

DETECTION DE FEUX DE FORET A PARTIR D'IMAGES SATELLITAIRES IRT PAR ANALYSE STATISTIQUE D'EVENEMENTS RARES

Florent LAFARGE¹, Xavier DESCOMBES¹, Josiane ZERUBIA¹, Sandrine MATHIEU-MARNI²

¹Projet Ariana, commun INRIA - I3S
2004 route des Lucioles - BP93, 06902 Sophia Antipolis, Cedex France

²Alcatel Space
100, boulevard du Midi - BP99, 06156 Cannes La Bocca Cedex France

Florent.Lafarge@inria.fr, Xavier.Descombes@inria.fr, Josiane.Zerubia@inria.fr,
sandrine.marni@space.alcatel.fr

Résumé – Le canal IRT (InfraRouge Thermique) contient des longueurs d'onde particulièrement sensibles à l'émission de chaleur. Les feux de forêt sont alors caractérisés par des pics d'intensité sur ce type d'images. Nous proposons une méthode automatique de détection de feux de forêt fondée sur une analyse statistique des pics d'intensité de l'image. Pour ce faire, nous modélisons dans un premier temps l'image par une réalisation d'un champ gaussien, champ ayant des propriétés particulièrement intéressantes. Les zones de feux, minoritaires et de fortes intensités sont considérées comme des éléments étrangers à ce champ : ce sont des événements rares. Ensuite, par une analyse statistique, nous déterminons un jeu de probabilités définissant, pour une zone donnée de l'image, un degré d'appartenance au champ gaussien, et par complémentarité aux zones potentiellement en feux.

Abstract – The TIR (Thermal InfraRed) channel has a wavelength sensitive to the emission of heat : the higher the heat of an area, the higher the intensity of the corresponding pixel of the image. Then a forest fire can be characterized by a peak intensity in TIR images. We present an automatic classification method based on Gaussian field theory which allows forest fire detection from TIR images. First we model the image as a realization of a Gaussian field. The fire areas, which have high intensity and are supposed to be a minority, are considered as foreign elements of that field : they are rare events. Then we determine by a statistical analysis a set of probabilities which characterizes the degree of belonging to the Gaussian field of a small area of the image. So, we estimate the probability that the considered area is a potential fire.

1 Introduction

Les dégâts résultant des incendies de forêts présentent des conséquences très néfastes sur l'environnement. Aujourd'hui, les espoirs se concentrent sur les technologies spatiales apportées par les satellites. L'enjeu est notamment de détecter de façon précoce les feux de forêt grâce à l'imagerie satellitaire.

Dans ce domaine, les méthodes de détection existantes sont principalement fondées sur des analyses radiométriques d'images thermiques et moyen infrarouge. Des algorithmes tels que le CFDA ("Canadian Fire Detection Algorithm") fondé sur les travaux de Flannigan et Vonder Haar [1], permettent, par des jeux de tests sur des mesures et des seuillages, d'obtenir de bons résultats quant à la détection des feux. Cependant, la plupart de ces méthodes nécessitent des connaissances a priori sur les capteurs employés, les caractéristiques du satellite ou bien encore les spécificités du terrain observé.

Nous proposons une nouvelle méthode de détection de feux de forêt qui n'utilise pas ces connaissances a priori. Cette méthode permet, en effet, de détecter les feux de manière automatique à partir d'images thermiques moyenne résolution. La seule contrainte de cette méthode réside dans l'hypothèse suivante : les feux doivent être des événements rares dans l'image (c'est-à-dire très minoritaires en terme de pixels : cette contrainte signifie, entre autre, que la résolution de l'image ne doit pas être

trop fine). Cette méthode est fondée sur la théorie des champs aléatoires. Nous modélisons les images par des réalisations de champs gaussiens afin d'en extraire, par une analyse statistique, les éléments étrangers pouvant correspondre aux feux de forêt. Une approche comparable a été développée en IRM fonctionnelle par K.J Friston, K.J. Worsley et J.B. Poline. Nous nous sommes appuyés sur leurs travaux en imagerie médicale, notamment sur [2] et [3] pour proposer cette nouvelle technique de détection de feu.

