

Nouvelle Méthode en Cascade pour la Classification Hiérarchique Multi-Temporelle ou Multi-Capteur d'Images Satellitaires Haute Résolution

Ihsen Hedhli, Gabriele Moser, Josiane Zerubia

► **To cite this version:**

Ihsen Hedhli, Gabriele Moser, Josiane Zerubia. Nouvelle Méthode en Cascade pour la Classification Hiérarchique Multi-Temporelle ou Multi-Capteur d'Images Satellitaires Haute Résolution. *Revue Française de Photogrammétrie et de Télédétection*, Société Française de Photogrammétrie et de Télédétection, 2018, 216, pp.3-17. hal-01632923

HAL Id: hal-01632923

<https://hal.archives-ouvertes.fr/hal-01632923>

Submitted on 10 Nov 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NOUVELLE MÉTHODE EN CASCADE POUR LA CLASSIFICATION HIÉRARCHIQUE MULTI-TEMPORELLE OU MULTI-CAPTEUR D'IMAGES SATELLITAIRES HAUTE RÉOLUTION

Ihsen Hedhli^{1,2}, Gabriele Moser², Josiane Zerubia¹

1: INRIA, 2004 Route des Lucioles, BP 93, 06902 Sophia-Antipolis Cedex, France

2: Université de Gênes, Dept. DITEN, Via Opera Pia 11a, 16145 Gênes, Italie

Résumé

Ce papier présente un modèle de classification multi-résolution, multi-date et éventuellement multi-capteur fondé sur une modélisation statistique explicite au travers d'un modèle hiérarchique de champs de Markov construit sur une structure quad-arbre. L'approche proposée consiste en un classifieur bayésien supervisé qui combine un modèle statistique conditionnel par classe et un champ de Markov hiérarchique fusionnant l'information spatio-temporelle et multi-résolution. La méthode proposée intègre des informations pixel par pixel à la même résolution. Cela en se basant sur le critère des Modes Marginales a Posteriori (MPM en anglais), qui vise à affecter à chaque pixel l'étiquette optimale en maximisant récursivement la probabilité marginale *a posteriori*, étant donné l'ensemble des observations multi-temporelles ou multi-capteur. Une des originalités de l'approche proposée est l'utilisation en cascade de plusieurs quad-arbres, chacun étant associé à une nouvelle image disponible, en vue de caractériser les corrélations associées à des images distinctes.

Mots clés : Images satellitaires, séries temporelles, multi-résolution, multi-capteur, quad-arbre, classification, champs de Markov hiérarchiques, MPM.

Abstract

This paper describes a method dedicated to multi-resolution, multi-date and eventually multi-sensor classification based on explicit statistical modeling through hierarchical Markov random field modeling based on quad-tree. The proposed approach consists of a supervised Bayesian classifier that combines a joint class-conditional statistical model for pixelwise information and a hierarchical Markov random field for spatio-temporal and multiresolution contextual information fusion based on the Marginal Posterior Mode (MPM). The aim is to recursively maximize the posterior marginal at each pixel, which associates the most probable class label given the entire input information. Within this framework, an interesting novel element of the proposed approach is the use of multiple quad-trees in cascade, each associated with a new image in the available set in order to characterize the correlations associated with distinct images in the data set.

Keywords : Satellite images, image time series, multi-resolution, multi-sensor, quad-tree, classification, hierarchical Markov random fields, MPM.

1. Introduction

Les moyens mis en oeuvre pour surveiller la surface de la Terre, notamment les zones urbaines, en cas de catastrophes naturelles telles que les inondations ou les tremblements de terre, et pour évaluer l'impact de ces événements, jouent un rôle primordial du point de vue sociétal, économique et humain. De très nombreuses applications (scientifiques, civils et militaires) sont faites de ces données. Les satellites ont eu une contribution importante et unique dans certains domaines dont l'évaluation et la surveillance de l'environnement (déforestation, pollutions, pollution lumineuse, urbanisation et périurbanisation, fragmentation écologique, érosion, etc.). Dans ce cadre, des méthodes de classification précises et efficaces sont des outils particulièrement importants pour aider à l'évaluation rapide et fiable des changements au sol et des dommages provoqués.

Étant données l'énorme quantité et la variété des don-

nées disponibles grâce aux missions satellitaires de dernière génération et de différents types (Lillesand et al., 2014) (voir figure 1), la principale difficulté est de trouver un classifieur qui puisse prendre en compte des données multi-bande, multi-résolution (Kato et al., 1996; Laferté et al., 2000; Chardin, 2000), multi-date (Hoberg et al., 2015; Hedhli et al., 2014; Melgani et Serpico, 2003), et éventuellement multi-capteur (Prendes et al., 2015a,b; Hedhli et al., 2015a,b; Voisin, 2012), tout en gardant un temps de calcul acceptable (Tupin et al., 2014).

Considérant une série d'images satellitaires représentant la même surface terrestre, un modèle multi-source est proposé afin de fusionner les informations spatiales (texture), temporelles et multi-résolution. À cette fin, deux approches bayésiennes peuvent être adoptées : un classifieur en cascade (Swain, 1978) obtenu en supprimant le couplage entre les dimensions spatiales et temporelles, et en classifiant chaque image de la série sur la base

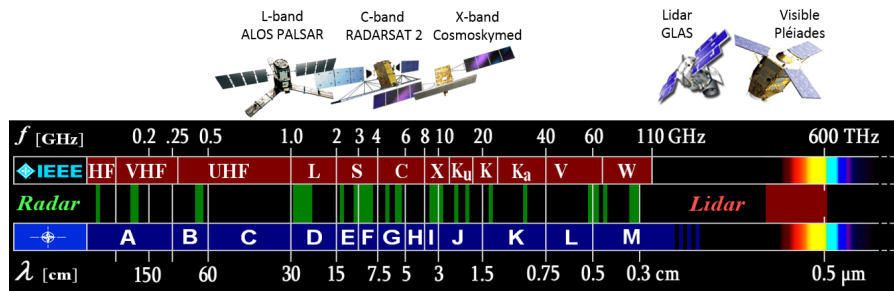


FIGURE 1: Classification d'un ensemble de satellites selon le spectre électromagnétique

d'elle-même et des images précédentes, ou bien l'approche mutuelle qui classe chaque image sur la base des images précédentes et ultérieures (Melgani et Serpico, 2003; Bruzzone et al., 1999; Petitjean et al., 2014). L'approche en cascade est adoptée dans cet article afin de préserver l'ordre temporel des images et d'améliorer la classification à chaque nouvelle donnée disponible. La méthode proposée aborde le problème de la classification multi-temporelle et multi-résolution dans laquelle, la fusion multi-date et multi-résolution est fondée sur une modélisation statistique explicite au travers d'un modèle hiérarchique de champs de Markov (MRF) (Kato et Zerubia, 2012; Li, 2009; Winkler, 2003; Chellappa et Jain, 1993) fusionnant l'information spatio-temporelle et multi-résolution.

Plus précisément, la méthode proposée traite conjointement deux problèmes de fusion de données : fusion de données multi-résolution et fusion de données multi-source (multi-temporelle ou multi-capteur). À notre connaissance, ce problème de fusion est rarement abordé dans la littérature de classification des images de télédétection. Pour résoudre le problème d'utilisation de données hétérogènes, des techniques antérieures de classification d'image multi-résolution ont été fondées, par exemple, sur les transformations en ondelettes (Deng et al., 2014), la segmentation hiérarchique multi-échelle en bas niveau (Akbas et Ahuja, 2014), ou les modèles stochastiques de classification d'images à plusieurs résolutions (Voisin et al., 2014; Kurtz et al., 2012; Laferté et al., 2000). En revanche, ces méthodes utilisent principalement des données mono-temporelles. D'autre part, des méthodes classiques de détection de changement fonctionnent avec des images multi-temporelles, mais en utilisant une seule résolution spatiale. Par exemple, les approches actuelles d'analyse multi-temporelle qui se concentrent généralement sur des séquences assez longues d'images à une seule résolution grossière (de quelques dizaines à une centaine de mètres) et ne visent pas la classification de l'image (Gao et al., 2015; Petitjean et al., 2014). Afin d'utiliser toutes les informations disponibles simultanément, d'autres méthodes proposent un cadre où différents algorithmes peuvent utiliser différentes sources. Ces classificateurs travaillent ensemble de manière collaborative grâce à un raffinement automatique et mutuel de leurs résultats (Forestier et al., 2008). D'autres méthodes du

type apprentissage profond (réseau de neurones profond et réseau de neurones convolutif) sont une autre façon de travailler en multirésolution en tentant de classer les images avec un haut niveau d'abstraction (cf., couches de convolution et de max-pooling) (Chen et al., 2014; Zhao et Du, 2016; Mura et al., 2015).

Tout d'abord, ce papier présente brièvement les données utilisées avant de se focaliser d'avantage sur la structure hiérarchique employée pour organiser ces données hétérogènes provenant des différentes résolutions et de différents capteurs et/ou dates. Ensuite, nous expliquons les choix méthodologiques. Finalement, nous montrons certains résultats de classification sur une série temporelle d'images optiques et un ensemble d'images provenant de différents capteurs (optique et radar) avant de conclure et de présenter quelques perspectives pour le futur.

2. Description des données utilisées

Les images traitées dans le présent travail sont essentiellement de deux types : des images optiques (Prasad et al., 2010) et des images radar (Henderson et Lewis, 1998). Les complémentarités entre ces deux types de données sont multiples. L'imagerie radar est particulièrement sensible aux propriétés géométriques des cibles qu'elles soient microscopiques (rugosité, effets de surface) ou macroscopiques (orientation, multiréflexion), tandis que l'imagerie optique est plus sensible aux propriétés physico-chimiques des cibles (émissivité, albédo, couleur). Parmi les propriétés auxquelles le radar est sensible, la nature des matériaux (cibles métalliques, par exemple), ou leur état (humidité des sols, sécheresse d'une végétation) sont des paramètres importants pouvant être imperceptibles en optique. Nous donnons ci-dessous les caractéristiques des satellites Pléiades (optique) et COSMO-SkyMed (radar) qui correspondent aux images que nous avons utilisées.

