

# Segmentation Multi-Image : Une Approche Collaborative pour la Construction d'Arbres Binaire de Partitions

## Multi-Image Segmentation: A Collaborative Approach For Binary Partition Tree Construction \*

Jimmy Francky Randrianasoa

Université de Reims Champagne-Ardenne, CReSTIC

UFR Sciences Exactes et Naturelles, Moulin de la Housse, BP 1039, 51687 Reims, FRANCE  
jimmy.randrianasoa@univ-reims.fr

### Résumé

La segmentation d'une image est généralement effectuée suivant le paradigme "une image, un algorithme". Cependant, il est parfois pertinent d'utiliser plusieurs images ou plusieurs algorithmes pour extraire plus d'informations. Pour pallier les problèmes induits par la fusion de segmentations, plusieurs travaux ont déjà été effectués afin de permettre un consensus entre plusieurs résultats de segmentation. Cet article propose une contribution portant sur la fusion de segmentations, se concentrant particulièrement sur la partie "n images" du paradigme. Contrairement à la majorité des approches proposées dans la littérature, notre but n'est pas de fusionner plusieurs résultats de segmentation mais de travailler en amont, au niveau de l'espace de recherche. Pour cela, nous proposons un schéma algorithmique permettant de créer un arbre binaire de partitions à partir de plusieurs images tout en passant par une étape de collaboration. Ce schéma algorithmique a été adapté de façon à pouvoir appliquer des politiques de consensus initialement proposées dans le domaine de l'apprentissage automatique. Ainsi, l'approche proposée permet d'obtenir un espace de recherche modélisant une hiérarchie de segmentations. D'un point de vue applicatif, des tests ont été effectués sur des images de télédétection, illustrant l'utilité potentielle de notre approche pour des images satellites.

### Mots Clef

Fusion de segmentations, hiérarchie morphologique, multi-image, stratégies collaboratives, arbre binaire de partitions, télédétection.

### Abstract

Image segmentation is generally performed in a "one image, one algorithm" paradigm. However, it is sometimes required to consider several images of a same scene, or to

carry out several (or several occurrences of a same) algorithm(s) to fully capture relevant information. To solve the induced segmentation fusion issues, various strategies have been already investigated for allowing a consensus between several segmentation outputs. This article proposes a contribution to segmentation fusion, with a specific focus on the "n images" part of the paradigm. Its main originality is to act on the segmentation research space, i.e., to work at an earlier stage than standard segmentation fusion approaches. To this end, an algorithmic framework is developed to build a binary partition tree in a collaborative fashion, from several images, thus allowing to obtain a unified hierarchical segmentation space. This framework is, in particular, designed to embed consensus policies inherited from the machine learning domain. Application examples proposed in remote sensing emphasise the potential usefulness of our approach for satellite image processing.

### Keywords

Segmentation fusion, morphological hierarchies, multi-image, collaborative strategies, binary partition tree, remote sensing.

## 1 Introduction

En analyse / traitement d'image, la segmentation est une tâche cruciale. Le concept de segmentation est, au demeurant, très générique, en termes de définition, de niveau sémantique, ou encore d'algorithmique.

Le paradigme "une image, un algorithme" est le principal invariant de la segmentation. En effet, pour une application donnée, un algorithme spécifique est généralement choisi suivant le problème considéré ; paramétré suivant les propriétés physiques / sémantiques des images cibles ; et enfin appliqué sur une image, ou sur chaque image d'un jeu de données.

Le problème de la segmentation étant intrinsèquement mal posé, il est clair que le résultat d'une segmentation obtenu à partir d'une image ne peut pas être complètement satis-

\* Cette recherche a été partiellement financée par l'Agence Nationale de la Recherche (Projet n° ANR-12-MONU-0001).

faisant. Il est donc parfois pertinent d'étendre le paradigme "une image, un algorithme", soit en appliquant un algorithme sur plusieurs images de la même scène – dans le but d'enrichir / améliorer les informations fournies en entrée – soit en appliquant plusieurs algorithmes, ou le même algorithme instancié avec divers ensembles de paramètres – dans le but d'enrichir / améliorer les informations obtenues en sortie.

Dans la littérature, les paradigmes " $n$  images, un algorithme" et "une image,  $n$  algorithmes" sont traités sous la terminologie commune de fusion de segmentations. La plupart du temps, la stratégie consiste à obtenir  $n$  résultats de segmentation, puis appliquer une stratégie de consensus pour rassembler et unifier ces  $n$  résultats en un seul, supposé plus précis. Cette stratégie de fusion a également été proposée dans le domaine de l'apprentissage automatique, spécialement la classification non supervisée (*clustering*) [18]. Intrinsèquement, le problème est le même, avec néanmoins pour différence que la classification ne prend généralement pas en compte l'organisation spatiale des données.

Dans cet article, nous proposons une nouvelle approche de fusion de segmentations, qui se concentre sur le paradigme " $n$  images, un algorithme". La pertinence et la nouveauté de cette approche, qui est basée sur les hiérarchies morphologiques et les opérateurs connexes [16], réside au niveau de l'étape où la fusion a lieu. Contrairement à la majorité des approches de fusion de segmentations, notre but n'est pas de fusionner plusieurs résultats de segmentation obtenus à partir de plusieurs images, mais d'obtenir directement un consensus lors de la définition de l'espace de recherche, c'est-à-dire, durant la construction de la hiérarchie modélisant les  $n$  images.

Cette stratégie, qui agit au niveau de l'espace de recherche, induit deux effets secondaires. Premièrement, elle nous permet de nous baser sur des approches hiérarchiques de segmentation, et ainsi proposer un schéma riche et polyvalent, qui peut être facilement instancié suivant le domaine d'application. Deuxièmement, nous pouvons directement bénéficier des stratégies de fusion de données "non-spatiales", précédemment proposées par la communauté de l'apprentissage automatique, en procédant à la fusion au niveau interne des structures de données impliquées dans l'algorithme de construction de la hiérarchie, au lieu des régions spatiales de divers résultats de segmentation.

