



HAL
open science

Extraction de réseaux linéiques à partir d'images satellitaires et aériennes par processus ponctuels marqués

Caroline Lacoste

► **To cite this version:**

Caroline Lacoste. Extraction de réseaux linéiques à partir d'images satellitaires et aériennes par processus ponctuels marqués. Interface homme-machine [cs.HC]. Université Nice Sophia Antipolis, 2004. Français. NNT: . tel-00261397

HAL Id: tel-00261397

<https://theses.hal.science/tel-00261397>

Submitted on 6 Mar 2008

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.

UNIVERSITÉ DE NICE-SOPHIA ANTIPOLIS - UFR Sciences

École Doctorale : Sciences et Technologies de l'Information et de la Communication

THÈSE

pour obtenir le titre de

DOCTEUR EN SCIENCES

DE L'UNIVERSITÉ DE NICE-SOPHIA ANTIPOLIS

Mention : Automatique, Traitement du signal et des images

par

Caroline LACOSTE

*Extraction de réseaux linéiques
à partir d'images satellitaires et aériennes
par processus ponctuels marqués*

Thèse dirigée par Josiane ZERUBIA

et préparée à l'INRIA Sophia Antipolis, projet ARIANA

Soutenue le 30 Septembre 2004 devant la commission d'examen

Michel	BARLAUD	Professeur, UNSA	Président
Michel	SCHMITT	Professeur, ENSM	Rapporteurs
Henri	MAITRE	Professeur, ENST	
Nicolas	BAGHDADI	Chercheur, BRGM	Examineur
Xavier	DESCOMBES	Chargé de recherche, INRIA	Co-encadrant
Josiane	ZERUBIA	Directrice de recherche, INRIA	Directrice

Remerciements

Ce travail a été réalisé à l'INRIA (Institut National de Recherche en Informatique et en Automatique), au sein du projet ARIANA. Il a été financé par l'INRIA et le BRGM (Bureau de Recherche Géologique et Minière).

Je remercie Michel BARLAUD, Professeur à l'UNSA, d'avoir accepté de présider ce jury de thèse.

Je suis reconnaissante à Michel SCHMITT, Professeur à l'ENSM, et Henri MAITRE, Professeur à l'ENST, d'avoir apporté leur point de vue sur cette thèse en acceptant d'en être les rapporteurs.

Je tiens à remercier Nicolas BAGHDADI, Chercheur au BRGM, d'avoir suivi régulièrement l'avancement de cette thèse, d'avoir participé à la phase de validation des résultats, et d'avoir bien voulu juger ce travail.

Je souhaite exprimer ma gratitude à Josiane ZERUBIA et Xavier DESCOMBES pour m'avoir encadré avec disponibilité et clairvoyance tout au long de cette thèse. Merci à Josiane pour m'avoir encouragée, poussée à publier, et m'avoir permis des rencontres scientifiques intéressantes et variées. Merci à Xavier pour m'avoir guidée tout au long de cette thèse, pour ses conseils avisés et ses bonnes idées.

Je remercie Hervé LE MEN (IGN) pour ses commentaires pertinents relativement à l'extraction de réseaux routiers.

Je remercie également Radu STOICA (Université Jaume I, Castellón) et Marie-Colette VAN LIESHOUT (CWI, Amsterdam) pour les échanges scientifiques et les conseils prodigués lors de nos rencontres.

Je voudrais enfin remercier tous ceux qui m'ont entourée durant ces trois années.

Table des matières

Introduction	5
1 État de l'art en extraction de réseaux	9
1.1 Les différents types de réseaux	9
1.1.1 Quels sont les réseaux d'intérêt ?	9
1.1.2 Caractéristiques communes	11
1.1.3 Variabilité des réseaux	12
1.2 Méthodes semi-automatiques	14
1.2.1 Algorithmes de suivi	14
1.2.2 Programmation dynamique	16
1.2.3 Contours actifs	17
1.2.4 Analyse multi-résolution de profils transversaux	18
1.3 Méthodes automatiques	18
1.3.1 Détection de lignes par optimisation locale	19
1.3.2 Couplage d'algorithmes	24
1.3.3 Apport de données cartographiques	26
1.3.4 Approche multi-résolution	27
1.3.5 Systèmes d'interprétation	28
1.3.6 Processus objet	29
2 Extraction d'objets par processus ponctuels marqués	31
2.1 Motivations	31
2.1.1 Des pixels aux objets	31
2.1.2 Cadre stochastique	32
2.2 Processus spatiaux : définitions et notations	32
2.2.1 Processus ponctuels	32
2.2.2 Processus ponctuels marqués	34
2.2.3 Processus spécifié par une densité	35
2.2.4 Processus de Markov	36
2.2.5 Conditions de stabilité	39
2.3 Construction du processus pour l'extraction d'objet	39
2.4 Chaînes de Markov et convergence	41
2.4.1 Définitions et notations	41

2.4.2	Stationnarité	42
2.4.3	Convergence	42
2.5	Échantillonnage	46
2.5.1	Méthodes MCMC	47
2.5.2	Méthodes MCMC pour les processus ponctuels	48
2.5.3	Construction efficace de l'algorithme MCMC à sauts réversibles	54
2.5.4	Quel algorithme choisir ?	54
2.5.5	Sous quelles conditions l'algorithme converge-t-il ?	55
2.5.6	Détecter la convergence en pratique	58
2.6	Optimisation	59
2.6.1	Recuit simulé	60
2.6.2	Schéma de décroissance géométrique	60
2.6.3	Schéma de décroissance adaptatif	61
3	Extraction des réseaux linéiques par processus de segments	63
3.1	Modèles a priori	64
3.1.1	Processus de référence	64
3.1.2	Modèle "Candy"	65
3.1.3	Modèle "Quality Candy"	68
3.1.4	Modèle "IDQ"	73
3.1.5	Noyau de proposition pour un échantillonnage efficace	75
3.1.6	Échantillons des modèles a priori	83
3.2	Incorporation des propriétés radiométriques des données	88
3.2.1	Approche par champ externe	88
3.2.2	Approche bayésienne	91
3.2.3	Noyau de proposition fondé sur les données	93
3.3	Réglage des paramètres	95
3.3.1	Calibrage des paramètres pour une approche par champ externe.	97
3.3.2	Calibrage des paramètres pour le modèle bayésien	99
3.4	Évaluation quantitative des résultats	101
3.5	Choix du modèle a priori	103
3.6	Robustesse des paramètres	108
3.7	Résultats de l'extraction du réseau linéique par processus de segments	111
3.7.1	Approche par champ externe	111
3.7.2	Évaluation de l'apport des pré-calculs	118
3.7.3	Approche bayésienne	121
4	Extension aux lignes brisées	127
4.1	Motivations	127
4.2	Processus de référence	128
4.3	Modèle a priori	129
4.3.1	A priori sur la forme d'une ligne brisée	129
4.3.2	Interactions entre les lignes brisées	130
4.3.3	Densité a priori	132

4.4	Terme d'attache aux données	133
4.5	Échantillonnage approprié	135
4.5.1	Naissance et mort uniforme	135
4.5.2	Naissance et mort de lignes brisées réduites à un segment	136
4.5.3	Ajout et retrait d'un segment	137
4.5.4	Translation d'un point de contrôle	139
4.5.5	Division et fusion de segments	140
4.5.6	Division et fusion de lignes brisées	143
4.6	Résultats	148
5	Extension à l'extraction de réseaux hiérarchiques	153
5.1	Exemple de réseau hiérarchique	153
5.2	Segmentation par champ de Markov	157
5.2.1	Modèle a priori	157
5.2.2	Vraisemblance des observations	158
5.2.3	Optimisation	159
5.2.4	Résultat de l'extraction du réseau par champ de Markov	161
5.3	Modélisation du réseau par processus objet	165
5.3.1	Modélisation hiérarchique du réseau	165
5.3.2	Processus définis dans le voisinage d'un objet déjà détecté	167
5.3.3	Incorporation des propriétés radiométriques	169
5.4	Extraction du réseau par une modélisation hiérarchique	172
5.4.1	Initialisation du réseau	172
5.4.2	Génération de nouvelles branches	177
	Conclusion	183
	Publications	189
	Bibliographie	190

Introduction

L'interprétation des images constitue pour les cartographes un outil très important, et parfois indispensable, pour optimiser le temps passé sur le terrain tout en améliorant d'une façon très sensible la précision du document cartographique final. L'étendue des surfaces à cartographier et les délais de mise à jour font du développement d'outils algorithmiques de détection d'items cartographiques un enjeu important. Nous nous intéressons ici à l'utilisation de données satellitaires et aériennes pour la détection d'objets linéiques. L'item cartographique recherché est constitué des différents réseaux présents dans l'image tels que le réseau routier ou le réseau hydrographique.

De nombreuses méthodes ont été développées pour répondre au problème de l'extraction des réseaux à partir d'images satellitaires et aériennes, notamment dans le cas particulier des réseaux routiers. Une grande partie des travaux sur l'extraction des réseaux se situent dans une problématique d'aide à la saisie des interprètes d'images. Ces méthodes sont semi-automatiques et permettent une extraction précise et efficace d'une route spécifiée par un utilisateur. Mais le gain de productivité apporté par ces méthodes est faible par rapport au travail d'extraction mené par un expert. Nous nous situons ici dans la seconde catégorie d'approches, qui se propose de fournir une extraction complètement automatique du réseau. La qualité de l'extraction se mesure alors plus en termes d'exhaustivité et de limitation des fausses alarmes qu'en termes de précision, comme cela peut l'être pour les méthodes semi-automatiques.

L'objectif de cette thèse est donc l'extraction complètement automatique des réseaux routiers ou hydrographiques à partir d'images satellitaires et aériennes, en minimisant dans la mesure du possible les fausses alarmes et les omissions. Nous proposons de fournir le résultat de l'extraction sous la forme d'une collection d'objets, chaque objet correspondant à une portion de route ou de rivière. Cette représentation par objet permet une interprétation aisée de la scène observée et facilite la correction du résultat par un interprète d'image, si besoin est.

Du point de vue méthodologique, nous nous plaçons dans un cadre de géométrie stochastique et utilisons des processus objet, ou processus ponctuels marqués, comme modèles *a priori*. Ces modèles, d'une utilisation récente en analyse d'image, permettent de bénéficier des propriétés des approches stochastiques du type champ de Markov, tout en manipulant des contraintes géométriques fortes. Ainsi, ils permettent de manipuler

des objets géométriques simples mais génériques, car dépendant d'un certain nombre de paramètres variables. Des interactions entre objets distincts peuvent être prises en compte dans le modèle, ce qui nous permet d'injecter des contraintes sur la géométrie et la topologie du réseau (courbure, densité des embranchements, etc.). De plus, nous disposons d'algorithmes d'exploration de l'espace d'état très utiles dès lors que l'espace est de très grande dimension. Ainsi, un recuit simulé sur un algorithme de type Monte Carlo par Chaîne de Markov (MCMC) permet une optimisation globale sur l'espace des configurations d'objets, indépendamment de l'initialisation.

Dans ce manuscrit, nous proposons tout d'abord une modélisation du réseau linéique présent dans des images aériennes ou satellitaires par une collection de segments, chaque segment correspondant à l'axe central d'une section de route ou de rivière. Le modèle *a priori* choisi est un processus objet dans lequel les objets sont des segments interagissant entre eux. Un premier modèle de ce type, le modèle "Candy", a été introduit par R. Stoica pour l'extraction du réseau routier [Stoica, 2001]. Nous proposons deux extensions de ce modèle qui prennent en compte la qualité des interactions entre les objets du processus de façon continue, afin d'exploiter au mieux les caractéristiques géométriques et topologiques du réseau recherché. Les propriétés radiométriques des données sont prises en compte dans un terme d'attache aux données fondé sur des mesures d'homogénéité et de contraste avec l'environnement. Nous proposons trois approches pour construire ce terme d'attache aux données, la première étant précise et les deux autres efficaces car fondées sur des pré-calculs. Le choix des paramètres du modèle se fait selon la méthode des boîtes qualitatives [Azencott, 1992], assurant que la configuration optimale vérifie de bonnes propriétés. L'optimisation est réalisée par recuit simulé sur un algorithme MCMC à sauts réversibles [Geyer et Møller, 1994, Green, 1995], construit de façon à accélérer la convergence par l'ajout de perturbations pertinentes, telles que la naissance d'un segment dans le voisinage d'un autre segment ou la naissance d'un segment fondée sur les données.

Nous proposons ensuite une extension de cette modélisation à des objets plus complexes. Les objets du processus sont alors des lignes brisées composées d'un nombre inconnu de segments. La connexion entre segments est ainsi incorporée dans la définition même des objets. De plus, les jonctions entre les routes ou les rivières peuvent être modélisées de façon simple via la définition d'une interaction entre deux objets. La complexité des objets du modèle rend nécessaire l'utilisation de perturbations appropriées dans l'algorithme MCMC, telles que le mouvement d'un point de contrôle, l'ajout et le retrait d'un segment, et des perturbations de type "division et fusion" de lignes brisées.

Enfin, nous avons mené une étude sur le cas de l'extraction d'un réseau constitué de fleuves et de leurs affluents. Nous exploitons la structure arborescente de ce réseau en proposant une modélisation hiérarchique du réseau. Cette modélisation nous permet de définir un algorithme récursif de génération de nouvelles branches à partir de celles

qui ont déjà été détectées. Cet algorithme est fondé sur une modélisation des affluents d'un fleuve par un processus objet défini sur la zone d'influence du fleuve considéré.

Ces travaux sont réalisés avec la collaboration du Bureau de Recherche Géologique et Minière (BRGM) qui a partiellement financé cette thèse. Le BRGM nous a fourni plusieurs images satellitaires en vue de l'extraction du réseau hydrographique. Le BRGM a en outre participé à la phase de validation des résultats, notamment en nous fournissant des extractions manuelles des zones traitées.

Concernant l'extraction du réseau routier, nous remercions le Centre National d'Etudes Spatiales (CNES) pour nous avoir fourni des données satellitaires et l'Institut Géographique National (IGN) pour nous avoir fourni des données aériennes.

Ce document est organisé de la façon suivante :

- Nous dressons au **chapitre 1** un panorama de l'extraction de réseaux à partir d'images satellitaires et aériennes.
- Avant de décrire les différents modèles proposés pour l'extraction de réseaux, nous considérons dans le **chapitre 2** le problème plus général d'extraction d'objets à partir de données spatiales. Nous expliquons la pertinence du choix d'une approche par processus objet et donnons les points-clés à connaître pour construire un processus ponctuel marqué en vue d'une extraction d'objets.
- Le **chapitre 3** concerne l'extraction du réseau linéique par processus de segments. Chaque modèle proposé y est détaillé et testé sur des images aériennes et satellitaires, optiques et radar.
- Le **chapitre 4** propose une extension de cette modélisation par l'utilisation d'objets plus complexes : les lignes brisées. Les résultats expérimentaux permettent d'évaluer la pertinence de cette nouvelle modélisation.
- Enfin, le **chapitre 5** concerne l'extraction d'un réseau hiérarchique constitué de fleuves et de leurs affluents sur une image de Guyane fournie par le BRGM.

Chapitre 1

État de l'art en extraction de réseaux

Ce chapitre propose un panorama de l'extraction de réseaux (routes ou rivières) à partir d'images satellitaires et aériennes. Sans être exhaustif, il illustre les diverses méthodes utilisées pour l'extraction des réseaux linéiques et surfaciques par des exemples de références et des exemples originaux. Par analogie, des travaux en imagerie médicale concernant le réseau vasculaire seront également cités. Après une description des caractéristiques principales des réseaux d'intérêt et de leur variabilité, nous présenterons les deux catégories de méthodes proposées dans la littérature : les méthodes semi-automatiques par opposition aux méthodes complètement automatiques.

1.1 Les différents types de réseaux

Avant la présentation des différentes approches proposées pour l'extraction de réseaux, il convient de définir ce que nous entendons par réseau, comment apparaissent les réseaux dans les images, quelles en sont les principales caractéristiques et dans quelle mesure ces caractéristiques varient.

1.1.1 Quels sont les réseaux d'intérêt ?

Un réseau se définit comme un ensemble de lignes qui s'entrecroisent plus ou moins régulièrement. En particulier, nous considérons l'ensemble des réseaux qui apparaissent dans les images aériennes et satellitaires tels que le réseau routier et le réseau hydrographique. Des exemples de réseaux observés en télédétection sont donnés par la figure 1.1. De nombreuses méthodes ont été développées pour répondre à l'extraction de ces items cartographiques, notamment pour le cas particulier des réseaux routiers. L'objectif final de ces approches est souvent la production ou la mise à jour cartographique. En effet, l'étendue des surfaces à cartographier et les délais de mise à jour font de l'extraction de réseaux un enjeu important. C'est également dans ce contexte que nous nous situons et c'est pourquoi la majorité des approches citées dans ce chapitre concernent l'extraction

de ces items cartographiques.

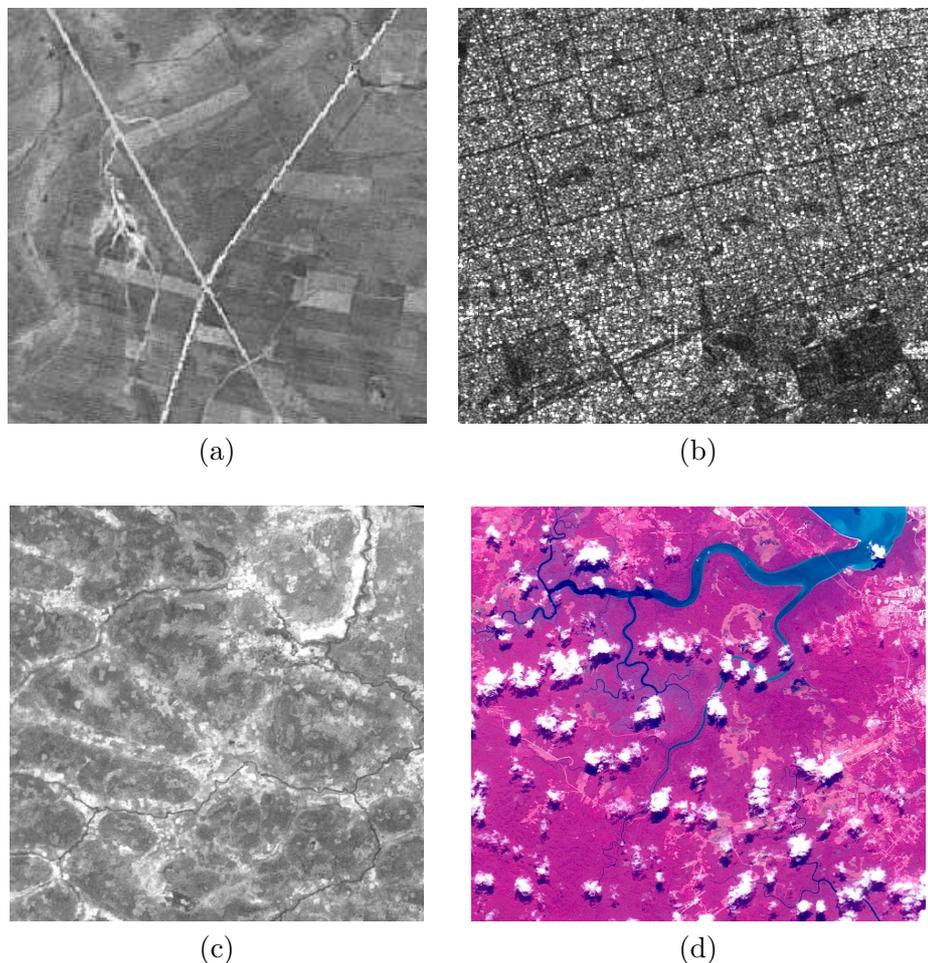


FIG. 1.1 – Exemples de réseaux observés en télédétection : des réseaux routiers sur une image SPOT (a) et sur une image radar ERS (b) ; des réseaux hydrographiques sur des images SPOT dont les forêts galeries, repérables par la végétation en bordure de rivière (c) et des fleuves et leurs affluents (d).

Dans un tout autre contexte mais avec des problématiques similaires, nous citerons également des méthodes d'extraction de réseaux vasculaires en imagerie médicale. Ces réseaux peuvent être observés au travers d'angiogrammes, un angiogramme étant une image radiographique des vaisseaux sanguins du corps ou de toute partie de l'organisme. L'acquisition d'un volume de données (séquences 3D) se fait le plus souvent après l'injection intraveineuse d'un produit de contraste qui accentue la différence de signal entre le compartiment vasculaire et les tissus environnants (graisse, muscle). La matérialisation d'un angiogramme se fait par l'algorithme MIP (Maximum Intensity

Projection) : sélection des voxels d'intensité maximale et fusion de toutes les coupes. Deux exemples d'angiogrammes sont donnés par la figure 1.2. L'extraction des vaisseaux constitue une étape essentielle pour de nombreuses applications pratiques telles que le diagnostic automatique de certaines malformations et le recalage d'image. Les nombreuses méthodes proposées dans ce contexte reposent sur des méthodologies similaires à celles utilisées pour l'extraction des réseaux en observation de la Terre.

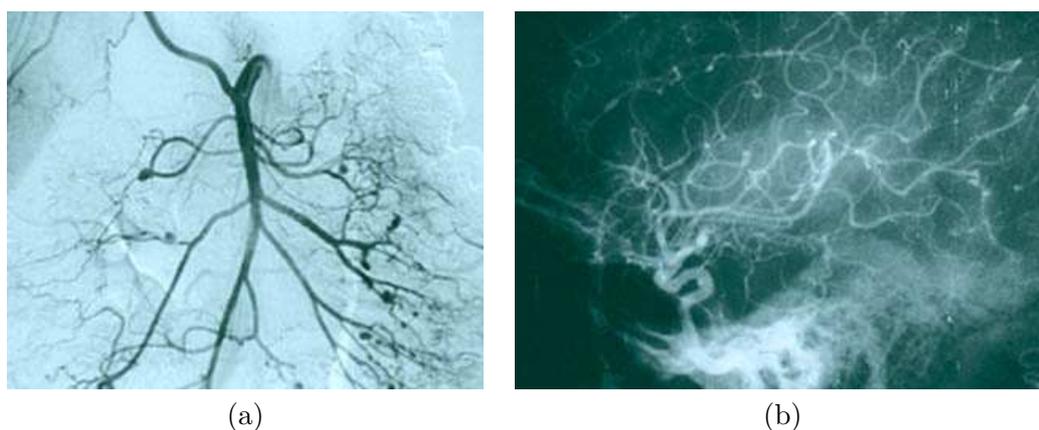


FIG. 1.2 – Les vaisseaux sanguins - (a) angiogramme abdominal - (b) angiogramme du système nerveux central.

1.1.2 Caractéristiques communes

Dans tout problème de reconnaissance de forme, il est essentiel de bien définir l'objet ou les objets que l'on cherche à extraire de l'image. Bien qu'il existe une variabilité importante parmi les réseaux d'intérêt, des caractéristiques majeures peuvent être dégagées. Ainsi, les réseaux sont généralement caractérisés par les contraintes **géométriques** suivantes :

- (G_1) la courbure du réseau est faible (en particulier pour les réseaux routiers) ;
- (G_2) la forme des branches est allongée et de longueur importante ;
- (G_3) l'épaisseur des branches varie peu (voire pas du tout) et de façon progressive.

Des caractéristiques **topologiques** peuvent également être exploitées :

- (T_1) le réseau présente peu d'extrémités libres (i.e. non connectées) ;
- (T_2) le réseau peut présenter des intersections ;
- (T_3) il est peu redondant (pas de superposition mis à part au niveau des intersections).

Du point de vue **radiométrique**, deux hypothèses sont constantes quelque soit le type de réseau à extraire :

- (H_1) le niveau de gris du réseau est localement homogène ;
- (H_2) le réseau contraste fortement avec son environnement.

Certaines méthodes d'extraction utilisent également des informations issues du **contexte**, comme le fait que plusieurs types d'objets peuvent se trouver à proximité des routes, et des informations **fonctionnelles**, comme le fait que les routes relient plusieurs agglomérations ou que les routes sont des voies carrossables.

1.1.3 Variabilité des réseaux

Les caractéristiques présentées dans le paragraphe précédent sont soumises à de grandes variations selon le type d'application, la zone d'observation et le mode d'acquisition des données. Cette variabilité explique en partie la diversité des méthodes proposées pour l'extraction de réseau.

1.1.3.1 Variabilité intrinsèque des différents réseaux

Malgré certaines constantes, les caractéristiques géométriques et radiométriques des différents réseaux sont très variables selon le type d'application. Ainsi, le réseau à extraire pourra être :

- **linéique** comme les axes routiers ou le centre des vaisseaux ;
- **surfacique** comme les routes apparaissant comme des rubans de radiométrie homogène (*cf.* paragraphe 1.1.3.2) ou les vaisseaux sanguins et les fleuves dont la largeur varie progressivement ;
- **hiérarchique** comme les réseaux hydrographiques constitués de fleuves et de leur affluents. Ce type de réseau peut être également qualifié de fractal, un fleuve et ses affluents étant caractérisés par des propriétés invariantes par changement d'échelle.

Les caractéristiques géométriques et topologiques sont plus ou moins pertinentes selon le type de réseau que l'on cherche à extraire. Par exemple, la faible courbure propre aux réseaux routiers ne se retrouve pas dans le réseau hydrographique sinueux présenté dans la figure 1.1(c).

On observe également une forte variabilité entre réseaux routiers. Les propriétés des réseaux routiers varient selon le type de routes que l'on cherche à extraire (autoroutes, routes, rues ou chemins), leur contexte (rural, péri-urbain, urbain ou forestier) et leur date de construction. A titre d'exemple, on constate que les réseaux routiers présents dans les villes américaines diffèrent fortement de ceux présents dans les villes européennes. En effet, les réseaux américains sont caractérisés par une structure de grille comme on peut le voir sur l'image de la ville Mexico présentée dans la figure 1.1(b). Cette particularité peut être exploitée de façon efficace si l'on ne s'intéresse qu'à ce type de réseaux.

1.1.3.2 Variabilité due au mode d'acquisition

L'apparence du réseau dans l'image est directement liée au mode d'acquisition.

Les propriétés radiométriques du réseau et de son environnement sont fortement dépendantes du matériel de prise de vue et de numérisation. Ainsi, un même réseau apparaîtra différemment selon l'angle de prise de vue, le type de capteur, la résolution spatiale ou encore la bande spectrale considérée. Par exemple, le bruit présent dans une image radar diffère fortement de celui présent dans les images optiques. La figure 1.1(b) illustre ces propos par une image radar ERS dans laquelle l'apparence du réseau est affectée par un bruit granulaire appelé chatolement (speckle, en anglais) mais n'est pas sensible aux perturbations atmosphériques comme en imagerie optique (*cf.* figure 1.1(d)).

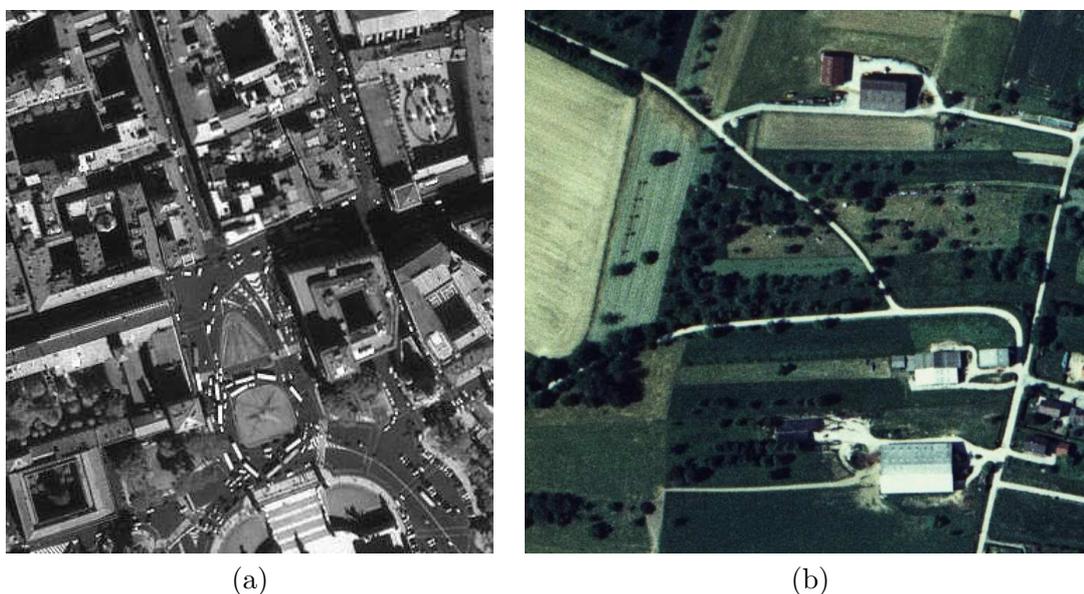


FIG. 1.3 – Exemples de réseaux routiers sur des images haute résolution - (a) réseau routier urbain sur une image IKONOS d'un mètre de résolution © Space Imaging - (b) réseau routier rural sur une image aérienne de cinquante centimètres de résolution fournie par l'IGN.

Les propriétés géométriques du réseau dépendent également du mode d'acquisition des données et plus particulièrement de la résolution du capteur. Ainsi, le réseau routier, qui apparaît comme un réseau linéique de un à trois pixels de large sur les images de faible résolution, apparaît comme un réseau surfacique à haute résolution. Cette variabilité d'apparence induit une variabilité dans les méthodes proposées. Bien que la plupart des travaux en extraction du réseau routier concernent l'extraction des réseaux linéiques, la récente augmentation de la résolution spatiale en imagerie satellitaire a fait naître un intérêt grandissant pour l'extraction précise des routes, celle-ci permettant notamment une reconnaissance des voies détectées (autoroute, rue, chemin). En contrepartie, cette précision rend le problème d'extraction beaucoup plus complexe. En

effet, dans les images haute résolution, un “**bruit géométrique**”, dû à la présence d’objets perturbant l’apparence de la route, s’ajoute au bruit du capteur. La figure 1.3 illustre ces propos par deux images haute résolution : une image satellitaire d’une zone urbaine où des véhicules sont clairement visibles et peuvent nuire à la bonne détection des rues ; une image aérienne d’une zone rurale où des arbres engendrent des occlusions de portions de route.

1.2 Méthodes semi-automatiques

Une grande partie des travaux sur l’extraction des réseaux se situent dans une problématique d’aide à la saisie des interprètes d’image. Les méthodes proposées dans ce cadre sont semi-automatiques : un opérateur doit fournir des points de contrôle pour initialiser un algorithme de suivi (paragraphe 1.2.1), une recherche de chemin optimal par programmation dynamique (paragraphe 1.2.2), un contour actif (paragraphe 1.2.3), ou encore un algorithme fondé sur une analyse de profils transversaux (paragraphe 1.2.4). Ces méthodes permettent en général d’effectuer une extraction rapide et d’une grande précision. Il faut néanmoins se poser la question de savoir si le gain de productivité apporté par les techniques semi-automatiques est significatif par rapport au travail d’extraction mené par un expert. Ces techniques peuvent également être utilisées dans un contexte complètement automatique en utilisant une pré-détection automatique des points d’amorce.

1.2.1 Algorithmes de suivi

1.2.1.1 Filtrage

Le problème de suivi de route peut être vu comme un problème de filtrage. Étant donné une trajectoire dans l’espace d’état et les observations passées, le filtrage correspond à un processus d’estimation de l’état courant sachant son passé et les observations courantes. Ainsi, le suivi de route peut être vu comme la poursuite d’un véhicule supposé circuler au milieu de la chaussée à vitesse constante.

[Vosselman et de Knecht, 1995] proposent d’utiliser un filtrage de Kalman pour estimer la position de la route à chaque instant. L’algorithme est initialisé par un segment fourni par un opérateur à partir duquel un profil caractéristique de la route est défini. La prédiction d’un nouveau segment se fait à partir des estimations de la direction et la courbure de la route. Une mise en correspondance entre le profil caractéristique de la route et le profil courant permet d’évaluer la distance entre les deux profils. La position du nouveau segment est mise à jour selon cette distance. En incorporant les positions obtenues antérieurement, les autres paramètres tels que la direction et la courbure de la route peuvent être également mis à jour.

Le filtrage particulaire, technique non paramétrique fondée sur des simulations séquentielles de type Monte Carlo, fournit un cadre statistique rigoureux pour approximer des distributions *a posteriori*, avec presque aucune restriction sur les composants du modèle. Les distributions *a posteriori* sont estimées par un ensemble d'échantillons (appelés particules) qui sont successivement pondérés (en général, selon la vraisemblance aux observations) et propagés (selon des équations pouvant être non-linéaires). [Pérez *et al.*, 2001] proposent d'appliquer cette technique à l'extraction de routes à partir d'images aériennes. La flexibilité des approches par filtrage particulaire permet d'incorporer des éléments importants dans le modèle *a priori*, tels que la régularité des contours, une structure de ruban de largeur variable et une gestion des coins. L'algorithme développé par [Pérez *et al.*, 2001] permet ainsi de suivre les routes de façon précise même dans le cas d'un changement brusque d'orientation. Cet algorithme peut facilement prendre en compte de multiples informations. Ainsi, [Bicego *et al.*, 2003] modifient cet algorithme pour une extraction dans un contexte urbain. Le terme de vraisemblance est amélioré de façon à prendre en compte des informations de couleur et pour gérer les problèmes d'occlusion. Une automatisation y est proposée par une recherche de points d'amorce utilisant la transformée de Hough. Les résultats obtenus présentent néanmoins de nombreuses omissions dues à la complexité des images.

Cependant, la performance du filtrage particulaire diminue dès lors que l'espace d'état augmente et le support de la vraisemblance diminue. [Vermaak *et al.*, 2003] proposent une alternative reposant sur une approximation variationnelle de la distribution *a posteriori* (dont les paramètres sont estimés de façon itérative selon un algorithme de type EM). Cette méthode est comparée à deux techniques de filtrage particulaire sur un exemple de synthèse et deux applications réelles : le suivi d'objet dans une séquence vidéo et le suivi de contour dans une image unique, dont le suivi de contours parallèles pour l'extraction d'objets ayant une forme de ruban. Les résultats obtenus montrent que la performance de l'algorithme variationnel est supérieure à celle du filtrage particulaire.

1.2.1.2 Théorie de l'information

[Geman et Jedynak, 1996] proposent une approche originale fondée sur la théorie de l'information pour le suivi rapide de routes à partir d'un point d'amorce et d'une direction. Des tests à effectuer sont choisis afin de réduire autant que possible l'incertitude sur l'hypothèse de "position correcte de la route" selon les résultats des premiers tests. Ce choix est effectué de façon dynamique selon la distribution jointe des tests et des hypothèses. Le problème d'optimisation correspond à un problème de minimisation d'entropie (ce qui revient à une minimisation d'incertitude). Cette méthode est performante pour ce qui est du suivi des autoroutes sur de grandes distances (de l'ordre de cent kilomètres sur des images SPOT panchromatiques de 10 mètres de résolution).

1.2.1.3 Suivi par homogénéité directionnelle

Considérant les routes comme des surfaces sans discontinuité apparente plutôt que des lignes, [Airault et Jamet, 1994, Ruskoné, 1996] proposent d'effectuer un suivi reposant sur un critère d'homogénéité directionnelle. Ce critère est fondé sur une hypothèse de variance minimale dans la direction longitudinale de la route. A partir d'un point et d'une orientation d'origine, l'algorithme recherche le chemin optimal parmi un arbre de chemin possible. Chaque chemin est composé de segments de droite choisis localement suivant le critère d'homogénéité. L'inconvénient d'un tel choix est que le suivi peut être fortement dérouté par la présence d'arbres à proximité des routes, le mauvais état de la chaussée, la présence de flaques d'eau... Cela provoque l'apparition à l'intérieur de l'arbre de recherche des tracés qui s'égarer dans des zones localement homogènes comme des champs.

1.2.2 Programmation dynamique

De nombreux travaux d'extraction de réseaux linéiques reposent sur une optimisation par programmation dynamique, technique efficace permettant de minimiser une fonction de coût dans un graphe (en général, en un temps polynomial). Une première étape consiste à définir une fonction de coût. La route correspond alors au chemin de moindre coût partant d'un point initial A et arrivant à un point final B , fixés par un utilisateur ou par une recherche automatique des points extrémaux.

[Fischler *et al.*, 1981] utilisent cette technique d'optimisation pour la détection précise des routes et des structures linéaires dans des scènes rurales. Les auteurs proposent de combiner l'information locale de plusieurs opérateurs de détection de ligne et de contour afin d'établir un masque permettant de restreindre la recherche et de fournir plusieurs cartes de coût définies sur ce masque. Une constante est ajoutée à chaque coût positif afin de minimiser la courbure de la route. L'algorithme de programmation dynamique proposé, appelé F^* , est appliqué à chaque carte fournissant plusieurs chemins optimaux. Le chemin de coût normalisé minimal est alors sélectionné. Cet algorithme, référence dans le domaine de l'extraction de réseaux linéiques, donne de très bons résultats sur des images aériennes de faible résolution.

[Merlet et Zerubia, 1996] proposent une amélioration de cet algorithme avec une modélisation sous-jacente par champs de Markov. Ce formalisme permet d'intégrer simultanément les différentes sources d'information. Le contraste de la route par rapport au fond proche est pris en compte via la définition de potentiel sur des cliques de plus de deux points. De plus, des voisinages de taille supérieure sont considérés afin de mieux prendre en compte la courbure du réseau mais cela reste encore très local du fait du caractère pixélique de la méthode.

[Barzohar et Cooper, 1996] proposent un modèle de route géométrique pour l'extraction de route dans les images aériennes avec une hypothèse de variation faible et

graduelle de la largeur de route. Les points centraux des routes et les distances aux bords associées, le contraste au niveau des bords, les niveaux de gris adjacents aux bords et du fond de l'image sont modélisés par un processus auto-régressif. L'estimation au sens du Maximum A Posteriori (MAP) des routes sachant les points extrémaux est réalisée par programmation dynamique. Une automatisation est proposée via un partitionnement de l'image en fenêtres de taille réduite sur lesquelles une pré-détection des routes candidates est réalisée par estimation locale utilisant la programmation dynamique. Une fenêtre pouvant contenir plusieurs routes (au niveau d'une jonction par exemple), les auteurs proposent d'utiliser un second partitionnement qui correspond à une translation du premier. De bons résultats sont obtenus sur des images de synthèse.

[Grün et Li, 1995] proposent un schéma d'extraction de routes semi-automatique qui combine une transformée en ondelettes et un algorithme de programmation dynamique. La transformée en ondelettes est construite de façon à rehausser le contraste des routes par rapport à leur environnement et est calculée de façon efficace par un algorithme pyramidal. La route est modélisée par modèle générique fondé sur des propriétés géométriques et photométriques. Contrairement aux méthodes décrites précédemment, le chemin n'est plus décrit par une chaîne de pixels mais par un polygone décrit par n points de contrôle. Le polygone est initialisé grossièrement par un nombre réduit de points de contrôle fournis par un utilisateur. La position des n points de contrôle est optimisée par programmation dynamique, chaque point pouvant bouger dans une fenêtre de taille 5×5 dans la direction normale à la courbure. Des nouveaux points sont proposés au centre de chaque nouveau segment de longueur supérieure à un certain seuil. L'optimisation est ensuite réitérée sur le nouveau polygone. Les points non pertinents sont supprimés. Et ainsi de suite jusqu'à convergence. Cette méthode a été appliquée avec succès pour l'extraction de route à partir d'images satellitaires SPOT et aériennes.

1.2.3 Contours actifs

Les contours actifs sont largement utilisés pour tout ce qui concerne l'extraction d'objet dans les images. En effet, c'est une méthode précise et efficace. Les contours actifs doivent néanmoins être initialisés près de l'objet à détecter. Le contour évolue ensuite selon un algorithme de minimisation d'énergie qui favorise en général une faible courbure et un fort gradient au niveau du contour.

[Fua et Leclerc, 1990] utilisent les contours actifs pour l'extraction des routes et des bâtiments. Les contours sont modélisés par des rubans définis par des points centraux successifs et une troisième composante correspondant à la largeur du ruban. Les auteurs montrent que les courbes optimisées sont de bonnes approximations des contours.

[Neuenschwander *et al.*, 1997] présentent une stratégie d'optimisation originale permettant à l'utilisateur de ne fournir que les extrémités du contour à détecter. Le contour est divisé en deux parties : une active et une passive. Au cours de l'optimisation, seule

la partie active prend en compte le potentiel d'attache aux données. La partie active est initialisée au niveau des extrémités du contour et se propage progressivement vers le centre du contour selon le fonctionnement d'une fermeture éclair (ziplock, en anglais).

[Rochery *et al.*, 2003] introduisent une nouvelle classe de contours actifs d'ordre supérieur pour la détection du linéique (routes, rivières, ...). Ces contours actifs reposent sur la définition d'énergies quadratiques, contrairement aux énergies classiquement utilisées qui sont linéaires. Ces énergies permettent de définir des interactions non triviales entre les différents points du contour. Elles donnent naissance à des forces non locales, permettant ainsi d'introduire une information géométrique forte dans le modèle. D'un point de vue algorithmique, ils utilisent une méthodologie par courbes de niveau afin de trouver le minimum de l'énergie, la présence de forces non locales nécessitant une extension des méthodes standard. Cette méthode est nettement moins sensible à l'initialisation que les méthodes usuelles : les contours initiaux peuvent être très distants de la ligne à détecter et ne recouvrir que très partiellement cette ligne ; dans le cas où l'image considérée ne contient qu'un réseau non fermé (par exemple, deux routes qui se croisent), aucune initialisation n'est nécessaire : le contour initial correspondant aux bords de l'image. Cependant, cette méthode est nettement moins efficace en termes de temps de calcul que les méthodes usuelles reposant sur les contours actifs.

1.2.4 Analyse multi-résolution de profils transversaux

La plupart des méthodes décrites précédemment sont applicables au milieu rural, voire péri-urbain, mais montrent leurs limites en environnement urbain où le "bruit géométrique" (*cf.* paragraphe 1.1.3.2) prend toute son ampleur. Dans ce contexte, [Couligner et Ranchin, 2000, Péteri *et al.*, 2001] présentent une méthode pour l'extraction surfacique des réseaux en grille à partir d'images de moyenne et haute résolution. L'intérêt de cette méthode réside dans l'utilisation de plusieurs résolutions pour réduire les effets du bruit. L'utilisateur doit fournir les points extrémaux de la rue à extraire. L'extraction des bords se fait par une recherche des points d'intersections des profils transversaux de la rue provenant de l'image et des deux premières approximations de l'image. Une transformée en ondelettes est utilisée pour extraire (s'ils existent) le terre-plein central et les terre-pleins secondaires. Notons que cette méthode n'est applicable que sur les réseaux de rues rectilignes.

1.3 Méthodes automatiques

L'objectif final des méthodes d'extraction complètement automatique est de remplacer le travail d'extraction mené par un interprète d'image. C'est un problème mal-posé (ne dépend pas de façon continue des données et pas de solution unique) dans lequel il est difficile de trouver le bon compromis entre les fausses alarmes et les omissions. Les méthodes automatiques sont le plus souvent fondées sur une première détection reposant sur une optimisation locale (paragraphe 1.3.1). Ces premières méthodes ne fournissent en général qu'une extraction très grossière du réseau. Elles sont donc suivies d'une

étape de reconstruction du réseau par l'utilisation de méthodes semi-automatiques ou une modélisation par champ de Markov sur graphe (paragraphe 1.3.2). L'extraction du réseau peut être guidée par des données cartographiques (paragraphe 1.3.3). Le problème d'extraction se complexifie lorsque l'on cherche à extraire des réseaux surfaciques à partir d'images haute résolution. Dans ce contexte, des approches multi-résolutions (paragraphe 1.3.4) sont proposées pour réduire l'effet du bruit tout en fournissant une extraction précise. Récemment, des systèmes d'interprétation très complets (paragraphe 1.3.5), intégrant des informations sémantiques, permettent une extraction de réseau dans des scènes complexes. Nous terminerons cet état de l'art avec les processus objet (paragraphe 1.3.6) qui fournissent un formalisme rigoureux pour une extraction du réseau sans étape d'initialisation.

1.3.1 Détection de lignes par optimisation locale

Les méthodes de détection de lignes consistent à rechercher dans l'image les pixels ou les zones présentant les principales caractéristiques des routes, rivières ou autre structure linéaire. Elles reposent sur une optimisation locale, ce qui induit une importante sensibilité au bruit, en particulier pour les images haute résolution. Elles n'assurent donc pas l'exhaustivité ni l'absence de fausses alarmes. Elles sont souvent utilisées en tant que pré-traitement permettant d'extraire les pixels ou les zones ayant une probabilité plus ou moins forte d'appartenir à l'item considéré.

1.3.1.1 Morphologie mathématique

Partant de la constatation que les structures linéaires à détecter sont plus claires que leur environnement, la transformation dite du "chapeau haut de forme" ("top hat", en anglais) peut être appliquée. Dans le cas de structures plus foncées que leur environnement, le "chapeau haut de forme" est appliquée à l'image inversée. Cet opérateur issu de la théorie de morphologie mathématique permet d'extraire les pics d'intensité de l'image. Il consiste en la soustraction de l'image par son ouverture. Une ouverture par un élément structurant B est obtenue par une érosion par B (en chaque pixel, on prend la valeur minimale sur le domaine défini par B) suivi d'une dilatation par B (en chaque pixel, on prend la valeur maximale sur le domaine défini par B). L'ouverture supprime ainsi les pics tout en préservant les vallées. Pour la détection des structures linéaires, l'élément structurant utilisé est lui-même une structure linéaire (conditionnée par la largeur des structures linéaires à détecter). Plusieurs orientations doivent donc être considérées. L'ouverture est obtenue en procédant à des ouvertures successives par l'élément structurant orienté selon chaque direction possible et en assignant à chaque pixel la valeur minimale obtenue sur l'ensemble des ouvertures directionnelles. L'efficacité du traitement dépend de la dynamique de l'image et notamment du bruit puisque les pics du bruit sont détectés par le "chapeau haut de forme". [Destival, 1987, Serendero, 1989] appliquent cet opérateur suivi d'un seuillage sur des images panchromatiques. Le résultat dépend de la largeur et/ou hauteur des sommets à conserver.

[Zhang *et al.*, 1999] présentent une méthode fondée sur la morphologie mathématique pour une extraction automatique du réseau routier dans des images haute résolution. Après une pré-détection grossière des régions contenant des routes, des ouvertures triviales morphologiques sont utilisées successivement pour filtrer les perturbations provoquées par des objets aux caractéristiques spectrales semblables à celles des surfaces routières (par exemple les maisons de même radiométrie que les routes). Pour cela, ils utilisent un concept de granulométrie permettant de déterminer la distribution de taille et de forme des objets dans l'image. Ils présentent leur méthode comme un bon point de départ à une réelle extraction permettant une réduction de l'espace de recherche en fournissant des positions approximatives. Bien qu'adaptée à l'imagerie haute résolution, la méthode reste sensible au problème d'occlusion partielle de la route, due par exemple à la présence d'arbres à proximité des routes.

1.3.1.2 Détecteurs de lignes

Un des opérateurs les plus connus pour la détection de ligne est l'opérateur DUDA, proposé par [Duda et Hart, 1973]. Il utilise quatre masques à convoluer avec l'image correspondant à des structures linéaires de trois pixels de long et les pixels voisins associés. Des scores sont calculés pour chaque masque mesurant conjointement l'homogénéité du niveau de gris le long de la structure et le contraste avec le voisinage. Le score maximal est assigné au pixel central du masque. Cet opérateur est sensible à l'orientation de la route, à la discrétisation, aux changements brutaux d'orientation ainsi qu'au faible contraste par rapport à l'environnement. De plus, cet opérateur détecte les profils de type contour simple et engendre donc un nombre significatif de fausses détections. Le résultat est cependant moins chahuté que dans le cas d'une transformation "chapeau haut de forme". [Fischler *et al.*, 1981] utilisent les scores fournis par cet opérateur en combinaison avec des opérateurs plus sélectifs (peu de fausses alarmes mais moins exhaustifs) pour le calcul des cartes de coût sur lesquelles ils appliquent l'algorithme de programmation dynamique F^* (*cf.* paragraphe 1.2.2). Afin d'éviter la détection des contours conjointement aux lignes, [Roux, 1992] propose une version modifiée de l'opérateur DUDA en ne tenant compte que du groupe de voisins le moins contrasté avec l'axe central dans le calcul du score. [Gurney, 1980] propose un détecteur de ligne pour l'extraction de la Tamise à partir d'une image Landsat. Celui-ci repose sur 14 masques et permet de considérer toutes les lignes de largeur 1 et de longueur 3. Ces masques sont efficaces pour détecter les lignes droites et les courbes faibles mais ne le sont pas pour les courbes prononcées et les coins.

Certains auteurs se sont ensuite intéressés au cas des lignes épaisses. Un premier choix consiste à étendre les masques utilisés pour les lignes fines à des masques plus grands permettant la détection des lignes plus épaisses. Ainsi, [Dhérété et Desachy, 1999] étendent l'opérateur DUDA à des fenêtres de dimension 7×7 pour l'extraction de structures plus épaisses. [Huber et Lang, 2001] proposent une extension de l'opérateur DUDA pour la détection de routes à partir d'images SAR haute résolution. Les pixels sont rem-

placés par des sous-régions et le score est adapté aux images SAR. [Wang *et al.*, 1996] étendent le détecteur de ligne proposé par [Gurney, 1980] pour la détection du réseau routier urbain. Ce type d'opérateur étant très sensible au bruit dans les images, [Wang *et al.*, 1996] proposent de réduire le bruit par un pré-traitement de l'image reposant sur des filtres morphologiques.

Un autre choix consiste à s'intéresser à la texture plutôt qu'à des simples différences de niveau de gris. Ainsi, [Haverkamp, 2002] propose un détecteur de route dans le milieu urbain fondé sur la texture. Plus précisément, ce détecteur est fondé sur l'analyse de signatures angulaires de texture contenues dans une image panchromatique IKONOS. La signature angulaire de texture associée à un pixel est constituée des variances (après normalisation) calculées sur un ensemble de fenêtres rectangulaires autour du pixel en question. Afin de réduire les surdétectations, un masque de végétation est construit à partir d'une image multispectrale IKONOS. Un pixel n'appartenant pas au masque de végétation sera classifié en tant que "route" si plusieurs minima locaux significatifs sont observés dans la signature de texture. Un intérêt de cette approche est qu'elle permet de distinguer les pixels correspondant à des intersections des autres pixels à partir du nombre de minima observés. Cette classification est suivie d'une étape "haut niveau" permettant de passer des pixels classifiés au réseau composé de segments connectés entre eux. La méthode a été testée sur une image de San Diego et donne de bons résultats en termes d'identification (environ 80% de routes identifiées) et de localisation (écart-type d'un pixel). Néanmoins, cette méthode n'est applicable que dans le cas de réseaux en forme de grille tels que ceux rencontrés dans les villes américaines et est sensible au bruit géométrique dû par exemple à la présence de voitures sur les routes.

1.3.1.3 Banc de filtres

Les méthodes par banc de filtres permettent une analyse multi-échelle et ont donc l'avantage de fournir une extraction des lignes d'épaisseur variable, ce qui s'avère très utile dès lors que le réseau présente une variabilité non négligeable d'épaisseur (vaisseau sanguin, fleuves, etc.).

[Koller *et al.*, 1995] utilisent une combinaison non linéaire de filtres linéaires pour la segmentation et la description locale de structures linéaires en deux et trois dimensions. Le banc de filtre permet de détecter les structures allongées et symétriques tout en supprimant la réponse des contours simples. Un maximum aigu est obtenu le long du profil des structures linéaires et à travers le changement d'échelle. La réponse finale ne dépend pas de la largeur. Ce filtrage ne nécessite pas de paramétrisation et fournit une description locale du contraste, la position de la ligne centrale, la largeur et l'orientation de la ligne. La nature générique de cet algorithme est montrée au travers de plusieurs applications dont la détection du réseau routier sur une image aérienne et la segmentation de vaisseaux sanguins cérébraux à partir de données volumiques obtenues par résonance magnétique (IRM). Cependant, l'utilisation de filtres de détection de contour directionnels rend la méthode coûteuse en temps de calcul.

[Poli et Valli, 1996] utilisent un ensemble de filtres linéaires pour la détection des vaisseaux sanguins en imagerie médicale. Les filtres, fondés sur des noyaux Gaussiens directionnels, sont construits de façon à différencier les diverses épaisseurs et orientations des vaisseaux avec une efficacité maximale. Cette méthode permet ainsi de réaliser la détection des vaisseaux en temps réel. Les sorties des différentes orientations et échelles sont intégrées et validées de façon à interdire les structures différentes des vaisseaux. La segmentation est réalisée par un seuillage par hystérésis [Canny, 1986]. Les résultats obtenus sur des angiogrammes coronaires sont prometteurs.

1.3.1.4 Géométrie différentielle

Considérant l'image comme une hypersurface (ou plus simplement comme un relief), certains auteurs proposent d'en extraire les lignes en utilisant les propriétés géométriques différentielles de cette hypersurface. L'extraction de lignes revient à localiser les positions des crêtes et des ravins dans l'hypersurface. Les points de crête (resp. ravins) sont définis comme les maxima (resp. minima) locaux des principales courbures de l'hypersurface. Pour localiser les points de crête, on s'intéresse aux dérivées premières et secondes en chaque point de l'image. Les lignes centrales des routes, rivières ou vaisseaux sanguins sont ensuite obtenues en reliant les points de crêtes. Les points forts de ces modèles sont leurs bases mathématiques, l'excellente localisation des points de crête et la non-détection des contours. De plus, les images 2D et 3D sont traitées de façon similaire, étant modélisées respectivement par des hypersurfaces 3D et 4D. Ces méthodes sont néanmoins connues pour leur sensibilité au bruit.

Pour extraire les points de crête, [Eberly *et al.*, 1994] donne une formule explicite de la courbure et de sa direction et recherchent les maxima dans l'image de courbure. Cependant, cette méthode échoue dès que les crêtes sont plates.

[Wang et Pavlidis, 1993] proposent une détection des arrêtes et ravins en utilisant une approximation locale de l'hypersurface par son polynôme de Taylor. La direction de la ligne est déterminée à partir du Hessien du polynôme et les points de la ligne sont les pixels vérifiant une forte dérivée seconde directionnelle perpendiculaire à la direction de la ligne. Les lignes sont détectées avec une précision sous-pixélique sans avoir à construire de filtres directionnels spécialisés. Cependant, l'approximation par coefficients polynomiaux sur des masques conduit à de multiples réponses pour une seule ligne et ne permet pas la détection de ligne d'épaisseur supérieure à la taille des masques.

C'est pourquoi [Steger, 1996, Steger, 1998, Lindeberg, 1998] proposent d'utiliser des masques gaussiens pour la détection des crêtes. Ainsi, en sélectionnant une variance appropriée σ , ces masques peuvent être calibrés pour une certaine épaisseur. Pour détecter les lignes d'épaisseur arbitraire, σ peut être sélectionnée en chaque point en itérant sur l'espace d'échelle. [Steger, 1996] propose une analyse multi-échelle de lignes de différents types (parabolique, en forme de barre) permettant une sélection de

paramètres appropriés pour l'algorithme de détection de point de crête. Un algorithme de chaînage permet de relier les points de crête en lignes et jonctions en préservant le maximum de points. Parallèlement à l'extraction des points de crête, [Steger, 1998] propose une extraction des points de contour par une recherche localisée autour des points de crêtes utilisant des masques de taille réduite. Cela rend la méthode efficace en terme de temps de calcul. Le biais entre la ligne extraite et la position des contours peut être prédite analytiquement grâce à une modélisation explicite de l'interaction entre la ligne et la position des contours. Par retrait de ce biais, les contours des lignes sont donc extraits avec précision. L'algorithme a été testé avec succès sur des images aériennes et médicales contenant des lignes de différentes largeurs et asymétries. Les positions des lignes et contours correspondent à des items sémantiques dans l'image : à l'axe principal et les bords des routes dans les images aériennes et à des vaisseaux dans les images médicales. Cependant, l'algorithme détecte également les toits des maisons dans les images aériennes.

1.3.1.5 Les réseaux de neurones

Le problème d'extraction de structures linéaires peut être vu comme un problème de classification où deux labels peuvent être assignés aux pixels de l'image : le label "fond" et le label "item cartographique". Dans ce cadre, les réseaux de neurones peuvent s'avérer un outil puissant de classification puisqu'ils ne supposent aucune connaissance sur la distribution sous-jacente et permettent d'éviter les phases de rehaussement et de seuillage. De plus, une fois entraînés ils peuvent réaliser l'extraction en temps réel.

[Bhattacharya et Parui, 1997] proposent d'utiliser un réseau de neurones multicouches pour la détection de routes. Les entrées du réseau sont des fenêtres de taille $n \times n$ centrées sur chaque pixel p de l'image. Le réseau calcule par lui-même des valeurs discriminantes dans les noeuds de l'unique couche cachée à partir des n^2 entrées, puis classe le pixel p à partir des noeuds cachés dans le noeud de sortie. Ce réseau est entraîné par un algorithme de type rétro-propagation à partir de quelques échantillons de type "fond" et "route" sélectionnés par un opérateur humain à partir d'une image provenant d'un certain type de capteur. Ils proposent une amélioration de l'algorithme de "rétro-propagation" usuel pour diminuer le temps de calcul en utilisant des taux d'apprentissage adaptatifs. Le réseau entraîné est ensuite utilisé pour l'extraction de routes dans d'autres images provenant du même capteur. Les auteurs insistent sur le fait que chaque sous-classe de route (autoroute, route urbaine, etc.) et de fond (champs, lac, etc.) doit être représentée dans les échantillons utilisés dans la phase d'apprentissage pour que le réseau puisse classifier correctement tout type de route dans la phase de reconnaissance. La taille des fenêtres est également cruciale : elle doit être juste assez large pour contenir une route et quelques pixels voisins correspondant au fond de l'image. Enfin, la direction des routes dans l'ensemble d'apprentissage peut influencer de façon non négligeable sur la performance de l'algorithme. Pour réduire la dépendance à la direction des routes, les auteurs présentent chaque échantillon selon quatre orientations différentes dans la base d'apprentissage, ainsi que leurs images miroir. Les résultats de

la classification sur des images SPOT et IRS sont relativement bons malgré quelques interruptions et quelques fausses alarmes. Remarquons néanmoins que le réseau n'a été appliqué qu'à des images de taille 128×128 issues de l'image dans laquelle les échantillons d'apprentissage ont été sélectionnés. Qu'en est-il avec une autre image ?

1.3.2 Couplage d'algorithmes

Les algorithmes de reconstruction du réseau reposent généralement sur une initialisation fournie par un algorithme d'optimisation locale. Ainsi, les méthodes semi-automatiques peuvent être utilisées dans un cadre complètement automatique et les combinant avec des détecteurs de bas niveau. Cette reconstruction de réseau à partir d'une pré-détection peut également être réalisée via des cartes auto-organisatrices ou une modélisation globale du réseau par champ de Markov sur graphe.

1.3.2.1 Méthodes semi-automatiques pour la reconstruction de réseau

Les méthodes semi-automatiques décrites dans le paragraphe 1.2 ont l'avantage de reposer sur une optimisation globale et d'incorporer des informations sur la topologie du réseau (contrainte de faible courbure, continuité). Par conséquent, de nombreux auteurs proposent d'étendre ces méthodes à des méthodes complètement automatiques utilisant une pré-détection imprécise fondée sur une optimisation locale. Une détection automatique de points d'amorce est donc souvent proposée [Zlotnick et Carnine, 1993, Barzohar et Cooper, 1996, Bicego *et al.*, 2003] ou même seulement suggérée [Fischler *et al.*, 1981].

Les méthodes semi-automatiques peuvent aussi être employées en tant que post-traitement par regroupement perceptuel, comme cela est fait pour le groupement de contours par [Urago *et al.*, 1994]. Concernant la détection du réseau routier, [Serendero, 1989] propose d'utiliser des algorithmes de chaînage et de prolongement des éléments détectés par programmation dynamique afin d'améliorer le résultat d'une détection par "chapeau haut de forme".