2.1. Pléiades

Ce capteur optique permet d'enregistrer les rayonnements diffusés de la surface de la Terre dans certaines longueurs d'ondes. En effet, en fonction de la nature et des caractéristiques intrinsèques des objets à la surface

de la Terre, chaque surface possède une signature spectrale (quantité d'énergie émise ou réfléchiée en fonction de la longueur d'onde) qui lui est propre et qui permet son identification sur les images satellitaires. L'utilisation conjointe de l'information spectrale et de l'information spatiale (texture) présente alors l'intérêt d'exploiter pleinement les possibilités du capteur. Pour chaque prise de vue, le satellite Pléiades capte systématiquement le rayonnement électromagnétique de la surface de la Terre provenant de quatre bandes spectrales (XS) et une bande panchromatique (P) à une résolution spatiale améliorée par rapport aux autres canaux spectraux, et permet une observation bien plus détaillée des surfaces avec une profondeur radiométrique de 16 bits (voir la figure 2).

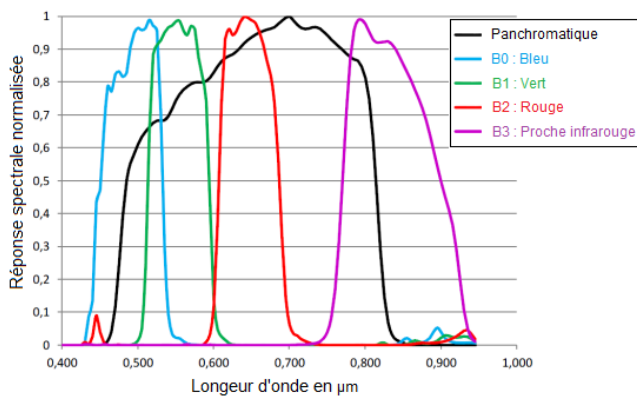


FIGURE 2: Bandes spectrales Pléiades [<http://www.geo-airbusds.com/pleiades/>]

En plus de leurs capacités à discerner des objets distincts, les données de quatre bandes multispectrales sont utiles pour afficher des images en fausse-couleur, et la bande panchromatique peut être utilisée pour améliorer le détail des données multispectrales. En fait, les images conçues peuvent être de type « bundle » (le panchromatique et le XS séparés, fournis comme 2 produits distincts) ou fusionné PXS (« pansharpened » en anglais). Les images panchromatiques sont acquises à une résolution spatiale de 70 cm, et rééchantillonnées au sol à 50 cm. Les images XS sont acquises à une résolution spatiale de 2,8 m, et rééchantillonnées à 2 m. Les images fusionnées sont générées au sol à 50 cm.

2.2. COSMO-SkyMed

Les satellites radar utilisent une antenne d'émission et de réception spécifique de type radar. Le principe repose sur l'émission d'une impulsion courte d'une onde radar à la fréquence porteuse, et l'écoute de son retour. Pour l'imagerie radar, une technologie spécifique, la synthèse d'ouverture (RSO), aussi couramment caractérisé par l'acronyme anglo-saxon SAR pour « Synthetic Aperture Radar », est employée. Il s'agit d'une technique de radar spéciale qui permet d'obtenir des images radar haute résolution d'objets très éloignés. Avec le radar, les micro-ondes servent à mesurer les distances et l'amplitude du signal reçu. Contrairement à un altimètre radar dirigé vers le nadir, le système RSO émet des impulsions

radar latérales. Grâce à ce principe de balayage latéral, le radar reçoit des signaux réfléchis par différents objets sur Terre à différents moments. Cela permet de distinguer les objets composés par des matériaux différents et de caractériser leur texture 3D. La constellation COSMO-SkyMed (CSK) se compose de 4 satellites de moyenne dimension, chacun équipé avec un système RSO à micro-ondes haute résolution en bande X. Le RSO employé a la capacité de changer l'orientation de façon à effectuer l'acquisition des images soit à droite, soit à gauche de la trace au sol du satellite. L'information produite est fondée sur les caractéristiques fondamentales décrites dans la table 1.

3. Structure Hiérarchique

Dans cette partie, nous présentons la structure dans laquelle nous allons organiser les données. Cette structure doit être hautement parallélisable pour traiter une grande quantité de données, et doit fournir une topologie qui simplifie les interactions entre les différentes images. La structure dite « pyramidale » (voir figure 3) peut-être utilisée dans le modèle proposé en insérant les images suivant leurs résolutions respectives (Jolion et Rosenfeld, 2004). Les topologies qui modélisent l'inter-

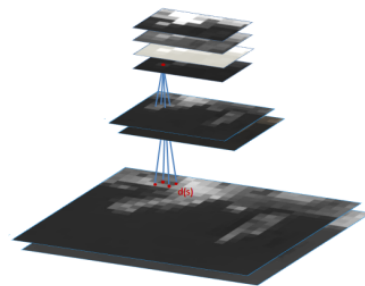


FIGURE 3: Structure Pyramidale

action inter-échelle ont été largement étudiées et leurs représentations conduisent naturellement à des modèles utilisant les arbres. Entre autres, les quad-arbres (Luetgen et al., 1993; Pérez, 1993) ont été proposés pour la modélisation de ces interactions. L'utilisation des quad-arbres permet de bénéficier de leurs bonnes propriétés analytiques (par ex. la causalité (Basseville et al., 1992)) et donc d'appliquer des algorithmes non itératifs. En revanche, il faut noter que cette structure hiérarchique impose un facteur 2 entre les résolutions spatiales dans les niveaux successifs de la pyramide.

Par la suite, nous notons les noeuds sur le quad-arbre avec un indice s qui appartient à l'ensemble des sites $S = S^0 \cup S^1 \cup \dots \cup S^r$ où r désigne la résolution la plus grossière (racine) et 0 correspond la résolution la plus fine. Nous définissons des relations de parenté entre les différents niveaux n du quad-arbre, de telle sorte que si nous considérons un site s pour un niveau n quelconque, nous définissons un opérateur de décalage vers le haut μ tel que $s^- = \mu(s)$ est le parent

	Spotlight	Stripmap		ScanSAR	
		HIMAGE	Ping Pong	Wide Region	Huge Region
Polarisation	Single	Single	Dual	Single	Single
Ampleur [kmxkm]	10X10	40X40	30X30	100X100	200X200
Résolution [m]	1	3	15	30	100

TABLE 1: Les caractéristiques fondamentales suivant les modalités de CSK. [<http://www.cosmo-skymed.it/fr/index.htm>]

du noeud s . Définissons également l'opérateur de décalage vers le bas β tel que $s^+ = \beta(s)$ est un descendant de s , l'opérateur d'échange α entre les noeuds appartenant à la même échelle, et finalement, notons par $d(s)$ l'ensemble contenant s et tous ses descendants (jusqu'à $n = 0$) comme illustré sur la Figure 4.

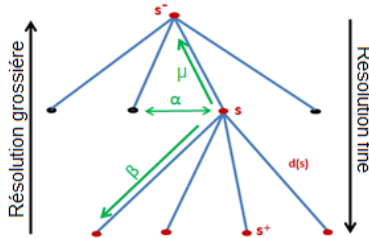


FIGURE 4: Quad-arbre

L'originalité de l'approche proposée est l'extension de la structure hiérarchique présentée ci-dessus pour pouvoir intégrer des données multi-temporelles (voir la Figure 5) ou multi-capteur (voir la Figure 6), cela en utilisant en cascade plusieurs quad-arbres, chacun étant associé à de nouvelles images acquises à une date différente ou provenant d'un autre capteur, en vue de caractériser les corrélations temporelles ou la complémentarité inter-capteur. À cet effet, nous définissons un autre opérateur de décalage vers le haut ω tel que $s^- = \omega(s)$ est l'équivalent du parent du noeud s mais dans le quad-arbre précédent. Nous définissons aussi un opérateur σ d'échange entre les noeuds d'une même échelle et appartenant à des quad-arbres consécutifs pour caractériser la corrélation entre les images. Cette structure permet, d'une manière très naturelle, l'utilisation d'une modélisation statistique explicite à travers un champ de Markov hiérarchique (HMRF) en utilisant un ensemble de champs aléatoires à différentes échelles et pour différentes sources (dates ou capteurs). Les transitions inter-échelle et/ou source (σ , ω , α , β et δ dans les Figures 5 et 6) déterminent la structure hiérarchique du champ de Markov puisqu'elles conservent la causalité des interactions statistiques concernées.

4. Classifieur hiérarchique proposé

L'objectif de la classification est d'estimer un ensemble d'étiquettes cachées \mathcal{X} sachant un ensemble d'observations \mathcal{Y} attachées à l'ensemble des noeuds (pixels) \mathcal{S} .

Chaque étiquette a une valeur dans l'ensemble des étiquettes, $\Lambda = \{0, 1, \dots, M - 1\}$. L'espace des configurations $\Omega = \Lambda^{|\mathcal{S}|}$ est l'ensemble de toutes les configurations possibles.