Notre approche repose sur les Arbres Binaires de Partitions (BPTs) [14] comme modèle hiérarchique. Ce choix est motivé d'une part par la possibilité d'adapter la construction de BPT, contrairement aux autres structures d'arbres qui sont étroitement déterminées à partir des images (à savoir, les arbres de coupes et les arbres de formes). D'autre part, les BPTs sont souvent appliqués en télédétection, où l'utilisation de plusieurs images d'une même scène est fréquente, comme l'usage des stratégies de fusion pour les problèmes de classification non supervisée.

Cet article est organisé comme suit. Dans la Section 2, une brève vue d'ensemble est proposée, décrivant les précé-

dentes stratégies de fusion de segmentations et l'application des BPTs et des hiérarchies pour l'analyse multiple d'images de télédétection. Dans la Section 3, le schéma algorithmique et structurel pour la construction de BPT à partir de plusieurs images est décrite, et différentes stratégies de consensus pour l'instanciation du schéma sont discutées. Dans la Section 4, des expérimentations sont effectuées sur des images de télédétection, afin d'illustrer la pertinence et l'utilité potentielle de la segmentation à partir d'un BPT représentant plusieurs images, en considérant deux cas d'application en télédétection. Dans la Section 5, une discussion conclut cet article en soulignant les perspectives et les futurs travaux potentiels.

## 2 Travaux Connexes

### 2.1 Fusion de Segmentations

La fusion de segmentations consiste à établir un consensus entre plusieurs résultats de segmentation. La taxonomie de la fusion de segmentations est directement calquée sur celle de la segmentation.

D'une part, la segmentation peut être vue comme un processus ayant pour but d'extraire une structure (objet) du reste de l'image (le fond). Des exemples typiques de telles stratégies de segmentation sont les modèles déformables, les coupes de graphe, *etc.* Dans ce contexte, la fusion de segmentations peut être interprétée comme un problème d'interpolation géométrique (voir, par exemple, [13, 19]). Cette sous-famille de méthodes de fusion de segmentations, principalement utilisée en imagerie médicale, sort du cadre de notre étude.

D'autre part, la segmentation peut être vue comme un processus qui a pour but de définir une partition de la totalité de l'image, afin d'extraire des zones homogènes significatives. Des exemples de telles stratégies de segmentation sont les *watersheds*, le *split-and-merge* et, plus généralement, les opérateurs connexes. Les méthodes de classification non supervisée entrent aussi dans cette catégorie, avec la différence qu'elles ne prennent pas en compte l'organisation spatiale des données images.

Du point de vue de l'apprentissage automatique, certaines méthodes procèdent à une fusion de segmentations à travers des ensembles de *clustering*; une étude comparative peut être trouvée dans [8]. Il a aussi été proposé d'interpréter les informations obtenues en chaque pixel dans différents résultats de segmentation comme des vecteurs caractéristiques qui sont ensuite impliqués dans des procédures d'optimisation. En particulier, une approche basée sur deux étapes successives de K-MEANS a été considérée dans [11], ou un schéma bayésien dans [21].

D'autres approches prennent en compte explicitement l'organisation spatiale des résultats de segmentation. Dans les travaux pionniers [7], des images d'une même scène obtenues à partir de diverses modalités ont été fusionnées et optimisées suivant les informations relatives aux bords des objets. Plus tard, l'analyse de la co-occurrence des pixels voisins a permis d'améliorer la précision d'une partition is-

sue de plusieurs versions avec de légères perturbations des frontières [6].

Plus récemment, les opérateurs connexes ont aussi été impliqués. Les *watersheds* stochastiques [3] ont été introduits comme une solution pour améliorer la robustesse des *watersheds* basés sur des marqueurs, par la fusion des résultats obtenus à partir de différentes initialisations. Cette approche a notamment inspiré une autre méthode de *stochastic minimum spanning forest* [5] pour la segmentation / classification d’images de télédétection hyper-spectrales. Les *random walkers* [22] ont enfin été considérés pour la segmentation d’un graphe généré à partir du degré d’homogénéité entre différents résultats de segmentation.

## 2.2 Hiérarchies Morphologiques, Multi-Images et Télédétection

Les hiérarchies morphologiques associées aux opérateurs connexes [16] ont été impliquées avec succès en segmentation d’image suivant le paradigme “une image, un algorithme”. Dans les structures d’arbres associées, les nœuds modélisent des régions homogènes dans l’image tandis que les arêtes représentent leurs relations d’inclusion.

Les arbres classiques, comme les arbres de coupes [15] ou les arbres de formes [12], nous permettent d’effectuer une segmentation hiérarchique en fusionnant des zones plates. Ces structures produisent en sortie des partitions partielles d’une image avec des niveaux de détails modifiables. Cependant, elles s’appuient fortement sur l’intensité spectrale de l’image, ce qui n’est pas compatible avec les spécificités des images satellites.

Une première solution pour faire face à ce problème repose sur la connectivité contrainte [17]. La relation de connectivité génère un partitionnement des zones de l’image, de la plus fine vers la plus grossière ; les hiérarchies de partitions sont alors formées en variant un seuil qui est associé à chaque contrainte de connectivité. D’une manière différente, le BPT [14] reflète une mesure de similarité (choisie) entre les régions voisines, et modélise une hiérarchie entre ces régions. Les BPTs ont été utilisés pour segmenter différents types d’images satellites [20, 4]. Cependant, les objets contenus dans les images satellites sont souvent trop hétérogènes pour être correctement segmentés à partir d’une seule image. Il semble alors pertinent de considérer  $n$  images d’une même scène pour enrichir l’espace de données et améliorer la capacité de la segmentation hiérarchique.

Dans ce contexte, des efforts ont été menés pour étendre les hiérarchies morphologiques à traiter  $n$  images. Dans [10], une extension du modèle de BPT a été proposée pour traiter des images satellites multi-résolution en considérant une hiérarchie par résolution d’image. Dans la même veine, une approche basée sur des hiérarchies morphologiques multiples a été développée dans [1] pour segmenter des images multispectrales.