Les images étudiées proviennent du satellite BIRD, fournies par l'Agence Spatiale Allemande (DLR) à Alcatel Space. Elles représentent le Portugal durant le mois de mai 2003 à une résolution de 300 mètres.

2 Méthode proposée

2.1 Champs aléatoires

Dans cette partie, nous rappelons quelques principes de la théorie des champs aléatoires. Nous nous intéressons, en particulier, aux champs gaussiens. Pour des informations complémentaires, le lecteur est invité à consulter [4] et [5].

Définition 1 Soient (Ω, \mathcal{F}, P) un espace probabilisé complet et \mathcal{T} un espace topologique. Un **champ aléatoire** à valeurs

réelles est une application mesurable $X : \Omega \rightarrow \mathbb{R}^T$. Si $\mathcal{T} \subset \mathbb{R}^N, N \in \mathbb{N}$, X est appelée champs aléatoire de dimension N .

De cette définition, nous obtenons un ensemble de mesures F_{t_1, \dots, t_n} définies par $F_{t_1, \dots, t_n}(B) = P((X(t_1), \dots, X(t_n)) \in B)$ pour tout $B \in \mathcal{B}^n$, où \mathcal{B} est la tribu borélienne engendrée par les ouverts de \mathbb{R} . L'ensemble de toutes ces mesures est appelé la famille des \mathcal{F}_d distributions.

Un champ aléatoire est gaussien si ses \mathcal{F}_d distributions sont des vecteurs gaussiens.

Un champ aléatoire $X(t)$ de dimension N est dit **homogène** si $\forall \tau, t_1, \dots, t_k \in \mathbb{R}^N$, le vecteur $(X(t_1), \dots, X(t_k))$ a même loi que le vecteur $(X(t_1 + \tau), \dots, X(t_k + \tau))$.

Définition 2 Soit $F(t) : \mathbb{R}^N \rightarrow \mathbb{R}$ une fonction. $\forall u \in \mathbb{R}$ et $S \subset \mathbb{R}^N$, on définit l'ensemble d'excursion de la fonction F au dessus du seuil u dans S par :

$$A_u(F, S) = \{t \in S : F(t) \geq u\}$$

En terme d'imagerie, A_u représente l'ensemble des "clusters" d'une image au seuil u , c'est-à-dire l'ensemble des surfaces définies, dans l'espace 3D, par l'intersection entre la silhouette de l'image (i.e. la surface dont la cote correspond à l'intensité des pixels de l'image) et le plan de cote u . L'expression (1), établie par Adler [4], permet d'estimer le nombre m_u d'ensembles connexes de l'ensemble d'excursion A_u (i.e. le nombre de "clusters" au seuil u) d'un champ gaussien homogène de dimension 2, centré et réduit.

$$E[m_u] = S(2\pi)^{\frac{3}{2}} |\Lambda|^{\frac{1}{2}} \exp\left(-\frac{u^2}{2}\right) \quad (1)$$

où S représente le nombre de pixels de l'image et Λ , la matrice de covariance des dérivées de l'image.

Pour que cette formule soit applicable, l'image doit vérifier deux conditions principales:

- Les dérivées partielles jusqu'au second ordre du champ aléatoire doivent être continues, ce qui signifie, en terme d'imagerie, que l'image doit avoir une certaine homogénéité dans l'intensité de ses pixels (i.e. un certain lissage) et notamment qu'elle ne doit pas être trop bruitée.
- Pour qu'une image soit modélisable par une réalisation d'un champ gaussien homogène, il faut que la distribution des intensités des pixels de cette image suive une loi normale (cela revient à avoir une distribution normale de l'histogramme).

2.2 Prétraitements

Les données utilisées, c'est-à-dire les images TIR des feux, doivent subir une série de prétraitements afin que l'équation (1) soit applicable. En effet, elles doivent pouvoir être modélisées par des réalisations de champs gaussiens : leur distribution doit être gaussienne.

