



Approche non supervisée par processus ponctuels marqués pour l'extraction d'objets à partir d'images aériennes et satellitaires

Saima Ben Hadj, Florent Chatelain, Xavier Descombes, Josiane Zerubia

► To cite this version:

Saima Ben Hadj, Florent Chatelain, Xavier Descombes, Josiane Zerubia. Approche non supervisée par processus ponctuels marqués pour l'extraction d'objets à partir d'images aériennes et satellitaires. *Revue Française de Photogrammétrie et de Télédétection*, Société Française de Photogrammétrie et de Télédétection, 2010, 194, pp.2-15. hal-00638665

HAL Id: hal-00638665

<https://hal.inria.fr/hal-00638665>

Submitted on 8 Nov 2011

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.

APPROCHE NON SUPERVISEE PAR PROCESSUS PONCTUELS MARQUES POUR L'EXTRACTION D'OBJETS A PARTIR D'IMAGES AERIENNES ET SATELLITAIRES

Saima Ben Hadj, Florent Chatelain, Xavier Descombes, Josiane Zerubia

Equipe-Projet Ariana, INRIA Sophia Antipolis – Méditerranée/I3S

2004, route des Lucioles, BP 93

06902 SOPHIA ANTIPOLIS cedex - France

Tel: 04 92 38 78 65 - Fax : 04 92 38 76 43

E-mail: Saima.Ben_hadj@inria.fr, florent.chatelain@gipsa-lab.inpg.fr, Xavier.Descombes@inria.fr, Josiane.Zerubia@inria.fr

Résumé

Les modèles de processus ponctuels marqués ont été précédemment appliqués avec succès pour résoudre des problèmes d'extraction de réseaux de formes à partir d'images aériennes et satellitaires haute résolution. L'avantage de ces modèles est qu'ils prennent en compte la géométrie des objets à extraire. En particulier, les objets de l'image sont représentés par un ensemble de formes géométriques dont la disposition est gouvernée par deux types d'énergies : une énergie d'attache aux données qui lie les objets à l'image traitée, et une énergie de régularisation qui permet d'introduire des connaissances a priori sur le réseau d'objets. Par ailleurs, des paramètres qui traduisent l'influence de ces énergies sont introduits dans la définition de ce modèle. Afin d'assurer une extraction automatique des objets, une méthode d'estimation des paramètres en question, fondée sur une version stochastique de l'algorithme Espérance-Maximisation, a été étudiée et a conduit à des résultats prometteurs sur un exemple simple de processus ponctuels d'objets circulaires. Nous proposons de prolonger cette étude afin d'extraire des formes géométriques plus générales, i.e. des formes elliptiques et rectangulaires. Nous abordons ainsi plusieurs applications ; à savoir l'extraction de flamants roses, de houppiers d'arbres, de bateaux dans un port maritime, de la trace au sol du bâti ainsi que la détection de tentes de réfugiés. L'originalité majeure de ce travail réside dans la définition des composantes de l'énergie qui permettent de modéliser l'alignement des navires à quai, ainsi que les interactions qui existent entre les bâtis tout en restant dans un cadre automatique.

Keywords: processus ponctuel marqué, RJMCMC, recuit simulé, SEM, pseudo-vraisemblance, extraction d'objet.

Abstract

Previously, marked point process models have been successfully applied to solve problems of feature network extraction from aerial and satellite high resolution photographs. The main advantage of these models is that they take into account object geometry. In particular, objects in the image are represented by a set of geometric shapes which are governed by two types of energy: a data energy term that links objects to the processed image, and a regularizing energy term which introduces some prior knowledge about the observed scene. Furthermore, some parameters reflecting the importance of these energies are involved in the definition of this model. Thus, to ensure automatic object extraction, an estimation method of these parameters, based on the Stochastic Expectation-Maximization algorithm was studied and showed promising results on a simple example of a marked point process where the considered objects are disks. We propose to extend this study to extract more general geometric shapes, such as ellipses and rectangles. Hence, we deal with several applications, namely flamingo extraction, tree crown extraction, detection of boats in a seaport, building footprint detection and refugee tent detection. The main originality of this work consists in developing new energy components that allow to model boat alignment and building interactions within an unsupervised framework.

Keywords: marked point process, RJMCMC, simulated annealing, SEM, pseudo-likelihood, object extraction.

1. Introduction

L'analyse des données en télédétection est devenue l'ambition de nombreuses communautés : publiques, environnementales, industrielles et aussi militaires. L'extraction de l'information à partir de l'imagerie aérienne et satellitaire représente alors un enjeu important. Répondre à un tel problème, et plus particulièrement extraire les objets de la scène à partir d'images de télédétection, nécessite la mise en œuvre de méthodes spécifiques adaptées aux propriétés de haute, voire très

haute résolution des images. Par exemple, la résolution des images aériennes est passée de plusieurs dizaines de centimètres à quelques centimètres. Par ailleurs, les satellites d'observation les plus récents ont une résolution de l'ordre du mètre. Ainsi, la résolution des images de SPOT 5 en supermode est de 2,5 m, la résolution des images issues du satellite IKONOS est de l'ordre de 1 m et celle du satellite PLEIADES sera de 0,7 m. Afin de mieux prendre en compte les propriétés de ces images, des modèles de processus ponctuels marqués ont été récemment introduits et ont prouvé leur pertinence en télédétection pour l'extraction d'objets [Chatelain et al,

2009a, Perrin et al., 2005]. L'avantage majeur de ces modèles est qu'ils permettent d'exploiter la géométrie des objets à extraire ainsi que leurs interactions spatiales. En effet, dans le cadre des processus ponctuels marqués, les objets de l'image sont vus comme un ensemble de points, de lignes ou de surfaces. Par ailleurs, ces processus sont régis par deux types d'énergies : une énergie d'attache aux données qui lie les objets à l'image traitée, et une énergie de régularisation qui contrôle la répartition des objets sur la scène en fonction de nos a priori sur la scène recherchée.

Néanmoins, cette modélisation fait intervenir certains paramètres (facteurs de pondération des fonctions d'énergie) dont la valeur adéquate dépend entièrement de l'image traitée. De ce fait, la nécessité de calibrer ces paramètres est une limite à l'utilisation intensive et automatique de ce type de modèle. Afin d'obtenir une extraction des données non-supervisée, une étude des méthodes d'estimation des paramètres en question a été menée. Une méthode fondée sur une version stochastique de l'algorithme « Espérance-Maximisation » («Stochastic Expectation-Maximization algorithm» en anglais) [Celeux et al., 1996] a été étudiée. Cette méthode d'estimation a prouvé sa pertinence pour l'estimation des paramètres d'un modèle simple où les objets sont des disques [Chatelain et al., 2009a].

Nous prolongeons ici cette approche d'extraction d'objets à des modèles de formes géométriques plus générales telles que les ellipses et les rectangles. Nous abordons ainsi plusieurs applications. Tout d'abord, nous envisageons un modèle de processus d'ellipses pour résoudre des problèmes d'extraction de flamants roses, de houppiers (couronnes des arbres) dans une forêt ainsi que le décompte des navires dans un port. En plus de cette généralisation, notre contribution principale réside dans la définition d'un nouveau terme d'énergie pour la détection des bateaux, et l'extraction de la trace au sol du bâti. Ensuite, nous définissons un modèle d'objets rectangulaire afin d'effectuer la détection de tentes de réfugiés dans un campement d'urgence après un séisme, ainsi que l'extraction de la trace au sol du bâti dans un milieu urbain. Nous modifions l'énergie du modèle pour chacune de ces applications. Notamment, nous proposons une nouvelle énergie qui reflète l'alignement des bâtiments dans les zones urbaines.

Pour ce faire, nous consacrons la première partie de cet article à présenter la méthode non supervisée proposée pour l'extraction d'objets. Ensuite, nous évoquons, dans la deuxième partie, des applications de cette méthode permettant l'extraction d'objets elliptiques. Dans la troisième partie, nous présentons des applications permettant l'extraction d'objets rectangulaires. Pour finir, nous proposons certaines perspectives permettant d'étendre ce travail de recherche.

2. Méthode non supervisée pour l'extraction d'objets

Les modèles de processus ponctuels marqués ont reçus une attention soutenue en statistiques spatiales [Møller, Waagepetersen, 2004] et ont été appliqués avec succès en analyse d'image. Nous reprenons un modèle particulier qui a été développé dans [Chatelain et al., 2009a] et qui a été essentiellement appliqué en vue d'extraire des objets circulaires. L'avantage de ce

modèle est la simplicité de sa fonction d'énergie et l'existence d'une interprétation physique des paramètres associés. Afin d'automatiser l'extraction des objets, nous nous focalisons ensuite sur la méthode d'estimation des paramètres du modèle.

