

Représentation d’images hyperspectrales sous forme d’arbres α

François MERCIOL, Laetitia CHAPEL, Sébastien LEFÈVRE

Université de Bretagne Sud — IRISA

Campus de Tohannic, BP 573, 56017 Vannes Cedex, France

{francois.merciol, laetitia.chapel, sebastien.lefevre}@univ-ubs.fr

Résumé – Les arbres α fournissent une représentation hiérarchique d’une image en partitions de régions d’hétérogénéité croissante. Ce modèle, inspiré du principe de saut minimum, a été revisité récemment en morphologie mathématique pour des images à niveaux de gris, et a été utilisé avec succès dans le domaine de la télédétection. Cet article montre comment cette représentation peut être adaptée à des données plus complexes, ici des images hyperspectrales, suivant différentes stratégies. Nous savons que la mesure de distance entre deux pixels voisins est un élément essentiel dans la qualité de l’arbre construit, mais les métriques usuelles ne sont pas satisfaisantes. Nous montrons ici qu’une solution intéressante pour appréhender des données hyperspectrales consiste en l’apprentissage préalable de la métrique à utiliser et la prise en compte de connaissances du domaine.

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.

1 Introduction

Avec l’avènement des images satellite à haute et très haute résolution spatiale, les méthodes usuelles d’interprétation d’image, opérant à l’échelle des pixels, ne suffisent plus. En effet, il est nécessaire d’avoir recours à des représentations de plus haut niveau des images, par exemple sous la forme de régions issues d’un processus de segmentation. Cette approche est connue en télédétection sous l’appellation *approche orientée objet* [2]. Face à la difficulté d’obtenir automatiquement une segmentation pertinente d’une image satellite, différentes représentations multi-échelles ont été proposées dans la littérature et nous nous intéressons ici au modèle des arbres α revisité récemment par Ouzounis et Soille [8], inspiré du principe de saut minimum bien connu en reconnaissance de formes [12]. La représentation d’une image panchromatique (à niveaux de gris) par son arbre α est utile pour identifier les bâtiments endommagés après un séisme [8] ou sélectionner des zones d’intérêt dans l’image (ou dans un espace de caractéristiques) [9].

Dans cet article, nous nous intéressons à l’extension des arbres α à la représentation de données plus complexes, ici des images hyperspectrales. Cette question a fait l’objet d’un travail récent [4] où les auteurs ont proposé d’utiliser un critère d’information mutuelle. Nous pensons qu’il est préférable de tirer parti d’une connaissance préalable du domaine, fournie par un expert sous forme d’un ensemble d’apprentissage. Nous montrons ainsi comment, dans un cadre supervisé, l’apprentissage de métrique permet de produire une mesure de dissimila-

rité plus pertinente et ainsi d’améliorer la qualité de la représentation hiérarchique qui en découle.

L’organisation de cet article est la suivante. Nous rappelons dans la section 2 le concept d’arbre α en introduisant les notations nécessaires. La section 3 traite spécifiquement de l’extension des arbres α au cas hyperspectral. L’intérêt de la méthode proposée est illustré par des résultats expérimentaux en section 4 avant de dresser des conclusions et perspectives à ce travail.

2 Arbres α et hiérarchies

Nous rappelons ici la définition d’un arbre α , qui est une représentation multi-échelle d’une image au travers de ses zones α , en utilisant les notations de [8]. Soit I une image définie sur un domaine de définition E . La segmentation d’une image est une partition \mathbf{P} de E , ou projection $x \rightarrow \mathbf{P}(x)$ de E dans $\mathcal{P}(E)$ telle que

$$\begin{aligned}\forall x \in E &\Rightarrow x \in \mathbf{P}(x) \\ \forall x, y \in E &\Rightarrow \mathbf{P}(x) = \mathbf{P}(y) \text{ ou } \mathbf{P}(x) \cap \mathbf{P}(y) = \emptyset\end{aligned}$$

avec $\mathbf{P}(x)$ représentant une cellule de \mathbf{P} contenant un point $x \in E$. On a donc $\bigcup_{x \in E} \mathbf{P}(x) = E$.

