging  $\alpha$ -CCs.

## 4. EXPERIMENTS

We consider two widely benchmarked hyperspectral images acquired by the ROSIS sensor: a part of *Pavia center* (Fig. 1),

similarly to our previous work [2], and *Pavia University*, respectively made of 102 and 103 spectral bands. The first is a  $1096 \times 715$  image, containing 27,019 labeled pixels. The second contains  $610 \times 340$  pixels, among which 42,776 are labeled. Both contain nine reference classes of urban, soil and vegetation features. In order to classify the pixels, we consider two settings for pixel feature computation:

- features are directly the values in the spectral bands;
- features are derived from an  $\alpha$ -tree: we extract  $(\alpha, \omega)$ -CCs from the  $\alpha$ -tree using a regular sample of  $\omega$  values, and describe each of them with its size, its within-cluster variance, and the mean spectral signature (averaged over all pixels of the  $(\alpha, \omega)$ -CCs) of all its bands. Concatenation of these descriptors for all sampled  $\omega$  values constitutes the complete set of features.

We focus on *one-against-one* SVM (e.g., see [11]) to perform the classification, using the R implementation Kernlab [12]. We restrict ourselves to a Gaussian kernel, for which admissible couple of parameters  $C$ , that controls the effect of misclassified pixels, and  $\sigma$ , that controls the spread of the kernel, have been defined using a grid search. The best set of parameters is then chosen using a repeated random sub-sampling scheme by running 10 experiments on each couple of parameters. On both cases, we use 5% of randomly chosen pixels to learn the classifier. The  $\alpha$ -tree has been computed using a metric learned from the data, where the same 5% of the pixels have been integrated into sets of *must-link* and *cannot-link* constraints in order to learn the distance metric. 10 values of  $\omega$  have been chosen to construct the tree-derived features.

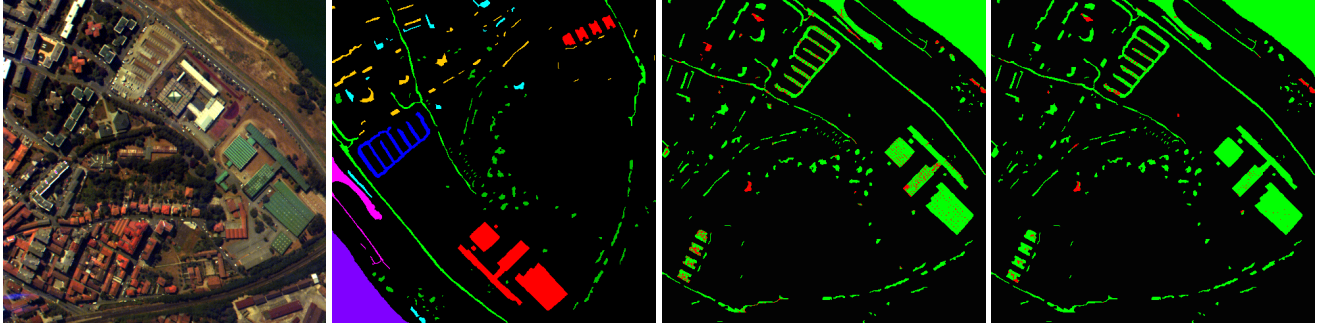
Method	<i>Pavia University</i>	<i>Pavia Center</i>
	OA (%)	OA (%)
Spectral bands features	$93.19 \pm 0.19$	$93.37 \pm 0.90$
Tree-derived features	$98.14 \pm 0.25$	$97.00 \pm 0.62$

**Table 1.** Mean  $\pm$  standard deviation of overall accuracies (OA) computed over 10 repetitions for the 2 tested methods.

Tab. 1 reports the overall accuracies and Fig. 1 shows the pixels that have been correctly or mis-classified. We note that the tree-derived features gives the best classification results. It allows one to implicitly take into account spatial information, leading to less isolated pixels in the classification map.

## 5. ACKNOWLEDGMENTS

The authors acknowledge the support of the French Agence Nationale de la Recherche (ANR) under reference ANR-13-JS02-0005-01 (Asterix project). They also wish to thank Paolo Gamba from University of Pavia and the HySens project for providing the data included in the Pavia datasets.



**Fig. 1.** *Pavia center* dataset and classification results. From left to right: false color composition, ground truth, and results obtained using the pixelwise spectral bands dataset and those with the tree-derived features dataset: pixels correctly classified are depicted in green while the misclassified ones are in red.

## 6. CONCLUSION

Representation of an image through its  $\alpha$ -connected components can lead to a powerful hierarchical model known as  $\alpha$ -tree. Furthermore, it can be improved using the constrained connectivity paradigm, leading to  $(\alpha, \omega)$ -connected components. These representations allow multiscale image analysis of an image and have been successfully applied in various applications in remote sensing, but mostly for panchromatic images. Nevertheless, their extension to hyperspectral data is not trivial. The two main stumbling blocks that are to be overcome in such high dimension spaces are i) the definition of a sensible distance, ii) the design of efficient schemes to support constrained connectivity. While we have addressed the first question in a recent study [2], we focus here on the second issue and propose new approximation schemes for computing  $(\alpha, \omega)$ -connected components and extracting them from an  $\alpha$ -tree. First results are promising but still require a deeper investigation, including a comparison with other spatial-spectral classification techniques, as well as using such trees for other challenges raised in remote sensing.

## 7. REFERENCES

- [1] G. K. Ouzounis and P. Soille, "Pattern spectra from partition pyramids and hierarchies," in *International Symposium on Mathematical Morphology*. 2011, vol. 6671 of *LNCS*, pp. 108–119, Springer.
- [2] F. Merciol, L. Chapel, and S. Lefèvre, "Hyperspectral image representation through  $\alpha$ -trees," in *ESA Conference on Image Information Mining*, 2014, pp. 37–40.
- [3] P. Soille, "Constrained connectivity for hierarchical image partitioning and simplification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 7, pp. 1132–1145, 2008.
- [4] L. Bruzzone and L. Carlin, "A multilevel context-based system for classification of very high spatial resolution images," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 44, no. 9, pp. 2587–2600, Sept 2006.
- [5] P. Soille, "Constrained connectivity for hierarchical image partitioning and simplification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 7, pp. 1132–1145, July 2008.
- [6] S. Xiang, F. Nie, and C. Zhang, "Learning a mahalanobis distance metric for data clustering and classification," *Pattern Recognition*, vol. 41, no. 12, pp. 3600–3612, 2008.
- [7] P. Soille, "Preventing Chaining through Transitions While Favouing It within Homogeneous Regions," in *International Symposium on Mathematical Morphology*. 2011, vol. 6671 of *LNCS*, pp. 96–107, Springer.
- [8] P. Soille, "Constrained connectivity for the processing of very high resolution satellite images," *International Journal of Remote Sensing*, vol. 31, no. 22, pp. 5879–5893, 2010.
- [9] J. Weber, E. Aptoula, and S. Lefèvre, "Vectorial quasi-flat zones for color image simplification," in *International Symposium on Mathematical Morphology*. May 2013, vol. 7883 of *LNCS*, pp. 231–242, Springer.
- [10] E. Aptoula and S. Lefèvre, "A comparative study on multivariate mathematical morphology," *Pattern Recognition*, vol. 40, no. 11, pp. 2914–2929, 2007.
- [11] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 8, pp. 1778–1790, 2004.
- [12] A. Karatzoglou, A. Smola, K. Hornik, and A. Zeileis, "kernlab – An S4 package for kernel methods in R," *Journal of Statistical Software*, vol. 11, no. 9, pp. 1–20, 2004.