2.1. Modèle de processus ponctuels marqué

Dans notre modélisation de processus ponctuels marqués, nous définissons un objet $u = (p, m)$ par sa position $p = (x, y)$ dans le plan de l'image et par un ensemble de paramètres m qui décrit la forme géométrique de l'objet. Soit P un sous-espace de \mathbb{R}^2 qui désigne l'espace des positions des objets, et M un sous-espace de \mathbb{R}^k qui désigne l'espace des marques. Une configuration d'objets appartenant à l'espace d'objets $W = P \times M$ est une liste non-ordonnée d'objets telle que :

$$x = \{ u_1, u_2, u_3, \dots \} \in \Omega_n, u_i = (p_i, m_i) \in W$$

Un processus ponctuel marqué ou processus d'objets défini sur $W = P \times M$ est une variable aléatoire dont les réalisations sont des configurations d'objets de $\Omega = \bigcup_n \Omega_n$. Dans la suite, nous nous intéressons à deux cas particuliers de processus correspondants à différentes formes d'objets. Le premier est un processus d'ellipses que nous exploitons pour extraire les flamants roses, les houppiers d'arbres et les bateaux. Ce processus est défini sur l'espace d'objets suivant :

$$W = [0, X_{max}[\times [0, Y_{max}[\times [a_{min}, a_{max}] \times [a_{min}, a_{max}] \times [0, \pi[$$

avec $P = [0, X_{max}[\times [0, Y_{max}[$ l'espace des positions données par les centres des ellipses, limité par la dimension de l'image X_{max} et Y_{max} . $M = [a_{min}, a_{max}] \times [a_{min}, a_{max}] \times [0, \pi[$ décrit l'espace des marques d'une ellipse. En effet, une ellipse est définie par son demi-petit-axe a , son demi-grand-axe b et son orientation ω . Cette paramétrisation est illustrée par la Figure 1 (gauche) :

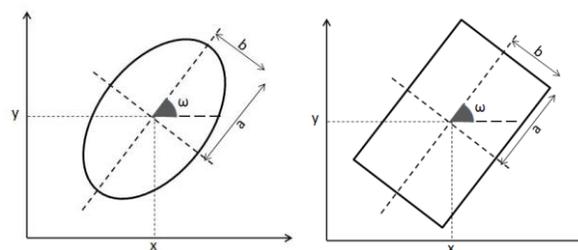


Figure 1: gauche : paramétrisation d'une ellipse, droite : Paramétrisation d'un rectangle

Le deuxième processus auquel nous nous intéressons dans ce papier est un processus de rectangle. Un rectangle est également donné par la position de son centre sur l'image $p \in [0, X_{max}[\times [0, Y_{max}[$, par sa demi-largeur $l \in [l_{min}, l_{max}]$, sa demi-longueur $L \in [L_{min}, L_{max}]$, et son orientation $\omega \in [0, \pi[$. La Figure 1 (droite) décrit une telle paramétrisation. Un des processus ponctuels les plus simples est le processus de Poisson qui décrit une structure spatiale des objets complètement aléatoire (les

objets sont indépendants les uns des autres). D'une manière générale, un processus ponctuel marqué est défini par une mesure de densité de probabilité par rapport à une mesure de Poisson de référence. Dans la suite, nous nous intéressons à un processus particulier, dit processus de Gibbs, défini par une densité de probabilité s'exprimant sous la forme énergétique suivante :

$$f_{\theta}(x, y) = \frac{\exp(-U_{\theta}(x, y))}{c(\theta)} \quad (1)$$

où $U_{\theta}(x, y)$ est l'énergie du modèle avec y la radiométrie de l'image et x la configuration d'objets. $c(\theta) = \int \exp(-U_{\theta}(x, y)) \cdot \mu(dx) \cdot \nu(dy)$ est la constante de normalisation du modèle considéré où $\mu(\cdot)$ désigne la mesure de probabilité du processus de Poisson de référence, $\nu(\cdot)$ définit la mesure de probabilité associée aux observations et θ est le vecteur de paramètres du modèle permettant l'adaptabilité du modèle à différentes images. Vu le nombre important de configurations possibles, il est généralement difficile de calculer la fonction de normalisation $c(\theta)$.

La fonction d'énergie incluse dans le modèle proposé permet de simuler les interactions entre les objets, à savoir une attraction ou une répulsion. En effet, l'énergie $U_{\theta}(x, y)$ est la somme de deux types d'énergies : une énergie d'attache aux données (également appelée «énergie externe») $U_{\theta}^d(x, y)$ qui reflète l'adéquation de la configuration x aux données y et une énergie a priori (ou une énergie interne) $U_{\theta}^p(x)$ qui traduit nos connaissances a priori sur la structure de la configuration d'objets de l'image. Ainsi, la configuration d'objets x qui permet l'extraction des objets de l'image est celle qui maximise la vraisemblance jointe $f_{\theta}(x, y)$ i.e. la configuration x qui minimise l'énergie totale $U_{\theta}(x, y)$:

$$x = \text{Arg Max}_x f_{\theta}(x, y) = \text{Arg Min}_x U_{\theta}(x, y)$$

L'optimisation de cette fonction d'énergie est réalisable grâce à l'algorithme de recuit simulé [Azencott, 1992, Van Laarhoven, Aarts, 1987].

2.1.1. Energie externe

Deux approches de calcul de l'énergie externe sont décrites dans le rapport de recherche [Ben Hadj et al., 2010b]. La première, dite bayésienne, est calculée à partir de la vraisemblance $f_{\theta}(y/x)$ des données y pour une configuration d'objets x connue. Elle ne repose pas, ou peu, sur les propriétés morphologiques des objets. Elle peut aboutir à plusieurs sur-détections, notamment lorsque l'image traitée contient des formes variées. Nous considérons ici une approche dite détecteur, qui consiste à définir une énergie locale $U_{\theta}^d(u, y)$ au niveau de l'objet u de la configuration x . C'est un terme similaire à un détecteur d'objets. En effet, l'énergie d'attache d'un objet u est calculée à partir d'une mesure de la distance radiométrique que l'on note $d(u, F^{\rho}(u))$, entre l'ensemble des pixels de l'objet u et celui de sa frontière $F^{\rho}(u)$ (où ρ désigne la largeur de cette frontière, souvent fixée à 1 ou 2 pixels, en pratique). Pour ce faire, nous considérons

l'expression de la distance suivante qui est similaire à la distance de Bhattacharya :

$$d(u, F^{\rho}(u)) = \frac{(\mu_u - \mu_{F^{\rho}(u)})^2}{4\sqrt{(\sigma_u^2 + \sigma_{F^{\rho}(u)}^2)}} - \frac{1}{2} \log \left(\frac{2 \cdot \sigma_u \cdot \sigma_{F^{\rho}(u)}}{\sigma_u^2 + \sigma_{F^{\rho}(u)}^2} \right) \quad (2)$$

où (μ_u, σ_u) et $(\mu_{F^{\rho}(u)}, \sigma_{F^{\rho}(u)})$ sont respectivement la moyenne et la variance des pixels appartenant à u et celles qui correspondent aux pixels appartenant à sa frontière $F^{\rho}(u)$. En effet, pour nos applications, les objets à détecter sont supposés entourés par une zone homogène, plus (ou moins) sombre que l'objet, i.e. de radiométrie inférieure (ou supérieure) à celle de l'objet. L'énergie d'attache aux données d'un objet u est ainsi construite comme étant une qualification de la distance $d(u, F^{\rho}(u))$ mesurée à travers la fonction $Q(\cdot)$:

$$U_{\theta}^d(u, y) = Q \left(\frac{d(u, F^{\rho}(u))}{d_0} \right)$$

avec

$$Q(x) = \begin{cases} 1 - x^{\frac{1}{3}} & \text{si } x < 1 \\ \exp\left(-\frac{x-1}{3}\right) - 1 & \text{si } x \geq 1 \end{cases}$$

La fonction $Q(\cdot)$ permet de favoriser les objets bien placés pour lesquels la distance $d(u, F^{\rho}(u))$ est supérieure à un seuil d_0 en leur attribuant une valeur négative de l'énergie, et elle pénalise les objets mal placés (distance $d(u, F^{\rho}(u))$ inférieure à d_0) en leur attribuant une valeur positive. La Figure 2 permet d'illustrer le comportement d'une telle fonction $Q(\cdot)$. L'énergie d'attache de la configuration x aux données y est alors proportionnelle à la somme des énergies locales définies au niveau de chaque $u \in x$:

$$U_{\theta}^d(x, y) = \gamma_d \sum_{u \in x} U_{\theta}^d(u, y)$$

Le paramètre $\gamma_d > 0$ représentant le poids de l'énergie externe est crucial dans la décision sur la présence des objets.

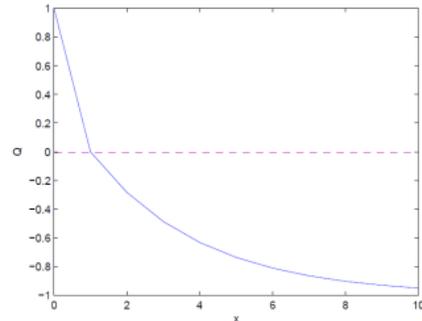


Figure 2: Comportement de la fonction $Q(x)$.

2.1.2. Energie interne

Ce terme a priori formalise nos connaissances a priori sur les interactions qui existent entre les objets. Dans un premier temps, nous considérons les interactions de type recouvrement entre les objets. En d'autres termes, nous pénalisons les paires d'objets ayant une grande proportion de recouvrement, afin de ne pas détecter le même objet plusieurs fois. Nous proposons plus loin d'autres formes d'énergie interne en fonction de l'application traitée. Pour ce faire, nous avons eu recours à un processus de type "Hard Core" qui présente l'avantage d'avoir une fonction simple d'énergie. Ce processus attribue une énergie infinie (donc une probabilité nulle) aux paires d'objets qui dépassent le taux de recouvrement toléré s et une énergie nulle aux paires d'objets qui le respectent. La densité non normalisée d'un tel processus s'écrit comme suit :

$$h_{\theta}(x) = \beta^{n(x)} \exp \left(- \sum_{1 < i < j < n(x)} U_s(x_i, x_j) \right)$$

β est le paramètre d'activité du processus, $n(x)$ est le nombre d'objets dans la configuration x et $U_s(x_i, x_j)$ est l'énergie de recouvrement associée aux paires d'objets (x_i, x_j) , qui s'exprime comme suit dans le cadre d'un processus Hard Core :

$$U_s(x_i, x_j) = \begin{cases} 0 & \text{si } p(x_i, x_j) < s \\ +\infty & \text{sinon} \end{cases}$$

avec $p(x_i, x_j)$ le taux de recouvrement entre les objets x_i et x_j . Ce taux peut être calculé comme étant l'aire commune des objets x_i et x_j , divisée par l'aire maximale de ces objets. Le seuil s représente le taux de recouvrement autorisé entre deux objets.

