HYPERSPECTRAL IMAGE REPRESENTATION THROUGH α-TREES

François Merciol, Laetitia Chapel and Sébastien Lefèvre

Université de Bretagne Sud – IRISA
Campus de Tohannic, BP 573, 56017 Vannes Cedex, France

ABSTRACT

α-trees provide a hierarchical representation of an image into partitions of regions with increasing heterogeneity. This model, inspired from the single-linkage paradigm, has recently been revisited for grayscale images and has been successfully used in the field of remote sensing. This article shows how this representation can be adapted to more complex data here hyperspectral images, according to different strategies. We know that the measure of distance between two neighbouring pixels is a key element for the quality of the underlying tree, but usual metrics are not satisfying. We show here that a relevant solution to understand hyperspectral data relies on the prior learning of the metric to be used and the exploitation of domain knowledge.

***Index Terms*—** Hyperspectral Images, Metric Learning, Hierarchical Representations, α-Trees

1. INTRODUCTION

With the continuous growth of high and very high spatial resolution sensors, standard pixel-based methods for image understanding are not adapted anymore. Indeed, it is mandatory to rely on image representation of higher level, for instance through regions produced by a segmentation process. This approach is known as *object-based image analysis* [1] in remote sensing. 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 α-tree model recently revisited by Ouzounis and Soille [2, 3], inspired by the well-known principle of single linkage in pattern recognition [4]. Describing a panchromatic (greyscale) image through its α-tree is useful to identify damaged buildings after natural hazards [2] or to select interest zones in an image (or a feature space) [5], while dealing with multispectral images (e.g. [6]) remains mainly unexplored.

In this article, we propose to extend α-trees to deal with more complex data, and particularly hyperspectral images. This issue has been tackled recently in [7] where the authors have proposed to use a mutual information criterion. We suggest here another strategy using prior knowledge, brought by an expert through labeling of learning samples. We show how, in an unsupervised context, metric learning can help to elaborate a more relevant dissimilarity measure and then to improve the quality of the subsequent hierarchical representation.

This paper is organized as follows. We recall in section 2 the concept of α-tree through introducing required notations. Section 3 deals specifically with extending α-trees to hyperspectral images. The interest of the proposed method is illustrated through experimental results in section 4 before concluding and suggesting directions for future work.

2. α-TREES AND HIERARCHIES

We recall here the definition of an α-tree, which is a multiscale representation of an image through its α-zones, using notations from [2]. Let I be an image defined on a domain E. The segmentation of an image is a partition $\mathbf{P}$ of E, or projection $x \to \mathbf{P}(x)$ of E in $\mathcal{P}(E)$ such as $\forall x \in E \Rightarrow x \in \mathbf{P}(x)$ and $\forall x, y \in E \Rightarrow \mathbf{P}(x) = \mathbf{P}(y)$ or $\mathbf{P}(x) \cap \mathbf{P}(y) = \emptyset$ with $\mathbf{P}(x)$ representing a cell of $\mathbf{P}$ containing a point $x \in E$. We have thus $\bigcup_{x\in E} \mathbf{P}(x) = E$.

We also write $\pi(x \rightsquigarrow y)$ a path of length N between two elements $x, y \in E$, i.e. a chain of adjacent elements $\langle x = x_0, x_1, \ldots, x_{N-1} = y\rangle$. $\Pi \neq \emptyset$ is the set of all paths linking x to y. Minimal dissimilarity between x and y is defined by

$$\hat{d}(x, y) = \bigwedge_{\pi\in\Pi} \left\{ \bigvee_{i\in[0,\ldots,N-1]} \{d(x_i, x_{i+1}) \mid x_i, x_{i+1} \in \Pi\} \right\}$$

where $d(x, y)$ is a measure of dissimilarity between attributes from pixels x and y (e.g. their intensities or grey values). For a given pixel x, its α-zone written α-$\mathcal{Z}(x)$ is composed of all pixels linked to x by a path with local steps not higher than α:

$$\alpha\text{-}\mathcal{Z}(x) = \{x\} \cup \{y \mid \exists\pi(x \rightsquigarrow y) : \forall x_i \in \pi(x \rightsquigarrow y) \wedge x_i \neq y \Rightarrow d(x_i, x_{i+1}) \leq \alpha\}$$