Les images satellitaires possèdent, bien souvent, plusieurs classes dominantes (mer, terre, nuages,...), ce qui rend leur distribution multimodale. La première étape consiste à rendre la distribution unimodale. Il faut donc séparer les différents modes de l'image pour n'en conserver qu'un : celui le plus proche de la modélisation des feux de forêt, c'est-à-dire celui qui aura l'intensité moyenne la plus élevée. Pour ce faire, nous utilisons

l'algorithme du "K-means" à deux classes de manière itérée. Comme nous souhaitons conserver le mode de l'image qui a la plus forte intensité moyenne, nous allons éliminer, par un principe semblable à la dichotomie, les classes de moyenne d'intensité les plus faibles. A chaque itération, nous séparons l'image en deux classes en utilisant l'algorithme du "K-means" à deux classes et nous gardons la classe de moyenne d'intensité la plus élevée. Nous testons alors si la classe conservée suit une distribution unimodale par un test sur le kurtosis [5].

Nous réalisons ensuite sur la sous-image obtenue une fermeture, puis une ouverture par morphologie mathématique [6], afin d'améliorer sa connexité, c'est-à-dire en faisant disparaître les petits éléments isolés du masque et en rebouchant les trous présents dans le masque. Cette opération nous permet également d'améliorer la normalité de la distribution : nous récupérons, en effet, la queue de la distribution associée à la classe d'intérêt.

La dernière étape consiste à rendre la distribution parfaitement gaussienne par une spécification d'histogramme. Pour plus d'information sur cette opération, le lecteur est invité à consulter [7]. Enfin, nous centrons et réduisons la distribution gaussienne obtenue.

2.3 Etude statistique sur l'appartenance d'un "cluster" à un champ gaussien

Les feux de forêt correspondent aux pixels de très forte intensité. Ainsi, suite aux prétraitements de la partie précédente, il est légitime de considérer l'image comme l'union d'une réalisation d'un champ gaussien et d'un ensemble de points étranger à ce champ : les feux de forêt. La réalisation du champ gaussien est un ensemble de pixels très majoritaire, que nous pouvons considérer comme le "fond de l'image". Les feux de forêt sont, eux, représentés par des ensembles de pixels de taille relativement petite par rapport aux dimensions de l'image; ils sont surtout très rares (hypothèse \mathcal{H}_0).

Nous considérons, par ailleurs, que ces pixels rares ayant une forte intensité ne peuvent représenter que des feux de forêt, cela signifie que les seuls événements naturels ou artificiels pouvant expliquer une intensité thermique très élevée sur la surface terrestre imagée sont forcément un feu de forêt (hypothèse \mathcal{H}_1). Nous allons, dans ce cas, pouvoir utiliser les résultats de la théorie des champs aléatoires, et particulièrement l'équation (1) afin d'établir des lois probabilistes sur l'appartenance d'un cluster à un champ gaussien. Nous pourrions alors estimer si un ensemble connexe de pixels appartient au fond de l'image ou, par négation, appartient à l'ensemble des pixels représentant les feux de forêt.

2.3.1 Pics d'intensité

Dans un premier temps, nous estimons l'appartenance d'un "cluster" à une réalisation d'un champ gaussien en fonction de l'intensité maximale des pixels de ce "cluster".

La probabilité qu'un "cluster", pris au seuil u et ayant un pic d'intensité égal à $u + H_0$, appartienne à la réalisation d'un champ gaussien G est estimée par le quotient de l'espérance du nombre de "clusters" au seuil $u + H_0$ divisée par l'espérance du nombre de "clusters" au seuil u , espérances données par

l'équation (1). En notant $C_u^{H_0}$, un "cluster" pris au seuil u et ayant un pic d'intensité égal à $u + H_0$, nous avons

$$P(C_u^{H_0} \in G) = \frac{E[m_{u+H_0}]}{E[m_u]} = \left(1 + \frac{H_0}{u}\right) e^{-(uH_0 + \frac{H_0^2}{2})} \quad (2)$$

Nous notons P_H cette probabilité. Il est intéressant d'observer le comportement de P_H en fonction du seuil u connaissant la valeur du pic d'intensité $u + H_0$: cette fonction admet un minimum unique. Le seuil u^* correspondant permettra de fixer un seuil de référence u_0 que l'on utilisera ultérieurement.