2.2. Estimation des paramètres du modèle

Dans le cas des données complètes, c'est-à-dire lorsque la configuration d'objets x est connue, des méthodes d'estimation telles que le Maximum de Vraisemblance par méthodes de Monte Carlo par Chaîne de Markov (MVMCMC), et le Maximum de Pseudo Vraisemblance (MPV) ont été étudiées et comparées [Ben Hadj et al., 2010b]. Dans le cas des données incomplètes, c'est-à-dire lorsque seule l'image y est donnée, la configuration d'objets x étant inconnue, une méthode fondée sur la version stochastique de l'algorithme Espérance-Maximisation (SEM) a été proposée [Chatelain et al., 2009b]. Dans cette méthode, nous cherchons à estimer θ par la maximisation du logarithme de la vraisemblance jointe de la configuration x et des données y . Comme la configuration x est inconnue, il n'est pas possible de maximiser directement la vraisemblance jointe $f_{\theta}(x, y)$. Nous proposons alors de maximiser son espérance $E[f_{\theta}(x, y)]$ par un algorithme Espérance-Maximisation. Cependant, le calcul de cette espérance est numériquement difficile car l'espace des configurations d'objets Ω est très grand. Une variante de l'algorithme EM dite Monte-Carlo EM permet de surmonter ce problème en approchant l'espérance du logarithme de la vraisemblance jointe par une méthode de Monte Carlo.

Ceci revient à simuler n configurations d'objets $x^{(1)}, x^{(2)}, \dots, x^{(n)}$ suivant la loi conditionnelle aux observations $f_{\theta}(x/y)$ puis à approcher l'espérance $E[\log f_{\theta}(x, y) / y]$ par une moyenne empirique de la log-vraisemblance étendue :

$$E[\log f_{\theta}(x, y) / y] = \frac{1}{n} \sum_{i=1}^n \log f_{\theta}(x^{(i)}, y)$$

La simulation de n échantillons est coûteuse du point de vue temps de calcul. C'est pourquoi, nous proposons d'utiliser la version stochastique de l'algorithme EM [Celeux et al., 1996] qui consiste à ne considérer qu'un seul échantillon tiré conditionnellement aux données y selon le processus de paramètre θ . Le principe de cet algorithme consiste à itérer les trois étapes suivantes :

1. étape S : simuler une configuration d'objets $x^{(k)}$,
2. étape E : évaluer le logarithme de la vraisemblance jointe : $Q(\theta, \theta^k; y) = f_{\theta}(x^{(k)}, y)$,
3. étape M : maximiser le logarithme de la vraisemblance jointe : $\theta = \text{Arg Max}_{\theta} Q(\theta, \theta^k; y)$.

La convergence de cet algorithme a été discutée dans plusieurs ouvrages, mais aucun ne prouve sa convergence vers un maximum global. Cette dernière a été prouvée uniquement sous des hypothèses assez contraignantes sur la régularité du modèle [Delyon et al., 1999, Dempster et al., 1997]. Dans nos simulations, nous considérons que l'algorithme converge lorsque l'on atteint la stabilité du vecteur de paramètres θ . En outre, l'optimum obtenu par l'algorithme SEM dépend des valeurs initiales du vecteur de paramètre θ . Notamment, lorsque la solution initiale est éloignée de la solution limite, l'algorithme converge extrêmement lentement. Pour le modèle considéré de processus ponctuel, nous avons fixé à la main certains paramètres puisqu'ils ont une interprétation physique telle que, par exemple, la largeur de la frontière d'un objet ρ , ou encore la proportion de recouvrement maximale s . Nous avons également fixé le seuil de différence de contraste d_0 d'une manière déterministe. Dans [Chatelain, 2009b], nous avons proposé une méthode d'estimation de ce paramètre. De plus, il est prouvé dans [Chatelain et al., 2009a] qu'une faible valeur du paramètre β est récompensée par une grande valeur du poids d'attache aux données γ_d et vice versa. Ainsi, nous proposons de fixer le paramètre β à la main et d'estimer le poids γ_d par l'algorithme SEM. De la sorte, l'ensemble des paramètres θ à estimer est réduit à $\theta = \gamma_d$. Par ailleurs, la vraisemblance $f_{\theta}(x, y)$, donnée par l'équation (1), est inaccessible étant donné que la fonction de partition $c(\theta)$ est inconnue. C'est pourquoi, nous proposons d'approcher cette vraisemblance par la pseudo-vraisemblance [Baddeley, Turner, 2000] dont l'expression sur l'espace borné W est donnée par l'équation suivante :

$$PL_W(\theta; x, y) = \left[\prod_{x_i \in x} \lambda_{\theta}(x_i; x, y) \right] \cdot \exp \left(- \int_W \lambda_{\theta}(x_i; x, y) \Lambda(du) \right) \quad (3)$$

où $\lambda_{\theta}(u; x, y) = \beta \exp(-\gamma_d U_{\theta}^d(u, y) - \sum_{x_i \in x / x_i \neq u} U_s(u, x_i))$ est l'intensité de Papangelou étendue. $\Lambda(\cdot)$, défini sur W , est la mesure d'intensité de processus de Poisson de

référence. Bien que la densité de probabilité $f_{\theta}(x/y)$ soit inaccessible, il est possible de tirer des échantillons suivant cette loi grâce à des méthodes de type MCMC (Markov Chain Monte Carlo). Dans nos simulations, nous avons utilisé un algorithme de Metropolis-Hastings à sauts réversibles [Green, 1995].

2. Extraction d'objets par un processus d'ellipses

L'approche d'extraction d'objets, précédemment décrite, a été validée sur un processus simple de cercles et a montré des résultats prometteurs [Chatelain et al., 2009a]. Nous appliquons ce modèle afin d'extraire des objets de forme elliptique. Pour ce faire, nous avons considéré un processus ponctuel marqué d'ellipses tel que défini dans la première partie de cet article. Notons que l'espace d'objets associé à ce processus est de grande dimension (cinq) de sorte que le temps de calcul de la procédure d'extraction d'objets est très important. En effet, la complexité de calcul de la pseudo-vraisemblance donnée par l'équation (3) est proportionnelle au cardinal de W . Néanmoins, ce coût peut être réduit en envisageant une programmation parallèle de la méthode d'estimation.

Dans ce qui suit, nous appliquons l'approche proposée aux problèmes suivants :

- le comptage de flamants roses qui se regroupent lors de la période de migration ou de reproduction afin d'évaluer l'évolution d'une population,
- la détection d'arbres dans une zone boisée afin d'évaluer la dynamique des ressources forestières,
- la détection de navires à quai.

Par ailleurs, nous proposons une modification de l'énergie du modèle suivant le problème abordé. Toutes nos simulations ont été effectuées sur une machine avec un processeur de fréquence 1,86 GHz.

2.1. Extraction de flamants roses

Considérons l'image de taille 274×269 pixels, présentée sur la Figure 3 (gauche). Cette image représente une photographie aérienne d'un îlot de Camargue où est regroupée une colonie de flamants roses. La résolution d'une telle image est d'environ 20 cm.

Nous avons initialisé manuellement le paramètre d'activité β comme étant une évaluation grossière du nombre d'objets possiblement présents dans l'image ($\beta = 1000$). Comme les flamants roses se chevauchent, nous avons toléré jusqu'à 30% de recouvrement entre les ellipses ($s = 0,3$). L'algorithme d'estimation converge après 12 min et aboutit à une valeur du poids d'attache aux données égale à $\gamma_d = 16,25$. La configuration d'objets obtenue à la dernière itération de l'algorithme est présentée sur la Figure 3 (milieu).

Une fois les paramètres du modèle estimés, les objets sont extraits grâce à un algorithme de recuit simulé [Azencott, 1992, Van Laarhoven, Aarts, 1987]. Ce dernier contribue à la détection de 387 flamants roses, soit une erreur inférieure à 5% visualisés sur la Figure 3 (droite). Notons qu'une telle erreur est acceptable par les écologistes. Une vérité de terrain générée à la main est également présentée sur la figure 3 (droite). Les points

noirs sur les flamants correspondent aux bonnes détections. Les croix présentées sur les ellipses rouges correspondent aux fausses détections (faux positifs). Celles-ci sont au nombre de 5. Les ellipses noires correspondent aux flamants non détectés par la méthode proposée (4 sous-détections).