Nous notons également $\pi(x \rightsquigarrow y)$ un chemin de longueur N entre deux éléments $x, y \in E$, c’est-à-dire une chaîne d’éléments adjacents deux à deux $\langle x = x_0, x_1, \dots, x_{N-1} = y \rangle$. $\Pi \neq \emptyset$ est l’ensemble de tous les chemins reliant x à y . La

dissimilarité minimale entre x et y est définie par

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

où $d(x, y)$ est une mesure de dissimilarité entre les attributs des pixels x et y (par exemple leurs intensités ou niveaux de gris). Pour un pixel donné x , sa zone α notée $\alpha\text{-}\mathcal{Z}(x)$ est composée de tous les pixels reliés à x par un chemin dont les sauts intermédiaires ne sont pas supérieurs à α :

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

L'arbre α est alors une pyramide de partitions de E en zones α , ou projections $\Delta^A : E \rightarrow \Pi^A(E)$ définies par

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

avec $\Pi^A(E)$ l'ensemble de toutes les partitions α de E et $A = [0, 1, \dots, \alpha_{\max}]$ l'ensemble des valeurs de α . La relation \preceq correspond à la notion d'ordre selon $\alpha \in A$:

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

Un niveau de la pyramide $\Delta_{\alpha}^A \in \Delta^A$ est une partition \mathbf{P}^{α} de E , avec $\alpha \in A$. Soit $j \in J^{\alpha}$, dans lequel $J^{\alpha} \subseteq \mathbb{Z}$ est un ensemble d'indices, utilisé pour accéder aux zones α composant \mathbf{P}^{α} . La hiérarchie de partitions α notée Λ^A est une famille de projections ordonnées $\Lambda_{\alpha}^A : J^{\alpha} \rightarrow K^{\alpha}$ avec $K^{\alpha} \subseteq J^{\alpha}$, soit

$$\forall \alpha, \alpha' \in A, \alpha' < \alpha, \\ \Lambda^A = \{\Lambda_{\alpha=0}^A, \Lambda_{\alpha=1}^A, \dots, \Lambda_{\alpha=\alpha_{\max}}^A\} \mid \Lambda_{\alpha'}^A \prec \Lambda_{\alpha}^A$$

et

$$\forall \alpha \in A \setminus 0, \forall j \in J^{\alpha}, \\ \Lambda_{\alpha}^A = \{\alpha_j\text{-}\mathcal{Z} \mid (\alpha_j\text{-}\mathcal{Z} \in \Delta_{\alpha}^A) \wedge (\alpha_j\text{-}\mathcal{Z} \notin \Delta_{\alpha-1}^A)\}$$

Différents modèles hiérarchiques de représentation des images ont été proposés dans la littérature [13], par exemple les arbres de type *min-* ou *max-tree*, et les arbres de partition binaires (*binary partition tree*). Les premiers présentent l'inconvénient d'imposer une structure d'ordre sur les données initiales, ce qui n'est pas un problème trivial sur des données vectorielles telles que des images multi- ou hyperspectrales [1]. Les seconds, même s'ils offrent une représentation plus riche que les arbres α , nécessitent un temps de calcul bien plus important et supposent la définition de critères pour fusionner des noeuds [10, 14]. Les ondelettes et les pyramides gaussiennes font partie d'un autre ensemble de représentations multi-échelles qui ne fournissent pas intrinsèquement des partitions de l'image à différentes résolutions.

De par leur simplicité, les arbres α offrent en outre l'avantage de pouvoir être construits rapidement à l'aide d'algorithmes adaptés (à peine plus d'une minute pour une image satellite de plus de 20 millions de pixels) [5]. Une fois ces représentations construites, il est possible d'analyser l'image dans un

cadre interactif (temps-réel). Ainsi, la segmentation interactive de séquences vidéo couleur en s'appuyant sur cette représentation a été proposée dans [7]. L'extension de ces arbres à des données complexes reste cependant un problème ouvert.