Dans un autre contexte, [Bobillet *et al.*, 2003] proposent d'utiliser un réseau de contours actifs pour la détection de rang de cultures en télédétection haute résolution. L'objectif final est de fournir une image de la vigueur de la végétation à partir de mesures effectuées sur les rangs détectés. L'extraction des rangs se doit donc d'être précise, ce qui justifie le choix d'une modélisation par contours actifs connus pour leur précision. Ce réseau est initialisé à partir d'une estimation grossière de l'orientation des rangs et leur fréquence dans le domaine spectral après une transformée de Fourier. La détection et la localisation du pic principal permettent d'obtenir ces deux paramètres. L'optimisation fondée sur une modélisation par un réseau de contours actifs (composé de plusieurs lignes quasiment parallèles) abouti à des résultats très proches des rangs réels dans le cas de rangs rectilignes, non-enherbés, et sans ombres portées.

1.3.2.2 Cartes auto-organisatrices

[Doucette *et al.*, 2001] proposent d'utiliser une méthode inspirée des cartes auto-organisatrices (ou Self Organizing Maps, SOM) pour extraire l'axe central des routes sur des images hyperspectrales haute résolution. Les cartes auto-organisatrices correspondent à une classe de réseaux neuronaux permettant de visualiser les relations non-linéaires d'un ensemble d'exemples dans un contexte d'apprentissage non supervisé. Proposé par Kohonen en 1982, l'algorithme d'apprentissage répond à un double objectif de construction de groupes de données homogènes et de représentation graphique de ces groupes : à la tâche de groupement (clustering, en anglais) s'ajoute une problématique de préservation d'information topologique. L'algorithme proposé par [Doucette *et al.*, 2001] prend en entrée les coordonnées des pixels étiquetés en tant que "route" à partir d'une classification supervisée de l'image hyperspectrale. La méthode a l'avantage de s'affranchir de la définition des contours des routes tout en fournissant une localisation précise de l'axe central. Bien que les auteurs situent leur approche dans la classe des méthodes semi-automatiques du fait de l'intervention d'un opérateur humain dans la phase de pré-détection, cette méthode pourrait être facilement utilisée dans un contexte d'extraction complètement automatique.

1.3.2.3 Champs de Markov sur graphe

Les champs de Markov sur graphe sont un outil puissant pour la détection automatique de réseau. Le réseau est modélisé dans sa totalité par un graphe où chacun des arcs correspond à une section du réseau. On rentre dans le cadre d'une approche objet par opposition aux approches par champ de Markov usuelles, dans lesquelles le champ est défini sur une grille de pixels. Tout en bénéficiant des bonnes propriétés des approches par champ de Markov, comme la robustesse au bruit, cette modélisation permet d'introduire des contraintes géométriques fortes sur le réseau. L'inconvénient de cette méthode est qu'elle nécessite une phase d'initialisation exhaustive : un arc non détecté dans la phase d'initialisation ne pourra pas l'être par la suite.

Un premier choix consiste à initialiser l'algorithme par une pré-détection obtenue par une optimisation locale, comme par exemple dans [Tupin *et al.*, 1998] où une détection des structures linéaires initialise un champ de Markov sur graphe pour la détection des routes en imagerie radar. Remarquons que les liens reliant les structures linéaires détectées sont incorporés dans le graphe ce qui permet de combler certaines omissions. De plus, la méthode permet d'éliminer de nombreuses fausses détections.

[Géraud, 2003] propose quant à lui une phase d'initialisation fondée sur des pré-traitements morphologiques. Plus exactement, un traitement de bas-niveau fournissant pour chaque pixel un potentiel d'appartenance à une route est tout d'abord appliqué. Puis, il propose de réaliser une sur-segmentation de l'image par la méthode de "ligne de partage des eaux" ("watershed", en anglais) appliquée sur l'image de potentiel (filtrée pour réduire la présence de minima locaux). Un champs de Markov est ensuite défini

sur un graphe d'adjacence de courbes, calculé sur les lignes de partage des eaux et supposé contenir le réseau routier. L'intérêt de cette initialisation est qu'elle est fondée sur des primitives de haut niveau (approche région).

1.3.3 Apport de données cartographiques

L'extraction de réseaux guidée par des données cartographiques est une tâche ardue du fait du caractère exogène des différentes sources. En effet, les cartographes peuvent parfois introduire des distortions dans les cartes dans un souci de simplification et/ou pour souligner certains détails, comme par exemple les lacets d'une route de montagne. Ces distortions sont communément appelées erreurs de généralisation. Deux stratégies sont proposées pour résoudre le problème d'extraction de route guidée par un réseau routier imprécis. Soit le réseau initial est recalé comme un tout avec l'image, soit il est divisé en une collection d'objets linéaires, où chaque objet est utilisé pour initialiser un algorithme d'extraction de route, puis le réseau est reconstruit à partir de la topologie initiale du réseau. La première approche est utilisée pour le recalage de réseau à basse résolution alors que la seconde est utilisée dès lors qu'il s'agit de travailler avec des images haute résolution.

[Relier *et al.*, 2002] se situent dans le cadre d'un recalage direct du réseau sur des images satellitaires basse résolution. La modélisation du réseau se fait par champs de Markov sur graphe. Les sommets du graphe sont des points de jonction entre deux routes ou les points de forte courbure et les arrêtes sont les routes. Ces attributs sont considérés comme des variables aléatoires et l'objectif est de trouver les attributs qui minimisent une fonction de coût fondée sur des contraintes topologiques, les données de télédétection et la distance entre le réseau détecté et le réseau issu de la carte. La méthode permet de réaliser un recalage non rigide du réseau routier de la carte sur une image satellitaire en éliminant tout effet de généralisation. En effet, le recalage des arrêtes se fait par un algorithme de programmation dynamique [Merlet et Zerubia, 1996] et est donc très précis. L'inconvénient majeur de cette méthode est un temps de calcul important dû au fréquent calcul de chemin entre deux noeuds.

[Dhéréte et Desachy, 1999] proposent d'utiliser la logique floue pour résoudre le problème de l'imprécision lors du recalage automatique des objets géographiques linéaires issus de bases de données sur des images SPOT. Une première détection de ligne est obtenue en fusionnant (pour obtenir un résultat plus fiable) les résultats de plusieurs algorithmes reposant sur une optimisation locale appliqués à différents canaux SPOT. Puis une fusion de type Dempster-Shafer permet d'apporter une information sémantique à partir des couches thématiques de la base de données. Enfin, un recalage est réalisé pour un objet cartographique donné en utilisant l'algorithme de programmation dynamique proposé par [Fischler *et al.*, 1981] suivi d'un algorithme de contour actif pour lisser le premier résultat selon une contrainte de faible courbure.

[Guigues et Vilgino, 2000] présentent une approche originale d'extraction de réseaux

routiers à partir d'images aériennes guidée par un réseau routier imprécis. Ils se situent dans le deuxième type d'approche et se focalisent sur la procédure d'extraction de routes. Celle-ci se décompose en deux étapes. La première est une étape de détection de structures parallèles via une simulation de propagations et de réflexion de rayons lumineux. La deuxième est une étape de construction de routes hypothétiques par une technique de programmation dynamique utilisant le parallélisme précédemment détecté. Un vote est ensuite utilisé pour sélectionner le parcours le plus probablement emprunté par la route. Ce système fonctionne correctement en zone rurale et se révèle prometteur pour l'interprétation de situations complexes, en particulier pour les routes à plusieurs voies.

1.3.4 Approche multi-résolution

Les images haute résolution permettent la détection d'éléments qui étaient jusqu'à présent difficiles à extraire. Cependant cette précision rend la tâche d'extraction complexe. Notamment, un bruit géométrique inhérent à la scène observée (par exemple, dû à la présence d'arbres à proximité des routes) vient s'ajouter au bruit du capteur. Afin de réduire cette sensibilité au bruit, certains auteurs proposent d'utiliser des techniques multi-échelles, par exemple en fusionnant le résultat de l'extraction de l'axe central à basse résolution, moins sensible au bruit géométrique, avec le résultat de l'extraction des bordures de route à haute résolution fournissant une meilleure localisation du réseau [Heipke *et al.*, 1995, Baumgartner *et al.*, 1996].

[Heipke *et al.*, 1995] proposent de réaliser une première détection à basse résolution qui est fondée sur l'hypothèse usuelle de fort contraste entre les routes et le fond, et le fait que les routes présentes sont claires (\sim intensité élevée). Un squelette est ensuite extrait des zones de pixels répondant bien aux tests. Puis, un chaînage de pixels est effectué, permettant d'éliminer certaines fausses alarmes. Cette partie est très rapide mais il reste des problèmes au niveau des jonctions et les contours de maisons sont détectés. C'est là qu'intervient le modèle haute résolution fondé sur les hypothèses de contours parallèles et d'homogénéité interne. Ce modèle permet de réaliser une détection des contours approximatés par des polygones, une vérification du parallélisme et une extension des contours parallèles là où il y a poursuite d'homogénéité (notamment au niveau des embranchements où il n'y a plus de contours). Cette deuxième extraction engendre de nombreuses omissions, notamment quand le bruit géométrique perturbe la détection. Une combinaison des deux résultats est proposée.

[Baumgartner *et al.*, 1996] proposent une méthode similaire avec une détection des axes des routes à basse résolution par géométrie différentielle [Steger, 1996]. Une détection des marques au sol à haute résolution est proposée.

[Baumgartner *et al.*, 1999] combinent cette approche multi-résolution avec l'utilisation d'une information contextuelle. Le réseau est décrit par des objets (sections, intersections,...), les relations entre ces objets et les autres objets (bâtiments, arbres,...).

Les bords des routes sont obtenus en fusionnant les résultats de l'extraction de ligne à basse résolution et de contour à haute résolution. Les sections et les intersections sont construits à partir de cette extraction. La reconstruction se fait ensuite par groupement de segments et fermeture des espaces entre segments en utilisant une information contextuelle. Une segmentation en régions selon le contexte (rural, forêt ou péri-urbain) est utilisée pour restreindre les interactions possibles avec les objets du fond. Le procédé donne de bons résultats en milieu rural mais pour les milieux péri-urbains ou forestiers les auteurs avouent qu'il serait nécessaire d'introduire plus d'information contextuelle.

[Laptev *et al.*, 2000] proposent d'utiliser l'extraction de l'axe central à basse résolution pour initialiser une extraction à haute résolution par contours actifs. Les fausses alarmes sont retirées à partir d'une évaluation de la largeur du contour à haute résolution. Un post-traitement utilisant une optimisation par contours actifs permet de combler les zones d'occlusion. L'optimisation est réalisée par une stratégie "ziplock" utilisant comme initialisation les points extrémaux des routes déjà détectées et des hypothèses de faible courbure et de même largeur que celle des routes adjacentes. Les résultats obtenus sur des images de scènes rurales sont d'une grande précision (sous-pixélique) et montrent la faible sensibilité de la méthode à la présence d'occlusions.

1.3.5 Systèmes d'interprétation

Les systèmes d'interprétation se fondent sur le principe de pensée d'un interprète d'image face à une recherche d'objets dans les images. Ainsi, l'extraction du réseau est réalisée via une interprétation de la scène observée.

Ainsi, [Ruskoné, 1996] propose un système d'extraction du réseau routier reposant sur deux phases : une phase ascendante guidée par les données et une phase descendante guidée par l'objectif. La phase ascendante, reposant sur des critères de bas niveau, se décompose en trois étapes usuelles pour ce qui est de l'extraction automatique de réseaux :

- la détection des amorces ;
- le prolongement des routes par un algorithme de suivi ;
- la reconstruction du réseau par groupement perceptuel.

La phase descendante est une étape de validation reposant sur des modules successifs de haut niveau :

- la "valuation" qui consiste à décomposer le réseau en segments et à attribuer à chaque segment un vecteur de mesures (variance et moyenne radiométrique, courbure, etc.) ;
- la "pré-validation" qui permet de formuler des hypothèses quant à la nature des segments (arbre, route, etc.) ;
- le "validation" qui permet de valider la présence des segments dans le réseau par une analyse des hypothèses.

Cette étape de validation passe par une interprétation locale du contexte qui par la reconnaissance d'un objet donné permet de déduire la position de la route.

[S.Hinz et Baumgartner, 2003] proposent d'incorporer une connaissance détaillée des routes et de leur contexte dans un système d'extraction automatique des routes sur des images aériennes haute résolution. Les stratégies automatiques d'extraction sont adaptées en fonction de leur contexte qui peut être urbain, forêt ou rural comme cela avait été proposé précédemment dans [Baumgartner *et al.*, 1999]. Les modèles utilisés prennent en compte des informations sémantiques de type marquage au sol ou présence de véhicule. L'extraction incorpore des composants pour l'auto-évaluation qui évaluent la pertinence des hypothèses pour la suite du traitement. Plusieurs vues de la même scène permettent l'exploitation des redondances, la prédiction et le traitement des occlusions, et la description de la scène par objet dans l'espace 3D. Les résultats présentés sont relativement bons (75% des routes sont détectées et 5% de fausses alarmes) étant donnée la complexité des scènes traitées, notamment lorsque l'apparence des routes est fortement perturbée par d'autres objets.

[Zhang, 2004] propose un système de reconstruction 3D du réseau routier à partir d'images stéréo aériennes, se focalisant sur les aires rurales. Le système incorpore des informations sémantiques, contextuelles, des règles et des modèles pour restreindre l'espace de recherche et traite les sous-classes de route de façon spécifique à chacune. Une vérification de la plausibilité des hypothèses est effectuée pour obtenir des résultats fiables. Un ensemble d'outils de traitements de données est utilisé pour extraire des indices variés sur la présence de routes. Ces indices et les sources d'informations disponibles sont fusionnés de façon complémentaire et redondante pour permettre de traiter les erreurs et résultats incomplets. Le fait de travailler avec des images stéréo permet de générer directement les hypothèses de route dans l'espace 3D. Le système permet de reconstruire plus de 90% des routes sur des zones rurales avec de bons résultats même sur des zones où la présence d'ombres et d'occlusion rend la détection problématique.

1.3.6 Processus objet

D'une utilisation récente en analyse d'image, les processus objet (ou processus ponctuels marqués) permettent de bénéficier des propriétés des approches stochastiques, du type champs de Markov, tout en manipulant des contraintes géométriques fortes. Ce formalisme issue de la géométrie stochastique permet de déterminer simultanément le nombre d'objets dans la scène observée, leurs positions et leurs caractéristiques (longueur, orientation,...) dans un cadre rigoureux. De façon similaire aux approches par champs de Markov sur graphe, cette approche repose sur une modélisation de la scène par une collection d'objets et une minimisation globale d'une énergie définie sur l'ensemble du réseau. La principale différence avec les champs de Markov sur graphe vient du fait que les positions des objets et le nombre d'objets évoluent au cours de l'algorithme. Cette approche permet donc de se soustraire à la phase d'initialisation ou, s'il y en a une, de réduire la sensibilité à l'initialisation.

Récemment, [Stoica, 2001, Stoica *et al.*, 2004] ont introduit un tel processus, ap-

pelé “Candy”, pour l'extraction non supervisée du réseau linéique à partir d'images satellitaires ou aériennes. Ce modèle est fondé sur l'idée qu'un réseau routier peut être assimilé à une réalisation d'un processus Markov objet, où les objets correspondent à des segments en interaction. Outre les propriétés géométriques du réseau (connectivité, continuité, etc.), prises en compte dans un terme *a priori*, les propriétés radiométriques et de texture sont prises en compte dans un terme d'attache aux données. L'optimisation du modèle se fait par un algorithme de type Monte Carlo par chaîne de Markov à sauts réversibles [Geyer et Møller, 1994, Geyer, 1999, Green, 1995]. Aucune initialisation n'est nécessaire. Les résultats sont prometteurs et nous proposons de poursuivre dans cette voie en proposant :

- des améliorations faisant intervenir des coefficients mesurant la qualité des interactions entre objets ;
- des extensions à d'autres objets ;
- la construction de noyaux de proposition pertinents pour diminuer le temps de calcul (principal inconvénient de la méthode).

Chapitre 2

Extraction d'objets par processus ponctuels marqués

Notre principal objectif est l'extraction du réseau linéique ou surfacique (routes, rivières) à partir d'images aériennes ou satellitaires sous la forme d'une collection d'objets, chaque objet correspondant à une portion de route ou de rivière. Dans ce chapitre, nous considérons le problème plus général d'extraction d'objets à partir de données spatiales : cellules à partir d'images microscopiques, bâtiments à partir de modèles numériques d'élévation, arbres à partir d'images aériennes, etc. Nous expliquerons dans un premier temps la pertinence du choix d'une modélisation par processus ponctuels marqués. Nous donnerons ensuite les points-clés à connaître sur les processus spatiaux pour comprendre comment construire un processus ponctuel marqué pour l'extraction d'objets, *i.e.* comment exploiter les données radiométriques et les contraintes topologiques et géométriques caractérisant les objets que l'on cherche à extraire, tout en respectant des conditions de stabilité nécessaire à la bonne définition du processus. Enfin, des techniques d'échantillonnage et d'optimisation sont données après un bref rappel sur les chaînes de Markov.

2.1 Motivations

2.1.1 Des pixels aux objets

Les processus ponctuels marqués, ou processus objet, permettent une modélisation de la scène observée par une collection d'objets. Cette approche objet fournit divers avantages qui sont désormais d'actualité avec la récente possibilité de travailler sur des images haute résolution. En effet, il est désormais possible de détecter des objets jusqu'à présent non détectables et de le faire avec une précision de l'ordre du mètre voire du centimètre. Mais cette même précision induit un bruit géométrique inhérent à la scène observée, ce bruit géométrique prenant toute son ampleur en zone urbaine. Pour répondre à ces nouvelles problématiques d'extraction à partir d'images haute résolution, une modélisation au niveau du pixel n'est plus pertinente. On préférera

donc une modélisation de la réalité sous-jacente observée par des objets. Les processus objet fournissent un formalisme rigoureux pour cela. De fortes contraintes géométriques et topologiques peuvent ainsi être introduites :

1. sur la **forme des objets**, en manipulant des objets géométriques simples mais génériques car dépendant d'un certain nombre de paramètres variables ;
2. sur l'**agencement des objets** dans la scène, au travers de la prise en compte d'interactions entre objets distincts.

L'exploitation de telles contraintes permet d'être plus robuste au bruit que les approches pixéliques. Nous verrons par la suite que la construction d'un processus exploitant ces contraintes se fait aisément via la définition d'un processus de points marqués de référence et d'une densité par rapport à ce processus. De plus, un terme de vraisemblance ou plus généralement d'attache aux données, pourra être calculé de façon directe selon l'appartenance d'un pixel à tel ou tel objet.

2.1.2 Cadre stochastique

En plus des avantages d'une approche objet, une modélisation par processus Markov objet bénéficie des avantages d'un cadre stochastique. En effet, les modèles stochastiques sont connus pour leur robustesse au bruit. De plus, les interactions entre objets peuvent être incorporées dans le modèle à travers la définition de probabilités conditionnelles locales faciles à calculer. Du point de vue algorithmique, on dispose d'algorithmes d'exploration de l'espace d'état qui s'avèrent très utiles lorsque l'espace d'état est grand. Ainsi, un recuit simulé sur un algorithme de type Monte Carlo permet une optimisation globale sur l'espace des configurations d'objets indépendamment de l'initialisation. En plus de l'échantillonnage, les techniques de Monte Carlo par chaîne de Markov permettent de faire de l'estimation de paramètres.

2.2 Processus spatiaux : définitions et notations

2.2.1 Processus ponctuels

Considérons tout d'abord le cas plus simple où l'objectif consiste à décrire la scène par un ensemble non ordonné de points de \mathbb{R}^d :

$$\mathbf{x} = \{x_1, \dots, x_n\}, n \in \mathbb{N}$$

correspondant aux positions des objets à extraire. Pour cela, nous proposons une modélisation du mécanisme stochastique sous-jacent qui a permis de générer les données par une configuration aléatoire de points. Dans cette optique, les processus ponctuels fournissent un cadre rigoureux fondé sur la théorie de la mesure. Un processus ponctuel se définit de la façon suivante :

Définition 1 Soit (χ, d) un espace métrique complet et séparable (en général, \mathbb{R}^d muni de la distance euclidienne). Un **processus ponctuel** χ est une application X d'un

espace probabilisé $(\Omega, \mathcal{A}, \mathcal{P})$ vers la famille des configurations de points de χ localement finies¹, telle que pour tout borélien borné $A \subset \chi$, le nombre $N(A) = N_X(A)$ de points dans A soit une variable aléatoire (presque sûrement finie).

En pratique, les données sont observées dans une région bornée F de \mathbb{R}^d . Le nombre de points observés est donc fini. Un moyen simple de construire un **processus ponctuel fini**² consiste à se donner :

- une distribution de probabilité discrète $(p_n)_{n \in \mathbb{N}}$ pour le nombre de points ;
- une famille de densités de probabilité symétriques $j_n(x_1, \dots, x_n)$ sur F^n pour les positions des points.

On se donne une mesure borélienne ν sur F (en général, la mesure de Lebesgue) afin que les densités j_n puissent être définies par la mesure produit ν^n . La construction du processus X revient à définir $N(F)$ comme une variable aléatoire de loi $(p_n)_n$ puis à conditionner par rapport aux événements $\{N(F) = n\}$. Sachant $\{N(F) = n\}$, le vecteur aléatoire (X_1, \dots, X_N) correspondant à la position des objets peut alors être distribué selon j_n indépendamment de $N(A)$. Remarquons que la symétrie des densités j_n est nécessaire puisqu'un processus ponctuel est indifférent à l'ordre dans lequel les points sont placés.

Le plus connu et le plus élémentaire des processus ponctuels est le **processus uniforme de Poisson**, aussi appelé **processus homogène de Poisson**. C'est un processus spatial qui vérifie une indépendance stochastique complète. Ce processus "complètement aléatoire" se définit comme suit pour $\chi = \mathbb{R}^d$:

Définition 2 *Un processus ponctuel X sur \mathbb{R}^d est un processus uniforme de Poisson d'intensité $\lambda > 0$ si :*

- (P1) $N(A)$ suit une loi de Poisson d'espérance $\lambda|A|$ pour tout borélien borné $A \subseteq \mathbb{R}^d$, où $|A|$ est la mesure de Lebesgue de A (i.e. l'aire de A dans \mathbb{R}^2 ou son volume dans \mathbb{R}^3);
- (P2) Si A_1, \dots, A_k sont des boréliens de \mathbb{R}^d disjoints, alors les variables aléatoires $N(A_1), \dots, N(A_k)$ sont indépendantes.

De ces deux propriétés, il découle la propriété conditionnelle suivante :

Propriété 1 *Conditionnellement à $N(A) = n$, les n points présents dans A sont indépendants et uniformément distribués dans A .*

¹Une configuration $\mathbf{x} \subseteq \chi$ est dite **localement finie** si elle place un nombre fini de points dans tout borélien borné $A \subseteq \chi$.

²Un processus ponctuel sur χ est dit **fini** si $N_X(\chi)$ est fini presque sûrement.

La loi d'un processus de Poisson d'intensité λ sur une fenêtre d'observation $F \subset \mathbb{R}^d$ peut donc être définie par la mesure de probabilité suivante :

$$\mu(A) = \sum_{n=0}^{\infty} \underbrace{\frac{(\lambda|F|)^n e^{-\lambda|F|}}{n!}}_{p_n} \underbrace{\int_F \dots \int_F}_{n \text{ fois}} \mathbf{1}_A(\{x_1, \dots, x_n\}) \underbrace{\frac{1}{|F|^n}}_{j_n(x_1, \dots, x_n)} dx_1 \dots dx_n \quad (2.1)$$

$$\mu(A) = \sum_{n=0}^{\infty} \frac{e^{-\nu(F)}}{n!} \int_{F^n} \mathbf{1}_A(\mathbf{x}) d\nu^n(x_1, \dots, x_n) \quad (2.2)$$

où A est un ensemble de la tribu \mathcal{B} associée à l'espace d'état (\equiv ensemble des configurations de points finies sur F), $\mathbf{1}_A$ est la fonction indicatrice d'appartenance à A , ν correspond à la mesure de Lebesgue multipliée par λ , et \mathbf{x} la configuration de points (non ordonnée) correspondant au vecteur (x_1, \dots, x_n) .

Plus généralement, un **processus de Poisson non homogène** peut être défini par une **mesure d'intensité** ν sur $F \subseteq \chi$. La propriété **P1** devient alors :

(**P1'**) *Pour tout borélien borné $A \subseteq F$, $N(A)$ suit une loi de Poisson d'espérance :*

$$\nu(A) = \int_A \lambda(x) dx < \infty \quad (2.3)$$

où λ est une fonction d'intensité positive et mesurable sur F .

La **propriété 1** devient alors : *Conditionnellement à $N(A) = n$, les n points présents dans A sont indépendants et distribués dans A suivant une distribution de probabilité proportionnelle à ν : $P(X_i \in B) = \nu(B \cap A) / \nu(A)$.*

La loi d'un processus de Poisson non homogène sur un ensemble borné $F \subset \mathbb{R}^d$ est donnée par l'équation (2.2) où ν est la mesure d'intensité du processus définie par l'équation (2.3).

2.2.2 Processus ponctuels marqués

Travailler avec des objets plutôt qu'avec des points revient à travailler avec des points marqués où les points correspondent à la position d'un objet et les marques aux attributs permettant de décrire l'objet (taille, orientation, couleur, etc.). Définir un **processus objet** revient donc à définir un **processus ponctuel marqué**.

Définition 3 *Un processus ponctuel marqué sur \mathbb{R}^d dont les marques sont dans M est un processus de point sur $\mathbb{R}^d \times M$ tel que $N(A \times M)$ soit p. s. fini pour tout borélien borné $A \subset \mathbb{R}^d$.*

Un processus ponctuel marqué de **Poisson** est un processus ponctuel marqué où les points sont distribués selon la loi du processus de Poisson de mesure Λ sur \mathbb{R}^d et les marques associées à ces points sont indépendamment et identiquement distribuées

selon P_M dans M . Ceci est équivalent à définir un processus de Poisson sur $\mathbb{R}^d \times M$ de mesure d'intensité $\nu = \Lambda \otimes P_M$, *i.e.*

$$\nu(A \times B) = \Lambda(A)P_M(B) = \int_A \lambda(x) dx \int_B dP_M(m) \quad (2.4)$$

pour A borélien de \mathbb{R}^d et B borélien de M , où Λ et P sont des mesures d'intensité sur \mathbb{R}^d et M respectivement avec $P_M(M) = 1$.

La loi d'un processus de Poisson de mesure d'intensité $\nu = \Lambda \otimes P_M$ sur une fenêtre d'observation $F \subset \mathbb{R}^d$ peut donc être définie par la mesure de probabilité suivante :

$$\mu(A) = \sum_{n=0}^{\infty} \frac{e^{-\Lambda(F)}}{n!} \int_{(F \times M)^n} \mathbf{1}_A(\{z_1, \dots, z_n\}) d\nu^n(z_1, \dots, z_n) \quad (2.5)$$

où $A \in \mathcal{B}$.

2.2.3 Processus spécifié par une densité

Les processus de Poisson ne constituent pas en eux-mêmes un modèle pertinent pour la modélisation d'objets dans une scène, les structures spatiales induisant la dépendance stochastique. Ils sont néanmoins utilisés en tant que processus de référence permettant une construction relativement simple de modèles plus complexes. Ainsi, les différentes connaissances sur les objets à détecter et leur organisation spatiale peuvent être incorporées au travers d'une **densité de probabilité** (appelée **dérivée de Radon-Nikodym**) par rapport à un processus de Poisson de référence. Cette approche est particulièrement utile pour les **processus ponctuels finis**. En effet, il est difficile de définir des densités pour des processus ponctuel non finis [Stoyan *et al.*, 1987, van Lieshout, 2000]. De plus, étant donné une collection de densités de probabilité consistante sur une séquence d'ensembles croissant sur \mathbb{R}^d , une distribution limite n'existe pas toujours et, si elle existe, elle n'est pas nécessairement unique [Kerstan *et al.*, 1978].

Dans la suite, nous considérerons les configurations de points restreintes à une région bornée $F \subset \mathbb{R}^d$ (ce qui est toujours le cas dans la pratique). L'espace d'état E est donc l'ensemble des configurations finies de points sur F :

$$E = \bigcup_{n=0}^{\infty} E_n, \quad n < \infty \quad (2.6)$$

$$\text{avec } E_n = \begin{cases} \{\{x_1, \dots, x_n\}, x_i \in F\} & , \text{ dans le cas de points simples} \\ \{\{z_1, \dots, z_n\}, z_i \in F \times M\} & , \text{ dans le cas de points marqués} \end{cases}$$

Définition 4 Une **densité de probabilité** f par rapport au processus de Poisson de mesure μ finie et non-atomique est une fonction positive et mesurable sur E telle que :

$$\int_E f(\mathbf{u}) d\mu(\mathbf{u}) = 1 \quad (2.7)$$

Cette densité définit un processus ponctuel fini et simple (puisque le processus de référence l'est), un processus ponctuel étant **simple** s'il prend ses valeurs dans l'ensemble des configurations de points localement finies et constituées de **points distincts** presque sûrement.

Pour interpréter l'équation (2.7), il faut décomposer l'intégrale en somme sur les E_n et passer des espaces non ordonnés E_n des configurations aux espaces ordonnés $\chi^n = \chi \times \dots \times \chi$ (n fois), où $\chi = F$ ou $F \times M$. La ν -mesure de E_n est égale à $\nu(\chi)^n/n!$, où $n!$ intervient puisque χ^n est ordonné tandis que E_n ne l'est pas. La masse de E est donc égale à

$$\nu(E) = \sum_{n=0}^{\infty} \frac{\nu(\chi)^n}{n!} = e^{\nu(\chi)}$$

Pour obtenir une distribution normalisée, le facteur de normalisation $e^{\nu(\chi)}$ doit donc être pris en compte. L'équation (2.7) peut donc s'écrire comme suit :

$$\sum_{n=0}^{\infty} \frac{e^{-\nu(\chi)}}{n!} \int_{\chi^n} f(\mathbf{u}) d\nu^n(u_1, \dots, u_n) = 1 \quad (2.8)$$

où \mathbf{u} est la configuration non ordonnée associée à (u_1, \dots, u_n) . Remarquons que $\nu(\chi)$ est la même si $\chi = F$ ou $\chi = F \times M$, puisque la mesure d'intensité associée aux marques P_M est une mesure de probabilité et donc :

$$\nu(F \times M) = \Lambda(F)$$

où Λ est la mesure d'intensité associée aux points.

De ces considérations, on déduit :

1. la distribution du nombre total de points dans F :

$$p_n = \frac{e^{-\Lambda(F)}}{n!} \int_{\chi^n} f(\mathbf{u}) d\nu^n(u_1, \dots, u_n) \quad (2.9)$$

où ν est la mesure d'intensité du processus de Poisson de référence ; dans le cas de points simples, $\chi = F$ et $\nu = \Lambda$; dans le cas de points marqués, $\chi = F \times M$ et $\nu = \Lambda \otimes P_M$.

2. la densité jointe des n points conditionnellement aux événements $\{N(\chi) = n\}$:

$$j_n(u_1, \dots, u_n) = \frac{f(\mathbf{u})}{\int_{\chi^n} f(\mathbf{u}) d\nu^n(u_1, \dots, u_n)} \quad (2.10)$$

2.2.4 Processus de Markov

Comme pour les champs de Markov, une notion de voisinage et la propriété de Markov peuvent être introduites pour définir un **processus ponctuel de Markov**.

Aussi connus sous le nom de **processus ponctuels de Gibbs**, ce sont des processus définis par une densité qui peut s'écrire sous une forme énergétique incorporant des potentiels d'interaction. Ce type de processus permet ainsi de prendre en compte de façon efficace et explicite les interactions entre les points (ou objets) de la configuration, et est donc largement utilisé en physique statistique et, plus récemment, en traitement d'image. Un processus ponctuel de Markov se définit de la façon suivante [Ripley, 1977] :

Définition 5 Soit X un processus ponctuel sur un espace borné $F \subset \mathbb{R}^d$ spécifié par une densité f par rapport à μ la mesure du processus de Poisson de référence. X est un **processus de Markov** par rapport à la relation symétrique et réflexive \sim sur F si, pour tout $\mathbf{x} \in E$ tel que $f(\mathbf{x}) > 0$,

- (a) $f(\mathbf{y}) > 0$ pour tout $\mathbf{y} \subseteq \mathbf{x}$ (hérédité) ;
- (b) pour tout $u \in F$, $f(\mathbf{x} \cup \{u\})/f(\mathbf{x})$ (intensité conditionnelle de Papangelou) ne dépend que de u et de son voisinage $\partial(\{u\}) \cap \mathbf{x} = \{x \in \mathbf{x} : u \sim x\}$.

Remarque 1 Cette définition peut être généralisée à un processus ponctuel de référence défini sur un espace métrique complet et séparable (χ, d) de mesure d'intensité finie et non-atomique (i.e. processus de Poisson fini et simple). Elle est notamment valide pour les processus ponctuels marqués. Le processus de référence est alors un processus ponctuel marqué de Poisson sur $F \times M$ et la mesure \sim est une mesure symétrique et réflexive \sim sur $F \times M$.

L'équivalence entre processus de Markov et processus de Gibbs est donnée par le théorème dit de "**Hammersley-Clifford**" par similarité avec les champs de Markov :

Théorème 1 Une densité de processus ponctuel f est markovienne par rapport à la relation de voisinage \sim si et seulement si il existe une fonction mesurable $\phi : E \rightarrow [0, \infty[$ telle que :

$$f(\mathbf{x}) = \prod_{\text{cliques } \mathbf{y} \subseteq \mathbf{x}} \phi(\mathbf{y}) \quad \forall \mathbf{x} \in E \quad (2.11)$$

où \mathbf{y} est une clique si tous les éléments de \mathbf{y} sont voisins les uns des autres par rapport à \sim .

L'équation (2.11) peut également s'écrire sous forme énergétique :

$$f(\mathbf{x}) \propto \exp \left(- \sum_{\text{cliques } \mathbf{y} \subseteq \mathbf{x}} V(\mathbf{y}) \right) \quad \forall \mathbf{x} \in E \quad (2.12)$$

où \propto signifie "proportionnel à" et $V : E \rightarrow \mathbb{R}$ est une fonction de potentiel : $V(\mathbf{y})$ est appelé le **potentiel** de la clique \mathbf{y} . La somme des potentiels sur toutes les cliques de \mathbf{x} est appelée l'**énergie** de la configuration \mathbf{x} .

Ce théorème est utile pour décomposer une distribution jointe de grande dimension en fonctions d'interaction de dimension réduite qui sont plus faciles à interpréter. Pour illustrer ces propos, nous donnons l'exemple des **modèles d'interaction par paire**, largement utilisés en physique statistique. Cette classe spécifique de processus de Markov est spécifiée par une densité de la forme suivante :

$$f(\mathbf{x}) \propto \prod_{x \in \mathbf{x}} \beta(x) \prod_{u, v \in \mathbf{x}: u \sim v} \gamma(u, v) \quad (2.13)$$

où $\beta : F \rightarrow [0, \infty[$ est la fonction d'intensité et $\gamma : F \times F \rightarrow [0, \infty[$ est la fonction d'interaction par paire. Dans le cas où β et γ sont constantes, l'expression (2.13) de la densité peut s'écrire de la façon suivante :

$$f(\mathbf{x}) \propto \beta^{n(\mathbf{x})} \gamma^{s(\mathbf{x})} \quad (2.14)$$

où $n(\mathbf{x})$ et $s(\mathbf{x})$ représentent respectivement le nombre de points et le nombre de paires de points en interaction (c'est-à-dire vérifiant la relation \sim) dans la configuration \mathbf{x} . Différents types de processus dérivent de cette équation :

- Si $\gamma = 1$, les interactions inter-points n'interviennent plus et le processus ponctuel spécifié par la densité donnée par l'équation (2.14) est un **processus de Poisson** sur F de mesure d'intensité $\beta \nu$ (où ν est la mesure d'intensité du processus de référence). Ainsi, β est donc un paramètre permettant de contrôler le nombre de points, puisque l'espérance du nombre de points est multipliée par le facteur β .
- Lorsque $\gamma = 0$, le processus, appelé **processus "hard core"**, est construit de façon à interdire toute interaction. En effet, si deux points de \mathbf{x} vérifient \sim , alors $f(\mathbf{x}) = 0$.
- Lorsque $\gamma \in]0, 1[$, le processus induit une répulsion entre les points proches au sens de la relation \sim . En effet, plus le nombre de paires de points en interaction est grand, plus la densité de probabilité est faible. Ce processus est connu sous le nom de **processus de Strauss**.
- Par similarité, une première idée pour la construction d'un processus induisant une attraction entre points pourrait être de prendre $\gamma > 1$. Cependant, dans ce cas là, le processus est **mal défini**. En effet, l'équation (2.14) ne définit plus une densité par rapport au processus de référence puisque $h(\mathbf{x}) = \beta^{n(\mathbf{x})} \gamma^{s(\mathbf{x})}$ avec $\gamma > 1$ n'est pas intégrable par rapport à la mesure du processus de référence μ :

$$\int h(\mathbf{x}) d\mu(\mathbf{x}) = \infty$$

Il est néanmoins possible d'utiliser une telle expression pour la densité en la multipliant par $\mathbf{1}\{n(\mathbf{x}) \leq n_0\}$ où $\mathbf{1}$ est la fonction indicatrice et n_0 est le nombre maximal de points autorisé.

2.2.5 Conditions de stabilité

Comme nous venons de le voir dans l'exemple des modèles d'interactions par paire, la construction d'un processus de Markov ne peut se faire via la définition d'une densité non normalisée h quelconque : h doit être normalisable par rapport au processus de référence. Un processus spécifié par h par rapport à un processus de Poisson de référence sera bien défini si h vérifie la condition de Ruelle, initialement donnée par [Ruelle, 1970] pour vérifier le comportement thermodynamique d'un processus ponctuel.

Condition 1 *Un processus ponctuel spécifié par une densité non normalisée h par rapport à la mesure μ du processus de Poisson de référence, est **stable au sens de Ruelle** s'il existe $M \leq 1$ tel que :*

$$h(\mathbf{x}) \leq M^{n(\mathbf{x})} \quad \forall \mathbf{x} \in E \quad (2.15)$$

En effet, cette condition est clairement suffisante pour que h soit normalisable par rapport à μ :

$$\int h(\mathbf{x}) d\mu(\mathbf{x}) \leq \sum_{n=0}^{\infty} \frac{M^n \nu(\chi)^n}{n!} = e^{M\nu(\chi)}$$

Une deuxième condition de stabilité, impliquant la première (2.15), est très utile dès que l'on cherche à échantillonner un processus spatial. En effet, les preuves de convergence ergodique des méthodes de Monte Carlo par chaînes de Markov pour l'échantillonnage de processus spatiaux reposent sur cette condition. Dans la suite, nous supposons toujours cette condition satisfaite et nous la vérifierons lors de la proposition de nouveaux modèles.

Condition 2 *Un processus ponctuel spécifié par une densité non normalisée h par rapport à la mesure μ du processus de Poisson de référence, est dit **localement stable** s'il existe $M \in \mathbb{R}$ tel que :*

$$h(\mathbf{x} \cup u) \leq Mh(\mathbf{x}) \quad \forall \mathbf{x} \in E, \forall u \in \chi \quad (2.16)$$

2.3 Construction du processus pour l'extraction d'objet

Revenons au problème considéré qui est de modéliser la scène par une collection d'objets afin d'en réaliser l'extraction. Comment définir le processus pour exploiter à la fois les propriétés topologiques et géométriques des objets d'intérêt et les propriétés radiométriques et de texture des données ?

Tout d'abord, on se donne un processus de Poisson de référence adapté aux objets que l'on cherche à extraire. Afin de manipuler des objets simples mais génériques, on définit un processus de points marqués où les points correspondent aux positions des objets dans la fenêtre d'observation et les marques aux divers attributs des objets (longueur, largeur, orientation, etc.). Ce processus de référence n'est pas en lui-même un a

priori pertinent puisqu'il traduit la notion d'indépendance : les points n'interagissent pas entre-eux. Bien que l'on puisse y incorporer une information de localisation (processus non homogène) ou même un *a priori* sur les marques, on le prend en général uniforme.

Pour intégrer un *a priori* sur les interactions entre les objets (et sur la forme des objets si le processus de référence est homogène), il suffit de spécifier le processus par une densité non normalisée h_p par rapport au processus de Poisson de référence. Celle-ci doit au moins vérifier la condition de stabilité de Ruelle pour que le processus soit bien défini.

Ensuite, pour la bonne localisation des points, un terme d'attache aux données h_d doit être défini. Considérons le cas où les données dont nous disposons correspondent à une image Y . La configuration d'objet \mathbf{x} étant définie dans une région $F \subset \mathbb{R}^2$, une projection de \mathbf{x} dans la grille pixélique devra donc être effectuée. Soit $S(\mathbf{x})$ la silhouette de \mathbf{x} sur la grille, où $s_p(\mathbf{x})$, la valeur de $S(\mathbf{x})$ au pixel p , est strictement positive si la projection de \mathbf{x} sur la grille atteint p et est nulle sinon. Une première approche consiste à définir le terme d'attache aux données comme la **vraisemblance des observations** par rapport à cette silhouette [Baddeley et van Lieshout, 1993, Rue et Hurn, 1999, Perrin *et al.*, 2004]. Si l'on suppose que les valeurs y_p de l'image Y sont conditionnellement indépendantes sachant \mathbf{x} et si le niveau de gris y_p ne dépend que de la valeur $s_p(\mathbf{x})$, alors on peut écrire la vraisemblance des observations de la façon suivante :

$$h_d(\mathbf{x}) = h_d(Y|\mathbf{x}) \propto \prod_{p \in G} g(y_p | s_p(\mathbf{x})) \quad (2.17)$$

où G désigne la grille pixélique et $g(y_p | s_p(\mathbf{x}))$ est la vraisemblance de y_p sachant $s_p(\mathbf{x})$, *i.e.* sachant les paramètres d'un modèle de fond si $s_t(\mathbf{x}) = 0$ ou les paramètres d'un modèle associé aux objets sinon. Par exemple, dans le cas le plus simple où le fond et les objets sont décrits par deux niveaux de gris distincts m_F et m_O , seulement dégradés par un bruit gaussien additif, le modèle de fond sera une vraisemblance gaussienne de moyenne m_F et variance σ_F^2 et celui des objets sera une vraisemblance gaussienne de moyenne m_O et variance σ_O^2 . Cette approche permet de travailler dans un cadre bayésien ce qui permet, entre autre, de réaliser une estimation des paramètres des modèles. Il faut néanmoins pouvoir définir un modèle de fond, ce qui devient une tâche ardue dès que le fond est constitué de zones hétérogènes (ex : différentes cultures dans une zone agricole) ou que des objets non pertinents pour l'application considérée sont présents dans l'image et interfèrent avec la bonne détection des objets d'intérêt. C'est pourquoi ce terme d'attache aux données n'est utilisé que dans les cas simples où l'on peut définir aisément les deux modèles, comme, par exemple, pour les cellules dans des images de microscopiques confocales [Rue et Hurn, 1999]. Dans les cas plus complexes, on préférera utiliser les données en tant que **champ externe** permettant d'évaluer la qualité de chaque objet compte tenu des hypothèses d'homogénéité et de contraste avec l'environnement que l'on a pu émettre [Garcin *et al.*, 2001, Ortner, 2001, Stoica *et*

al., 2004]. Cette deuxième approche consiste à définir l'énergie d'attache aux données pour une configuration \mathbf{x} comme la somme de potentiels associés à chaque objet de la configuration :

$$h_d(\mathbf{x}) = \exp \left(- \sum_{x \in \mathbf{x}} v_d(x) \right) \quad (2.18)$$

Chaque potentiel $v_d(x)$ est évalué sur un masque pixélique associé à l'objet x . Le potentiel sera d'autant plus faible que les hypothèses radiométriques seront plausibles. L'inconvénient d'une telle approche est que l'on sort du cadre bayésien et que deux objets se superposant comptent double. Il faudra donc fortement pénaliser les superpositions d'objets.

Finalement, la densité complète du processus se définit comme le produit de la densité *a priori* et du terme d'attache aux données :

$$f(\mathbf{x}) \propto h_p(\mathbf{x}) h_d(\mathbf{x}) \quad (2.19)$$

Dans le cas où h_d correspond à la vraisemblance des observations (équation 2.17), la densité complète correspond à la densité *a posteriori* : $f(\mathbf{x}) = f(\mathbf{x}|Y)$.

2.4 Chaînes de Markov et convergence

La construction des algorithmes d'échantillonnage présentés dans le paragraphe 2.5 reposent sur la construction d'une chaîne de Markov convergeant vers la mesure à échantillonner. Pour comprendre le principe de ces algorithmes, nous donnons dans ce paragraphe les propriétés nécessaires à la convergence de la chaîne.

2.4.1 Définitions et notations

Définition 6 Une suite de variables aléatoires $\{X_n\}$ à valeurs dans E muni de sa tribu \mathcal{B} est une **chaîne de Markov** si :

$$p(X_{t+1} \in A | X_0 = x_0, \dots, X_t = x_t) = p(X_{t+1} \in A | X_t = x_t) \quad \forall A \in \mathcal{B} \quad (2.20)$$

En d'autres termes, une chaîne de Markov a la propriété que son évolution (passage de X_t à X_{t+1}) ne dépend que de l'état courant ($X_t = x_t$) et pas de son passé. Cette propriété facilite donc la mise en oeuvre informatique d'un tel processus car il n'est pas nécessaire de garder en mémoire tout le passé du processus pour effectuer des calculs. Cette chaîne est **homogène** si cette évolution ne dépend pas de la date t , mais seulement des états concernés. Nous ne considérerons dans cette partie que ce type de chaîne de Markov.

Définition 7 On appelle **noyau de transition** une fonction P définie sur $E \times \mathcal{B}$ telle que :

1. $\forall x \in E, P(x, \cdot)$ est une mesure de probabilité;

2. $\forall A \in \mathcal{B}$, $P(\cdot, A)$ est mesurable.

Le noyau de transition P associé à une chaîne de Markov homogène est donné par :

$$P(x, A) = p(X_{t+1} \in A | X_t = x) \quad (2.21)$$

2.4.2 Stationnarité

La propriété essentielle requise par tout échantillonneur de type Monte Carlo par chaîne de Markov (cf. paragraphe 2.5.1) est que la chaîne $\{X_t\}$ ait une mesure stationnaire π . Cette propriété est équivalente à l'invariance de π par rapport à $\{X_t\}$:

Propriété 2 Une mesure π est **stationnaire** ou **invariante** pour la chaîne de Markov de noyau de transition P si :

$$\pi(A) = \int P(x, A) d\pi(x) \quad \forall A \in \mathcal{B} \quad (2.22)$$

La vérification de la réversibilité d'un noyau stochastique P par rapport à π est un outil pour établir la stationnarité de π pour P .

Propriété 3 Une chaîne de Markov sur l'espace mesurable est **réversible** si son noyau de transition P vérifie l'équation :

$$\int_A P(x, B) d\pi(x) = \int_B P(y, A) d\pi(y) \quad \forall A, B \in \mathcal{B} \quad (2.23)$$

Cette propriété signifie que sous π la probabilité de passer de A à B est la même que de passer de B à A . Certains auteurs parlent aussi de la condition d'équilibre minutieux ($\pi(dx)P(x, dy) = \pi(dy)P(y, dx)$) qui est souvent équivalente à la réversibilité de la chaîne, comme, par exemple, dans le cas discret. La réversibilité étant plus facile à vérifier que la stationnarité, la plupart des algorithmes d'échantillonnage vérifient cette propriété, cette dernière impliquant l'invariance pour π .

Théorème 2

$$REVERSIBILITE \implies STATIONNARITE$$

2.4.3 Convergence

La vérification de la stationnarité ne suffit pas à garantir la convergence de la chaîne :

$$X_n \xrightarrow[n \rightarrow \infty]{D} \pi$$

c'est-à-dire que :

$$P^t(x, A) \xrightarrow[t \rightarrow \infty]{} \pi(A) \quad \forall x \in E, A \in \mathcal{B} \quad (2.24)$$

où $P^t(x, A) = p(X_t \in A | X_0 = x)$.

A. Irréductibilité

Pour vérifier l'équation (2.24), il faut au moins vérifier la propriété suivante :

Propriété 4 *La chaîne Markov $\{X_n\}$ sur un espace mesurable (E, \mathcal{B}) est ϕ -irréductible s'il existe une mesure non nulle ϕ sur \mathcal{B} telle que pour tout $x \in E$ et $\forall A \in \mathcal{B}$:*

$$\phi(A) > 0 \Rightarrow \exists t \in \mathbf{N} : P^t(x, A) > 0 \quad (2.25)$$

L'intérêt de cette propriété est qu'elle implique les propriétés suivantes :

Proposition 1 *Si $\{X_n\}$ est ϕ -irréductible et si elle possède une mesure invariante π , alors :*

1. π est l'unique mesure invariante pour le noyau P ;
2. $\{X_n\}$ est π -irréductible ;
3. $\pi(A) = 0 \Rightarrow \phi(A) = 0$

On peut donc montrer la π -irréductibilité (ou irréductibilité) de la chaîne via une étude sur une autre mesure. L'irréductibilité signifie que la chaîne a une probabilité non nulle d'atteindre en temps fini tout ensemble π -probable quelque soit la condition initiale.

B. Apériodicité

En plus de l'irréductibilité, l'équation (2.24) implique l'apériodicité définie comme suit :

Propriété 5 *Soit $\{X_t\}$ une chaîne π -irréductible. Les ensembles A_1, \dots, A_m de \mathcal{B} forment un m -cycle si :*

$$\begin{aligned} x \in A_1 &\Rightarrow P(x, A_2) = 1 \\ &\vdots \\ x \in A_{m-1} &\Rightarrow P(x, A_m) = 1 \\ x \in A_m &\Rightarrow P(x, A_1) = 1 \end{aligned}$$

et $\pi(\cup_{i=1}^m A_i) = 1$. Le plus grand $d \in \mathbf{N}$ pour lequel un d -cycle est formé est appelé la période de la chaîne. Quand $d = 1$, la chaîne est **apériodique**.

Une condition suffisante pour qu'il y ait apériodicité est que

$$\exists x \in E : P(x, \{x\}) > 0 \quad (2.26)$$

Si la condition (2.26) est vérifiée, on dit que la chaîne est **fortement apériodique**. Cette propriété est donc toujours vérifiée dès que la probabilité de rester dans l'état courant est positive (dans les algorithmes de type Metropolis-Hastings, probabilité de

rejet $\delta > 0$).

Lorsque la chaîne est π -irréductible et apériodique on obtient la convergence pour presque tout $x \in E$:

$$\|P^t(x, \cdot) - \pi\| \xrightarrow[t \rightarrow \infty]{} 0 \quad \forall x \in E \setminus N \quad (2.27)$$

où $N \in \mathcal{B}/\pi(N) = 0$ et $\|\cdot\|$ désigne la norme en variation totale : $\|\mu_1 - \mu_2\| = \sup_A |\mu_1(A) - \mu_2(A)|$.

C. Convergence ergodique

Pour obtenir la convergence indépendamment de toute condition initiale, on introduit la notion de récurrence au sens de Harris :

Propriété 6 Une chaîne de Markov est récurrente au sens de **Harris** si :

$$\forall x \in E, \forall A \in \mathcal{B} / \pi(A) > 0 : p(\{\exists t / X_t \in A\} | X_0 = x) = 1 \quad (2.28)$$

Cette propriété implique l'irréductibilité, *i.e.* la π -irréductibilité avec π mesure stationnaire de la chaîne. Dans le cas d'un espace d'état fini, l'irréductibilité et la récurrence au sens de Harris sont équivalentes.

Finalement, une chaîne de Markov (de mesure stationnaire π) apériodique et récurrente au sens de **Harris** converge ergodiquement vers π :

Théorème 3

$$HARRIS + APERIODICITE \implies ERGODICITE$$

La notion d'ergodicité concerne la convergence d'une mesure vers une autre indépendamment de la condition initiale.

Propriété 7 L'ergodicité est équivalente à :

$$\|P^t(x, \cdot) - \pi\| \xrightarrow[t \rightarrow \infty]{} 0 \quad \forall x \in E \quad (2.29)$$

Une fois la stationnarité et l'apériodicité vérifiées, il reste donc à vérifier la récurrence au sens de Harris pour une bonne définition de l'algorithme d'échantillonnage. Pour cela, le concept des ensembles petits s'avère utile.

Définition 8 Un ensemble $C \in \mathcal{B}$ est dit **petit** s'il existe un entier m , un réel $\epsilon > 0$ et une mesure de probabilité κ sur \mathcal{B} tels que :

$$P^m(x, A) \geq \epsilon \kappa(A) \quad \forall x \in C, \forall A \in \mathcal{B} \quad (2.30)$$

Proposition 2 *Supposons $\{X_t\}$ irréductible et apériodique et qu'il existe un ensemble petit $C \in \mathcal{B}$ et une fonction $V : E \rightarrow \mathbb{R}$ telle que $\{x : V(x) \leq n\}$ est petit pour n'importe quel $n \in \mathbb{N}$ et que :*

$$\forall x \in E \setminus C : \mathbb{E}[V(X_1)|X_0 = x] \leq V(x) \quad (2.31)$$

où $\mathbb{E}[V(X_1)|X_0 = x]$ désigne l'espérance de $V(X_1)$ sous la mesure de probabilité $P(x, \cdot)$. Alors la chaîne est récurrente au sens de **Harris**.

L'équation (2.31) est appelée **condition de drift** pour la récurrence.

Pour montrer la convergence de la chaîne pour toute condition initiale (après avoir vérifié la stationnarité), il faut donc montrer qu'elle est apériodique et qu'elle est ϕ -irréductible. Dans le cas où l'irréductibilité et la récurrence au sens de Harris ne sont pas équivalentes, une condition de drift pourra être établie pour montrer la récurrence au sens de Harris. Par le théorème 3, l'ergodicité sera finalement prouvée.

D. Ergodicité géométrique

L'ergodicité géométrique, propriété plus forte que l'ergodicité (propriété 7), se définit comme suit :

Propriété 8 *Une chaîne de Markov est géométriquement ergodique s'il existe une constante $r > 1$ telle que*

$$\sum_{t=1}^{\infty} r^t \|P^t(x, \cdot) - \pi(\cdot)\| < \infty \quad \forall x \in E \quad (2.32)$$

L'ergodicité géométrique impose donc que la convergence de P^n vers π se fasse avec une vitesse géométrique puisque l'équation (2.32) implique :

$$\|P^t(x, \cdot) - \pi\| \leq M r^{-t} \quad (2.33)$$

où $M = \sum_{t=1}^{\infty} r^t \|P^t(x, \cdot) - \pi\|$.

L'équation (2.32) peut être démontrée par une **condition de drift géométrique** qui se définit comme suit :

Condition 3 *Il existe une fonction $V : E \rightarrow [1, \infty[$, des constantes $b < \infty$ et $\lambda < 1$, et un ensemble petit $C \in \mathcal{B}$ tel que :*

$$\mathbb{E}[V(X_1)|X_0 = x] \leq \lambda V(x) + b \mathbf{1}_C(x) \quad \forall x \in E \quad (2.34)$$

Remarque 2 *Cette condition implique la condition de drift pour la récurrence.*

L'ergodicité géométrique est une propriété "qualitative" et son importance réside principalement dans l'établissement du théorème de la limite centrale (théorème 4). Elle permet également d'établir des bornes qualitatives pour les taux de convergence.

E. Théorème de la limite centrale

Soit $g(X)$ une statistique d'intérêt telle que g soit π intégrable :

$$\mu_g = E_\pi[g(X)] = \int g(x) d\pi(x) \quad (2.35)$$

où π est la distribution stationnaire d'une chaîne de Markov récurrente au sens de Harris. Considérons également une estimée \bar{g}_t obtenue par l'échantillonneur :

$$\bar{g}_t = \frac{1}{t} \sum_{i=1}^t g(x_i) \quad (2.36)$$

La loi des grands nombres (assurée par la récurrence au sens de Harris) impose que $\bar{g}_t \rightarrow \mu_g$ presque sûrement. Supposons que la variance asymptotique

$$\sigma_g^2 = \lim_{t \rightarrow \infty} t \operatorname{Var}_\pi(\bar{g}_t) \quad (2.37)$$

existe et est donnée par :

$$\sigma_g^2 = \operatorname{Var}_\pi(g(X)) + 2 \sum_{k=1}^{\infty} \operatorname{Cov}_\pi(g(X_t)g(X_{t+k})) \quad (2.38)$$

Si la chaîne est géométriquement ergodique et que la fonction g satisfait une condition de Lyapunov donnée par :

$$\int |g(x)|^{2+\epsilon} d\pi(x) < \infty \quad (2.39)$$

avec $\epsilon > 0$, alors le Théorème de la Limite Centrale (TLC) s'applique :

Théorème 4 *Si la chaîne est géométriquement ergodique et g vérifie la condition de Lyapunov (2.39) alors :*

$$\sqrt{t} \frac{\bar{g}_t - \mu_g}{\sqrt{\sigma_g^2}} \xrightarrow[t \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, 1) \quad (2.40)$$

Remarque 3 *L'équation (2.40) peut également être montrée par la vérification de la condition de drift géométrique pour une fonction finie V , avec la condition $g^2 \leq V$, remplaçant la condition de Lyapunov.*

2.5 Échantillonnage

Une façon simple de modéliser la scène observée par une collection d'objets est de construire un processus ponctuel marqué spécifié par une densité de probabilité f . Cependant, cette densité n'est généralement connue qu'à un facteur près. Que ce soit en inférence statistique ou en extraction d'objets dans une scène, des techniques de type Monte Carlo sont utilisées du fait du calcul impossible de la constante de normalisation. Ce sont des techniques d'échantillonnage qui utilisent les échantillons aléatoires en tant qu'outil de calcul et permettent ainsi l'évaluation d'intégrales non calculables analytiquement.

2.5.1 Méthodes MCMC

Les méthodes de type Monte Carlo par Chaîne de Markov (MCMC) répondent au problème d'échantillonnage de la façon suivante : on construit une chaîne de Markov discrète $\{X_0, X_1, X_2, \dots\}$, ayant un espace d'état E , qui converge vers la distribution recherchée π . La chaîne est conçue pour être ergodique, c'est-à-dire que la distribution de probabilité sur E converge asymptotiquement vers π indépendamment de l'initialisation. De plus, ses transitions correspondent à des perturbations simples de structures sur E et sont donc simples à simuler. Nous pouvons donc échantillonner selon π de la façon suivante : partir d'un état arbitraire dans E , simuler la chaîne pendant un nombre d'itérations N suffisamment grand. A partir de N , les échantillons successifs sont distribués suivant une loi proche de π .

Les algorithmes de type MCMC les plus connus sont les algorithmes de type Metropolis-Hastings et l'échantillonneur de Gibbs, largement utilisés en analyse d'images (notamment pour l'échantillonnage de champs de Markov). Les algorithmes de type Metropolis-Hastings sont fondés sur la construction d'une chaîne de Markov à temps discret dont les transitions sont définies en deux étapes : une phase de proposition de perturbation suivie de l'acceptation ou du rejet de la perturbation selon le rapport de vraisemblance entre le nouvel état et l'ancien état et la façon dont les propositions sont faites. L'algorithme MCMC le plus simple est l'algorithme de Metropolis proposé par [Metropolis *et al.*, 1953] dans lequel seul le rapport de vraisemblance intervient dans la phase d'acceptation. Une formulation plus générale a ensuite été donnée par [Hastings, 1970]. La méthodologie pour simuler une distribution spécifiée par une densité non normalisée h se décrit comme suit (avec les notations utilisées pour un espace discret, par soucis de simplification d'écriture) :

- Se donner une **matrice de probabilité de proposition de perturbation** $Q = (q_{ij})_{i,j \in E}$, où $q_i = q(i \rightarrow \cdot)$ est la densité de probabilité de proposition de perturbation à partir de i .
- Proposer un nouvel état j selon $q(i \rightarrow \cdot)$.
- Accepter la perturbation de i à j avec une probabilité α_{ij} , définie de façon à vérifier la **condition de micro-réversibilité** ou **d'équilibre minutieux** ("detailed balance", en anglais) :

$$P(k, l) h(l) = P(l, k) h(k) \quad \forall k, l \in E \quad (2.41)$$

où P est le noyau de transition de la chaîne ; ce qui revient à vérifier que :

$$q_{kl} \alpha_{kl} h(l) = q_{lk} \alpha_{lk} h(k) \quad \forall k, l \in E$$

La condition d'équilibre minutieux exprime le fait que, dans l'état stationnaire, la probabilité d'aller d'un état d'équilibre i vers un état j est la même que celle d'aller d'un

état d'équilibre j vers un état i . C'est une condition suffisante pour vérifier que h est la mesure stationnaire de la chaîne $\{X_t\}$. Pour vérifier cette condition, le choix usuel est le suivant : prendre $\alpha_{ij} = \min\{R_{ij}, 1\}$ avec un **taux d'Hastings** défini par :

$$R_{ij} = \frac{h(j) q_{ji}}{h(i) q_{ij}} \quad (2.42)$$

On retrouve l'algorithme proposé par Metropolis en prenant un noyau de perturbation symétrique $q_{ji} = q_{ij}$.