2.2. Extraction de houppiers

Nous nous focalisons maintenant sur le problème de la détection des houppiers (i.e. couronnes d'arbres). Nous considérons l'image de la Figure 4 (gauche), de taille 229×196 pixels. Elle représente un extrait d'une plantation de peupliers en Saône et Loire. Pour cette image, nous avons considéré un seuil de recouvrement entre les ellipses identique à celui de l'exemple précédent, étant donné la possibilité de recouvrement des houppiers. Nous avons également utilisé la valeur du paramètre d'activité utilisé dans la simulation précédente ($\beta = 1000$). D'autre part, le seuil de différence de contraste est fixé à $d_0 = 2$. La largeur de la couronne circonscrite considérée est $\rho = 1$. Le poids d'attache aux données estimé par l'algorithme SEM est $\gamma_d = 15,14$. Le coût calculatoire de cette phase d'estimation est environ de 38 min. La configuration pour laquelle l'algorithme a convergé est exposée sur la Figure 4 (milieu). La phase d'extraction par l'algorithme de recuit simulé a permis de localiser 598 arbres présentés sur la Figure 4 (droite). L'erreur de détection (contenant les faux positifs et les faux négatifs) que nous avons calculée à partir du résultat obtenu est de l'ordre de 2%. Nous avons également employé un processus de cercles pour la détection des arbres à partir de cette image. Le résultat de détection est aussi de bonne qualité, avec toutefois une erreur de détection plus élevée que celle obtenue avec un modèle d'ellipses. Ceci nous a encouragé à traiter d'autres applications, à savoir la détection de bateaux.

2.3. Détection de navires

Considérons la Figure 5 (gauche) qui présente un extrait d'une photographie aérienne d'un port maritime en France, de taille 385×385 pixels. Les paramètres du modèle considéré sont les suivants : le paramètre d'activité est fixé à 1000, et le seuil de différence de contraste à $d_0 = 6$. Etant donné que les navires sont assez proches et contigus, nous avons autorisé jusqu'à 30% de recouvrement entre les ellipses. L'exécution de l'algorithme SEM sur cette image a duré 1 h et 36 min et a abouti à une valeur du poids d'attache aux données égale à $\gamma_d = 38,65$. La phase d'extraction des objets par recuit simulé a permis de localiser 440 bateaux parmi 508 objets détectés. Le résultat de cette phase est exposé sur la Figure 5 (droite). Ce résultat présente environ 125 sous-détections et 68 fausses alarmes dont la plupart (65%) sont situées sur les quais du port. Nous donnerons plus loin une analyse statistique plus détaillée des résultats obtenus afin d'évaluer la précision de la solution proposée.

2.3. Evaluation des résultats

Dans le tableau 1, nous résumons les résultats statistiques des différents problèmes traités. Comme l'approche proposée est stochastique, nous donnons le temps d'exécution moyen de la phase la plus coûteuse, qui s'avère être l'estimation des paramètres par l'algorithme SEM. Nous présentons également la moyenne de l'estimée ainsi que sa déviation standard afin d'évaluer la stabilité et l'efficacité de la méthode d'estimation proposée. De plus, nous calculons l'erreur de détection comme étant le pourcentage des faux positifs et des faux négatifs. D'après ce tableau, nous pouvons conclure que l'estimateur proposé est efficace et stable puisque l'écart type de l'estimée $\sigma[\gamma_d]$ est faible par rapport à sa moyenne $\bar{\gamma}_d$ pour les trois applications. Néanmoins, il est coûteux en termes de temps de calcul moyen.

	Flamants roses	Arbres	Navires
Taille de l'image	274x269	229x196	385 x 275
Temps de calcul moyen	41 min	24 min	1h 41 min
$\bar{\gamma}_d$	17,87	14,85	38,62
$\sigma[\gamma_d]$	2,58	0,54	0,54
Nombre de faux positifs	5	< 5	> 68
Nombre de faux négatifs	4	< 5	> 125
Nombre d'objets détectés	387	598	508
Erreur de détection	2,32 %	1,67 %	> 37,99 %

Tableau 1: Résultats statistiques des simulations effectuées

D'autre part, nous pouvons déduire que l'approche proposée de processus ponctuel marqué d'ellipses est bien adaptée au problème de détection de flamants roses et d'arbres. En revanche, elle a échoué pour la détection des navires. En effet, l'erreur de détection des bateaux dépasse 37%. Dans la suite, nous nous focalisons sur le problème de détection des navires et nous proposons de modifier les deux termes d'énergie du modèle, c'est-à-dire, l'énergie a priori et l'énergie d'attache aux données.

2.3. Modification du modèle pour la détection des navires à quai

2.3.1. Modification de l'énergie d'attache aux données

Du fait que les navires sont très proches et adjacents, la couronne circonscrite à un objet n'est plus homogène et ne permet pas de distinguer les différents objets de l'image. Ceci explique les fausses alarmes présentes sur les quais du port (la Figure 5 (droite)). Comme solution à

ce problème, nous proposons de modifier la définition de la frontière d'une ellipse $F^p(u)$ impliquée dans le terme d'attache aux données. En effet, au lieu de mesurer la différence de contraste entre l'ensemble des pixels d'un objet et celui de sa couronne circonscrite, nous proposons de calculer la différence de contrastes entre l'ensemble des pixels d'un objet et l'ensemble des pixels appartenant aux deux extrémités de sa couronne circonscrite comme le montre la Figure 6 (gauche). Le calcul de cette distance est effectué grâce à la fonction de l'attache aux données (2). En considérant cette nouvelle définition de l'attache aux données, l'algorithme SEM correspondant converge après 1 h 59 min pour les mêmes valeurs de paramètres que ceux considérés lors de la simulation précédente. Il a abouti à une valeur de l'estimée de poids d'attache aux données égale à $\gamma_d = 11,95$. Sur la Figure 7 (gauche), nous présentons la configuration d'objets permettant l'extraction des bateaux, obtenue par recuit simulé. D'après ce résultat, nous pouvons remarquer une légère amélioration de la détection des navires.

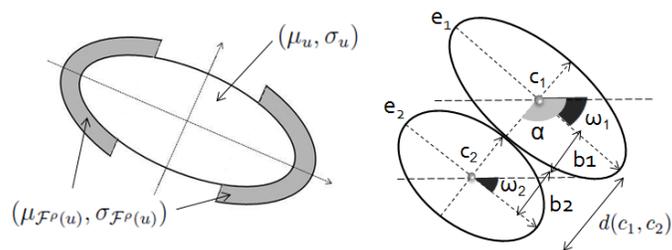


Figure 6 : gauche : Nouvelle définition de la frontière d'une ellipse pour la détection des navires, droite : alignement de deux ellipses

En effet, le nombre d'omissions a diminué par rapport à la simulation précédente. De plus, les fausses alarmes qui apparaissent sur les quais sont éliminées. Néanmoins, certaines fausses détections qui correspondent aux ellipses non alignées apparaissent. Dans ce cas, et afin d'éviter la présence de tels objets, nous proposons d'introduire une deuxième composante de régularisation permettant de favoriser l'alignement des ellipses.

2.3.2. Modification de l'énergie a priori

Comme proposé dans [Perrin et al., 2005.] pour la détection des arbres alignés, nous introduisons ici un terme de régularisation qui permet de favoriser les ellipses proches et alignées, étant donnée l'organisation des bateaux à quai. Pour ce faire, nous définissons les grandeurs suivantes qui permettent d'indiquer si deux ellipses e_1 et e_2 sont tangentes :

1. $d_\omega(e_1, e_2) = |\omega_1 - \omega_2|$ où ω_1 et ω_2 sont respectivement l'orientation de l'ellipse e_1 et l'orientation de l'ellipse e_2 . En effet, les ellipses tangentes ont la même orientation.
2. $d_\alpha(e_1, e_2) = \left| \alpha - \left(\frac{\omega_1 + \omega_2}{2} + \frac{\pi}{2} \right) \right|$ où α est l'angle que fait la droite passant par les centres des deux ellipses avec l'horizontale. En effet, pour les ellipses alignées, la droite passant par leurs centres doit être perpendiculaire à chacun des grands-axes des

ellipses comme illustré sur la Figure 6 (droite). Cette mesure nous permet de vérifier que les ellipses ne sont pas décalées et se trouvent sur le même niveau.

3. $d_c(e_1, e_2) = |d(c_1, c_2) - (b_1 + b_2)|$ où $d(c_1, c_2)$ est la distance euclidienne entre le centre c_1 de l'ellipse e_1 et le centre c_2 de l'ellipse e_2 . Les grandeurs b_1 et b_2 sont respectivement le demi-petit-axe de l'ellipse e_1 et le demi-petit-axe de l'ellipse e_2 . D'ailleurs, la distance $d_c(e_1, e_2)$ nous permet d'évaluer le rapprochement de deux ellipses.

