

Approche *a contrario* pour la détection de changements à partir d’images satellite basse résolution

Amandine ROBIN^{1,2}, Lionel MOISAN¹, Sylvie LE HÉGARAT-MASCLE²

¹Laboratoire MAP5, Université Paris-Descartes,
45, rue des Saints-Pères 75006 Paris, France

²Laboratoire CETP, Institut Pierre Simon Laplace
10-12, avenue de l’Europe, 78 140 Vélizy, France

{robin,moislan}@math-info.univ-paris5.fr, mascle@cetp.ipsl.fr

Résumé – Nous considérons le problème du suivi de végétation et de la détection de changements à partir de séquences d’images satellite présentant une forte répétitivité temporelle mais une faible résolution spatiale. Nous dérivons un critère probabiliste *a contrario* mesurant la cohérence entre la séquence d’images “basse résolution” et un état antérieur de référence (classification). La mesure ainsi obtenue est au cœur d’un algorithme stochastique qui sélectionne automatiquement les zones contenant les changements apparus au cours de la période étudiée. Le principe de cette approche est validé d’un point de vue théorique : nous montrons que tout changement, aussi local soit-il, peut être détecté pour peu que le niveau de *contraste* de l’image soit suffisamment élevé. Enfin, nous présentons des expériences sur données simulées, qui affichent des performances prometteuses en vue d’une future application à des données réelles.

Abstract – In this study, we focus on the problem of vegetation monitoring and change detection from remote sensing data. Multitemporal information required by such an application is provided by coarse resolution series presenting high time-repetitiveness. We derive a probabilistic *a contrario* criterion analyzing the coherence between a coarse resolution image time series and a previous knowledge of the land-cover state (through a classification). The measure obtained is the core of a stochastic algorithm that selects automatically the regions with the most likely changes. The principle of this approach is validated from a theoretical viewpoint, since we prove that even very local changes may be detected, provided that the image *contrast* level is high enough. We conclude by showing promising experiments on simulated data.

1 Introduction

Les méthodes automatiques de détection de changements en imagerie satellitaire font l’objet d’un intérêt croissant, notamment en raison des nombreuses applications liées à l’analyse de la surface terrestre ou de l’environnement (suivi de la végétation, mise à jour de cartographies, gestion des risques, etc). Les satellites permettant le suivi des zones d’intérêt présentent généralement soit une forte résolution spatiale, soit une forte résolution temporelle, mais ces deux qualités sont incompatibles, puisque le délai entre deux acquisitions d’une même zone (qui dépend de l’orbite du satellite) est, en première approximation, inversement proportionnel à la largeur de sa trace sur la surface terrestre.

La plupart des méthodes existantes (voir [1, 2] par exemple) sont adaptées à des données haute résolution spatiale (h.r.), typiquement Landsat/TM ou Spot/HRV (pixel de largeur inférieure à 30 m). Elles ne sont donc valables que pour des applications nécessitant très peu d’images par an. Dans le cas du suivi de la végétation, l’évolution temporelle étant l’un des critères des plus discriminants, l’exploitation conjointe de données haute et basse résolution devient nécessaire.

Les premières méthodes de détection de changements à basse résolution (b.r.), c’est-à-dire pour une largeur de pixel allant de 300 m pour les images Meris à 1 km pour les images Spot/VGT ou NOAA/AVHRR, ont été adaptées de méthodes haute résolution (h.r.) [3, 4]. Des travaux plus récents [5] parviennent à

détecter des changements de taille inférieure à celle du pixel b.r. en utilisant des informations *a priori* sur l’état précédent pour contraindre une approche de type *Expectation Maximization* (EM). Ces dernières méthodes atteignent leurs limites dès que le changement considéré concerne une surface de taille inférieure à un certain pourcentage du pixel b.r. et lorsque le nombre de pixels affectés par des changements dépasse une certaine proportion de l’image.