L'échantillonneur de Gibbs est un cas particulier de l'algorithme d'Hastings, applicable dans le cas les états sont partitionnés en coordonnées $i = (i_1, \dots, i_n)$. A chaque étape, une seule des coordonnées i_k est mise à jour conditionnellement aux valeurs des autres coordonnées. Les probabilités de propositions sont donc définies comme suit :

$$q_{i_k j_k} = P(i_k \rightarrow j_k \mid i_1, \dots, i_{k-1}, i_{k+1}, \dots, i_n) = \frac{h(i_1, \dots, i_{k-1}, j_k, i_{k+1}, \dots, i_n)}{\sum_l h(i_1, \dots, i_{k-1}, l, i_{k+1}, \dots, i_n)} \quad (2.43)$$

L'intérêt d'un tel choix est que la probabilité d'acceptation est toujours égale à 1. L'étape d'acceptation/rejet est donc retirée de l'algorithme. Il faut néanmoins pouvoir calculer la distribution de proposition donnée par l'équation (2.43). Cet algorithme est particulièrement adapté à l'échantillonnage des champs de Markov [Geman et Geman, 1984], pour lesquels la probabilité de proposition d'une nouvelle valeur j_k en k ne dépend que de j_k et des valeurs des coordonnées voisines de la coordonnée k (*i.e.* des pixels voisins du pixel k , en image).

2.5.2 Méthodes MCMC pour les processus ponctuels

Rappelons que la distribution π d'un processus Markov objet est définie sur un espace d'état E de dimension variable, le nombre d'objets étant une variable aléatoire. La chaîne de Markov $\{X_t\}_{t=0,1,\dots}$ permettant la simulation d'un processus objet devra donc être construite en conséquence. Ainsi, des perturbations de type naissance et mort d'un objet seront définies, permettant à la chaîne de réaliser de petits sauts entre les espaces E_n (espaces définis dans l'équation 2.6). Deux types algorithmes sont proposés dans la littérature pour l'échantillonnage de processus spatiaux : les algorithmes de type Naissance-Mort (NM) et les algorithmes de type Metropolis-Hastings (MH).

L'échantillonneur de type naissance-mort

L'échantillonneur de type naissance-mort est fondé sur les processus naissance et mort [Preston, 1976, Ripley, 1977]. Initialement définie en temps continu, la simulation d'un processus naissance et mort est la première technique utilisée en statistique pour simuler les processus ponctuels de Markov. Les transitions à partir d'un état \mathbf{x} sont des naissances ou des morts, *i.e.* des ajouts ou des suppressions d'un point (ou d'un

Initialisation par $X_0 = \mathbf{x}_0$ avec $\mathbf{x}_0 \in E / f(x_0) > 0$. Au temps t , $X_t = \mathbf{x}$:	
1.	T_t est distribué exponentiellement de moyenne $1/(B(\mathbf{x}) + D(\mathbf{x}))$ où $B(\mathbf{x}) = \int_{\chi} b(\mathbf{x}, u) d\nu(u)$ est le taux de naissance global et $D(\mathbf{x}) = \sum_{u \in \mathbf{x}} d(\mathbf{x}, u)$ est le taux de mort global
2.	La transition suivante est une :
• naissance	avec la probabilité $B(\mathbf{x})/(B(\mathbf{x}) + D(\mathbf{x}))$ → ajout d'un nouveau point (marqué) u , généré par rapport à la densité de probabilité $b(\mathbf{x}, u)/B(\mathbf{x})$
• mort	avec la probabilité $D(\mathbf{x})/(B(\mathbf{x}) + D(\mathbf{x}))$ → retrait d'un point (marqué) u , choisi avec la probabilité $d(\mathbf{x}, u)/D(\mathbf{x})$

TAB. 2.1 – Algorithme de type naissance-mort.

objet) u , avec des taux respectifs $b(\mathbf{x}, u)$ et $d(\mathbf{x}, u)$. Le plus souvent, les taux sont définis comme suit :

$$\begin{aligned} b(\mathbf{x}, u) &= \frac{f(\mathbf{x} \cup u)}{f(\mathbf{x})} , & u \in \chi \text{ avec } \chi = F \text{ ou } \chi = F \times M \\ d(\mathbf{x}, u) &= 1 , & u \in \mathbf{x} \end{aligned}$$

où f est la densité spécifiant le processus spatial à échantillonner par rapport au processus de Poisson de référence. Toutes les propositions de transitions sont acceptées avec la probabilité 1 mais le processus reste dans l'état X_t pendant T_t , un temps de séjour distribué exponentiellement. La simulation d'un processus de naissance mort consiste en la génération des état successifs et du temps de séjour comme cela est décrit dans la table 2.1. Un inconvénient majeur de cet algorithme est que le calcul du taux global de naissance est, en général, difficile à effectuer. Une solution est de remplacer $b(\mathbf{x}, u)$ par une constante β telle que le taux global de naissance soit borné par $\beta\nu(\chi)$, et d'accepter l'ajout d'un élément u avec la probabilité $\lambda(u; \mathbf{x})/\beta$, où $\lambda(u; \mathbf{x}) = f(\mathbf{x} \cup u)/f(\mathbf{x})$ (intensité conditionnelle de Papangelou). Cette stratégie marche relativement bien lorsque les interactions sont faibles. Si l'intensité conditionnelle est fortement piquée, les probabilités d'acceptation auront tendance à être trop petites et l'algorithme perdra en efficacité.

Algorithme de type Metropolis-Hastings

[Geyer et Møller, 1994, Geyer, 1999, Green, 1995] reprennent le formalisme introduit par Hastings (décrit dans le paragraphe 2.5.1) et l'adaptent au cas des pro-

cessus spatiaux en construisant une chaîne pouvant réaliser de petits sauts entre des espaces de dimensions différentes; d'où le nom de méthodes de **Monte Carlo par chaîne de Markov à sauts réversibles** (Reversible Jump Markov Chain Monte Carlo (RJMCMC), en anglais). La méthodologie reste la même que celle de l'algorithme de Metropolis-Hastings classique mais ce sont des mesures qui sont considérées plutôt que des densités.

- la densité non normalisée h est remplacée par une mesure π sur $E = \bigcup_{n=0}^{\infty} E_n$;
- la densité de probabilité $q(\mathbf{x} \rightarrow \cdot)$ est remplacée par un noyau de proposition $Q(\mathbf{x} \rightarrow \cdot)$;
- le calcul du taux d'acceptation (taux de Green) se fait via la définition d'une mesure symétrique ψ sur $E \times E$ par rapport à laquelle on peut définir la dérivée de Radon-Nikodym de $\pi Q(A, B) = \int_A Q(\mathbf{x} \rightarrow B) d\pi(\mathbf{x})$.

Initialisation par $X_0 = \mathbf{x}_0$ avec $\mathbf{x}_0 \in E / f(x_0) > 0$. Au temps t , $X_t = \mathbf{x}$:

1. Proposer $\mathbf{y} \sim Q(\mathbf{x} \rightarrow \cdot)$
2. Évaluer le taux de Green $R(\mathbf{x}, \mathbf{y})$ donné par l'équation (2.51)
3. Accepter \mathbf{y} avec la probabilité $\alpha(\mathbf{x}, \mathbf{y}) = \min(1, R(\mathbf{x}, \mathbf{y}))$

TAB. 2.2 – Algorithme de type Metropolis-Hastings pour les processus ponctuels.

Les étapes de l'algorithme MCMC à sauts réversibles suivant une dynamique de Metropolis-Hastings sont résumées dans la table 2.2. A chaque itération, une perturbation de l'état courant \mathbf{x} vers un nouvel état \mathbf{y} est proposée suivant un noyau de proposition $Q(\mathbf{x} \rightarrow \cdot)$. La perturbation est acceptée avec une probabilité $\alpha(\mathbf{x}, \mathbf{y})$. Celle-ci est calculée de façon à vérifier la condition de réversibilité de la chaîne (impliquant que π est la mesure stationnaire de la chaîne X_t). L'expression de cette condition est donnée par l'équation suivante :

$$\int_A P(\mathbf{x}, B) d\pi(\mathbf{x}) = \int_B P(\mathbf{y}, A) d\pi(\mathbf{y}) \quad \forall A, B \in \mathcal{B} \quad (2.44)$$

où \mathcal{B} désigne la tribu associée à E et P est le noyau de transition de X_t , défini comme suit :

$$\begin{aligned} P(\mathbf{x}, A) &= P(X_{t+1} \in A \mid X_t = \mathbf{x}) \\ &= \underbrace{\int_A \alpha(\mathbf{x}, \mathbf{y}) dQ^{\mathbf{x}}(\mathbf{y})}_{\text{Acceptation}} + \underbrace{\mathbf{1}(\mathbf{x} \in A) \int_E (1 - \alpha(\mathbf{x}, \mathbf{y})) dQ^{\mathbf{x}}(\mathbf{y})}_{\text{Rejet}} \end{aligned} \quad (2.45)$$

où $\mathbf{1}(\cdot)$ correspond à la fonction indicatrice et $Q^{\mathbf{x}}(\cdot) = Q(\mathbf{x} \rightarrow \cdot)$. Pour vérifier (2.44), il suffit de vérifier que :

$$\int_A \int_B \alpha(\mathbf{x}, \mathbf{y}) dQ^{\mathbf{x}}(\mathbf{y}) d\pi(\mathbf{x}) = \int_B \int_A \alpha(\mathbf{y}, \mathbf{x}) dQ^{\mathbf{y}}(\mathbf{x}) d\pi(\mathbf{y}) \quad \forall A, B \in \mathcal{B} \quad (2.46)$$

Pour définir une fonction α vérifiant cette égalité, [Green, 1995] propose de se donner une **mesure symétrique** ψ sur $E \times E$ telle que pour tout borélien A et B :

$$\psi(A, B) = 0 \implies \pi Q(A, B) = \int_A Q(\mathbf{x} \rightarrow B) d\pi(\mathbf{x}) = 0$$

i.e. telle que πQ soit **absolument continu** par rapport à ψ . On peut alors définir une fonction unique D telle que :

$$\int_A Q(\mathbf{x} \rightarrow B) \pi(d\mathbf{x}) = \int_{A \times B} D(\mathbf{x}, \mathbf{y}) d\psi(\mathbf{x}, \mathbf{y}) \quad (2.47)$$

La fonction D est appelée la **dérivée de Radon-Nikodym** (ou la densité) de πQ par rapport à ψ . Une fois ψ définie, on peut donc récrire l'équation (2.46) de la façon suivante :

$$\int_A \int_B \alpha(\mathbf{x}, \mathbf{y}) D(\mathbf{x}, \mathbf{y}) d\psi(\mathbf{x}, \mathbf{y}) = \int_B \int_A \alpha(\mathbf{y}, \mathbf{x}) D(\mathbf{y}, \mathbf{x}) d\psi(\mathbf{y}, \mathbf{x}) \quad (2.48)$$

Puisque ψ est symétrique, il suffit de vérifier :

$$\alpha(\mathbf{x}, \mathbf{y}) D(\mathbf{x}, \mathbf{y}) = \alpha(\mathbf{y}, \mathbf{x}) D(\mathbf{y}, \mathbf{x}) \quad \forall \mathbf{x}, \mathbf{y} \in E \quad (2.49)$$

Plusieurs choix sont possibles pour la probabilité d'acceptation. Dans le cas d'un espace d'état fini, [Peskun, 1973] a montré que le choix optimal pour α est de prendre cette probabilité aussi grande que possible afin de réduire l'auto-corrélation de la chaîne $\{X_t\}$. Le choix usuel pour α vérifiant l'équation (2.49) est alors :

$$\alpha(\mathbf{x}, \mathbf{y}) = \min \{1, R(\mathbf{x}, \mathbf{y})\} \quad (2.50)$$

où R est le **taux de Green** donné par :

$$R(\mathbf{x}, \mathbf{y}) = \frac{D(\mathbf{y}, \mathbf{x})}{D(\mathbf{x}, \mathbf{y})} \quad (2.51)$$

où D est la dérivée de Radon-Nikodym de πQ par rapport à la mesure symétrique ψ .

Le premier algorithme de type MH pour la simulation d'un processus de point spécifié par une densité f par rapport à la mesure μ d'un processus de Poisson uniforme est celui proposé par [Geyer et Møller, 1994]. Les perturbations proposées consistent en l'ajout d'un point uniformément dans la région observée $F \subset \mathbb{R}^2$ ou la suppression d'un point choisi uniformément dans la configuration courante. Ces perturbations correspondent au mouvement le plus simple pour passer de E_n à E_{n+1} et réciproquement : le mouvement réversible de type "naissance et mort" uniforme. Le noyau de proposition Q utilisé correspond à une "naissance et mort" uniforme où la naissance est proposée avec la probabilité p_b et la mort avec la probabilité $p_d = 1 - p_b$:

$$Q(\mathbf{x} \rightarrow A) = p_b Q_b(\mathbf{x} \rightarrow A) + p_d Q_d(\mathbf{x} \rightarrow A) \quad (2.52)$$

où Q_b et Q_d sont des sous-noyaux, correspondant respectivement à la naissance et à la mort d'un point, et définis comme suit :

$$Q_b(\mathbf{x} \rightarrow A) = \int_F \mathbf{1}_A(\mathbf{x} \cup u) \frac{1}{|F|} du \quad (2.53)$$

$$Q_d(\mathbf{x} \rightarrow A) = \sum_{u \in \mathbf{x}} \mathbf{1}_A(\mathbf{x} \setminus u) \frac{1}{n(\mathbf{x})} \quad (2.54)$$

Remarque 4 *Le choix usuel pour les probabilités de choix de mouvement est de prendre $p_b = p_d = 0.5$.*

Il reste à trouver une mesure symétrique sur $E \times E$ par rapport à laquelle dériver πQ . Remarquons que les deux parties du noyau n'incluent que des mouvements de E_n à E_{n+1} et réciproquement. Nous pouvons donc restreindre l'étude aux ensembles de la tribu associée à E tels que $A_n \subseteq E_n$ et $B_{n+1} \subseteq E_{n+1}$. L'expression des mesures $\pi Q(A_n, B_{n+1})$ et $\pi Q(B_{n+1}, A_n)$ sont les suivantes :

$$\int_{A_n} Q(\mathbf{x}, B_{n+1}) d\pi(\mathbf{x}) = p_b \int_{A_n} \left[\int_F \mathbf{1}_{B_{n+1}}(\mathbf{x} \cup u) \frac{du}{|F|} \right] f(\mathbf{x}) d\mu(\mathbf{x}) \quad (2.55)$$

$$\int_{B_{n+1}} Q(\mathbf{x}, A_n) d\pi(\mathbf{x}) = p_d \int_{B_{n+1}} \left[\sum_{u \in \mathbf{x}} \mathbf{1}_{A_n}(\mathbf{x} \setminus u) \frac{1}{n(\mathbf{x})} \right] f(\mathbf{x}) d\mu(\mathbf{x}) \quad (2.56)$$

Soit ψ_n^+ définie sur $E_n \times E_{n+1}$ et ψ_n^- sur $E_{n+1} \times E_{n+1}$ par :

$$\psi_n^+(A_n, B_{n+1}) = \int_{A_n} \int_F \mathbf{1}_{B_{n+1}}(\mathbf{x} \cup u) d\nu(u) d\mu(\mathbf{x}) \quad (2.57)$$

$$\psi_n^-(B_{n+1}, A_n) = \int_{B_{n+1}} \sum_{u \in \mathbf{x}} \mathbf{1}_{A_n}(\mathbf{x} \setminus u) d\mu(\mathbf{x}) \quad (2.58)$$

où ν est la mesure d'intensité du processus ponctuel de référence. Celui-ci étant pris uniforme, ν est égale à la mesure de Lebesgue multipliée par l'intensité λ . Il est facile de

montrer que ces deux mesures définissent une mesure symétrique ψ_n sur $\{E_n \times E_{n+1}\} \cup \{E_{n+1} \times E_n\}$ en décomposant la mesure μ du processus de Poisson de référence de la même façon que cela est fait dans l'équation (2.2) :

$$\begin{aligned}\psi_n^-(B_{n+1}, A_n) &= \frac{e^{-\lambda|F|}}{(n+1)!} \int_{F^{n+1}} \mathbf{1}_{B_{n+1}}(\mathbf{x}^+) \underbrace{\sum_{u \in \mathbf{x}} \mathbf{1}_{A_n}(\mathbf{x}^+ \setminus u)}_{(n+1)\mathbf{1}_{A_n}(\{x_1, \dots, x_n\})} d\nu^{n+1}(x_1, \dots, x_{n+1}) \\ \psi_n^-(B_{n+1}, A_n) &= \frac{e^{-\lambda|F|}}{n!} \int_F \int_{F^n} \mathbf{1}_{B_{n+1}}(\mathbf{x}^- \cup x_{n+1}) \mathbf{1}_{A_n}(\mathbf{x}^-) d\nu^n(x_1, \dots, x_n) d\nu(x_{n+1}) \\ \psi_n^-(B_{n+1}, A_n) &= \psi_n^+(A_n, B_{n+1})\end{aligned}$$

ψ est ensuite définie sur $E \times E$ comme la mesure concentrée sur $\bigcup_{n=0}^{\infty} \{\{E_n \times E_{n+1}\} \cup \{E_{n+1} \times E_n\}\}$ et égale à ψ_n sur $\{E_n \times E_{n+1}\} \cup \{E_{n+1} \times E_n\}$. Elle est symétrique et vérifie bien la propriété : $\psi(A, B) = 0 \implies \pi Q(A, B) = 0$. Finalement, la dérivée de Radon-Nikodym dans le cas d'une naissance uniforme sur F se déduit directement des équations (2.55) et (2.57) :

$$D(\mathbf{x}, \mathbf{x} \cup u) = \frac{p_b f(\mathbf{x})}{\lambda|F|} \quad (2.59)$$

De même, la dérivée de Radon-Nikodym dans le cas d'une mort uniforme dans la configuration \mathbf{x} se déduit directement des équations (2.56) et (2.58) :

$$D(\mathbf{x}, \mathbf{x} \setminus u) = \frac{p_d f(\mathbf{x})}{n(\mathbf{x})} \quad (2.60)$$

Finalement, les taux de Green $R(\mathbf{x}, \mathbf{x} \cup u)$ et $R(\mathbf{x}, \mathbf{x} \setminus u)$ respectivement associés à une naissance et à une mort sont :

$$R(\mathbf{x}, \mathbf{x} \cup u) = \frac{p_d}{p_b} \frac{\lambda|F|}{n(\mathbf{x}) + 1} \frac{f(\mathbf{x} \cup u)}{f(\mathbf{x})} \quad (2.61)$$

$$R(\mathbf{x}, \mathbf{x} \setminus u) = \frac{p_b}{p_d} \frac{n(\mathbf{x})}{\lambda|F|} \frac{f(\mathbf{x} \setminus u)}{f(\mathbf{x})} \quad (2.62)$$

Remarque 5 La plupart du temps, le rapport $\frac{p_d}{p_b}$ n'intervient pas, le choix usuel étant de prendre $p_d = p_b = 0.5$ ("probabilité de proposer une naissance" = "probabilité de proposer une mort").

Remarque 6 La même méthodologie peut être appliquée pour un processus de référence non homogène de mesure d'intensité ν . Dans ce cas, si l'on propose les nouveaux points avec une probabilité proportionnelle à ν , on retrouve les mêmes taux d'acceptation.

Remarque 7 Dans le cas de processus ponctuels marqués, ces taux sont encore valables si le processus de référence est un processus ponctuel marqué de Poisson de mesure d'intensité $\nu \otimes P_M$ (donnée par l'équation (2.4)) et si les marques associées à l'objet proposé (dans le cas d'une naissance) sont générées suivant la mesure de probabilité associée aux marques P_M .

2.5.3 Construction efficace de l'algorithme MCMC à sauts réversibles

Bien qu'il soit suffisant de définir un noyau de type "naissance et mort" uniforme dans l'algorithme MH pour échantillonner un processus Markov objet [Geyer et Møller, 1994], la formulation proposée par [Green, 1995] permet de définir d'autres types de perturbations rendant l'algorithme plus efficace, comme des petites perturbations d'objets ou des fusions et divisions d'objets. Ainsi, le noyau de proposition Q peut être décomposé en sous-noyaux q_i , chacun correspondant à un mouvement réversible (naissance-mort, fusion-division, transformations symétriques, etc.). L'algorithme utilisant cette décomposition de noyaux est donné dans la table 2.3. L'algorithme est le

<p>Initialisation par $X_0 = \mathbf{x}_0$ avec $\mathbf{x}_0 \in E / f(x_0) > 0$. Au temps t, $X_t = \mathbf{x}$:</p> <ol style="list-style-type: none"> 1. Choisir aléatoirement un type de perturbation i 2. Proposer $\mathbf{y} \sim q_i(\mathbf{x} \rightarrow \cdot)$ 3. Évaluer le taux de Green $R_i(\mathbf{x}, \mathbf{y})$ donné par l'équation (2.63) 4. Accepter \mathbf{y} avec la probabilité $\alpha_i(\mathbf{x}, \mathbf{y}) = \min(1, R_i(\mathbf{x}, \mathbf{y}))$
--

TAB. 2.3 – Algorithme MH avec une décomposition du noyau de proposition.

même que dans la table 2.2, en remplaçant Q par un noyau choisi aléatoirement q_i . A chaque noyau q_i , est associée la dérivée de Radon-Nikodym correspondante : $D_i = \frac{d\pi q_i}{d\psi_i}$. Et de la même façon que précédemment, le taux de Green associé à une perturbation de type i de \mathbf{x} à \mathbf{y} est donné par :

$$R_i(\mathbf{x}, \mathbf{y}) = \frac{D_i(\mathbf{y}, \mathbf{x})}{D_i(\mathbf{x}, \mathbf{y})} \quad (2.63)$$

En effet, il suffit que chaque sous-noyau vérifie la condition de réversibilité (2.44) pour que le noyau global vérifie lui-même cette condition.

2.5.4 Quel algorithme choisir ?

Grâce au calcul des taux de naissance, l'échantillonneur de type Naissance-Mort (NM) a l'avantage de rejoindre rapidement les configurations de plus grande vraisemblance. A l'inverse, l'étape d'acceptation/refus de l'algorithme Metropolis-Hastings (MH) est pénalisante en terme de vitesse de convergence : de nombreuses propositions de perturbation sont refusées lorsque la vraisemblance est piquée (température faible) [Clifford et Nicholls, 1994]. Cependant, les taux de naissance sont en général

difficiles à calculer et doivent être remis à jour à chaque itération de l’algorithme NM. Par conséquent, une itération de l’échantillonneur NM est nettement plus lourde en temps de calcul qu’une itération de type MH. Même pour des modèles simples tels que le processus de Strauss, pourtant bien adaptés à un échantillonnage NM, [Imberty et Descombes, 2000] montrent que l’algorithme NM n’est efficace que dans le cas d’une intensité faible (peu de points) et d’interactions à faible portée (faible rayon d’interaction). Plus simple à implanter, l’algorithme MH est en général plus efficace en termes de temps de calcul (même avec un unique noyau “naissance et mort” uniforme); et il l’est d’autant plus que les interactions sont fortes. Or, l’intérêt des processus Markov objets pour l’extraction d’objets en imagerie est de pouvoir facilement intégrer des connaissances *a priori* sur l’agencement des objets dans la scène (alignement des arbres, connectivité des routes, alignement des bâtiments, etc.). Les modèles proposés dans ce cadre sont donc caractérisés par de fortes interactions entre les objets. On aura donc tout intérêt à utiliser un algorithme de type MH.

De plus, ce type d’algorithme est beaucoup plus flexible que l’algorithme NM. Ainsi, comme nous l’avons indiqué dans le paragraphe précédent, le noyau de proposition peut se décomposer en sous-noyaux, chacun correspondant à un mouvement réversible. Ainsi, de petites perturbations d’objets, telles que des rotations ou des translations, sont souvent proposées puisqu’elles correspondent à un raccourci entre une mort suivi d’une naissance et sont très utiles lorsqu’un objet est approximativement bien placé. D’autres perturbations peuvent être particulièrement adaptées à un type de modèle. Par exemple, dans le cas de modèles faisant intervenir des interactions attractives, il est pertinent de proposer des naissances dans un voisinage des objets de la configuration courante. Dans le cas d’objets complexes tels que des polygones, des perturbations de type fusion/division seront également très utiles pour sortir d’optima locaux. Enfin, des champs externes fondés sur une pré-détection de bas niveau, peuvent d’être utilisés pour définir une naissance fondée sur ce champs externe, plus pertinente qu’une naissance uniforme. L’efficacité de l’algorithme sera donc conditionnée par la façon dont on a construit l’algorithme, ou plutôt le noyau de proposition.

Par conséquent, l’algorithme à utiliser pour l’échantillonnage de processus construits pour la modélisation de la scène par une collection d’objets (objets dont l’agencement dans la scène est caractérisé par de fortes interactions) est, sans aucun doute, l’algorithme de type **Metropolis-Hastings avec décomposition du noyau**, où chaque sous-noyau est adapté aux objets à extraire et correspond à des perturbations pertinentes (guidées par les contraintes, les données, etc.).

2.5.5 Sous quelles conditions l’algorithme converge-t-il ?

Tous les algorithmes que nous avons présentés vérifient la stationnarité de π (mesure d’intérêt) par rapport à $\{X_t\}$, la chaîne de Markov construite. En effet, ils sont construits de façon à vérifier la réversibilité du noyau de transition de la chaîne par rapport à π , ce qui implique la stationnarité de π . Cependant, la stationnarité ne suffit pas

à garantir la convergence vers la mesure π . Des conditions supplémentaires, fondées sur les propriétés des chaînes de Markov décrites dans le paragraphe 2.4, sont à vérifier. La condition de stabilité locale d'un processus ponctuel est particulièrement intéressante pour montrer la convergence ergodique. Considérons en premier lieu l'algorithme MH (table 2.2) avec un noyau de type "naissance et mort" uniforme dont les taux de Green associés à une naissance et une mort sont donnés par les équations 2.61 et 2.62 avec $p_b = p_d = 1$. Soit A un tel algorithme avec n'importe quelle condition initiale (de densité de probabilité non nulle). [Geyer et Møller, 1994, Geyer, 1999] montrent que :

Proposition 3 *Si la condition de stabilité locale (équation (2.16)) est vérifiée par la densité non normalisée du processus, alors l'algorithme A simule une chaîne de Markov géométriquement ergodique.*

Autrement dit, la chaîne ainsi construite converge ergodiquement vers π avec une vitesse géométrique. Pour un autre noyau, la preuve de convergence devra être établie (au moins de la convergence ergodique). Regardons alors plus en détail comment [Geyer et Møller, 1994, Geyer, 1999] établissent la convergence de la chaîne sous la condition de stabilité locale du processus.

Irréductibilité

Si la condition de stabilité locale (équation (2.16)) est vérifiée par la densité non normalisée h du processus, alors la probabilité d'accepter le retrait d'un élément u de la configuration courante \mathbf{x} est plus grande que :

$$\min \left(1, \frac{n(\mathbf{x})}{\lambda|F|} \frac{h(\mathbf{x} \setminus u)}{h(\mathbf{x})} \right) \geq \frac{1}{M\lambda|F|} \quad (2.64)$$

où M est la constante de la condition (2.16)), prise assez grande pour que $\frac{1}{M\lambda|F|} < 1$. Rappelons que la probabilité p_d de choisir une mort est égale à $1/2$. Aussi, $P(\emptyset, \{\emptyset\}) \geq \frac{1}{2}$ où \emptyset dénote la configuration vide et $\{\emptyset\}$ l'ensemble composé de la configuration vide ($\{\emptyset\} = E_0$). Si $t \geq n(\mathbf{x})$, on a donc :

$$P^t(\mathbf{x}, \{\emptyset\}) \geq P^{n(\mathbf{x})}(\mathbf{x}, \{\emptyset\}) P^{t-n(\mathbf{x})}(\emptyset, \{\emptyset\}) \geq \left(\frac{1}{2M\lambda|F|} \right)^t \quad (2.65)$$

Cela implique alors que la probabilité d'atteindre l'ensemble $\{\emptyset\}$ en un nombre fini d'itérations ($t \geq n(\mathbf{x})$), à partir de n'importe quelle configuration finie \mathbf{x} , est positive. Posons ϕ mesure sur \mathcal{B} (tribu de E) définie par :

$$\phi(A) = \mathbf{1}(\emptyset \in A)$$

L'équation (2.65) implique donc la ϕ -irréductibilité. Par la proposition 1, si π est la mesure stationnaire, alors la chaîne est π -irréductible.

Apériodicité

La propriété la plus facile à vérifier est l'apériodicité puisqu'il suffit de vérifier qu'il existe un état tel que la probabilité de rester dans cet état soit supérieure à 0 (la chaîne est alors fortement apériodique). Ainsi, l'apériodicité est vérifiée par n'importe quel échantillonneur qui a des probabilités non nulles de rejet, dont l'algorithme A .

Ensembles petits

Soit $C = \{\mathbf{x} \in E : n(\mathbf{x}) \leq m\}$ et c la partie droite de l'équation (2.65), alors :

$$P^m(\mathbf{x}, A) \leq c \phi(A) \quad \forall A \in \mathcal{B} \quad \forall \mathbf{x} \in C$$

Ainsi, l'équation (2.65) implique également que n'importe quel ensemble pour lequel $n(\mathbf{x})$ est borné est petit.

Ergodicité géométrique

La probabilité d'accepter l'ajout d'un élément u vérifie :

$$\min \left(1, \frac{\lambda|F|}{n(\mathbf{x})+1} \frac{h(\mathbf{x} \cap u)}{h(\mathbf{x})} \right) \leq \frac{M\lambda|F|}{n(\mathbf{x})+1} \quad (2.66)$$

Pour tout $\epsilon \in]0, 1[$, cette probabilité est inférieure à ϵ quand $n(\mathbf{x}) \geq K_\epsilon$ avec $K_\epsilon = \frac{M\lambda|F|}{\epsilon}$. De plus, si $n(\mathbf{x}) \geq K_\epsilon$, la probabilité d'accepter un retrait vérifie :

$$\min \left(1, \frac{n(\mathbf{x})}{\lambda|F|} \frac{h(\mathbf{x} \setminus u)}{h(\mathbf{x})} \right) \geq \min \left(1, \frac{n(\mathbf{x})}{M\lambda|F|} \right) = 1 \quad (2.67)$$

Posons $V(\mathbf{x}) = A^{n(\mathbf{x})}$, où $A > M\lambda|F|$. Soit $\mathbf{x} \in E_n$. Remarquons que l'état suivant ne pourra appartenir qu'à E_{n+1} ou E_{n-1} dans le cas de l'acceptation d'une naissance ou d'une mort ou E_n dans le cas d'un rejet. On a alors pour : $n \geq K_\epsilon$

$$\begin{aligned} \mathbb{E}[V(X_1)|X_0 = \mathbf{x}] &= \int V(\mathbf{y})P(\mathbf{x}, d\mathbf{y}) \\ &= \int_{E_{n+1}} A^{n+1}P(\mathbf{x}, d\mathbf{y}) + \int_{E_n} A^n P(\mathbf{x}, d\mathbf{y}) + \int_{E_{n-1}} A^{n-1}P(\mathbf{x}, d\mathbf{y}) \\ &= A^n \left(A P(x, E_{n+1}) + P(x, E_n) + \frac{1}{A} P(x, E_{n-1}) \right) \\ &\leq V(\mathbf{x}) \underbrace{\left(\frac{1}{2}\epsilon A + \frac{1}{2} + \frac{1}{2} \frac{1}{A} \right)}_{B(\epsilon)} \end{aligned}$$

le facteur $1/2$ correspondant aux probabilités de proposer une mort ou une naissance. Puisque le choix de ϵ est libre et que $B(\epsilon)$ converge vers $\frac{1}{2}(1 + \frac{1}{A})$ quand ϵ tend vers 0, on peut choisir ϵ tel que :

$$\mathbb{E}[V(X_1)|X_0] \leq \lambda V(\mathbf{x}), \quad n(\mathbf{x}) \geq K_\epsilon$$

avec $\lambda < 1$. Pour $C = \{\mathbf{x} \in E : n(\mathbf{x}) \leq K_\epsilon\}$, la condition de drift pour la récurrence (2.31) est vérifiée. Il y a donc convergence ergodique puisque l'irréductibilité et l'apériodicité ont été vérifiées. Pour $\mathbf{x} \notin C$, la condition de drift géométrique (2.34) est également vérifiée. Pour $\mathbf{x} \in C$, $\mathbb{E}[V(X_1)|X_0 = \mathbf{x}] \leq A^{K_\epsilon+1}$. La condition de drift géométrique (2.34) est donc vérifiée pour $b = A^{K_\epsilon+1}$. La chaîne ainsi construite est donc géométriquement ergodique.

Ajout de sous-noyaux

Supposons à présent que l'on définisse un noyau composé du noyau "naissance et mort" uniforme et de sous-noyaux correspondant à des perturbations où aucun saut entre les espace E_n (par exemple : translation d'un point, rotation d'un objet, etc.). Sous la condition de stabilité locale, on obtient avec la même procédure la convergence ergodique (ergodicité géométrique). En revanche, avec tout autre sous-noyau permettant de réaliser des petits sauts, il faudra vérifier que les probabilités d'accepter des ajouts et des retraits d'objet sont telles que la convergence ergodique puisse être vérifiée.

2.5.6 Détecter la convergence en pratique

Une fois que l'échantillonneur a été défini, il reste le problème du critère d'arrêt de l'algorithme. Nous avons montré comment vérifier la convergence ergodique de l'échantillonneur. Cependant, même dans le cas d'une ergodicité géométrique, ces résultats théoriques ne nous disent pas quand arrêter l'algorithme. Parfois, il est néanmoins possible que l'algorithme lui-même indique que l'équilibre a été atteint. Des échantillons exacts plutôt qu'approximés sont alors obtenus. Ce sont des méthodes appelées méthodes de simulation exacte [Kendall et Møller, 2000, van Lieshout, 2000]. Elles sont utilisées en inférence statistique où il est important d'être sûr d'avoir convergé pour le calcul des statistiques d'intérêt. Pour la détection d'objet, l'importance est moindre et la complexité des modèles rend la mise en oeuvre des méthodes de simulation exacte délicate. Nous utiliserons plutôt un critère d'arrêt fondé sur la détection de la convergence de moyenne(s) empirique(s), ce type de détection fournissant un diagnostic fiable pour algorithmes MCMC [Robert et Casella, 1999].

Considérons la moyenne empirique d'une fonction g obtenue par l'échantillonneur au temps t :

$$\bar{g}_t = \frac{1}{t} \sum_{i=1}^t g(\mathbf{x}_i)$$

Par exemple, un choix typique pour g est de considérer la fonction mesurant le nombre de points d'une configuration : $g(\mathbf{x}) = n(\mathbf{x})$. Pour une plus grande fiabilité, on pourra également considérer la convergence de plusieurs moyennes empiriques : nombre de points total, nombre de points en interaction, etc. Nous donnons ici une façon d'effectuer le diagnostic pour une fonction mesurable g quelconque. Après un nombre minimal

d'itérations t_{min} , la moyenne empirique est évaluée après toutes les N itérations :

$$\bar{g}_k = \frac{1}{k} \sum_{i=1}^k g(\mathbf{x}_{l(i)})$$

où $l(i) = t_{min} + 1 + N(i - 1)$. Un tel sous-échantillonnage permet de limiter l'effet d'auto-corrélation de la chaîne. Or, une chaîne suivant une dynamique de Metropolis-Hastings est particulièrement corrélée puisque, d'une part, les perturbations proposées sont de petites modifications de la configuration courante (petits sauts, perturbation d'un seul élément) et, d'autre part, ces perturbations sont souvent refusées : $\mathbf{x}_{t+1} = \mathbf{x}_t$. Soit M_{ref} la moyenne empirique évaluée à l'instant t_{ref} . La différence absolue entre M_{ref} et la moyenne suivante, *i.e.* calculée à l'itération $(t_{ref} + N)$, est évaluée. Si cette différence est inférieure à une petite constante ϵ (=test vérifié), on poursuit le test entre M_{ref} et la moyenne calculée à l'itération $(t_{ref} + 2N)$. On procède ainsi tant que le test est vérifié. Si le test est vérifié pour n_{test} itérations successives, alors la convergence est supposée atteinte. Si le test n'est pas vérifié pour une itération j , alors la procédure de test est réinitialisée avec $t_{ref} = j$. L'algorithme correspondant est donné dans la table 2.4.

1. Effectuer $t_{min} + 1$ itérations de l'algorithme d'échantillonnage
2. Initialiser : $t = t_{min} + 1, k = 1, M_{ref} = g(x_t), T = 0$
3. Évaluer : $\bar{g}_k = \frac{(k-1)\bar{g}_{k-1} + g(x_t)}{k}$
4. Tester : $\begin{cases} \text{Si } \bar{g}_k - M_{ref} < \epsilon, & T = T + 1 \\ \text{Sinon, } & T = 0 \text{ et } M_{ref} = \bar{g}_k \end{cases}$
5. Si $T = n_{test}$, arrêt de l'algorithme
Sinon : a. Effectuer N itérations de l'algorithme d'échantillonnage
b. Mettre à jour $k = k + 1, t = t + N$
c. Aller en 3

TAB. 2.4 – Critère d'arrêt fondé sur la convergence de moyenne empirique.

2.6 Optimisation

Afin d'extraire la collection d'objets présente dans l'image dont nous disposons, nous cherchons une configuration qui maximise la densité f du processus par rapport

au processus de Poisson de référence. C'est un problème non convexe dont l'optimisation directe est impossible du fait de la grande taille de l'espace d'état E , défini par l'équation 2.6. L'optimisation se fait donc généralement par un recuit simulé permettant d'estimer ce maximum.

2.6.1 Recuit simulé

Cet algorithme est donné dans la table 2.5. Il permet d'accéder aux pics de la

Simulations successives de processus spécifiés par une densité f_T par rapport au processus de référence :

$$f_T(\mathbf{x}) \propto [h_d(\mathbf{x}) h_p(\mathbf{x})]^{1/T}$$

avec T diminuant pas à pas vers zéro.

TAB. 2.5 – Recuit simulé.

densité en simulant successivement des processus définis par la densité f à la puissance $1/T$, la température T diminuant pas à pas vers zéro. Les simulations successives sont réalisées par un algorithme d'échantillonnage décrit dans le paragraphe 2.5. Remarquons que, si $T = \infty$, alors le processus simulé correspond au processus de référence. Au début de l'algorithme (température élevée), le processus n'est pas très sélectif. Cela permet de passer d'un mode à l'autre de la densité assez facilement. Plus la température diminue, plus les configurations ayant une densité élevée sont favorisées. Enfin, pour une température nulle la densité du processus correspond à des diracs au niveau des maxima globaux de la densité. Une preuve de convergence pour une décroissance logarithmique de la température T est donnée dans [van Lieshout, 1993] (avec un échantillonneur de type naissance-mort). Ce schéma théorique est de la forme suivante :

$$T(t) = \frac{c}{\log(1+t)} \quad (2.68)$$

où c est plus grand que la profondeur du plus profond minimum local de l'énergie U qui n'est pas un minimum global. Rappelons qu'un minimum de l'énergie correspond à un maximum de la densité : $f(\mathbf{x}) = \exp(-U(\mathbf{x}))$.

2.6.2 Schéma de décroissance géométrique

Le schéma de décroissance théorique ne garantit la convergence qu'en un temps infini. Mais en pratique, l'algorithme est itéré en un nombre fini d'itérations. Or, une décroissance logarithmique est très lente comparativement à un schéma de décroissance géométrique défini par :

$$T(t+1) = cT(t) \quad (2.69)$$

où c est une constante proche de 1. C'est un des schéma les plus largement utilisés du fait de sa simplicité. La constante choisie pour les processus ponctuels devra être plus élevée que celles généralement choisies pour les champs de Markov. En effet, pour les processus ponctuels marqués une itération ne correspond qu'à la perturbation d'un objet ou deux, alors que pour les champ de Markov une itération correspond à un balayage sur tous les pixels de l'image. Ainsi, dans le chapitre 3, nous utiliserons majoritairement la constante $c = 0.999999$ avec une décroissance à chaque itération. Dans les cas difficiles, où il est important de décroître lentement en température, nous pourrons utiliser une décroissance par palier : la décroissance n'est alors appliquée que toutes les N itérations.

2.6.3 Schéma de décroissance adaptatif

Un choix plus pertinent consiste à utiliser un schéma adaptatif de la température qui ajuste la température suivant le comportement de l'algorithme. L'idée est de rester le plus proche de l'équilibre tout en diminuant la température le plus vite possible. Une implémentation de cette idée est décrite par [Hoffmann *et al.*, 1991]. Elle consiste à implanter un schéma géométrique par palier où la longueur des paliers dépend des fluctuations de la moyenne empirique de l'énergie. Remarquons que lorsque l'équilibre est atteint, cette moyenne est supposée fluctuer autour de la vraie valeur. L'énergie moyenne empirique est évaluée sur des intervalles de temps réguliers. Pour l'intervalle $i = [t_i, t_i + N]$, l'énergie moyenne est :

$$\langle U \rangle_i = \frac{1}{N} \sum_{k=t_i}^{t_i+N} U(\mathbf{x}_k) \quad (2.70)$$

La température reste constante sur chaque intervalle de longueur N . Lorsque la fin d'un intervalle i est atteinte, on compare l'énergie moyenne avec celle de l'intervalle $i - 1$. La décroissance est acceptée seulement si la nouvelle moyenne est supérieure à la précédente :

$$T_i = \begin{cases} T_{i-1} & \text{si } \langle U \rangle_{i+1} \leq \langle U \rangle_i \\ cT_{i-1} & \text{si } \langle U \rangle_{i+1} > \langle U \rangle_i \end{cases} \quad (2.71)$$

Ainsi, on est assuré de remonter au moins une fois en énergie à la même température et donc de ne pas être trop éloigné de l'équilibre. C'est ce schéma que nous utilisons pour l'échantillonnage des processus d'objets complexes (chapitres 4 et 5).

Chapitre 3

Extraction des réseaux linéiques par processus de segments

Nous proposons dans ce chapitre de réaliser une extraction non supervisée du réseau linéique (routes, rivières) présent dans des images aériennes ou satellitaires via une modélisation du réseau par une collection de segments, où chaque segment correspond à une section de route ou de rivière. Pour cela, nous plaçons dans un cadre stochastique et utilisons des processus ponctuels marqués comme modèles *a priori*. Nous décrivons tout d’abord trois modèles *a priori* permettant d’incorporer de fortes contraintes géométriques et topologiques qui caractérisent les réseaux d’intérêt au travers de la définition d’interaction entre segments. Les deux derniers modèles sont des extensions du modèle “Candy”, introduit par [Stoica, 2001] pour l’extraction du réseau routier. Ces deux extensions font intervenir des coefficients mesurant la qualité des interactions entre objets. Nous expliquons ensuite comment construire le terme d’attache aux données. Plusieurs approches sont proposées, chacune d’entre elles étant fondée sur une évaluation de l’homogénéité locale du réseau et du contraste avec son environnement proche. Le choix de l’approche pourra se faire selon l’importance donnée à la précision de l’extraction par rapport à celle donnée à l’efficacité de l’algorithme. Une fois le modèle complet défini, il reste le problème du choix des paramètres. Nous proposons ici un calibrage des paramètres assurant que la configuration optimale vérifie de bonnes propriétés. L’optimisation est réalisée par recuit simulé sur un algorithme MCMC à sauts réversibles, construit de façon à accélérer la convergence par l’ajout de perturbations pertinentes. Les premiers résultats expérimentaux obtenus permettent de vérifier l’intérêt d’une prise en compte de la qualité des interactions de façon continue. La performance de cette modélisation par processus de segments est finalement montrée sur de nombreux exemples. En effet, le réseau obtenu par recuit simulé est en général constitué de longues branches continues et de faible courbure, et présente relativement peu de surdétectations et d’omissions étant donné le caractère complètement automatique de la méthode.

3.1 Modèles a priori

Comme nous l'avons vu dans le chapitre 2, une façon simple de procéder pour introduire un *a priori* sur l'agencement des objets dans la scène est de définir un processus de Poisson de référence et de construire une densité h_p incorporant des connaissances *a priori* sur la forme des objets et leur agencement et permettant de spécifier un processus par rapport au processus de référence. Après une brève description du processus de référence, nous présentons dans ce paragraphe trois densités *a priori* initialement construites pour la modélisation de réseaux routiers.

3.1.1 Processus de référence

Le processus de référence utilisé dans ce chapitre est un processus de Poisson homogène (ou uniforme) sur la fenêtre d'observation F . Comme illustré par la figure 3.1, les objets (*i.e.* les points marqués) du processus sont des segments. Chaque segment $s \in \mathbf{s}$ est décrit par :

- son centre $p = (x, y) \in F = [0, X_{max}] \times [0, Y_{max}]$
- sa longueur $L \in [L_{min}, L_{max}]$
- son orientation $\theta \in [0, \pi]$

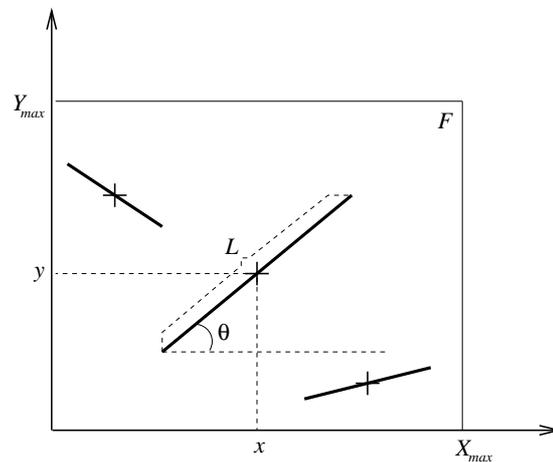


FIG. 3.1 – Objet type du processus.

Sous la mesure μ du processus de référence, le nombre de points suit une loi de Poisson ; les points (*i.e.* centres des segments) sont uniformément distribués dans F ; et les marques associées (*i.e.* leurs longueurs et leur orientations) sont uniformément distribuées dans $M = [L_{min}, L_{max}] \times [0, \pi]$. Les trois modèles qui suivent sont des

processus ponctuels marqués spécifiés par une densité h_p par rapport à μ .

3.1.2 Modèle “Candy”

Le modèle “Candy” a été introduit par [Stoica, 2001] pour l’extraction du réseau routier. Ce modèle est fondé sur l’idée qu’un réseau routier peut être assimilé à une réalisation d’un processus Markov objet, où les objets correspondent à des segments en interaction. Trois types d’interactions sont définis :

- une interaction fondée sur une relation de connexion, favorisant la continuité du réseau ;
- une interaction fondée sur une relation de proximité, pénalisant les regroupements de segments ;
- une interaction fondée sur une relation de mauvais alignement, pénalisant les fortes courbures.

La relation de **connexion** \sim_c est définie à une constante ϵ près puisqu’une connexion exacte est de probabilité nulle sous la mesure de référence μ . Deux segments vérifient cette relation si la différence entre exactement deux de leur extrémités est inférieure à ϵ . La relation de connexion permet de définir trois états possibles pour chaque segment comme cela est illustré par la figure 3.2. L’état de chaque segment de la configuration

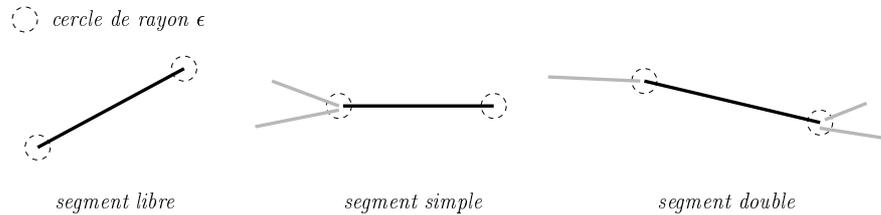


FIG. 3.2 – Trois états dérivant de la relation de connexion \sim_c .

intervient dans la densité. Ainsi, les segments libres (*i.e.* non connectés) et les segments simples (*i.e.* connectés par une seule de leurs extrémités) sont pénalisés par des potentiels constants et positifs, ω_f et ω_s , afin de favoriser un réseau continu.

Une relation de **proximité** \sim_p est définie afin d’éviter les regroupements de segments. Deux segments s_1 et s_2 vérifient $s_1 \sim_p s_2$ si les deux conditions suivantes sont vérifiées :

- (C₁) leurs centres se situent à une distance inférieure à la demi-longueur du segment le plus long ;
- (C₂) la différence d’orientation τ_{12} (en valeur absolue et modulo $\pi/2$) entre s_1 et s_2 vérifie :

$$\tau_{12} = \min\{|\theta_1 - \theta_2|, \pi - |\theta_1 - \theta_2|\} < \tau_p \quad (3.1)$$

où τ_p est un seuil à fixer par l’utilisateur.

La figure 3.3 illustre cette définition. Un potentiel constant et positif ω_p est assigné aux paires de segments vérifiant la relation de proximité \sim_p .

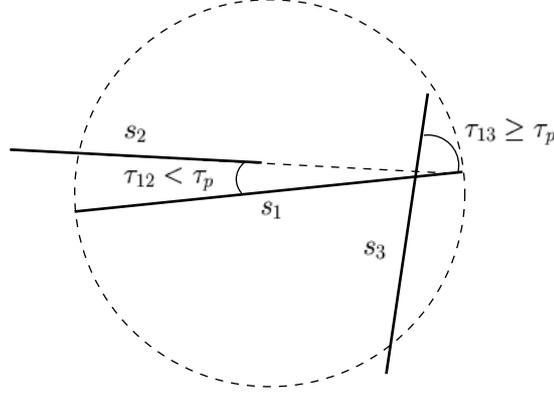


FIG. 3.3 – Relation de proximité \sim_p . Le cercle décrit la zone d’influence interne du segment s_1 . Les deux paires de segments (s_1, s_2) et (s_1, s_3) vérifient la condition (C_1) , *i.e.* les centres des petits segments s_2 et s_3 sont dans la zone d’influence du grand segment s_1 . Mais seuls s_1 et s_2 vérifient $s_1 \sim_p s_2$.

La relation de **mauvais alignement** \sim_a est introduite afin de contrôler la courbure du réseau. Une zone d’influence externe $Z(s)$ est définie pour chaque segment $s = (p, L, \theta)$ par l’union de deux disques de rayon $L/4$ autour des deux extrémités e_1 et e_2 du segment s :

$$Z(s) = b(e_1, \frac{L}{4}) \cup b(e_2, \frac{L}{4}) \quad (3.2)$$

Deux segments s_1 et s_2 vérifient $s_1 \sim_a s_2$ si les trois conditions suivantes sont vérifiées :

- (C_4) leurs centres se situent à une distance supérieure à la demi-longueur du segment le plus long ;
- (C_5) exactement une extrémité de s_1 se trouve dans $Z(s_2)$ ou exactement une extrémité de s_2 se trouve dans $Z(s_1)$;
- (C_6) la différence d’orientation τ_{12} entre s_1 et s_2 vérifie :

$$\tau_{12} = \min\{|\theta_1 - \theta_2|, \pi - |\theta_1 - \theta_2|\} > \delta \quad (3.3)$$

où δ est un seuil à fixer par l’utilisateur.

La figure 3.4 illustre cette définition. Un potentiel constant et positif ω_a est assigné aux paires de segments vérifiant \sim_a .

Finalement, le modèle “Candy” est spécifié par la **densité a priori** suivante :

$$h_p(\mathbf{s}) \propto \beta^{n(\mathbf{s})} \exp[-(\omega_f n_f(\mathbf{s}) + \omega_s n_s(\mathbf{s}) + \omega_p n_p(\mathbf{s}) + \omega_a n_a(\mathbf{s}))] \quad (3.4)$$

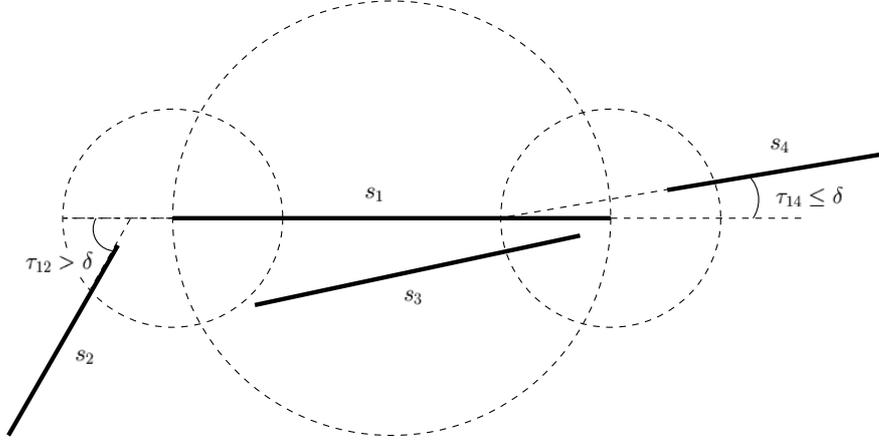


FIG. 3.4 – Relation de mauvais alignement \sim_a . Le cercle interne correspond à la zone d'influence interne mais aussi à une zone non prise en compte par la relation de mauvais alignement avec s_1 . Ainsi, s_3 dont le centre est dans ce cercle, ne sera pas pris en compte. Les deux cercles externes correspondent à la zone d'influence externe de s_1 . Remarquons que s_1 et s_4 sont dans cette zone, mais seuls s_1 et s_2 sont mal-orientés : $s_1 \sim_a s_2$.

où β est un facteur d'intensité, les ω_i sont des potentiels constants et strictement positifs, $n(\mathbf{s})$ est le nombre total de segments dans la configuration \mathbf{s} , $n_l(\mathbf{s})$ le nombre de segments libres dans \mathbf{s} , $n_s(\mathbf{s})$ le nombre de segments simples dans \mathbf{s} , $n_p(\mathbf{s})$ le nombre de paires de segments vérifiant \sim_p dans \mathbf{s} , et $n_a(\mathbf{s})$ le nombre de paires de segments vérifiant \sim_a dans \mathbf{s} .

Remarque 8 Nous avons en réalité simplifié le modèle proposé initialement par Radu Stoica en supprimant l'a priori sur la longueur des segments.

Cette densité spécifie un processus de point marqué bien défini, puisque la condition de stabilité de Ruelle est vérifiée. La condition plus forte de **stabilité locale** est également vérifiée. Plus précisément, [Stoica, 2001] montre que :

$$\frac{h_p(\mathbf{s} \cup u)}{h_p(\mathbf{s})} \leq \beta \exp[-12 (\omega_f + \omega_s)] \quad (3.5)$$

où h_p est donnée par l'équation (3.4). Les propriétés markoviennes de ce modèle ont été étudiées par [van Lieshout et Stoica, 2001], qui montrent que c'est un **processus de Markov** par rapport à la relation \sim définie par :

$$s \sim s' \iff d(p_s - p_{s'}) \leq 2 (L_{max} + \epsilon) \quad (3.6)$$

où p_s désigne le centre du segment s et d la distance euclidienne dans \mathbb{R}^2 . Autrement dit, la probabilité locale d'un segment s sachant le reste de la configuration ne dépend que

des segments dont le centre appartient au disque centre p_s et de rayon égal à $2(L_{max} + \epsilon)$.

La densité de ce modèle a été construite de façon à incorporer des contraintes fondées sur les caractéristiques géométriques et topologiques des réseaux routiers (cf. paragraphe 1.1.2). En effet, la connexion entre segments est favorisée, les regroupements de segments sont limités et les points de forte courbure sont pénalisés. Le modèle “Candy” présente néanmoins un défaut en termes de distinction de configurations de qualités différentes, en termes de continuité et de courbure. En effet, la densité donnée par l'équation (3.4) prendra la même valeur pour des configurations présentant des interactions de qualités différentes, comme cela est illustrée par la figure 3.5. Un réseau obtenu par optimisation de la densité h_p (équation (3.4)) pourra donc présenter de petites interruptions et la courbure du réseau ne sera pas minimale. C'est pourquoi nous proposons deux nouveaux modèles prenant en compte des coefficients mesurant la qualité des interactions entre segments.

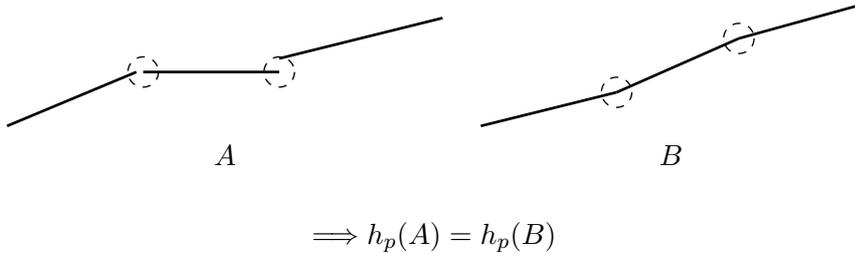


FIG. 3.5 – Un défaut du modèle “Candy”.

3.1.3 Modèle “Quality Candy”

Le modèle “Quality Candy” est une extension du modèle “Candy” obtenu en remplaçant les potentiels d'interaction constants par des fonctions de potentiel fondés sur des mesures de qualité des interactions. La forme générale de la densité est :

$$h_p(\mathbf{s}) \propto \beta^{n(\mathbf{s})} \exp \left[- \left(\omega_f n_f(\mathbf{s}) + \omega_s n_s(\mathbf{s}) + \sum_{r \in \mathcal{R}} \omega_r \sum_{\langle s_i, s_j \rangle_r} g_r(s_i, s_j) \right) \right] \quad (3.7)$$

où \mathcal{R} est l'ensemble des relations permettant la définition d'interactions par paire, $\langle s_i, s_j \rangle_r$ désigne une paire de segments vérifiant $s_i \sim_r s_j$, et g_r est une fonction de potentiel par rapport à r .

Remarque 9 Si $\mathcal{R} = \{\sim_p, \sim_a\}$ et si g_p et g_a sont constantes ($g_p \equiv g_a \equiv 1$), on retrouve la densité de type “Candy” donnée par l'équation (3.4).

Deux interactions par paire interviennent dans le modèle “Quality Candy”. La première interaction est fondée sur une relation de connexion $\sim_{c'}$. La relation de connexion $\sim_{c'}$ est définie de la même façon que \sim_c avec la condition supplémentaire que les deux segments connectés forment un angle grave afin d’accélérer l’optimisation (cette deuxième condition est retirée pour l’extraction de réseaux très sinueux). Par exemple, dans la figure 3.6, s_1 et s_3 ne sont pas connectés. La deuxième interaction est fondée sur la relation de proximité \sim_p .

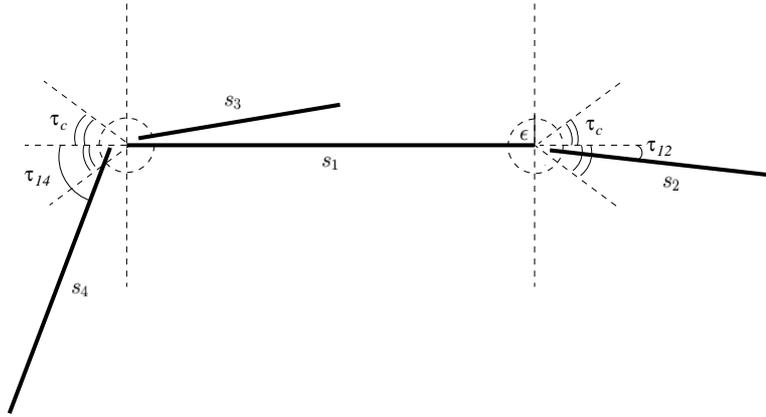


FIG. 3.6 – Différents types de potentiels associées à la relation de connexion $\sim_{c'}$. La connexion entre s_1 et s_2 est favorisée par un potentiel négatif, d’autant plus faible que τ_{12} sera petit. La connexion entre s_1 et s_4 est pénalisée par un potentiel positif. Bien que $s_1 \sim_c s_3$, $s_1 \not\sim_{c'} s_3$.