Ainsi, deux ellipses e_1 et e_2 sont dites tangentes, si elles vérifient les trois conditions suivantes :

$$e_1 \sim_{al} e_2 \Leftrightarrow \begin{cases} d_\omega(e_1, e_2) \leq d_{\omega_{max}} \\ d_\alpha(e_1, e_2) \leq d_{\alpha_{max}} \\ d_c(e_1, e_2) \leq d_{c_{max}} \end{cases}$$

c'est-à-dire, chacune des mesures définies précédemment est limitée par un seuil donné. Les distances limites $d_{\omega_{max}}$, $d_{\alpha_{max}}$ et $d_{c_{max}}$ définissent la précision de la relation d'alignement \sim_{al} , et sont mises à une valeur proche de 0 dans nos tests. Pour qualifier une telle relation, nous avons eu recours à une fonction dite de récompense qui a été employée dans [Ortner et al., 2008] afin de favoriser l'alignement des bâtiments. Elle s'écrit comme suit :

$$\varpi(x, x_{max}) = -\frac{1}{x_{max}^2} \left[\frac{1 + x_{max}^2}{1 + x^2} - 1 \right], \text{ pour } x \leq x_{max} \quad (4)$$

Cette fonction a été construite de sorte à favoriser les cas suivants :

- lorsque $x = 0 \Rightarrow \varpi(0, x_{max}) = -1$
- lorsque $x = x_{max} \Rightarrow \varpi(x_{max}, x_{max}) = 0$

Nous avons ensuite défini l'énergie a priori liée à l'alignement de deux ellipses comme suit :

$$U_{al}(e_1, e_2) = \begin{cases} \delta(e_1, e_2) \cdot \varpi(d_\alpha(e_1, e_2), d_{\alpha_{max}}) & \text{si } e_1 \sim_{al} e_2 \\ 0 & \text{sinon} \end{cases}$$

$\delta(e_1, e_2)$ est un coefficient qui qualifie le rapprochement des deux ellipses e_1 et e_2 . Il varie en fonction de la distance $d_c(e_1, e_2)$. Lorsque les ellipses e_1 et e_2 sont tangentes, ce coefficient est maximal, il est égal à 1. En revanche, lorsque les ellipses sont éloignées, ce coefficient est nul. Pour formaliser cette idée, nous avons considéré la fonction linéaire suivante :

$$\delta(e_1, e_2) = 1 - \frac{1}{2 b_{max}} d_c(e_1, e_2)$$

avec b_{max} désignant le demi-petit-axe maximal d'une ellipse. Ainsi, l'énergie a priori liée à l'alignement des ellipses d'une configuration d'objets x est la somme des énergies a priori d'alignement définies au niveau de chaque paire d'objets (e_i, e_j) de x :

$$U_{al}^p(x) = \gamma_{al} \sum_{0 < i < j < n(x)} U_{al}(e_i, e_j) \quad (5)$$

Le paramètre γ_{al} est également un paramètre de notre modèle. Nous proposons alors de l'estimer par l'algorithme SEM. Dans ce cas, il doit être initialisé soigneusement. En effet, l'estimation par la méthode SEM est très dépendante du choix des paramètres initiaux. Notamment, lorsque le vecteur de paramètres de départ est éloigné de l'estimateur de Maximum de Vraisemblance (MV), l'algorithme converge extrêmement lentement. Empiriquement, en effectuant plusieurs simulations de l'algorithme SEM sur l'image de navires de la Figure 5 (gauche), nous avons conclu que la valeur de γ_{al} la plus convenable, qui assure une meilleure détection des objets, est celle qui est inférieure à $\frac{\gamma_d^{(0)}}{2}$. D'ailleurs, nous avons initialisé le paramètre du poids de l'énergie a priori d'alignement comme étant le tiers de la valeur initiale du poids d'attache aux données, les autres paramètres du modèle ayant les mêmes valeurs que pour la simulation précédente. L'algorithme d'estimation a abouti aux valeurs des estimées suivantes $\gamma_d = 25.34$ et $\gamma_{al} = 10.53$ après 2h 36 min de temps de calcul. Ces paramètres ont contribué à l'image d'extraction montrée en Figure 7 (milieu). En comparant ce résultat à celui des simulations précédentes, nous pouvons constater une légère amélioration de la détection des bateaux. En effet, les ellipses de la Figure 7 (milieu) sont plus ordonnées que celles des Figures 5 (droite) et 7 (gauche). Toutefois, certaines fausses alarmes persistent telles que celles qui correspondent aux ellipses alignées, mais qui se placent sur le même navire (Figure 7 (milieu)). Ces ellipses ont généralement une orientation différente de celles des navires dans l'image. Ceci s'explique par le fait que les ellipses alignées ont une énergie plus petite (respectivement une probabilité plus importante) que celle d'une seule ellipse se plaçant sur le navire. Une solution à ce problème consiste à favoriser les ellipses ayant une orientation privilégiée, puisque les navires dans l'image traitée ont une direction commune. Pour ce faire, nous avons estimé cette direction de façon empirique, soit une orientation ω_N . En d'autres termes, nous avons effectué une simulation rapide d'une configuration d'objets. Ensuite, nous avons calculé l'angle ω_N comme étant la moyenne des orientations des ellipses les plus adaptées aux données (ie. les ellipses qui se positionnent bien sur les navires de l'image). Nous avons ensuite modifié l'énergie a priori de sorte à inclure la contrainte sur l'orientation des ellipses alignées. En effet, seuls les navires à quai qui correspondent aux ellipses alignées suivent la direction donnée par l'angle ω_N . Ainsi, une ellipse est favorisée via l'équation (5) uniquement si cette ellipse a une orientation proche de ω_N . Si l'ellipse ne vérifie pas cette condition, elle n'est favorisée par aucune énergie a priori. Ainsi, l'énergie a priori liée à l'alignement des objets est désormais définie comme suit :

$$U_{al_{\omega_N}}(e_i, e_j) = \begin{cases} U_{al}(e_i, e_j) & \text{si } |\omega_{e_i} - \omega_N| < d_{\omega_{max}} \text{ et } |\omega_{e_j} - \omega_N| < d_{\omega_{max}} \\ 0 & \text{sinon} \end{cases}$$

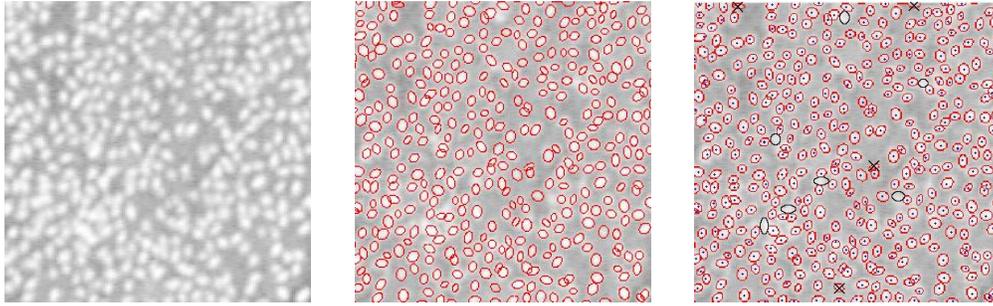


Figure 3 : gauche : colonie de flamants roses en Camargue, France © La Tour du Valat, milieu : configuration d'ellipses simulée à l'étape S de l'algorithme SEM pour les paramètres ($\beta = 1000, \gamma_d = 16,25, s = 0,3, d_0 = 1,33$), droite : résultat d'extraction des flamants roses, 387 ellipses.

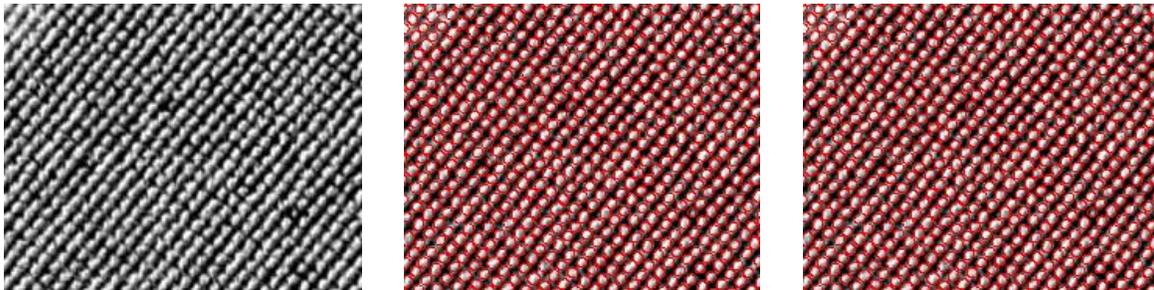


Figure 4 : gauche : extrait d'une plantation de peupliers en Saône et Loire, France © IFN, milieu : configuration d'ellipses obtenue à l'étape S de l'algorithme SEM pour les paramètres ($\beta = 1000, \gamma_d = 15,14, s = 0,2, d_0 = 2$), droite : résultat d'extraction des arbres, 598 ellipses.

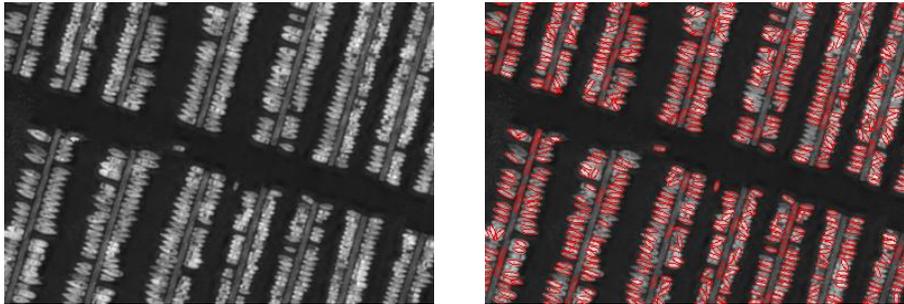


Figure 5 : gauche : image de navires en France © CNES, droite : résultat d'extraction des bateaux pour les paramètres ($\beta = 1000, \gamma_d = 38,65, s = 0,3, d_0 = 6$), 538 ellipses.