Ces travaux récents montrent l’intérêt de considérer un paradigme multi-image pour améliorer la segmentation hié-

rarchique des structures complexes que l’on trouve dans les images de télédétection. Les méthodes de fusion de segmentations multi-image tirent partie de la complémentarité des données disponibles, et peuvent être adaptées au cas des hiérarchies morphologiques par l’interprétation de “ $n$  images, un algorithme” comme “ $n$  images, une hiérarchie”. En établissant des collaborations où les fusions s’accordent relativement aux différents contenus des images, il est alors possible de construire graduellement une hiérarchie consensuelle unique, qui peut être utilisée pour détecter des objets complexes tout en évitant le bruit (spectral, sémantique, ...) présents dans une seule image.

Nous proposons, dans la section suivante, une méthode de segmentation hiérarchique collaborative, étendant les BPTs pour le multi-image. Contrairement aux approches classiques de fusion de segmentations, nous ne comptons pas fusionner plusieurs résultats de segmentation, mais directement développer un consensus durant la construction de la hiérarchie qui modélise les  $n$  images.

## 3 Construction d’un Arbre Binaire de Partitions Multi-Images

Nous rappelons d’abord l’algorithme de construction de BPT. Notre point de vue se concentre particulièrement sur les structures de données, qui sont la pierre angulaire de notre contribution. Ensuite, nous décrivons notre généralisation de cet algorithme pour traiter plusieurs images. Plusieurs familles de stratégies de consensus sont finalement discutées pour l’instanciation de ce schéma.

### 3.1 Construction Classique d’un BPT

**Résumé de l’algorithme [14].** Un BPT est une représentation hiérarchique d’une image. Plus précisément, c’est un arbre binaire, dont chaque nœud forme une région connexe. Chacun de ces nœuds est soit une feuille – donc une région “élémentaire” – soit un nœud interne, modélisant l’union des régions de ses deux fils. La racine est le nœud correspondant au support de l’image. En pratique, un BPT est construit à partir de ses feuilles – provenant d’une partition initiale du support de l’image – vers sa racine, d’une manière ascendante, en choisissant et fusionnant itérativement deux régions adjacentes qui minimisent un critère de fusion reflétant, par exemple, les similarités spectrales et / ou géométriques des régions.

**Description structurelle de l’algorithme.** Une image est une fonction  $I : \Omega \rightarrow V$  qui associe pour chaque point  $x$  de l’ensemble fini  $\Omega$  une valeur  $I(x)$  de l’ensemble  $V$ . Un point crucial pour la construction de BPT – et plus généralement pour effectuer un filtrage connexe – est la prise en compte de la structure de  $\Omega$ . Plus précisément, il est primordial de modéliser le fait que deux points  $x$  et  $y$  de  $\Omega$  soient voisins. Cela est effectué grâce à la définition d’une adjacence  $A_\Omega$ , c’est-à-dire, une relation binaire irréflexive, symétrique sur  $\Omega$ . En d’autres termes,  $(\Omega, A_\Omega)$  est un graphe qui modélise la structure de l’image.

Considérons maintenant une partition initiale  $\mathcal{L}$  de  $\Omega$ . (Chaque nœud  $L \subseteq \Omega$  de  $\mathcal{L}$  est généralement supposé connexe par rapport à  $A_\Omega$ .) Cette partition  $\mathcal{L}$  définit l'ensemble de feuilles du BPT que nous allons construire.

Pour une partition  $\mathcal{P}$  de  $\Omega$  (et en particulier pour  $\mathcal{L}$ ) nous pouvons définir une adjacence héritée de celle de  $\Omega$ . Plus précisément, nous pouvons dire que deux nœuds distincts  $N_1, N_2 \in \mathcal{P}$  sont adjacents s'il existe  $x_1 \in N_1$  et  $x_2 \in N_2$  tels que  $(x_1, x_2)$  est une arête de  $A_\Omega$ , c'est-à-dire,  $x_1$  et  $x_2$  sont adjacents dans  $(\Omega, A_\Omega)$ . Cette nouvelle relation d'adjacence  $A_{\mathcal{P}}$  est aussi irréflexive et symétrique. Dans le cas de  $\mathcal{L}$ , cela nous permet de définir un graphe  $\mathfrak{G}_{\mathcal{L}} = (\mathcal{L}, A_{\mathcal{L}})$  qui modélise la structure de la partition de l'image  $I$ .

Le BPT est une structure de données qui décrit la simplification progressive de  $\mathfrak{G}_{\mathcal{L}}$  en un graphe trivial  $(\Omega, \emptyset)$ . Cette procédure consiste à définir une séquence  $(\mathfrak{G}_i = (N_i, A_{N_i}))_{i=0}^n$  (avec  $n = |\mathcal{L}| - 1$ ) comme suit. Premièrement, posons  $\mathfrak{G}_0 = \mathfrak{G}_{\mathcal{L}}$ . Ensuite, pour chaque  $i$  de 1 à  $n$ , nous choisissons les deux nœuds  $N_{i-1}$  et  $N'_{i-1}$  de  $\mathfrak{G}_{i-1}$  liés par une arête  $(N_{i-1}, N'_{i-1}) \in A_{N_{i-1}}$  qui minimisent un critère de fusion, et nous définissons  $\mathfrak{G}_i$  tel que  $N_i = N_{i-1} \setminus \{N_{i-1}, N'_{i-1}\} \cup \{N_{i-1} \cup N'_{i-1}\}$ ; en d'autres termes, nous remplaçons ces deux nœuds par leur union. L'adjacence  $A_{N_i}$  est définie suivant  $A_{N_{i-1}}$  : nous supprimons l'arête  $(N_{i-1}, N'_{i-1})$ , et nous remplaçons chaque arête  $(N_{i-1}, N''_{i-1})$  et / ou  $(N'_{i-1}, N''_{i-1})$  par une arête  $(N_{i-1} \cup N'_{i-1}, N''_{i-1})$  (en particulier, deux anciennes arêtes pourraient être fusionnées en une seule).