2.3.2 Extension spatiale

Dans un deuxième temps, nous estimons l'appartenance d'un "cluster" à une réalisation d'un champ gaussien en fonction de la surface de ce "cluster".

Soient N_u le nombre de pixels d'intensité supérieure à u et n_u , le nombre de pixels d'un "cluster" au seuil u .

Nous cherchons à estimer la probabilité qu'un "cluster" pris à un seuil u ait une certaine surface n_0 , c'est-à-dire à établir une loi sur n_u .

Nosko a établi expérimentalement que n_u suivait une distribution exponentielle [8]. Il nous suffit donc d'estimer son paramètre, c'est-à-dire d'estimer l'inverse de l'espérance de n_u . On utilise pour cela les lois de N_u et m_u à travers l'expression suivante :

$$E[n_u] = \frac{E[N_u]}{E[m_u]} \quad (3)$$

La distribution des pixels suivant une loi normale, l'espérance de N_u est donnée par :

$$E[N_u] = S \int_u^\infty (2\pi)^{-\frac{1}{2}} e^{-\frac{x^2}{2}} dx = S \cdot \Phi(-u) \quad (4)$$

L'espérance de m_u est, elle, donnée par l'équation (1). Nous en déduisons une expression de l'espérance de n_u par :

$$E[n_u] = \frac{E[N_u]}{E[m_u]} = \frac{\Phi(-u)}{(2\pi)^{\frac{3}{2}} |\Lambda|^{\frac{1}{2}} \exp-\frac{u^2}{2}} \quad (5)$$

De la même façon que dans la partie précédente où nous connaissions la valeur du pic d'intensité des "clusters", nous allons pouvoir, à partir de la loi de n_u , estimer la probabilité d'appartenance d'un "cluster" $C_u^{S_0}$ pris au seuil u et ayant une surface S_0 à une réalisation d'un champ gaussien G . Nous avons en effet :

$$P(C_u^{S_0} \in G) = P(n_u \geq S_0) = e^{-\frac{S_0}{E[n_u]}} = e^{-\frac{(2\pi)^{\frac{3}{2}} |\Lambda|^{\frac{1}{2}} S_0 \exp-\frac{u^2}{2}}{\Phi(-u)}} \quad (6)$$

Nous appelons P_S cette probabilité.

2.3.3 Seuillage contextuel

Les équations (2) et (6) permettent d'estimer la probabilité d'appartenance d'un "cluster" à une réalisation d'un champ gaussien de deux manières différentes. Il s'agit alors d'utiliser ces deux estimations afin de dégager une loi de décision qui soit pertinente et robuste.

Cette loi de décision consiste à *accepter le "cluster" comme élément étranger à une réalisation du champ gaussien* si la probabilité P_H est inférieure à une certaine probabilité limite p_l ou

si P_H est supérieure à p_l et P_S est inférieure à p_l .

La probabilité limite p_l s'apparente alors à une "p-value", c'est-à-dire à la probabilité de commettre une erreur dans notre estimation. Le résultat sera donc d'autant plus robuste que p_l sera faible. Par cette loi de décision, nous testons la non-appartenance d'un "cluster" à un champ gaussien par la probabilité P_H ; si le test est rejeté, nous accordons une deuxième chance en testant cette fois-ci par la probabilité P_S .