$\mathcal{X} = \{x_s\}_{s \in \mathcal{S}}, x_s \in \Lambda$ et $\mathcal{Y} = \{y_s\}_{s \in \mathcal{S}}$ sont considérés comme des processus aléatoires. Dans ce contexte, nous étudions le problème d'inférence de la « meilleure » configuration \hat{x} de \mathcal{X} dans Ω . La formulation bayésienne de ce problème d'inférence consiste à minimiser le coût d'erreur de segmentation selon un critère choisi (Schlutter et al., 2012; Ferguson, 1967) :

$$\hat{x} = \arg \min_{x' \in \Omega} \sum_{x \in \Omega} \mathcal{C}(x, x') p(x|y) \quad (1)$$

où $\mathcal{C}(\cdot, \cdot)$ est le coût pénalisant l'écart entre la configuration estimée et la configuration aléatoire idéale.

Parmi les différents algorithmes de classification bayésienne à minimum d'erreur, deux critères ont été largement utilisés. Le premier vise à estimer le « maximum a posteriori » (MAP) où la configuration optimale est obtenue par :

$$\hat{x}_{MAP} = \arg \min_{x \in \Omega} p(x|y) \quad (2)$$

Cela en incluant dans (1) la fonction de coût suivante :

$$\begin{aligned} \mathcal{C}(x, x') &= 1 - \delta(x, x') \\ &= 1 - \prod_{n=0}^R \delta(x^n, x'^n) \end{aligned} \quad (3)$$

(dans le cadre hiérarchique)

où R est le nombre de niveaux de la pyramide et δ est le symbole de Kronecker.

Cette fonction de coût est égale à 1 si une erreur de classification apparaît dans un niveau quelconque de la pyramide. Par conséquent, cette fonction pénalise les erreurs indépendamment de l'échelle à laquelle elles se produisent, ce qui n'est pas très intéressant pour des méthodes de classification hiérarchiques et ne compte pas le nombre total d'erreurs.

Le second critère très utilisé dans la littérature est celui des « Modes Marginales a Posteriori » (MPM en anglais) dont la fonction du coût est définie par :

$$\mathcal{C}(x, x') = \sum_{s \in \mathcal{S}} 1 - \delta(x_s, x'_s) \quad (4)$$

L'utilisation de cette fonction offre l'avantage de pénaliser les erreurs en fonction de leur nombre et de l'échelle à laquelle elles se produisent : par exemple, une erreur à une

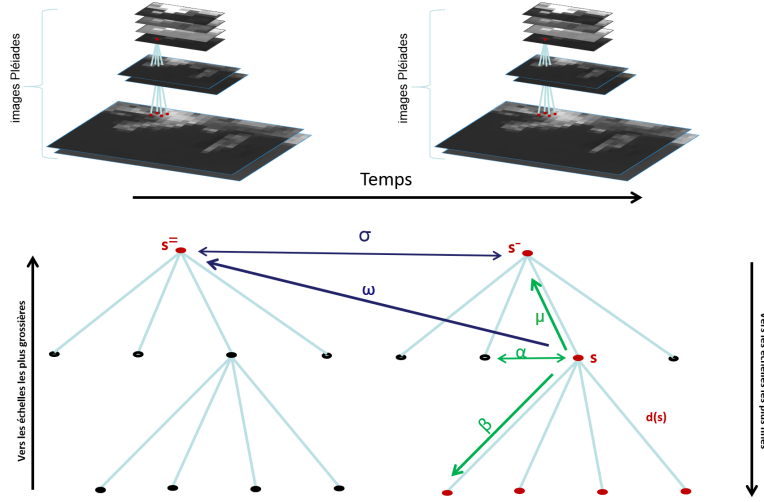


FIGURE 5: Structure hiérarchique multi-temporelle

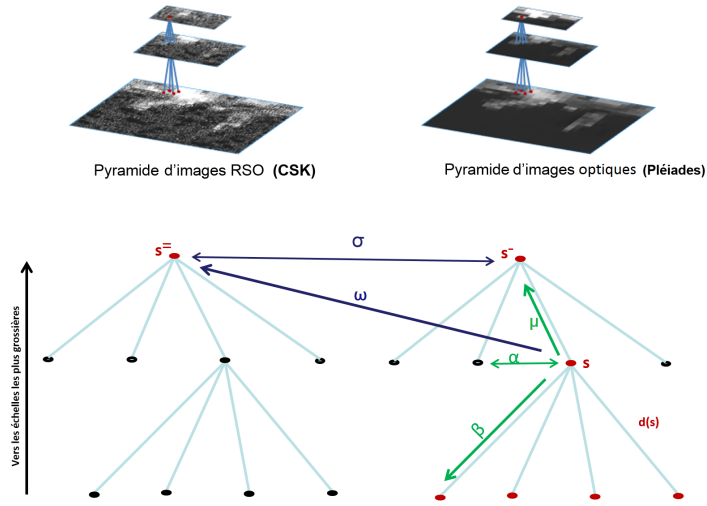


FIGURE 6: Structure hiérarchique multi-capteur

échelle grossière est plus fortement pénalisée qu'une erreur à une échelle plus fine. L'estimateur bayésien résultant est défini pour chaque pixel s par :

$$\forall s \in \mathcal{S}, \hat{x}_s = \arg \max_{x_s \in \Lambda} p(x_s | y) \quad (5)$$

Parmi les méthodes utilisées pour résoudre ce problème d'optimisation, on distingue la méthode fondée sur la structure quad-arbre employée par (Laferté et al., 2000) qui vise à affecter à chaque pixel s l'étiquette optimale \hat{x}_s , en maximisant récursivement la probabilité marginale *a posteriori* $p(x_s | y)$, cela en exprimant la probabilité marginale du pixel s à travers la probabilité marginale $p(x_{s^-} | y)$ du parent s^- . Malgré ses performances en temps de calcul et en précision, l'utilisation de cet algorithme conduit à des artefacts de blocs (Laferté et al., 2000). Ce phénomène peut être expliqué par le fait que, à une échelle donnée, deux pixels voisins peuvent ne pas avoir le même parent. Dans ce cas, une limite apparaît beaucoup plus

facilement que s'ils étaient liés au même parent ou s'il existait d'autres liaisons entre les pixels à des échelles consécutives. Afin d'éviter ces artefacts, nous pouvons utiliser une structure hiérarchique plus connectée dans laquelle la liaison causale ascendant-descendant dans la même pyramide n'est pas l'unique liaison entre les niveaux consécutifs (cf. par exemple (Bouman et Shapiro, 1994)).

Dans ce contexte, une nouvelle formulation de l'algorithme fondé sur le critère MPM est proposée dans cet article en utilisant la structure hiérarchique en cascade définie dans la partie 3. La probabilité marginale *a posteriori* $p(x_s | y)$ de chaque pixel s , est exprimée non seulement à travers la probabilité marginale *a posteriori* $p(x_{s^-} | y)$ du parent s^- , mais aussi à travers la probabilité marginale *a posteriori* $p(x_{s^=} | y)$ du parent à une source antérieure $s^=$ afin de caractériser les corrélations entre des images consécutives de la structure en cascade définie

et éviter ainsi les artéfacts de blocs :

$$\begin{aligned}
p(x_s|y) &= \sum_{s^-, s=} p(x_s|x_{s^-}, x_{s=}, y).p(x_{s^-}, x_{s=} | y) \\
&= \sum_{s^-, s=} p(x_s|x_{s^-}, x_{s=}, y_{d(s)}).p(x_{s^-}, x_{s=} | y) \\
&= \sum_{s^-, s=} \mathcal{A} \cdot p(x_{s^-}, x_{s=} | y) \\
&= \sum_{s^-, s=} \mathcal{A} \cdot p(x_{s^-} | y).p(x_{s=} | y) \quad (6)
\end{aligned}$$

où

$$\begin{aligned}
\mathcal{A} &= p(x_s|x_{s^-}, x_{s=}, y_{d(s)}) \\
&= \frac{p(x_s, x_{s^-}, x_{s=} | y_{d(s)})}{\sum_s p(x_s, x_{s^-}, x_{s=} | y_{d(s)})} \quad (7)
\end{aligned}$$

Il est à noter que les équations dans (6) impliquent deux hypothèses d'indépendance conditionnelle :

- pour le passage de la ligne 1 à la ligne 2 dans (6) : étant donné les étiquettes x_{s^-} et $x_{s=}$, l'étiquette x_s ne dépend pas de tout l'ensemble d'observations y , elle dépend seulement des observations attachées à s et à ses descendants ce qui implique que :

$$p(x_s|x_{s^-}, x_{s=}, y) = p(x_s|x_{s^-}, x_{s=}, y_{d(s)})$$

- pour le passage de la ligne 3 à la ligne 4 dans (6) : les étiquettes x_{s^-} et $x_{s=}$ sont indépendantes, étant donné l'ensemble d'observation y , ce qui implique :

$$p(x_{s^-}, x_{s=} | y) = p(x_{s^-} | y).p(x_{s=} | y)$$

Il reste à calculer les probabilités jointes $p(x_s, x_{s^-}, x_{s=} | y_{d(s)})$ évoquées dans (7) qui peuvent être déterminées en estimant les probabilités introduites dans la factorisation suivante :

$$\begin{aligned}
p(x_s, x_{s^-}, x_{s=} | y_{d(s)}) &= \quad (8) \\
p(x_s|x_{s^-}, x_{s=}) \cdot \frac{p(x_{s^-} | x_{s=}) \cdot p(x_{s=} | x_{s^-})}{p(x_{s^-}, x_{s=})} \cdot p(x_s | y_{d(s)})
\end{aligned}$$

incluant :

- des probabilités *a priori* ($p(x_{s=})$ et $p(x_{s^-})$),
- des probabilités inter-échelle et/ou multi-source ($p(x_s|x_{s^-})$ et $p(x_{s^-}|x_{s=}, x_{s=})$),
- des probabilités *a posteriori* partielles $p(x_s | y_{d(s)})$.