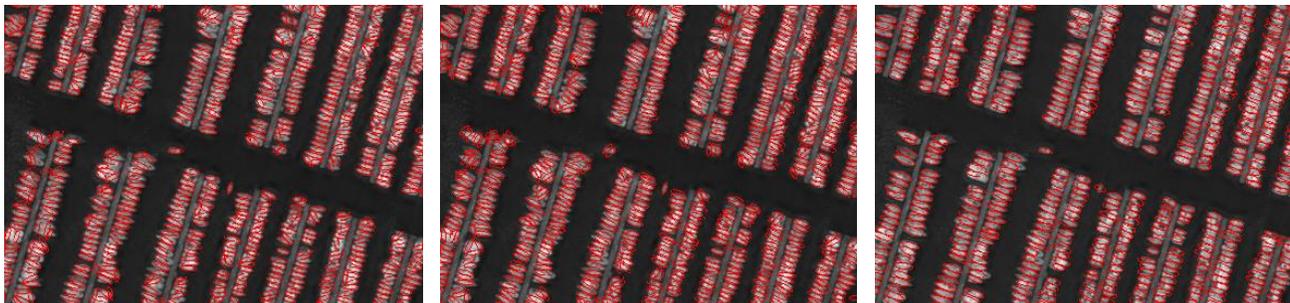


Figure 7 : gauche : résultat d'extraction des bateaux avec modification de l'énergie d'attache aux données, 635 ellipses obtenues pour les paramètres ($\beta = 1000, \gamma_d = 11,95, s = 0,3, d_0 = 6$), milieu : résultat d'extraction des bateaux avec favorisation des ellipses tangentielles, 715 ellipses obtenues pour les paramètres ($\beta = 1000, \gamma_d = 25,34, \gamma_{al} = 10,53, s = 0,3, d_0 = 6$), droite : résultat d'extraction des navires avec favorisation des ellipses alignées ayant une direction privilégiée, 523 ellipses obtenues pour les paramètres ($\beta = 1000, \gamma_d = 27,56, \gamma_{al} = 9,18, s = 0,3, d_0 = 6$).

La phase d'estimation est accomplie en 1h 38 min et a abouti aux valeurs suivantes des paramètres $\gamma_d = 27,56$ et $\gamma_{al} = 9,18$. Ces paramètres ont permis la détection de 523 objets présentés sur la Figure 7 (droite). Cette approche a résulté en une meilleure détection des bateaux avec une erreur inférieure à 9,5 %.

Une limite de cette approche est qu'elle est contrainte par une organisation particulière des bateaux dans l'image. En effet, elle n'est valable que pour des images de navires ayant une direction commune. Néanmoins, elle peut être généralisée à des images de navires ayant des orientations différentes, en effectuant une estimation locale de la direction des navires. En effet, l'estimation locale des orientations des navires dans l'image est envisageable grâce à la connaissance d'une carte portuaire ou par une estimation de l'orientation du gradient. Dans ce cadre, nous pouvons également nous inspirer des travaux développés dans [Perrin et al., 2006] où une méthode fondée sur la transformée de Fourier pour l'estimation de la direction des arbres dans une forêt est proposée.

3. Extraction d'objets par un processus de rectangle

3.1. Extraction de la trace au sol du bâti

Considérons une partie du MNE de la ville d'Amiens, de taille 231×194 pixels, représentée sur la Figure 8 (gauche). Pour cette image, nous avons autorisé jusqu'à 20% de recouvrement entre les objets ($s = 0,3$) car on s'intéresse uniquement à la silhouette des bâtiments. De plus, nous avons fixé β à 500. La valeur du seuil de différence de contraste considérée est $d_0 = 4$.

En considérant ces paramètres, l'algorithme d'estimation converge après 51 min. Le résultat est donné sur la Figure 14. Cette phase a été accomplie avec un poids d'attache aux données γ_d égal à 29,40. Ce paramètre ainsi trouvé a permis d'entamer la phase d'extraction dont le résultat est donné sur la Figure 8 (milieu). Certains rectangles ne sont pas bien alignés avec leurs voisins et sont éloignés, ce qui entraîne une légère sous-détection des objets. Afin d'éviter ce problème, nous proposons d'introduire un terme de régularisation qui permet de favoriser les rectangles proches, alignés ou orthogonaux. En effet, les bâtiments dans les grandes villes présentent usuellement cette structure géométrique (bâtiments alignés et orthogonaux).

3.2. Régularisation du modèle pour l'extraction des empreintes de bâtiments

Nous décrivons ici un terme qui favorise l'alignement et l'orthogonalité des rectangles. Pour ce faire, nous nous sommes inspirés des travaux présentés dans [Ortner et al., 2008] sur l'extraction automatique des caricatures de bâtiments. Nous définissons ainsi ces deux types de relations qui lient les bâtiments des grandes villes. Notons qu'une interaction orthogonale entre un couple de rectangles n'est pas vue comme étant une simple rotation de $\frac{\pi}{2}$ entre les rectangles. Une contrainte sur les cotés des rectangles concernés doit être prise en compte comme détaillé dans ce qui suit.

3.2.1. Relation d'alignement

Deux rectangles u et v sont dits alignés, si l'un des petits cotés de u est opposé à un petit côté de v (cf. Figure 9 (gauche)). Une formulation de cette interaction entre rectangles est donnée par les trois conditions suivantes : d'abord les rectangles alignés doivent être suffisamment proches de sorte que la distance euclidienne $d_C(i, j)$ entre le coin indexé i du rectangle u et le coin indexé j du rectangle v soit faible avec i et j inclus dans l'ensemble $\{0, 1, 2, 3\}$. De plus, les rectangles concernés doivent avoir la même orientation de sorte que $d_\omega(u, v) = |\omega_u - \omega_v| \leq d_{\omega_{max}}$, où ω_u et ω_v sont respectivement l'orientation du rectangle u et l'orientation du rectangle v . Enfin, la condition $i + j = 3$ permet de garantir que les petits cotés appropriés des deux rectangles se regardent. Ainsi, une telle relation d'alignement \sim_{al_i} peut être définie comme suit :

$$u \sim_{al_i} v, i \in \{0, 1, 2, 3\} \Leftrightarrow \begin{cases} d_C(i, j) \leq d_{C_{max}} \\ d_\omega(u, v) \leq d_{\omega_{max}} \\ i + j = 3 \end{cases}$$

$d_{C_{max}}$ est la distance maximale qui sépare les coins concernés de deux rectangles alignés et $d_{\omega_{max}}$ désigne le seuil au dessous duquel deux rectangles sont considérés comme parallèles. Ces paramètres ($d_{C_{max}}$ et $d_{\omega_{max}}$) contrôlent la précision de la relation d'alignement. Notons également que la relation ainsi définie n'est ni symétrique, ni réflexive.

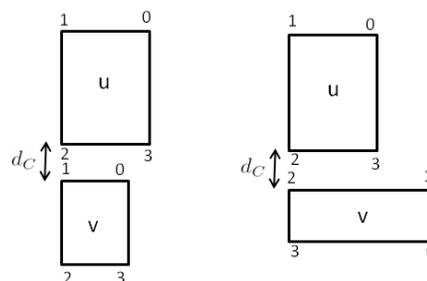


Figure 9 : gauche : rectangles alignés, droite : rectangles orthogonaux.

En outre, nous avons exploité la fonction décrite par l'équation (4) pour qualifier une telle relation. Ainsi, l'énergie a priori d'une paire d'objets (u, v) , liée à la relation \sim_{al_i} , est définie par l'expression suivante :

$$U_{al_i}(u, v) = \begin{cases} \frac{1}{2} \varpi(d_C(i, j), d_{C_{max}}) + \frac{1}{2} \varpi(d_\omega(u, v), d_{\omega_{max}}) & \text{si } u \sim_{al_i} v \\ 0 & \text{sinon} \end{cases} \quad (6)$$

Cette fonction attribue une valeur négative aux paires d'objets alignés en fonction de la différence de leur orientation ainsi que de la distance entre les coins appropriés. Par conséquent, l'énergie a priori de la configuration de rectangles x , qui correspond à la

contrainte d'alignement, est la somme de toutes les énergies définies au niveau de chaque paire d'objets :

$$U_{al}^p(\mathbf{x}) = \gamma_{al} \sum_{0 < j < k < n(\mathbf{x})} \sum_{0 \leq i \leq 3} U_{ali}(x_j, x_k)$$

où la pondération γ_{al} est un paramètre du modèle qui quantifie l'importance de la contrainte d'alignement par rapport aux autres contraintes du modèle.

3.2.2 Relation d'orthogonalité

Comme les bâtiments dans les zones urbaines n'ont pas forcément la même taille, ces derniers ne peuvent pas être disposés de manière strictement parallèle. Nous proposons de favoriser également les objets orthogonaux. La relation d'orthogonalité qui lie une paire de rectangles est définie ainsi : deux rectangles u et v sont dits orthogonaux, si l'une des petites arêtes de l'un des rectangles est opposée à une grande arête de l'autre rectangle (Figure 9 (droite)). Une formulation de cette relation consiste à satisfaire les conditions suivantes :

$$u \sim_{orth_i} v, i \in \{0, 1, 2, 3\} \Leftrightarrow \begin{cases} d_c(i, j) \leq d_{c_{max}} \\ \left| d_\omega(u, v) - \frac{\pi}{2} \right| \leq d_{\omega_{max}} \\ i + j = 0, 2, 4 \text{ ou } 6 \end{cases}$$

Par analogie avec l'expression de l'énergie d'alignement donnée par l'équation (6), l'énergie a priori d'une paire de rectangles qui correspond aux connexions orthogonales suivant la relation d'orthogonalité \sim_{orth_i} s'écrit comme suit :

$$U_{orth_i}(u, v) = \begin{cases} \frac{1}{2} \left[\varpi(d_c(i, j), d_{c_{max}}) + \varpi \left(\left| d_\omega(u, v) - \frac{\pi}{2} \right|, d_{\omega_{max}} \right) \right] & \text{si } u \sim_{orth_i} v \\ 0 & \text{sinon} \end{cases}$$