Cette loi de décision permet de réaliser des analyses locales sur les "clusters". Il nous faut maintenant proposer un algorithme général sur toute une image. Pour ce faire, nous définissons une famille de seuils $U = (u_{-n}, \dots, u_0, \dots, u_m)$, fixée relativement au seuil de référence u_0 , décrit dans la partie 2.3.1. L'algorithme général consiste alors à :

1. fixer la probabilité limite p_l de la loi de décision
2. établir l'ensemble des "clusters" pris au seuil u_0
3. **pour chacun de ces "clusters"**,
 - étendre le "cluster" à une famille de "clusters" relativement à l'ensemble de seuils U .
 - pour chaque élément de cette famille, établir P_S et P_H
4. garder les "clusters" acceptés par la loi de décision.

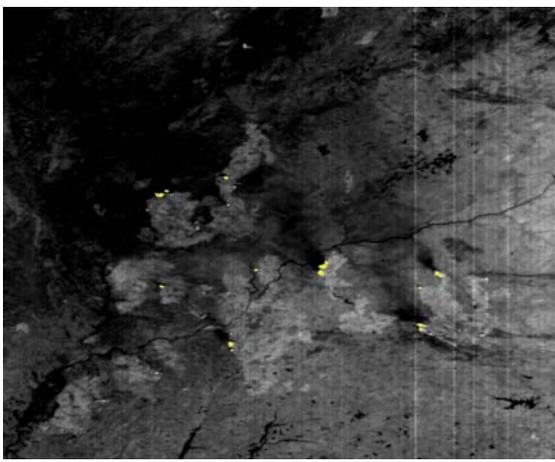
3 Résultats

Les "clusters" acceptés ont été répertoriés en trois catégories : en jaune, les "clusters" acceptés avec une probabilité limite $p_l = 0.01$; en bleu, $p_l = 0.05$ et enfin en rouge, $p_l = 0.1$. Cela signifie que les zones en jaune, par exemple, n'appartiennent pas au "fond de l'image" et ce avec un degré de certitude de 99%. En tenant compte de l'hypothèse (\mathcal{H}_1), nous pouvons affirmer que ces zones sont des feux de forêt à 99%.

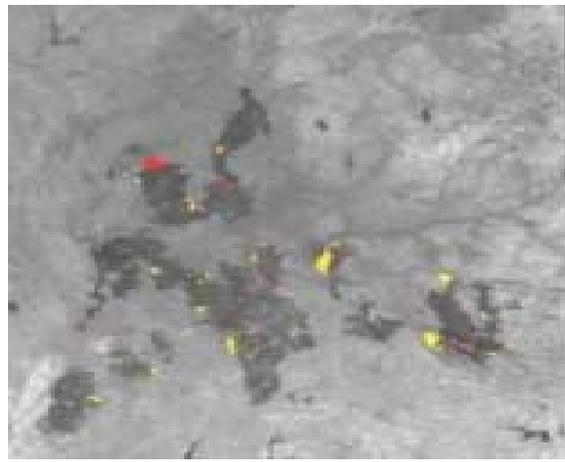
Les résultats obtenus sont très satisfaisants (figure 1). Le taux de détection des feux est supérieur à 85% et le taux de fausses alarmes est nul avec une probabilité limite $p_l = 0.05$. Sur la figure 1-a, le résultat obtenu apparaît très correct comparativement à la vérité de terrain associée, fournie par Alcatel Space. En effet, cette dernière fait état de 14 feux répartis dans le centre du Portugal. Sur ces 14 feux, nous en détectons 12 (11 feux détectés avec $p_l = 0.01$ + 1 feu détecté avec $p_l = 0.05$). Par ailleurs, l'algorithme n'a pas détecté de feux qui n'existent pas sur la vérité de terrain : le nombre de fausses alarmes est nul. Enfin, en terme de temps de calcul, l'algorithme a mis 136 secondes sur un P4-2Ghz pour une image 4000×1000 pixels.