Par la suite, à chaque nouvelle donnée t (nouvelle image à une date différente ou acquise par un capteur différent), pour estimer la probabilité marginale *a posteriori* $p(x_s|y)$, il suffit de calculer les probabilités évoquées dans (8). A cet effet, l'approche proposée se déroule en trois passes récursives, « descendante n° 1 », « ascendante n° 1 » et « descendante n° 2 » (voir Algorithme 1 et Figure 7).

4.1. Passe n° 0 : MPM sur la première pyramide

Selon l'approche en cascade, la classification est effectuée d'abord à la première pyramide (i.e., l'instant $t = 0$ ou le premier capteur dans le cas multi-capteurs) en

utilisant une formulation *MPM* à une seule pyramide, dans laquelle une segmentation est obtenue récursivement sur les échelles à travers des passes ascendantes et descendantes. Les détails de cette formulation (i.e., une seule pyramide) peuvent être trouvés dans (Laferté et al., 2000). Nous rappelons seulement que le processus est initialisé par la définition de la distribution de probabilité *a priori* sur la racine du quad-arbre correspondant à la première pyramide. Cette initialisation est nécessaire pour commencer une récursion descendante et calculer les probabilités *a priori* dans tous les niveaux du quad-arbre. Une stratégie d'initialisation simple consiste à utiliser une distribution *a priori* uniforme. À la suite de ce traitement à la première pyramide, les probabilités marginales *a posteriori* $p(x_s|y)$ seront connues pour chaque pixel du quad-arbre.

4.2. Passe descendante n° 1 : estimation des probabilités *a priori*

Afin de calculer les probabilités *a priori* dans la pyramide \mathcal{P} (à une nouvelle date t dans le cadre multi-temporel ou avec un nouveau capteur c dans le cadre multi-capteur), nous procédons à une passe descendante, en commençant par estimer les probabilités *a priori* à la résolution la plus grossière (étape 1 dans Algorithme 1 et Figure 7). Pour ce faire, nous utilisons la carte de classification obtenue en utilisant la structure hiérarchique en cascade jusqu'à la pyramide $\mathcal{P} - 1$ en considérant un modèle markovien spatial (Li, 2009) qui conduit à une meilleure estimation de l'*a priori*. Ainsi, en utilisant le théorème de Hammersley-Clifford, (Hammersley et Clifford, 1971; Moussouris, 1974) nous pouvons définir un *a priori* local pour chaque pixel s à la résolution la plus grossière comme étant un modèle de Potts (Winkler, 2003) décrit ci-dessous :

$$p(x_s|x_{s'}, s' \sim s) \propto \exp\left(-\beta \sum_{s \sim s'} \delta(x_s, x_{s'})\right) \quad (9)$$

Où $\delta(\cdot)$ est le symbole de Kronecker, $s \sim s'$ signifie que s et s' sont voisins par rapport à un système de voisinage donné et β est un paramètre de lissage, de signe positif pour faire décroître l'énergie lorsque les étiquettes sont identiques.

Puis, une passe descendante est effectuée pour calculer les probabilités *a priori* à chaque niveau du quad-arbre (étape 2 dans Algorithme 1 et Figure 7) en utilisant la formulation récursive suivante :

$$p(x_s) = \sum_{s^-} p(x_s|x_{s^-}) \cdot p(x_{s^-}) \quad (10)$$

Les probabilités de transition inter-échelle au sein de l'arbre $p(x_s|x_{s^-})$ peuvent être calculées indépendamment de telle sorte que x_s appartienne à la même classe que son ascendant x_{s^-} .

4.3. Passe ascendante n° 1 : estimation des probabilités jointes

En se basant sur (8), les probabilités jointes $p(x_s, x_{s^-}, x_{s=} | y_{d(s)})$ sont déterminées par :

- les probabilités *a priori* $p(x_s)$, déjà estimées lors de la passe descendante précédente,
- les probabilités inter-échelle et/ou inter-pyramide $p(x_s|x_{s-})$ et $p(x_s|x_{s-}, x_{s=})$, qui peuvent être calculées en favorisant l'appartenance des pixels concernés à une même classe ; L'hypothèse fondamentale est de considérer le processus aléatoire \mathcal{X} **comme markovien en échelle et entre les pyramides** (chaque pixel dépend seulement de son ascendant dans la même pyramide et/ou son ascendant dans la pyramide précédente). Ce sont ces probabilités de transition inter-échelles qui spécifient le champ de Markov, car elle traduisent la causalité des interactions statistiques entre les différentes échelles et/ou pyramides. Nous avons opté pour la forme de la probabilité de transition adopté dans (Bouman et Shapiro, 1994; Hedhli et al., 2015a)
- les probabilités *a posteriori* partielles $p(x_s|y_{d(s)})$ qui peuvent être calculées moyennant une passe ascendante en utilisant la formulation introduite par (Laferté et al., 2000) :

$$p(x_s|y_{d(s)}) \propto p(y_s|x_s) \cdot p(x_s) \cdot \prod_{\tilde{s} \in s^+} \sum_{\tilde{s}} \frac{p(x_{\tilde{s}}|y_{d(\tilde{s})})}{p(x_{\tilde{s}})} \cdot p(x_{\tilde{s}}, x_s) \quad (11)$$

Les probabilités *a posteriori* partielles nécessitent l'estimation d'un modèle statistique conditionnel par classe combinant des informations pixel par pixel, à la même résolution pour le calcul de $p(y_s|x_s)$. Cela peut se faire en traitant la modélisation des probabilités marginales des images acquises à chaque résolution, conditionnées à chaque classe (voir la partie 4.5).

4.4. Passe descendante n° 2 : probabilité *a posteriori* et optimisation

Connaissant les probabilités jointes calculées dans la partie précédente (4.3), nous procédons à une passe descendante dans laquelle les probabilités marginales *a posteriori* $p(x_s|y)$ sont calculées en utilisant (6). En principe, l'étiquette x_s qui maximise $p(x_s|y)$ sur l'ensemble des classes Λ peut être sélectionnée et attribuée à x_s . En revanche, cette procédure est souvent évitée dans la littérature sur les MRF hiérarchiques en raison de son temps de calcul (linéaire par rapport au nombre de classes, nombre de pixels à chaque échelle et au nombre de pyramides utilisés) et aussi pour éviter les artéfacts de blocs (Laferté et al., 2000). En tant qu'approche alternative, ici, une formulation spécifique de l'algorithme de Metropolis modifié (Kato, 1994) (MMD) est appliquée séparément pour chaque échelle de chaque pyramide. Plus précisément, à chaque échelle r et pyramide p le MMD est utilisé pour minimiser l'énergie $U(\mathcal{X}_t^r|y)$ par rapport à l'ensemble \mathcal{X}_t^r de configurations d'étiquettes possibles :

$$U(\mathcal{X}_t^r|y) = \sum_{s \in S_p^r} \log(p(x_s|y)) - \beta \sum_{s \sim s'} \delta(x_s, s_{s'}) \quad (12)$$

Où le premier terme de l'équation (12) est exprimé en termes de probabilités *a posteriori* calculées par le

critère MPM et le deuxième terme est dû au modèle de Potts sur la couche r du pyramide p .

4.5. Estimation des fonctions de densité de probabilité

Nous cherchons à modéliser les distributions statistiques de chaque classe m ($m \in \Lambda$) considérée pour la classification, étant donnée une base d'apprentissage. Pour chaque classe, les fonctions de probabilités (FDP) $p_m(y_s|x_s = m)$ suivent un modèle de mélange fini de K distributions de niveaux de gris indépendantes :

$$p_m(y_s|x_s = m) = \sum_{i=1}^{K^{mnt}} \Pi_i^{mnt} \mathcal{F}_i^{mnt}(y_s|\theta_i^{mnt}), \quad (13)$$

où Π_i^{mnt} sont les proportions du mélange, θ_i^{mnt} est l'ensemble des paramètres de la $i^{\text{ème}}$ composante du mélange pour la $m^{\text{ème}}$ classe, la $n^{\text{ème}}$ échelle et la $t^{\text{ème}}$ pyramide. \mathcal{F}_i^{mnt} est la loi de probabilité correspondante dont les paramètres θ_i^{mnt} sont estimés en ayant recours à un algorithme stochastique d'espérance-maximisation (SEM) (Celeux et al., 1996). Les lois choisies, en revanche, varient selon la nature de l'image. En fait, dans le cadre des images optiques, les FDP \mathcal{F}_i^{mnt} sont modélisées au travers d'un mélange gaussien (Richards, 2013). Contrairement aux images RSO où l'information radar est mélangée à un bruit multiplicatif (Henderson et Lewis, 1998). C'est pourquoi, nous utilisons la distribution Gamma généralisée comme modèle empirique pour la modélisation des images RSO (H.C. Li et Fan, 2011).

5. Résultats

Avant de comparer les résultats du modèle développé avec ceux obtenus via des méthodes de référence, il faut noter que la structure hiérarchique proposée suppose une contrainte sur les résolutions des images pour les différents niveaux de la pyramide. En fait, le choix de la topologie quad-arbre impose un facteur deux entre les résolutions spatiales des images dans les niveaux successifs de la pyramide. Cela implique que des niveaux manquants peuvent survenir. Ces niveaux « vides » sont remplis grâce à une décomposition en ondelettes (Mallat, 2008) de l'image intégrée au niveau de la résolution la plus fine. De nombreuses familles d'ondelettes ont été testées. Les résultats sont relativement proches, et la principale différence se situe au niveau de la finesse de la classification finale. Nous remarquons que la plus grande précision est atteinte en utilisant des ondelettes de type Daubechies 10 (db10) (Daubechies, 1988) voir Figure 8.