La méthode que nous proposons ici est entièrement automatique et vise à extraire d’une image b.r. le sous-domaine le plus cohérent avec une classification h.r. donnée (les pixels de changement correspondent alors au domaine complémentaire). À partir d’une classification h.r. correspondant à la date t_0 et d’une séquence d’images b.r. de la même région acquise ultérieurement, nous proposons de déterminer dans quel sous-domaine spatial cette classification est encore correcte à un temps $t > t_0$ donné. Pour chaque sous-domaine de l’image b.r., nous mesurons la cohérence entre la classification h.r. et la séquence b.r. observée par le degré de contradiction qu’elle implique sur un modèle naïf très simple. Cette méthode, inspirée de la modélisation *a contrario* introduite en analyse d’images dans [6], permet ainsi de calculer un niveau de significativité sans avoir à quantifier les écarts attendus (bruit, distortions, variabilité intrinsèque, etc.) entre la classification h.r. initiale et les observations b.r. ultérieures.

Dans la section 2, nous commençons par préciser le système de formation des images (lien avec les caractéristiques du sol,

changement de résolution), avant de dériver une expression explicite du critère de détection *a contrario* (nombre de fausses alarmes) et de discuter ses performances théoriques. Dans la section 3, nous décrivons un algorithme stochastique associé à ce critère, avant de commenter les résultats obtenus sur des données simulées (section 4).

2 Modèle de détection

2.1 Modèle d'image

Supposons que les données suivantes sont disponibles :

- une classification h.r. C de la région d'intérêt à la date t_0 . Cette classification associe un label l à chaque pixel du domaine h.r., où $l \in L$ et L représente l'ensemble des types d'occupation du sol possibles.

- une séquence d'images b.r. de la même zone, acquises à des dates ultérieures à t_0 . Par la suite, on considère le cas d'une séquence b.r. réduite à une image unique, notée v , mais la généralisation au cas multi-temporel est immédiate en considérant un vecteur d'images.

Pour commencer, on modélise l'image b.r. en fonction de l'information spatiale haute résolution et des caractéristiques des classes (intensité moyenne). Supposons que chaque pixel h.r. représente une surface dont le type d'occupation est décrit par un label unique (classe *pure*). La valeur moyenne d'un pixel b.r. correspond alors à la somme des moyennes caractéristiques de chaque type d'occupation du sol, pondérée par leur taux d'occupation au sein du pixel. Soit y un pixel appartenant au domaine b.r. Ω et $\alpha_l(y)$ la surface relative occupée par le label l dans le pixel y (par construction, $\sum_{l \in L} \alpha_l(y) = 1$). En notant $e = (e_l)_{l \in L}$, l'intensité moyenne caractéristique de chaque label, une estimation de l'intensité en un pixel b.r. y est donnée par

$$\hat{v}(y) = \sum_{l \in L} \alpha_l(y) e_l. \quad (1)$$

Cette hypothèse est bien connue sous le nom de *modèle linéaire de mélange* et largement utilisée en télédétection [7, 8].

2.2 Détection *a contrario*

La différence entre l'image estimée (\hat{v}) par (1) et observée (v) sur un sous-domaine ω de Ω peut être mesurée, en norme L^2 , par

$$E(\omega, e) = \sum_{y \in \omega} (v(y) - \hat{v}(y))^2. \quad (2)$$

Dans notre problème, on suppose la famille (e) des moyennes caractéristiques des classes inconnue et l'on cherche l'erreur minimale

$$\bar{E}(\omega) = \min_e E(\omega, e). \quad (3)$$

Cette formulation est couramment utilisée comme schéma de régression pour l'estimation de paramètres. Dans l'objectif de décider de l'apparition de changements dans un domaine ω , la principale difficulté est la définition un seuil *a priori* sur $\bar{E}(\omega)$. De plus, pour caractériser le domaine le plus inchangé ω , il est nécessaire de normaliser la valeur de l'erreur en fonction de la taille du domaine considéré. La modélisation *a contrario* offre un moyen d'accomplir ces deux tâches, en formalisant

la détection d'une structure donnée comme un événement très improbable par rapport à une hypothèse naïve sur les données :