3 Extensions au cas hyperspectral

La construction d'un arbre α nécessite la définition d'une mesure de dissimilarité entre les attributs de tout couple de pixels adjacents x et y . Cette dissimilarité peut simplement s'écrire sous la forme d'une différence de niveaux de gris dans le cas panchromatique (images à niveaux de gris), ou plus généralement avec toute norme L_q (avec $q = 1, 2$ ou encore ∞). Notons néanmoins qu'une telle mesure est très sensible au problème de réaction en chaîne observé lors de la discrétisation des contours dans une image : il est alors possible d'enrichir cette mesure en tenant compte des corrélations entre valeurs de pixel voisins [4].

Le recours à des métriques plus complexes s'avère nécessaire dès lors qu'on analyse des images multi- ou hyperspectrales, où les pixels sont plongés dans des espaces à haute voire très haute dimensionalité. Ainsi, il a été proposé dans [7] de construire l'arbre α avec la distance de Chebyshev (ou norme L_{∞}) pour traiter des séquences vidéo couleur. Dans le cas hyperspectral, il a cependant été montré dans [6] que la distance euclidienne pouvait être avantageusement remplacée par la distance angulaire ou *spectral angle mapper* (SAM)

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

avec $\|\cdot\|$ la norme euclidienne et \cdot le produit scalaire.

Les différentes distances évoquées précédemment sont calculées sans tenir compte de possibles a priori sur les données. Or dans le cas de l'analyse d'images satellite hyperspectrales, il est envisageable de s'appuyer sur des ensembles d'apprentissage, composés de pixels étiquetés par un expert (cela correspond à la problématique de classification supervisée, fréquente en télédétection). Nous pouvons alors faire appel à une distance guidée ou apprise, définie par exemple par la norme euclidienne utilisant une métrique de Mahalanobis apprise à partir des données [15], soit

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

L'apprentissage de la métrique $M = WW^T$ est ici guidé par les étiquettes des données, et plus particulièrement dans un cadre d'apprentissage semi-supervisé à partir de contraintes de paires *must-link*

$$\mathcal{S} = \{(\mathbf{a}, \mathbf{b}) \mid \mathbf{a} \text{ et } \mathbf{b} \text{ appartiennent à la même classe}\}$$

et de contraintes *cannot-link*

$$\mathcal{D} = \{(\mathbf{a}, \mathbf{b}) \mid \mathbf{a} \text{ et } \mathbf{b} \text{ appartiennent à des classes différentes}\}.$$

La matrice W est alors construite telle que la somme des distances au carré des points de \mathcal{S} soit minimale et celle des points

de \mathcal{D} soit maximale, ce qui revient à considérer la fonction objectif suivante :

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

où \hat{S}_S et \hat{S}_D sont les matrices de covariance des ensembles \mathcal{S} et \mathcal{D} . La métrique de Mahalanobis finale est alors construite comme $M = W^*(W^*)^T$.

4 Expérimentations

Nous évaluons ici l'apport de l'apprentissage de la métrique sous-jacente à la construction d'un arbre α lorsque celui-ci est construit à partir d'une image satellite hyperspectrale. Dans ce contexte, l'objectif est ensuite de procéder à une classification supervisée de l'image, où chaque région (ou zone α) est associée à une classe d'occupation du sol.

4.1 Procédure d'évaluation

Contrairement à [14], nous n'avons pas introduit ici d'heuristique pour sélectionner une segmentation ou partition optimale à partir de la représentation hiérarchique d'une image. La définition d'une telle heuristique est un problème ouvert et d'actualité [11]. Il n'est donc pas possible d'utiliser une mesure de qualité basée sur la précision d'une classification supervisée des régions issues de cette partition.

Nous avons choisi de pallier l'absence d'un critère optimal de segmentation en exploitant la vérité terrain dont nous disposons dans un contexte de classification supervisée, comme il a été proposé dans [3]. Ainsi, nous pouvons procéder à une évaluation exhaustive de l'ensemble des nœuds de l'arbre, en mesurant pour chacun d'eux un score local de qualité par la F-mesure (inspiré de l'indice de Jaccard modifié J' de [3]). Cette mesure est un compromis entre rappel et précision, et est définie par $F = 2 \times VP / (2 \times VP + FN + FP)$ où VP , FN , et FP représentent respectivement le nombre de pixels correctement classés, le nombre de faux négatifs, et le nombre de faux positifs. Ces quantités sont évaluées par comparaison avec les régions de référence fournies par la vérité terrain. Nous conservons ensuite les régions pertinentes les plus significatives (ou autrement dit, les nœuds les plus proches de la racine qui optimisent le F-score). La segmentation partielle qui en résulte est alors considérée comme une partition optimale théorique de l'image, dont les nœuds s'approchent le plus des régions de l'image de référence.