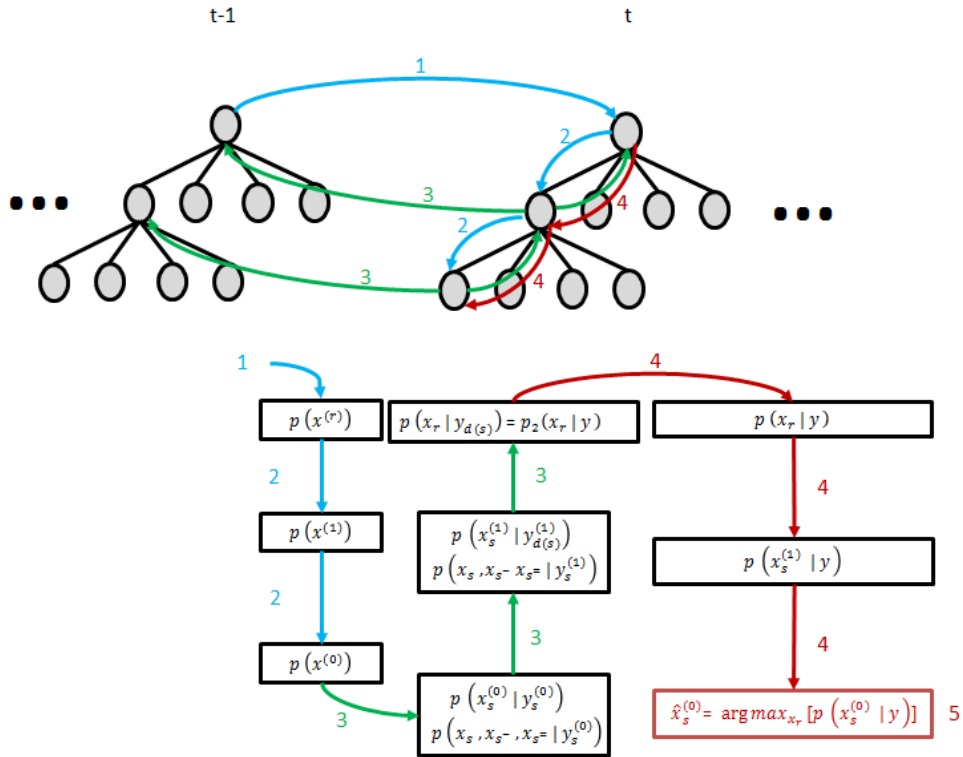


FIGURE 7: Champ de Markov hiérarchique multi-temporel ou multi-capteur et estimation des probabilités utilisées dans l'algorithme 1

Algorithm 1 MPM multi-source (i.e. multi-temporel ou multi-capteur)

Conditions initiales :

- carte de classification à la racine du quad-arbre à la pyramide \mathcal{P}_0
- \mathfrak{P} , le nombre total des pyramides.

à une pyramide t :

- 1: Initialisation de l'*a priori* en utilisant la carte de classification à la pyramide $\mathcal{P} - 1$ via (9).
 - 2: Passe descendante n° 1 : estimation des probabilités *a priori* via (10)
 - 3: Passe ascendante n° 1 : estimation des probabilités $p(x_s | y_{d(s)})$, puis $p(x_s, x_{s-}, x_s = | y_{d(s)})$ via (8) et (11)
 - 4: Passe descendante n° 2 : estimation des probabilités *a posteriori* $p(x_s | y)$ via (6)
 - 5: Résultat : obtention de la carte de classification par la maximisation des probabilités $p(x_s | y)$ à la résolution la plus fine.
 - 6: si $\mathcal{P} < \mathfrak{P}$ aller à 1.
-

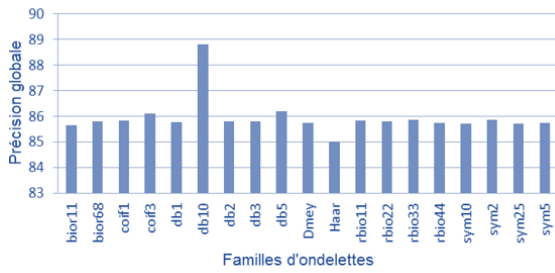


FIGURE 8: Précision globale en utilisant différentes familles d'ondelettes

Le classifieur hiérarchique décrit précédemment est assez générique pour pouvoir être utilisé dans un cadre multi-temporel ou multi-capteur. Pour ce faire, nous proposons de le valider expérimentalement sur une série temporelle d'images Pléiades (voir la partie 5.1). Puis, nous utilisons deux ensembles d'images RSO (CSK) et optique (Pléiades) dans le cadre muticapteur (comme illustré dans la partie 5.2).

5.1. Cas multi-temporel

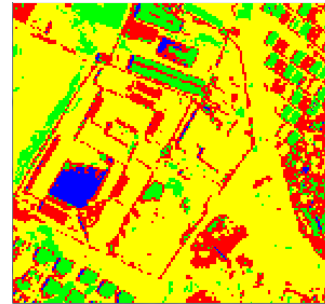
Dans cette partie, une série temporelle de trois images Pléiades panchromatiques (rééchantillonnées à 0.5 m) et multi-spectrales (rééchantillonnées à 2 m) acquises sur Port-au-Prince (Haïti) en 2011, 2012 et 2013 sont utilisées. L'algorithme proposé est donc appliqué en utilisant trois pyramides de trois niveaux ($R = 2$). Ce choix de configuration semble être optimum. En effet, un niveau de décomposition plus élevé pourrait supprimer certains détails pertinents de l'image et un choix de niveau plus faible ne permet pas d'exploiter pleinement la structure hiérarchique. Quant aux images Pléiades utilisées dans cet article, la résolution la plus fine de la pyramide multi-résolution (niveau 0) correspond à la résolution la plus fine des images d'entrée (i.e. panchromatique). Les bandes multi-spectrales enregistrées simultanément sont intégrées dans le niveau 2 de la pyramide (la racine). Le niveau 1 manquant est complété par la décomposition en ondelettes db10 de l'image panchromatique.

Cinq classes d'occupation du sol ont été considérées : urbain (rouge), eau (bleu), végétation (vert), sol nu (jaune) et conteneurs (mauve). L'étape d'apprentissage est faite sur une vérité terrain construite manuellement par photo-interprétation. La méthode proposée donne une précision globale de la classification de valeur élevée avec un faible temps de calcul (voir figure 9, table 2).

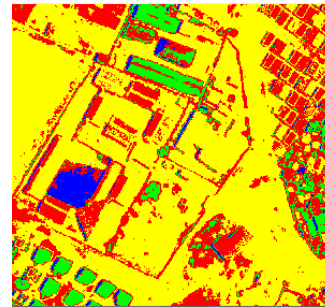
Nous comparons le modèle présenté avec les résultats de la classification obtenus en utilisant les techniques suivantes :

- Une approche hiérarchique à une seule date proposée dans (Laferté et al., 2000) où la classification est obtenue récursivement en appliquant le critère MPM dans une structure quad-arbre. L'introduction de l'aspect multi-temporel permet non seulement de contourner les artéfacts de blocs causés par l'utilisation de la structure quad-arbre

(voir Figure 11), mais aussi d'obtenir une meilleure discrimination des classes (voir les figures 9 et 10 et le tableau 2). L'une des principales sources d'erreur de classification dans les résultats utilisant une seule date est la confusion entre les classes « végétation » et « sol nu ». Cette erreur de classification est réduite dans les résultats obtenus par la méthode multi-temporelle proposée grâce à l'aspect spatio-temporel.



(1)



(2)

FIGURE 11: (1) Artéfact de blocs apparaissant avec la méthode proposée dans (Laferté et al., 2000), (2) réduction des effets de blocs en utilisant la méthode proposée

- Une approche markovienne multi-temporelle et mono-échelle proposée dans (Melgani et Serpico, 2003). Contrairement à l'approche en cascade proposée où l'information temporelle se propage dans un seul sens pour préserver la causalité du modèle, la méthode consiste à réaliser un échange bidirectionnel de l'information temporelle entre des modèles MRF définis à partir des images consécutives de la série temporelle en utilisant un perceptron multicouche (MLP en anglais) de réseau de neurones. Dans une forme de fonction d'énergie appropriée, chaque modèle MRF (associé à chaque image de la série temporelle) intègre trois types d'information : spectrale, contextuelle et spatio-temporelle. L'approche mutuelle réduit le risque de la propagation des erreurs de classification entre deux dates consécutives de la série temporelle en comparaison avec l'approche en cascade ci-dessus où les erreurs peuvent se propager en raison de la causalité de la méthode. Toutefois, en raison de l'aspect non itératif de l'algorithme hiérarchique utilisé, seulement quelques minutes