Finalement, l'expression de l'énergie a priori d'une configuration d'objets \mathbf{x} qui favorise les interactions orthogonales est donnée par l'équation suivante :

$$U_{orth}^p(\mathbf{x}) = \gamma_{orth} \sum_{0 < j < k < n(\mathbf{x})} \sum_{0 \leq i \leq 3} U_{orth_i}(x_j, x_k)$$

avec γ_{orth} la pondération de cette énergie. Ainsi, l'énergie d'une configuration de rectangles qui favorise les interactions entre les bâtiments peut être définie comme étant la somme des énergies d'alignement et d'orthogonalité, pondérée par le coefficient γ_{bat} :

$$U_{bat}^p(\mathbf{x}) = \gamma_{bat} [U_{al}^p(\mathbf{x}) + U_{orth}^p(\mathbf{x})]$$

L'introduction du paramètre γ_{bat} nous permet de réduire le nombre de paramètres à estimer. En effet, au lieu d'estimer les deux paramètres γ_{al} et γ_{orth} , liés respectivement à la contrainte d'alignement et d'orthogonalité des rectangles, nous pouvons normaliser la valeur de ces paramètres et la compenser par la valeur du paramètre γ_{bat} défini sur \mathbb{R}^+ . En particulier, nous fixons ces deux coefficients γ_{al} et γ_{orth} à 1 pour l'image

de la ville d'Amiens de la Figure 8 (gauche) puisque les connexions orthogonales et alignées sont présentes avec la même importance. En conséquence, l'énergie associée aux interactions entre les bâtiments est désormais exprimée comme suit :

$$U_{bat}^p(\mathbf{x}) = \gamma_{bat} \sum_{0 < j < k < n(\mathbf{x})} \sum_{0 \leq i \leq 3} [U_{ali}(x_j, x_k) + U_{orth_i}(x_j, x_k)]$$

L'initialisation du paramètre γ_{bat} comme pour les autres paramètres du modèle, est cruciale pour la convergence et la rapidité de l'algorithme SEM. Dans [Ben Hadj et al., 2009b], nous étudions la fonction d'énergie totale du modèle afin de restreindre le domaine de définition de ce paramètre. D'autre part, l'expérience a prouvé que la valeur la plus appropriée doit être inférieure à celle du poids d'attache aux données. D'ailleurs, nous avons considéré la moitié de la valeur du poids d'attache aux données initiale ($\gamma_{bat}^0 = \frac{\gamma_d^0}{2}$), les autres paramètres étant les mêmes que pour le précédent test effectué sur l'image d'Amiens. L'algorithme d'estimation converge après 6 h 21 min et a abouti aux valeurs des estimées suivantes : $\gamma_d = 30,05$ et $\gamma_{bat} = 16,84$. Ces paramètres nous ont permis d'entamer la phase d'extraction de la trace au sol du bâti qui s'est achevée par la génération de 83 rectangles présentés sur la Figure 8 (droite). En comparant les résultats d'extraction des Figures 8 (milieu) et 8 (droite), nous constatons qu'effectivement, le terme a priori sur les connexions entre les bâtiments a permis de rapprocher les rectangles et d'imposer ceux qui sont alignés ou orthogonaux à l'exception de 5 paires distantes. La présence de tels objets distants se justifie par le fait que la distance qui les sépare ne suffit pas pour créer un nouvel objet sans qu'il ne se superpose avec ces voisins. En effet, un recouvrement qui dépasse 20 % est strictement interdit. Ce problème peut être corrigé en tolérant un taux de recouvrement plus grand ou en effectuant des simulations avec des rectangles de taille très variable. Cette taille est contrôlée par les bornes a_{min} , a_{max} , b_{min} et b_{max} de l'espace des objets. En effet, dans nos tests, nous avons choisi un espace d'objets restreint.

3.3. Détection de tentes

Nous proposons maintenant d'appliquer le modèle de rectangle afin de détecter des tentes de réfugiés sur l'image de taille 301×301 pixels, présentée sur la Figure 10 (gauche et milieu). Il s'agit d'une photographie de tentes dans un campement d'urgence après un tremblement de terre à Boumerdès en Algérie. Les tentes sont les objets rectangulaires rouges et bleus, on en dénombre 52 tentes sur cette image. Elles ont quasiment la même taille. Ceci représente un avantage en termes de temps de calcul de la pseudo-vraisemblance lors de la phase d'estimation. En effet, le calcul de la pseudo-vraisemblance représente la tâche la plus coûteuse de l'algorithme d'estimation. Par ailleurs, puisque les tentes ne sont pas parfaitement alignées, nous n'avons pas pris en compte la composante d'énergie qui contrôle l'alignement des objets. Les énergies considérées sont

l'énergie d'attache aux données et l'énergie a priori qui correspond à une pénalisation du recouvrement entre les objets. En effet, nous n'avons autorisé aucun chevauchement entre les objets du fait que les tentes sont espacées. Nous avons alors fixé la valeur de la proportion de recouvrement maximale entre les objets s à 0. En outre, nous avons choisi un paramètre d'activité initial β égal à 500. Quant à la valeur du seuil initial de différence de contrastes d_0 , elle doit permettre à la fois la détection des tentes bleues (tentes claires sur l'image en niveaux de gris (cf. Figure 10)) ainsi que des tentes rouges (tentes sombres sur l'image en niveaux de gris). Les tests ont montré qu'une valeur égale à 3 permet la détection des deux types de tentes. En considérant ces valeurs de paramètres, l'algorithme SEM converge après 26 min et aboutit à une valeur du poids d'attache aux données égale à $\gamma_d = 21,22$. La phase d'extraction a permis de détecter 117 objets dont 55 % sont des fausses alarmes, visualisées sur la Figure 10 (droite). Nous remarquons que bien que le nombre de fausses alarmes soit important, nous avons réussi à détecter quasiment toutes les tentes. Ces fausses détections s'expliquent par le fait que la composante d'attache aux données ne prend pas en considération, ou peu, l'homogénéité des pixels de l'objet et ceux de son environnement. Ainsi, lorsque l'image traitée contient une variété d'objets, le modèle considéré contribue inévitablement à des fausses détections. Par exemple, nous remarquons la présence d'objets dont la frontière n'est pas homogène comme la série de rectangles qui se place sur l'ombre du bâtiment. Par ailleurs, plusieurs fausses alarmes correspondent à des rectangles ayant un fond non homogène, tels que ceux qui se placent sur les véhicules. Pour réduire ces fausses alarmes, nous effectuons un post-traitement qui permet de supprimer les objets qui sont mal adaptés aux données. Nous redéfinissons la composante d'attache aux données en subdivisant l'environnement d'un rectangle en quatre régions comme illustré par la Figure 11.

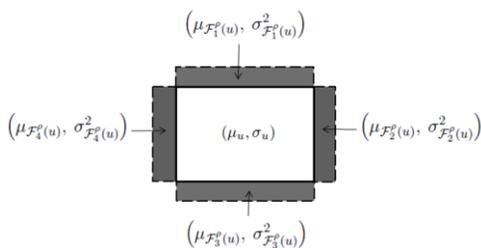


Figure 11 : Nouvelle définition de l'attache aux données pour la détection des tentes.

Nous obtenons ainsi quatre petits rectangles qui entourent le rectangle u , ayant chacun une moyenne $\mu_{F_i^p(u)}$ et une variance $\sigma_{F_i^p(u)}$, avec $i \in \{0, 1, 2, 3\}$.

Nous mesurons la similarité entre l'ensemble des pixels de l'objet u et celui des quatre régions de sa frontière via la distance donnée par l'équation (2). Ainsi, nous construisons l'énergie d'attache aux données comme étant une qualification de la distance minimale des quatre distances calculées. Autrement-dit, l'expression de la

nouvelle énergie d'attache aux données associée à un objet u s'écrit comme suit :

$$U_{\theta}^d(u, \mathcal{Y}) = Q \left(\frac{\text{Min}_{0 \leq i \leq 3} d(u, F_i^p(u))}{d_0} \right)$$

$F_i^p(u)$ désigne un petit rectangle dans la frontière de l'objet u . Ensuite, nous supprimons d'une manière déterministe tous les objets de la configuration obtenue qui ne correspondent pas bien aux données suivant cette nouvelle définition de l'attache aux données, soit tous les objets dont l'énergie d'attache aux données est positive. Ce traitement a permis de supprimer 60 fausses alarmes à partir de la configuration d'objets visualisées sur la Figure 10 (droite). La nouvelle configuration obtenue par ce traitement est présentée par la Figure 12 (gauche). Nous constatons, néanmoins, la présence de certaines fausses alarmes qui correspondent aux rectangles dont le fond est non homogène. Dans une deuxième étape de ce post-traitement, nous supprimons ces objets de fond non homogènes en effectuant un simple test sur leur variance. Ceci a abouti à la suppression de 5 objets et a donné lieu à la configuration d'objets de la Figure 12 (droite).

En résumé, l'approche proposée pour la détection des tentes est fondée sur la géométrie des objets et n'exploite pas ou peu d'autres informations telles que la couleur des objets à détecter, ce qui explique son échec quant à la détection d'objets à partir d'images non homogènes, contenant une variété d'objets. Par ailleurs, le post-traitement proposé a permis de minimiser les fausses alarmes. Il pourrait être intéressant d'injecter une information de couleur dans le modèle afin de localiser les tentes sur une telle image, en associant une marque de couleur aux objets. Pour ce faire, nous pourrions effectuer, dans un premier temps, une extraction des objets suivant leur forme sur l'image en niveau de gris. Dans une deuxième étape, nous traiterions, l'image en couleur associée en sélectionnant à partir de la configuration obtenue les objets ayant une marque de couleur qui correspond à celle des tentes (la couleur des tentes est supposée connue : soit bleue, soit rouge).