D'un point de vue structurel, un BPT  $\mathfrak{T}$  est le diagramme de Hasse d'un ensemble  $(\bigcup_{i=0}^n N_i, \subseteq)$  qui est partiellement ordonné. D'un point de vue algorithmique,  $\mathfrak{T}$  est construit parallèlement à la simplification progressive de  $\mathfrak{G}_0$  vers  $\mathfrak{G}_n$ ; en d'autres termes,  $\mathfrak{T}$  stocke l'historique des fusions de nœuds. Plus précisément, nous définissons une séquence  $(\mathfrak{T}_i)_{i=0}^n$  comme suit. Posons  $\mathfrak{T}_0 = (N_0, \emptyset) = (\mathcal{L}, \emptyset)$ . Ensuite, pour chaque  $i$  de 1 à  $n$ , nous construisons  $\mathfrak{T}_i$  à partir de  $\mathfrak{T}_{i-1}$  par l'ajout d'un nouveau nœud  $N_{i-1} \cup N'_{i-1}$ , et de deux arêtes  $(N_{i-1} \cup N'_{i-1}, N_{i-1})$  et  $(N_{i-1} \cup N'_{i-1}, N'_{i-1})$ . Le BPT  $\mathfrak{T}$  est finalement défini par  $\mathfrak{T}_n$ .

**Remarque.** La description classique de l'algorithme de construction de BPT considère en entrée : l'image  $I$ ; un modèle de région, qui permet de "décrire" les nœuds; et un critère de fusion, qui permet de quantifier l'homogénéité des nœuds avant et après une fusion éventuelle. Ces informations sont importantes d'un point de vue applicatif. Cependant, d'un point de vue algorithmique, leur seule utilité est de définir une valuation des arêtes qui permet de choisir quels nœuds sont à fusionner à une étape donnée. Par la suite, nous pourrions ainsi considérer qu'un BPT est totalement défini seulement par deux informations en entrée : (1) le graphe  $\mathfrak{G}_{\mathcal{L}} = (\mathcal{L}, A_{\mathcal{L}})$  qui modélise la partition initiale de l'image; (2) une fonction de valuation  $W : (2^\Omega)^2 \times V^\Omega \rightarrow \mathbb{R}$  qui permet de choisir, pour chaque étape du processus, la prochaine paire de nœuds à fusionner.

**Structures de données.** La description précédente de l'algorithme de construction de BPT implique la définition

et la mise à jour, durant la totalité du processus, de plusieurs structures de données, à savoir : le graphe  $\mathfrak{G}$ , qui nous permet de connaître les nœuds restant à fusionner et quels sont leurs liens d'adjacence; et l'arbre  $\mathfrak{T}$  qui est construit progressivement. Afin de calculer efficacement la valuation  $W$ , il est aussi important d'associer chaque nœud de  $\mathfrak{G}$  à la partie correspondante de l'image  $I$ , c'est-à-dire, à travers une mise en correspondance entre  $\mathfrak{G}$  (en fait,  $\mathcal{N}$ ) et  $\Omega$ .

La dernière structure de données requise est une liste triée  $\mathcal{W}$  qui rassemble les valuations de chaque arête restante de  $\mathfrak{G}$ . Cette liste contient les informations qui vont permettre, pour chacune des  $n$  itérations du processus, le choix du couple de nœuds à fusionner. Comme  $\mathcal{W}$  est triée, ce choix est fait en temps constant  $O(1)$ . Après l'opération de fusion,  $\mathcal{W}$  doit être mise à jour : (1) en supprimant l'arête entre les deux nœuds; (2) en mettant à jour les arêtes affectées par l'opération de fusion; et (3) en réordonnant ces arêtes mises à jour. L'opération (1) est effectuée en temps constant  $O(1)$ . L'opération (2) est effectuée en  $O(\alpha \cdot T_W)$ , où  $T_W$  est le coût de calcul de  $W$  pour une arête, et  $\alpha$  est le nombre de voisins des nœuds fusionnés ( $\alpha$  est généralement majoré par une constante de faible valeur). L'opération (3) est effectuée en  $O(\alpha \cdot \log_2 |\mathcal{W}|)$ .

### 3.2 Généralisation de la Construction de BPT pour Plusieurs Images

Considérons maintenant  $k > 1$  images  $I_j$ , au lieu d'une seule. En partant de l'hypothèse qu'elles correspondent à une même scène, nous considérons – quitte à ré-échantillonner – que toutes sont définies sur un même support. Nous avons alors un ensemble d'images  $\{I_j : \Omega \rightarrow V_j\}_{j=1}^k$ . Le but est maintenant de construire un BPT à partir de ces  $k$  images, en généralisant le schéma algorithmique décrit dans la Section 3.1. Une étape du processus de cet algorithme est illustrée en Figure 1.

**Évolutions structurelles.** Comme indiqué ci-dessus, nous avons premièrement besoin d'un graphe qui modélise la partition initiale  $\mathcal{L}$  des images. Comme toutes les images  $I_j$  partagent le même support  $\Omega$ , un tel graphe peut être obtenu facilement, soit par la subdivision de  $\Omega$  en un ensemble de singletons – le graphe  $\mathfrak{G}_{\mathcal{L}}$  induit est alors isomorphe à  $(\Omega, A_\Omega)$  – soit en considérant des zones plates, du produit cartésien de toutes les images.

Les  $k$  images  $I_j$  partagent le même support, mais elles prennent leurs valeurs dans différents ensembles  $V_j$ . En conséquence, suivant le paradigme classique de création de BPT, chacune d'elles est associée à une fonction de valuation spécifique  $W_j : (2^\Omega)^2 \times V_j^\Omega \rightarrow \mathbb{R}$  qui est définie en particulier par rapport aux valeurs de l'ensemble  $V_j$ .