La faible courbure et la proximité des extrémités connectées sont favorisées au travers d’un **potentiel de connexion** g_c associé à chaque paire de segments connectés. Celui-ci est défini pour deux segments connectés s_i et s_j comme la moyenne de deux fonctions, l’une concernant la différence d’orientation τ_{ij} entre s_i et s_j (en valeur absolue) et l’autre la distance d_{ij} entre les deux extrémités connectées :

$$g_c(s_i, s_j) = \frac{g_\tau(\tau_{ij}) + g_\epsilon(d_{ij})}{2} \quad (3.8)$$

$$\text{avec } \begin{cases} g_\tau(\tau_{ij}) = \begin{cases} -\sigma(\tau_{ij}, \tau_c) & \text{si } \tau_{ij} < \tau_c \\ 1 & \text{sinon} \end{cases} \\ g_\epsilon(d_{ij}) = -\sigma(d_{ij}, \epsilon) \end{cases}$$

La fonction g_τ , qui concerne la différence d’orientation, donne un poids négatif aux paires de segments dont la différence d’orientation est inférieure à un seuil τ_c et un poids positif dans le cas contraire. Notons que cette expression correspond au cas où la connexion est définie pour des angles graves ; dans le cas où les connexions à angle aigu

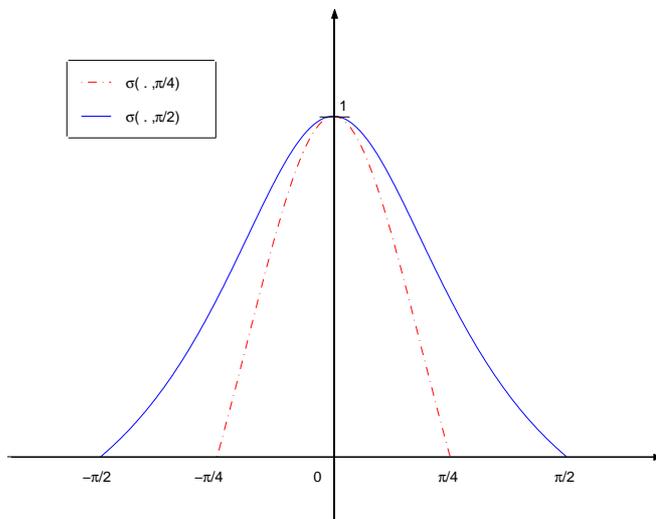


FIG. 3.7 – Fonction de qualité pour deux bornes maximales : $M = \pi/2$ et $M = \pi/4$.

sont considérées, il suffit d'assigner une valeur positive (par exemple, égale à 1) aux paires de segments connectés formant un angle aigu. La fonction g_ϵ , qui concerne la distance entre les extrémités connectées, est quant à elle toujours négative. Le potentiel en résultant $g_c(s_i, s_j)$ est du même signe que $g_\tau(\tau_{ij})$ comme cela est illustré par la figure 3.6. Les termes négatifs, favorisant les connexions de bonne qualité, dérivent de la **fonction de qualité** $\sigma(\cdot, M)$ paramétrée par son support $[-M, M]$. Elle est définie comme suit :

$$\begin{aligned} \sigma(\cdot, M) : [-M, M] &\longrightarrow [0, 1] \\ x &\longmapsto \sigma(x, M) = \frac{1}{M^2} \left(\frac{1 + M^2}{1 + x^2} - 1 \right) \end{aligned} \quad (3.9)$$

C'est une fonction positive qui prend son maximum (égal à 1) en 0, comme illustré par la figure 3.7. La qualité d'une paire de segment connectés sera donc maximale pour une différence d'orientation nulle et une distance de connexion nulle.

Ce nouveau potentiel nous permet de travailler sans la relation de mauvais alignement \sim_a puisque la partie positive du potentiel de connexion concerne le même type d'interaction.

Le **potentiel de proximité** correspondant aux paires de segments vérifiant la relation de proximité \sim_p pénalise plus ou moins les paires de segments vérifiant \sim_p selon leur différence d'orientation. L'expression de ce potentiel est donnée par l'équation suivante pour deux segments s_i et s_j concernés par cette relation et dont la différence

d'orientation est égale à τ_{ij} :

$$g_p(s_i, s_j) = \begin{cases} \infty & \text{si } \tau_{ij} < \delta_p \\ 1 - \sigma(\tau_{ij}, \tau_p) & \text{sinon} \end{cases} \quad (3.10)$$

où τ_p est le paramètre de l'équation (3.1), *i.e.* la différence d'orientation maximale telle que deux segments s_i et s_j de centres proches vérifient $s_i \sim_p s_j$; δ_p est la différence d'orientation minimale autorisée entre deux segments s_i et s_j de centres proches (pour une différence inférieure à ce seuil, la configuration sera de densité nulle). Le potentiel de type "hard core", associé aux paires de segments vérifiant \sim_p et quasiment parallèles, a été introduit pour des raisons de stabilité.

Finalement, la **densité** du modèle "Quality Candy" s'écrit sous la forme suivante :

$$h_p(\mathbf{s}) \propto \beta^{n(\mathbf{s})} \exp[-(\omega_f n_f(\mathbf{s}) + \omega_s n_s(\mathbf{s}) + \omega_c \sum_{\langle s_i, s_j \rangle_c} g_c(s_i, s_j) + \omega_p \sum_{\langle s_i, s_j \rangle_p} g_p(s_i, s_j))] \quad (3.11)$$

où les ω_i sont des poids constants et strictement positifs. Cette densité spécifie un processus de Markov par rapport à la relation \sim donnée par l'équation (3.6). De plus, ce processus Markov objet est **localement stable**.

Preuve de stabilité locale

Pour montrer la stabilité locale d'un processus spécifié par h_p par rapport à la mesure de référence μ , il suffit de trouver une borne supérieure M pour la densité de Papangelou pour tout $\mathbf{s} \in E$ tel que $h_p(\mathbf{x}) > 0$. L'expression de la densité de Papangelou du modèle "Quality Candy" est la suivante :

$$\frac{h_p(\mathbf{s} \cup u)}{h_p(\mathbf{s})} = \beta \exp \left[-\omega_f \underbrace{(n_f(\mathbf{s} \cup u) - n_f(\mathbf{s}))}_A - \omega_s \underbrace{(n_s(\mathbf{s} \cup u) - n_s(\mathbf{s}))}_B \right. \\ \left. - \omega_c \underbrace{\sum_{\substack{\langle s_i, u \rangle_c \\ s_i \in \mathbf{s}}} g_c(s_i, u)}_C - \omega_p \underbrace{\sum_{\substack{\langle s_i, u \rangle_p \\ s_i \in \mathbf{s}}} g_p(s_i, u)}_D \right] \quad (3.12)$$

Trouver des bornes inférieures pour les parties A , B , C et D est suffisant pour prouver la stabilité locale.

L'ajout d'un nouveau segment u dans la configuration \mathbf{s} donne :

$$n_f(\mathbf{s} \cup u) - n_f(\mathbf{s}) = \begin{cases} 1 & \text{si } u \text{ est libre} \\ -\#\{\{s_i \text{ libre dans } \mathbf{s}/u \sim_c s_i\}\} & \text{si } u \text{ est simple ou double} \end{cases}$$

où $\#(\cdot)$ désigne le cardinal d'un ensemble. Ici, le nombre de segments libres dans \mathbf{s} qui peut être connecté à un nouveau segment u dans \mathbf{s} ne peut être borné comme cela peut l'être pour le modèle "Candy" du fait de la modification de la définition de la relation

de connexion¹. En effet, une infinité de segments pourrait avoir une extrémité dans un même disque de rayon ϵ tout en étant libre, puisque la nouvelle connexion $\sim_{\mathcal{C}}$ ne prend pas en compte les segments formant un angle aigu. C'est la raison pour laquelle nous avons introduit un potentiel "hard core" pour les paires de segments proches (i.e. vérifiant \sim_p) formant un angle trop aigu (équation (3.10)). Si $\tau_{ij} < \delta_p \ll \pi/2$ pour s_i et $s_j \in \mathbf{s}$, on a donc $h_p(\mathbf{s}) = 0$. En ne considérant que les configurations \mathbf{s} telles que $h_p(\mathbf{s}) > 0$, le nombre de segments connectés à une extrémité e_u de u est borné par n_{max} , le nombre maximal de segments dont une extrémité est dans le disque de centre e_u et de rayon ϵ , formant un angle large avec u , et tels que les angles formés soient supérieurs ou égales à δ_p . Ce nombre n_{max} est borné par :

$$n_{max} \leq \left\lceil \frac{\pi}{\delta_p} \right\rceil + 1$$

où $\lceil \cdot \rceil$ désigne la partie entière et δ_p est en radian. La borne minimale de A est obtenue dans le cas d'une naissance d'un segment double u , connecté à n_{max} segments libres dans \mathbf{s} , en chacune de ces deux extrémités :

$$A \geq -2 n_{max}$$

La partie B peut être décomposée comme suit :

$$n_s(\mathbf{s} \cup u) - n_s(\mathbf{s}) = \begin{cases} 1 & \text{si } u \text{ est libre} \\ 1 + \#(\{s_i \text{ libre dans } \mathbf{s}/u \sim_{\mathcal{C}} s_i\}) \\ - \#(\{s_i \text{ simple dans } \mathbf{s}\} \cap \{s_i \text{ double dans } (\mathbf{s} \cup u)\}) & \text{si } u \text{ est simple} \\ \#(\{s_i \text{ libre dans } \mathbf{s}/u \sim_{\mathcal{C}} s_i\}) \\ - \#(\{s_i \text{ simple dans } \mathbf{s}\} \cap \{s_i \text{ double dans } (\mathbf{s} \cup u)\}) & \text{si } u \text{ est double} \end{cases}$$

La partie B sera minimale quand un nombre maximal de segments simples dans \mathbf{s} deviennent des segments doubles dans $(\mathbf{s} \cup u)$. Dans ce cas, B vérifie :

$$B \geq -2 n_{max}$$

Étant donné l'équation (3.8) et le fait que la fonction de qualité soit majorée par 1, le potentiel associé aux paires de segments $\langle s_i, s_j \rangle_{\mathcal{C}}$ vérifie : $g_c(s_i, s_j) \geq -1$. La somme C est donc bornée comme suit :

$$C \geq -2 n_{max}$$

Comme g_p est une fonction positive, l'expression D vérifie :

$$D \geq 0$$

¹Si l'on considère la relation de connexion \sim_c (Candy), le nombre de segments libres dont une extrémité se situe dans un cercle de rayon ϵ est inférieur à 6.

Finalemment, $\forall u$ et $\forall \mathbf{s} \in E / h_p(\mathbf{s}) \neq 0$,

$$\frac{h_p(\mathbf{s} \cup u)}{h_p(\mathbf{s})} \leq \beta \exp \left[2 (\omega_f + \omega_s + \omega_c) \left(\left[\frac{\pi}{\delta_p} \right] + 1 \right) \right] \quad (3.13)$$

ce qui prouve que le modèle “Quality Candy” est localement stable.

3.1.4 Modèle “IDQ”

Le modèle “IDQ” a été introduit par [Ortner, 2001] pour l’extraction des bâtiments à partir de modèles numériques d’élévation. Des interactions par paire, par triplet ou plus peuvent être définies via une formulation générale de la densité accordant la même importance aux interactions faisant intervenir des nombres différents d’objets. Chaque interaction, définie par rapport à une relation \sim_r et pour un nombre donné d’objets c_r , intervient dans le modèle au travers de trois termes globaux :

- l’**intensité** I qui correspond à la moyenne des nombres de groupes d’objets en interaction auxquelles un objet appartient. L’intensité I_r des interactions définies par rapport à une relation \sim_r est donnée par :

$$I_r(\mathbf{s}) = \frac{c_r N_g(r, \mathbf{s})}{N_o(r, \mathbf{s})} \quad (3.14)$$

où $N_g(r, \mathbf{s})$ est le nombre de groupes d’objets en interaction dans la configuration \mathbf{s} et $N_o(r, \mathbf{s})$ est le nombre d’objets en interaction dans la configuration \mathbf{s} . Si $N_o(r, \mathbf{s}) = 0$, on pose $I_r(\mathbf{s}) = 0$.

- la **diffusion** D qui correspond à la proportion d’objets en interaction dans la configuration \mathbf{s} . La diffusion D_r des interactions définies par rapport à une relation \sim_r est donnée par :

$$D_r(\mathbf{s}) = \frac{N_o(r, \mathbf{s})}{n(\mathbf{s})} \quad (3.15)$$

où $n(\mathbf{s})$ est le nombre total d’objets dans \mathbf{s} . Si $n(\mathbf{s}) = 0$, on pose $D_r(\mathbf{s}) = 0$.

- la **qualité** Q qui correspond à la qualité moyenne des groupes d’objets en interaction. La qualité Q_r des interactions définies par rapport à une relation \sim_r est donnée par :

$$Q_r(\mathbf{s}) = \frac{1}{N_g(r, \mathbf{s})} \sum_{\mathbf{g} \in \mathcal{G}_r(\mathbf{s})} \sigma_r(\mathbf{g}) \quad (3.16)$$

où $\mathcal{G}_r(\mathbf{s})$ est l’ensemble des groupes d’objets en interaction dans la configuration \mathbf{s} et σ_r est une fonction mesurant la qualité d’un groupe d’objet en interaction. C’est une fonction à valeurs dans $[0, 1]$, croissante par rapport à la qualité des interactions.

Pour contrôler l'intensité d'une interaction définie par rapport à \sim_r , [Ortner, 2001] propose de se donner la valeur objectif I_r^o pour l'intensité I_r et de définir une fonction de potentiel de la forme suivante :

$$V_{I_r}(\mathbf{s}) = \left(\tilde{I}_r(\mathbf{s}) - I_r^o \right) \log \left(\frac{\tilde{I}_r(\mathbf{s}) + 1}{I_r^o + 1} \right) \quad (3.17)$$

où $\tilde{I}_r(\mathbf{s})$ correspond à la valeur tronquée de $I_r(\mathbf{s})$ entre I_{min} et I_{max} :

$$\tilde{I}_r(\mathbf{s}) = \begin{cases} I_{min} & \text{si } I_r(\mathbf{s}) \leq I_{min} \\ I_r(\mathbf{s}) & \text{si } I_r(\mathbf{s}) \in [I_{min}, I_{max}] \\ I_{max} & \text{si } I_r(\mathbf{s}) \geq I_{max} \end{cases}$$

Ce tronquage permet de rendre la mesure d'intensité bornée et donc de montrer facilement la stabilité locale du modèle "IDQ". De même, pour contrôler la diffusion D_r , une fonction de potentiel est définie par rapport à la valeur objectif D_{opt} :

$$V_{D_r}(\mathbf{s}) = (D_r(\mathbf{s}) - D_r^o) \log \left(\frac{D_r(\mathbf{s}) + 1}{D_r^o + 1} \right) \quad (3.18)$$

Enfin, la fonction de qualité étant à valeur dans $[0, 1]$, la fonction de potentiel utilisée pour maximiser la qualité Q_r est la suivante :

$$V_{Q_r}(\mathbf{s}) = Q_r(\mathbf{s}) - 1 \quad (3.19)$$

L'énergie d'interaction est ensuite définie comme une somme pondérée de ces potentiels. Ainsi, la forme générale de la densité de type "IDQ" est donnée par :

$$h_p(\mathbf{s}) \propto \beta^n \exp \left[- \sum_{\sim_r \in \mathcal{R}} \gamma_I^r V_I^r(\mathbf{s}) + \gamma_D^r V_D^r(\mathbf{s}) + \gamma_Q^r V_Q^r(\mathbf{s}) \right] \quad (3.20)$$

où les $\gamma_I^r, \gamma_D^r, \gamma_V^r$ sont des poids positifs. Cette densité spécifie un processus ponctuel marqué bien défini et **localement stable** et cela pour n'importe quelle relation intervenant dans le modèle. En effet, puisque tous les potentiels V sont bornés, la densité de Papangelou l'est aussi. De plus, les fonctions de potentiels proposées par [Ortner, 2001] permettent de travailler avec des quantités qui sont du même ordre que le facteur d'intensité β , quelque soit la cardinalité c_r des interactions. Remarquons qu'une densité de type "IDQ" ne spécifie pas un processus de Markov, mais cela ne pose aucun problème du point de vu algorithmique puisque les quantités intervenant dans la densité h_p sont faciles à calculer.

Ici, nous utilisons ce formalisme dans le cas simple des interactions par paire. Les relations considérées sont les mêmes que celles utilisées dans le modèle "Quality Candy" : la relation de connexion $\sim_{c'}$, et la relation de proximité \sim_p . Afin d'introduire des contraintes de connectivité et de non superposition, nous avons choisi les valeurs objectifs suivantes :

- La densité optimale de connexion est prise égale à 1 :

$$D_c^o = 1$$

puisque nous voulons que tous les segments soient connectés.

- L'intensité optimale de connexion est prise égale à 2 :

$$I_c^o = 2$$

puisque nous voulons que les segments soient doublement connectés, donc qu'ils appartiennent à deux paires d'objets connectés.

- La densité optimale de proximité est prise nulle :

$$D_p^o = 0$$

puisque nous voulons limiter la présence de paire de segments vérifiant \sim_p . Dans ce cas, il est suffisant de définir la densité et nous fixons donc le poids de l'intensité à 0 :

$$\gamma_D^p = 0$$

La fonction de qualité utilisée pour la connexion est la même que celle utilisée pour le modèle "Quality Candy", *i.e.* :

$$\sigma_c(\langle s_i, s_j \rangle_{c'}) = -g_c(s_i, s_j)$$

où g_c est la fonction de potentiel associée à la connexion et donnée par l'équation (3.8). Le potentiel fondé sur la qualité Q_c de la configuration permet ainsi de contrôler la courbure moyenne du réseau et la moyenne des distances entre extrémités du réseau. Une fonction de qualité est également définie pour la proximité, malgré sa faible influence par rapport à la densité de proximité : l'énergie d'interaction sera plus faible pour une configuration sans interaction de proximité qu'une configuration présentant des interactions de proximité de qualité optimale. Elle est donnée par :

$$\sigma_p(\langle s_i, s_j \rangle_p) = 1 - \sigma(\tau_{ij}, \tau_p)$$

où τ_{ij} est la différence d'orientation (en valeur absolue et modulo $\pi/2$) des deux segments s_i et s_j , τ_p le seuil utilisé dans la définition de la proximité, et σ la fonction de qualité donnée par l'équation (3.9).

3.1.5 Noyau de proposition pour un échantillonnage efficace

Avant d'utiliser ces modèles en tant que modèle *a priori* pour l'extraction des réseaux dans les images, il convient de vérifier que les échantillons des processus, spécifiés par les densités *a priori* données dans les paragraphes précédents, présentent les mêmes caractéristiques géométriques et topologiques que les réseaux d'intérêt. Nous

proposons de réaliser l'échantillonnage des processus de segments par un algorithme de type Metropolis-Hastings, dont la structure générale est donnée dans la table 2.3. Comme nous l'avons expliqué dans le paragraphe 2.5.4, un point intéressant de cet algorithme est que le noyau de proposition Q peut être décomposé en plusieurs sous-noyaux, chacun correspondant à un mouvement réversible. En effet, bien qu'il soit suffisant de définir un noyau de type "naissance et mort" uniforme, il est important de définir des transformations pertinentes pour accélérer la convergence de la chaîne de Markov. Nous présentons ici différents sous-noyaux : le noyau de "naissance et mort" uniforme dont l'utilisation garantit l'irréductibilité de la chaîne, et des noyaux permettant un échantillonnage efficace, construits de façon à vérifier la réversibilité de la chaîne.

Naissance et mort uniforme

Le noyau de "naissance et mort" uniforme (NMU) est le noyau le plus simple permettant de réaliser des petits sauts entre les différents espaces $E_n = \{\mathbf{s} \in E : n(\mathbf{s}) = n\}$. Ce noyau de proposition consiste à proposer l'ajout d'un segment, choisi uniformément dans $F \times M$, ou le retrait d'un segment choisi uniformément dans la configuration \mathbf{s} . Dans le paragraphe 2.5.2, nous donnons l'exemple d'un tel noyau pour le cas d'un processus ponctuel non marqué. La méthodologie pour le calcul du taux de Green est la même pour les processus ponctuels marqués et l'on retrouve les mêmes taux si les marques associées à l'objet proposé dans le cas d'une naissance sont générées suivant la mesure de probabilité associée aux marques (remarque 7). C'est le cas ici puisque la mesure de probabilité P_M associée aux marques est la mesure uniforme sur $M = [L_{min}, L_{max}] \times [0, \pi]$.

Ainsi, les **taux de Green** associés à une NMU pour l'échantillonnage d'un processus de segments sur F spécifié par une densité h par rapport à un processus de poisson d'intensité λ sont les suivants :

$$R(\mathbf{s}, \mathbf{s} \cup u) = \frac{p_d}{p_b} \frac{\lambda |F|}{n(\mathbf{s}) + 1} \frac{h(\mathbf{s} \cup u)}{h(\mathbf{s})} \quad (3.21)$$

$$R(\mathbf{s}, \mathbf{s} \setminus u) = \frac{p_b}{p_d} \frac{n(\mathbf{s})}{\lambda |F|} \frac{h(\mathbf{s} \setminus u)}{h(\mathbf{s})} \quad (3.22)$$

où p_b et p_d correspondent respectivement aux probabilités de proposer une naissance et de proposer une mort. En pratique, nous prendrons $p_b = p_d = \frac{1}{2}$.

Perturbations simples de segments

Les petites perturbations d'objets sont des mouvements plus efficaces qu'une mort suivi d'une naissance et sont très utiles dès lors que des objets sont approximativement bien placés. Ces mouvements reposent sur des **transformations symétriques** afin de garantir la réversibilité de la chaîne. Soit $\mathcal{T} = \{T_a : a \in A\}$ une famille de transformations symétriques paramétrées par un vecteur $a \in A$. La symétrie de ces transformations est nécessaire pour respecter la réversibilité de la chaîne (probabilité d'aller \equiv probabilité de retour). Le noyau associé à cette famille consiste à choisir aléatoirement un objet

s dans la configuration courante \mathbf{s} et à proposer une perturbation en appliquant T_a à s , où a est choisi aléatoirement dans A . Si le choix de l'objet s et le choix du vecteur a sont faits **uniformément** dans \mathbf{s} et A , alors le taux de Green correspond au **taux de Metropolis** égal au rapport des vraisemblances :

$$R(\mathbf{s}, (\mathbf{s} \setminus s) \cup s') = \frac{h((\mathbf{s} \setminus s) \cup s')}{h(\mathbf{s})} \quad (3.23)$$

Nous définissons trois transformations symétriques de ce type : rotation, translation et dilatation (*i.e.* modification de la longueur d'un segment). Celles-ci sont illustrées par la figure 3.8.

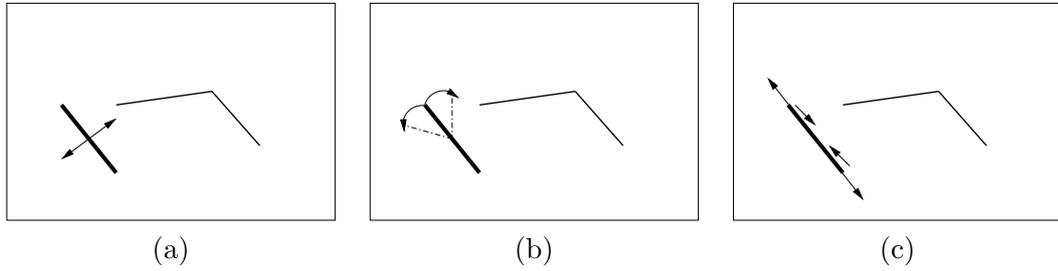


FIG. 3.8 – Perturbations symétriques : (a) translation - (b) rotation - (c) dilatation.

Considérons la famille des **rotations** définies dans $[-\Delta_\theta, \Delta_\theta]$. Une rotation T_{d_θ} consiste à modifier l'orientation θ du segment considéré en ajoutant $d_\theta \in [-\Delta_\theta, \Delta_\theta]$:

$$T_{d_\theta} \left(s = \begin{bmatrix} x \\ y \\ L \\ \theta \end{bmatrix} \right) = \begin{bmatrix} x \\ y \\ L \\ (\theta + d_\theta) [\pi] \end{bmatrix}$$

où $[\cdot]$ désigne la fonction modulo.

De la même façon, la famille des **translations** est paramétrée par un vecteur $[d_x, d_y]$, $d_x \in [-\Delta_x, \Delta_x]$, $d_y \in [-\Delta_y, \Delta_y]$. Une translation $T_{[d_x, d_y]}$ correspond à la translation du centre (x, y) du segment considéré s avec la condition que le nouveau point soit dans $F = [0, X_{max}] \times [0, Y_{max}]$:

$$T_{[d_x, d_y]} \left(\begin{bmatrix} x \\ y \\ L \\ \theta \end{bmatrix} \right) = \begin{bmatrix} (x + d_x) [X_{max}] \\ (y + d_y) [Y_{max}] \\ L \\ \theta \end{bmatrix}$$

Enfin, la famille des **dilatations** est paramétrée par un vecteur $d_L \in [-\Delta_L, \Delta_L]$. Une dilatation T_{d_L} correspond à la modification d'un segment de longueur L par l'ajout

de d_L avec la condition que la nouvelle longueur soit comprise entre L_{min} et L_{max} :

$$T_{d_L} \left(\begin{bmatrix} x \\ y \\ L \\ \theta \end{bmatrix} \right) = \begin{bmatrix} x \\ y \\ L_{min} + (L - L_{min} + d_L) [L_{max} - L_{min}] \\ \theta \end{bmatrix}$$

Perturbations de segments améliorées

Les propositions de perturbations de segment décrites précédemment ne sont pas pertinentes quand le segment est connecté. En effet, la connexion est souvent éliminée par la perturbation. C'est pourquoi nous définissons un nouveau sous-noyau qui dépend de l'état du segment par rapport à la connexion. Le noyau est alors constitué des étapes suivantes :

1. Choisir uniformément un segment s dans la configuration \mathbf{s} .
2. Si s est **libre**, choisir uniformément une perturbation simple de segment parmi les perturbations définies précédemment : rotation, translation ou dilatation.
3. Si s est **simple**, choisir uniformément un mouvement parmi les trois mouvements suivants :
 - (a) Translation de l'extrémité non connectée dans un carré centré au niveau de cette extrémité de telle sorte que la nouvelle longueur soit dans $[L_{min}, L_{max}]$.
 - (b) Translation de l'extrémité connectée dans un carré de côté inférieur à ϵ centré au niveau de cette extrémité de telle sorte que le nouveau segment reste connecté par cette extrémité.
 - (c) Translation d'une connexion relative à s . Premièrement, on choisit uniformément une paire de segments $\langle s, s' \rangle$ parmi les paires de segments connectés auxquelles s appartient ; deuxièmement, on choisit uniformément un vecteur $[d_x, d_y]$ dans $[-\Delta, \Delta] \times [-\Delta, \Delta]$; finalement, on translate les deux extrémités connectées par l'ajout du vecteur $[d_x, d_y]$ de telle sorte que les deux nouvelles longueurs soient dans $[L_{min}, L_{max}]$.
4. Si s est **double**, choisir uniformément un mouvement parmi les mouvements (b) ou (c) décrits précédemment pour un segment simple :
 - (b') Choisir uniformément une des extrémités de s et appliquer la translation (b) à cette extrémité.
 - (c') Procéder à la translation de connexion de la même façon que cela est fait en (c) pour les segments simples.

Remarque 10 *Si une de ces perturbations entraîne un changement d'état, la proposition de perturbation doit être refusée puisque, dans ce cas, le mouvement inverse n'est jamais proposé.*

Le taux de Green est encore réduit au rapport des vraisemblances (équation 3.23) puisque toutes les transformations sont symétriques et proposées uniformément.

Naissance et mort dans un voisinage

Étant donné que les segments du réseau sont supposés être connectés, nous avons introduit une “naissance et mort” dans un voisinage par rapport à la relation de connexion. Une naissance d'un segment connecté est proposée avec une probabilité p_{bv} et une mort d'un segment connecté avec la probabilité $p_{dv} = 1 - p_{bv}$. L'expression du noyau de proposition est donc de la forme suivante :

$$Q_{NMV}(\mathbf{s} \rightarrow A) = p_{bv} Q_{bv}(\mathbf{s} \rightarrow A) + p_{dv} Q_{dv}(\mathbf{s} \rightarrow A) \quad (3.24)$$

La **naissance dans un voisinage** consiste à choisir aléatoirement un segment s dans la configuration courante \mathbf{s} et une de ces extrémité $e_{s,i}$, puis à proposer un nouveau segment à partir de $e_{s,i}$. La génération du nouveau segment u se fait de la façon suivante :

1. une extrémité e_u du nouveau segment est générée par la sélection uniforme d'un point dans la boule de centre $e_{s,i}$ et de rayon ϵ ;
2. la longueur L_u du nouveau segment est générée de façon uniforme dans $[L_{min}, L_{max}]$;
3. la direction d_u du nouveau segment est générée de façon uniforme dans :
 - $[0, 2\pi]$ si la relation de connexion est \sim_c , définie indépendamment de l'angle formé par les deux segments connectés,
 - $[\theta - \frac{\pi}{2}, \theta + \frac{\pi}{2}]$ où θ est l'orientation de s si la relation de connexion est $\sim_{c'}$, qui n'est définie que pour les segments formant un angle grave.

La partie du noyau correspondant à une naissance peut s'écrire comme suit :

$$Q_{bv}(\mathbf{s} \rightarrow A) = \sum_{s \in \mathbf{s}} p^s(\text{choisir } s) \sum_{i=1}^2 p^s(\text{choisir } i) P_Z(\mathbf{s} \cup \eta_{s,i}(Z)) \quad (3.25)$$

où $Z \sim \mathcal{U}[\mathcal{B}(0, \epsilon) \times [L_{min}, L_{max}] \times [0, 2\pi]]$ dans le cas où la relation de connexion est \sim_c et $Z \sim \mathcal{U}[\mathcal{B}(0, \epsilon) \times [L_{min}, L_{max}] \times [0, \pi]]$ dans le cas où la relation de connexion est $\sim_{c'}$, et $\eta_{s,i}$ est le difféomorphisme permettant de passer du vecteur ainsi généré z à $u = (p_u, L_u, \theta_u)$ en connaissant le segment s et l'extrémité $e_{s,i}$. Ceci est équivalent à définir le difféomorphisme η de (e_u, L_u, d_u) à (p_u, L_u, θ_u) :

$$\eta(e_u, L_u, d_u) = \left(e_u + \frac{L_u}{2} \begin{bmatrix} \cos(d_u) \\ \sin(d_u) \end{bmatrix}, L_u, d_u[\pi] \right)$$

Le choix le plus simple pour la probabilité p^s de choix d'un segment dans \mathbf{s} est de prendre une distribution uniforme :

$$p^s(\text{choisir } s) = \frac{1}{n(\mathbf{s})}$$

Nous considérons qu'il n'est pas judicieux de proposer un segment à partir d'une extrémité située en dehors de la fenêtre d'observation F , nous construisons la probabilité de choisir l'extrémité $e_{s,i}$ comme suit :

$$p^s(\text{choisir } i) = \begin{cases} \frac{1}{2} & \text{si les deux extrémités de } s \text{ sont dans } F \\ 0 & \text{si } e_{s,i} \text{ n'est pas dans } F \\ 1 & \text{si } e_{s,i} \text{ est l'unique extrémité de } s \text{ présente dans } F \end{cases}$$

L'équation 3.25 devient :

$$Q_{bv}(\mathbf{s} \rightarrow A) = \sum_{s \in \mathbf{s}} \frac{1}{n(\mathbf{s})} \sum_{i=1}^2 p^s(\text{choisir } i) \int_{\Sigma} \mathbf{1}_A(\mathbf{s} \cup \eta_{s,i}(z)) \frac{dz}{|\Sigma|} \quad (3.26)$$

$$\int_{F \times M} \mathbf{1}_{V(s,i)}(u) \mathbf{1}_A(\mathbf{s} \cup u) \underbrace{|J_{\eta_{s,i}^{-1}}(u)|}_{1} \frac{du}{|\Sigma|}$$

où Σ est l'espace dans lequel est généré Z , $\mathbf{1}_{V(s,i)}$ est la fonction indicatrice d'appartenance au voisinage de s par rapport à une connexion avec l'extrémité $e_{s,i}$, $|J_{\eta_{s,i}^{-1}}(u)|$ est le déterminant du jacobien de la réciproque de $\eta_{s,i}$. Si $n(\mathbf{s}) > 0$, l'équation 3.26 devient :

$$Q_{bv}(\mathbf{s} \rightarrow A) = \frac{1}{n(\mathbf{s})} \int_{F \times M} \sum_{s \in V_{\mathbf{s} \cup u}(u)} p^s(\text{choisir } i : u \in V(s, e_i)) \mathbf{1}_A(\mathbf{s} \cup u) \frac{du}{|\Sigma|} \quad (3.27)$$

où $V_{\mathbf{s} \cup u}(u)$ désigne l'ensemble des voisins de u dans la configuration $(\mathbf{s} \cup u)$. Remarquons que la somme sur i est supprimée puisque un segment s ne peut être connecté à u que par une seule de ces extrémités, l'extrémité i telle que $u \in V(s, e_i)$.

La **mort dans un voisinage** consiste à choisir aléatoirement une paire de segments $\langle u, v \rangle$ connectés uniformément dans la configuration et retirer un des deux segments selon $p^{\langle u, v \rangle}$ défini comme suit :

$$p^{\langle u, v \rangle}(\text{choisir } u) = \begin{cases} \frac{1}{2} & \text{si les deux extrémités concernées par la connexion entre } \\ & u \text{ et } v \text{ sont dans } F \\ 0 & \text{si l'extrémité de } v \text{ concernée par la connexion n'est pas} \\ & \text{dans } F \\ 1 & \text{si l'extrémité de } v \text{ concernée par la connexion est dans} \\ & F \text{ et celle de } u \text{ ne l'est pas.} \end{cases}$$

Ce choix s'explique par le fait que nous ne proposons que des naissances dans un voisinage des extrémités qui sont situées dans F . La partie du noyau correspondant à une mort peut s'écrire comme suit :

$$Q_{dv}(\mathbf{s} \rightarrow A) = \sum_{\langle s_i, s_j \rangle_{c^*}} \frac{1}{N_{pc}(\mathbf{s})} \sum_{k \in \{i, j\}} p^{\langle s_i, s_j \rangle} (\text{choisir } s_k) \mathbf{1}_A(\mathbf{s} \setminus s_k) \quad (3.28)$$

$$= \frac{1}{N_{pc}(\mathbf{s})} \sum_{u \in \mathbf{s}} \sum_{v \in V_{\mathbf{s}}(u)} p^{\langle v, u \rangle} (\text{choisir } u) \mathbf{1}_A(\mathbf{s} \setminus u) \quad (3.29)$$

où c^* désigne la relation de connexion choisie ($c^* = c$ ou c' selon le modèle choisi) et $N_{pc}(\mathbf{s})$ désigne le nombre de paire de segments en interaction de connexion \mathbf{s} .

Le calcul du **taux de Green**, permettant d'assurer la réversibilité de la chaîne, se fait de la même façon que pour le noyau "naissance et mort" uniforme (paragraphe 2.5.2). La mesure symétrique ψ choisie pour dériver la mesure πQ_{NMV} est la même que celle utilisée pour le noyau naissance mort. C'est une mesure symétrique concentrée sur $\bigcup_{n=0}^{\infty} \{\{E_n \times E_{n+1}\} \cup \{E_{n+1} \times E_n\}\}$ et égale à ψ_n sur $\{E_n \times E_{n+1}\} \cup \{E_{n+1} \times E_n\}$ où les mesures ψ_n sont définies de la façon suivante :

$$\psi_n(A_n, B_{n+1}) = \int_{A_n} \int_{F \times M} \mathbf{1}_{B_{n+1}}(\mathbf{s} \cup u) d\nu(u) d\mu(\mathbf{s}) \quad (3.30)$$

$$\psi_n(B_{n+1}, A_n) = \int_{B_{n+1}} \sum_{u \in \mathbf{s}} \mathbf{1}_{A_n}(\mathbf{s} \setminus u) d\mu(\mathbf{s}) \quad (3.31)$$

où ν est la mesure d'intensité du processus ponctuel marqué de mesure μ .

La dérivée de Radon-Nikodym de πQ_{NMV} dans le cas d'une naissance d'un segment u dans un voisinage d'un segment de \mathbf{s} se déduit des équations (3.27) et (3.30) :

$$D(\mathbf{s}, \mathbf{s} \cup u) = \frac{p_{bv} f(\mathbf{s}) \sum_{s \in V_{\mathbf{s} \cup u}(u)} p^s (\text{choisir } i : u \in V(s, i))}{n(\mathbf{s}) |\Sigma| \lambda |F| dP_M(L_u, \theta_u)} \quad (3.32)$$

où λ est l'intensité du processus de référence et P_M la mesure de probabilité associée aux marques (loi uniforme sur M). De même, la dérivée de Radon-Nikodym dans le cas d'une mort d'un segment connecté $u \in \mathbf{s}$ se déduit des équations (3.29) et (3.31) :

$$D(\mathbf{s}, \mathbf{s} \setminus u) = \frac{p_{dv} f(\mathbf{s}) \sum_{v \in V_{\mathbf{s}}(u)} p^{\langle u, s \rangle} (\text{choisir } u)}{N_{pc}(\mathbf{s})} \quad (3.33)$$

Finalement, les taux de Green $R(\mathbf{s}, \mathbf{s} \cup u)$ et $R(\mathbf{s}, \mathbf{s} \setminus u)$ respectivement associés à une naissance et une mort sont :

$$R(\mathbf{s}, \mathbf{s} \cup u) = \frac{f(\mathbf{s} \cup u)}{f(\mathbf{s})} \frac{p_{dv}}{p_{bv}} \frac{\gamma \pi \epsilon^2}{\lambda |F|} \frac{n(\mathbf{s})}{N_{pc}(\mathbf{s} \cup u)} \frac{\sum_{v \in V_{\mathbf{s} \cup u}(u)} p^{\langle u, v \rangle} (\text{choisir } u)}{\sum_{v \in V_{\mathbf{s} \cup u}(u)} p^v (\text{choisir } i : u \in V(v, i))} \quad (3.34)$$

$$R(\mathbf{s}, \mathbf{s} \setminus u) = \frac{f(\mathbf{s} \setminus u)}{f(\mathbf{s})} \frac{p_{bv}}{p_{dv}} \frac{\lambda|F|}{\gamma \pi \epsilon^2} \frac{N_{pc}(\mathbf{s})}{n(\mathbf{s} \setminus u)} \frac{\sum_{v \in V_{\mathbf{s}}(u)} p^v(\text{choisir } i : u \in V(v, i))}{\sum_{v \in V_{\mathbf{s}}(u)} p^{<u,v>}(\text{choisir } u)} \quad (3.35)$$

où γ est le rapport entre la densité de probabilité $dP_M(L_u, \theta_u)$ et la densité de probabilité $dP_G(L_u, d_u)$ correspondant à la génération de la longueur et de la direction du nouveau segment. Dans le cas d'une connexion définie pour un angle large le rapport γ est égal à 1 et dans le cas d'une définition pour tout angle γ est égal à 2.

Remarque 11 *Dans le cas où les extrémités de u vérifiant une relation de connexion sont dans la fenêtre F et où toutes les extrémités (de segments de \mathbf{s}) concernées par une relation de connexion avec u sont également dans la fenêtre F , on a :*

$$\frac{\sum_{v \in V_{\mathbf{s}}(u)} p^v(\text{choisir } i : u \in V(v, i))}{\sum_{v \in V_{\mathbf{s}}(u)} p^{<u,v>}(\text{choisir } u)} = \frac{\frac{1}{2} \#(V_{\mathbf{s}}(u))}{\frac{1}{2} \#(V_{\mathbf{s}}(u))} = 1$$

En pratique, on approximera ce rapport, qui est en général très proche de 1, par la valeur 1.

Convergence ergodique

L'irréductibilité de la chaîne est assurée si le noyau de "naissance mort" uniforme intervient dans le noyau de proposition global de la chaîne et si la condition de stabilité locale est vérifiée, ce qui est le cas pour tous les modèles présentés ici. L'ergodicité de la chaîne (convergence pour toute condition initiale $\mathbf{x} \in E : h(\mathbf{x}) > 0$) reste vraie par l'ajout de sous-noyaux réversibles concernant des perturbations gardant le nombre de segments constant. Dans le cas de l'ajout d'un sous-noyau réversible de type "naissance et mort", l'ergodicité reste vrai si le taux associé à une naissance tend vers zéro quand $n(\mathbf{s})$ tend vers l'infini.

Pour le modèle "Quality Candy", la convergence ergodique est assurée par le fait que ce modèle présente un potentiel de type "hard core" pour la relation de proximité \sim_p . Or, on peut montrer qu'il existe un nombre maximal n_M de segments positionnés dans la fenêtre F tel que aucune interaction de type "hard core" ne soit vérifiée. Toute configuration \mathbf{s} telle que $n(\mathbf{s}) > n_M$ présentera au moins une interaction de type "hard core". Or, l'algorithme ne considère pas les configurations présentant des interactions de type "hard core" puisque celles-ci sont de densité de probabilité nulle. La chaîne de Markov ainsi construite a donc pour espace d'état $\{\mathbf{s} \in E : h(\mathbf{s}) > 0\} \subset \{\mathbf{s} \in E : n(\mathbf{s}) \leq n_M\}$. L'apériodicité et l'irréductibilité suffisent donc à montrer l'ergodicité de la chaîne. Pour le modèle "Quality Candy", l'ergodicité de la chaîne reste donc vraie par l'ajout de tout sous-noyau de type naissance-mort réversible, tel que le noyau NMV.

Les modèles “Candy” et “IDQ” ne présentent pas de potentiel de type “hard core”. Dans le cas où toutes les relations mises en jeu par un modèle sont réflexives, on peut montrer que le rapport n/N_g du nombre d’objets sur le nombre de groupes d’objets en interaction tend vers zéro quand n tend vers l’infini. Sous la condition de stabilité locale, on peut alors montrer que le taux associé à une naissance tend vers 0 quand n tend vers l’infini. Mais, ce n’est pas le cas des relations de connexion définies ici. On propose de définir le noyau Q de la chaîne de Markov de la façon suivante :

$$Q(\mathbf{s}, A) = p_{\text{NMU}}^{\mathbf{s}} Q_{\text{NMU}} + p_{\text{NMV}}^{\mathbf{s}} Q_{\text{NMU}} + \sum_i p_i^{\mathbf{s}} q_i$$

$$\text{avec } \begin{cases} p_{\text{NMU}}^{\mathbf{s}} + p_{\text{NMV}}^{\mathbf{s}} + \sum_i p_i^{\mathbf{s}} = 1 & \forall \mathbf{s} \\ p_{\text{NMU}}^{\mathbf{s}} > 0 & \forall \mathbf{s} \\ p_{\text{NMV}}^{\mathbf{s}} = 0 & \forall \mathbf{s} : n(\mathbf{s}) \leq n_M \end{cases} \quad (3.36)$$

où $p_{\text{NMU}}^{\mathbf{s}}$, $p_{\text{NMV}}^{\mathbf{s}}$ et $p_i^{\mathbf{s}}$ sont respectivement les probabilités de choisir un mouvement de type “naissance et mort” uniforme, un mouvement de type “naissance et mort” dans un voisinage et une perturbation i de segments (rotation, translation, etc.) connaissant \mathbf{s} . En effet, un tel choix pour le noyau de proposition nous assure la convergence ergodique sous la condition de stabilité locale de la densité du processus.

Remarque 12 *En pratique, nous n’observons jamais le problème de partir à l’infini puisque informatiquement le nombre de points d’une configuration est limité à un nombre de points donné. Ceci revient à multiplier la densité par $\mathbf{1}\{n(\mathbf{x}) \leq n_0\}$ où n_0 est le nombre maximal de points autorisé. Dans ce cas là, l’irréductibilité et l’apériodicité suffisent à montrer la convergence du noyau suivant :*

$$Q(\mathbf{s}, A) = p_{\text{NMU}} Q_{\text{NMU}} + p_{\text{NMV}} Q_{\text{NMV}} + \sum_i p_i q_i \quad (3.37)$$

avec $p_{\text{NMU}} + p_{\text{NMV}} + \sum_i p_i = 1$ et $p_{\text{NMU}} > 0$.

3.1.6 Échantillons des modèles a priori

Dans ce paragraphe, nous présentons des échantillons obtenus par un algorithme de type Metropolis-Hastings après détection automatique de la convergence en utilisant le critère d’arrêt donné dans la table 2.4. Le noyau de proposition Q choisi est composé de deux noyaux de proposition équiprobables : une “naissance et mort” uniforme (NMU) et une “naissance et mort” dans un voisinage (NMV) par rapport à la relation de connexion :

$$Q(\mathbf{s}, A) = \frac{1}{2} Q_{\text{NMU}} + \frac{1}{2} Q_{\text{NMV}} \quad (3.38)$$

Nous avons également réalisé des tests avec un noyau unique de type NMU mais dans ce cas la stationnarité est atteinte avec largement plus d’itérations, ou n’est pas détectée avant l’arrêt de l’algorithme par le nombre d’itérations (nous avons fixé un nombre maximal d’itérations égal à plusieurs millions d’itérations). Ceci s’explique par

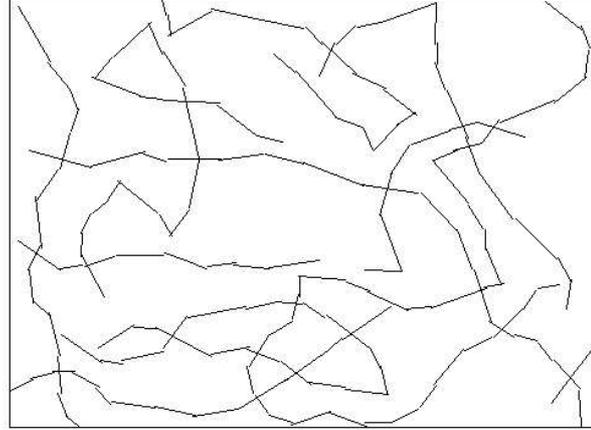


FIG. 3.9 – Échantillon du modèle “Candy” spécifié par la densité h_p donnée par l’équation (3.4) par rapport au processus de Poisson d’intensité 40000 en posant $|F| = 1$ et avec la paramétrisation suivante : $\beta = 1$, $w_f = \frac{15}{3}$, $w_s = \frac{35}{3}$, $w_p = \frac{50}{3}$, $w_a = \frac{50}{3}$.

le fait que les interactions mises en jeu par les trois modèles présentés dans le paragraphe 3.1 sont très fortes. En effet, ce sont des modèles construits de façon à exploiter de fortes contraintes géométriques et topologiques.

La convergence de la chaîne est évaluée au travers de la convergence de différentes statistiques pour chaque modèle. Pour le modèle “Candy”, la convergence est évaluée à partir de la détection de la convergence des moyennes empiriques des statistiques suffisantes $n(\mathbf{s})$, $n_s(\mathbf{s})$, $n_f(\mathbf{s})$, $n_p(\mathbf{s})$ et $n_a(\mathbf{s})$ selon le critère d’arrêt donné dans la table 2.4. Pour le modèle “Quality Candy” et le modèle “IDQ”, le nombre d’interactions de mauvais alignement $n_a(\mathbf{s})$ n’intervient pas mais est remplacé par la qualité moyenne de la configuration. Pour le modèle “IDQ”, $n_s(\mathbf{s})$, $n_f(\mathbf{s})$ sont remplacés par $N_{pc}(\mathbf{s})$ le nombre de paires de segments en interaction de connexion. Pour calculer ces moyennes empiriques un sous-échantillonnage est utilisé, comme proposé dans le paragraphe 2.5.6, en sélectionnant un échantillon toutes les 1000 itérations (après un nombre minimal d’itérations fixé à 10^4 itérations).

La figure 3.9 présente le dernier échantillon obtenu par l’algorithme MH pour le modèle “Candy”. L’échantillon obtenu est satisfaisant dans le sens où tous les segments sont connectés et forment de longues lignes brisées. Cependant, le réseau obtenu n’est pas aussi lisse qu’un réseau routier peut l’être en réalité : il présente des connexions entre segments de mauvaise qualité et des points de forte courbure. La convergence a été détectée après 1.6×10^6 itérations, ce qui correspond à un temps de calcul de 1 minute et 46 secondes avec un processeur Pentium III, 1 GHz, et 1 Go de mémoire vive.

La figure 3.10 présente des échantillons obtenus avec des paramétrisations différentes de l'énergie de type "Quality Candy". L'échantillon (a) correspond à une paramétrisation nulle de l'énergie d'interaction et correspond donc à une réalisation d'un processus uniforme de Poisson. On vérifie que la moyenne empirique du nombre de points $n(\mathbf{s})$ est égale à $\beta \lambda |F|$, où λ est l'intensité du processus de référence et $|F|$ est l'aire de la fenêtre d'observation (ici, on pose $|F| = 1$). Cela permet de vérifier le bon fonctionnement de l'algorithme proposé. L'échantillon (b) concerne une réalisation d'un processus faisant intervenir le nombre de segments libres et de segments simples sans la prise en compte de la relation de proximité. Dans ce cas, la preuve de la stabilité du processus (paragraphe 3.1.3) n'est plus valable. En pratique, cela se traduit par une prolongation perpétuelle du réseau : l'acceptation d'un nouveau segment connecté est généralement acceptée même s'il en recouvre d'autres. En fait, il n'y a pas convergence et l'algorithme qui a fourni l'échantillon (b) ne s'est pas arrêté par détection de la convergence des moyennes empiriques mais par le nombre d'itérations. On a ensuite incorporé au modèle les interactions de proximité sans la prise en compte de la qualité des interactions de connexion. L'échantillon (c) obtenu présente des points de forte courbure. Enfin, l'échantillon (d) correspond au modèle complet et semble plus adapté à la modélisation du réseau routier que peut l'être le modèle "Candy", la faible courbure et la continuité étant mieux respectées. L'algorithme MH pour le modèle complet s'est arrêté par détection de la convergence après 1.3×10^6 itérations, soit 2 minutes en temps de calcul. La convergence est plus rapide en termes d'itérations mais s'obtient en plus de temps que pour le modèle "Candy". En effet, le coût du calcul d'une itération est plus important pour le modèle "Quality Candy" (8.86 secondes pour 10^5 itérations) que pour le modèle "Candy" (6.37 secondes pour 10^5 itérations). Cela peut s'expliquer par le fait que la moyenne empirique du nombre de segments est de l'ordre de 140 pour le modèle "Candy" (avec la paramétrisation choisie ici) alors qu'elle est de l'ordre de 150 pour le modèle "Quality Candy". De plus, le calcul des mesures de qualité n'est pas effectué pour le modèle "Candy". Par contre, une interaction de plus est à considérer dans le modèle "Candy" : l'interaction de mauvais alignement.

La figure 3.11 concerne le modèle "IDQ". L'échantillon (a) montre un défaut de ce modèle : si le poids associé à la qualité de connexion est réduit à zéro, un segment aura tendance à être connecté à deux autres segments par une seule de ces extrémités. En effet, ce n'est plus le nombre de segments libres et de segments simples qui intervient dans la densité mais l'intensité de connexion, *i.e.* la proportion de segments par rapport au nombre de paires de segments en interaction de connexion. Nous avons construit l'intensité de connexion pour avoir autant de segments que de paires de segments connectés. Trois segments ayant chacun une extrémité dans un cercle de rayon ϵ sont donc optimaux dans ce sens (dans le cas où ils ne forment pas d'angle aigu). L'échantillon (b), obtenu avec le modèle complet, est toutefois de très bonne qualité par rapport à l'objectif que l'on s'est fixé, à savoir fournir un modèle permettant la génération d'un réseau formé de branches connectées entre elles et de faible courbure.

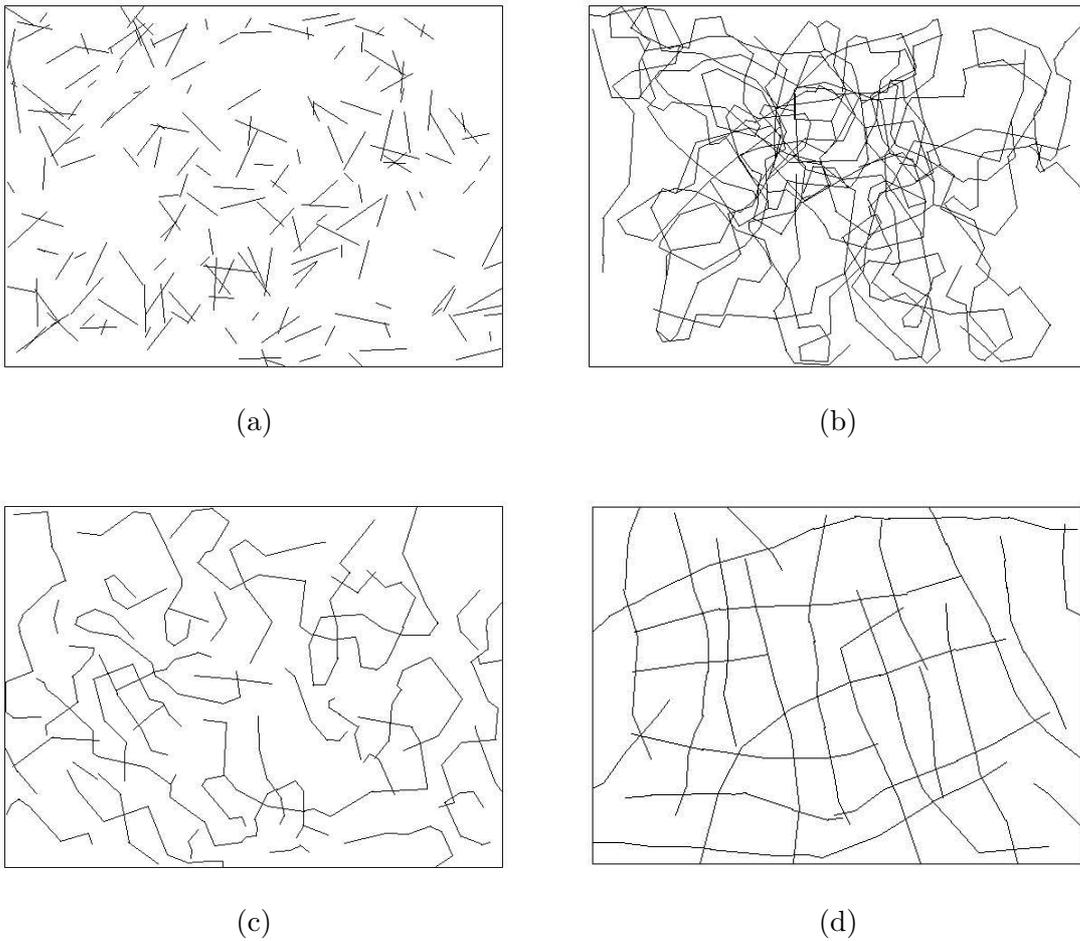


FIG. 3.10 – Échantillons du modèle “Quality Candy” - (a) Processus sans interaction : $\omega_i = 0 \forall i$ - (b) Processus ne dépendant que des états des segments : $\omega_f \neq 0, \omega_s \neq 0, \omega_r = \omega_c = 0$ - (c) Processus dépendant des états des segments et du mauvais alignement : $\omega_c = 0$ - (d) Modèle complet.

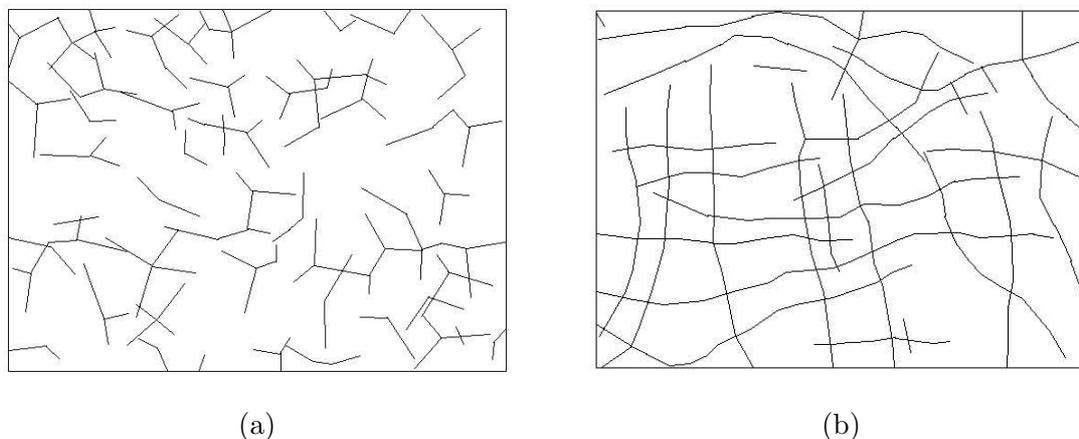


FIG. 3.11 – Échantillons du modèle "IDQ" - (a) Modèle avec un poids nul associé à la qualité de connexion : $\gamma_Q^c = 0$ - (b) Modèle complet.

Le coût d'une itération pour ce modèle est de 6.19 secondes pour 10^5 itérations alors que la moyenne empirique du nombre de segments est de l'ordre de 200 à la fin de l'algorithme d'échantillonnage. Comme pour le "Quality Candy" des mesures de qualité sont à calculer, mais ceci est compensé par le fait que le nombre de segments simples et de segment libres n'interviennent pas dans le modèle. La mise à jour de $n_f(\mathbf{s})$ et $n_s(\mathbf{s})$ lors de l'ajout d'un nouveau segment est en effet coûteuse en temps de calcul puisqu'il faut mettre à jour l'état des voisins de ce nouveau segment par rapport à la relation de connexion, et cela en fonction de toutes les interactions de connexion auxquelles les voisins participent. L'algorithme MH pour le modèle complet s'est arrêté par détection de la convergence après 1.5×10^6 itérations, soit 1 minute et 36 secondes en temps de calcul. Ce faible temps de calcul est à contrebalancer avec le fait que la configuration initiale choisie dans ce cas n'est pas la configuration vide comme pour les deux autres modèles mais une configuration de n points distribués uniformément dans la fenêtre. En effet, pour obtenir une configuration pertinente les poids associés aux potentiels doivent être pris nettement plus élevés que ceux des deux autres modèles puisque les quantités intervenant dans la densité I , D et Q sont des moyennes sur la totalité de la configuration : lorsque le nombre d'objet est important la modification des quantités I , D et Q lors d'une perturbation est très faible. Or, dans le cas où les objets sont peut nombreux la variation est nettement plus importante. C'est pourquoi le choix de poids élevés ne permet pas d'utiliser la configuration vide comme initialisation : l'état vide étant un optimum local de la densité, la chaîne Markov aura tendance à rester dans cet état. Une autre solution pour obtenir un échantillon de forte densité aurait été d'utiliser un recuit simulé : l'initialisation par une configuration vide n'aurait alors pas posé de problème puisque au début du recuit on simule suivant une distribution proche de la distribution uniforme.

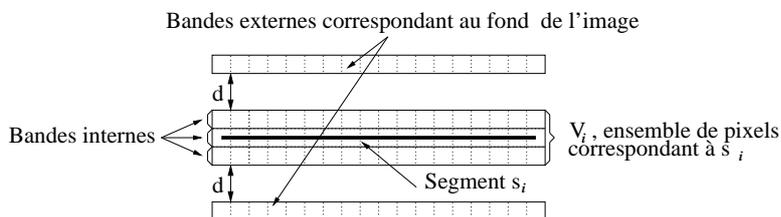


FIG. 3.12 – Masque de pixels associé à un segment

3.2 Incorporation des propriétés radiométriques des données

3.2.1 Approche par champ externe

La complexité des images que l'on souhaite traiter rend la modélisation du fond de l'image complexe, ce qui rend délicate la définition d'une vraisemblance. C'est pourquoi nous proposons dans un premier temps une approche par champ externe. L'image satellitaire ou aérienne est alors utilisée en tant que champ externe permettant l'évaluation de la qualité de chaque segment de la configuration \mathbf{s} compte tenu des deux hypothèses usuellement faites en détection de réseau :

(H_1) le niveau de gris du réseau est localement homogène ;

(H_2) le réseau contraste fortement avec son environnement.