H₀ : *Le modèle naïf pour l'image basse résolution est un processus aléatoire de variables aléatoires gaussiennes i.i.d. $\mathcal{N}(\mu, \sigma^2)$ où $\mu \in \mathbb{R}$ et $\sigma > 0$ sont fixés.*

Dans l'esprit de [6], nous pouvons alors définir le *nombre de fausses alarmes (NFA)* associé au sous-domaine $\omega \subset \Omega$ par

$$NFA(\omega, \delta, \sigma, \mu) = \eta(\omega, \Omega) \cdot \mathbb{P}[\bar{E}(\omega) \leq \delta \mid H_0]. \quad (4)$$

Le terme $\mathbb{P}[\bar{E}(\omega) \leq \delta \mid H_0]$ représente la probabilité, pour le domaine ω , d'obtenir une erreur quadratique particulièrement faible dans une image b.r. aléatoire. Les $\eta(\omega, \Omega)$ sont des coefficients de pondération vérifiant $\sum_{\omega \subset \Omega} \eta(\omega, \Omega)^{-1} \leq 1$. Ils permettent de répartir le nombre de fausses alarmes attendu sur l'ensemble des sous-domaines $\omega \subset \Omega$.

Définition 1 (et théorème) *Un sous-domaine $\omega \subset \Omega$ est dit ε -significatif si*

$$NFA(\omega, \bar{E}(\omega), \sigma, \mu) \leq \varepsilon. \quad (5)$$

L'espérance du nombre de domaines ε -significatifs sous l'hypothèse H_0 est inférieure ou égale à ε .

Plus le *NFA* est faible, plus le domaine ω est cohérent avec le modèle de l'image. Classiquement, on choisit $\varepsilon = 1$ pour garantir, en moyenne, au plus une fausse détection.

Le choix le plus simple pour $\eta(\omega, \Omega)$ est $\eta(\omega, \Omega) = 2^{|\Omega|}$, ce qui répartit le risque de fausse détection uniformément sur tous les sous-domaines, mais rend du même coup très improbable la détection par hasard de domaines très petits ($|\omega| \ll |\Omega|$) ou très grands ($|\omega| \approx |\Omega|$). Pour comparer équitablement des domaines de taille différente, il est plus judicieux de répartir le risque par taille, en prenant

$$\eta(\omega, \Omega) = |\Omega| \binom{|\Omega|}{|\omega|}. \quad (6)$$

Avec l'hypothèse *a contrario* donnée par H_0 , le *NFA* peut être calculé explicitement :

Théorème 1 *Soit $P(a, x) = \frac{1}{\Gamma(a)} \int_0^x e^{-t} t^{a-1} dt$, alors*

$$NFA(\omega, \delta, \sigma, \mu) = \eta(\omega, \Omega) \cdot P\left(\frac{|\omega| - |\Omega|}{2}, \frac{\delta}{2\sigma^2}\right). \quad (7)$$

Avec cette nouvelle mesure, on peut maintenant sélectionner le domaine ω le plus *cohérent* avec la classification initiale, en tant que domaine ω minimisant $NFA(\omega, \bar{E}(\omega), \sigma, \mu)$. Remarquons que le nombre de fausses alarmes obtenu est en fait indépendant du choix de μ (nous omettons donc désormais le dernier argument de la fonction *NFA*). En pratique, la variance σ^2 du modèle naïf (H_0) est fixée égale à la variance empirique de l'image b.r. analysée de façon à garantir l'absence de toute détection dans une image de bruit blanc.