D'un point de vue structurel de données, la généralisation de l'algorithme de construction de BPT dépendra toujours d'une seule graphe  $\mathfrak{G}$ , qui sera progressivement simplifié; et d'un arbre  $\mathfrak{T}$  qui sera progressivement construit pour donner le BPT. Une application unique entre  $\mathcal{N}$  et  $\Omega$  permettra toujours d'avoir accès aux valeurs d'un nœud pour

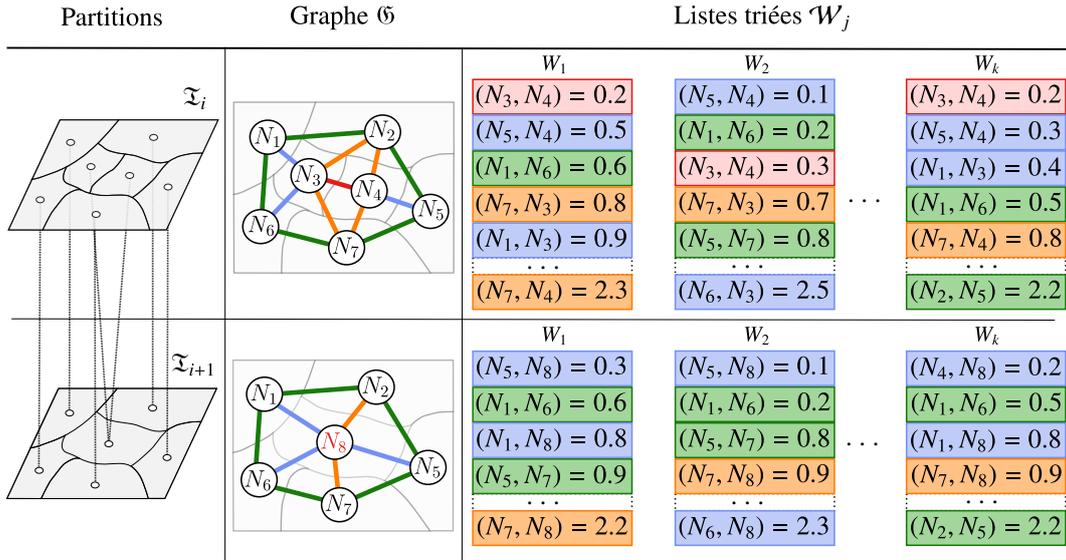


FIGURE 1 – Une étape de la construction de BPT à partir de  $k$  images. À gauche : les partitions de  $\Omega$  avant et après la fusion de deux nœuds. Au centre : les graphes  $\mathfrak{G}$  associés, avant et après la fusion de  $N_3$  et  $N_4$ , formant le nouveau nœud  $N_8$ . L’arête rouge est supprimée. Les arêtes bleues et oranges sont mises à jour, par exemple,  $(N_1, N_3)$  devient  $(N_1, N_8)$ ; celles qui sont oranges sont fusionnées par paires, par exemple,  $(N_7, N_3)$  et  $(N_7, N_4)$  deviennent  $(N_7, N_8)$ . Les arêtes vertes ne sont pas affectées. À droite : les  $k$  listes  $W_j$ . Les cellules rouges sont supprimées, comme l’arête  $(N_3, N_4)$ ; cette arête a été choisie suivant une politique de consensus donnée, grâce à sa “haute” position dans les  $k$  listes. Les scores des cellules bleues et oranges sont mis à jour suivant  $N_8$ ; les cellules oranges sont fusionnées par paires. Les positions des cellules bleues et oranges sont mises à jour suivant leurs nouveaux scores. Les scores des cellules vertes ne sont pas affectés.

les  $k$  images. La principale différence repose maintenant sur le fait que chaque fonction  $W_j$  induit – en première approximation – une liste spécifique triée  $\mathcal{W}_j$  pour recueillir les évaluations des arêtes restantes de  $\mathfrak{G}$ , calculées par  $W_j$ .

**Conséquences algorithmiques.** D’un point de vue algorithmique, chaque itération du processus de construction conserve la même structure. Une arête est choisie et les deux nœuds correspondants du graphe sont fusionnés. Cette opération implique de mettre à jour des nœuds et des arêtes de  $\mathfrak{G}$ , et ajoute un nouveau nœud en plus de deux arêtes dans  $\mathfrak{T}$ . Les principales différences sont : (i) que plusieurs listes triées sont à mettre à jour au lieu d’une seule; et (ii) que le choix de l’arête optimale doit être fait en considérant les informations portées par ces  $k$  listes triées au lieu d’une seule, pour un BPT classique.

D’un point de vue informatique, le choix d’une arête à supprimer n’est plus une opération à temps constant, mais dépendra de la manière dont les informations sont utilisées et comparées. Ensuite, les opérations (1–3) décrites dans l’algorithme de construction classique de BPT, pour la maintenance des listes triées, doivent être dupliquées pour chaque liste. Ces opérations sont alors exécutées respectivement en  $O(k)$ ,  $O(k \cdot \alpha \cdot T_{W_*})$  et  $O(k \cdot \alpha \cdot \log_2 |\mathcal{W}_*|)$ .

Cependant, cette généralisation initiale de l’algorithme de construction de BPT peut être raffinée en étudiant plus précisément les politiques de consensus qui sont à considérer pour le choix d’une arête, notamment par rapport aux va-

luations  $W_j$  et / ou aux listes triées  $\mathcal{W}_j$ .

### 3.3 Stratégies de Consensus

À chaque itération, le choix de l’arête optimale à supprimer, menant à la fusion des deux nœuds concernés, dépend d’un consensus entre les informations des  $k$  images. Pour chaque image  $I_j$ , les informations utiles sont détenues, d’une part, par la fonction de valuation  $W_j : (2^\Omega)^2 \times V^\Omega \rightarrow \mathbb{R}$  qui fournit une valeur *absolue* pour chaque arête et, de l’autre, par la liste triée  $\mathcal{W}_j$ , qui fournit des informations *relatives* sur les arêtes, notamment par rapport à leur ordre dans  $W_j$ . Ces informations sont de natures distinctes; nous étudions leur pertinence suivant les différentes politiques de consensus considérées.