Un potentiel fondé sur une mesure de qualité est alors associé à chaque segment, ce qui permet de définir le terme d'attache aux données de la façon suivante :

$$h_d(\mathbf{s}) = \exp \left(-\omega_d \sum_{s_i \in \mathbf{s}} v_d(i) \right) \quad (3.39)$$

où $v_d(i)$ est le potentiel associé au segment s_i et ω_d est un poids positif traduisant le degré de confiance accordé aux données.

Construction du terme d'attache aux données

Pour le calcul du terme d'attache aux données, nous associons à chaque segment de la configuration un masque de pixels comme le montre la figure 3.12. Ce masque est composé de :

- un ensemble de pixels V_i correspondant à s_i dans l'image. Cet ensemble est composé d'un nombre fixé n_b de bandes internes (le nombre n_b dépend directement de la largeur du réseau dans l'image) ;
- deux ensembles de pixels R_i^1 et R_i^2 correspondant au fond proche de s_i dans l'image. Ces deux ensembles consistent en deux régions adjacentes et colinéaires à V_i , prises à une distance fixée d de V_i pour permettre une légère variation de la largeur du réseau.

La vérification des hypothèses H_1 et H_2 se fait alors par la vérification des deux hypothèses suivantes pour chaque segment s_i de la configuration \mathbf{s} :

- (H'_1) la variation de niveau de gris entre deux bandes internes de V_i successives est faible ;
- (H'_2) la variation de niveau de gris entre V_i et les régions externes, R_i^1 et R_i^2 , est importante.

Pour évaluer la différence de moyennes de niveau de gris entre deux ensembles de pixels (disjoints), nous utilisons la mesure t sur laquelle se fonde le **t-test de Student**. Ce test statistique est utilisé pour déterminer si deux échantillons x et y sont issus de deux populations de moyennes différentes, *i.e.* si les moyennes de x et y sont significativement différentes. Si les distributions des populations sont normales, ce test sera particulièrement adapté à des problèmes d'inférence fondés sur de petits échantillons. Dans le cas d'une population quelconque, alors les échantillons doivent être représentatifs ($n > 30$). La mesure t est donnée par la formule ² suivante :

$$t(x, y) = \frac{|\bar{x} - \bar{y}|}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}} \quad (3.40)$$

où \bar{x} , σ_x^2 et n_x (resp. \bar{y} , σ_y^2 et n_y) correspondent respectivement à la moyenne, la variance, et le nombre d'éléments de l'échantillon x (resp. y).

La valeur statistique utilisée pour évaluer l'hypothèse d'homogénéité H'_1 pour un segment s_i de la configuration \mathbf{s} est alors le maximum des valeurs de la mesure t entre deux bandes internes successives :

$$T_1(s_i) = \max_{j \in \{1, \dots, n_b - 1\}} [t(b_j^i, b_{j+1}^i)] \quad (3.41)$$

où b_j^i est la bande interne j de l'ensemble de pixel V_i associé à s_i . Notons que, si le réseau présent dans l'image ne fait qu'un pixel de largeur, ce test ne pourra pas être effectué et la valeur 1 sera assignée à $T_1(s_i)$. Le choix du maximum revient à ne considérer que les deux bandes qui sont le moins en adéquation avec l'hypothèse de moyennes égales. De même, la valeur choisie pour évaluer l'hypothèse de contraste H'_2 est le minimum des valeurs de la mesure t entre V_i et chacune des bandes externes R_1^i et R_2^i :

$$T_2(s_i) = \min_{l \in \{1, 2\}} [t(R_l^i, V_i)] \quad (3.42)$$

Ces deux hypothèses sont évaluées conjointement à partir du rapport entre les valeurs T_2 et T_1 , avec la condition que T_1 soit supérieure à 1 pour ne pas favoriser excessivement les régions très homogènes et éviter ainsi les fausses alarmes. On a donc :

$$T_i = \frac{T_{H_2}(s_i)}{\max[1, T_{H_1}(s_i)]} \quad (3.43)$$

²Plusieurs formules peuvent être utilisées pour le t-test. L'équation (3.40) est la formule utilisée lorsque les variances sont distinctes et les échantillons sont non corrélés.

les segments dont la valeur T_i est inférieure à τ_1 sont considérés comme des segments aberrants par rapport aux données. Un potentiel positif V_{max} leur est assigné. Le potentiel V_{max} a été introduit après une étude sur les contraintes à respecter entre l'*a priori* et l'attache aux données afin d'obtenir une configuration optimale vérifiant de bonnes propriétés (voir paragraphe 3.3). Pour les autres segments, on procède à un seuillage empirique de T_i par une valeur maximale τ_2 , suivi d'une transformation linéaire de $[\tau_1, \tau_2]$ vers $[-1, 1]$ pour obtenir le potentiel d'attache aux données. Le potentiel correspondant au segment s_i est alors défini comme suit :

$$v_d(i) = \begin{cases} V_{max} & \text{si } T_i < \tau_1 \\ 1 - 2 \frac{T_i - \tau_1}{\tau_2 - \tau_1} & \text{si } \tau_1 \leq T_i \leq \tau_2 \\ -1 & \text{si } T_i > \tau_2 \end{cases} \quad (3.44)$$

L'énergie d'attache aux données est alors la somme des potentiels $v_d(i)$ associés à chacun des segments s_i de la configuration \mathbf{s} multipliée par un poids positif ω_d . Le terme d'attache aux données est donné par l'équation (3.39).

Nous avons donc construit un terme d'attache aux données reposant sur des hypothèses à la fois réalistes et assez générales pour représenter différents types de réseaux. Celles-ci sont évaluées à partir de mesures statistiques, largement utilisées en inférence statistique pour tester si les moyennes de deux ensembles d'échantillons sont significativement différentes. Cependant, l'utilisation de ces potentiels s'avère coûteuse en temps de calcul, et d'autant plus coûteuse que le potentiel v_d est calculé à chaque proposition d'un nouvel élément et que l'algorithme d'optimisation proposé dans le paragraphe 2.6 nécessite souvent quelques millions de propositions de nouveaux éléments avant d'atteindre la configuration optimale. Nous proposons donc de réaliser un pré-calcul des potentiels permettant une approximation du terme d'attache aux données.

Approximation du terme d'attache aux données par pré-calcul

Nous proposons une approximation du terme d'attache aux données fondée sur un pré-calcul des potentiels pour des segments de longueur minimale L_{min} , positionnés en chaque pixel de l'image, pour un nombre fixé d'orientations N_θ .

Pour chaque orientation de cet espace discrétisé $\tilde{\theta}_k, k=1, \dots, N_\theta$, nous assignons à chaque pixel pix de l'image la valeur $w_k(pix)$ égale au potentiel du segment $s_j = (p_j, L_{min}, \theta_k)$, où p_j est tel que l'ensemble de pixel V_j associé à s_j contienne pix , qui répond le mieux aux tests statistiques, *i.e.* qui minimise le potentiel donné par l'équation (3.44) :

$$w_k(pix) = \min_{\substack{s_j : V_j \ni pix \\ L_j = L_{min} \\ \theta_j = \tilde{\theta}_k}} v_d(j) \quad (3.45)$$

Nous obtenons alors N_θ cartes, définies pour chacune des orientations de l'espace discrétisé et permettant d'évaluer en chaque pixel la possibilité de la présence d'une

structure linéaire d'une orientation donnée. Pour un segment s_i d'orientation θ_i de la configuration \mathbf{s} , le potentiel d'attache aux données sera donc calculé à partir de la carte $W_k = (w_k(pix))_{pix=1..P}$ correspondant à l'orientation $\tilde{\theta}_k$ la plus proche de θ_i . Ce potentiel est défini par la moyenne des valeurs des pixels de l'ensemble V_i prises dans la carte W_k :

$$\tilde{v}_d(i) = \frac{1}{\text{card}(V_i)} \sum_{pix \in V_i} w_k(pix), \quad k = \arg \min_j |\theta_i - \tilde{\theta}_j| \quad (3.46)$$

L'expression du terme d'attache aux données est finalement la même que dans l'équation (3.39) en remplaçant v_d par $\tilde{v}_d(i)$:

$$\tilde{h}_d(\mathbf{s}) = \exp \left(-\omega_d \sum_{s_i \in \mathbf{S}} \tilde{v}_d(i) \right) \quad (3.47)$$

Notons que l'utilisation de ce terme d'attache aux données sera moins précise que la précédente puisqu'une seule longueur de segment est considérée pour les tests statistiques et que l'on doit souvent se résoudre à un petit nombre d'orientations pour un gain significatif en temps de calcul. Néanmoins, cette méthode peut s'avérer très efficace si l'on y associe un noyau de proposition fondé sur ces pré-calculs comme cela est décrit dans le paragraphe 3.2.3.

3.2.2 Approche bayésienne

Nous avons dans un deuxième temps étudié la possibilité de travailler dans un cadre bayésien. Mais pour rentrer dans ce cadre, nous devons en premier lieu définir la vraisemblance des observations par rapport à une configuration de segment \mathbf{s} . On se heurte alors à la complexité des images. En effet, le fond de l'image est souvent non homogène, des contours autres que ceux du réseau à extraire peuvent être présents dans l'image, et des objets (arbres, voitures, bâtiments, etc.) peuvent interférer avec le réseau. Au vu de ces considérations, il devient difficile, voire impossible, de définir la vraisemblance des données brutes via la construction d'un modèle de fond et d'un modèle associé au réseau. Une solution serait de travailler sur les contours de l'image plutôt que l'image elle-même. Mais, comment expliquer la présence de contours n'appartenant pas au réseau? Plutôt qu'un détecteur de contours, nous proposons alors d'utiliser un filtre adapté à la détection des structures linéaires d'intérêt. Les sorties de ce filtre devront être des mesures évaluant la possibilité de la présence d'une structure linéaire en chaque pixel de l'image. Ces mesures pourront alors être utilisées en tant que données. Si ce filtre est de bonne qualité, une vraisemblance de la sortie de ce filtre Y par rapport à une configuration \mathbf{s} peut être définie comme suit.

Soit S la silhouette de la configuration de segments \mathbf{s} sur la grille pixélique : $S(i)$ est égale à 1 s'il existe un segment s_j de \mathbf{s} tel que l'ensemble de pixel associé V_j contienne le pixel i , et est nulle sinon. L'image cachée S est alors considérée comme une réalisation d'un champ caché \mathbf{S} correspondant au champ bruité \mathbf{Y} dont Y est une réalisation. Le

processus de bruit \mathbf{B} est supposé additif :

$$\mathbf{Y} = \mathbf{S} + \mathbf{B}$$

Supposons que la sortie du filtre Y prenne ces valeurs dans $[0, 1]$ et que celles-ci soient croissantes selon la confiance accordées à la présence d'une structure linéaire. Nous faisons alors les hypothèses suivantes sur le bruit :

$$B(i) = \begin{cases} |Z(i)| & \text{si } S(i) = 0 \\ -|Z(i)| & \text{si } S(i) = 1 \end{cases}$$

où Z est la réalisation d'un bruit blanc gaussien. La vraisemblance de Y peut alors s'écrire comme suit :

$$h_d(Y/S) = \prod_{i=1}^P 2 p_Z(Y(i) - S(i)) \quad (3.48)$$

$$h_d(Y/S) = \prod_{i=1}^P \frac{2}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(Y(i) - S(i))^2}{2\sigma^2}\right) \quad (3.49)$$

$$h_d(Y/S) = \left(2\sqrt{\frac{\lambda_d}{\pi}}\right)^P \prod_{i=1}^P \exp(-\lambda_d (Y(i) - S(i))^2) \quad (3.50)$$

où P est le nombre de pixels de la grille et $\lambda_d = \frac{1}{2\sigma^2}$.

En ce qui concerne le filtre de détection de structures linéaires, nous proposons de construire un filtre de la façon dont nous avons construit les cartes de pré-calcul du terme d'attache aux données. Il repose sur la définition d'une mesure de confiance σ_d mesurant la possibilité qu'un segment corresponde à une structure linéaire d'intérêt. La mesure σ_d est une mesure dans $[0, 1]$ définie comme l'opposée d'une mesure de potentiel à valeurs dans $[-1, 0]$. En prenant une mesure de potentiel du type de celle proposée dans l'équation (3.44), on définit la mesure de confiance σ_d comme suit :

$$\sigma_d(i) = \begin{cases} 0 & \text{si } T_i < \tau_1 \\ \frac{T_i - \tau_1}{\tau_2 - \tau_1} & \text{si } \tau_1 < T_i < \tau_2 \\ 1 & \text{si } T_i > \tau_2 \end{cases} \quad (3.51)$$

où τ_1 et τ_2 sont deux seuils à fixer selon le contraste de l'image, T_i est la valeur statistique donnée par l'équation (3.43) évaluant conjointement les hypothèses de contraste et d'homogénéité pour le segment s_i . De la même façon que pour le pré-calcul des potentiels, nous calculons les valeurs de confiance σ_d des segments de longueur minimale L_{min} , positionnés en chaque pixel de l'image, pour un nombre fixé d'orientations N_θ . En chaque pixel i , nous prenons la réponse maximale obtenue par tous les segments dont le masque associé contient i , et ceci indépendamment de l'orientation (contrairement,

au cas du pré-calcul des potentiels). La réponse au filtre Y s'écrit donc comme suit :

$$Y(i) = \max_{\substack{s_j : V_j \ni i \\ L_j = L_{min} \\ \theta_j = \tilde{\theta}_1 \dots \tilde{\theta}_{N_\theta}}} \sigma_d(j) \quad (3.52)$$

Par construction, ce filtre permet d'évaluer la présence de structures linéaires en chaque pixel par rapport aux hypothèses de contraste avec l'environnement proche et d'homogénéité interne (dans le cas de structures larges). L'évaluation de ces hypothèses repose sur des mesures statistiques, utilisées généralement pour le t-test de Student connu pour être adapté aux petits échantillons (*i.e.* dans notre cas, des petits ensembles de pixels). Il a également l'avantage de pouvoir être utilisé pour différents types de réseaux, la largeur et la longueur minimale des segments étant modulables. Tout autre filtre adapté à la détection de structures linéaires est utilisable ici et peut être combiné au filtre que nous venons de décrire dans un cadre de fusion de données.

3.2.3 Noyau de proposition fondé sur les données

Dans le paragraphe 3.2.1, nous avons proposé un pré-calcul du potentiel d'attache aux données pour différentes orientations en chaque pixel de l'image. Nous proposons ici d'exploiter ces pré-calculs pour construire un noyau de proposition fondé sur les données. La proposition de nouveaux segments se fait alors en fonction d'une mesure non homogène fondée sur des cartes de probabilités de présence de structures linéaires en chaque pixel. Par l'utilisation d'un tel noyau, nous évitons de nombreuses propositions de segments non pertinentes qui seront rejetées dans la phase de rejet/acceptation mais qui coûtent autant de temps de calcul qu'une proposition pertinente. L'algorithme d'optimisation devrait donc converger nettement plus rapidement que dans le cas d'une "naissance et mort uniforme".

Que ce soit pour le calcul du potentiel d'attache aux données ou pour le calcul de la sortie du filtre de détection de structures linéaires proposé dans le paragraphe 3.2.2, les pré-calculs que nous avons construits nous fournissent un indice d'appartenance d'un pixel à une structure linéaire. Ici, nous cherchons à définir des probabilités qu'un pixel corresponde au centre d'un segment. La procédure de pré-calcul s'en trouve simplifiée. En effet, il suffit de stocker pour chaque pixel i et chaque orientation $\tilde{\theta}_k$ de la discrétisation, le potentiel $v_d(i, \tilde{\theta}_k) = v_d(s)$ correspondant au segment $s = (p, L_{min}, \tilde{\theta}_k)$ dont le centre est p est positionné au niveau du pixel i (*i.e.* au centre du pixel i). Nous obtenons ainsi, pour chaque orientation, une carte C_k définissant un noyau de naissance non homogène :

$$C_k(i) = \frac{m - v_d(i, \tilde{\theta}_k)}{\sum_{j=1}^{N_{pix}} (m - v_d(j, \tilde{\theta}_k))} \quad , \quad i = 1, \dots, N_{pix} \quad (3.53)$$

où N_{pix} est le nombre de pixel de la grille et m est une constante choisie telle que C_k définisse une mesure de probabilité strictement positive. Par exemple, on pourra prendre :

$$m = V_{max} + 1$$

où V_{max} correspond à la borne supérieure du potentiel défini par l'équation (3.44). Plus le potentiel associé à un pixel i sera faible, plus la probabilité $C_k(i)$ de proposer le centre d'un segment à l'intérieur de ce pixel sera forte. La procédure de proposition d'ajout d'un nouveau segment u selon les cartes C_1, \dots, C_{θ_N} est la suivante :

- la longueur et l'orientation du nouveau segment θ_u et L_u sont tirées uniformément dans l'espace des marques $M = [L_{min}, L_{max}] \times [0, \pi]$;
- un pixel i est alors tiré selon la carte C_k , correspondant à l'orientation la plus proche de θ_u : $k = \arg \min_j [|\theta_u - \theta_j|]$;
- le centre du segment p_u est alors tiré uniformément dans le carré de F correspondant au pixel i de l'image.

Remarque 13 *Le tirage aléatoire du centre p_u dans le pixel i n'est pas réalisé en pratique puisqu'une différence de position sous-pixélique n'induit pas de différence de potentiel d'attache aux données. La différence énergétique est donc négligeable.*

Pour la mort, un segment est retiré uniformément de la configuration. Cette différence de symétrie dans le noyau "naissance et mort" fondé sur les données (NMD) est contre-balançée par le taux de Green qui dépend de la carte correspondant à l'orientation la plus proche du segment considéré.

Pour calculer le taux de Green nous utilisons la même mesure symétrique que pour le noyau "mort et naissance uniforme" pour dériver la mesure πQ_{NMD} où le noyau Q_{NMD} est défini comme suit :

$$Q_{\text{NMD}}(\mathbf{s} \rightarrow A) = p_{bD} Q_{bD}(\mathbf{s} \rightarrow A) + p_{dD} Q_{dD}(\mathbf{s} \rightarrow A) \quad (3.54)$$

où $Q_{bD} = Q_b$ est le noyau associé à la mort uniforme et Q_{dD} correspond au noyau associé à la naissance non homogène et peut s'écrire comme suit :

$$Q_{bD}(\mathbf{s} \rightarrow A) = \int_M \sum_{i=1}^{N_{pix}} C_{k_\theta}(i) \int_{[i] \subset F} \mathbf{1}_A(\mathbf{s} \cup s) \frac{dp}{|[i]|} dP_M(L, \theta) \quad (3.55)$$

où $k_\theta = \arg \min_j [|\theta - \tilde{\theta}_j|]$ et $[i]$ est le carré de F correspondant au pixel i . Si l'on suppose que tous les pixels sont représentés par des carrés de même taille dans F , l'aire de ce carré est alors égale à : $|[i]| = \frac{|F|}{N_{pix}}$. La somme des intégrales sur les carrés composant F pouvant s'écrire comme une intégrale sur F , l'équation 3.55 devient :

$$Q_{bD}(\mathbf{s} \rightarrow A) = \int_{F \times M} \frac{N_{pix} C_{k_\theta}(i_p)}{|F|} \mathbf{1}_A(\mathbf{s} \cup s) dp dP_M(L, \theta) \quad (3.56)$$

où le pixel i_p correspond au pixel où le point p est situé.

Dans le cas d'une naissance d'un segment $s = (p, L, \theta)$ fondée sur les cartes C_k , la dérivée de Radon-Nikodym de πQ_{NMD} par rapport à ψ , donnée par l'équation (3.30), est alors :

$$D(\mathbf{s}, \mathbf{s} \cup s) = \frac{f(\mathbf{s}) p_{bD} N_{pix} C_{k_\theta}(i_p)}{\lambda |F|} \quad (3.57)$$

où λ est l'intensité du processus de référence. Dans le cas d'une mort, la dérivée est donnée par :

$$D(\mathbf{s}, \mathbf{s} \setminus s) = \frac{p_{dD} f(\mathbf{s})}{n(\mathbf{s})} \quad (3.58)$$

Finalement, les taux de Green respectivement associés à une naissance et une mort d'un segment $s = (p, L, \theta)$ sont :

$$R(\mathbf{s}, \mathbf{s} \cup s) = \frac{f(\mathbf{s} \cup s)}{f(\mathbf{s})} \frac{p_{dD}}{p_{bD}} \frac{\lambda |F|}{N_{pix} C_{k_\theta}(i_p) (n(\mathbf{s}) + 1)} \quad (3.59)$$

$$R(\mathbf{s}, \mathbf{s} \setminus s) = \frac{f(\mathbf{s} \setminus s)}{f(\mathbf{s})} \frac{p_{bD}}{p_{dD}} \frac{N_{pix} C_{k_\theta}(i_p) n(\mathbf{s})}{\lambda |F|} \quad (3.60)$$

De la même façon que pour la “naissance et mort uniforme”, la “naissance et mort fondées sur les données” permet de vérifier, dans le cas où les probabilités $C_k(i)$ sont strictement positives, la **convergence ergodique** de la chaîne. S'il existe des pixels tels que $C_k(i) = 0$, ce noyau ne pourra pas garantir la convergence à lui seul, puisque l'espace d'état ne pourra être complètement exploré (on perd donc l'irréductibilité de la chaîne).

L'utilisation d'un tel noyau est pertinente pour l'échantillonnage à température faible (*i.e.* densité piquée) d'un processus spécifié par la densité complète $h_T(\mathbf{s}) = (h_p(\mathbf{s}) h_d(\mathbf{s}))^{\frac{1}{T}}$, et donc pour la recherche de la configuration de segments qui maximise la densité complète. En effet, les segments bien placés sont proposés plus souvent que les autres et sont acceptés dans les cas où ils vont dans le sens de *l'a priori* (*i.e.* s'ils ne sont pas mal situés par rapport aux autres segments de la configuration) car le rapport de densité est prépondérant par rapport à la division par la probabilité $C_k(i)$ de proposer un segment dont l'orientation est proche de $\tilde{\theta}_k$ au niveau du pixel i . Par contre, à température haute, les propositions de naissance seront souvent refusées et les propositions de mort souvent acceptées puisque dans ce cas le rapport des densités est proche de 1. Une idée serait alors de changer de mesure de référence μ : une mesure non homogène fondée sur les données permettrait de n'éliminer que très rarement les segments répondant bien aux données, et cela, même à température faible.

3.3 Réglage des paramètres

Afin d'extraire le réseau linéique présent dans les données, nous avons à estimer la configuration de segments \hat{s} qui maximise la densité complète du processus f :

$$\hat{s} = \arg \max_{\mathbf{s} \in E} \underbrace{f(\mathbf{s})}_{h_p(\mathbf{s}) h_d(\mathbf{s})} \quad (3.61)$$

où $E = \bigcup_{n=0}^{\infty} E_n$, avec E_n défini comme l'ensemble des configurations composées de n segments, h_p est la densité *a priori* et h_d est le terme d'attache aux données. Le problème du choix des paramètres définissant h_p et h_d se pose alors. Nous distinguons les paramètres de type "physique" ou "radiométriques" des paramètres de pondération.

Les **paramètres physiques** sont des paramètres utilisés dans la définition des objets du processus, dans la définition d'interactions entre objets et dans la construction du terme d'attache aux données. Plus précisément, on entend par paramètres physiques :

- les paramètres définissant l'espace des marques : L_{min} et L_{max} ;
- les paramètres de seuillage permettant de définir une relation (ex : ϵ , pour la connexion) ou permettant de différencier des interactions définies par rapport à la même relation (ex : τ_c , pour la connexion) ;
- les paramètres utilisés pour la définition du masque de pixels associé à un segment.

Ces paramètres dépendent directement du réseau que l'on cherche à extraire ou de certaines considérations plus générales, et sont relativement faciles à fixer. Par exemple, le nombre de bandes internes du masque associé à un segment dépend directement de la largeur du réseau dans l'image, *i.e.* de la résolution de l'image et du type de réseau d'intérêt (route, chemin ou rivière). Le paramètre ϵ doit être inférieur à $L_{min}/2$ pour une bonne définition de la relation de connexion. S'il est pris inférieur à 1 pixel, le réseau ne sera pas très flexible puisque les connexions seront concentrées sur 1 pixel. De plus, plus le paramètre ϵ est petit, moins les connexions sont probables par rapport à la mesure de référence et on aura donc tendance à souvent refuser le prolongement du réseau à température haute. En général, nous prenons ϵ entre deux et cinq pixels selon la résolution de l'image.

Nous entendons par **paramètres radiométriques** les paramètres de seuillage du terme d'attache aux données. Ceux-ci pourront être choisis en fonction du contraste du réseau avec son environnement proche. Nous étudierons leur robustesse dans le paragraphe 3.6.

Le choix des **paramètres de pondération** est plus délicat. Ce sont à ces paramètres que font référence les méthodes d'estimation de paramètres, les paramètres physiques étant supposés fixés. Dans le cas de données complètes (*i.e.* connaissant \hat{s}), [Stoica, 2001, van Lieshout et Stoica, 2001] proposent d'estimer les paramètres du modèle *a priori* par un algorithme de gradient stochastique. Celui-ci converge vers un optimum local situé dans le voisinage de la paramétrisation initiale. Cette méthode reste d'application limitée, étant donné que l'objectif principal est, en général, de trouver la configuration \hat{s} qui maximise la densité. Dans le cas de données incomplètes (\hat{s} inconnu), on cherche à la fois à estimer \hat{s} et les paramètres w_i des différents termes. L'algorithme

d'estimation le plus populaire est l'algorithme EM (Expectation - Maximisation), reposant sur la maximisation de l'espérance de la log-vraisemblance. Cependant, il n'est pas toujours adapté : il ne résout pas le problème de la fonction de partition et ne garantit pas la convergence vers le maximum de vraisemblance, tout comme les variantes de l'EM permettant de travailler avec des densités non connues à un facteur près. Nous proposons une alternative à l'estimation des paramètres en proposant de choisir les paramètres de pondération à l'aide de la méthode des **boîtes qualitatives** [Azencott, 1992]. Cette méthode permet de fournir des estimations grossières, souvent suffisantes, des paramètres. On utilise des spécifications locales afin d'obtenir un jeu d'inégalités. Sur la base de considérations heuristiques, on obtient un système d'inéquations linéaires qui fournit un ensemble de paramètres admissibles. Cet ensemble peut être vide si le modèle est trop restrictif. Nous donnons deux exemples de calibrage des paramètres de pondération de la densité complète avec une densité *a priori* h_p de type "Quality Candy" : le premier exemple concerne une approche par champ externe, le deuxième concerne une approche bayésienne.

3.3.1 Calibrage des paramètres pour une approche par champ externe.

À la sortie de l'algorithme d'optimisation, nous souhaitons obtenir une configuration vérifiant les propriétés suivantes :

1. Pas de segment libre
2. Pas de paire de segments simples non connectés au reste du réseau
3. Pas d'interruption dans le réseau qui pourrait être comblée par l'ajout d'un segment double s_i n'induisant pas de points de forte courbure ($g_c(s_i, \cdot) \leq 0$) et n'interagissant pas avec un segment proche ($s_i \not\sim_p \cdot$)
4. Pas de branches composées de plusieurs segments consécutifs aberrants par rapport aux données ($v_d(i) = V_{max}$)
5. Toute branche doit être prolongée au maximum : pas de fin de branche pouvant être prolongée par un segment s_i de bonne qualité, *i.e.* tel que $v_d(i) \leq 0$, $g_c(s_i, \cdot) \leq 0$ et $s_i \not\sim \cdot$
6. Pas de segments trop proches

Rappelons que la densité complète du modèle “Quality Candy” avec une approche par champ externe est donnée par les équations (3.11) et (3.39) :

$$h(\mathbf{s}) = \exp - \underbrace{\left[\omega_\beta n + \omega_f n_f + \omega_s n_s + \omega_c \sum g_c(s_i, s_j) + \omega_p \sum g_p(s_i, s_j) + \omega_d \sum v_d(i) \right]}_{U(\mathbf{s})}$$

avec $\begin{cases} \omega_\beta = \log(\beta) \\ \omega_f, \omega_s, \omega_c, \omega_p, \omega_d > 0 \end{cases}$ (3.62)

où $U(\mathbf{s})$ est l'énergie de la configuration \mathbf{s} . Les propriétés sur la configuration finale peuvent être traduites par les contraintes suivantes sur les paramètres énergétiques :

1. L'énergie d'un segment libre s_1 doit être plus élevée que l'énergie de la configuration vide. Le cas extrême vérifiant cette contrainte est celui où $v_d(1)$ est égal à -1 . D'où l'inégalité suivante :

$$\omega_\beta + \omega_f - \omega_d > 0 \quad (3.63)$$

2. L'énergie de deux segments simples s_1 et s_2 doit être plus élevée que l'énergie de la configuration vide. Considérant le cas extrême où $v_d(1), v_d(2)$ et $g_c(s_1, s_2) = 1$, on a :

$$2\omega_\beta + 2\omega_s - \omega_c - 2\omega_d > 0 \quad (3.64)$$

3. L'énergie doit décroître par l'ajout d'un segment double s_i reliant deux segments simples tel que les potentiels de connexion associés $g_c(s_i, \cdot)$ soit inférieurs ou égaux à zéro et n'interagissant pas selon la relation de proximité. Le cas extrême pour lequel cette condition est vérifiée est atteint pour $g_c(s_i, s_j) = 0$ pour tout segment s_j connecté avec s_i et pour un potentiel d'attache aux données maximal : $v_d(i) = V_{max}$. L'expression de cette contrainte est alors :

$$V_{max} \omega_d + \omega_\beta < 2\omega_s \quad (3.65)$$

Nous imposons néanmoins une augmentation d'énergie quand les connexions sont de mauvaise qualité : $g_c(s_i, \cdot) = 1$. On a donc également :

$$2\omega_c + V_{max} \omega_d + \omega_\beta > 2\omega_s \quad (3.66)$$

4. L'énergie doit augmenter par l'ajout de deux segments doubles consécutifs qui ont un potentiel égal à V_{max} . En considérant le cas extrême où les potentiels de connexion sont égaux à -1 , on obtient :

$$-3\omega_c + 2V_{max} \omega_d + 2\omega_\beta > 2\omega_s \quad (3.67)$$

5. L'énergie doit diminuer par l'ajout d'un segment simple s_i à un autre segment simple s_j si $g_c(s_i, s_j) \leq 0$, $v_d(i) \leq 0$ et $s_i \not\sim s_j$:

$$\omega_\beta < 0 \quad (3.68)$$

A l'inverse, elle doit augmenter si $v_d(i) = V_{max}$:

$$\omega_\beta - \omega_c + V_{max} \omega_d > 0 \quad (3.69)$$

6. La dernière propriété (pas de segments trop proches) est vérifiée dès que $\omega_p > 0$ du fait de l'utilisation d'un potentiel de type "hard core" éliminant les couples de segments proches et quasiment parallèles. Le poids ω_p agit uniquement sur les interactions de type proximité dont l'orientation est suffisamment grande. S'il est pris grand par rapport aux autres poids, les interactions de type proximité seront fortement pénalisées, et d'autant plus que l'angle formé sera faible.

Les équations (3.67), (3.64) et (3.68) impliquent la contrainte suivante sur le potentiel maximal d'attache aux données :

$$\begin{aligned} V_{max} &> \frac{1}{\omega_d} \left(\omega_s - \omega_\beta + \frac{3\omega_c}{2} \right) \\ V_{max} &> \frac{1}{\omega_d} \underbrace{\left(\omega_s + \omega_\beta - \frac{\omega_c}{2} \right)}_{>\omega_d} \underbrace{\left(-2\omega_\beta + 2\omega_c \right)}_{>0} \\ V_{max} &> 1 \end{aligned}$$

Nous fixons alors le paramètre V_{max} à 2 par soucis de simplification du système d'inéquations :

$$V_{max} = 2 \quad (3.70)$$

Les inéquations ci-dessus permettent de définir les encadrements des paramètres tels que ceux-ci définissent des densités qui atteignent leur maximum pour des configurations respectant chacune des contraintes énoncées ci-dessus. En posant $\omega_d = x \omega_c$, on obtient les encadrements donnés dans la table 3.1 où a , b et c sont des constantes positives. Plus le coefficient x est grand, plus le poids des données sera grand par rapport à la qualité des connexions. Un exemple est donné avec la contrainte supplémentaire que le potentiel d'un segment libre soit plus grand ou égal au potentiel de deux segments simples connectés par une connexion de qualité positive (dans le cas où l'on considère des segments répondant correctement aux données) afin d'accepter souvent la naissance dans le voisinage des segments libres :

$$2\omega_s + \omega_\beta \leq \omega_f \quad (3.71)$$

3.3.2 Calibrage des paramètres pour le modèle bayésien

De même que pour l'approche par champ externe, on peut définir des contraintes sur l'énergie *a posteriori* dans un cadre bayésien. L'énergie *a posteriori* est donnée par l'équation suivante :

$$U(\mathbf{s}/Y) = U_p(\mathbf{s}) + \lambda_d \sum (Y(i) - S(i))^2 \quad (3.72)$$

$V_{max} = 2$	ω_p	ω_β	x	ω_s	ω_c	ω_d	ω_f
Encadrement	> 0	$-a$	> 2	$b > (3x + \frac{5}{2})a$	$\frac{2b+a}{2(x+1)} < c < \frac{2b-2a}{2x+1}$	xc	$> xc + a$
Exemple	200	-1	4	33	7	28	67

TAB. 3.1 – Choix des paramètres énergétiques du modèle “Quality Candy” avec incorporation des données par champ externe par la méthode des boîtes qualitatives.

où U_p est l'énergie *a priori*, Y correspond aux données (sortie du filtre de détection de structures linéaires), S est la silhouette de s sur la grille de pixel, et λ_d est un poids positif dérivant de la variance d'un processus de bruit. Nous cherchons à minimiser l'énergie $U(\mathbf{s}/Y)$ avec un modèle *a priori* de type “Quality Candy”. Celle-ci doit vérifier des contraintes sur la forme de réseau que nous désirons obtenir.

Tout d'abord un segment libre ne doit pas apparaître à la fin de la configuration à moins que son masque soit composé d'un nombre de pixel supérieur à n_{min} fixé par l'utilisateur. Nous avons donc la contrainte suivante sur les paramètres énergétiques du modèle *a posteriori* avec un modèle *a priori* de type “Quality Candy” :

$$\omega_f + \omega_\beta - n_{min} \lambda_d \geq 0 \quad (3.73)$$

où ω_f est le poids positif associé à un segment libre, ω_β le poids (positif ou négatif) associé à tout segment. L'équation (3.73) se justifie car nous ne travaillons ici qu'avec des valeurs comprises entre 0 et 1 (ce qui implique que la différence $Y(i) - S(i)$ est supérieure ou égale à -1).

La même inégalité s'impose pour deux segments simples connectés :

$$2\omega_s + 2\omega_\beta - \omega_c - \lambda_d n_{min} \geq 0 \quad (3.74)$$

où w_c est le poids positif associé à la fonction de potentiel g_c à valeurs dans $[-1, 1]$ définie pour les paires de segments connectés.

Nous souhaitons de plus favoriser un prolongement du réseau. L'énergie doit alors décroître par l'ajout d'un segment double reliant deux segments simples tel que les potentiels associés aux connexions avec ce segment soient négatifs ou nuls. Nous ne considérons que les segments de masque de taille inférieure à n_{min} et ne vérifiant pas une relation de mauvaise orientation avec un autre segment de la configuration s .

$$n_{min} \lambda_d + \omega_\beta \leq 2\omega_s \quad (3.75)$$

Un segment simple répondant correctement aux données ajouté à l'extrémité non connectée d'un autre segment simple fera diminuer l'énergie si le potentiel associé à la clique est négatif et s'il n'y a pas de relation de mauvaise orientation entrant en jeu :

$$\omega_\beta \leq 0$$

Par contre, si ce segment est situé sur le fond de l'image (potentiel positif), l'énergie doit augmenter :

$$\omega_\beta \geq 0$$

Il vient donc :

$$\omega_\beta = 0 \tag{3.76}$$

Dans le cas d'un modèle bayésien, nous pouvons imposer la contrainte supplémentaire que deux petits segments simples connectés entre eux par une connexion de qualité maximale sont équivalent à un long segment libre u s'il induisent la même silhouette sur l'image que u . Dans ce cas, ils induisent le même potentiel d'attache aux données que u puisque celui-ci est calculé à partir de la silhouette. Ceci permet d'introduire la contrainte suivante :

$$2\omega_s - \omega_c = \omega_f \tag{3.77}$$

Remarquons que, sous la contrainte (3.77), les équations (3.73) et (3.74) sont équivalentes et impliquent l'équation (3.75) en prenant un potentiel ω_β nul.

Pour réduire l'espace des paramètres admissibles, nous imposons qu'un segment libre u soit accepté si l'ajout de u implique une diminution de l'énergie d'attache aux données supérieure à $n_{min} \lambda_d$. Ceci implique :

$$\omega_f = n_{min} \lambda_d \tag{3.78}$$

Finalement, avec un calibrage des paramètres selon la table 3.2, la densité atteindra son maximum pour des configurations présentant de bonnes propriétés. Un exemple est donné pour un nombre minimal de pixels n_{min} fixé à 40.

	ω_p	ω_β	ω_s	ω_c	ω_f	n_{min}	λ_d
Encadrement	$]0, \infty[$	0	$]0, \infty[$	$]0, 2\omega_s[$	$2\omega_s - \omega_c$	$\gg 1$	$\frac{\omega_f}{n_{min}}$
Exemple	10	0	7	4	10	40	0.25

TAB. 3.2 – Choix des paramètres énergétiques du modèle bayésien de type “Quality Candy” par la méthode des boîtes qualitatives.

3.4 Évaluation quantitative des résultats

Afin de mieux comparer les résultats fournis par différents algorithmes, il est important de définir des mesures de performance quantitatives par rapport à un réseau de référence (extraction manuelle, vérité terrain ou carte géographique), comme cela est proposé dans [Harvey, 1999, Sampère, 2001]. De plus, ces mesures quantitatives peuvent être intégrées dans un système de mise à jour cartographique automatique. Par exemple, dans [Hivernat *et al.*, 1999], une qualification automatique des résultats

permet d'améliorer itérativement la qualité d'une mise en correspondance entre un réseau routier issu d'une carte et un réseau extrait d'une image satellitaire, avant de fournir des informations utiles pour la mise à jour cartographique.

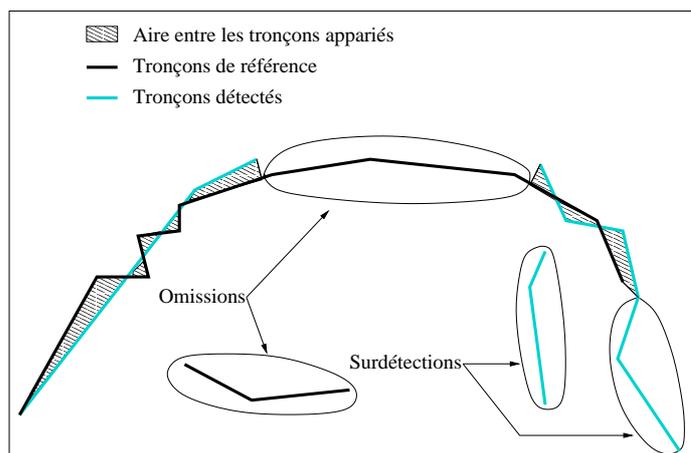


FIG. 3.13 – Mise en correspondance du réseau extrait par l'algorithme avec le réseau de référence.

Ainsi, dans les cas où nous disposons d'un réseau de référence, nous proposons une évaluation quantitative des résultats par rapport au réseau de référence extrait manuellement. Les segments fournis par l'algorithme sont mis en correspondance avec les branches de la référence (fournie sous forme de lignes brisées). Un segment est apparié à la référence si chacune de ses extrémités est à une distance du réseau de référence inférieure à un seuil fixé par l'utilisateur (on pourra prendre, par exemple, un seuil égal à dix pixels). Cet appariement, illustré par la figure 3.4, permet le calcul de :

- la longueur $L_0(\mathbf{s}_{ref})$ des tronçons du réseau de référence \mathbf{s}_{ref} qui ont été omis par l'extraction automatique ;
- la longueur $L_S(\mathbf{s})$ des tronçons du réseau extrait \mathbf{s} automatiquement qui correspondent à des fausses alarmes ;
- l'aire $A(\mathbf{s}_{ref}, \mathbf{s})$ comprise entre les branches appariées du réseau de référence \mathbf{s}_{ref} et du réseau détecté \mathbf{s} , comme illustré par les zones hachurées de la figure 3.4.

Trois critères d'évaluation du réseau extrait, dérivant de ces calculs, seront utilisés pour comparer quantitativement les résultats obtenus par les différents algorithmes d'extraction du réseau linéique :

θ , le pourcentage d'omissions par rapport à $L(\mathbf{s}_{ref})$, la longueur totale du réseau

de référence :

$$O = 100 \frac{L_0(\mathbf{s}_{ref})}{L(\mathbf{s}_{ref})} \quad (3.79)$$

S , le pourcentage de surdétectations par rapport à la longueur totale du réseau de référence :

$$S = 100 \frac{L_S(\mathbf{s})}{L(\mathbf{s}_{ref})} \quad (3.80)$$

D , une erreur géométrique correspondant à la distance D entre les tronçons appariés égale à l'aire comprise entre ces tronçons divisée par la longueur des tronçons de la référence qui ont été appariés :

$$D = \frac{A(\mathbf{s}_{ref}, \mathbf{s})}{L(\mathbf{s}_{ref}) - L_0(\mathbf{s}_{ref})} \quad (3.81)$$

3.5 Choix du modèle a priori

Nous nous sommes tout d'abord intéressé, dans le cadre de l'extraction non supervisée du réseau linéique, au cas le plus largement étudié dans la littérature, celui de l'extraction du réseau routier. Nous proposons de modéliser le réseau d'intérêt par un processus Markov objet où les objets sont des segments interagissant entre eux. Les réseaux routier sont généralement caractérisés par une faible courbure, de longues branches bien distinctes, continues et reliées entre elles. Les modèles *a priori*, définis dans le paragraphe 3.1, ont été construits pour incorporer ces fortes contraintes géométriques et topologiques. Les performances du modèle "Candy", introduit par [Stoica, 2001], pour la simulation et l'extraction des réseaux routiers ont déjà été montrées : les résultats sont prometteurs, présentant relativement peu de fausses alarmes et d'omissions pour une méthode complètement automatique. Notre premier objectif a été de déterminer si l'incorporation de mesures de qualité des interactions entre segments dans les densités *a priori* améliorerait les résultats et, si oui, de quelle façon est-il préférable de les incorporer.

Nous considérons ici une approche par champ externe. Les données sont donc incorporées par le terme d'attache aux données défini par l'équation (3.39) où le potentiel associé à chaque segment est donné par l'équation (3.44). Les seuils τ_1 et τ_2 associés à ce potentiel sont choisis en fonction du contraste entre le réseau et son environnement. Nous réalisons l'optimisation de chaque modèle par un recuit simulé sur un algorithme MCMC à sauts réversibles utilisant un noyau de "naissance et mort" uniforme pour l'exploration de tout l'espace d'état et des noyaux pertinents pour une exploration efficace : "naissance et mort" dans un voisinage par rapport à la relation de connexion et des petites perturbations de segments : rotation, translation et dilatation.

La figure 3.14 présente les résultats de l'extraction du réseau routier sur une image satellitaire par un recuit simulé avec une décroissance géométrique de la température : $T_{t+1} = c T_t$ avec $c = 0.99999$. L'image traitée est une image SPOT Panchro de moyenne résolution (10 mètres), de taille 256×256 pixels. Les routes principales présentes dans

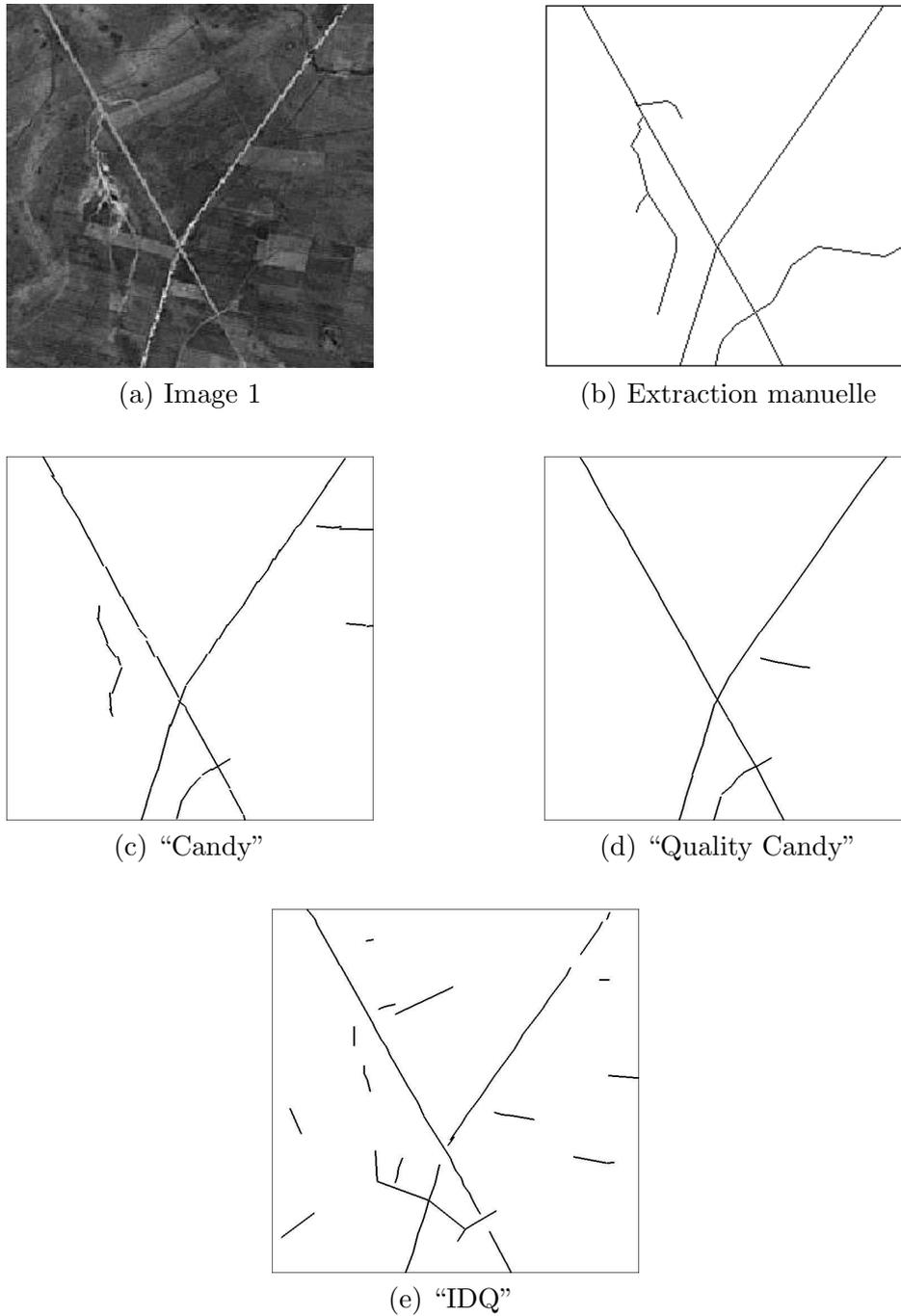


FIG. 3.14 – Résultats de l'extraction du réseau routier par recuit simulé à partir d'une image SPOT de 256×256 pixels fournie par le Centre National d'Etudes Spatiales (CNES) pour les trois modèles *a priori*.

cette image sont rectilignes et contrastent fortement avec le fond de l'image. Par contre, le réseau secondaire est sinueux et certaines parties présentent un très faible contraste (l'extraction manuelle en est rendue délicate). Le choix empirique des seuils τ_1 et τ_2 pour cette image est le suivant : $\tau_1 = 4$ et $\tau_2 = 8$.

Les résultats obtenus avec les modèles "Candy" et "Quality Candy" sont satisfaisants : ils présentent peu de surdétectations et l'on peut constater que l'image présente effectivement des structures linéaires au niveau de ces surdétectations ; les omissions sont principalement dues au fait que le réseau secondaire est sinueux pas endroit et très faiblement contrasté. Nous verrons par la suite qu'en diminuant les seuils τ_1 et τ_2 on peut réduire fortement ces omissions. L'utilisation de potentiels continus fondés sur des mesures de qualité des interactions dans le modèle "Quality Candy" permet d'obtenir un réseau de meilleure qualité que celui obtenu avec le modèle "Candy" : la courbure locale du réseau est plus faible (en général), la distance entre les extrémités connectées est sous-pixélique, et le croisement des deux routes principales est de meilleure qualité. Ces deux résultats ont été obtenus en un temps équivalent : un peu plus de 12 minutes pour le "Candy" modèle et un peu plus de 11 minutes pour le "Quality Candy" modèle avec un processeur Pentium III, 1 GHz, et 1 Go de mémoire vive.

Le résultat correspondant au modèle "IDQ" montre qu'il ne suffit pas qu'un modèle fournisse des échantillons de bonne qualité par un échantillonnage sans information radiométrique, pour être adapté à l'extraction du réseau routier. Le réseau extrait en utilisant le modèle "IDQ" est, en effet, de qualité médiocre alors que les échantillons correspondant au modèle *a priori* "IDQ" présentés dans le paragraphe 3.1.6 sont de meilleure qualité que ceux du modèle "Candy". De nombreux segments libres sont présents à la fin de l'algorithme induisant de nombreuses surdétectations et des interruptions sont visibles au niveau des deux routes principales. Ceci s'explique par la définition de l'intensité de connexion (moyenne des nombres de groupes d'objets en interaction auxquelles un objet appartient). Celle-ci correspondant à une moyenne sur toute la configuration, la présence de segments libres est contrebalancée par la présence de segments connectés à plus de deux segments. De plus, la qualité des interactions intervient dans le modèle "IDQ" au travers d'un potentiel positif auxquels les segments libres ne participent pas, et se voient donc pénalisés. La diffusion, *i.e.* la proportion d'objet connectés dans la configuration courante, permet de limiter la présence de segments libres mais un poids trop fort par rapport à la qualité induira un réseau de mauvaise qualité. Enfin, nous avons constaté que, dans les cas où le poids de la qualité de connexion est fort par rapport à celui de l'attache aux données, l'algorithme avait du mal à proposer de nouvelles branches de qualité correctes mais inférieure à la qualité des branches de très bonne qualité déjà détectées, ces nouvelles branches faisant baisser la qualité moyenne de la configuration. Ces constatations nous amènent à préférer un modèle faisant intervenir des potentiels locaux plutôt que globaux. De plus, ce résultat a été obtenu en plus de 17 minutes, ce qui est bien plus long que le temps nécessaire à l'obtention du réseau en utilisant les deux autres modèles.

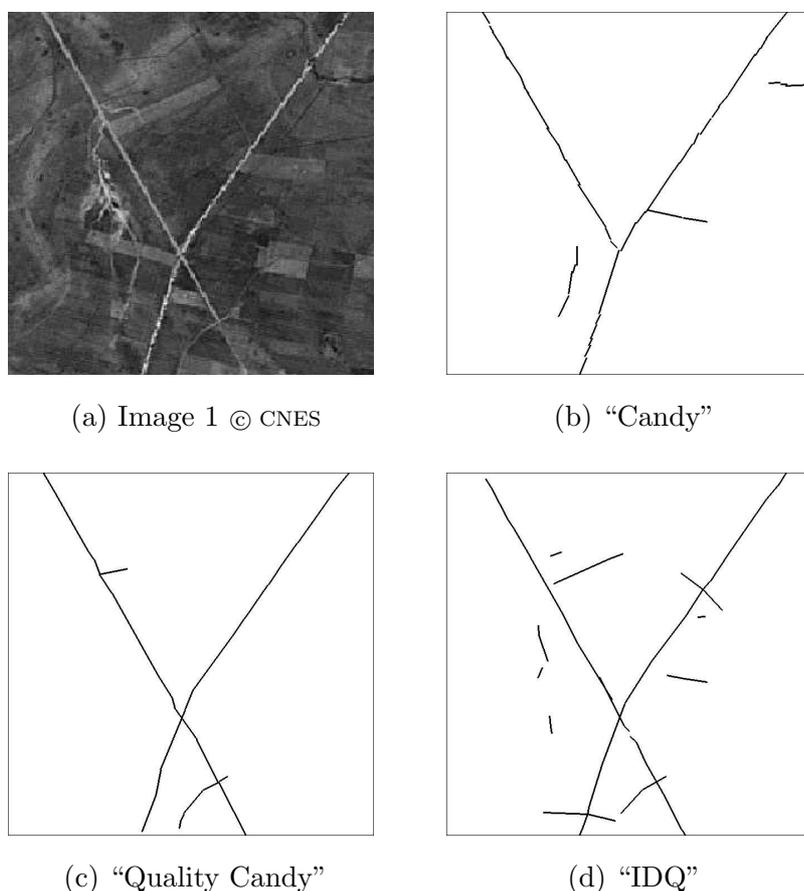


FIG. 3.15 – Résultats de l’extraction du réseau routier avec une décroissance rapide de la température.

La figure 3.15 présente les résultats de l’extraction du réseau routier obtenus sur la même image mais avec une décroissance plus rapide de la température ($c = 0.99999$). Cela permet d’obtenir des résultats beaucoup plus rapidement comme cela est montré dans le tableau 3.3 : en moins de 6 minutes pour le modèle “Candy”, moins de 4 minutes pour le modèle “Quality Candy” et moins de 16 minutes pour le modèle “IDQ”. Cependant, dans le cas du modèle “Candy”, le réseau obtenu est moins complet que le réseau obtenu avec une décroissance plus lente (figure 3.15). Notamment, toute une section d’une route principale est omise. En effet, il existe (au moins) une température critique par laquelle il est important de passer lentement pour atteindre le maximum global de la densité, ou au moins une bonne approximation de ce maximum global. Le réseau obtenu en utilisant le modèle “Candy” correspond à un optimum local très loin du maximum global. Par contre, les deux autres modèles sont moins sensibles à la vitesse de décroissance de la température que le modèle “Candy” : la qualité des

résultats est similaire à celle de ceux obtenus avec une décroissance de la température plus lente (figure 3.15).

	“Candy”	“Quality Candy”	“IDQ”
décroissance rapide	1.9×10^6 itérations 5 min 31 s	1.6×10^6 itérations 3 min 56 s	6.1×10^6 itérations 15 min 21 s
décroissance lente	3.6×10^6 itérations 12 min 28 s	4.5×10^6 itérations 11 min 17 s	7.2×10^6 itérations 17 min 30 s

TAB. 3.3 – Temps de calcul du recuit simulé sur l’image 1 selon la vitesse de décroissance de la température pour chaque modèle.

Finale­ment, nous avons choisi d’utiliser le modèle “Quality Candy” en tant que modèle *a priori*. L’étude comparative menée dans [Lacoste *et al.*, 2002] nous a conforté dans ce choix et nous a permis, à partir de résultats obtenus sur des images satellitaires et aériennes, d’établir le classement, donné dans le tableau 3.4, des trois modèles *a priori* selon le temps de calcul, le lissage induit par l’utilisation du modèle (*i.e.* faible courbure, continuité), et les taux d’omissions et de surdétections. Globalement, les modèles “Candy” et “Quality Candy” fournissent de bons résultats avec relativement peu de fausses alarmes et de surdétections, étant donné qu’aucun point d’amorce n’est fourni par un opérateur (l’algorithme étant initialisé avec la configuration vide). Le modèle “Quality Candy” donne des résultats de meilleure qualité en moins de temps, ce qui justifie l’utilisation de potentiels continus fondés sur des mesures de qualités. Par contre, le modèle “IDQ”, dont la densité est fondée sur des valeurs moyennes, n’est pas adapté à l’extraction des réseaux, mis à part pour les réseaux en forme de grille. De plus, le temps de calcul a pratiquement doublé par l’utilisation d’un tel modèle, dû à une convergence bien plus lente.

	“Candy”	“Quality Candy”	“IDQ”
Temps de calcul	+	+	--
Lissage	-	++	++
Exhaustivité	+	++	-
Fausses Alarmes	+	++	--

TAB. 3.4 – Résumé des performances pour chacun des modèles. (–) désigne de mauvaises performances et (+) de bonnes performances.

3.6 Robustesse des paramètres

Nous considérons dans ce paragraphe un modèle complet issu d'une approche par champ externe où le modèle *a priori* est le modèle choisi dans le paragraphe précédent, à savoir le modèle "Quality Candy". L'extraction des réseaux linéiques consiste alors à chercher la configuration de segments maximisant la densité complète du modèle. La densité complète dépend de nombreux paramètres plus ou moins faciles à fixer. Les paramètres physiques, tels que la longueur maximale d'un segment ou la différence d'orientation tolérée entre deux segments proches, sont faciles à fixer. De plus, mis à part les bornes L_{min} et L_{max} associées à la longueur d'un segment, la distance de connexion ϵ , tous les paramètres physiques restent identiques entre deux images. Finalement, la difficulté réside dans le choix des paramètres énergétiques ω_i et des seuils τ_1 et τ_2 correspondant aux paramètres de l'attache aux données. Nous vérifions dans ce paragraphe que :

- un même jeu de paramètres énergétiques est utilisable pour différentes images ;
- une faible variation des seuils τ_1 et τ_2 ne modifie que faiblement les résultats ;
- on peut utiliser les mêmes seuils pour des images présentant des réseaux linéiques de même contraste.

Nous avons choisi d'utiliser la paramétrisation donnée à titre d'exemple dans le tableau 3.1. Celle-ci devrait fournir des configurations de segments présentant de bonnes propriétés (pas de segment libre, pas de proximité, etc.). Nous avons tout d'abord testé cette paramétrisation sur des images 300×300 extraites d'une image satellitaire (SPOT XS2, de 20 mètres de résolution) de la Guinée-Conakry. Cette image a été fournie par le Bureau de Recherche Géologique et Minière (BRGM) qui a participé financièrement à ces travaux. Le BRGM a en outre participé à la phase de validation des résultats obtenus sur l'extraction du réseau hydrographique. L'item cartographique recherché est le réseau hydrographique. Ce dernier est repérable par les arbres présents au niveau des rivières. Étant donné le caractère sinueux de ce type de réseau, la relation connexion est définie pour toute différence d'orientation entre les segments concernés. Après plusieurs expérimentations sur la première image, nous avons choisi les valeurs 3 et 7 pour les deux seuils τ_1 et τ_2 . Les résultats du recuit donnés par la figure 3.16 montrent que les paramètres sont robustes pour des images de radiométrie similaire concernant le même réseau. En effet, les paramètres sont identiques pour les trois images, fournissant à chaque fois des résultats satisfaisants. De plus, une petite perturbation des paramètres τ_1 et τ_2 conduit à un réseau contenant la même structure principale (voir (a) et (c) comparativement à (b) dans la figure 3.16).

Les paramètres sont également robustes par rapport à la taille de l'image, comme le montre le résultat présenté dans la figure 3.17. Celui-ci concerne l'extraction du réseau à partir d'une image extraite de la même image mais plus grande que les trois premières. Avec les mêmes paramètres que précédemment, on obtient un réseau de bonne qualité.