Pour évaluer la segmentation produite par la sélection des nœuds pertinents dans l'arbre, et puisque l'image de référence n'est que partiellement étiquetée (nous ne pouvons donc pas faire d'hypothèse sur les étiquettes réelles des pixels non étiquetés par l'expert), nous l'avons comparée à la vérité terrain à l'aide de la mesure de rappel, soit $VP / (VP + FN)$. Afin d'éviter le biais évidemment introduit par le choix de cette mesure, les pixels isolés ont au préalable été retirés de la carte de segmentation.

4.2 Données et résultats

Le jeu de données considéré est une sous-partie de l'image hyperspectrale *Pavia center*, composée de 102 bandes spectrales et acquise par le capteur ROSIS. L'image est de taille 492×492 pixels et contient 8 classes d'intérêt. La vérité terrain est partiellement connue : 27019 pixels sont étiquetés (figure 1 (a)). Les métriques évaluées sont Chebyshev, SAM et la métrique apprise (figures 1 (c) à (e)). Afin d'apprendre la métrique de distance liée aux données, 1% des pixels de chaque classe ont été intégrés dans un jeu de contraintes *must-link* et *cannot-link*.

Le tableau 1 présente une synthèse des résultats obtenus. Nous pouvons formuler les observations suivantes :

- l'apport de la distance angulaire SAM sur des métriques usuelles telles que celle de Chebyshev reste à démontrer lorsque ces distances sont utilisées pour construire des arbres α à partir d'images hyperspectrales ;
- l'apprentissage de métrique (ici une métrique de Mahalanobis), et plus généralement la prise en compte au plus tôt des connaissances de l'expert, est une approche prometteuse pour construire des arbres α adaptés à l'analyse et la classification d'images hyperspectrales.

TABLE 1 – Evaluation quantitative de la classification obtenue à partir d'une segmentation de l'extrait de *Pavia center* selon un arbre construit avec différentes métriques.

Méthode utilisée	Rappel (en %)
Chebyshev	84.71
SAM	82.46
Métrique apprise	86.47

5 Conclusion

La représentation hiérarchique d'une image par ses zones quasi-plates α offre différents avantages : description compacte, représentation rapide à calculer, applications en télédétection particulièrement prometteuses (pour des images panchromatiques ou à niveaux de gris). L'extension de ces représentations à des images plus complexes, telles que des images hyperspectrales, n'est cependant pas triviale.

Dans cet article, nous avons souligné les limites des métriques usuelles, telles que la distance de Chebyshev [7] ou la distance angulaire [6]. Nous avons également montré l'apport d'une métrique guidée par des connaissances a priori, dans un cadre d'apprentissage semi-supervisé. Ces premiers résultats, bien que prometteurs, nécessitent d'être validés sur un jeu de données plus complet et un plus large panel de métriques (notamment l'utilisation très récente d'un critère d'information mutuelle locale [4]), ainsi que des méthodes reconnues en hyperspectral (SVM, ACP). Une étude plus générale des différentes représentations hiérarchiques permettrait également