2.3 Performances attendues

Le modèle de détection étant maintenant établi, nous proposons ici d'évaluer ses performances théoriques dans le cas du modèle simple pour l'image observée $v = \hat{v} + b$, où \hat{v} représente l'image idéale décrite par (1) et b un processus de variables

aléatoires gaussiennes i.i.d. $\mathcal{N}(0, \sigma_b^2)$ modélisant la variabilité intra-classe. Sous cette hypothèse, la variance de v s'écrit

$$\sigma^2 = \sigma_0^2 + \sigma_b^2, \quad (8)$$

où σ_0^2 représente la variance empirique de \hat{v} (variance inter-classe). Le résidu moyen vaut alors

$$\bar{\delta} = \mathbb{E} [\overline{E}(\omega)] = |\omega| \sigma_b^2,$$

et l'on peut estimer la performance moyenne de la méthode (seuil de détectabilité) grâce à

$$NFA(\omega, \bar{\delta}, \sigma) = |\Omega| \binom{|\Omega|}{|\omega|} \cdot P \left(\frac{|\omega| - |L|}{2}, \frac{|\omega|}{2(\frac{\sigma_0}{\sigma_b})^2 + 2} \right). \quad (9)$$

Ce nombre de fausses alarmes typique dépend de la taille de ω et du rapport $c = \frac{\sigma_0}{\sigma_b}$, appelé *contraste*. L'étude de ses variations en fonction du contraste et de la taille des domaines $|\Omega|$ et $|\omega|$ conduit aux propriétés suivantes :

- le NFA typique étant une fonction décroissante du *contraste* de l'image, plus la valeur du contraste est élevée, plus le domaine ω est détectable ;

- pour un domaine Ω de taille fixée, tout sous-domaine ω peut être détecté dès que le contraste c est suffisamment fort ;

- pour un niveau de contraste $c > 0$ fixé, il existe une proportion critique $p(c) \in [0, 1[$ telle que tout sous-domaine ω de taille $|\omega| > p(c)|\Omega|$ peut être détecté dès que le domaine Ω est suffisamment grand.

3 Algorithme

L'algorithme suivant prend en entrée une classification h.r. et une image b.r. de la même scène, et retourne une estimation du sous-domaine ω le plus cohérent avec la classification (i.e. de NFA minimal). Etant donné la taille des images considérées, une exploration des $2^{|\Omega|}$ sous-domaines n'est pas envisageable. On utilise par conséquent une stratégie de type *Random Sample Consensus* [9], comme dans [10].

```

affecter à  $\sigma^2$  la variance de l'image b.r.
initialiser  $\delta_{min}[]$  et  $NFA_{min}$  à  $+\infty$ 
répéter  $N$  fois
  tirer aléatoirement un ensemble  $I$  de  $|L|$  pixels de  $\Omega$ 
  calculer  $e$  en résolvant (1) pour  $y \in I$  (matrice carrée)
  calculer  $r(y) = (\hat{v}(y) - v(y))^2$  pour  $y \in \Omega$  en utilisant (1)
  trier  $\Omega$  en  $(y_i)_{1 \leq i \leq |\Omega|}$  par  $r(y_i)$  croissants
  poser  $\delta = 0$ 
  pour  $k = |L| + 1$  à  $k = |\Omega|$ 
    poser  $\delta = \delta + r(y_k)$ 
    si  $\delta < \delta_{min}[k]$  alors
      mettre à jour  $\delta_{min}[k]$ 
      si  $NFA(k, \delta, \sigma) < NFA_{min}$  alors
        mettre à jour  $NFA_{min}$  et poser  $\omega = \{y_i\}_{i=1..k}$ 
  fin
fin
fin

```

Le seul paramètre de l'algorithme est le nombre total d'itérations N . Pour les expériences présentées section 4, une bonne convergence de l'algorithme a été obtenue pour $N = 10000$.