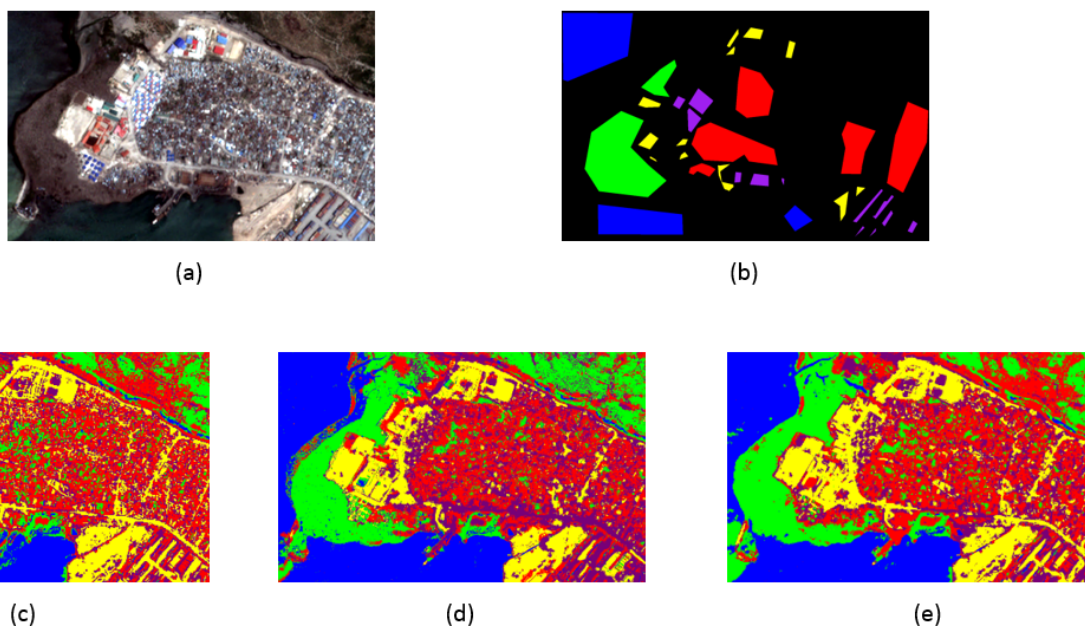


FIGURE 9: Première parcelle de test : (a) Image Pléiades, (© CNES distribution Airbus DS, 2013). (b) Vérité Terrain. (c) Carte de classification en utilisant la méthode de (Laferté et al., 2000). (d) Carte de classification en utilisant une série temporelle de trois images Pléiades en appliquant la méthode de (Melgani et Serpico, 2003). (e) Carte de classification en utilisant la méthode proposée. (urbain, eau, végétation, sol nu, conteneurs)

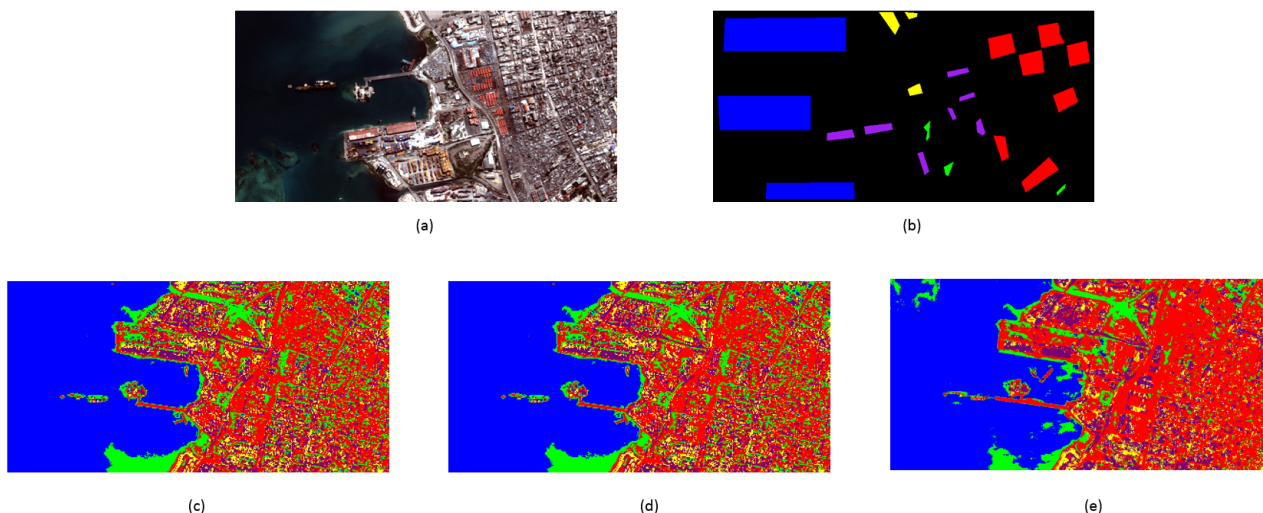


FIGURE 10: Deuxième parcelle de test : (a) Image Pléiades, (© CNES distribution Airbus DS, 2013). (b) Vérité Terrain. (c) Carte de classification en utilisant la méthode de (Laferté et al., 2000). (d) Carte de classification en utilisant une série temporelle de trois images Pléiades en appliquant la méthode de (Melgani et Serpico, 2003). (e) Carte de classification en utilisant la méthode proposée. (urbain, eau, végétation, sol nu, conteneurs)

	Port-au-Prince, Haïti					total	temps de calcul
	urbain	eau	végétation	sol nu	conteneurs		
La méthode proposée	81.62%	100%	96.69%	92.82%	62.82%	85,59 %	8 minutes
(Laferté et al., 2000)	77.45%	88.62%	72.59%	86.02%	57.73%	76,34%	3 minutes
(Melgani et Serpico, 2003)	80.63%	100%	86.33%	87.61%	69.61%	84,83 %	≈ 1 heure

TABLE 2: Taux de bonne classification pour chacune des 5 classes considérées et taux moyen de bonne classification sur les deux parcelles de test utilisées

ont été nécessaires pour obtenir des résultats satisfaisants très similaires à ceux obtenus par l'approche « mutuelle » pour laquelle un long temps de calcul est nécessaire.

5.2. Cas multi-capteur

Nous testons maintenant l'algorithme de classification hiérarchique proposé sur un jeu de données multi-capteur composé de deux images recalées de Port-au-Prince : (i) une image CSK mono-polarisation et mono angle d'incidence avec une polarisation HH, une modalité Stripmap et une résolution de 2,5 m (© ASI ; figure 12 (a)), et (ii) une image multispectrale Pléiades de résolution initiale 2 m, illustrée dans la Figure 12 (b) (© CNES distribution Airbus DS, 2013). Les caractéristiques techniques des images considérées ont été décrites dans la partie 2, la méthode proposée est comparée avec les résultats de classification obtenus avec les techniques suivantes :

- Les classifications hiérarchiques pour chaque capteur séparément (voir la figure 12 (e) pour l'image Pléiades et la Figure 12 (d) pour l'image CSK) en utilisant la méthode de (Laferté et al., 2000).
- L'approche multi-capteur proposée par (Storvik et al., 2009) dans laquelle la vraisemblance est construite en fusionnant les lois marginales : gamma généralisée (pour l'image RSO) et gaussienne (pour l'image optique) pour obtenir une distribution méta-gaussienne (Fang et al., 2002). La classification est obtenue ensuite par l'estimation du maximum de vraisemblance.
- L'approche multi-capteur et multi-échelle proposée dans (Voisin et al., 2014) qui prend en compte à la fois une représentation statistique marginale des images RSO et optique par modèle de mélanges finis et de copules pour calculer les distributions jointes ainsi qu'une modélisation contextuelle à partir de champs de Markov hiérarchiques.

une première analyse visuelle des résultats de classifications suggère que la méthode hiérarchique proposée conduit à de meilleurs résultats, en comparaison avec les classifications hiérarchiques de chaque image séparée. En effet, les résultats expérimentaux obtenus en utilisant uniquement l'image radar détecte mieux les routes et les conteneurs, tandis que le résultat généré en utilisant seulement l'image optique discrimine mieux les classes, qui sont spatialement homogènes. La méthode proposée bénéficie à la fois des caractéristiques de l'imagerie optique et de l'imagerie radar en générant des résultats de classification qui, visuellement, discriminent toutes les classes dans les images à haute résolution considérées. Comparée à la méthode présentée dans (Storvik et al., 2009) dans laquelle la fusion des données optique et radar est obtenue via une distribution méta-gaussienne bivariable, la méthode proposée génère une carte de classification plus régulière grâce à l'aspect contextuel du champs de Markov (voir la figure 12 (g) et le table 3). D'autre part, en comparant notre méthode avec la méthode développée dans (Voisin et al., 2014) la méthode

proposée montre un gain en temps de calcul, car la méthode décrite dans (Voisin et al., 2014) utilise des copules pour extraire les informations des deux capteurs, et également un gain en précision en évitant les artefacts de blocs causés par l'utilisation d'un seul quad-arbre (voir la figure 12 (f) et le table 3).

5.3. Étude de faisabilité

Nous analysons dans cette partie la complexité des algorithmes proposés en terme de temps de calculs et de mémoire utilisée. Le classifieur supervisé décrit dans cet article est composé de deux parties exécutées séparément : l'apprentissage obtenu grâce à l'algorithme SEM (voir le début partie 4.5) et la classification fondée sur le critère MPM (voir la partie 4). Ces algorithmes sont des algorithmes de type Monte-Carlo appartenant à la famille des algorithmes randomisés dont le temps d'exécution est déterministe, mais dont le résultat peut être incorrect avec une certaine probabilité (généralement minime) (Motwani et Raghavan, 1995). Les algorithmes Monte-Carlo appartiennent à la classe de complexité BPP (pour « bounded-error probabilistic polynomial time ») qui décrit les problèmes de décision qui peuvent être résolus en temps polynomial, avec une probabilité bornée d'erreur. Les Figures 13 et 14 présentent les mesures quantitatives des temps d'exécution des algorithmes proposés suivant le nombre de pixels dans les images utilisées. Il faut noter ici que dans l'algorithme de classification nous utilisons deux pyramides à 3 niveaux chacune. En faisant varier les tailles des images utilisées, la Figure 13 montre que cet algorithme s'exécute en un temps polynomial. En ce qui concerne l'algorithme d'apprentissage, il est exécuté pour chaque image de la pyramide. La figure 14 montre l'exécution de cet algorithme est réalisée un temps logarithmique.

En ce qui concerne la mémoire utilisée, elle dépend essentiellement du nombre d'images et du nombre de classes utilisées à chaque niveau de la pyramide. En effet, il faut implémenter des structures de données pour manipuler les différentes probabilités à chaque pixel et pour chaque étiquette. L'optimisation du temps de calcul serait intéressante dans le but de pouvoir traiter des images bien plus grandes (par exemple, une acquisition Pléiades de 20 km de largeur et 20 km de longueur), correspondant à des tailles typiques d'acquisition. Actuellement, le code implémenté ne permet pas de traiter de telles images si on n'effectue pas un prédecoupage avant le traitement.