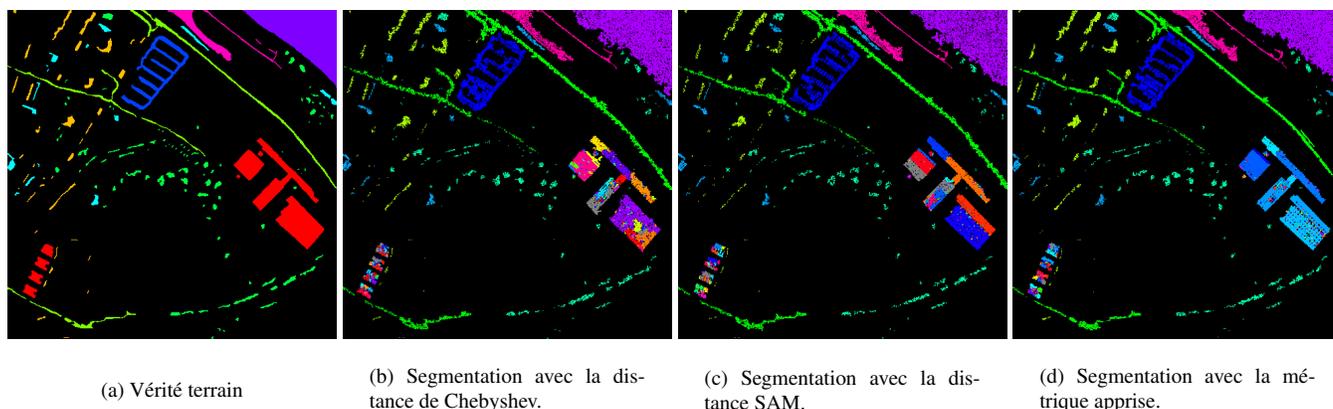


FIGURE 1 – Jeu de données *Pavia center* : vérité terrain et segmentations obtenues avec les différentes métriques.

d'évaluer l'intérêt des arbres α sur les autres modèles existants dans le contexte hyperspectral considéré ici.

Références

- [1] E. APTOULA et S. LEFÈVRE : A comparative study on multivariate mathematical morphology. *Pattern Recognition*, 40(11):2914–2929, 2007.
- [2] T BLASCHKE : Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1):2–10, 2010.
- [3] S. DERIVAUX, G. FORESTIER, C. WEMMERT et S. LEFÈVRE : Supervised image segmentation using watershed transform, fuzzy classification and evolutionary computation. *Pattern Recognition Letters*, 31(15):2364–2374, 2010.
- [4] Lionel GUEGUEN, Santiago VELASCO-FORERO et Pierre SOILLE : Local mutual information for dissimilarity-based image segmentation. *Journal of Mathematical Imaging and Vision*, pages 1–20, 2013.
- [5] J. HAVEL, F. MERCIOL et S. LEFÈVRE : Efficient schemes for computing α -tree representations. In *International Symposium on Mathematical Morphology*, volume 7883 de *Springer Lecture Notes in Computer Science*, pages 111–122, 2013.
- [6] 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*, 42(7):1552–1565, 2004.
- [7] F. MERCIOL et S. LEFÈVRE : Fast image and video segmentation based on alpha-tree multiscale representation. In *International Conference on Signal Image Technology and Internet Based Systems*, pages 336–342, 2012.
- [8] G. K. OUZOUNIS et P. SOILLE : Pattern spectra from partition pyramids and hierarchies. In *International Symposium on Mathematical Morphology*, volume 6671 de *Springer Lecture Notes in Computer Science*, pages 108–119, 2011.
- [9] G.K. OUZOUNIS, V. SYRRIS, L. GUEGUEN et P. SOILLE : The switchboard platform for interactive image information mining. In *Image Information Mining Conference : Knowledge Discovery from Earth Observation Data*, 2012.
- [10] P. SALEMBIER et L. GARRIDO : Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval. *IEEE Transactions on Image Processing*, 9(4):561–576, 2000.
- [11] J. SERRA et B.R. KIRAN : Optima on hierarchies of partitions. In *International Symposium on Mathematical Morphology*, volume 7883 de *Springer Lecture Notes in Computer Science*, pages 147–158, 2013.
- [12] R. SIBSON : SLINK : an optimally efficient algorithm for the single-link cluster method. *The Computer Journal*, 16(1):30–34, 1973.
- [13] P. SOILLE et L. NAJMAN : On morphological hierarchical representations for image processing and spatial data clustering. In *Applications of Discrete Geometry and Mathematical Morphology*, volume 7346 de *Springer Lecture Notes in Computer Science*, pages 43–67, 2012.
- [14] S. VALERO, P. SALEMBIER et J. CHANUSSOT : Hyperspectral image representation and processing with binary partition trees. *IEEE Transactions on Image Processing*, 22(4):1430–1443, 2013.
- [15] S. XIANG, F. NIE et C. ZHANG : Learning a mahalanobis distance metric for data clustering and classification. *Pattern Recognition*, 41(12):3600 – 3612, 2008.