4. Conclusion

L'approche présentée dans cet article a été développée durant le stage de master du premier auteur. Nous avons présenté une méthode non supervisée pour l'extraction d'une collection d'objets en se basant sur une modélisation par processus ponctuels marqués. Les paramètres impliqués dans ce modèle sont estimés grâce à une version stochastique de l'algorithme Espérance-Maximisation. En particulier, nous avons appliqué cette approche afin d'extraire des objets elliptiques et rectangulaires. Le modèle de processus d'ellipses a prouvé son efficacité quant à l'extraction des flamants roses et des houppiers d'arbres. Néanmoins, son application au problème de détection des navires a nécessité une redéfinition du terme d'attache aux données ainsi que l'introduction d'une contrainte a priori représentant l'alignement des navires. Une bonne détection des bateaux a été obtenue en favorisant les ellipses orientées suivant une direction privilégiée.

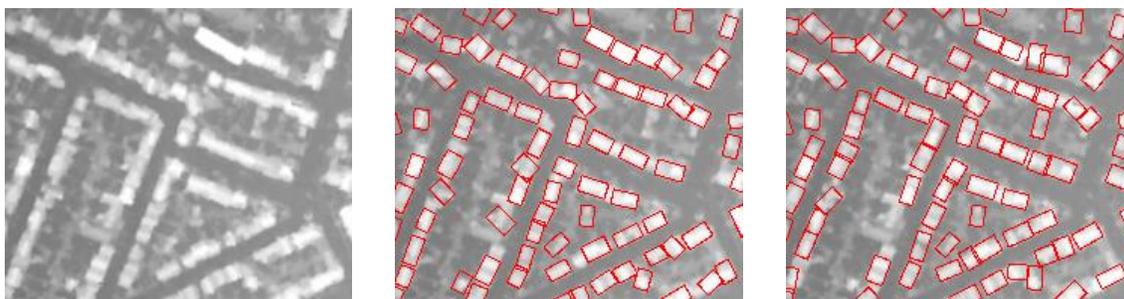


Figure 8 : gauche : Modèle Numérique d'Élévation d'une partie d'Amiens, France © IGN, milieu : résultat d'extraction de la trace au sol du bâti pour les paramètres ($\beta = 500, \gamma_d = 29,40, s = 0,2, d_0 = 4$), la configuration d'objets visualisée contient 83 rectangles, droite : résultat d'extraction de la trace au sol du bâti avec des rectangles alignés et orthogonaux obtenu pour les paramètres ($\beta = 500, \gamma_d = 30,05, \gamma_{bat} = 16,84, s = 0,2, d_0 = 4$), la configuration présentée contient 83 rectangles.



Figure 10 : gauche : extrait d'une image couleur de tentes à Boumerdès, Algérie © CNES, milieu : l'image de tentes qui correspond à un seul canal (quatrième canal), droite : résultat d'extraction des tentes pour les paramètres ($\beta = 500, \gamma_d = 21,22, s = 0, d_0 = 3$), 117 rectangles présentés.

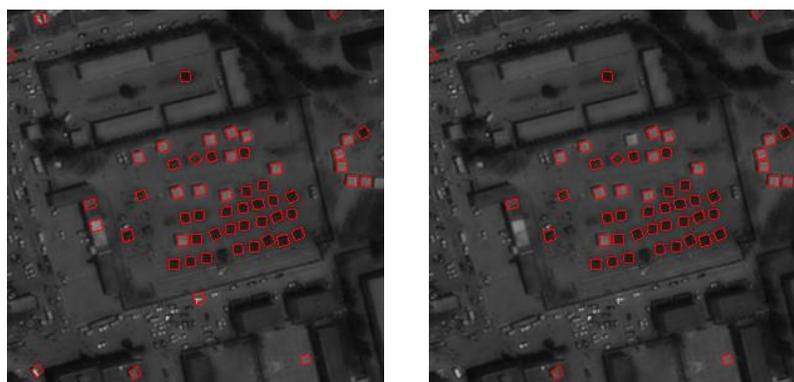


Figure 12 : gauche : résultat d'extraction des tentes après suppression des fausses alarmes (suivant la nouvelle définition de l'attache aux données), 57 rectangles présentés, droite : résultat d'extraction des tentes après suppression des objets non homogènes, 52 rectangles présentés.

Comme cette direction est choisie empiriquement dans nos tests, il peut être intéressant de l'estimer automatiquement par une estimation de la direction du gradient. En outre, la validation du modèle de rectangle sur des images de bâtiments a nécessité l'intervention d'un terme d'énergie qui favorise les connexions alignées ou orthogonales entre ces rectangles, ceci en raison de la structure organisée des bâtiments dans les zones urbaines. L'extraction de tentes à partir des images non homogènes peut être atteinte avec plus de précision en

envisageant un modèle qui prendrait en compte l'information de couleur des objets à extraire. Une des limites de l'approche proposée est le coût calculatoire important de la méthode d'estimation qui a été amplifié par l'utilisation d'objets de plus grande dimension (les ellipses et les rectangles). Pour faire face à ce problème, nous pouvons envisager l'utilisation de stratégies d'optimisation de la méthode d'estimation, telles que les techniques de « bootstrap » [Bougarradh, 2008] où l'estimation est effectuée sur des régions de l'image. De

plus, il pourrait être aussi intéressant de réaliser par une programmation parallélisée de la méthode d'estimation, notamment pour le calcul de la pseudo-vraisemblance qui constitue la tâche la plus coûteuse de l'algorithme SEM, (environ 70% du temps de calcul total). En outre, cette étude a visé essentiellement l'extraction d'objets surfaciques. Elle pourrait être étendue à l'extraction d'autres objets linéiques tels que les fleuves ou les routes. Dans ce cas, une représentation des objets par des segments serait alors adéquate [Lacoste et al., 2003, Lacoste et al., 2005].

Remerciement

Ce travail a été financé en partie par le Centre National d'Etudes Spatiales (CNES) dans le cadre du programme ORFEO. Nous tenons à remercier le CNES pour ce financement ainsi que pour les images fournies. Nous tenons à remercier également l'Inventaire Forestier National (IFN) et la station d'écologie de la Tour du Valat pour la fourniture des images aériennes et satellitaires sur lesquelles nous avons travaillé.

Références

Azencott R., 1992. *Simulated annealing: Parallelization techniques*. Springer-Verlag.

Baddeley A., Turner R., 2000. *Practical maximum pseudolikelihood for spatial point patterns (with discussion)*. Australian and New Zealand Journal of Statistics 42(3), pp. 283-322.

Ben Hadj S., Chatelain F., Descombes X., Zerubia J., 2010a. *Parameter estimation for a marked point process within a framework of multidimensional shape extraction from remote sensing images*. ISPRS, Paris 38, pp. 1-6.

Ben Hadj S., Chatelain F., Descombes X., Zerubia J., 2010b. *Estimation des paramètres de modèles de processus ponctuels marqués pour l'extraction d'objets en imagerie spatiale et aérienne haute résolution*. Rapport de recherche numéro 7350, INRIA.

Bougarradh A., M'hiri S., Ghorbel F., 2008. *Introduction of the bootstrap resampling in the generalized mixture estimation*. IEEE.

Casella G., Robert C., 1999. *Monte Carlo Statistical Methods*. Springer-Verlag.

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(4), pp. 287-314.

Chatelain, F., Descombes, X. and Zerubia, J., 2009a. *Estimation des paramètres de processus ponctuels marqués dans le cadre de l'extraction d'objets en imagerie de télédétection*. GRETSI'09, Dijon, France.

Chatelain, F., Descombes, X. and Zerubia, J., 2009b. *Parameter estimation for marked point processes. Application to object extraction from remote sensing*

images. EMMCVPR'09, LNCS, Springer-Verlag, pp. 221-234.

Delyon B., Lavielle M., Moulines E., 1999. *Convergence of a stochastic approximation version of the EM algorithm*. The Annals of Statistics 27(1), pp.94-128.

Dempster A. P., Laird N. M., Rubin D. B., 1997. *Maximum likelihood from incomplete data via the EM algorithm*. Journal of the Royal Stat. Society. Series B (Methodological) 39(1), pp. 1-38.

Green P., 1995. *Reversible jump Markov chain Monte Carlo computation and Bayesian model determination*. Biometrika 82(4), pp. 711-472.

Lacoste C., Descombes X., Zerubia J., Baghdadi N., 2003. *Extraction automatique des réseaux linéiques à partir d'images satellitaires et aériennes par processus Markov objet*. SFPT 170, pp. 13-22.

Lacoste C., Descombes X., Zerubia J., Baghdadi N., 2005. *Point Processes for Unsupervised Line Network Extraction in Remote Sensing*. IEEE Transactions on Pattern Analysis and Machine Intelligence 27(10), pp. 1568-1579.

Møller J., Waagepetersen R., 2004. *Statistical inference and simulation for spatial point processes*. CRC Press.

Ortner M., Descombes X., Zerubia J., 2008. *A marked point process of rectangles and segments for automatic analysis of Digital Elevation Models*. IEEE Transactions on Pattern Analysis and Machine Intelligence 30(1), pp. 105-119.

Perrin G., Descombes X., Zerubia J., 2005. *A Marked Point Process Model for Tree Crown Extraction in Plantations*. IEEE International Conference on Image Processing (ICIP), Gènes, Italie.

Van Laarhoven P.J., Aarts E. H., 1987. *Simulated annealing : theory and applications*. Springer-Verlag.