6. Conclusion et Perspectives

Dans la méthode que nous avons proposée, la fusion multi-résolution, multi-temporelle ou multi-capteur est fondée sur une modélisation statistique explicite au travers d'un modèle hiérarchique de champs de Markov. La nouveauté principale de l'approche est l'utilisation en cascade de plusieurs quad-arbres, chacun étant associé à une nouvelle image disponible, en vue de caractériser les corrélations temporelles ou spatiales associées à des

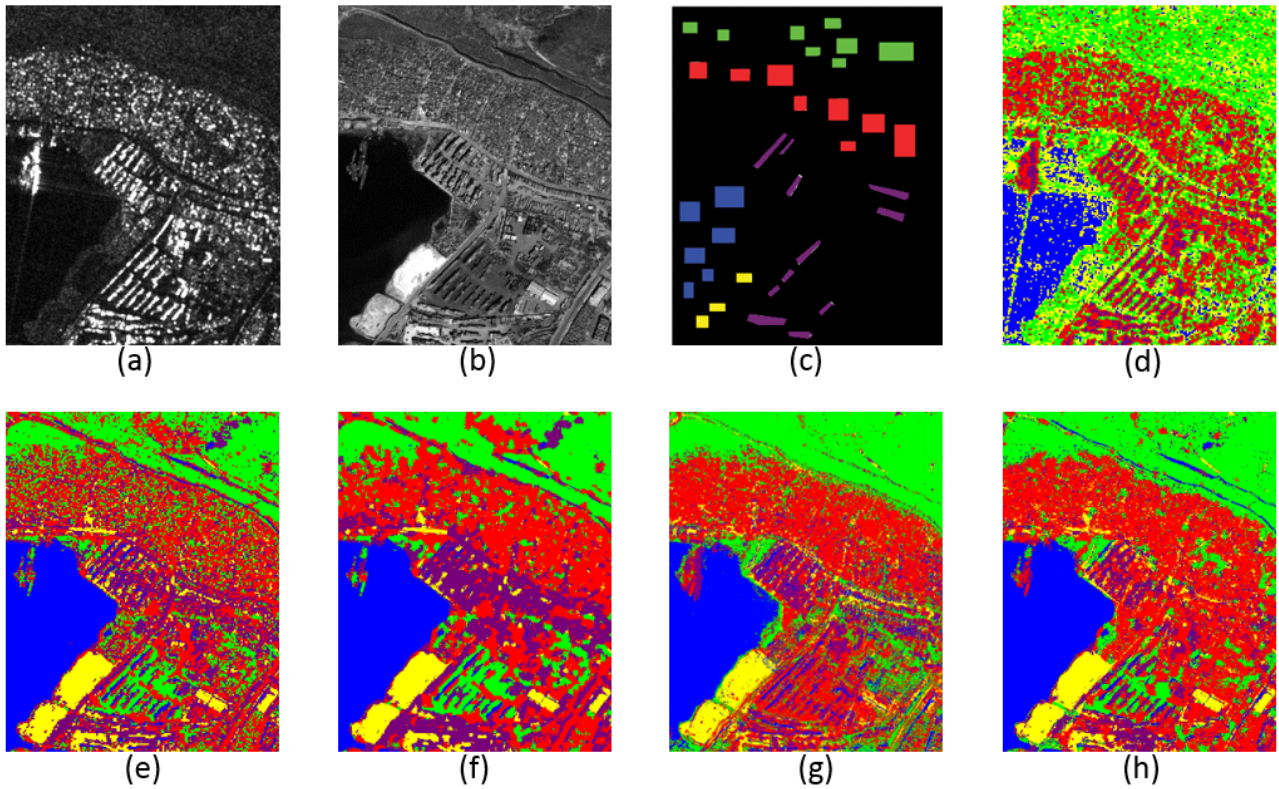


FIGURE 12: (a) CSK, (© ASI, 2013) (b) Pléiades, (© CNES distribution Airbus DS, 2013). (c) Vérité Terrain. (d) Carte de classification en utilisant la méthode de (Laferté et al., 2000) sur l'image RSO (e) Carte de classification en utilisant la méthode de (Laferté et al., 2000) sur l'image optique (f) Carte de classification en utilisant la méthode de (Voisin et al., 2014) (g) Carte de classification en utilisant la méthode de (Storvik et al., 2009) sur l'image optique (h) Carte de classification en utilisant la méthode proposée.

	Port-au-Prince, Haïti					total	temps de calcul
	eau	urbain	végétation	sol nu	conteneurs		
La méthode proposée	100%	78.12%	96.69%	89.78%	47.12%	82.69 %	4 minutes
(Storvik et al., 2009)	99.95%	97.32%	90.81%	96.22%	37.25%	79.44%	5 minutes
(Voisin et al., 2014)	100%	75.24%	87.16%	98.89%	49.31%	82,12 %	11 minutes

TABLE 3: Taux de bonne classification pour chacune des 5 classes considérées et taux moyen de bonne classification Les résultats sont obtenus en utilisant une image de taille (2400x1600) au niveau 0 de la pyramide en exécutant l'algorithme proposé avec un ordinateur dont les caractéristiques sont : un processeur Intel i7 quad-core (2.40 GHz) avec une RAM de 8-GB et une architecture 64-bit sous un système Linux.

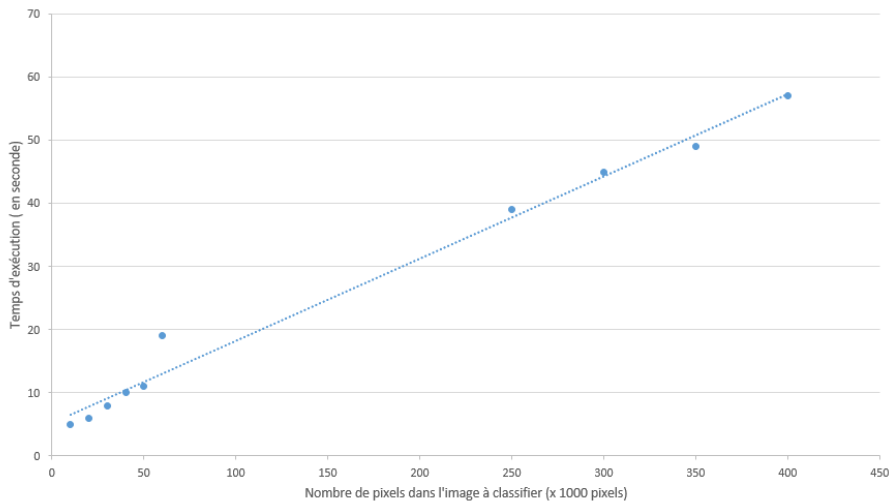


FIGURE 13: Temps d'exécution de l'algorithme de classification suivant le nombre de pixels dans l'image à la résolution la plus fine de la pyramide

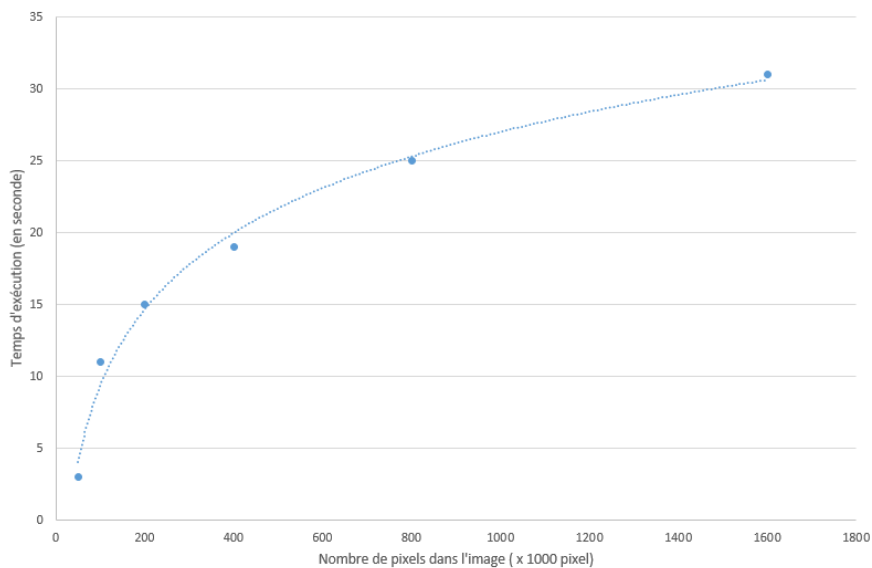


FIGURE 14: Temps d'exécution de l'algorithme d'apprentissage suivant le nombre de pixels pour chaque image de la pyramide

images distinctes. Lorsqu'elle est appliquée à des images à haute résolution, la méthode proposée donne une précision globale de la classification de valeur élevée avec un temps de calcul raisonnable grâce à la structure hiérarchique utilisée. Il convient de noter que le nombre de classes est identique pour tous les niveaux de la structure hiérarchique définie dans l'approche proposée. Cette hypothèse est stricte car les différents types d'occupation des sols dépendent naturellement de la résolution de l'image. Par conséquent, une extension de ce travail serait de définir un nombre différent de classes à chaque niveau de la pyramide et d'établir les liens hiérarchiques entre ces classes en fonction de leur signification sémantique respective, en se basant sur des vérités de terrain appropriées. De plus, une parallélisation de l'algorithme proposée pourrait être envisagé.