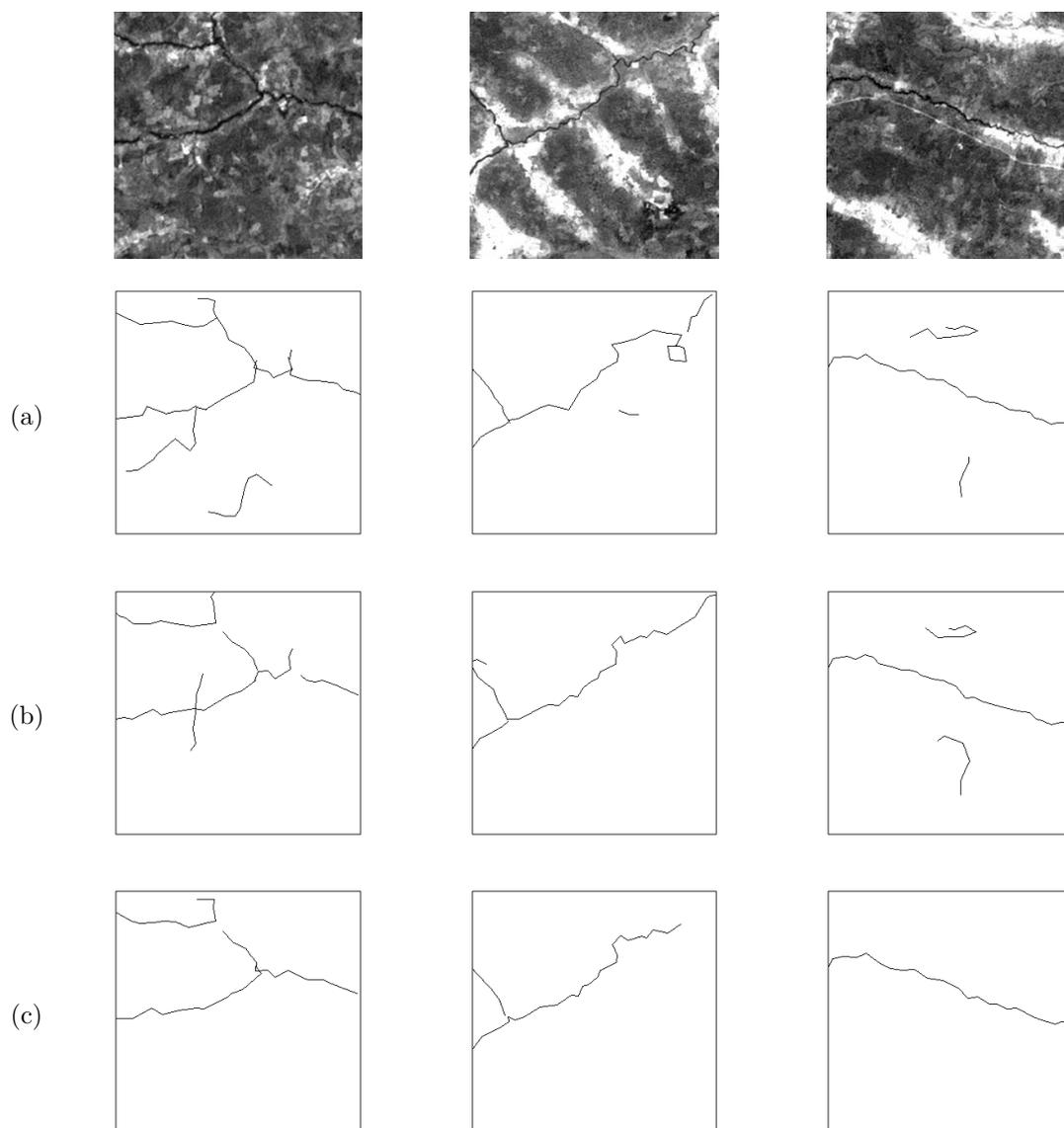


FIG. 3.16 – Résultats de l'extraction du réseau hydrographique à partir de trois images (300×300 pixels) extraites d'une image satellitaire (SPOT XS2) fournie par le BRGM, en utilisant les paramètres énergétiques donnés dans la table 3.1 et pour différentes valeurs des seuils τ_1 et τ_2 associés au terme d'attache aux données : (a) $\tau_1 = 2$ et $\tau_2 = 7$, (b) $\tau_1 = 3$ et $\tau_2 = 7$, (c) $\tau_1 = 3$ et $\tau_2 = 8$.

Omissions ~ 16% - Surdétectons ~ 11% - Distance ~ 1.7 pixels

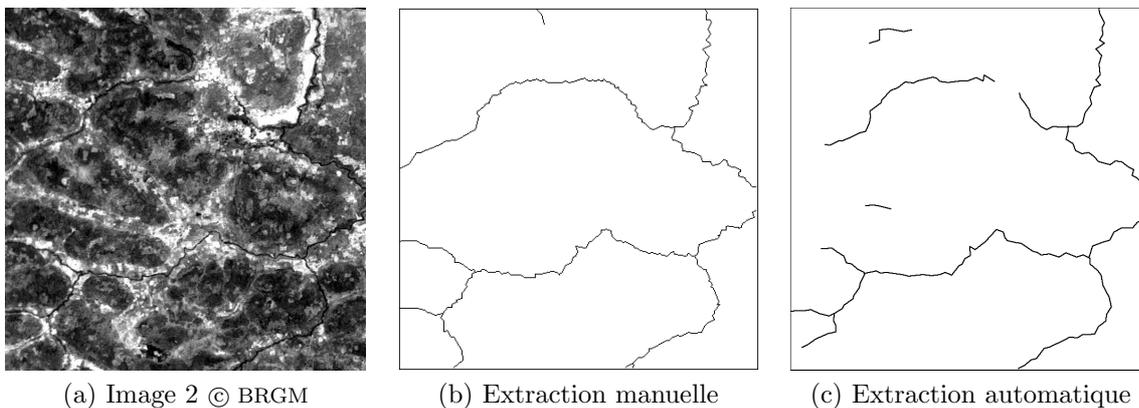


FIG. 3.17 – Résultat de l'extraction du réseau hydrographique à partir d'une image satellitaire (SPOT XS2) de taille 682×674 pixels, avec $\tau_1 = 3$ et $\tau_2 = 7$. Une extraction manuelle du réseau nous a été fournie par un expert du BRGM et nous permet d'évaluer quantitativement le résultat par rapport à ce réseau.

Omissions ~ 4% - Surdétectons ~ 20% - Distance ~ 1.0 pixels

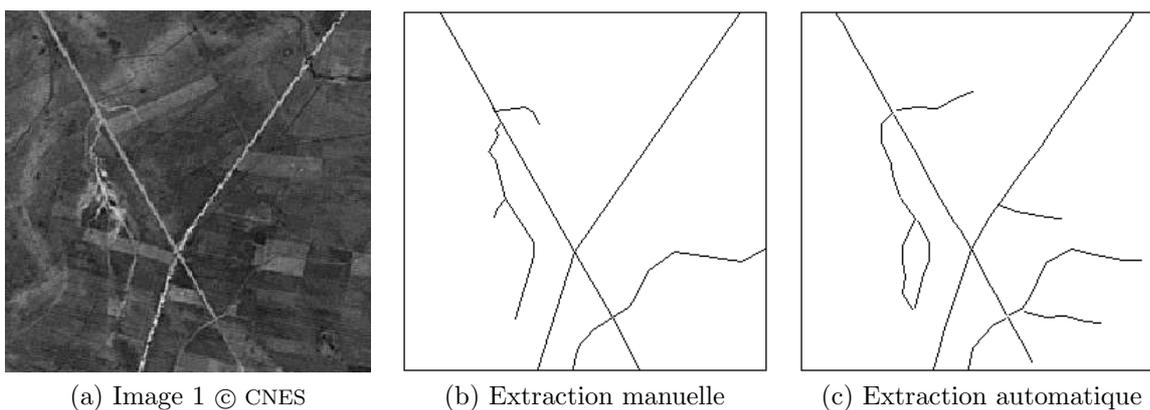


FIG. 3.18 – Résultats de l'extraction du réseau routier à partir d'une image SPOT (Panchro) de taille 256×256 pixels, avec $\tau_1 = 3$ et $\tau_2 = 7$.

Enfin, le même jeu de paramètres énergétiques a été testé sur une image SPOT de taille 256×256 pixels (Panchro et de résolution 10 m) où l'item cartographique recherché est le réseau routier. Le seul changement concerne le seuil τ_c de la relation de connexion dû à la différence de géométrie entre les réseaux routiers et les réseaux hydrographiques. Les seuils du terme d'attache aux données sont les mêmes que précédemment, *i.e.* $\tau_1 = 3$ et $\tau_2 = 7$. La principale difficulté ici est de détecter les chemins sinueux passant à travers les champs, qui ne sont pas rectilignes et ne contrastent que faiblement avec leur environnement. Un résultat de l'extraction avait déjà été fourni dans la figure 3.14 avec $\tau_1 = 4$ et $\tau_2 = 8$ dans lequel les chemins n'avait pratiquement pas été détectés. La figure 3.18 montre la pertinence de l'utilisation des seuils choisis empiriquement pour la détection des forêts galeries quand l'objectif est de détecter tout le réseau. Le résultat, obtenu en moins de 5 minutes avec un processeur Pentium de puissance 1 GHz, correspond à un réseau continu et avec un taux d'omissions inférieur à cinq pour cent mais aussi un taux de surdétectations de l'ordre de vingt pour cent. Ces importantes surdétectations sont à modérer par le fait que l'image présente effectivement des structures linéiques au niveau de ces surdétectations. Peut-être est-ce la référence qui comporte des omissions? Pour une évaluation quantitative plus pertinente, il serait intéressant de considérer plusieurs extractions manuelles.

3.7 Résultats de l'extraction du réseau linéique par processus de segments

3.7.1 Approche par champ externe

La méthode d'extraction proposée ici repose sur la définition d'un processus de segments par la densité complète donnée par :

$$h_p(\mathbf{s}) \propto h_p(\mathbf{s}) h_d(\mathbf{s}) \quad (3.82)$$

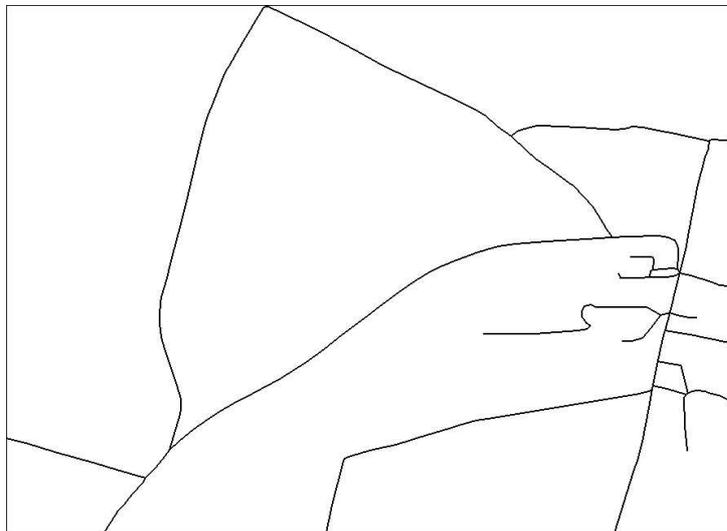
où $h_p(\mathbf{s})$ est la densité *a priori* donnée par l'équation (3.7) et $h_d(\mathbf{s})$ est le terme d'attache aux données donnés par l'équation (3.39). L'optimisation est réalisée via un recuit simulé sur un algorithme de type MCMC à sauts réversibles utilisant un noyau de "naissance et mort" uniforme et des noyaux pertinents pour une exploration efficace : "naissance et mort" dans un voisinage par rapport à la relation de connexion et des petites perturbations de segments. La décroissance de la température est géométrique :

$$T_{t+1} = c T_t$$

avec une constante c très proche de 1 (en général, $c = 0.999999$). Ce paragraphe présente les résultats de l'extraction du réseau linéique (routes, rivières) à partir d'images (haute et moyenne résolution, satellitaires et aériennes, optique et radar) sans utilisation de pré-calculs.



(a) Image 3 © IGN



(b) Extraction manuelle

FIG. 3.19 – Image aérienne de la région d’Altkirsh en Alsace de taille 1784×1304 pixels et de 50 centimètres de résolution fournie par l’Institut Géographique National.

Performance de la méthode en imagerie haute résolution

L'image présentée dans la figure 3.19 est une image aérienne de la région d'Alt-kirsh en Alsace fournie par l'IGN. La haute résolution de cette image (50 centimètres) induit un bruit géométrique, beaucoup plus difficile à prendre en compte dans les modèles pixéliques usuels que les bruits provenant du capteur. En effet, certains arbres interfèrent avec la bonne détection de certaines portions de route; certains champs sont de même niveau de gris que les routes; et les textures de certains champs sont en adéquation avec les hypothèses de contraste avec l'environnement proche et d'homogénéité interne sur lesquelles est fondé le potentiel d'attache aux données. On voit là tout l'intérêt d'avoir un bon *a priori*.

Omissions $\sim 16\%$ - *Surdétections* $\sim 4\%$ - *Distance* ~ 1.6 pixels



FIG. 3.20 – Résultats de l'extraction sur une image aérienne (figure 3.19).

Les résultats sont globalement très satisfaisants en termes de limitation des surdétctions et de précision : nous obtenons moins de 4% de surdétctions et les branches bien détectées sont à une distance de l'ordre du pixel du réseau extrait manuellement. La surdétction majeure correspond à une branche passant sur un bâtiment et reliant deux routes quasi-parallelèles. Or nous avons construit le modèle pour qu'il relie de telles branches. Si nous ne l'avions pas fait, nous n'aurions pu traiter les larges occlusions des routes dues à la présence d'arbres à proximité. Nous observons un taux d'omissions d'environ 16%. La figure 3.20 révèle que se sont en majorité les routes secondaires qui ont été omises. Cela peut s'expliquer par le fait que ces routes sont courtes et d'une courbure plus importante qu'au niveau des grands axes routiers, ce qui ne correspond pas au modèle a priori choisi. Remarquons aussi que la détection est difficile au niveau

des intersections à angle aigu : ceci est dû à la définition de la connexion par rapport aux extrémités des segments (et non d'une extrémité par rapport au reste du réseau) et à la pénalisation de la relation de proximité. A cela s'ajoute un faible contraste des routes avec leur environnement, au niveau des deux croisements où les routes n'ont pas été détectées.

Performance de la méthode en imagerie radar

La majorité des images que nous avons traitées sont des images optiques, mais nous avons également testé la méthode en imagerie radar. Nous présentons dans la figure 3.21 un résultat obtenu sur une image radar ERS de Mexico de taille 525×546 pixels. Le bruit de chatoiement associé à ce type d'image rend la détection du réseau problématique. Le résultat obtenu montre que le terme d'attache aux données permet la détection de routes dans les images très bruitées. De plus, le potentiel de connexion permet d'éviter les ruptures dans le réseau.

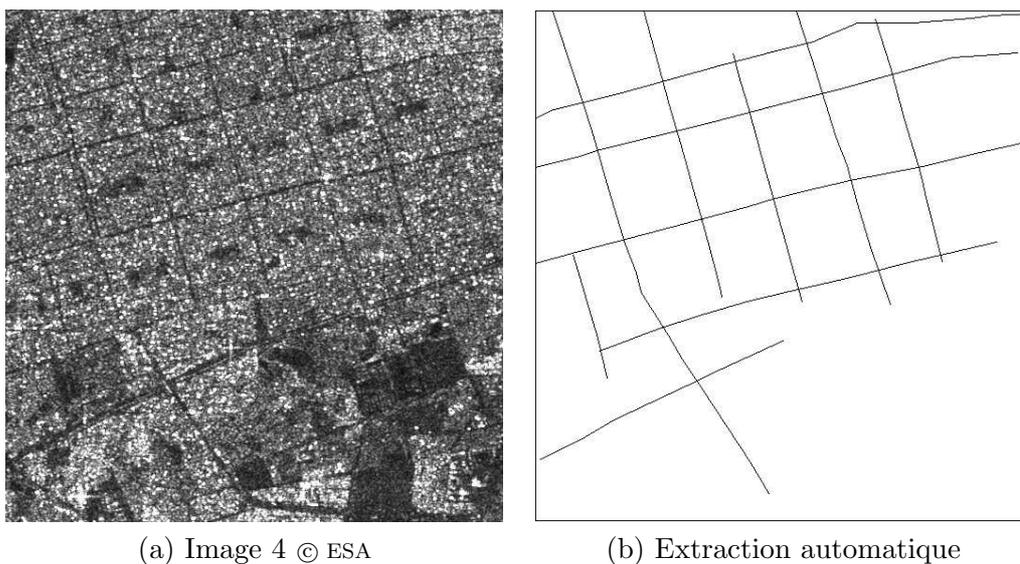


FIG. 3.21 – Résultats de l'extraction sur une image radar ERS (525×546 pixels) © European Space Agency.

Des résultats prometteurs pour l'extraction des réseaux sinueux

Ce modèle s'adapte aussi au cas des réseaux plus sinueux de façon encourageante, comme nous l'avons montré dans le paragraphe 3.6 par les résultats obtenus sur une image satellitaire (SPOT XS2, de 20 mètres de résolution) de la Guinée-Conakry où l'item cartographique recherché est le réseau hydrographique (figures 3.16 et 3.17). Nous montrons dans la figure 3.22 le meilleur résultat que nous avons obtenu sur l'image 2.

Les résultats (obtenus en 40 minutes) sont prometteurs puisque nous obtenons un

Omissions $\sim 13.7\%$ - Surdétectons $\sim 9.6\%$ - Distance ~ 1.5 pixels

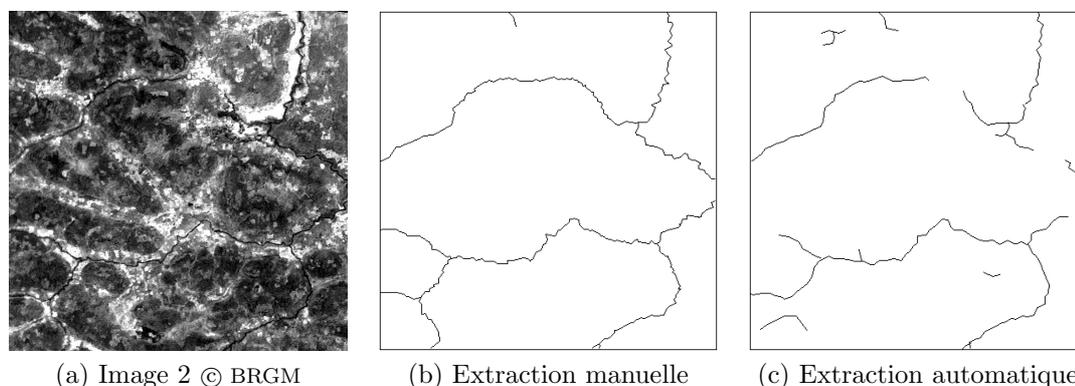


FIG. 3.22 – Résultat de l'extraction du réseau hydrographique à partir d'une image SPOT de taille 682×674 .

réseau continu avec des taux de surdétectons et d'omissions inférieurs à quinze pour cent malgré la difficulté de l'image (présence de zones peu contrastées, réseau très sinueux). Cependant, l'effet de lissage du potentiel de la clique de connexion, très utile dans le cas de l'extraction du réseau routier, présente ici un inconvénient puisque cela fournit un réseau moins sinueux que la vérité terrain. De plus, le fait d'utiliser un terme d'attache aux données fondé sur des potentiels en chaque segment plutôt que globalement ne nous permet pas de détecter les parties très sinueuses du réseau. En effet, la taille d'un segment ne peut être trop réduite, car, dans ce cas, les mesures de contraste et d'homogénéité ne seraient pas pertinentes. Bien que L_{min} soit fixée à 5 pixels pour cette image, cela ne suffit pas à suivre parfaitement le réseau.

Performance de la méthode comparativement à une méthode d'extraction fondée sur les champs de Markov sur Graphe

Comparer la méthode proposée avec des méthodes existantes est difficile étant donné que la plupart des méthodes d'extraction du réseau linéique performantes sont en général semi-automatiques ou reposent sur une exploitation de données externes comme les données géographiques. Les systèmes d'interprétation se situent quand à eux à un niveau au dessus, exploitant une information sémantique et réalisant l'extraction en combinant les résultats de différents algorithmes d'extraction et en incorporant des composants d'auto-évaluation. Nous nous situons dans une approche de plus bas niveau qui pourrait tout à fait s'intégrer dans un tel système. Parmi les méthodes les plus proches des approches par processus Markov objet, on peut citer les approches par champ de Markov sur graphe présentées dans le paragraphe 1.3.2. Rappelons que ces approches consistent à minimiser l'énergie définie sur un graphe construit par une pré-détection des structures linéaires, où chaque noeud correspond à une portion de route.

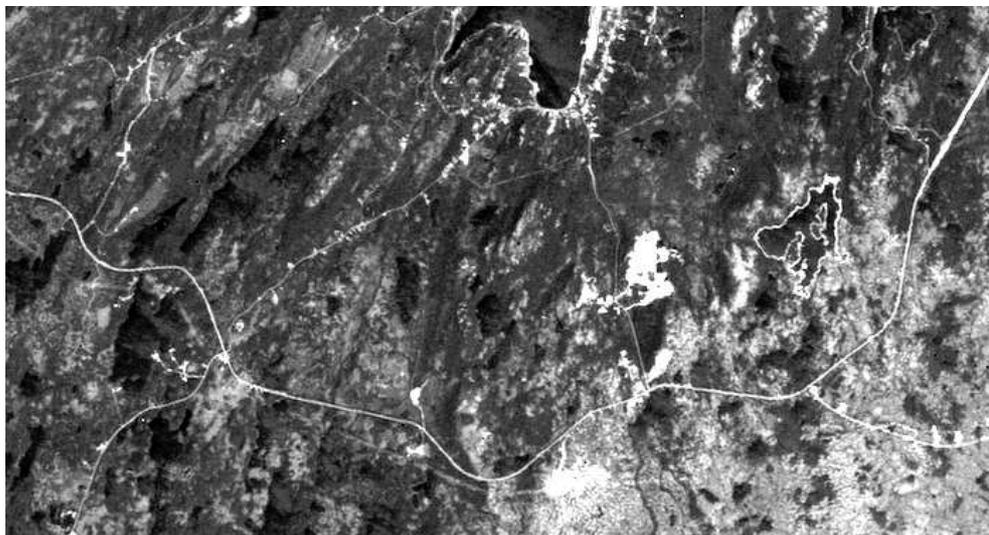
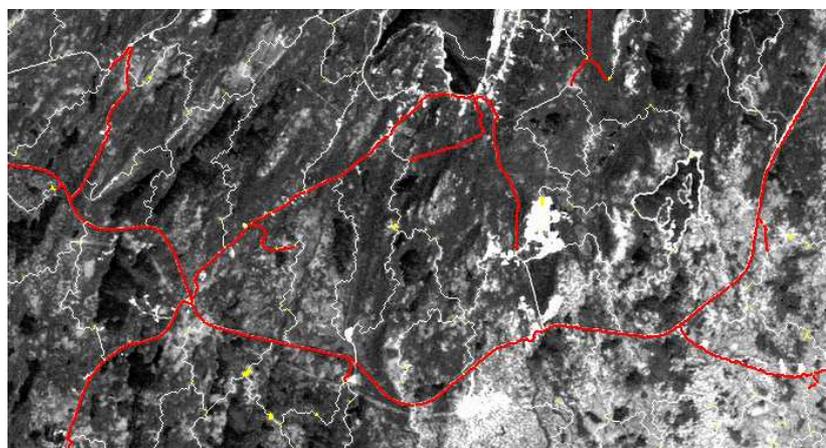


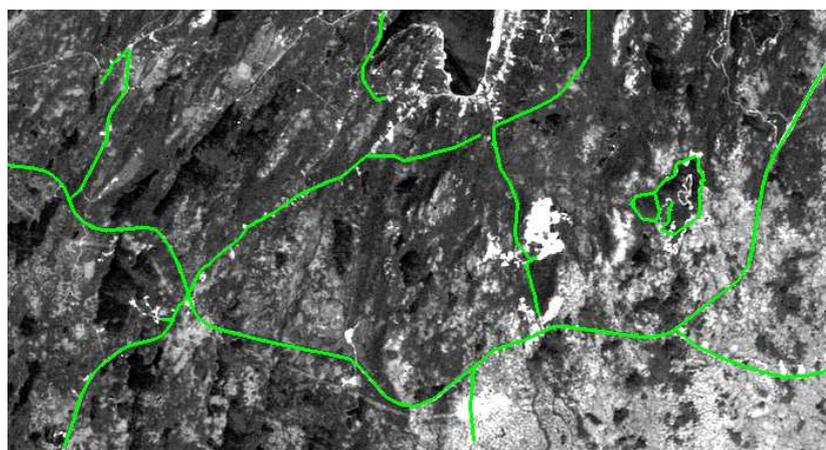
FIG. 3.23 – Extrait de 700×380 pixels d'une image Landsat de la ville St. Johns (Canada), de 25 m de résolution © Gouvernement of Canada with permission from Natural Resources Canada.

Nous comparons ici nôtre méthode à la méthode par champ de Markov sur graphe proposée par [Géraud, 2003]. L'initialisation du graphe est de très bonne qualité du fait de l'utilisation d'une approche région : une méthode de "ligne de partage des eaux" est appliquée à une image de potentiel (potentiel mesurant l'appartenance aux structures curvilignes à détecter) filtrée par une fermeture d'aire pour une suppression des minima locaux. [Géraud, 2003] a appliqué cette méthode sur une image extraite d'une image Landsat de la ville St. Johns (Canada), de 25 m de résolution et de 7 bandes spectrales. L'image utilisée est une image en niveau de gris donnée dans la figure 3.23.

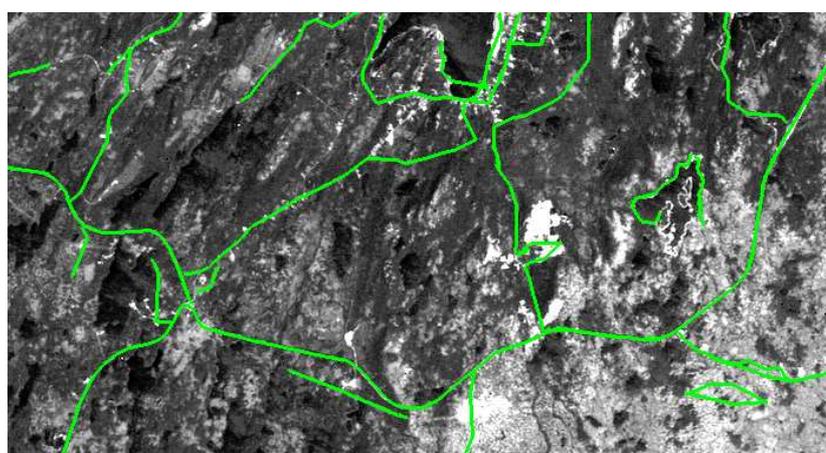
La figure 3.24 présente le résultat obtenu par [Géraud, 2003] et deux résultats obtenus par processus de segments, le premier correspondant à un seuil de contraste τ_2 plus élevé que le second. Globalement les trois résultats sont bons, fournissant les routes principales sans interruptions malgré la mauvaise qualité de l'image traitée : seulement 30 niveaux de gris sont utilisés. De plus, il n'y a pratiquement pas de fausses alarmes dans les trois cas. Les images (a) et (b), correspondant au résultat obtenu par [Géraud, 2003] et au premier résultat obtenu par processus de segment, présentent des réseaux de qualité similaire, l'extraction étant relativement sélective par rapport au contraste entre le réseau et son environnement. L'image (c), correspondant à un terme d'attache aux données moins sélectif par rapport au contraste entre le réseau et son environnement, fournit un résultat plus exhaustif que les deux premiers. On observe quelques fausses alarmes qui pourraient facilement être éliminées par un post-traitement. Les omissions correspondent à des sections très peu contrastées (presque non visible



(a) champ de Markov sur graphe



(b) processus de segments avec $\tau_1 = 2$ et $\tau_2 = 7$



(c) processus de segments avec $\tau_1 = 2$ et $\tau_2 = 6$

FIG. 3.24 – Résultats de l'extraction du réseau linéique sur l'image donnée dans la figure 3.23 obtenus par champ de Markov sur graphe (en rouge) et processus de segments (en vert). Les lignes blanches, obtenues par un algorithme de ligne de partage des eaux, correspondent aux noeuds du graphe initial. Les points jaunes correspondent aux arcs du graphe.

à l’œil nu) et au niveau des routes où les habitations présentes à proximité induisent une intensité très élevée et rendent la détection, via une évaluation de différence de moyenne, impossible. On remarque que certaines branches détectées dans l’image (c), correspondant bien au réseau réel, n’étaient pas présentes dans l’initialisation du graphe (en blanc, dans l’image (a)) et n’ont pu être détectées par optimisation sur ce graphe. C’est un inconvénient des méthodes par champs de Markov sur graphe, pour lesquelles l’initialisation est une étape cruciale. Au contraire, le résultat de l’algorithme MCMC à sauts réversibles ne dépend pas de l’initialisation. En pratique, ce n’est pas vraiment le cas puisque nous ne respectons pas la décroissance logarithmique de la température. De plus, la qualité de l’initialisation joue sur la vitesse de l’algorithme : plus l’initialisation est bonne, plus on peut se permettre de partir à une température faible et donc de proposer un algorithme plus rapide. Ici, nous ne nous sommes pas intéressés à cet aspect et l’initialisation de notre algorithme est réduite à la configuration nulle. Les deux résultats présentés en (b) et (c) ont été obtenus en 50 minutes pour le premier et en une heure pour le second avec un processeur de 1 GHz. Le résultat obtenu par [Géraud, 2003] sur une image originale de taille 2×10^6 pixels prend moins de 20 secondes avec un processeur de 1.7 GHz. Nous résumons les qualités et défauts des deux approches dans le tableau 3.5. Le principal défaut de notre algorithme est donc le temps de calcul nécessaire à l’optimisation. Nous donnons dans le prochain paragraphe quelques astuces pour réduire fortement ce temps de calcul.

	Processus de segments	Champ de Markov sur graphe
Temps de calcul	--	++
Exhaustivité	+	-
Fausses Alarmes	+	+
Connectivité	++	++
Précision	++	++

TAB. 3.5 – Performances relatives entre l’approche par champs de Markov sur Graphe proposée par Thierry Géraud et l’approche par processus Markov objet que nous proposons. (-) désigne de mauvaises performances et (+) de bonnes performances.

3.7.2 Évaluation de l’apport des pré-calculs

Dans ce paragraphe, nous présentons les résultats obtenus en utilisant des pré-calculs pour le terme d’attache aux données, donné par l’équation 3.47. Comme précédemment, l’optimisation est réalisée par un recuit simulé sur un algorithme MCMC à sauts réversibles. Nous avons tout d’abord utilisé le noyau décrit précédemment que nous noterons Q_1 . Puis, pour plus d’efficacité nous avons remplacé la “naissance et mort” uniforme (NMU) par une “naissance et mort” fondé sur les données (NMD). Nous noterons ce nouveau noyau Q_2 .

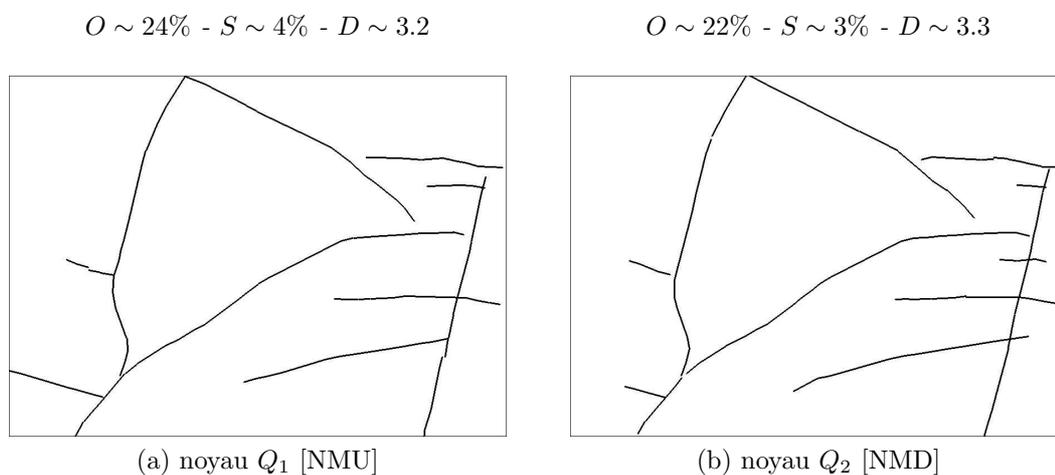


FIG. 3.25 – Résultats de l'extraction du réseau routier à partir de l'image 3 donnée dans la figure 3.19 avec pré-calcul des potentiels d'attache aux données. Le noyau Q_1 utilise un sous-noyau de type NMU qui est remplacé par un sous-noyau de type NMD dans le noyau Q_2 .

La figure 3.25 fournit les résultats de l'extraction à partir de l'image aérienne donnée dans la figure 3.19. Les résultats sont satisfaisants. Ils présentent moins de 5% de surdétections. Les routes principales sont toutes présentes, et de façon complète, dans le réseau détecté. Nous observons néanmoins un taux d'omissions important : plus de 20%. De plus, la précision est faible comparativement au résultat obtenu sans pré-calcul donné dans la figure 3.20. Comparativement, c'est donc le modèle sans pré-calcul qui fournit le meilleur résultat mais le temps de calcul nécessaire à l'obtention de ce résultat est relativement important : 23 minutes contre 19 minutes avec pré-calcul et le même noyau, et seulement 14 minutes quand nous ajoutons la naissance selon les cartes pré-calculées (*cf.* tableau 3.6). Ces résultats montrent qu'il est très intéressant d'utiliser le pré-calcul de l'attache aux données pour proposer de façon pertinente la naissance d'un segment puisque les résultats (a) et (b) de la figure 3.25 sont d'une qualité équivalente à la fois visuellement et quantitativement.

La figure 3.26 fournit les résultats de l'extraction avec pré-calcul sur l'image de forêt galerie (image 2). Ces réseaux obtenus sont de qualité équivalente à celle du résultat de l'extraction sans pré-calcul donné dans la figure 3.22 mais le temps de calcul est nettement plus faible dans le cas d'un pré-calcul (*cf.* tableau 3.7). Cela peut s'expliquer par le fait que la longueur moyenne d'un segment est très proche de la longueur minimale utilisée pour le pré-calcul. L'approximation est donc beaucoup moins grossière que dans le cas du réseau routier à haute résolution.

	Sans pré-calcul	Avec pré-calcul sous-noyau NMU	Avec pré-calcul sous-noyau NMD
Pré-calculs	-	8	8
Recuit simulé	23	11	6
Total	23	19	14

TAB. 3.6 – Temps de calcul de l'extraction du réseau routier (figure 3.19).

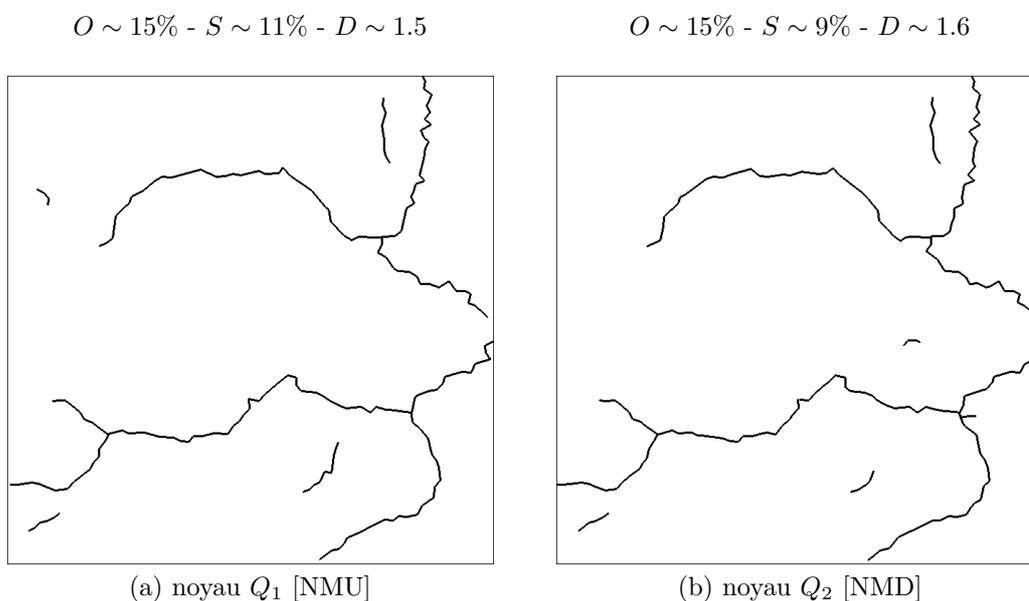


FIG. 3.26 – Résultats de l'extraction du réseau hydrographique à partir de l'image 2 donnée dans la figure 3.22 avec pré-calcul des potentiels d'attache aux données.

	Sans pré-calcul	Avec pré-calcul sous-noyau NMU	Avec pré-calcul sous-noyau NMD
Pré-calculs	-	3	3
Recuit simulé	40	11	5
Total	40	14	8

TAB. 3.7 – Temps de calcul de l'extraction du réseau hydrographique (figure 3.22).

3.7.3 Approche bayésienne

Ce paragraphe présente les résultats de l'extraction du réseau linéique par une approche bayésienne, dont l'objectif est d'extraire la configuration maximisant la densité *a posteriori* :

$$h_p(\mathbf{s}|Y) \propto h_p(\mathbf{s}) h_d(Y/S) \quad (3.83)$$

où $h_p(\mathbf{s})$ est la densité *a priori* donnée par l'équation (3.7) et $h_d(Y|S)$ est la vraisemblance de l'image filtrée Y donnée par l'équation (3.47), S étant la silhouette de \mathbf{s} sur la grille de pixels. Comme précédemment, l'optimisation se fait par un recuit simulé sur un algorithme MCMC à sauts réversibles.

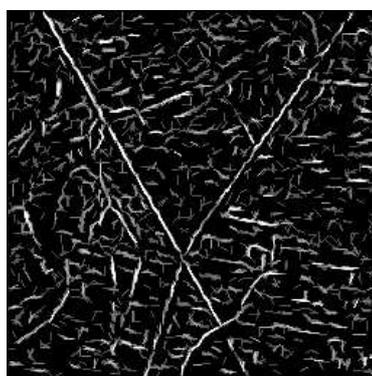
Ici, nous proposons d'ajouter, parallèlement au schéma de décroissance de température, un schéma de décroissance sur le poids de l'attache aux données ω_d . Soit $\omega_d^* = \frac{\omega_f}{n_{min}}$ le poids de l'attache aux données choisi selon la règle de calibrage donnée dans la table 3.2, qui a été établie pour obtenir une configuration finale présentant de bonnes propriétés. Soit $\omega_d(t)$ la valeur de ce poids à l'itération t . Nous commençons l'algorithme avec un poids $\omega_d(0)$ plus grand que ω_d^* et proposons une décroissance géométrique par plateau (avec une constante géométrique très proche de 1) de $\omega_d(t)$ jusqu'à l'itération t_f telle que $\omega_d(t_f) = \omega_d^*$. De cette façon, un grand nombre de segments libres, en adéquation avec les hypothèses radiométriques émises sur le réseau d'intérêt, peuvent être acceptés au début de l'algorithme, comme autant de points d'amorces ; alors que la contrainte sur l'absence de petits segments libres bien attachés peut être vérifiée à la fin de l'algorithme. Remarquons que cet algorithme n'est plus à proprement parler un recuit simulé permettant d'estimer le maximum *a posteriori*. C'est une façon de trouver une bonne initialisation pour faire un recuit à basse température t_f . Concernant le noyau de proposition de l'algorithme, nous utilisons le noyau le plus efficace construit jusqu'à présent, *i.e.* un noyau composé d'une "naissance et mort" fondée sur les données, une "naissance et mort" dans un voisinage par rapport à la relation de connexion, et un sous-noyau composé de perturbations de segments dépendant de l'état du segment à perturber.

Nous présentons tout d'abord deux résultats de cette approche sur des images moyenne résolution : le premier concerne l'extraction du réseau routier à partir d'une image SPOT Panchro de 10 mètres de résolution (figure 3.27) ; le second concerne l'extraction du réseau hydrographique à partir d'une image SPOT XS2 de 20 mètres de résolution (figure 3.28). Nous avons utilisé le même filtre de détection de structures linéaires pour ces deux images. Ce filtre correspond au filtre de détection proposé dans le paragraphe 3.2.2. Il repose sur le calcul de valeurs de confiance, traduisant la possibilité qu'un segment soit positionné sur une structure linéaire de l'image, pour des segments de taille minimale L_{min} , égale ici à 5, et de 8 orientations différentes. Le masque associé au segment est choisi d'épaisseur 1 (une seule bande interne), donc seul le contraste est évalué entre la bande interne et les deux bandes externes (d'épaisseur 3 pixels). Les deux seuils intervenant dans la mesure de confiance (ou mesure de contraste), donnée par l'équation 3.51, sont $\tau_1 = 2$ et $\tau_2 = 9$. L'image filtrée à été obtenue en moins de 10

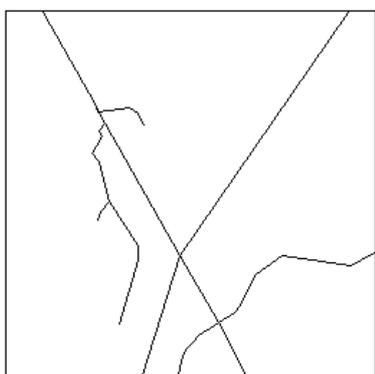
seconde pour l'image 1 de taille 255×255 , et un peu plus d'une minute pour l'image 2 de taille 682×674 avec un processeur de 2 GHz. Le filtre de détection fournit une image très bruitée : on voit apparaître des segments où les réponses ont été optimales. On voit ici que l'hypothèse d'un bruit non corrélé n'est pas valide. Or, grâce à l'utilisation de la méthode des boîtes qualitatives, les résultats de l'extraction sur les images filtrées sont de très bonne qualité, équivalente voire supérieure à la qualité des résultats obtenus par l'approche par champ externe. De plus, le temps de calcul est relativement faible : en moins de 5 minutes pour chacune des deux images avec un processeur 2 GHz.



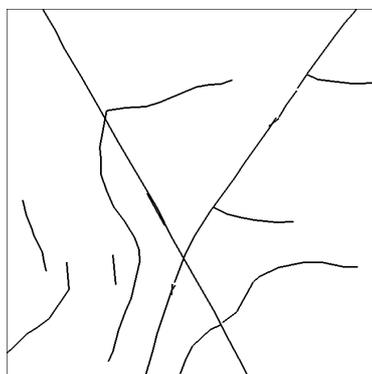
(a) Image 1 © CNES



(b) Image filtrée



(c) Extraction manuelle



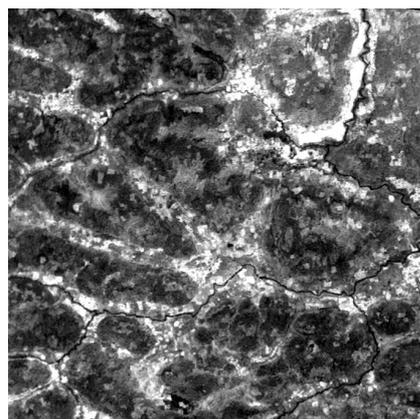
(d) Extraction automatique

$O \sim 3\%$ - $S \sim 33\%$ - $D \sim 1.0$ pixels

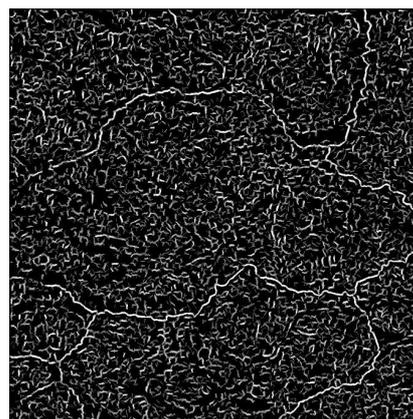
FIG. 3.27 – Résultat de l'extraction du réseau routier à partir d'une image SPOT de taille 255×255 et de résolution 10 mètres en utilisant le modèle bayésien.

Le résultat de l'extraction du réseau routier (figure 3.27) présente néanmoins 33% de surdétectés par rapport à l'extraction manuelle. Mais, visuellement, les surdétectés correspondent à des branches contrastées avec leur environnement. Ce ne sont donc

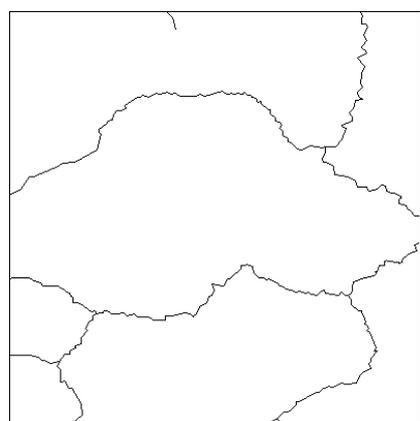
peut-être pas des surdétections par rapport à la scène observée. De plus, un simple post-traitement fondé sur une mesure de contraste permettrait de supprimer ces branches. Remarquons, de plus, qu'il n'y a que 3% d'omissions ce qui est légèrement inférieur au meilleur taux d'omissions (4%) obtenu par l'approche par champ externe (figure 3.18).



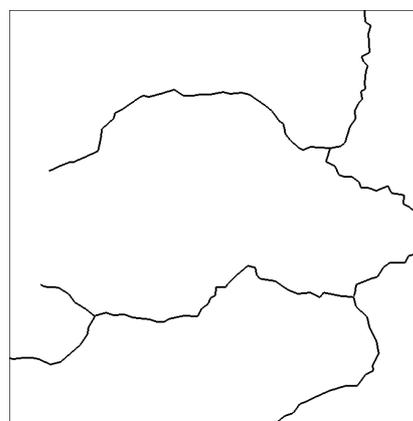
(a) Image 2 © BRGM



(b) Image filtrée



(c) Extraction manuelle



(d) Extraction automatique

$O \sim 11\%$ - $S \sim 0\%$ - $D \sim 1.7$ pixels

FIG. 3.28 – Résultat de l'extraction du réseau hydrographique à partir d'une image SPOT de taille 682×674 et de résolution 20 mètres en utilisant le modèle bayésien.

Concernant l'extraction des forêts galeries, l'approche bayésienne améliore les résultats de l'extraction déjà prometteurs par une approche par champs externe étant donné le faible contraste de l'image 2 et le caractère sinueux du réseau. En effet, les taux d'omissions et de surdétections obtenus par l'approche par champ externe sont respectivement de 14% et 10% sans pré-calculs (figure 3.22) et de 15% et 9% avec pré-calcul

(figure 3.26). Aucune fausse alarme et seulement 11% d'omissions sont observées dans le résultat obtenu par l'approche bayésienne (figure 3.28). On observe néanmoins une légère perte en précision.

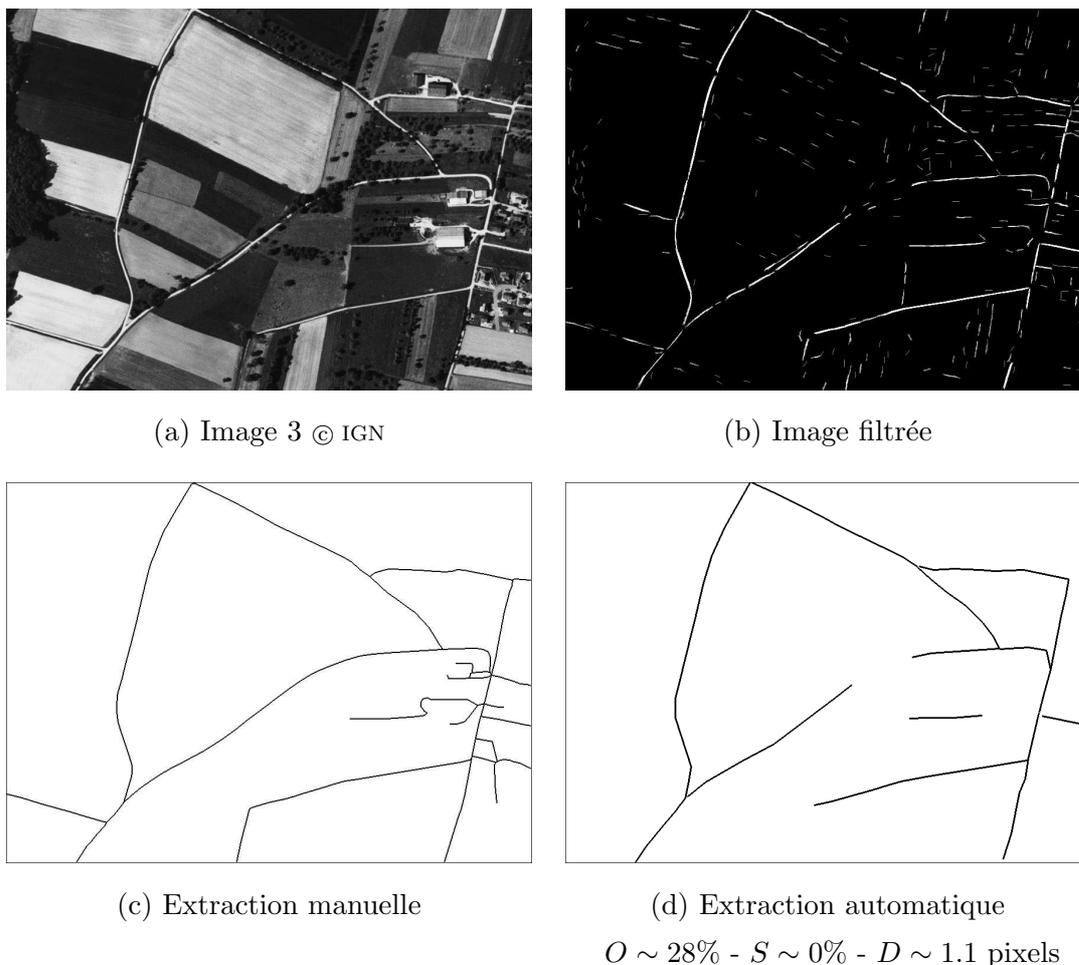


FIG. 3.29 – Résultat de l'extraction du réseau routier à partir d'une image aérienne de taille 1784×1304 pixels et de 50 centimètres de résolution en utilisant le modèle bayésien.

Bien que cette approche donne de bons résultats sur les deux précédents réseaux, il n'en est pas de même pour les réseaux dont un bruit géométrique perturbe fortement la détection. En effet, le modèle bayésien ne permet pas de combler de larges occlusions du réseau comme le résultat obtenu sur une image aérienne haute résolution présenté dans la figure 3.29. Ceci s'explique par le fait que plus une zone d'occlusion est grande, plus la probabilité d'accepter l'ajout d'un segment sur cette zone sera faible puisque le détecteur de structures linéaires a une réponse nulle dans les zones occlusions. Il est

néanmoins possible de combler ces occlusions en augmentant le poids de l'*a priori*, mais, dans ce cas, nous prolongeons également le réseau là où il n'y a rien. De ce point de vue, ce résultat est moins bon que ceux obtenus par champ externe. En effet, dans tous les résultats obtenus par champ externe nous avons pu combler les omissions dues à la présence d'arbres. Ceci s'explique par le fait que l'ajout d'un grand segment répondant mal aux données est aussi pénalisant que l'ajout d'un petit. C'est pourquoi nous pouvons, grâce à l'utilisation d'un jeu de paramètres énergétiques adéquate, détecter les longues occlusions. En contrepartie, dans le cas où nous détectons les longues occlusions, nous pouvons également relier des routes qui ne le devraient pas. Deux routes ont effectivement été reliées par un segment passant sur une maison dans les résultats donnés dans les figures 3.20 et 3.25. Finalement, le résultat obtenu en utilisant le modèle bayésien n'est pas à rejeter puisqu'il respecte plus les données dont nous disposons. Ainsi, le résultat présenté dans la figure 3.29 ne présente aucune fausse alarme et les jonctions du réseau y sont mieux détectées que dans les résultats obtenus par champ externe. Du point de vue de l'efficacité, la convergence est atteinte en 20 minutes sur un processeur 2 GHz, ce qui est supérieur au temps de calcul des résultats obtenus par champ externe.

Chapitre 4

Extension aux lignes brisées

Nous proposons dans ce chapitre une modélisation du réseau par une collection de lignes brisées décrites par un nombre inconnu de segments. Le modèle *a priori* est un processus de lignes brisées incorporant des contraintes sur la forme des lignes brisées et leur agencement dans la scène. Ainsi, les longues lignes brisées caractérisées par une faible courbure sont favorisées. De plus, la définition de deux interactions permet de pénaliser les lignes qui ne sont pas reliées au reste du réseau et de limiter les superpositions partielles de lignes brisées. Les propriétés radiométriques sont incorporées via la construction d'un terme d'attache aux données fondé sur une mesure du contraste du réseau avec son environnement proche et une mesure d'homogénéité entre deux sections successives du réseau. La complexité des objets du modèle rend nécessaire l'utilisation de perturbations appropriées dans l'algorithme d'échantillonnage. Ainsi, diverses perturbations sont proposées afin d'accélérer la convergence et sortir des maxima locaux. Les résultats expérimentaux montrent que la principale contribution de cette nouvelle modélisation est la possibilité de détecter de façon très précise les réseaux sinueux.

4.1 Motivations

Dans le chapitre 3, nous avons proposé une modélisation du réseau linéique par processus objet où les objets sont des segments interagissant entre eux et décrits par leur centre, leur orientation et leur longueur. La performance de cette modélisation a pu être vérifiée sur de nombreux exemples, notamment pour l'extraction du réseau routier. Néanmoins, nous avons pu observer :

- une détection imprécise des branches sinueuses (figures 3.22, 3.26, et 3.28) ;
- une détection difficile au niveau des intersections en "Y" (figure 3.20).

Nous proposons ici d'étendre cette modélisation à des objets plus complexes, comme cela est fait par [Rue et Hurn, 1999] pour un problème de reconnaissance de cellules dans les images biologiques. [Rue et Hurn, 1999] proposent une modélisation par un processus objet, le nombre de cellules étant inconnu, où les objets sont modélisés par des modèles déformables de résolution variable, qui correspondent à des polygones à n côtés, où n est inconnu.

Dans ce chapitre, nous proposons un nouveau processus objet pour l'extraction du réseau linéique (axe central des routes ou des rivières) où les objets sont des lignes brisées composées d'un nombre inconnu de segments. Ainsi, la connexion entre segments est directement incorporée dans la définition des objets du processus. Celle-ci pourra donc être exacte contrairement à la connexion définie dans les modèles précédents. Nous espérons donc gagner en continuité. De plus, les jonctions entre les routes ou les rivières peuvent être modélisées de façon simple via la définition d'une interaction entre deux objets du processus. Enfin, nous espérons pouvoir détecter avec plus de précisions les réseaux sinueux via la définition d'un terme d'attache aux données adapté.

4.2 Processus de référence

Le modèle que nous proposons dans ce chapitre est un processus objet dans lequel les objets sont des lignes brisées composées d'un nombre inconnu de segments. Chaque ligne brisée est décrite par les variables suivantes :

- son point initial $p^1 = (x, y) \in F \subset \mathbb{R}^2$;
- son épaisseur $e \in [e_{min}, e_{max}]$;
- un nombre de segments $n \in \{1, \dots, n_{max}\}$;
- les longueurs des segments $l_j \in [L_{min}, L_{max}]$, $j = 1, \dots, n$;
- les directions des segments $\alpha_j \in]-\pi, \pi]$, $j = 1, \dots, n$.

Un exemple pour $n = 3$ est donné dans la figure 4.1.

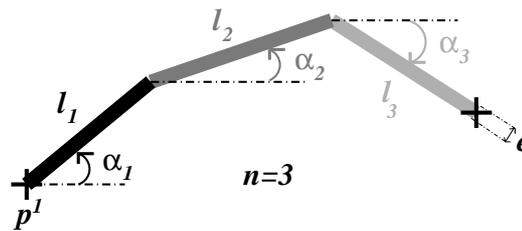


FIG. 4.1 – Objet type du processus

Remarque 14 Dans le cas des réseaux routiers, caractérisés par des branches de très faible courbure, on pourra considérer une direction initiale α_1 et les différences des directions successives $\{\Delta\alpha_j\}_{j=1, \dots, n-1}$ comprises dans $]-\pi/2, \pi/2]$, à la place des directions successives $\{\alpha_j\}_{j=1, \dots, n}$ des segments composant la ligne brisée. Un tel choix permettra de gagner en rapidité.

Comme pour les processus de segments, nous choisissons de considérer le processus de référence le plus simple possible : un **processus de Poisson homogène**. Sous la loi de ce processus, le nombre de lignes brisées suit une loi de Poisson d'espérance $\lambda|F|$ et les paramètres sont uniformément distribués dans leur espace d'état respectif. La mesure μ associée à ce processus est donnée par l'équation (2.5), avec une mesure d'intensité $\nu = \Lambda \otimes P_M$ où :

- Λ est la mesure de Lebesgue multipliée par λ ;
- P_M est la mesure de probabilité associée aux marques et dont l'expression est donnée par l'équation suivante :

$$P_M(B) = \sum_{n=1}^{n_{max}} \frac{1}{n_{max}} \int_{[e_{min}, e_{max}]} \int_{V^n} \mathbf{1}_B(e, v_1, \dots, v_n) \frac{d^n(v_1, \dots, v_n) de}{|V|^n (e_{max} - e_{min})} \quad (4.1)$$

où B est un ensemble de la tribu associée à l'espace des marques $M = [e_{min}, e_{max}] \times \bigcup_{n=1}^{n_{max}} V^n$, où $V = [L_{min}, L_{max}] \times]-\pi, \pi[$ et $V^n = V \times \dots \times V$ (n fois).

Pour introduire un *a priori* sur la forme des lignes brisées et les interactions entre lignes brisées, nous spécifions un processus par une densité h_p par rapport à cette mesure μ . L'expression de cette densité est donnée dans le paragraphe 4.3 après une description des connaissances *a priori* sur la forme d'une ligne brisée injectées dans le modèle et des interactions possibles entre les lignes brisées.

4.3 Modèle a priori

4.3.1 A priori sur la forme d'une ligne brisée

Nous introduisons tout d'abord un *a priori* sur le nombre de segments d'une ligne brisée. Ceci se fait au travers d'un potentiel U_{11} qui est d'autant plus faible que le nombre de segments n augmente :

$$U_{11}(n) = \frac{M_n}{(n+1)^2} \quad (4.2)$$

où M_n est le poids (constant) associé à ce potentiel. L'utilisation de ce potentiel entraîne une **pénalisation** des branches formées d'un **petit nombre de segments**.

Pour favoriser les **longs segments**, nous introduisons un potentiel U_{12} sur la longueur l d'un segment défini comme suit :

$$U_{12}(l) = M_l \frac{L_{max} - l}{L_{max} - L_{min}} \quad (4.3)$$

où M_l est le poids (constant) associé à ce potentiel.

Pour favoriser la **faible courbure** des branches composant le réseau, nous introduisons le potentiel U_{13} sur la courbure des lignes brisées au travers d'un potentiel U_{13} sur les différences entre les directions successives d'une ligne brisée :

$$U_{13}(\alpha_j, \alpha_{j+1}) = M_\alpha \left(\frac{1}{2} - \cos(\alpha_{j+1} - \alpha_j) \right) \quad (4.4)$$

où M_α est le poids (constant) associé à ce potentiel.

Enfin, nous introduisons un potentiel, de type “**hard core**”, qui interdit l'**intersection d'une ligne brisée avec elle-même**.

L'énergie *a priori* sur la forme d'une ligne brisée c composée de n segments est alors définie comme suit :

$$U_1(c) = \begin{cases} +\infty, & \text{si deux segments de } c \text{ s'intersectent} \\ U_{11}(n) + \sum_{j=1}^n U_{12}(l_j) + \sum_{j=1}^{n-1} U_{13}(\alpha_j, \alpha_{j+1}), & \text{sinon} \end{cases} \quad (4.5)$$

4.3.2 Interactions entre les lignes brisées

Nous considérons deux types d'interactions.