**Consensus des informations absolues.** Considérons que la politique de consensus consiste à choisir une arête ayant la plus petite moyenne de valuations parmi les  $k$  images, ou l’arête ayant la valuation minimale parmi toutes les images. Le premier type de consensus (nommé *min of mean*) est défini par une formulation linéaire :  $\arg_{(N, N') \in \mathcal{N}} \min \sum_{j=1}^k W_j((N, N'))$ , tandis que le second (nommé *min of min*) est défini par une formulation non-linéaire :  $\arg_{(N, N') \in \mathcal{N}} \min \min_{j=1}^k W_j((N, N'))$ . Cependant, dans les deux cas, la décision est prise par la considération des informations absolues portées par les arêtes. En d’autres termes, la connaissance des  $k$  valeurs de chaque point de  $\Omega$  est suffisante. En conséquence, les  $k$  listes triées

TABLE 1 – Coût en temps et en espace de la construction de BPT pour différentes familles de politiques de consensus. Les symboles  $\Theta$  and  $\Omega$  réfèrent ici aux notations de Landau. (Supp. : suppression ; Maj. : mise à jour).

Politiques de consensus	# $\mathcal{W}_*$	Choix d’arête	Supp. d’arête	Maj. d’arêtes	Tri des listes
Informations absolues	1	$\Theta(1)$	$\Theta(1)$	$\Theta(1)$	$\Theta(\log_2  \mathcal{W}_* )$
Inf. relatives locales	$k$	$\Omega(k)$	$\Theta(k)$	$\Theta(k)$	$\Theta(k \cdot \log_2  \mathcal{W}_* )$
Inf. relatives globales	$k$	$\Omega(k \cdot  \mathcal{W}_* )$	$\Theta(k)$	$\Theta(k)$	$\Theta(k \cdot \log_2  \mathcal{W}_* )$

$\mathcal{W}_j$  sont inutiles, et une simple liste triée  $\mathcal{W}$  contenant les informations de ces formulations – linéaires ou non-linéaires – est alors suffisante. La construction de BPT à partir de  $k$  images est alors équivalente à celle d’une seule image définie comme  $I : \Omega \rightarrow \prod_{j=1}^k V_j$ .

**Consensus des informations relatives locales.** Considérons maintenant que la politique de consensus consiste à choisir l’arête qui est le plus souvent en première position dans les  $k$  listes triées  $\mathcal{W}_j$ , ou la plus fréquemment présente parmi les  $r \ll |\mathcal{W}_*|$  premières positions dans les  $k$  listes triées  $\mathcal{W}_j$ . Ces politiques de consensus (nommées, *majority vote* et *most frequent*, potentiellement pondérées) n’interviennent pas au niveau des valuations absolues des arêtes, mais au niveau de leurs positions relatives dans les listes. Dans un tel cas, il est obligatoire de maintenir les  $k$  listes triées. Cependant, le processus de prise de décision ne requiert pas l’accès à l’ensemble de listes dans leur totalité, mais peut être réduit aux premiers (ou les  $r$  premiers) élément(s) de chacune, menant à un processus de prise de décision *locale*.

**Consensus des informations relatives globales.** Considérons finalement que la politique de consensus consiste à choisir l’arête qui a le meilleur rang global parmi les  $k$  listes triées  $\mathcal{W}_j$ . Une telle politique de consensus (par exemple, *best average*, ou *best median ranking*), agit sur les positions relatives des arêtes dans les listes. Contrairement au cas précédent, le processus de prise de décision requiert un accès à la totalité des contenus de toutes les listes, menant à un processus de prise de décision *globale* nécessitant un temps de calcul élevé. (Un tel coût peut être réduit en maintenant, dans les cas favorables, une  $(k + 1)$ -ème liste qui résume les informations globales, et / ou en adoptant des stratégies heuristiques qui mettent à jour les listes seulement après un nombre d’étapes donné).

**Conséquences algorithmiques et structurelles.** Le choix de stratégie de consensus est fortement dépendant de l’application. Ainsi, il est important de considérer un compromis entre le coût de l’approche et le bénéfice induit en terme de précision des résultats. Les coûts des différentes stratégies sont en particulier résumés dans la Table 1.

## 4 Expérimentations

Pour expérimenter notre schéma, deux applications ont été considérées dans le cadre de la télédétection : une application mono-date, mono-source, plusieurs images (bruitées),

pour évaluer la capacité à retrouver des informations malgré une dégradation de l’image ; et une application multi-date, multi-source, une image par date, pour évaluer la capacité à capter des informations redondantes qui sont indépendantes du temps.

Des stratégies simples de construction et de segmentation de BPT ont été volontairement choisies, afin d’éviter les biais que pourraient induire ces choix ; ainsi nous pouvons nous concentrer sur les effets de la structure du BPT multi-image actuel par comparaison avec un BPT standard.

### 4.1 Images Urbaines Bruitées

**Données.** Le jeu de données utilisé ici a été acquis pour la ville de Strasbourg. L’échantillon original (Figure 2(a)) est une image urbaine ( $1024 \times 1024$  pixels) acquise par les satellites PLÉIADES en 2012 (images fournies par le laboratoire LIVE, UMR CNRS 7263). C’est une image *pan-sharpening* ayant une résolution spatiale de 60 cm avec quatre bandes (NIR, R, G, B). À partir de cette image, une série de 7 images bruitées a été générée avec des bruits gaussien et impulsionnel (Figure 2(b)).