7. Remerciements

Une partie de ce travail de recherche a été soutenue par le Centre National d'Etudes Spatiales (CNES) via la convention CNES (131026) - INRIA (8361) dans le cadre de KAL-Haïti. Nous remercions le CNES pour les images Pléiades qui nous ont été fournies dans le cadre de cette étude et l'Agence Spatiale Italienne (ISA) pour les images Cosmo-SkyMed. Les auteurs remercient également, Alain Giros du CNES et Sebastiano B. Serpico de l'université de Gênes pour des discussions très constructives.

Références

Akbas, E., Ahuja, N., 2014. Low-level hierarchical multiscale segmentation statistics of natural images. IEEE Transactions

- on Pattern Analysis and Machine Intelligence 36 (9), 1900–1906.
- Basseville, M., Benveniste, A., Willsky, A. S., 1992. Multiscale representations of Markov random fields. *IEEE Trans. on Signal Processing* 40, 1915–1934.
- Bouman, C. A., Shapiro, M., 1994. A multiscale random field model for Bayesian image segmentation. *IEEE Trans. on Image Processing* 3, 162–177.
- Bruzzone, L., Fernandez, D., Serpico, S., 1999. A neural-statistical approach to multitemporal and multisource remote-sensing image classification. *IEEE Trans. on Geoscience and Remote Sensing* 37, 1350–1359.
- Celeux, G., Chauveau, D., Diebolt, J., 1996. Stochastic versions of the EM algorithm : an experimental study in the mixture case. *Journal of Statistical Computation and Simulation* 55, 287–314.
- Chardin, A., 2000. Modèles énergétiques hiérarchiques pour la résolution des problèmes inverses en analyse d'images application à la télédétection. Thèse de doctorat, Université de Rennes 1.
- Chellappa, R., Jain, A. K., 1993. *Markov Random Fields : Theory and Applications*. Academic Press.
- Chen, Y., Lin, Z., Zhao, X., Wang, G., Gu, Y., June 2014. Deep learning-based classification of hyperspectral data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7 (6), 2094–2107.
- Daubechies, I., 1988. Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics* 41, 909–996.
- Deng, J., Ban, Y., Liu, J., Li, L., Niu, X., Zou, B., 2014. Hierarchical segmentation of multitemporal radarsat-2 sar data using stationary wavelet transform and algebraic multigrid method. *IEEE Transactions on Geoscience and Remote Sensing* 52 (7), 4353–4363.
- Fang, H.-B., Fang, K.-T., Kotz, S., 2002. The meta-elliptical distributions with given marginals. *Journal of Multivariate Analysis* 82 (1), 1–16.
- Ferguson, T. S., 1967. *Mathematical Statistics. A Decision Theoretic Approach*. Academic Press.
- Forestier, G., Wemmert, C., Gançarski, P., 2008. Multisource images analysis using collaborative clustering. *EURASIP Journal on Advances in Signal Processing*, 133.
- Gao, F., Hilker, T., Zhu, X., Anderson, M., Masek, J., Wang, P., Yang, Y., 2015. Fusing landsat and MODIS data for vegetation monitoring. *IEEE Geoscience and Remote Sensing Magazine* 3 (3), 47–60.
- Hammersley, J. M., Clifford, P., 1971. *Markov fields on finite graphs and lattices*. unpublished.
- H.C. Li, W. Hong, Y. W., Fan, P. Z., 2011. On the empirical-statistical modeling of SAR images with generalized Gamma distribution. *IEEE J. Sel. Top. Signal Process* 5, 386–397.
- Hedhli, I., Moser, G., Serpico, S.-B., Zerubia, J., 2014. New cascade model for hierarchical joint classification of multitemporal, multiresolution and multisensor remote sensing data. Dans : *IEEE International Conference on Image Processing (ICIP)*. pp. 5247–5251.
- Hedhli, I., Moser, G., Serpico, S.-B., Zerubia, J., 2015a. New cascade model for hierarchical joint classification of multisensor and multiresolution remote sensing data. Dans : *IEEE International Geoscience and Remote Sensing Symposium*.
- Hedhli, I., Moser, G., Serpico, S.-B., Zerubia, J., 2015b. New hierarchical joint classification method of SAR-optical multiresolution remote sensing data. Dans : *IEEE/EURASIP European Signal Processing Conference*.
- Henderson, F. M., Lewis, A. J., 1998. *Manual of Remote Sensing, Third Edition, Volume 2 : Principles and applications of Imaging RADAR*. John Wiley and Sons Inc.
- Hoberg, T., Rottensteiner, F., Feitosa, R.-Q., Heipke, C., 2015. Conditional random fields for multitemporal and multiscale classification of optical satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing* 53, 659–673.
- Jolion, J. M., Rosenfeld, A., 2004. *A Pyramid Framework for Early Vision : Multiresolutional Computer Vision*. Kluwer Academic Publishers, Norwell, MA, USA.
- Kato, Z., 1994. Modélisations markoviennes multirésolutions en vision par ordinateur. application à la segmentation d'images SPOT. Thèse de doctorat, Université de Nice Sophia Antipolis.
- Kato, Z., Berthod, M., Zerubia, J., 1996. A hierarchical Markov random field model and multi-temperature annealing for parallel image classification. *Computer Vision, Graphics and Image Processing : Graphical Models and Image Processing* 58, 18–37.
- Kato, Z., Zerubia, J., 2012. *Markov Random Fields in Image Segmentation*. NOW publishers, Boston.
- Kurtz, C., Passat, N., Gancarski, P., Puissant, A., 2012. Extraction of complex patterns from multiresolution remote sensing images : A hierarchical top-down methodology. *Pattern Recognition* 45 (2), 685–706.
- Laferté, J., Perez, P., Heitz, F., 2000. Discrete Markov modeling and inference on the quad-tree. *IEEE Trans. on Image Processing* 9, 390–404.
- Li, S. Z., 2009. *Markov Random Field Modeling in Image Analysis*, 3rd Edition. Springer.
- Lillesand, M. L., Kiefer, R. W., Chipman, J. W., 2014. *Remote sensing and image interpretation*. John Wiley & Sons.
- Luetgen, M. R., Karl, W. C., Willsky, A. S., Tenney, R. R., 1993. Multiscale representations of Markov random fields. *IEEE Trans. on Signal Processing* 4, 3377–3396.
- Mallat, S., 2008. *A Wavelet Tour of Signal Processing*, 3rd ed. Academic Press.
- Melgani, F., Serpico, S. B., 2003. A Markov random field approach to spatio-temporal contextual image classification. *IEEE Trans. on Geoscience and Remote Sensing* 41, 2478–2487.
- Motwani, R., Raghavan, P., 1995. *Randomized Algorithms*. Cambridge University Press.
- Moussouris, J., 1974. Gibbs and Markov random system with constraints. *Journal of Statistical Physics* 10 (1), 11–33.
- Mura, M. D., Prasad, S., Pacifici, F., Gamba, P., Chanussot, J., Benediktsson, J. A., Sept 2015. Challenges and opportunities of multimodality and data fusion in remote sensing. *Proceedings of the IEEE* 103 (9), 1585–1601.
- Pérez, P., 1993. Champs markoviens et analyse multiresolution de l'image : application à l'analyse du mouvement. Thèse de doctorat, Université Rennes 1.
- Petitjean, F., Inglada, J., Gancarski, P., 2014. Assessing the quality of temporal high-resolution classifications with low-resolution satellite image time series. *International Journal of Remote Sensing* 35 (7), 2693–2712.
- Prasad, S., Bruce, L. M., Chanussot, J., 2010. *Optical Remote Sensing*. Springer.
- Prendes, J., Chabert, M., Pascal, F., Giros, A., Tourneret, J.-Y., 2015a. A new multivariate statistical model for change detection in images acquired by homogeneous and heterogeneous sensors. *IEEE Transactions on Image Processing* 24, 799–812.
- Prendes, J., Chabert, M., Pascal, F., Giros, A., Tourneret, J.-Y., Jan. 2015b. Performance assessment of a recent change detection method for homogeneous and heterogeneous images. *Revue Française de Photogrammétrie et de Télédétection* 209, 23–29.
- Richards, J. A., 2013. *Remote Sensing Digital Analysis*. Springer.
- Rosanolov, Y., 1982. *Markov Random Fields*. Springer.
- Schlueter, R., Nussbaum-Thom, M., Ney, H., 2012. Does the

- cost function matter in Bayes decision rule ? IEEE Transactions on Pattern Analysis and Machine Intelligence 34, 292–301.
- Storvik, B., Storvik, G., Fjortoft, R., 2009. On the combination of multisensor data using meta-Gaussian distributions. IEEE Trans. Geosci. Remote Sens. 7, 2372–2379.
- Swain, P., 1978. Bayesian classification in a time-varying environment. IEEE Trans. on Systems, Man and Cybernetics 12, 879–883.
- Tupin, F., Inglada, J., Nicolas, J.-M., 2014. Remote Sensing Imagery. John Wiley & Sons.
- Voisin, A., 2012. Classification supervisée d'images d'observation de la terre à haute résolution par utilisation de méthodes markoviennes. Thèse de doctorat, Université de Nice Sophia Antipolis.
- Voisin, A., Krylov, V., Moser, G., Zerubia, J., Serpico, S. B., 2014. Supervised classification of multi-sensor and multiresolution remote sensing images with a hierarchical copula-based approach. IEEE Trans. Geosci. Remote Sens. 2, 3346–3358.
- Winkler, G., 2003. Image Analysis, Random Fields and Markov chain Monte Carlo Methods 2nd edition. Springer.
- Zhao, W., Du, S., Aug 2016. Spectral and spatial feature extraction for hyperspectral image classification : A dimension reduction and deep learning approach. IEEE Transactions on Geoscience and Remote Sensing 54 (8), 4544–4554.