La première interaction est fondée sur une relation de **proximité** \sim_p entre lignes brisées. Cette relation se définit comme suit :

$$u \sim_p v \Leftrightarrow \left(\begin{array}{l} \exists j \in \{1, \dots, n_u\} : d(p_u^j, v) < d_{max} \text{ et } d(p_u^{j+1}, v) < d_{max} \\ \text{ou } \exists j \in \{1, \dots, n_v\} : d(p_v^j, u) < d_{max} \text{ et } d(p_v^{j+1}, u) < d_{max} \end{array} \right) \quad (4.6)$$

où d correspond à la distance euclidienne, n_u correspond au nombre de segments composant la ligne brisée u , et p_u^j désigne le point de contrôle numéro j décrivant u . Autrement dit deux lignes sont dites proches, si deux points consécutifs de l'une des deux lignes sont à une distance inférieure à d_{max} de l'autre ligne. Un exemple de deux lignes vérifiant cette relation est donnée dans la figure 4.2. Cette interaction est interdite par l'assignation d'un potentiel de type “**hard core**” aux lignes vérifiant cette relation. Ceci permet d'éviter un recouvrement partiel de lignes brisées.

La seconde interaction correspond à la **connexion** d'une ligne brisée, via une de ses extrémités, à une deuxième ligne brisée (pas forcément au niveau des extrémités de la deuxième ligne). Elle est fondée sur la distance euclidienne entre une extrémité p_u^k ($k = 1$ ou $n+1$) d'une ligne brisée u et une ligne brisée v appartenant à l'ensemble $\{\mathbf{c} \setminus u \cup \Gamma(F)\}$, composé des lignes de la configuration \mathbf{c} (autres que u) et de la ligne brisée, notée $\Gamma(F)$, correspondant au contour de la fenêtre d'observation F . Si la distance $d(p_u^k, v)$ entre p_u^k et v est inférieure à un seuil ϵ , alors u est dite connectée à v par son extrémité p_u^k . Soit $V_{\mathbf{c},F}(p_u^k)$ l'ensemble des lignes brisées voisines de u via une connexion par p_u^k :

$$V_{\mathbf{c},F}(p_u^k) = \{v \in \{\mathbf{c} \setminus u \cup \Gamma(F)\} : d(p_u^k, v) < \epsilon\}$$

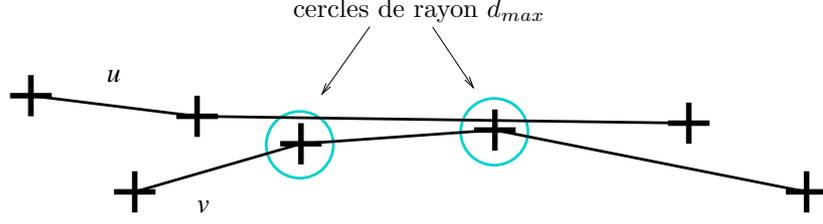


FIG. 4.2 – Relation de proximité entre deux lignes brisées. $u \sim_p v$ car deux points de contrôle consécutifs de la ligne brisée v sont à une distance inférieure à d_{max} de u .

On définit trois états d'une ligne brisée u selon la cardinalité des deux ensembles $V_{\mathbf{c},F}(p_u^1)$ et $V_{\mathbf{c},F}(p_u^{n+1})$. Une ligne brisée u est dite :

- **libre** si u n'est connectée par aucune de ses deux extrémités, *i.e.* :

$$V_{\mathbf{c},F}(u) = V_{\mathbf{c},F}(p_u^1) \cup V_{\mathbf{c},F}(p_u^{n+1}) = \emptyset$$

- **simple** si u est connectée par une seule de ses extrémités, *i.e.* :

$$V_{\mathbf{c},F}(u) \neq \emptyset, \quad \exists k : V_{\mathbf{c},F}(p_u^k) = \emptyset$$

- **double** si u est connectée par ses deux extrémités, *i.e.* :

$$V_{\mathbf{c},F}(p_u^1) \neq \emptyset, \quad V_{\mathbf{c},F}(p_u^{n+1}) \neq \emptyset$$

Ces trois états sont illustrés dans la figure 4.3. Nous proposons alors de gérer la connectivité du réseau au travers d'une **pénalisation des lignes brisées libres et simples** par des potentiels constants ω_f et ω_s :

$$U_{21}(u|V_{\mathbf{c},F}(u)) = \begin{cases} \omega_f, & \text{si } V_{\mathbf{c},F}(u) = \emptyset \\ \omega_s, & \text{si } V_{\mathbf{c},F}(u) \neq \emptyset, \quad \exists k : V_{\mathbf{c},F}(p_u^k) = \emptyset \\ 0, & \text{si } V_{\mathbf{c},F}(p_u^1) \neq \emptyset, \quad V_{\mathbf{c},F}(p_u^{n+1}) \neq \emptyset \end{cases} \quad (4.7)$$

Remarque 15 *Le fait de considérer le contour de la fenêtre F comme une ligne brisée à laquelle les lignes de la configuration courante peuvent être connectées revient à considérer qu'une extrémité est forcément connectée à une partie du réseau située à l'extérieur de F . Une ligne "sortante" pourra donc être considérée comme connectée via une extrémité proche de $\Gamma(F)$ et ne sera donc pas pénalisée outre mesure.*

Nous incorporons également au modèle un potentiel mesurant la qualité des connexions. La mesure de qualité est fondée sur la distance de connexion. Plus celle-ci sera faible,

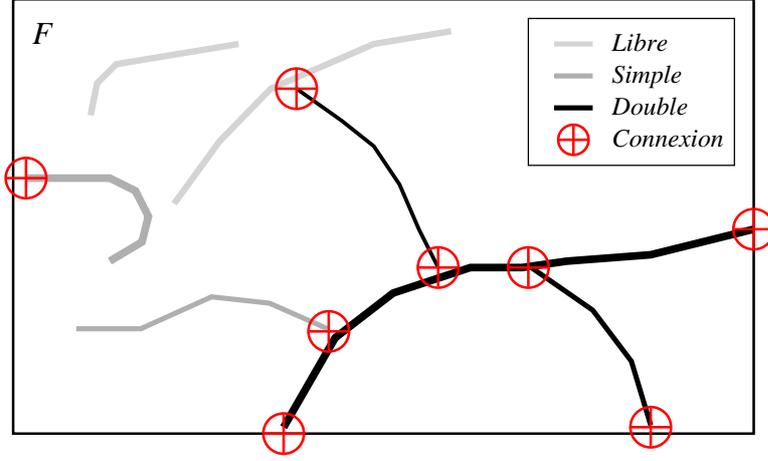


FIG. 4.3 – États des lignes brisées par rapport aux interactions de connexion.

plus la qualité sera proche de 1. Soit u une ligne brisée connectée à v par p_u^k . La qualité de la connexion correspondante $\langle u, v \rangle_{p_u^k}$ est donnée par :

$$\sigma(\langle u, v \rangle_{p_u^k}) = \frac{1}{\epsilon^2} \left(\frac{1 + \epsilon^2}{1 + d^2(p_u^k, v)} - 1 \right) \quad (4.8)$$

où ϵ est le seuil de connexion (il y a connexion si $d(p_u^k, v) < \epsilon$). Nous associons alors, à chaque ligne brisée u de la configuration, un potentiel fondé sur les qualités des connexions engendrées par ses deux extrémités :

$$U_{22}(u|V_{\mathbf{c}, F}(u)) = - \sum_{\substack{k=1, n+1 \\ v: d(p_u^k, v) < \epsilon}} \sigma(\langle u, v \rangle_{p_u^k}) \quad (4.9)$$

Les **distances faibles de connexion** sont ainsi favorisées par un potentiel négatif.

L'énergie *a priori* sur les interactions d'un objet u , avec les autres lignes brisées de la configuration \mathbf{c} ou le contour de la fenêtre $\Gamma(F)$, est alors définie comme suit :

$$U_2(u|\{\mathbf{c} \setminus u\}) = \begin{cases} +\infty, & \text{si } \exists v \in \mathbf{c} : u \sim_p v \\ U_{21}(u|V_{\mathbf{c}, F}(u)) + U_{22}(u|V_{\mathbf{c}, F}(u)) & \end{cases} \quad (4.10)$$

où U_{21} et U_{22} sont les potentiels de connexion donnés par les équations (4.7) et (4.9).

4.3.3 Densité a priori

Finalement, la densité *a priori* du processus de lignes brisées peut s'écrire sous la forme suivante :

$$h_p(\mathbf{c}) \propto \exp \left[- \sum_{i=1}^N U_1(c_i) + U_2(c_i | \{\mathbf{c} \setminus c_i\}) \right] \quad (4.11)$$

où $\mathbf{c} = \{c_1, \dots, c_N\}$ est une configuration composée de N lignes brisées, U_1 est le terme énergétique injectant un *a priori* sur la forme de chaque ligne brisée (équation (4.5)), et U_2 est le terme énergétique injectant un *a priori* sur l'agencement des lignes brisées dans la scène observée (équation (4.10)). Cette densité spécifie un processus de lignes brisées **localement stable** du fait de l'utilisation du potentiel de type "hard core" assigné aux paires de lignes vérifiant la relation de proximité. En effet, l'utilisation de ce potentiel induit (presque sûrement) un pavage maximal de tout borné F par un ensemble de lignes brisées : si une ligne est ajoutée à cet ensemble alors la relation de proximité sera vérifiée et la densité s'annulera. Il existe donc un seuil B tel que l'ajout d'une ligne brisée ne puisse entraîner plus de B connexions : la diminution de U_2 est donc bornée. Le terme U_1 étant borné inférieurement, la densité de Papangelou $h(\mathbf{c} \cup c)/h(\mathbf{c})$ est donc bornée.

4.4 Terme d'attache aux données

Nous proposons une approche par champ externe pour incorporer les propriétés radiométriques des données : l'image est utilisée en tant que champ externe permettant l'évaluation de la qualité de chaque ligne brisée de la configuration \mathbf{c} compte tenu des deux hypothèses usuellement faites en détection de réseau :

- (H_1) le niveau de gris du réseau est localement homogène ;
- (H_2) le réseau contraste fortement avec son environnement.

Pour le calcul du terme d'attache aux données, nous associons à chaque ligne brisée u un masque de pixels composé d'un ensemble de pixels V correspondant à la projection de u dans l'image et de deux ensembles de pixels R_1 et R_2 correspondant au fond proche de u dans l'image. Ces deux régions sont positionnées à une distance fixée d de V pour tolérer une légère variation de la largeur d'une branche donnée. Rappelons que, bien que l'épaisseur des lignes soit désormais une marque, *i.e* un paramètre variable, elle est supposée constante pour une ligne brisée donnée. Chaque masque est ensuite divisé en sections composées d'un nombre de pixels fixé, comme cela est illustré par la figure 4.4.

La vérification des hypothèses H_1 et H_2 se fait alors par la vérification des deux hypothèses suivantes :

- (H'_1) la variation de niveau de gris entre deux sections internes successives V^j et V^{j+1} est faible ;
- (H'_2) la variation de niveau de gris entre V_j et les régions externes correspondantes, R_1^j et R_2^j , est importante.

L'adéquation avec l'hypothèse d'homogénéité est évaluée pour chaque couple de sections internes successives $\{V^j, V^{j+1}\}$ par le calcul de la valeur t_h^j du t-test de Student

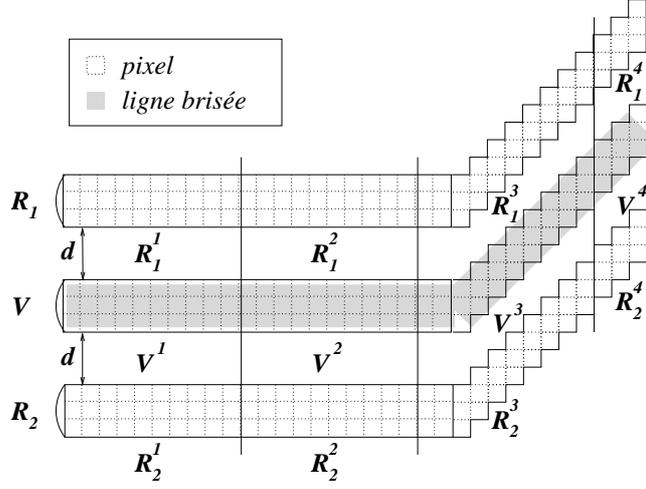


FIG. 4.4 – Masque de pixels associé à une ligne brisée.

(équation 3.40) entre V^j et V^{j+1} . On procède ensuite à un seuillage empirique de t_h^j entre 1 et τ_h , suivi d'une transformation linéaire de $[1, \tau_h]$ vers $[-1, 1]$ pour obtenir le potentiel $U_h(j, j+1)$, d'autant plus faible que $\{V^j, V^{j+1}\}$ est en adéquation avec H'_1 :

$$U_h(j, j+1) = \begin{cases} -1 & \text{si } t_h^j < 1 \\ 1 - 2 \frac{\tau_h - t_h^j}{\tau_h - 1} & \text{si } 1 \leq t_h^j \leq \tau_h \\ 1 & \text{si } t_h^j > \tau_h \end{cases} \quad (4.12)$$

L'adéquation avec l'hypothèse de contraste est évaluée pour chaque section $M^j = \{V^j, R_1^j, R_2^j\}$. La valeur statistique t_c^j associée à M^j est le minimum des deux valeurs du t-test de Student entre la section interne V^j et les deux sections externes R_1^j et R_2^j . On procède ensuite à un seuillage de t_c^j entre τ_1 et τ_2 , suivi d'une transformation linéaire de $[\tau_1, \tau_2]$ vers $[-1, 1]$ pour obtenir le potentiel $U_c(j)$, d'autant plus faible que M^j est en adéquation avec H'_2 :

$$U_c(j) = \begin{cases} 1 & \text{si } t_c^j < \tau_1 \\ 1 - 2 \frac{t_c^j - \tau_1}{\tau_2 - \tau_1} & \text{si } \tau_1 \leq t_c^j \leq \tau_2 \\ -1 & \text{si } t_c^j > \tau_2 \end{cases} \quad (4.13)$$

Finalement, le potentiel d'attache aux données associé à une ligne $u \in \mathbf{c}$ est défini comme suit :

$$U_d(u) = p_h \sum_{j=1}^{J-1} U_h(j, j+1) + p_c \sum_{j=1}^J U_c(j) \quad (4.14)$$

où I est le nombre de sections composant le masque de pixels associé à la ligne brisée u , p_h et p_c sont des poids positifs respectivement associés aux potentiels U_h et U_c

d’homogénéité et de contraste. Notons que p_c devra être nettement supérieur à p_h pour ne pas détecter les routes dans les zones très homogènes. L’énergie d’attache aux données est alors définie comme la somme des potentiels U_d sur chaque ligne brisée appartenant à \mathbf{c} . Le terme d’attache aux données h_d est donc donné par :

$$h_d(\mathbf{c}) \propto \exp\left(-\sum_{c_i \in \mathbf{c}} U_d(c_i)\right) \quad (4.15)$$

où U_d est donnée par l’équation (4.14).

4.5 Échantillonnage approprié

Nous proposons d’utiliser un algorithme de type Metropolis-Hastings pour réaliser l’échantillonnage de ce processus. La complexité des objets mis en jeu rend la définition de perturbations pertinentes nécessaire. En effet, on comprend bien qu’une naissance uniforme d’une ligne brisée ne pourra en aucun cas permettre la génération de branches complètes et bien positionnées. L’utilisation de ce seul noyau ne pourra donc être utilisé que dans le cas d’une température élevée (densité non piquée). C’est pourquoi nous avons défini une “naissance et mort” de ligne brisée ne contenant qu’un segment. Lorsque nous disposons d’informations de localisation pré-calculées, nous proposons une naissance fondée sur les données pour proposer des points d’amorces correctement positionnés. Nous combinons ce sous-noyau à un sous-noyau permettant le prolongement de ces points d’amorces qui consiste en l’ajout ou la suppression de segments au début ou à la fin de la ligne brisée. Nous utilisons également des petites perturbations de ligne brisée très utiles quand une ligne brisée est approximativement bien positionnée : modification de l’épaisseur d’une ligne brisée ; perturbation symétrique d’un point de contrôle ; “fusion et division” de segments. Enfin, nous utilisons deux sous-noyaux permettant la fusion de lignes brisées, ce qui permet de sortir aisément d’optima locaux où deux lignes brisées sont positionnées sur la même branche du réseau réel. Tous ces mouvements sont décrits dans les paragraphes suivants. Pour chaque mouvement, un taux de Green est calculé de la même façon que cela est décrit en détail dans le paragraphe 2.5.2 pour la naissance et mort uniforme d’un point de façon à assurer la réversibilité de la chaîne ainsi construite. Comme pour le modèle “Quality Candy”, l’apériodicité et l’irréductibilité suffisent donc à montrer l’ergodicité de la chaîne ainsi construite. En effet, on peut montrer qu’il existe un nombre maximal N_M de lignes brisées positionnées dans la fenêtre F tel qu’aucune interaction de type “hard core” ne soit vérifiée. Toute configuration \mathbf{c} telle que $N(\mathbf{s}) > N_M$ sera donc de densité nulle. La chaîne est donc construite sur un espace petit et il suffit donc de vérifier la réversibilité, l’apériodicité et l’irréductibilité pour converger ergodiquement vers la mesure d’intérêt.

4.5.1 Naissance et mort uniforme

La “naissance et mort” uniforme (NMU) d’une ligne brisée consiste à proposer un ajout uniforme d’une ligne brisée dans $F \times M$ ou un retrait uniforme d’une ligne brisée

dans la configuration courante \mathbf{c} . La naissance uniforme consiste à :

- générer un point initial uniformément dans F ;
- générer une épaisseur e dans $[e_{min}, e_{max}]$;
- choisir uniformément un nombre de segments n entre 1 et n_{max} ;
- et générer ensuite les paramètres décrivant chaque segment (longueur et direction), uniformément dans V^n .

Les taux de Green associés à une mort et une naissance sont alors les mêmes que ceux associés à une NMU pour l'échantillonnage d'un processus de segments sur F :

$$R(\mathbf{c}, \mathbf{c} \cup u) = \frac{p_d}{p_b} \frac{\lambda|F|}{N(\mathbf{c}) + 1} \frac{h(\mathbf{c} \cup u)}{h(\mathbf{c})} \quad (4.16)$$

$$R(\mathbf{c}, \mathbf{c} \setminus u) = \frac{p_b}{p_d} \frac{N(\mathbf{c})}{\lambda|F|} \frac{h(\mathbf{c} \setminus u)}{h(\mathbf{c})} \quad (4.17)$$

où h est la densité non normalisée du processus par rapport à μ , la mesure du processus de Poisson uniforme d'intensité λ .

Ce noyau ne sera utilisé qu'avec une probabilité très faible ou ne sera pas utilisé du tout. En effet, ce noyau peut être remplacé par l'utilisation des deux noyaux présentés dans les paragraphes 4.5.2 et 4.5.3 et qui permettent, par leur combinaison, d'assurer l'irréductibilité de la chaîne.

4.5.2 Naissance et mort de lignes brisées réduites à un segment

Le deuxième noyau de type naissance et mort correspond à la proposition d'une naissance d'une ligne brisée n'ayant qu'un seul segment avec la probabilité p_{b1s} et la proposition d'une mort d'une ligne brisée ne contenant qu'un seul segment avec la probabilité $p_{d1s} = 1 - p_{b1s}$. La proposition de mort ne sera donc effectuée que si la configuration courante contient des lignes contenant exactement un segment.

Dans le cas de l'échantillonnage du processus sans information radiométrique, la naissance correspond à une naissance uniforme : le segment proposé est généré uniformément dans le compact $Z = F \times [e_{min}, e_{max}] \times [L_{min}, L_{max}] \times]-\pi, \pi[$. La mort est également uniforme : un segment est tiré uniformément dans l'ensemble $E_1(\mathbf{c})$ des lignes brisées présentes dans la configuration courante \mathbf{c} composées d'un unique segment. Dans le cas d'une naissance d'une ligne composée d'un segment, le taux de Green est alors donné par :

$$R(\mathbf{c}, \mathbf{c} \cup u) = \frac{p_{d1s}}{p_{b1s}} \frac{\lambda|F|}{n_{max} (\#(E_1(\mathbf{c})) + 1)} \frac{h(\mathbf{c} \cup u)}{h(\mathbf{c})} \quad (4.18)$$

où $\#(E_1(\mathbf{c}))$ désigne le nombre de lignes brisées composées d'un unique segment dans la configuration \mathbf{c} . De même, dans le cas d'une mort de $u \in E_1(\mathbf{c})$:

$$R(\mathbf{c}, \mathbf{c} \setminus u) = \frac{p_{b1s}}{p_{d1s}} \frac{n_{max} \#(E_1(\mathbf{c}))}{\lambda|F|} \frac{h(\mathbf{c} \setminus u)}{h(\mathbf{c})} \quad (4.19)$$

Dans le cas où l'on dispose de cartes de probabilités de présence de structures linéaires en chaque pixel, nous proposons la naissance de nouvelles lignes composées d'un seul segment en fonction d'une mesure non homogène fondée sur ces cartes pré-calculées de la même façon que cela a été proposé dans le paragraphe 3.2.3. Soit C_1, \dots, C_{N_θ} les cartes de probabilités associées à chaque orientation $\tilde{\theta}_k \in [0, \pi[$, $k = 1, \dots, N_\theta$. La procédure de proposition d'ajout d'une nouvelle ligne brisée u selon les cartes C_1, \dots, C_{N_θ} est la suivante :

- l'épaisseur est générée dans $[e_{min}, e_{max}]$;
- la longueur l et la direction α du premier (et unique) segment sont générées uniformément dans $[L_{min}, L_{max}] \times]-\pi, \pi]$;
- un pixel i est alors tiré selon la carte C_{k_α} , correspondant à l'orientation $\tilde{\theta}_{k_\alpha}$ la plus proche de α [π] : $k_\alpha = \arg \min_j [|\alpha - \tilde{\theta}_j|]$;
- le point initial p de u est alors tiré uniformément dans le carré de F correspondant au pixel i de l'image.

Le mouvement inverse consiste à retirer uniformément une ligne brisée dans $E_1(\mathbf{c})$.

Les taux de Green respectivement associés à une naissance et une mort d'une ligne brisée $u = (p, e, l, \alpha)$ sont :

$$R(\mathbf{c}, \mathbf{c} \cup u) = \frac{p_{d1s}}{p_{b1s}} \frac{\lambda |F|}{N_{pix} C_{k_\alpha}(i_p) n_{max} (\#(E_1(\mathbf{c})) + 1)} \frac{h(\mathbf{c} \cup u)}{h(\mathbf{c})} \quad (4.20)$$

$$R(\mathbf{c}, \mathbf{c} \setminus u) = \frac{p_{b1s}}{p_{d1s}} \frac{N_{pix} C_{k_\alpha}(i_p) n_{max} \#(E_1(\mathbf{c}))}{\lambda |F|} \frac{h(\mathbf{c} \setminus u)}{h(\mathbf{c})} \quad (4.21)$$

où N_{pix} est le nombre de pixels dans la grille, i_p correspond au pixel de la grille où le point p se projette.

4.5.3 Ajout et retrait d'un segment

Le mouvement "ajout et retrait" d'un segment consiste à proposer l'ajout ou la suppression d'un segment en fin ou en début de ligne. Premièrement, une ligne u est choisie selon une loi uniforme sur la configuration courante. Deuxièmement, le choix du type de mouvement est fait selon une probabilité dépendant du nombre de segment composant c . Si la ligne est composée d'un seul segment, alors seul l'ajout d'un segment supplémentaire sera proposé. De même, si elle est composée d'un nombre maximal de segments n_{max} , seul le retrait d'un segment sera proposé. Dans les autres cas, le retrait et l'ajout d'un segment sont proposés avec une probabilité uniforme. Une fois le mouvement choisi, une extrémité de ligne est choisie avec la probabilité 1/2. Dans le cas d'un ajout, une longueur l et une direction α sont tirées uniformément dans $V = [L_{min}, L_{max}] \times]-\pi, \pi]$. Soit $u = (p, e, l_1, \alpha_1, \dots, l_n, \alpha_n)$ la ligne brisée choisie. Si l'extrémité choisie est le point initial $p = (x, y)$, alors la ligne perturbée est donnée par $u' = (p', e, l, \alpha, l_1, \alpha_1, \dots, l_n, \alpha_n)$ où le nouveau point initial est $p' = (x - l \cos(\alpha), y - l \sin(\alpha))$. Si l'extrémité choisie est le point final, la ligne perturbée est donnée par $u' = (p, e, l_1, \alpha_1, \dots, l_n, \alpha_n, l, \alpha)$. Dans le cas d'un retrait, si l'extrémité choisie est le point initial, alors la ligne perturbée est donnée par $u' = (p', e, l_2, \alpha_2, \dots, l_n, \alpha_n)$

où $p' = (x + l_1 \cos(\alpha_1), y + l_1 \sin(\alpha_1))$ et, sinon, $u' = (p, e, l_1, \alpha_1, \dots, l_{n-1}, \alpha_{n-1})$.

Dans les cas où la ligne brisée choisie est composée d'un nombre de segments compris entre 2 et $n_{max} - 1$, le taux de Green est donné par le rapport des densités :

$$R(\mathbf{c}, \mathbf{c}') = \frac{h(\mathbf{c}')}{h(\mathbf{c})} \quad (4.22)$$

où \mathbf{c}' est la configuration obtenue par la perturbation de la configuration courante \mathbf{c} .

Dans le cas d'un ajout d'un segment à une ligne composée d'un seul segment et comme celui d'un retrait d'un segment à une ligne composée de n_{max} segments, le taux de Green est donné par :

$$R(\mathbf{c}, \mathbf{c}') = \frac{h(\mathbf{c}')}{2 h(\mathbf{c})} \quad (4.23)$$

Le facteur $1/2$ intervient car l'ajout d'un segment à une ligne brisée contenant 1 segment (resp. le retrait d'un segment d'une ligne brisée contenant n_{max} segments) se fait avec la probabilité 1 et que le mouvement réciproque, *i.e.* le retrait d'un segment d'une ligne contenant 2 segments (resp. l'ajout d'un segment à une ligne brisée contenant $n_{max} - 1$ segments) est choisi avec la probabilité $1/2$.

La combinaison de ce noyau avec un des deux noyaux de type "mort et naissance" de lignes brisées composées d'un seul segment décrits précédemment permet d'assurer l'irréductibilité de la chaîne de Markov générée par l'algorithme MCMC. Cette combinaison permet ainsi de remplacer le noyau NMU qui, dans un cadre d'optimisation, n'est pas pertinent : une naissance de ligne brisée composée de plusieurs segments est alors majoritairement rejetée.

Soit q_1 le noyau de type "naissance et mort" de lignes composées d'un seul segment sans information radiométrique et q_2 le noyau "ajout et retrait" d'un segment. Vérifions le bon fonctionnement de l'algorithme MCMC utilisant le noyau $Q = 1/2q_1 + 1/2q_2$ pour l'échantillonnage d'un processus de Poisson uniforme de lignes brisées composées d'un nombre aléatoire de segments compris entre 1 et 10. Nous posons $\lambda|F| = 100$. L'algorithme MCMC, initialisé par la configuration nulle, est itéré jusqu'à un critère d'arrêt correspondant à la détection de convergence des moyennes empiriques de quantités mesurables sur la configuration courante, comme cela est décrit dans le paragraphe 2.5.6). Les quantités choisies ici sont N , le nombre total de lignes brisées dans la configuration courante, et $\{N_i\}_{i=1..10}$ le nombre de lignes brisées contenant i segments. Les moyennes sont calculées à partir d'un nombre fixé I_0 d'itérations (ici, $I_0 = 30000$). A partir de ce nombre I_0 , nous prenons 1 échantillon toutes les P itérations (ici, $P = 5000$). Ceci permet de réduire l'effet de la forte corrélation entre échantillons proches. La convergence est considérée atteinte si la différence des valeurs empiriques est suffisamment faible sur 5 étapes successives.

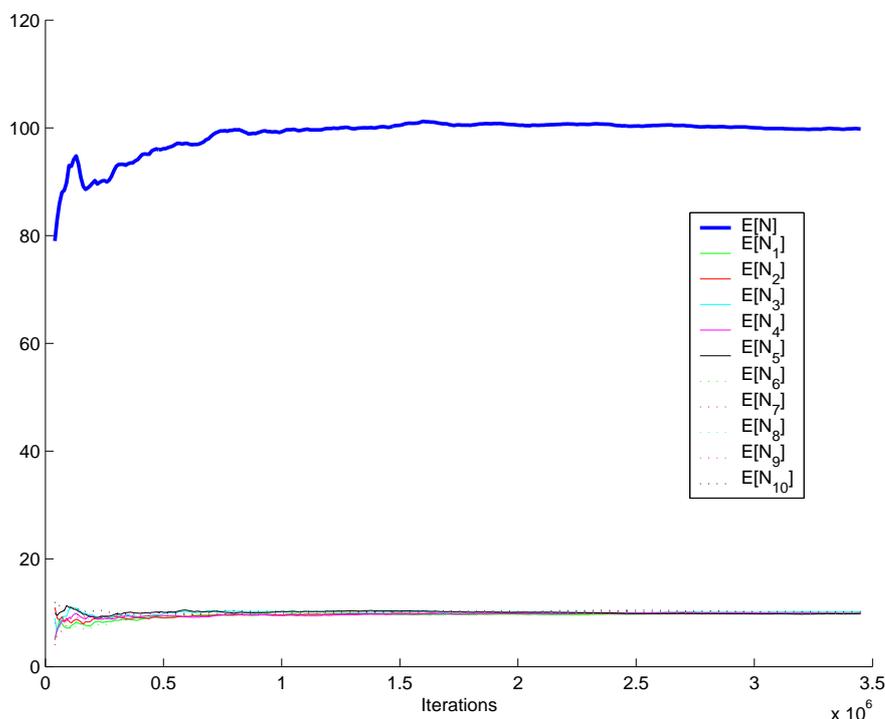


FIG. 4.5 – Moyennes empiriques du nombre total de lignes brisées et du nombre de lignes brisées par taille en fonction du nombre d’itérations de l’algorithme MCMC à sauts réversibles.

L’algorithme a convergé en moins de 34 secondes avec un processeur 2 GHz (3.5×10^6 itérations). Notons que le critère de convergence a été choisi très dur pour ne pas commettre d’erreur de diagnostic ; nous aurions pu nous arrêter entre 1 et 2 millions d’itérations comme le montre la figure 4.5. A la convergence, l’erreur empirique faite sur l’espérance du nombre de points est faible : elle est de l’ordre de 0.2 pour une valeur théorique $E(N) = \lambda|F| = 100$. De plus, les valeurs empiriques de l’espérance du nombre de lignes de taille i ($i = 1, \dots, N$) sont toutes proches de la valeur théorique obtenue par l’hypothèse d’une loi uniforme sur le nombre de points d’une ligne brisée :

$$P(i) = \frac{1}{10} \forall i = 1..10 \Rightarrow E[N_i] = \frac{E[N]}{10} = 10 \forall i = 1..10$$

4.5.4 Translation d’un point de contrôle

Nous utiliserons également le noyau de proposition q_3 qui consiste à proposer une translation d’un point de contrôle d’une ligne brisée. La translation correspond à une transformation symétrique T_a paramétrée par un vecteur a tiré uniformément dans un compact centré autour de l’origine. Le point de contrôle étant choisi uniformément dans

la configuration, le taux de Green est réduit au rapport des densités :

$$R(\mathbf{c}, \mathbf{c}') = \frac{h(\mathbf{c}')}{h(\mathbf{c})} \quad (4.24)$$

4.5.5 Division et fusion de segments

En plus de la translation, l'ajout et le retrait d'un point de contrôle au sein d'une ligne brisée permettent de perturber de façon pertinente une ligne brisée approximativement bien positionnée. Ces mouvements définissent un mouvement réversible appelé "division et fusion" de segments illustré par la figure 4.6.

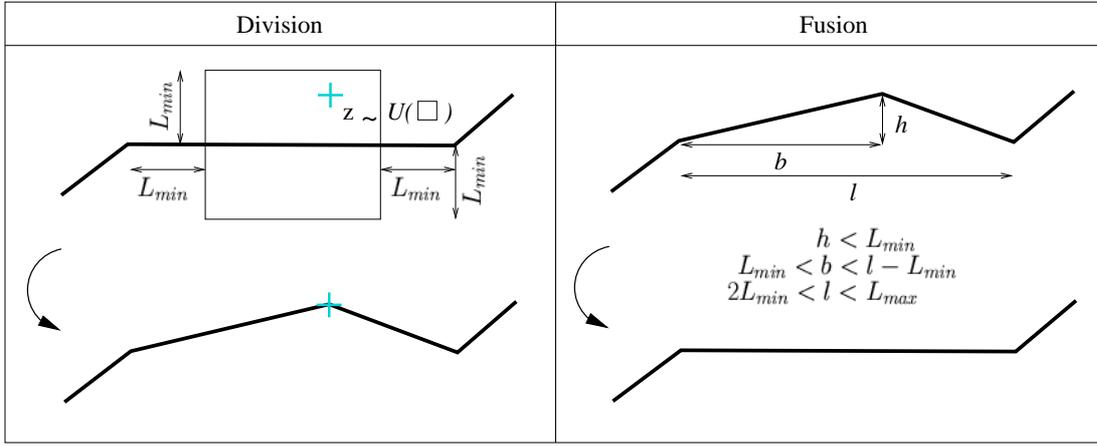


FIG. 4.6 – Division et fusion de segments.

a) Division

Soit $s = (p^j, p^{j+1})$ le segment dont les extrémités correspondent à deux points de contrôle consécutifs d'une ligne brisée de la configuration. Pour que s puisse être divisé, nous imposons que celui-ci soit de longueur l strictement supérieure à $2L_{min}$ et qu'il n'appartienne pas à une ligne brisée composée de n_{max} segments. Nous utilisons une variable auxiliaire Z définie ci-dessous pour obtenir deux nouveaux segments $s'_1 = (p^j, p')$ et $s'_2 = (p', p^{j+1})$ tels que p' soit situé dans le rectangle de longueur $l - 2L_{min}$ et de largeur $2L_{min}$ et dont l'axe principal correspond au segment initial. Remarquons que nous sommes ainsi assurés que les longueurs des nouveaux segments soient dans $[L_{min}, L_{max}]$. La variable auxiliaire Z correspond à une génération uniforme d'un point dans le rectangle de longueur $l - 2L_{min}$ et de largeur $2L_{min}$:

$$Z = \left[\begin{array}{l} H \sim \mathcal{U}([-L_{min}, L_{min}]) \\ B \sim \mathcal{U}([L_{min}, l - L_{min}]) \end{array} \right] \quad (4.25)$$

où l est la longueur de s . Du vecteur généré $z = (h, b)$, on peut passer aux paramètres $v_1 = (l_1, \alpha_1)$ décrivant le segment s'_1 par le difféomorphisme défini par :

$$v_1 = \eta_v(h, b) = \begin{bmatrix} \sqrt{h^2 + b^2} \\ \alpha + \arctan(\frac{h}{b}) \end{bmatrix} \quad (4.26)$$

où les paramètres $v = (l, \alpha)$ du segment s sont considérés fixés. Ensuite, on obtient les paramètres $v_2 = (l_2, \alpha_2)$ décrivant le segment s'_2 à partir de v_1 et $v = (l, \alpha)$ de la façon suivante :

$$v_2 = T(v, v_1) = \begin{bmatrix} \sqrt{(l \sin(\alpha) - l_1 \sin(\alpha_1))^2 + (l \cos(\alpha) - l_1 \cos(\alpha_1))^2} \\ \arctan(\frac{l \sin(\alpha) - l_1 \sin(\alpha_1)}{l \cos(\alpha) - l_1 \cos(\alpha_1)}) \end{bmatrix} \quad (4.27)$$

b) Fusion

Soit $s_j = (p^j, p^{j+1})$ et $s_{j+1} = (p^{j+1}, p^{j+2})$ deux segments consécutifs d'une ligne brisée, où p^j , p^{j+1} et p^{j+2} sont les trois points de contrôle consécutifs décrivant ces deux segments. La fusion de ces deux segments consiste à remplacer s_j et s_{j+1} par le segment $s'_j = (p^j, p^{j+2})$. La fusion de deux segments ne sera proposée que dans le cas où le point intermédiaire p^{j+1} est situé dans le rectangle qui pourrait être associé à $s_{l, l+1}$ pour une proposition de division. Les conditions à vérifier sont donc les suivantes :

- les deux points extrémaux p^j et p^{j+2} sont à une distance comprise entre $2L_{min}$ et L_{max} ;
- le point intermédiaire p^{j+1} est à une distance inférieure à L_{min} de $s'_j = (p^j, p^{j+2})$;
- le point correspondant à la projection orthogonale de p^{j+1} sur (p^j, p^{j+2}) est situé à une distance b de p^j comprise entre $2L_{min}$ et $l - L_{min}$.

c) Noyau de proposition

Soit \mathbf{c} la configuration courante. Il y a $N_D(\mathbf{c})$ segments pouvant être divisés dans cette configuration et $N_F(\mathbf{c})$ couple de segments pouvant être fusionnés. Il y a donc $N_T(\mathbf{c}) = N_D(\mathbf{c}) + N_F(\mathbf{c})$ transformations possibles de type "division et fusion" de segments. La première étape consiste à tirer uniformément une transformation parmi les $N_T(\mathbf{c})$ transformations. Si celle-ci est une division, nous procédons comme indiqué en a) ; si celle-ci est une fusion, nous procédons comme indiqué en b). Le noyau de proposition Q_{DFS} est donc donné par :

$$Q_{DFS}(\mathbf{c} \rightarrow A) = \sum_{i=1}^{N_D(\mathbf{c})} \frac{q_i^D(\mathbf{c}, A)}{N_T(\mathbf{c})} + \sum_{i=1}^{N_F(\mathbf{c})} \frac{q_i^F(\mathbf{c}, A)}{N_T(\mathbf{c})} \quad (4.28)$$

où $q_i^D(\mathbf{c}, A)$ correspond à la division du segment i et $q_i^F(\mathbf{c}, A)$ correspond à la fusion du couple i . La partie "division" $q_i^D(\mathbf{c}, A)$ s'écrit comme suit :

$$q_i^D(\mathbf{c}, A) = \int_{\Sigma_i} \mathbf{1}_A(D_i(\mathbf{c}, z)) \frac{dz}{2L_{min} (l_i - 2L_{min})} \quad (4.29)$$

où $\Sigma_i = [-L_{min}, L_{min}] \times [L_{min}, l_i - L_{min}]$ est le compact dans lequel la variable auxiliaire Z est générée et $D_i(\mathbf{c})$ est une configuration de lignes brisées où les paramètres (l_i, α_i) du segment i ont été remplacés par $(l_1, \alpha_1) = \eta_i(z)$ et $(l_2, \alpha_2) = T(i, \eta_i(z))$. La partie “fusion” $q_i^F(\mathbf{c}, A)$ du noyau de proposition est donnée par :

$$q_i^F(\mathbf{c}, A) = \mathbf{1}_A(F_i(\mathbf{c})) \quad (4.30)$$

$F_i(\mathbf{c})$ est une configuration de lignes brisées où les paramètres du couple i sont remplacés par les paramètres du segment résultat de la fusion de i .

d) Calcul du taux de Green

La mesure symétrique ψ sur $E \times E$ choisie pour dériver la mesure πQ_{DFS} est la suivante :

$$\begin{aligned} \psi(A, B) &= \int_A \sum_{s_i \in S(\mathbf{c})} \int_{\Sigma_i} \mathbf{1}_B(D_i(\mathbf{c}, z)) \frac{dz}{|V|} d\mu(\mathbf{c}) \\ &+ \int_A \sum_{(s_i, s_{i+1}) \in C(\mathbf{c})} \mathbf{1}_B(F_i(\mathbf{c})) |J_{\phi^{-1}}(v_i, v_{i+1})| d\mu(\mathbf{c}) \end{aligned} \quad (4.31)$$

où $S(\mathbf{c})$ désigne l'ensemble des segments pouvant être divisés, $C(\mathbf{c})$ désigne l'ensemble des segments consécutifs pouvant être fusionnés, $V = [L_{min}, L_{max}] \times [-\pi, \pi]$ est l'espace d'état des paramètres décrivant un segment, et ϕ est le difféomorphisme permettant de procéder au changement de variable suivant :

$$(v_1, v_2) \xleftarrow{\phi} (z, v)$$

où (v_1, v_2) correspondent aux paramètres des deux segments obtenus par la division du segment de paramètre v en utilisant la variable auxiliaire z . Plus précisément, ce difféomorphisme est donné par :

$$\phi(z, v) = (\eta_v(z), T(v, \eta_v(z))) \quad (4.32)$$

où η_v et T sont donnés par les équations (4.26) et (4.27). Le déterminant du jacobien de la fonction réciproque ϕ^{-1} est donné par :

$$|J_{\phi^{-1}}(v_1, v_2)| = \frac{l_1 l_2}{\sqrt{(l_1 \cos \alpha_1 + l_2 \cos \alpha_2)^2 + (l_1 \sin \alpha_1 + l_2 \sin \alpha_2)^2}} \quad (4.33)$$

où le dénominateur correspond à la longueur l du segment résultant de la fusion de v_1 et v_2 :

$$|J_{\phi^{-1}}(v_1, v_2)| = \frac{l_1 l_2}{l} \quad (4.34)$$

Finalement, dans le cas d'une division d'un segment de longueur l en deux segments de longueurs l_1 et l_2 , le taux de Green est donné par :

$$R(\mathbf{c}, \mathbf{c}') = \frac{N_T(\mathbf{c})}{N_T(\mathbf{c}')} \frac{L_{min}(l - 2L_{min})}{\pi (L_{max} - L_{min})} \frac{l}{l_1 l_2} \frac{h(\mathbf{c}')}{h(\mathbf{c})} \quad (4.35)$$

Dans le cas d'une fusion de deux segments de longueurs l_1 et l_2 en un segment de longueur l , le taux de Green est donné par :

$$R(\mathbf{c}, \mathbf{c}') = \frac{N_T(\mathbf{c})}{N_T(\mathbf{c}')} \frac{\pi (L_{max} - L_{min}) l_1 l_2}{L_{min}(l - 2L_{min})} \frac{h(\mathbf{c}')}{h(\mathbf{c})} \quad (4.36)$$

4.5.6 Division et fusion de lignes brisées

Nous proposons deux types de perturbations permettant la division d'une ligne brisée et la fusion de deux lignes brisées. Ces deux perturbations sont illustrées par la figure 4.7. La première perturbation (DFL1) consiste en l'ajout d'un segment reliant deux lignes brisées et le retrait d'un segment d'une ligne brisée. La deuxième perturbation (DFL2) consiste à briser une connexion entre deux segments consécutifs et à en créer une par un mouvement d'un point initial ou final d'une ligne.

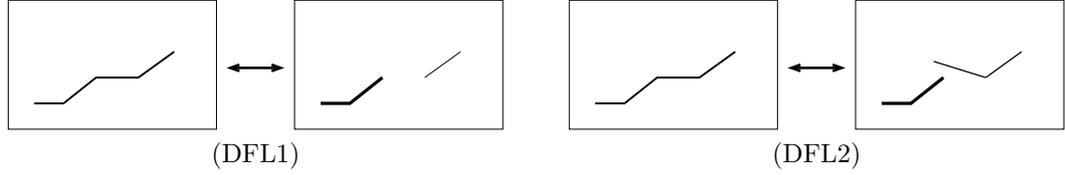


FIG. 4.7 – Perturbations de type “division et fusion” de lignes brisées.

4.5.6.1 Perturbation DFL1

Soit D_{j,δ_e} la transformation qui divise une ligne par retrait du segment j d'une ligne brisée d'épaisseur e composée d'au moins trois segments et qui permet de définir deux lignes brisées dont les épaisseurs sont égales à $e + \delta_e$ et $e - \delta_e$:

$$D_{j,\delta_e}(p, e, v_1, \dots, v_n) = \{(p^1, e + \delta_e, v_1, \dots, v_{j-1}), (p^{j+1}, e - \delta_e, v_{j+1}, \dots, v_n)\} \quad (4.37)$$

La variation de largeur δ_e est générée uniformément dans un compact tel que les deux nouvelles épaisseurs soient dans $[e_{min}, e_{max}]$:

$$\delta_e \sim \mathcal{U}([-M(e), M(e)]) , \text{ où } M(e) = \min\{e - e_{min}, e_{max} - e\} \quad (4.38)$$

La transformation réciproque F concerne les couples de lignes brisées (c_i, c_j) dont le point final $p_i^{n_i+1}$ de c_i et le point initial de c_j vérifient :

$$L_{min} \leq d(p_i^{n_i+1}, p_j^1) \leq L_{max} \quad (4.39)$$

De plus, la somme des segments les composant $(n_i + n_j)$ doit être inférieure à n_{max} :

$$n_i + n_j < n_{max} \quad (4.40)$$

La fusion F se définit de la façon suivante pour $c_i = (p_i^1, e_i, (v_i^k)_{k=1..n_i})$ et $c_j = (p_j^1, e_j, (v_j^k)_{k=1..n_j})$:

$$F(c_i, c_j) = (p_i^1, \frac{e_i + e_j}{2}, (v_i^k)_{k=1..n_i}, v_{ij}, (v_i^k)_{k=1..n_j}) \quad (4.41)$$

où v_{ij} correspond aux paramètres du segment $(p_i^{n_i+1}, p_j^1)$.

Le problème de la définition initiale pour la “division” est que la “fusion” ne peut s’appliquer que sur les couples de lignes brisées dont un point final et un point initial vérifie la condition de proximité (4.39). Or, il serait également intéressant de fusionner les couples dont les deux extrémités initiales ou les deux extrémités finales vérifient cette condition de proximité. Une première solution est de proposer un mouvement d’inversion de lignes brisées. Nous proposons ici d’intégrer ce mouvement d’inversion au noyau Q_{DFL1} . Ainsi, pour une division D_{j,δ_e} quatre paires de lignes brisées pourront être proposées avec une probabilité uniforme :

$$\begin{aligned} D_{j,\delta_e}^1(p^1, e, v_1, \dots, v_n) &= \{(p^1, e + \delta_e, v_1, \dots, v_{j-1}), (p^{j+1}, e - \delta_e, v_{j+1}, \dots, v_n)\} \\ D_{j,\delta_e}^2(p^1, e, v_1, \dots, v_n) &= \{(p^j, e + \delta_e, v'_{j-1}, \dots, v'_1), (p^{j+1}, e - \delta_e, v_{j+1}, \dots, v_n)\} \\ D_{j,\delta_e}^3(p^1, e, v_1, \dots, v_n) &= \{(p^j, e + \delta_e, v'_{j-1}, \dots, v'_1), (p^{n+1}, e - \delta_e, v'_n, \dots, v'_{j+1})\} \\ D_{j,\delta_e}^4(p^1, e, v_1, \dots, v_n) &= \{(p^1, e + \delta_e, v_1, \dots, v_{j-1}), (p^{n+1}, e - \delta_e, v'_n, \dots, v'_{j+1})\} \end{aligned} \quad (4.42)$$

où $v'_j = (l_j, \alpha_j - \pi)$ si $\alpha_j > 0$ et $v'_j = (l_j, \alpha_j + \pi)$ sinon. Soit $\{p_i^k, p_j^l\}$ une paire de points extrémaux pouvant engendrer une fusion de deux lignes c_i et c_j par l’ajout d’un segment les reliant. De même que pour la division, deux lignes pourront être proposées de façon équiprobable : une ligne $c' = F_{\{p_i^k, p_j^l\}}^1(\{c_i, c_j\})$ pour laquelle le nouveau segment est (p_i^k, p_j^l) et une ligne $c'' = F_{\{p_i^k, p_j^l\}}^2(\{c_i, c_j\})$ pour laquelle le nouveau segment est (p_j^l, p_i^k) .

Soit N_D le nombre de segments pouvant être supprimé pour une division de ligne et N_F le nombre de paire de points extrémaux de lignes brisées pouvant être reliés pour une fusion de lignes brisées (*i.e.* vérifiant les conditions (4.39) et (4.40)). Une perturbation est alors proposée de façon uniforme parmi le nombre total $N_T = N_D + N_F$ de perturbations de division et fusion. Le noyau de proposition Q_{DFL1} s’écrit donc de la façon suivante :

$$\begin{aligned} Q_{\text{DFL1}}(\mathbf{c} \rightarrow A) &= \sum_{s_i^j \in \mathcal{D}(\mathbf{c})} \frac{1}{N_T(\mathbf{c})} \sum_{k=1}^4 \frac{1}{4} \int_{-M(e_i)}^{M(e_i)} \mathbf{1}_A((\mathbf{c} \setminus c_i) \cup D_{j,\delta_e}^k(c_i)) \frac{d\delta_e}{2M(e_i)} \\ &+ \sum_{\{p_i^k, p_j^l\} \in \mathcal{F}(\mathbf{c})} \frac{1}{N_T(\mathbf{c})} \sum_{m=1}^2 \frac{1}{2} \mathbf{1}_A((\mathbf{c} \setminus \{c_i, c_j\}) \cup F_{\{p_i^k, p_j^l\}}^m(\{c_i, c_j\})) \end{aligned} \quad (4.43)$$

où s_i^j désigne le segment j de la ligne brisée c_i d’épaisseur e_i , p_i^k désigne le point k de la ligne brisée c_i , $\mathcal{D}(\mathbf{c})$ est l’ensemble des segments de la configuration \mathbf{c} qui, par leur retrait, peuvent engendrer une division de ligne brisée, et $\mathcal{F}(\mathbf{c})$ est l’ensemble des paires de points qui, par l’ajout d’un segment les reliant, peuvent engendrer une fusion de lignes brisées. Remarquons que la fusion d’une paire de lignes brisées $\{c_i, c_j\}$ peut se faire par quatre paires de points extrémaux : $\{p_i, p_j\}_1 = \{p_i^1, p_j^1\}$; $\{p_i, p_j\}_2 = \{p_i^1, p_j^{n_j+1}\}$;

$\{p_i, p_j\}_3 = \{p_i^{n_i+1}, p_j^{n_j+1}\}$; $\{p_i, p_j\}_4 = \{p_i^{n_i+1}, p_j^1\}$. Le noyau de proposition peut donc s'écrire de la façon suivante :

$$Q_{\text{DFL1}}(\mathbf{c} \rightarrow A) = \frac{1}{8N_T(\mathbf{c})} \sum_{c_i \in \mathbf{c}} \frac{1}{M(e_i)} \sum_{j=2}^{n_i-1} \sum_{k=1}^4 \int_{-M(e_i)}^{M(e_i)} \mathbf{1}_A((\mathbf{c} \setminus c_i) \cup D_{j,\delta_e}^k(c_i)) d\delta_e \\ + \frac{1}{2N_T(\mathbf{c})} \sum_{\{c_i, c_j\} \in \mathbf{c}} \sum_{k=1}^4 \sum_{m=1}^2 \mathbf{1}_A((\mathbf{c} \setminus \{c_i, c_j\}) \cup F_{\{p_i, p_j\}_k}^m(\{c_i, c_j\})) \quad (4.44)$$

La mesure ψ sur $E \times E$ choisie pour dériver la mesure πQ_{DFL1} est une mesure symétrique concentrée sur $\bigcup_{n=0}^{\infty} \{E_N \times E_{N+1}\} \cup \{E_{N+1} \times E_N\}$. Soit A et B des ensembles de la tribu associée à E tels que $A \subseteq E_N$ et $B \subseteq E_{N+1}$. La mesure ϕ est alors donnée par l'équation suivante :

$$\psi(A, B) = \int_A \sum_{i=1}^N \sum_{j=2}^{n_i-1} \sum_{k=1}^4 \int_{-M(e_i)}^{M(e_i)} \mathbf{1}_B((\mathbf{c} \setminus c_i) \cup D_{j,\delta_e}^k(c_i)) d\delta_e d\mu(\mathbf{c}) \\ \psi(B, A) = \int_B \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \sum_{k=1}^4 \frac{n_{\max}(e_{\max} - e_{\min})}{2d_k(c_i, c_j)\lambda|V|} \mathbf{1}_A((\mathbf{c} \setminus \{c_i, c_j\}) \cup F_k(c_i, c_j)) d\mu(\mathbf{c}) \\ = \int_B \sum_{\{c_i, c_j\}} \sum_{m=1}^2 \sum_{k=1}^4 \frac{n_{\max}(e_{\max} - e_{\min})}{2d_k(c_i, c_j)\lambda|V|} \mathbf{1}_A(\mathbf{c} \setminus \{c_i, c_j\}) \cup F_k^m(\{c_i, c_j\}) d\mu(\mathbf{c}) \quad (4.45)$$

où $c = F_k(c_1, c_2)$ est la ligne brisée résultant de la fusion de c_1 et c_2 par la paire d'extrémités $\{p_1, p_2\}_k$ et dont les paramètres de la ligne résultat sont ceux de c_1 suivi de ceux de c_2 . Σ et V désignent respectivement l'espace d'état associé à l'épaisseur d'une ligne et l'espace d'état associé aux paramètres d'un segment, et λ est le facteur d'intensité du processus de mesure μ . Le facteur $\frac{1}{2d_k(c_i, c_j)}$, où $d_k(c_i, c_j)$ désigne la distance entre les deux extrémités $\langle p_1, p_2 \rangle_k$, intervient du fait du changement de variable suivant :

$$(e_1, e_2, p') \longleftarrow (e, \delta_e, l_j, \alpha_j)$$

où e_1 et e_2 sont les épaisseurs de (c_1, c_2) obtenues par retrait du segment j d'une ligne d'épaisseur e , et p' correspond au nouveau point initial ainsi généré.

Finalement, le taux de Green, correspondant à une division d'une ligne brisée d'épaisseur e par le retrait d'un segment de longueur l , est donné par :

$$R(\mathbf{c}, \mathbf{c}') = \frac{N_T(\mathbf{c})}{N_T(\mathbf{c}')} \frac{8l\lambda|V| \min(e - e_{\min}, e_{\max} - e)}{n_{\max}(e_{\max} - e_{\min})} \frac{h(\mathbf{c}')}{h(\mathbf{c})} \quad (4.46)$$

où $|V| = 2\pi(L_{\max} - L_{\min})$. Le taux de Green, correspondant à une fusion de deux lignes brisées engendrant, par l'ajout d'un segment de longueur l , une nouvelle ligne d'épaisseur e , est donné par :

$$R(\mathbf{c}, \mathbf{c}') = \frac{N_T(\mathbf{c})}{N_T(\mathbf{c}')} \frac{n_{\max}(e_{\max} - e_{\min})}{8l\lambda|V| \min(e - e_{\min}, e_{\max} - e)} \frac{h(\mathbf{c}')}{h(\mathbf{c})} \quad (4.47)$$

4.5.6.2 Perturbation DFL2

La deuxième perturbation de fusion de lignes brisées a été construite pour permettre la fusion de deux lignes brisées proches dont les deux extrémités ne satisfont pas la condition de distance nécessaire à la proposition d'un segment entre ces deux extrémités. Nous proposons ici une fusion par le retrait du premier ou dernier segment d'une ligne brisée suivi de la fusion des deux lignes par rajout d'un segment les reliant. L'épaisseur de la nouvelle ligne est égale à la moyenne des épaisseurs des deux lignes fusionnées. Pour un couple de lignes brisées vérifiant $n_i + n_j \leq n_{max}$, nous avons donc huit possibilités de nouvelles lignes brisées non ordonnées comme cela est montré dans la figure 4.8. Parmi ces huit possibilités, seules les propositions de segments ayant une longueur comprise entre L_{min} et L_{max} permettent la fusion de lignes brisées bien définies. Ainsi, nous évitons de proposer de nombreuses fusions à rejeter. Pour une fusion donnée, deux lignes brisées composées des mêmes segments mais dans un ordre différent pourront être proposées de façon équiprobable.

Remarque 16 *L'exemple donné dans la figure 4.8 n'est pas représentatif des couples de lignes brisées que nous désirons fusionner mais permet de bien visualiser les différentes fusions possibles. Un couple de lignes brisées pour lequel la fusion est pertinente est donné figure 4.7. Remarquons qu'il ne peut pas être fusionné par une proposition du noyau DFL1 alors que DFL2 permet la fusion.*

La division concerne les lignes composées d'au moins deux segments. Un point de contrôle p_i^j est choisi uniformément parmi les points de contrôle $(p_i^2, \dots, p_i^{n_i})$ d'une ligne décrite par $n_i + 1$ points de contrôle. La ligne est alors coupée au niveau de ce point de contrôle. Nous obtenons deux lignes dont les épaisseurs sont égales à $e + \delta_e$ et $e - \delta_e$, où la variation de largeur δ_e est générée uniformément dans un compact $[-M(e), M(e)]$ défini dans l'équation (4.38) pour que les deux nouvelles épaisseurs soient dans $[e_{min}, e_{max}]$, e étant l'épaisseur de l'ancienne ligne. Pour une des deux lignes, choisie avec la probabilité 1/2, le point de contrôle choisi sera régénéré par la génération uniforme dans v d'une longueur l et d'une direction α . Comme précédemment, quatre lignes brisées ordonnées pourront être générées par inversion de l'ordre des deux lignes ainsi obtenues. Nous distinguons donc huit divisions à partir d'un point de contrôle p_i^j proposées de façon équiprobable (une fois l et α choisis).

Soit $N_F(\mathbf{c})$ le nombre de couples de points de contrôle permettant la fusion de deux lignes présentes dans \mathbf{c} et $N_D(\mathbf{c})$ le points de contrôle permettant la division d'une ligne présente dans \mathbf{c} . Une perturbation sera alors choisie uniformément parmi les $N_T(\mathbf{c}) = N_F(\mathbf{c}) + N_D(\mathbf{c})$ perturbations possibles. Le noyau de proposition s'écrit alors :

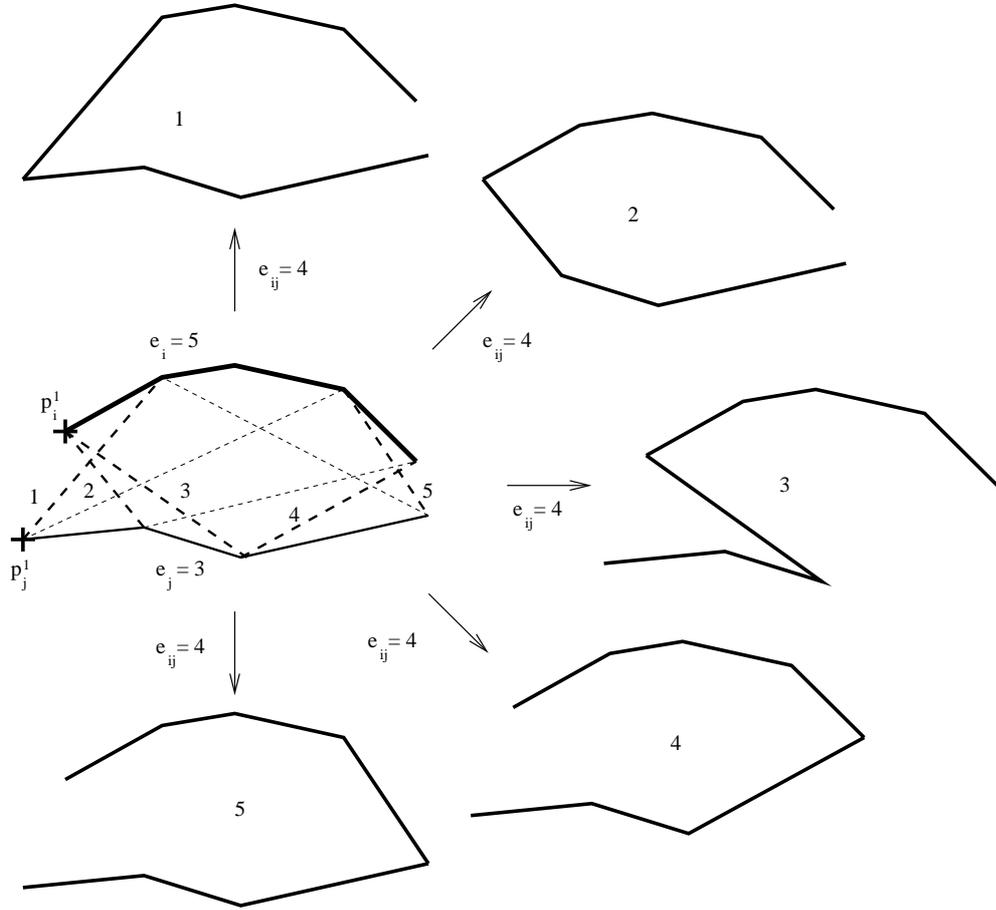


FIG. 4.8 – Fusions possibles pour une paire de lignes brisées. Seulement cinq seront proposées, les autres ne vérifiant pas la condition de l'ajout d'un segment de longueur inférieure à L_{max} .

$$\begin{aligned}
 Q_{DFL2}(\mathbf{c} \rightarrow A) &= \sum_{c_i \in \mathbf{c}} \sum_{j=2}^{n_i} \frac{1}{N_T(\mathbf{c})} \sum_{k=1}^8 \frac{1}{8} \int_{-M(e_i)}^{M(e_i)} \int_V \mathbf{1}_A((\mathbf{c} \setminus c_i) \cup D_{j,\delta_e}^k(c_i)) \frac{dv d\delta_e}{|V| 2M(e_i)} \\
 &+ \sum_{\{c_i, c_j\} \in \mathbf{c}} \sum_{k=1}^8 \frac{1}{N_T(\mathbf{c})} \sum_{m=1}^2 \frac{1}{2} \mathbf{1}_A((\mathbf{c} \setminus \{c_i, c_j\}) \cup F_{(p_i, p_j)_k}^m(\{c_i, c_j\}))
 \end{aligned} \tag{4.48}$$

Le taux de Green, correspondant à une division d'une ligne brisée d'épaisseur e via le remplacement d'un segment de longueur l , est donné par :

$$R(\mathbf{c}, \mathbf{c}') = \frac{N_T(\mathbf{c})}{N_T(\mathbf{c}')} \frac{16 l \lambda |V| \min(e - e_{min}, e_{max} - e)}{n_{max}(e_{max} - e_{min})} \frac{h(\mathbf{c}')}{h(\mathbf{c})} \tag{4.49}$$

Le taux de Green, correspondant à une fusion de deux lignes brisées via le remplacement d'un segment initial ou final par un nouveau segment de longueur l , est donné par :

$$R(\mathbf{c}, \mathbf{c}') = \frac{N_T(\mathbf{c})}{N_T(\mathbf{c}')} \frac{n_{max}(e_{max} - e_{min})}{16 l \lambda |V| \min(e - e_{min}, e_{max} - e)} \frac{h(\mathbf{c}')}{h(\mathbf{c})} \quad (4.50)$$

4.6 Résultats

Nous présentons dans ce paragraphe les résultats de l'extraction du réseau via la modélisation par un processus de lignes brisées spécifié par la densité complète donnée par :

$$h_p(\mathbf{c}) \propto h_p(\mathbf{c}) h_d(\mathbf{c}) \quad (4.51)$$

où $h_p(\mathbf{c})$ est la densité *a priori* donnée par l'équation (4.11) et $h_d(\mathbf{c})$ est le terme d'attache aux données donné par l'équation (4.15). L'optimisation est réalisée par recuit simulé sur un algorithme d'échantillonnage de type Metropolis-Hastings. Même si les perturbations définies dans le paragraphe 4.5 permettent d'accélérer la convergence et de sortir des minima locaux, l'obtention de l'intégralité du réseau nécessite une décroissance très lente de la température. C'est pourquoi nous utilisons ici une **décroissance adaptative de la température** qui permet de diminuer la température le plus vite possible tout en restant proche de l'équilibre. Le schéma de décroissance utilisé est décrit dans le paragraphe 2.6.3.