4 Conclusion

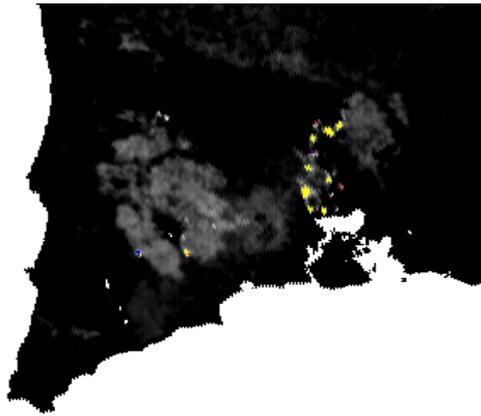
La méthode présentée permet d'extraire, de manière automatique, les feux de forêt sur des images satellitaires IRT. Elle fournit de très bons résultats concernant aussi bien le taux de détection des feux que le taux de fausses alarmes. Le principal avantage réside dans l'utilisation d'une probabilité limite p_l , agissant comme un coefficient de confiance sur le résultat. Il est envisagé de compléter, dans le futur, cette méthode par une étude topographique des pics d'intensité, dans le but de distinguer les zones en feux et les zones brûlées encore chaudes. Cette distinction nous permettrait de déterminer le sens de propagation des feux de forêt.



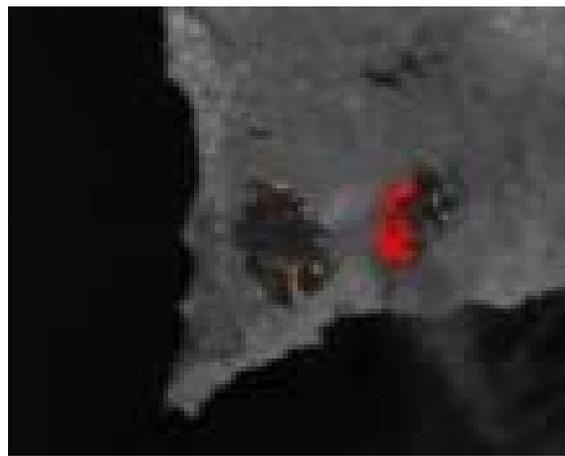
(a)



(b)



(c)



(d)

FIG. 1: (a) : Extrait du résultat obtenu sur l'image BIRD du 04/05/2003 en canal thermique. (c) : Extrait du résultat obtenu sur l'image BIRD du 14/05/2003 en canal thermique. (b) et (d) : Vérités de terrain associées

Remerciements

Les trois premiers auteurs remercient Alcatel Space pour le soutien financier de ce travail, ainsi que l'Agence Spatiale Allemande (DLR) pour la fourniture des données BIRD.

Références

- [1] M. Flannigan et T. Vonder Haar, « Forest fire monitoring using NOAA satellite AVHRR », *Canadian Journal of Forest Research*, vol. 16, 1986, p. 975–982.
- [2] K.J. Friston, K.J. Worsley, R.S.J Frackowiak, J.C Mazziotta, et A.C. Evans, « Assessing the significance of focal activations using their spatial extent », *Hum. Brain Map.*, vol. 1, 1994, p. 214–220.
- [3] J-B. Poline, K.J. Worsley, A.C. Evans, et K.J. Friston, « Combining spatial extent and peak intensity to test for activations in functional imaging », *Neuroimage*, vol. 5, 1997, p. 83–96.
- [4] R.J. Adler, *The Geometry of Random Fields*. Wiley, 1981.
- [5] F. Lafarge, X. Descombes, et J. Zerubia, « Détection de feux de forêt par analyse statistique de la radiométrie d'images satellitaires ». Rapport de recherche 5369, INRIA, nov 2004.
- [6] P. Soille, *Morphological Image Analysis*. Springer-Verlag, Heidelberg, 2nd édition, 2003.
- [7] J.P. Cocquerez et S. Philipp, éditeurs, *Analyse d'images : filtrage et segmentation*. Masson, 1995.
- [8] V.P. Nosko, « Local structure of Gaussian random fields in the vicinity of high level shines », *Soviet Mathematics : Doklady*, vol. 10, 1969, p. 1481–1484.