α-tree is then a pyramid of partitions of E into α-zones, or projections $\Delta^A : E \to \Pi^A(E)$ defined by

$$\forall\alpha, \alpha' \in A,\ \alpha' < \alpha, \quad \Delta^A = \left\{\mathbf{P}^{\alpha=\mathbf{0}}, \mathbf{P}^{\alpha=\mathbf{1}}, \ldots, \mathbf{P}^{\alpha=\alpha_{\max}}\right\} \mid \mathbf{P}^{\alpha'} \preceq \mathbf{P}^{\alpha}$$

with $\Pi^A(E)$ and $A = [0, 1, \ldots, \alpha_{\text{max}}]$ denoting respectively the sets of all α partitions of E and of α values. The relation $\preceq$ corresponds to the ordering w.r.t. $\alpha \in A$:

$$\forall x \in E, \alpha' < \alpha \Rightarrow \alpha'\text{-}\mathcal{Z}(x) \subseteq \alpha\text{-}\mathcal{Z}(x) \Rightarrow \mathbf{P}^{\alpha'} \preceq \mathbf{P}^{\alpha}$$

A level of the pyramid $\Delta_\alpha^A \in \Delta^A$ is a partition $\mathbf{P}^\alpha$ of E, with $\alpha \in A$. Let $j \in J^\alpha$, in which $J^\alpha \subseteq \mathbb{Z}$ is a set of indices, used to access α-zones composing $\mathbf{P}^\alpha$. The hierarchy of partitions α written $\mathbf{\Lambda}^A$ is a familly of ordered projections $\mathbf{\Lambda}_\alpha^A : J^\alpha \to K^\alpha$ with $K^\alpha \subseteq J^\alpha$, i.e.

$$\begin{aligned} &\forall \alpha, \alpha' \in A,\ \alpha' < \alpha, \\ &\qquad \mathbf{\Lambda}^A = \left\{ \mathbf{\Lambda}_{\alpha=0}^A, \mathbf{\Lambda}_{\alpha=1}^A, \ldots, \mathbf{\Lambda}_{\alpha=\alpha_{\text{max}}}^A \right\} | \mathbf{\Lambda}_{\alpha'}^A \prec \mathbf{\Lambda}_\alpha^A \end{aligned}$$

and

$$\begin{aligned} &\forall \alpha \in A \setminus 0, \forall j \in J^\alpha, \\ &\qquad \mathbf{\Lambda}_\alpha^A = \left\{ \alpha_j\text{-}\mathcal{Z} | (\alpha_j\text{-}\mathcal{Z} \in \Delta_\alpha^A) \wedge (\alpha_j\text{-}\mathcal{Z} \notin \Delta_{\alpha-1}^A) \right\} \end{aligned}$$

Different hierarchical models for image representation have been proposed in the literature [8], e.g. min-/max-tree or binary partition tree. The former requires to impose an ordering relation on the input data, which is a serious issue when dealing with vectorial data such as multi- or hyperspectral images [9]. The latter, while offering a richer representation than α-trees, requires a much higher computation time and assumes the existence of predefined criteria to merge nodes [10, 11]. Wavelets and Gaussian pyramids belong to another set of multiscale representations which do not intrinsically provide image partitions at different scales.

Thanks to their simplicity, α-trees also offer the advantage of being possibly built efficiently with adapted algorithms (ca. one minute for a satellite image of 20 millions pixels) [12]. Once these representations have been built, it is possible to analyse the underlying image in an interactive (i.e. real-time) framework. Interactive segmentation of color video sequences have been addressed with this model in [13]. However, extending these tree models to complex data is still an open problem. We propose here to address this issue by relying on a machine learning technique recently introduced in image processing but in a different context [14].

3. EXTENSION TO THE HYPERSPECTRAL CASE

Computing an α-tree requires to define a similarity measure between attributes of any pair of adjacent pixels x and y. This dissimilarity can be simply written as a difference of grey levels in the panchromatic case, or more generally as any norm L_q (with $q = 1, 2$ or even ∞). Let us note however that such a measure is very sensible to the chaining effect observed with edge discretization in digital images [15]: it is then possible to enrich this measure by taking into account correlations between values of neighbouring pixels [7].

Relying on more complex metrics is necessary as soon as multi- or hyperspectral images are considered, since pixels are then embedded in high and even very high dimensional spaces. Thus, it has been proposed in [13] to use the Chebyshev distance (or L_∞ norm) to build the α-tree, in order to deal with color video sequences. In the hyperspectral case, it has been however shown in [16] that angular distance or *spectral angle mapper* (SAM)

$$d(\mathbf{a}, \mathbf{b}) = \arccos \left(\frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \, \|\mathbf{b}\|} \right)$$

with $\|\cdot\|$ the Euclidean norm and $\cdot$ the scalar product, provides better results than Euclidean distance.