**Méthode et résultats.** Les BPTs ont été construits à partir d’une partition triviale  $\mathcal{L}$  composée de tous les ensembles de singletons, c’est-à-dire, un pixel par région. La fonction de valuation  $W_* : (2^\Omega)^2 \times V^\Omega \rightarrow \mathbb{R}$  est définie comme l’augmentation des intervalles des valeurs de l’intensité (pour chaque bande spectrale), potentiellement induite par la fusion des régions incidentes. Dans le cas du multi-image, la politique de consensus des informations relatives locales *most-frequent*, pondérée suivant les positions des arêtes dans les listes, est appliquée pour les premiers 10% de la liste  $\mathcal{W}_*$ .

Le BPT “standard” de l’image non bruitée de la Figure 2(a) a d’abord été segmenté en considérant une coupe horizontale prédéfinie, voir la Figure 2(c). Le BPT multi-image est alors segmenté de la même façon, menant à un même nombre de régions, voir la Figure 2(d). Dans ces figures, nous observons que les résultats obtenus à partir des images bruitées sont légèrement dégradés, mais de qualité comparable, par rapport à la Figure 2(c). Cela tend à confirmer la capacité de la segmentation basée sur les BPTs multi-image à générer des résultats précis, en discriminant des informations pertinentes par rapport au bruit grâce au consensus obtenu à partir des différentes images, même dans le cas d’un bas rapport signal sur bruit.

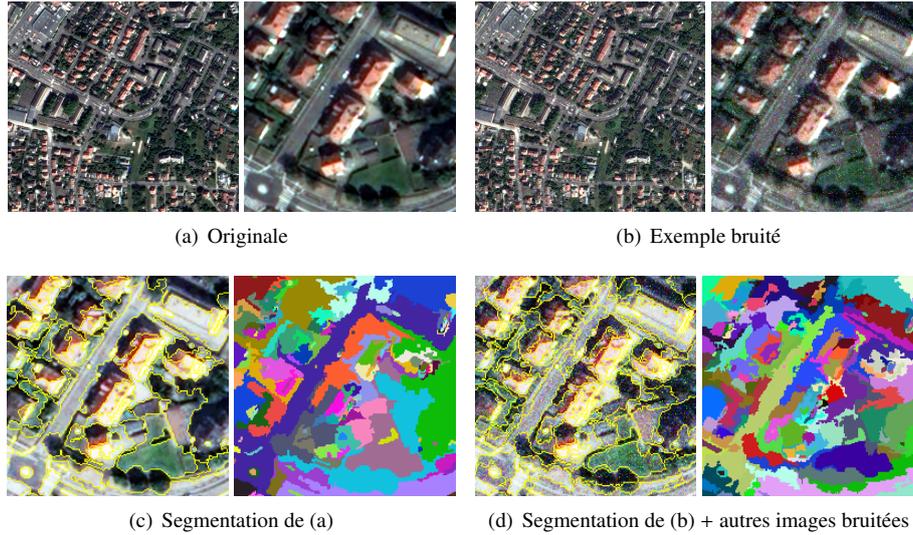


FIGURE 2 – (a) Image initiale, PLÉIADES, 2012 ; et un échantillon zoomé,  $200 \times 200$  pixels. (b) Un exemple d’image bruitée à partir de (a), bruit gaussien ( $\sigma = 10\%$ ) et bruit impulsif (5%), et un échantillon zoomé,  $200 \times 200$  pixels. (c) Segmentation à partir d’un BPT mono-image de (a). (d) Segmentation à partir d’un BPT multi-image de (b) avec 6 autres images bruitées.

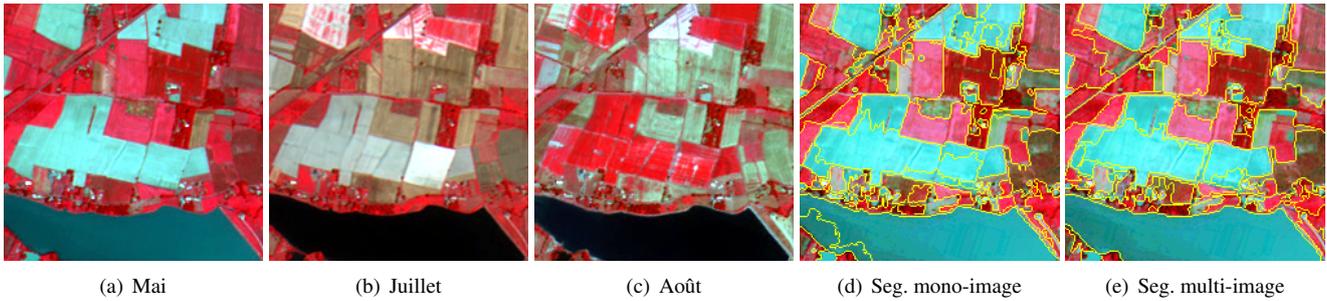


FIGURE 3 – (a–c) 3 échantillons ( $200 \times 200$  pixels) de la série temporelle d’images, FORMOSAT-2, 2007. (d) Segmentation obtenue à partir du BPT mono-image de (a). (e) Segmentation obtenue à partir du BPT multi-image de (a–c).

## 4.2 Images Agricoles Multi-Temporelles

**Données.** L’ensemble de données utilisé ici est une série temporelle d’images agricoles ( $1000 \times 1000$  pixels) correspondant à une zone localisée près de Toulouse. Les images ont été acquises par le satellite FORMOSAT-2 au cours de l’année agricole de 2007, voir la Figure 3 (a–c). Elles ont été orthorectifiées et ont une résolution spatiale de 8 m, avec quatre bandes (NIR, R, G, B).

**Méthode et résultats.** Les BPTs ont été construits à partir d’une même partition  $\mathcal{L}$  et une fonction de valuation  $W_\star : (2^\Omega)^2 \times V^\Omega \rightarrow \mathbb{R}$  telles que dans la Section 4.1. Ils sont aussi segmentés de la même manière.