4 Résultats

Dans cette section, une illustration de la méthode est présentée pour des images simulées à l'échelle Spot/HRV pour la h.r. (pixel de largeur 20 m) et, pour la b.r., à l'échelle Meris (pixel de largeur 300 m). La classification utilisée (Fig. 1) a été simulée à partir d'un extrait de taille 256×256 d'une image Spot/HRV de manière à présenter une géométrie réaliste pour la haute résolution. Elle résulte du tirage aléatoire d'un label $l \in \{0, 1, \dots, 8\}$ pour chaque région de la segmentation, chaque label représentant une classe de la scène observée (par exemple $l = 0$ pour le blé, $l = 1$ pour l'avoine, etc), à laquelle on associe aléatoirement (et injectivement) un intensité moyenne $e(l) \in \{0, 0.1, \dots, 0.8\}$. L'image h.r. est ensuite simulée en ajoutant à l'image \hat{v} ainsi obtenue un bruit blanc de moyenne nulle et d'écart-type $\sigma_b = 0.05$ pour modéliser la variabilité intra-classe. Deux types de changements sont ensuite simulés (Fig .2.a) : une classe est scindée en deux classes d'intensités moyennes très proches (apparition d'une nouvelle parcelle), et un large disque recouvrant plusieurs parcelles est superposé à l'image (simulation d'une inondation ou d'un incendie par exemple). Pour finir, cette image h.r. a été moyennée par blocs pour produire l'image b.r. Fig .2.b.

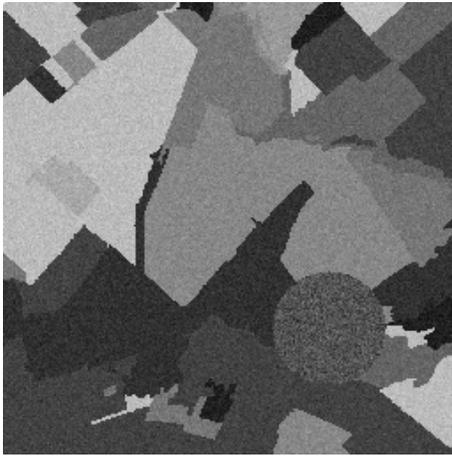
A partir de cette image b.r. et de la classification h.r. (Fig. 1), l'algorithme rejette 27 pixels ($|\Omega| - |\omega| = 27$), représentés en noir Fig. 2c. Les courbes blanches superposées aux domaines de changements correspondent à la trace des changements à h.r. et montrent la bonne correspondance entre les changements détectés et ceux réellement présents. On observe que les pixels b.r. ne contenant que des changements très partiels sont aussi détectés par l'algorithme. Notons aussi que le rejet de ces 27 pixels par une simple analyse de l'erreur quadratique, à la place du NFA , aurait été beaucoup plus délicat (cf. Fig. 3).



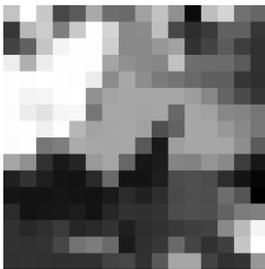
FIG. 1 – Classification h.r. (entrée du programme). Les classes sont numérotées de 0 à 8 comme indiqué sur la légende.

5 Conclusion

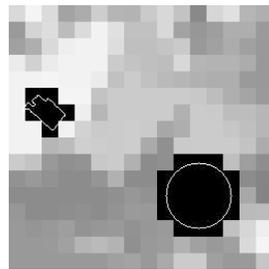
Nous avons proposé un modèle *a contrario* pour détecter des changements de types de couvert en comparant une image basse résolution à une classification haute résolution représen-



(a) Image h.r. simulée avec deux types of changements.



(b) Image b.r. simulée.



(c) Changements détectés.

FIG. 2 – L'image b.r. (b) est obtenue par moyennage de l'image h.r. (a) sur des blocks de 16×16 pixels. Les changements détectés (27 pixels) sont représentés en noir (c). Ils correspondent bien aux changements introduits à haute résolution (dont le contour est représentée en blanc, par superposition).

tant la même scène à une date antérieure. Ce modèle conduit à un algorithme stochastique (RANSAC) générique et complètement automatique, ne nécessitant aucune calibration préalable de paramètres en fonction du type de données observées. Il n'a pas encore été validé sur des données réelles, mais les résultats obtenus sur des données simulées réalistes sont très encourageants.