The various distances discussed previously are computed without relying on any prior knowledge provided with the data. In the case of hyperspectral satellite images, it is possible to exploit learning sets provided by an expert as training samples (pixels). This correspond to the general setting of supervised classification very commonly addressed in remote sensing. We can then rely on a distance built from a learning process, or in other words driven by prior knowledge, such as for instance the Euclidean norm using a Mahalanobis metric learnt from data [17], 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 here 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 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 [14]. The final Mahalanobis metric is then built such as $M = W^*(W^*)^T$.

4. EXPERIMENTS

We evaluate here the benefit of learning the metric required to build an α-tree when this tree is representing an hyperspectral satellite image. In this context, the goal is to perform a semi-supervised clustering of the image, where each region (or α-zone) is associated with a land cover class.

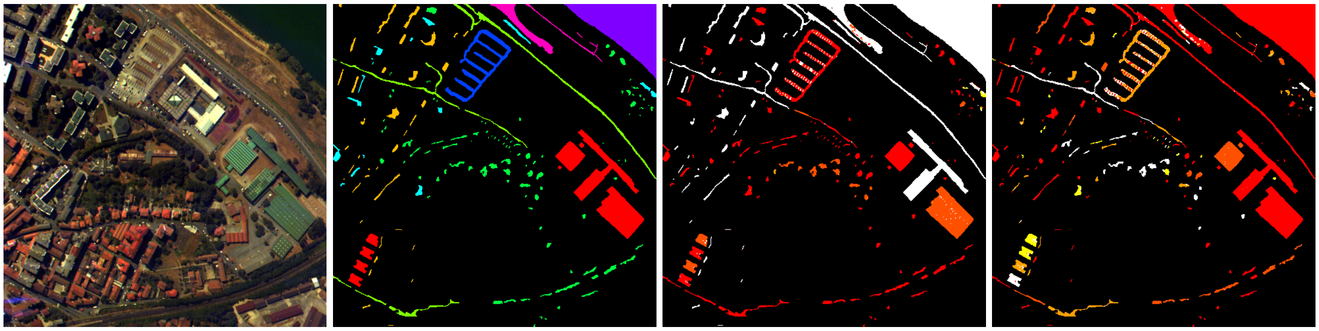

Fig. 1. *Pavia center* dataset and results. From left to right: false color composition, ground truth, and maps of relative F1-measure obtained with SAM and learnt metrics. For these two maps, the color code for the difference between pixel F1-measure and the best value among all the metrics is the following: 0 (red), less than 0.01 (orange-red), less than 0.05 (orange), less than 0.10 (yellow), higher than 0.10 (white).

4.1. Evaluation procedure

Contrary to [11], we have not introduced here any heuristic to select an optimal segmentation or partition from the hierarchical image representation. Defining such an heuristic is an open and topical problem [18]. It is thus not possible to use a quality measure based on the precision of a supervised classification applied on regions extracted from this partition.

In order to counter the lack of an optimal segmentation criterion, we exploit the ground truth available as in the context of supervised classification, as proposed in [19]. Thus we can perform an exhaustive evaluation of all tree nodes, by measuring for each of the node a local quality score using F1-measure (inspired from modified Jaccard index J' of [19]). This measure is a trade-off between recall and precision, and is given by $F1 = 2 \times TP/(2 \times TP + FN + FP)$ where TP, FN, and FP denote respectively the numbers of correctly classified pixels, false negatives and false positives. These quantities are evaluated through a comparison with reference regions provided by ground truth.

Each pixel for which the reference label is known is then assigned a best theoretical score, computed as the highest F1-measure among all nodes it is included in. These pixel scores are averaged over the whole image to provide a mean to assess the tree quality w.r.t. the available ground truth.

4.2. Data and results

Dataset considered here is a part of *Pavia center* hyperspectral image (Fig. 1), made of 102 spectral bands and acquired with ROSIS sensor. Image size is 492×492 pixels and 8 classes of interest have been defined. Ground truth is partially known: $27,019$ pixels are labelled. Metrics under evaluation are L_q (with $q = 1, 2, \infty$), SAM and the learnt metric. In order to learn the distance metric from the data, 1% of pixels in each class have been randomly integrated into sets of *must-link* and *cannot-link* constraints. The F1-measure provided here is averaged on 10 repetitions.