Les résultats de segmentation, représentés dans la Figure 3(e) fournissent des régions qui ne sont pas les mêmes que pour un BPT mono-image standard, construit à partir d’une image de la série (Figure 3(d)). D’une part, quelques effets produits par des bruits sémantiques de la segmentation mono-image sont parfois corrigés par les informations

redondantes obtenues à partir du multi-image. D’autre part, le BPT multi-image se concentre sur des détails qui sont seulement accessibles à travers une analyse temporelle, fournissant ainsi un outil potentiellement utile pour ce type de données.

## 5 Conclusion

Cet article a présenté un schéma (algorithmique et structuré) et différentes politiques de fusion basées sur des stratégies de consensus pour la construction d’un BPT unique à partir de plusieurs images. Cette contribution est, à notre connaissance, la première tentative de gérer une fusion de segmentations hiérarchique (morphologique), suivant le paradigme “ $n$  images, un algorithme”.

Les expérimentations réalisées sur des ensembles de multi-images satellites ont montré que la qualité des hiérarchies morphologiques induites est suffisante pour la segmentation, par exemple, à partir d’images bruitées ou multi-temporelles d’une même scène. Les stratégies de consen-

sus considérées dans cet étude restent cependant très basique (*most-frequent, majority vote, etc.*). L'intégration d'un consensus de plus haut niveau pourrait alors permettre l'amélioration de la qualité des hiérarchies et la segmentation induite.

Dans le cas des images mono-dates, les décisions de fusion sous-jacentes aux stratégies de consensus pourraient aussi être guidées par des informations sémantiques. Des progrès récents concernant la modélisation hiérarchique de telles informations sémantiques, dans le cadre de la classification des données de télédétection [9], pourraient faciliter de telles approches.

Dans le cas des images multi-dates, nous pourrions gérer la couverture terrestre de territoires observés en considérant un critère de fusion adapté au modèle de région [2]. De telles informations spatio-temporelles pourraient être utilisées suivant un consensus local entre les images, menant à des hyper-arbres où les branches modéliseraient des décisions de fusions temporelles locales.

## Références

- [1] H. G. Akcay and S. Aksoy. Automatic detection of geospatial objects using multiple hierarchical segmentations. *IEEE Transactions on Geoscience and Remote Sensing*, 46(7):2097–2111, 2008.
- [2] A. Alonso-González, C. López-Martínez, and P. Salembier. PolSAR time series processing with binary partition trees. *IEEE Transactions on Geoscience and Remote Sensing*, 52(6):3553–3567, 2014.
- [3] J. Angulo and D. Jeulin. Stochastic watershed segmentation. In *ISMM*, volume 1, pages 265–276. INPE, 2007.
- [4] J. A. Benediktsson, L. Bruzzone, J. Chanussot, M. Dalla Mura, P. Salembier, and S. Valero. Hierarchical analysis of remote sensing data: Morphological attribute profiles and binary partition trees. In *ISMM*, volume 6671 of *LNCS*, pages 306–319. Springer, 2011.
- [5] K. Bernard, Y. Tarabalka, J. Angulo, J. Chanussot, and J. A. Benediktsson. Spectral-spatial classification of hyperspectral data based on a stochastic minimum spanning forest approach. *IEEE Transactions on Image Processing*, 21(4):2008–2021, 2012.
- [6] K. Cho and P. Meer. Image segmentation from consensus information. *Computer Vision and Image Understanding*, 68(1):72–89, 1997.
- [7] C.-C. Chu and J. K. Aggarwal. The integration of image segmentation maps using region and edge information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(12):72–89, 1993.
- [8] L. Franek, D. Duarte Abdala, S. Vega-Pons, and X. Jiang. Image segmentation fusion using general ensemble clustering methods. In *ACCV*, pages 373–384, 2010.
- [9] C. Kurtz, B. Naegel, and N. Passat. Connected filtering based on multivalued component-trees. *IEEE Transactions on Image Processing*, 23(12):5152–5164, 2014.
- [10] C. Kurtz, N. Passat, P. Gançarski, and A. Puissant. Extraction of complex patterns from multiresolution remote sensing images: A hierarchical top-down methodology. *Pattern Recognition*, 45(2):685–706, 2012.
- [11] M. Mignotte. Segmentation by fusion of histogram-based K-means clusters in different color spaces. *IEEE Transactions on Image Processing*, 17(5):780–787, 2008.
- [12] P. Monasse and F. Guichard. Scale-space from a level lines tree. *Journal of Visual Communication and Image Representation*, 11(2):224–236, 2000.
- [13] T. Rohlfing and C. R. Maurer Jr. Shape-based averaging. *IEEE Transactions on Image Processing*, 16(1):153–161, 2007.
- [14] P. Salembier and 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.
- [15] P. Salembier, A. Oliveras, and L. Garrido. Antiextensive connected operators for image and sequence processing. *IEEE Transactions on Image Processing*, 7(4):555–570, 1998.
- [16] P. Salembier and M. H. F. Wilkinson. Connected operators: A review of region-based morphological image processing techniques. *IEEE Signal Processing Magazine*, 26:136–157, 2009.
- [17] P. Soille. Constrained connectivity for hierarchical image decomposition and simplification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(7):1132–1145, 2008.
- [18] A. Topchy, A. K. Jain, and W. Punch. Clustering ensembles: Models of consensus and weak partitions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(12):1866–1881, 2005.
- [19] J. Vidal, J. Crespo, and V. Maojo. A shape interpolation technique based on inclusion relationships and median sets. *Image and Vision Computing*, 25(10):1530–1542, 2007.
- [20] V. Vilaplana, F. Marques, and P. Salembier. Binary partition trees for object detection. *IEEE Transactions on Image Processing*, 17(11):2201–2216, 2008.
- [21] H. Wang, Y. Zhang, R. Nie, Y. Yang, B. Peng, and T. Li. Bayesian image segmentation fusion. *Knowledge-Based Systems*, 71:162–168, 2014.
- [22] P. Wattuya, K. Rothaus, J.-S. Praßni, and X. Jiang. A random walker based approach to combining multiple segmentations. In *ICPR*, pages 1–4, 2008.