Les méthodes classiques sont généralement limitées à une faible proportion de changements dans l'image b.r. L'analyse théorique du modèle *a contrario* montre au contraire que quelque soit leur proportion, les changements dans l'image b.r. peuvent être détectés à condition que le *contraste* de l'image soit suffisamment fort. Les premières simulations ont montré que les changements étaient bien détectés même pour des images contenant 80% de changements.

Une étude spécifique ultérieure pourrait analyser la robustesse de la méthode proposée. Elle permettrait une comparaison quantitative avec d'autres méthodes de détection de changements dans le cadre d'applications réelles : suivi de coupes de sylviculture, rotation des cultures sur un site agricole, etc.

Références

[1] J.F. Mas, "Monitoring land-cover changes : a comparison of change detection techniques," *Int. J. Rem. Sens.*, vol. 20, no. 1, pp. 139–152, 1999.

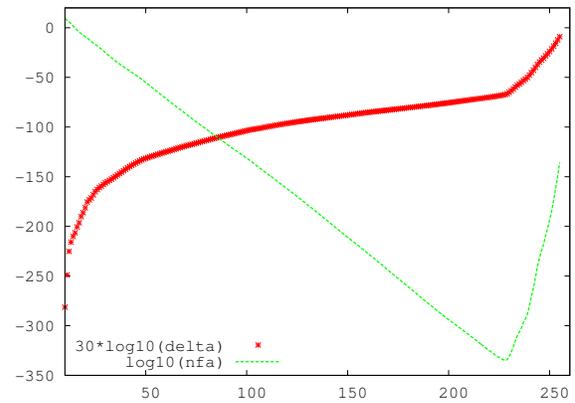


FIG. 3 – Les croix représentent, en fonction de k , l'erreur quadratique minimale obtenue sur tous les sous-domaines ω de taille k (en échelle log). Les pointillés représentent la courbe du NFA correspondant. Alors que la courbe de l'erreur quadratique seule permettrait difficilement de détecter le 27 pixels de changements, la courbe de NFA atteint un minimum marqué quand k correspond au domaine valide ω .

[2] L. Bruzzone and D.F. Prieto, "An adaptive semiparametric and context-based approach to unsupervised change detection in multitemporal remote-sensing images," *IEEE Trans. Geo. Rem. Sens.*, vol. 11, no. 4, pp. 452–466, 2002.

[3] E. Lambin and A. Stralher, "Change vector analysis in multitemporal space : a tool to detect and categorize land-cover change processes using high temporal-resolution satellite data," *Rem. Sens. Environment*, vol. 48, pp. 231–244, 1994.

[4] A. Nielsen, K. Conradsen, and J. Simpson, "Multivariate alteration detection (mad) and maf postprocessing in multispectral, bitemporal image data : New approaches to change detection studies, remote sensing of environment," *Rem. Sens. of Environment*, vol. 64, pp. 1–19, 1998.

[5] S. Le Hégarat, C. Ottlé, and C. Guérin, "Land cover change detection at coarse spatial scales based on iterative estimation and previous state information," *Rem. Sens. of Environment*, vol. 95, no. 4, pp. 464–479, 2005.

[6] A. Desolneux, L. Moisan, and J.M. Morel, "Meaningful alignments," *Int. J. Comp. Vision*, vol. 40, no. 1, pp. 7–23, 2000.

[7] R. Faivre and A. Fischer, "Predicting crop reflectances using satellite data observing mixed pixels," *J. Agric., Bio. Env. Stat.*, vol. 2, pp. 87–107, 1997.

[8] F. Van Der Meer, "Iterative spectral unmixing," *Int. J. Rem. Sens.*, vol. 20, no. 17, pp. 3431–3436, 1999.

[9] M. Fischler and R. Bolles, "Random sample consensus : a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, pp. 381–385, 1981.

[10] L. Moisan and B. Stival, "A probabilistic criterion to detect rigid point matches between two images and estimate the fundamental matrix," *Int. J. Comp. Vision*, vol. 57, no. 3, pp. 201–218, 2004.