Tab. 1 and Fig. 1 provide an overview of the obtained results. We can make the following observations:

- benefit of using SAM angular distance over standard metrics such as L_q is still to be demonstrated when these distances are used to build α-trees from hyperspectral images. As shown in the figure, some large areas have a rather low F1-measure in comparison to the other metrics;
- learning a Mahalanobis metric, and more generally taking into account expert knowledge as early as possible, is a promising approach to build α-trees adapted to hyperspectral image analysis and classification.

Metric	Average F1 (%)
SAM	68.44
L_1	83.79
L_2	87.49
L_∞	84.21
Learnt metric	**93.02**

Table 1. Average F1-measures for trees built with different metrics and considering a part of *Pavia center* dataset.

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.

6. CONCLUSION

Hierarchical representation of an image through its α-zones or α-connected components offers several advantages: compact description, low computational cost, and promising applications in remote sensing (mostly for panchromatic images). Extending these representations to more complex images, such as hyperspectral images, is however not trivial.

We have here underlined the limitations of standard metrics, such as L_q distance [13] or angular distance [16]. We have also shown the benefits brought by metrics driven by prior knowledge, in a semi-supervised learning framework. These first results, while promising, require a deeper validation on a larger dataset and considering a wider range of metrics (especially the local information criterion used very recently in this context [7]). Finally, a more general study of hierarchical image representations is also needed to evaluate the interest of α-trees over the other existing models in the context of hyperspectral images considered in this work.

7. REFERENCES

[1] T Blaschke, "Object based image analysis for remote sensing," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, no. 1, pp. 2–10, 2010.

[2] 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.

[3] G. Ouzounis and P. Soille, "The alpha-tree algorithm," Tech. Rep. 25500 EN, Joint Research Centre, European Commission, 2012.

[4] R. Sibson, "SLINK: an optimally efficient algorithm for the single-link cluster method," *The Computer Journal*, vol. 16, no. 1, pp. 30–34, 1973.

[5] G.K. Ouzounis, V. Syrris, L. Gueguen, and P. Soille, "The switchboard platform for interactive image information mining," in *Image Information Mining Conference*, 2012.

[6] 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.

[7] Lionel Gueguen, Santiago Velasco-Forero, and Pierre Soille, "Local mutual information for dissimilarity-based image segmentation," *Journal of Mathematical Imaging and Vision*, pp. 1–20, 2013.

[8] P. Soille and L. Najman, "On morphological hierarchical representations for image processing and spatial data clustering," in *Applications of Discrete Geometry and Mathematical Morphology*. 2012, vol. 7346 of *LNCS*, pp. 43–67, Springer.

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

[10] P. Salembier and L. Garrido, "Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval," *IEEE Transactions on Image Processing*, vol. 9, no. 4, pp. 561–576, 2000.

[11] S. Valero, P. Salembier, and J. Chanussot, "Hyperspectral image representation and processing with binary partition trees," *IEEE Transactions on Image Processing*, vol. 22, no. 4, pp. 1430–1443, 2013.

[12] J. Havel, F. Merciol, and S. Lefèvre, "Efficient schemes for computing α-tree representations," in *International Symposium on Mathematical Morphology*. 2013, vol. 7883 of *LNCS*, pp. 111–122, Springer.

[13] F. Merciol and S. Lefèvre, "Fast image and video segmentation based on alpha-tree multiscale representation," in *International Conference on Signal Image Technology and Internet Based Systems*, 2012, pp. 336–342.

[14] 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.

[15] P. Soille, "Preventing Chaining through Transitions While Favouring It within Homogeneous Regions," in *International Symposium on Mathematical Morphology*. 2011, vol. 6671 of *LNCS*, pp. 96–107, Springer.

[16] N. Keshava, "Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 7, pp. 1552–1565, 2004.

[17] L. Yang and R. Jin, "Distance metric learning: A comprehensive survey," Tech. Rep., Michigan State Universiy, 2006.

[18] J. Serra and B.R. Kiran, "Optima on hierarchies of partitions," in *International Symposium on Mathematical Morphology*. 2013, vol. 7883 of *LNCS*, pp. 147–158, Springer.

[19] S. Derivaux, G. Forestier, C. Wemmert, and S. Lefèvre, "Supervised image segmentation using watershed transform, fuzzy classification and evolutionary computation," *Pattern Recognition Letters*, vol. 31, no. 15, pp. 2364–2374, 2010.