Le **temps de calcul** reste néanmoins **important** comparativement au processus de segments. Cela est principalement dû à un calcul de l'attache aux données sur l'intégralité de la ligne brisée à chaque proposition de perturbation de cette ligne brisée même si la perturbation ne concerne qu'un ou deux segments. Nous procédons ainsi car le calcul du terme d'attache aux données ne correspond pas à une somme de potentiels évalués sur les segments composants les lignes brisées mais sur des sections issues d'un découpage uniforme du masque d'une ligne brisée. Plus les lignes brisées seront longues dans la configuration, plus le temps de calcul sera long. Ainsi, le coût d'une itération est élevé à la fin de l'algorithme, et d'autant plus que les routes ou les rivières présentes dans l'image sont longues et nombreuses.

En contrepartie, l'utilisation d'un tel procédé pour calculer le terme d'attache aux données nous permet de détecter avec **précision** les réseaux sinueux comme le montrent les résultats présentés dans les figures 4.9 et 4.10. En effet, nous avons désormais la possibilité d'utiliser des segments de taille très réduite. Ainsi, nous posons $L_{min} = 3$ pour la détection des rivières sur une image de 20 mètres de résolution et $L_{min} = 5$ pour la détection du réseau routier sur une image de 10 mètres de résolution. L'évaluation du potentiel d'attache aux données se fait dans les deux cas sur des sections de taille 20 pixels.

De plus, la **détection des jonctions est meilleure** par processus de lignes brisées que par processus de segments comme le montrent les figures 4.10, 4.11 et 4.12 . Ceci est

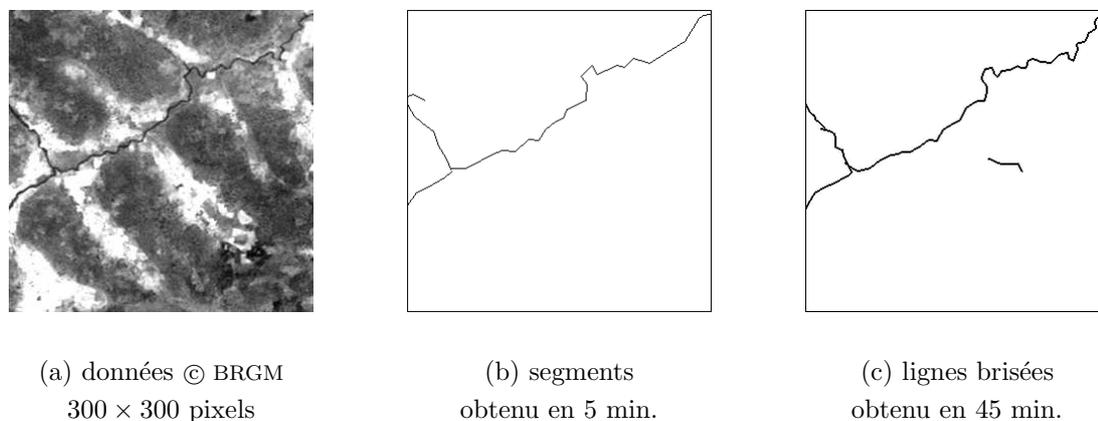


FIG. 4.9 – Extraction d'un réseau hydrographique à partir d'une image SPOT XS2 de 20 mètres de résolution : (b) par processus de segments ; (c) par processus de lignes brisées.

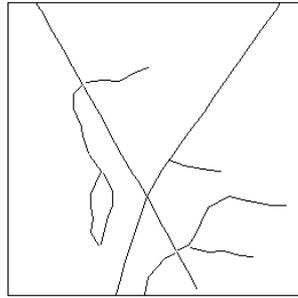
dû à l'utilisation d'une interaction de connexion favorisant la connexion de lignes brisées avec le reste du réseau. De plus, la distance de connexion est minimisée via l'utilisation d'un potentiel de connexion continu. Dans le cas des processus de segments, seule la connexion entre extrémités de segments est prise en compte. Les longues branches sont donc favorisées mais les jonctions entre ces branches ne sont pas favorisées explicitement. Néanmoins, des jonctions peuvent se former via la connexion avec une extrémité proche. Ceci explique en partie la mauvaise qualité de la jonction entre les deux routes sinueuses et une route principale dans la figure 4.10 : la connexion entre extrémités de segments a eu la primeur sur la qualité de l'attache aux données. Cette mauvaise qualité s'explique également par la pénalisation des segments proches.

L'interaction de proximité intervenant dans la modélisation par processus est en effet à revoir : elle induit une pénalisation des intersection en Y et interdit que deux segments, dont les centres sont à une distance inférieure à la demi-longueur maximale des deux segments, soient quasiment parallèles. Cette interdiction ne permet donc pas de détecter des routes proches. L'**interaction de proximité** entre lignes brisées est **mieux définie** : les intersections à angle aigu ne sont pas pénalisées et elle autorise que deux sections soient parallèles si elles sont situées à une distance supérieure à d . En prenant d petit, on peut donc détecter des routes très proches tout en interdisant les superpositions de segments. Cela est illustré par le résultat présenté dans la figure 4.12, où deux routes quasiment parallèles ont pu être détectées par processus de lignes brisées alors qu'une seule a pu être détectée par processus de segment.

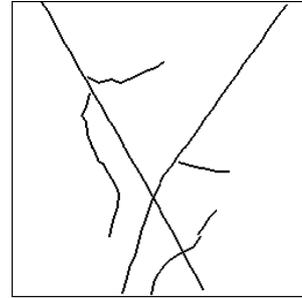
Du point de vue des omissions et des surdétectées, les processus de lignes brisées fournissent une extraction de qualité équivalente à une extraction obtenue pas processus



(a) Image 1 © CNES
256 × 256 pixels



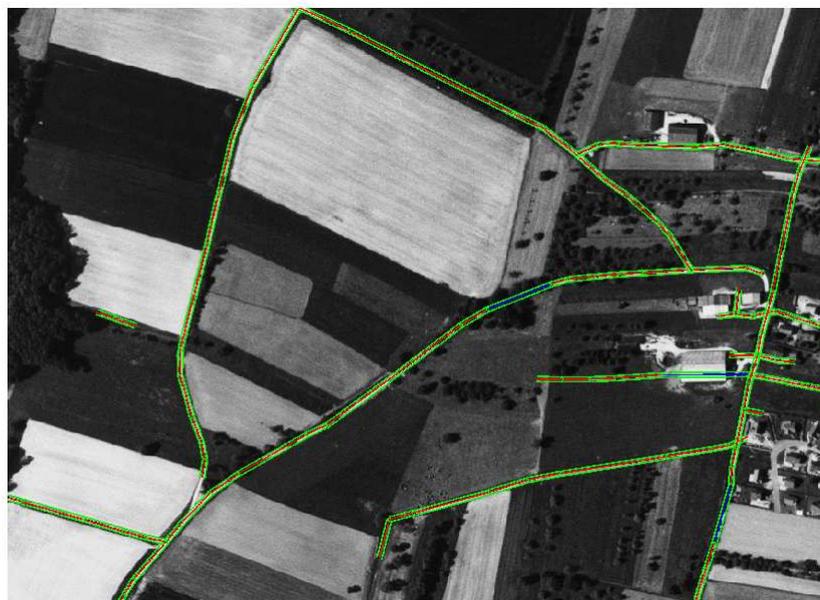
(b) segments
obtenu en 7 min.



(c) lignes brisées
obtenu en 1 h et 15 min.

FIG. 4.10 – Extraction d'un réseau routier à partir d'une image SPOT Panchro de 10 mètres de résolution : (b) par processus de segments ; (c) par processus de lignes brisées.

de segments. Une idée serait alors de combiner les deux approches en utilisant le résultat de l'extraction obtenu par processus de segments en tant qu'initialisation d'une extraction par processus de lignes brisées. Ce serait une façon d'obtenir un résultat précis en un temps de calcul raisonnable.



(a) processus de segments



(b) processus de lignes brisées

FIG. 4.11 – Résultats de l'extraction sur une image aérienne (figure 3.19) : (a) par processus de segments : les segments rouges répondent bien aux données (potentiel négatif) contrairement aux segments bleus (potentiel positif), les pixels verts sont les pixels utilisés pour l'évaluation du contraste avec le fond proche ; (b) par processus de lignes brisées : visualisation du masque associé aux lignes brisées en vert (partie interne) et bleu (partie externe) et des connexions de lignes brisées en rouge.



(a) processus de segments



(b) processus de lignes brisées

FIG. 4.12 – Résultats de l'extraction sur une image radar ERS (figure 3.21) : (a) par processus de segments : les segments rouges répondent bien aux données (potentiel négatif) contrairement aux segments bleus (potentiel positif), les pixels verts sont les pixels utilisés pour l'évaluation du contraste avec le fond proche ; (b) par processus de lignes brisées : visualisation du masque associé aux lignes brisées en vert (partie interne) et bleu (partie externe) et des connexions de lignes brisées en rouge.

Chapitre 5

Extension à l'extraction de réseaux hiérarchiques

Ce chapitre concerne l'extraction d'un réseau hiérarchique constitué de fleuves et de leurs affluents à partir d'une image radar de type ERS. Nous proposons un algorithme complètement automatique pour la détection d'un réseau hydrographique ayant une structure d'arbre. L'extraction du surfacique (branches de largeur supérieure à trois pixels) est réalisée par un algorithme efficace fondé sur une modélisation par champ de Markov. Ensuite, l'extraction du linéique se fait par un algorithme récursif fondé sur la définition d'un processus de lignes brisées par rapport à ce qui a déjà été détecté. Nous obtenons des résultats prometteurs en terme d'omissions et de surdétectons.

5.1 Exemple de réseau hiérarchique

Nous nous intéressons ici à l'extraction du réseau hydrographique à partir de données de télédétection dans une perspective d'aide à la mise à jour cartographique. En effet, l'imagerie satellitaire constitue pour les cartographes un outil très important pour optimiser le temps passé sur le terrain tout en améliorant la précision du document cartographique final. L'extraction du réseau sera menée sur une zone située en Guyane, où les mauvaises conditions météorologiques (nuages) rendent l'extraction du réseau à partir de l'imagerie optique difficile, comme cela est illustré par la figure 5.1. Nous proposons donc d'utiliser l'imagerie radar (de type radar à synthèse d'ouverture, RSO), dont le principal avantage est son indépendance aux sources d'illumination extérieures. Les capteurs radar sont ainsi opérationnels de jour comme de nuit et quelles que soient les conditions météorologiques, d'où l'appellation capteurs tout-temps. De plus, en imagerie radar, les surfaces lisses présentent une radiométrie faible, alors que les surfaces rugueuses présentent une radiométrie élevée. Ce type d'images se prête donc bien à la détection des fleuves, ceux-ci correspondant à des zones sombres dans un fond clair. Ainsi, notre objectif est l'extraction non supervisée du réseau hydrographique à partir de l'image radar ERS donnée dans la figure 5.2, en étant le plus exhaustif possible tout en limitant les fausses alarmes. Bien que la méthode proposée n'utilise qu'une seule

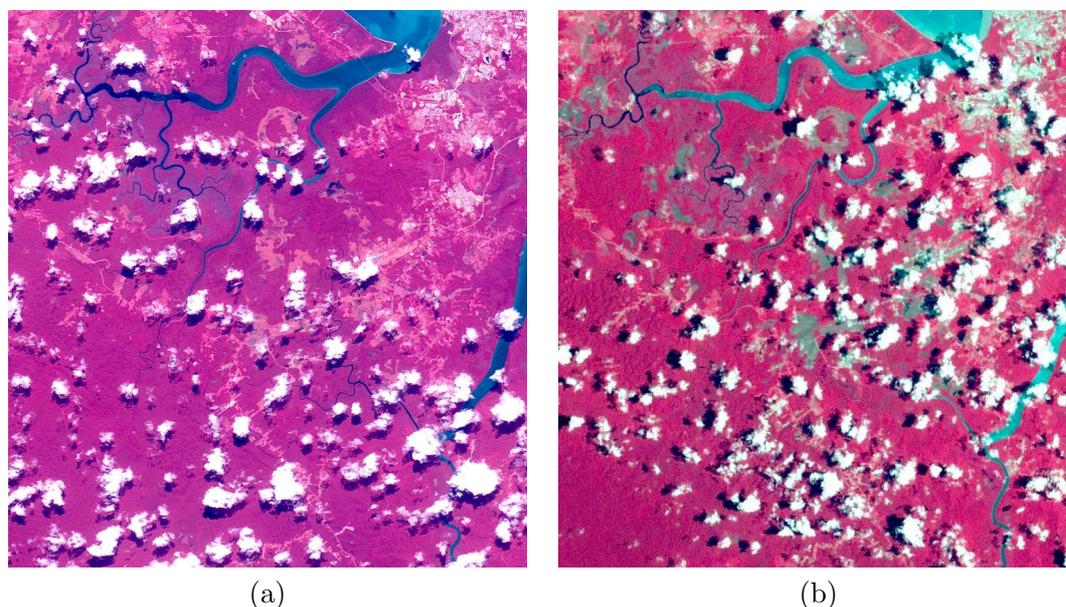


FIG. 5.1 – Sensibilité de l'imagerie optique aux conditions météorologiques : (a) image SPOT de la Guyane prise le 2 juillet 2001 ; (b) image Landsat de la même zone prise le 2 septembre 2002. Ces images nous ont été fournies par le BRGM.

image en entrée, ce travail pourrait facilement être étendu à un cadre de fusion de données pour bénéficier de l'apport des différents capteurs disponibles. Les données satellitaires, ainsi qu'une extraction manuelle du réseau (figure 5.3), nous ont été fournies par le Bureau de Recherche Géologique et Minière (BRGM).

Le réseau à extraire est constitué de deux fleuves et de leurs affluents. Pour l'extraction d'un tel réseau, nous nous devons de proposer une autre modélisation que celle proposée pour l'extraction de l'axe central des routes ou des rivières. En effet, ce réseau n'est plus caractérisé par des branches de largeur constante mais par des branches où l'on observe une augmentation progressive de la largeur de la source à l'embouchure. Dans un cadre de géométrie stochastique, cette caractéristique devra être incorporée à la forme des objets du modèle. Ainsi, un fleuve pourra être modélisé par un polygone dont la largeur diminue (ou augmente) progressivement du point initial de l'axe central de ce polygone censé représenter l'embouchure du fleuve (ou la source) jusqu'au point final de l'axe central de ce polygone représentant la source (ou l'embouchure). De plus, les branches sont reliées entre elles selon une structure arborescente : chaque fleuve principal constitue le tronc d'un arbre et les affluents de ce fleuve peuvent être assimilés à des branches à partir desquelles d'autres branches peuvent naître. Cette structure d'arbre pourra être exploitée de façon pertinente en proposant un algorithme récursif permettant la génération de nouvelles branches sachant ce qui a déjà été détecté. En effet, en supposant qu'un fleuve ait été détecté, la génération des affluents peut se faire

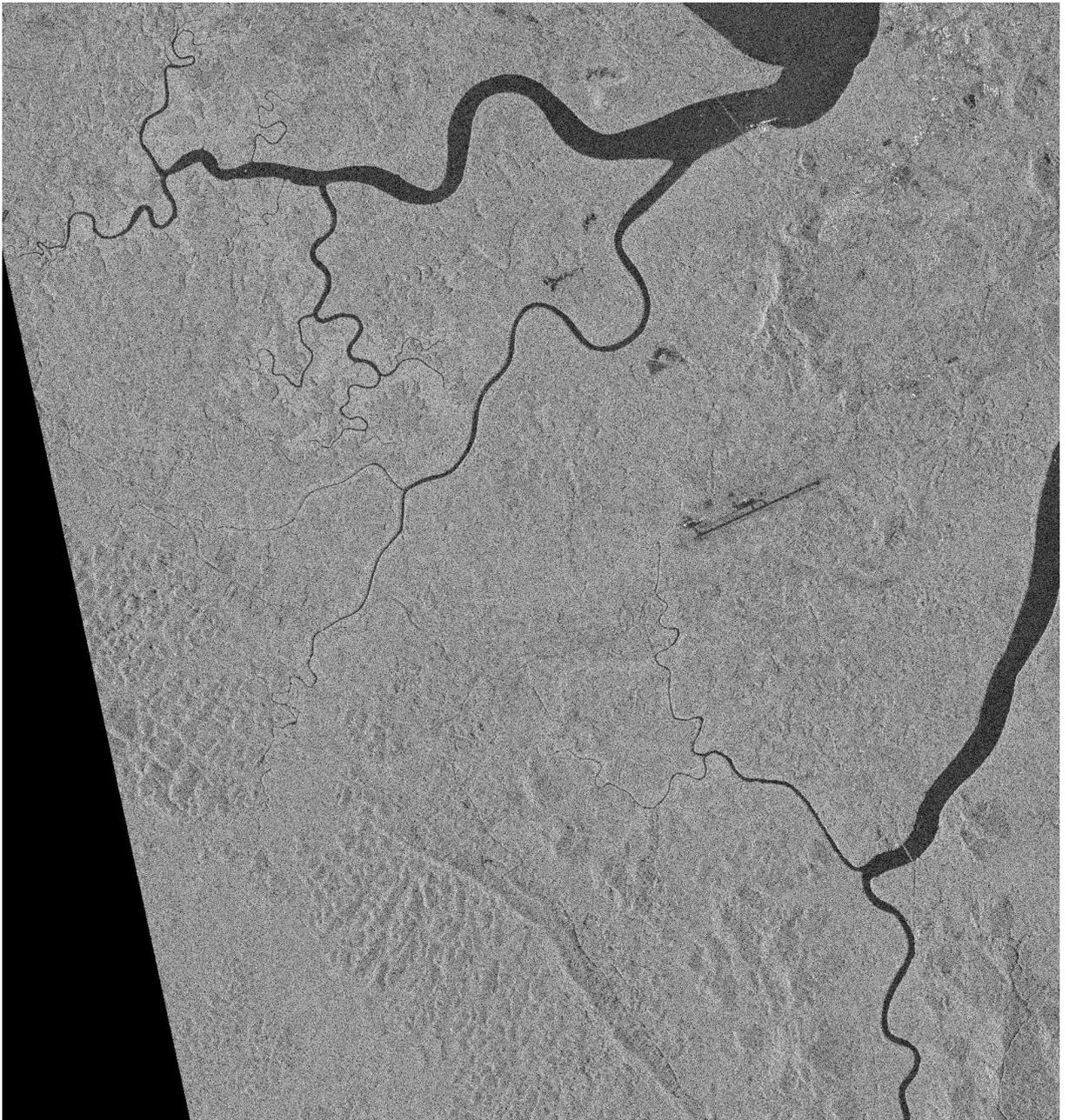


FIG. 5.2 – Image radar ERS de la Guyane fournie par le BRGM. La taille de l'image est 1709×1825 pixels avec une résolution de 12.5 mètres.



FIG. 5.3 – Extraction manuelle du réseau hydrographique fournie par le BRGM.

de façon efficace à proximité du fleuve.

5.2 Segmentation par champ de Markov

Avant de proposer une modélisation par processus objet, qui est une approche séduisante mais qui peut s'avérer très lourde en temps de calcul, notamment lorsque des objets complexes sont manipulés, nous proposons une modélisation pixélique de l'image dans un objectif de classification en deux classes : l'une correspondant au fond de l'image, l'autre correspondant au réseau hydrographique. Dans cette optique, nous proposons une modélisation par champ de Markov [Winkler, 2003] dans un cadre bayésien. En effet, ces modèles, connus pour leur robustesse au bruit, permettent d'introduire explicitement des connaissances *a priori* sur la structure spatiale des images analysées, au travers de probabilités conditionnelles locales, conjointement à la modélisation des mécanismes de dégradation des données.

5.2.1 Modèle a priori

L'image cachée X est considérée comme la réalisation d'un champ aléatoire $\mathbf{X} = \{X_s\}_{s \in S}$, où S est l'ensemble des sites de l'image (les pixels). Le champ \mathbf{X} est un champ de Markov si :

$$P(X_s = x_s | x_t, t \neq s) = P(x_s | \{x_t\}, t \in N_s)$$

où N_s est l'ensemble des sites voisins du site s . Une autre caractéristique importante vient du théorème de **Hammersley-Clifford** qui stipule qu'un **champ de Markov** tel que $P(X = x) > 0$ pour tout x , est un **champ de Gibbs** défini par :

$$P(X = x) = \frac{1}{Z} e^{-U(x)}$$

où Z est la constante de normalisation, appelée fonction de partition et U est une fonction d'énergie définie par :

$$U(x) = \sum_{c \in C} V(c)$$

où C est l'ensemble des cliques correspondant au voisinage choisi. Ici, nous considérons un voisinage d'ordre 1 (N_s est réduit aux 4 plus proches voisins) et des cliques d'ordre 1 et 2 (singletons et voisins deux à deux). Pour obtenir des zones homogènes dans l'image, l'énergie *a priori* est définie de façon à favoriser les pixels voisins ayant le même label. Par exemple, celle-ci peut simplement correspondre au nombre de cliques $\langle s, t \rangle$ ayant des labels différents (modèle de Potts) :

$$U(x) = \sum_{\langle s, t \rangle} \delta_{x_s \neq x_t} \quad (5.1)$$

où $\delta_{x_s \neq x_t}$ est égal à 1 si $x_s \neq x_t$ et à 0 sinon. Le problème d'un tel choix est qu'il induit une pénalisation des contours des objets. En utilisant un terme simple de lissage

assignant un potentiel positif aux paires de sites voisins ayant des labels différents, nous aurons par conséquent bien du mal à préserver les éléments linéiques. Afin de récupérer l'essentiel du réseau nous proposons de définir un champ booléen de contour. Ce champ, introduit par [Geman et Geman, 1984] pour la restauration d'image, représente explicitement la présence ou l'absence de discontinuités et vient donc rompre l'hypothèse de lissage faite par les méthodes classiques de régularisation. Le champ de contour est défini sur la grille duale de la grille pixélique S sur laquelle sont définis X et Y comme le montre la table 5.1.

o	+	o	+	o
+		+		+
o	+	o	+	o
+		+		+
o	+	o	+	o

TAB. 5.1 – Grille pixélique S : o, grille contour S^b : +

[Geman et Geman, 1984] considère le champ de contour inconnu, et propose d'estimer celui-ci parallèlement au champ caché X . Ici, dans un souci d'efficacité en terme de temps de calcul, nous proposons de considérer ce champ connu. Pour l'obtenir, nous utilisons un filtre de "Canny-Deriche" [Canny, 1986, Deriche, 1987] appliquée à une image filtrée par un filtre médian afin de réduire l'effet du bruit de chatoiement de l'image. Ce filtrage est suivi d'une extraction des maxima locaux (contours fins) et d'un seuillage par hystérésis, consistant à un seuillage bas suivi d'un chaînage des pixels et une conservation des chaînes contenant au moins un maximum local supérieur à un seuil haut. L'énergie *a priori* sur le champ caché X peut alors s'écrire comme suit :

$$U_1(X) = \sum_{\langle s,t \rangle} \delta_{x_s \neq x_t} (1 - b_{\langle s,t \rangle}) \quad (5.2)$$

où $b_{\langle s,t \rangle}$ désigne la valeur du champ de contour B au site de la grille de contour S^b situé entre deux sites voisins s et t sur la grille S . Celle-ci est égale à 1 si un contour est présent en ce site et à 0 sinon. Ce terme favorise donc les zones homogènes tout en préservant les discontinuités (*i.e.* les contours).

5.2.2 Vraisemblance des observations

Bien que le bruit des images radar soit corrélé, nous supposons que les valeurs de l'image Y sont conditionnellement indépendantes sachant X et que le niveau de gris y_s ne dépend que de la classe à laquelle il appartient, *i.e.* de la valeur x_s . Bien qu'erronée, cette hypothèse nous permet de définir simplement la vraisemblance des observations et, surtout, de gagner en efficacité, ce qui est l'objectif premier de la méthode d'extraction proposée dans ce paragraphe. Nous pouvons alors écrire la vraisemblance des

observations de la façon suivante :

$$f(Y|X) = \prod_{p \in S} g(y_s|x_s) \quad (5.3)$$

où $g(y_s|x_s)$ est la vraisemblance de y_s sachant x_s , *i.e.* sachant les paramètres d'un modèle de fond, si x_s est la valeur correspondant au fond, ou les paramètres d'un modèle associé aux fleuves sinon. Les deux modèles choisis sont des distributions gaussiennes de moyennes distinctes m_F et m_O et de même variance σ^2 . L'énergie d'attache aux données peut alors s'écrire comme suit :

$$U_2(Y|X) = \frac{1}{2\sigma^2} \sum_s (y_s - m_s)^2 \quad (5.4)$$

où $m_s = m_F$ si s appartient au fond de l'image ($x_s = \text{"fond"}$) ou $m_s = m_O$ sinon ($x_s = \text{"fleuve"}$).

Remarque 17 *Nous avons également testé l'algorithme de classification avec une hypothèse de bruit multiplicatif, théoriquement plus adapté à l'imagerie radar. Mais les meilleurs résultats ont néanmoins été obtenus avec une vraisemblance gaussienne.*

5.2.3 Optimisation

Si X est modélisé par un champ de Markov de système de voisinage $\{N_s, s \in S\}$ et d'énergie *a priori* U_1 , alors compte tenu de l'indépendance des $y_s|X$, on sait que le champ *a posteriori* $X|Y$ est un champ de Markov de même voisinage que X et d'énergie :

$$U(X|Y) = U_1(X) + U_2(Y|X) \quad (5.5)$$

Un candidat naturel pour X est la valeur qui maximise la densité *a posteriori* $P(X|Y)$:

$$\hat{X}_{MAP} = \arg \max_X P(X|Y) = \arg \min_X U(X|Y) \quad (5.6)$$

Plutôt que d'estimer ce maximum *a posteriori* par un recuit simulé sur un échantillonneur de Gibbs, nous proposons d'utiliser un **algorithme ICM** (Iterated Conditional Mode) [Besag, 1986]. Cet algorithme est un algorithme déterministe dans la mesure où il n'y a aucun tirage aléatoire. L'ICM ne comporte pas la caractéristique essentielle du recuit simulé qui est de visiter plusieurs puits d'énergie avant de se stabiliser dans l'un d'entre eux. Il ne réalise qu'une descente dans un puits d'énergie déterminé par la configuration initiale et le balayage. En ce sens, il ne peut être adapté à des énergies pour lesquelles on ne peut disposer d'une configuration initiale convenable. Néanmoins, cet algorithme reste une bonne solution si l'on veut gagner en efficacité. En effet, le nombre d'itérations nécessaire pour un ICM est nettement inférieur au nombre nécessaire pour un recuit simulé.

Parallèlement à l'estimation du champ caché X , se pose le problème de l'estimation des paramètres : σ , m_F et m_O . Remarquons qu'il aurait été plus rigoureux d'ajouter

un poids au terme *a priori*, mais, comme nous ne proposons pas d'estimer σ mais de le fixer empiriquement, cela est inutile. σ est alors vu comme un poids permettant de régler l'importance du terme d'attache aux données par rapport au terme *a priori*. Les deux paramètres m_F et m_0 , correspondant aux moyennes du niveau de gris du fond et du niveau de gris des fleuves, sont mis à jour au cours de l'algorithme par une estimation empirique de ces deux moyennes après chaque balayage de l'image. Partant de la constatation que l'intensité des fleuves est inférieure à celle du fond, nous initialisons m_F et m_0 par les valeurs suivantes : $m_F = 160$ et $m_0 = 100$. L'initialisation de X se fait par maximum de pseudo-vraisemblance :

$$x_s = \begin{cases} \text{“fond”} & \text{si } g(y_s|m_F) > g(y_s|m_0) \text{ , i.e. si } y_s > 130 \\ \text{“fleuve”} & \text{sinon} \end{cases} \quad (5.7)$$

ce qui nous permet d'avoir une configuration initiale proche de la configuration optimale. L'algorithme d'optimisation est donné dans la table 5.2.

<p>Initialisation de m_F, m_0 empiriquement. Initialisation de X_0 par maximum de pseudo-vraisemblance (équation 5.7) Au temps n, $X_n = X$</p> <ol style="list-style-type: none"> 1. Estimation des paramètres m_F et m_0 par $\overline{y_s^F}$ et $\overline{y_s^O}$: $\overline{y_s^F} = \frac{\sum_{s \in S} y_s \mathbf{1}_{\text{fond}}(x_s)}{\sum_{s \in S} \mathbf{1}_{\text{fond}}(x_s)} \quad , \quad \overline{y_s^O} = \frac{\sum_{s \in S} y_s \mathbf{1}_{\text{fleuve}}(x_s)}{\sum_{s \in S} \mathbf{1}_{\text{fleuve}}(x_s)}$ 2. Si les valeurs m_F et m_0 calculées à l'itération courante n sont égales à celles calculées à l'itération $n - 1$ et $n - 2$, arrêt de l'algorithme. 3. Estimation du champ X : Balayer l'ensemble des sites et en chaque site s, calculer la différence énergétique conditionnelle : $\Delta_U(s) = U(x_s = \text{fond} y_s, \{x_t\}, t \in N_s) - U(x_s = \text{fleuve} y_s, \{x_t\}, t \in N_s)$ Si $\Delta_U(s) < 0$, $x_s = \text{fond}$. Sinon, $x_s = \text{fleuve}$. 4. Si le nombre d'itérations maximal est atteint, arrêt de l'algorithme. Sinon, retourner en 1.
--

TAB. 5.2 – Algorithme ICM pour un étiquetage en deux classes.

5.2.4 Résultat de l'extraction du réseau par champ de Markov

Nous avons testé l'algorithme ICM sur l'image originale et sur une image à résolution réduite correspondant à l'image originale filtrée par un noyau Gaussien. Les résultats obtenus sont équivalents, voire meilleurs sur l'image réduite, le bruit y étant plus faible. De plus, le résultat de l'algorithme ICM a été obtenu en moins de 10 secondes sur l'image réduite par quatre (taille 911×853) contre 26 secondes sur l'image originale de taille 1709×1825 . La figure 5.4 montre le résultat obtenu sur l'image réduite. Celui-ci est satisfaisant dans le sens où **la majeure partie du réseau a été détectée**. Les quelques surdétectations peuvent être facilement supprimées par un post-traitement morphologique comme le montre la figure 5.5. Ce post-traitement consiste en une fermeture par un élément structurant linéaire suivi d'une extraction des grandes composantes connexes (*i.e.* contenant plus de 10000 pixels) en 8-connexité. Nous obtenons deux composantes connexes, respectivement représentées en gris et en blanc dans l'image 5.5. Chacune des deux composantes correspond à un fleuve et ses affluents présents dans l'image 5.2.

Néanmoins, une **large partie du réseau linéique**, *i.e.* de largeur inférieure à trois pixels, **a été omise** et ceci n'est pas dû à l'utilisation d'une image de taille réduite par rapport à l'image originale. En effet, cette omission est également vérifiée pour l'image originale du fait de la non détection des contours par le filtre de Canny au niveau des branches fines. En pratique, un filtre médian est nécessaire avant l'application du filtre de détection de contour si on veut éviter les nombreuses fausses alarmes dues au bruit de chatoiement. Les lignes fines de l'image disparaissent ou sont fortement atténuées par ce premier lissage. Le modèle *a priori* ne pourra donc pas préserver les discontinuités au niveau de ces lignes. Il est vrai que le modèle présenté ici est un des plus simples que l'on puisse proposer pour segmenter une image. Mais, nous avons également utilisé un autre modèle, appelé le chien-modèle, construit pour préserver non seulement les contours, mais aussi les lignes. Ce modèle a été introduit par [Descombes *et al.*, 1995] pour la restauration d'images binaires. Un résultat obtenu avec le chien-modèle est donné dans la figure 5.6. L'extraction est meilleure : on arrive à détecter plus loin dans la hiérarchie du réseau mais ce n'est pas encore optimal.

On voit là les **limites des champs de Markov**, qui ne permettent la définition des contraintes qu'à travers des interactions locales. S'il est vrai que ces contraintes peuvent s'avérer globales, elles doivent être définies localement. Dès lors, certaines contraintes géométriques sont difficiles à prendre en compte. Par exemple, la forme des objets ou la forme des régions lors d'une segmentation sont difficilement modélisées par des interactions locales. Ceci motive une approche par processus ponctuels, dans laquelle ce type d'informations géométriques peut être pris en compte. Néanmoins, les approches par processus objet peuvent s'avérer très lourdes en temps de calcul dès que l'on manipule des objets complexes. Or, la modélisation par champ de Markov permet de réaliser en peu de temps le gros du travail. Pourquoi ne pas l'exploiter en tant qu'initialisation d'un algorithme fondé sur une modélisation par processus objet ? C'est ce que nous proposons dans le paragraphe suivant en nous restreignant à la modélisation du linéique par processus de lignes brisées, le surfacique ayant été préalablement détecté par champ de

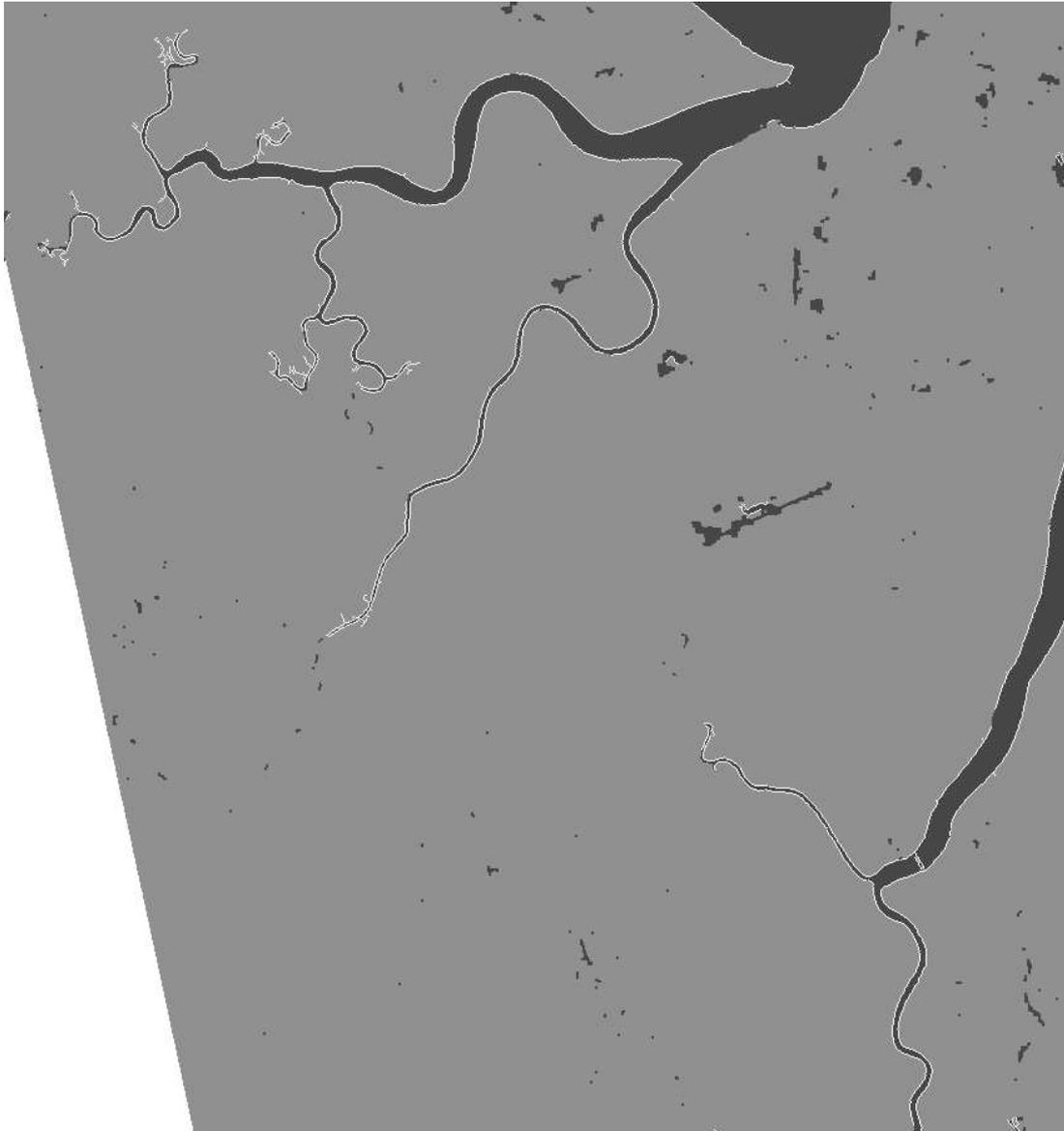


FIG. 5.4 – Classification obtenue par un ICM - Légende : NOIR = fleuves - GRIS = fond
- BLANC = bords obtenus par un filtrage de Canny suivi d'un seuillage par hystérésis.

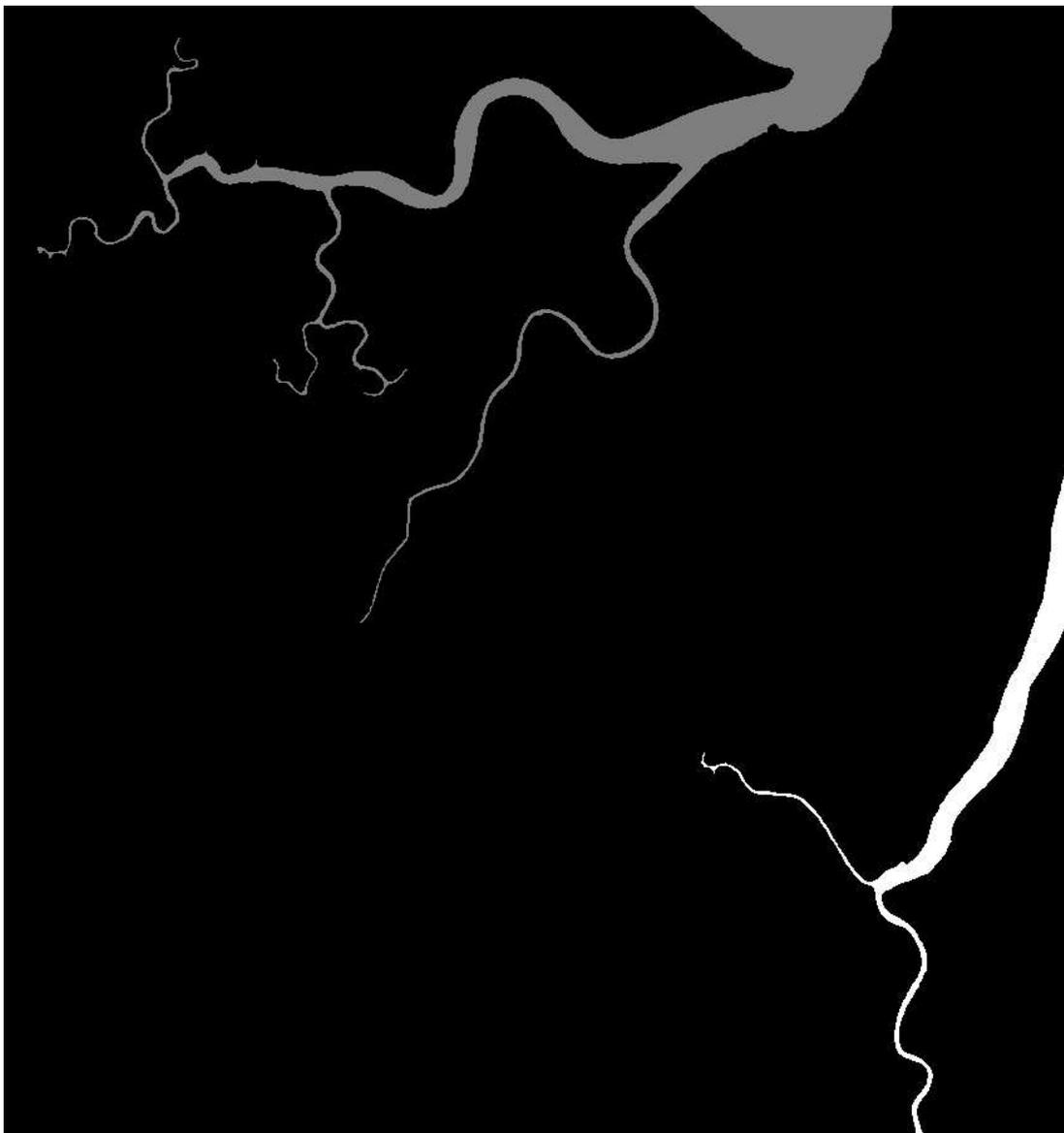


FIG. 5.5 – Post-traitement : fermeture morphologique où l'élément structurant est une ligne de 6 pixels, orientée selon 4 orientations, suivie de l'extraction des grandes composantes connexes (8-connexté). La plus grande composante connexe est en gris. La deuxième est en blanc.

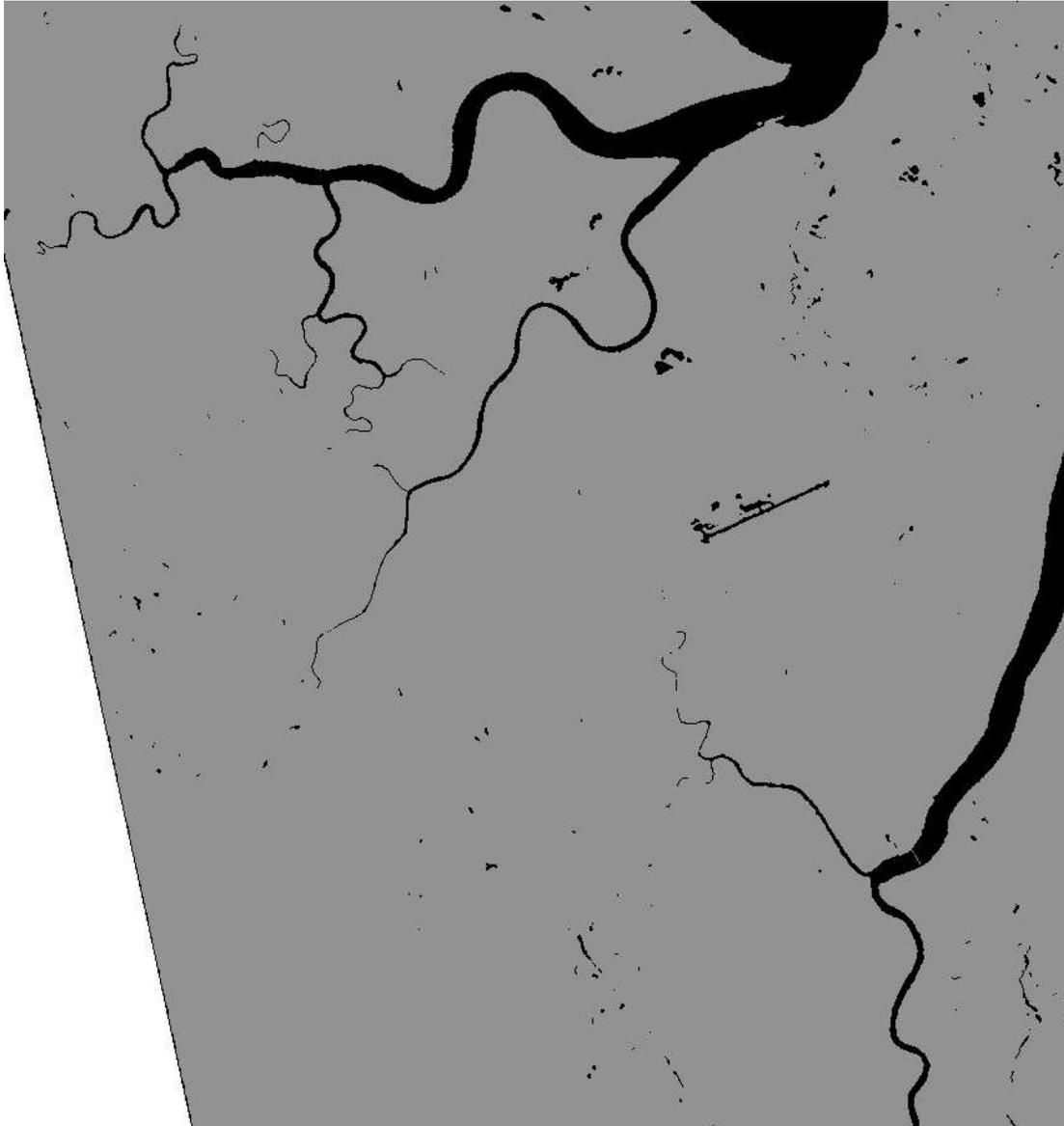


FIG. 5.6 – Classification obtenue par recuit simulé sur un échantillonneur de Gibbs, en utilisant le chien-modèle en tant que modèle *a priori*.

Markov.

5.3 Modélisation du réseau par processus objet

5.3.1 Modélisation hiérarchique du réseau

Comme nous l'avons déjà expliqué dans le paragraphe 5.1, une modélisation exploitant le caractère hiérarchique ou fractal du réseau peut s'avérer pertinente. Nous proposons donc une fondée sur la définition de processus ponctuels à différents niveaux ; le niveau 0 correspond aux fleuves principaux, le niveau 1 à leurs affluents directs, le niveau 2 aux affluents de ces affluents, etc. Cette représentation par niveau est illustrée dans la figure 5.7.

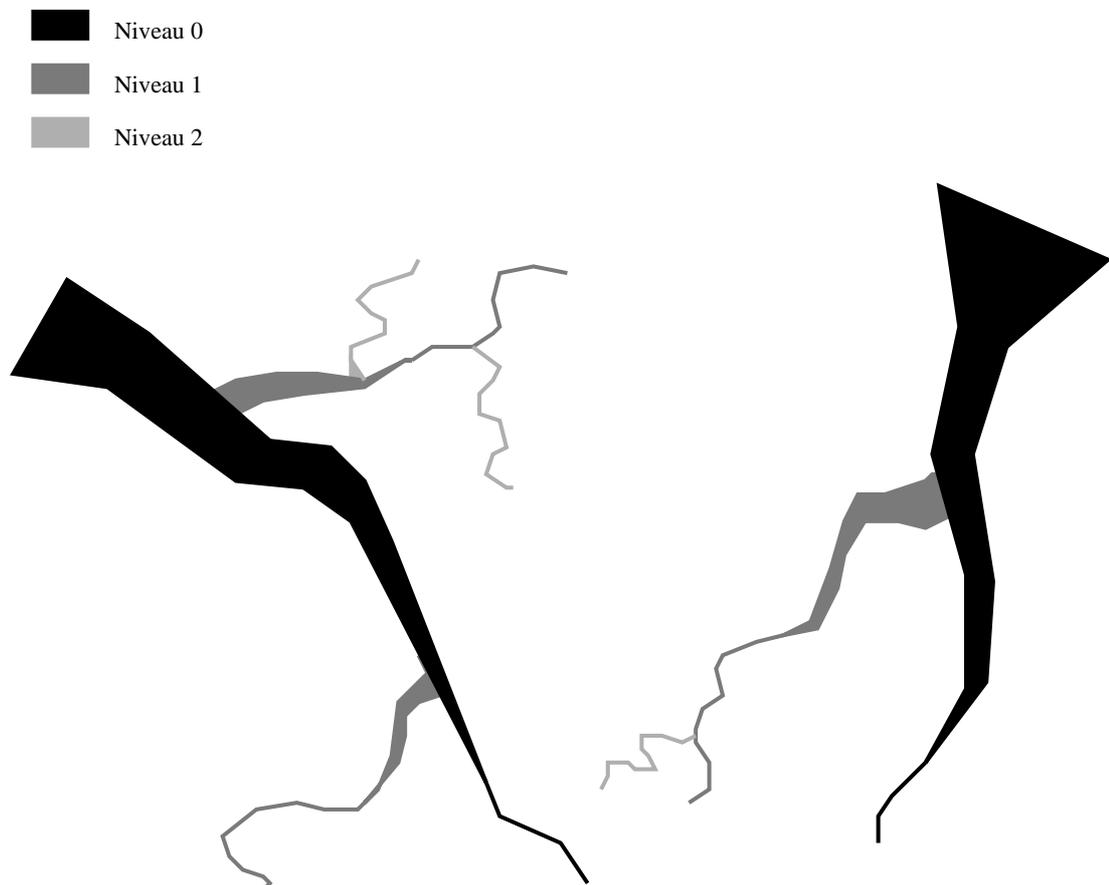


FIG. 5.7 – Modélisation hiérarchique du réseau.

Plutôt que de chercher à modéliser le réseau dans sa globalité, nous proposons de modéliser chaque niveau conditionnellement aux niveaux inférieurs. Cette modélisation

se fait via la définition d'un processus objet dans le voisinage d'un objet de la configuration d'un niveau inférieur. Les objets correspondent à des polygones représentant un fleuve dans sa totalité. La modélisation hiérarchique du réseau se définit comme suit :

- Au niveau 0 de la hiérarchie, nous définissons un processus \mathbf{C}_0 dans la fenêtre d'observation F . Le niveau 0 correspond aux fleuves principaux observables dans F .
- Au niveau 1 de la hiérarchie, le niveau 0 est considéré comme connu. Conditionnellement à $\{\mathbf{C}_0 = \{c_1, \dots, c_n\}\}$, on définit n processus "descendant" de la configuration "ascendante". Un processus descendant d'un objet c_i est un processus objet dans la zone d'influence $V(c_i) \subset F$ de l'objet c_i , correspondant à un voisinage de c_i disjoint des zones d'influence des autres objets. Ce processus correspond aux affluents d'un fleuve décrit par le polygone c_i . Remarquons que seuls les points du processus, qui correspondent aux points d'amorce des affluents, sont définis dans la zone d'influence. Les objets, décrits par un point d'amorce (position) et des marques (forme), peuvent déborder de la zone d'influence où vit le processus.
- Au niveau 2 de la hiérarchie, les niveaux 0 et 1 sont connus. Pour chaque configuration c du niveau 1, un processus descendant est défini dans une zone d'influence $V(c \subset F)$, et ainsi de suite.

Cette structure arborescente permettra de générer chacun des fleuves (niveau 0), de leurs affluents (niveau 1), des affluents de ces affluents (niveau 2), et ainsi de suite, de façon récursive. Le problème auquel nous risquons de nous heurter en choisissant une telle modélisation est le temps de calcul nécessaire pour obtenir les premiers niveaux. En effet, la manipulation d'objets épais, voire très épais au niveau des embouchures des fleuves, nécessite un temps de calcul important à chaque proposition de perturbation, l'attache aux données faisant alors intervenir un grand nombre de pixels. Une approche envisageable serait néanmoins de calculer ce terme d'attache aux données à différentes résolutions de l'image, obtenues par filtrages gaussiens successifs. Bien que séduisante, ce n'est pas l'approche que nous avons choisi d'adopter ici. En effet, l'étude menée sur l'extraction du réseau hydrographique par champ de Markov montre qu'il est possible d'extraire rapidement (environ 10 secondes) les branches de largeur supérieure à trois pixels. De plus, le post-traitement proposé nous assure d'éliminer toutes les surdétections. Nous proposons donc d'utiliser le résultat donné dans la figure 5.5 pour initialiser le réseau. Nous considérons alors que l'ensemble du réseau surfacique (de largeur supérieure à 3 pixels) est connu. Nous revenons ainsi à des objets linéiques, dont les perturbations seront nettement moins coûteuses en temps de calcul.

5.3.2 Processus définis dans le voisinage d'un objet déjà détecté

Soit \mathbf{C} , l'ensemble des objets déjà détectés. Mis à part pour le processus du niveau 0, chaque processus intervenant dans la modélisation hiérarchique est défini dans le voisinage d'un objet de $c \in \mathbf{C}$, conditionnellement à \mathbf{C} . Nous donnons, dans ce paragraphe, la définition d'un tel processus défini dans le voisinage d'un objet c décrit par son axe central (ligne brisée) et sa projection sur l'image $S(c)$. Nous noterons E_c l'équivalent en continu de la projection discrète $S(c)$ de c sur l'image. L'espace E_c est alors défini comme un compact inclu dans $F \subset \mathbb{R}^2$ délimité par les bords de l'objet c .

Le **processus de référence** \mathbf{X}_c par rapport à c est un processus ponctuel marqué dont les points suivent une loi de **Poisson** sur $V(c)$, un sous-ensemble de F correspondant à **zone d'influence** de c . La zone d'influence $V(c)$ se définit comme suit :

$$p \in V(c) \Leftrightarrow \begin{cases} d(p, c) < d_{max} \\ p \notin E_{\mathbf{C}} = \bigcup_{c \in \mathbf{C}} E_c \\ c = \arg \min_{\mathbf{C}} d(p, c) \end{cases} \quad (5.8)$$

où $d(p, c)$ désigne la distance entre le point p et les bords de l'objet c . Sur l'image, cette zone d'influence est définie par l'ensemble des pixels situés sur une bande fine située à proximité de la silhouette de l'objet c , tels que ces pixels n'appartiennent pas à la silhouette de l'ensemble \mathbf{C} des objets déjà détectés, et qu'ils ne soient pas plus proches d'un autre objet. Les zones d'influence associées à chaque objet de \mathbf{C} sont donc disjointes. La figure 5.8 illustre cette définition des zones d'influence associées à un ensemble d'objets.

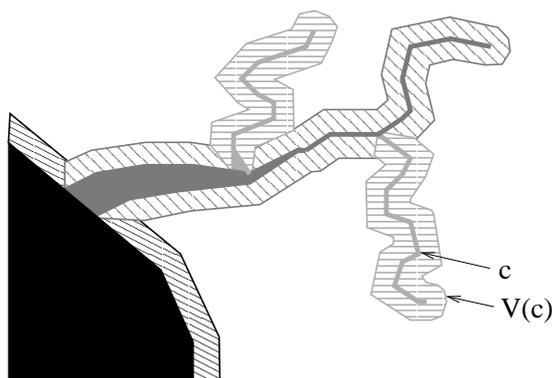


FIG. 5.8 – Zones d'influence associées aux objets déjà détectés (en uni : les objets ; en hachuré : les zones d'influence).

Les **objets** d'une configuration du processus \mathbf{X}_c sont de type **lignes brisées** où l'épaisseur n'est pas prise en compte. En effet, la projection de la ligne sur l'image est

supposée varier de 1 à 3 pixels puisque les lignes plus épaisses appartiennent par hypothèse aux niveaux supérieurs. Nous proposons alors de ne considérer que l'axe central de la ligne brisée. La projection de l'axe sur l'image correspond à une chaîne de pixel (épaisseur = 1). La variation de l'épaisseur entre 1 et 3 pixels est prise en compte au travers de l'ajout ou non des pixels voisins à cette chaîne en 4 connexité selon le rapport de vraisemblance entre une vraisemblance gaussienne de moyenne m_F et variance σ_F (fond) et une vraisemblance gaussienne de moyenne m_O et variance σ_O (fleuve) où les paramètres m_F, m_O, σ_F et σ_O ont été estimés à partir du résultat de l'ICM après post-traitement. Ainsi, l'épaisseur de la ligne est prise implicitement en compte au travers de la projection de cette ligne sur l'image qui est entièrement déterminée par son axe central. Les objets du processus \mathbf{X}_c sont donc définis par :

- un point initial $p^1 = (x, y) \in V(c)$;
- un nombre de segments $n \in \{1, \dots, n_{max}\}$;
- les longueurs des segments $l_j \in [L_{min}, L_{max}]$, $i = 1, \dots, n$;
- la direction initiale $\alpha_1 \in]\alpha_0 - \pi, \alpha_0 + \pi]$ de la ligne brisée, où α_0 correspond à la direction de la normale à l'axe central de c passant par p^1 .
- les directions des segments $\alpha_j \in]-\pi, \pi]$, $j = 2, \dots, n$.

La génération de la direction initiale α_1 proche de la direction normale à l'axe de c nous permet d'éviter la génération d'affluents à l'intérieur du fleuve.

Une **énergie a priori** est ensuite construite afin d'introduire un *a priori* sur la forme des lignes brisées et des interactions entre lignes brisées. Le terme énergétique sur la forme des lignes brisées est du même type que celui décrit pour le processus de lignes brisées présenté dans le chapitre 4. Ainsi, les **branches formées d'un petit nombre de segments** sont **pénalisées** via l'utilisation du terme U_{11} donné par l'équation (4.2). La **faible courbure du réseau** est **favorisée** au travers d'un potentiel U_{13} fondé sur les différences entre les directions successives d'une ligne brisée. En outre, ce potentiel induit l'**interdiction des angles trop aigus**. Son expression est la suivante :

$$U_{13}(\alpha_j, \alpha_{j+1}) = \begin{cases} +\infty & \text{si } \cos(\alpha_j - \alpha_{j-1}) < -0.8 \\ U_{13}(\alpha_j, \alpha_{j+1}) & \text{(équation 4.4) sinon} \end{cases} \quad (5.9)$$

Ce terme s'applique à tous les couples d'orientations successives de la ligne brisée et au couple (α_0, α_1) où α_0 est la direction normale à l'axe central du germe g au niveau du point initial de la graine.

De plus, nous **interdisons** toute **superposition** de plus de 50% du masque pixélique correspondant à un segment s d'une ligne brisée x avec le reste du réseau, c'est-à-dire la projection sur l'image de l'ensemble \mathbf{C} et de la configuration courante \mathbf{x} privée du

segment s . L'énergie *a priori* associée à une ligne brisée x composée de n segments est alors définie comme suit :

$$U_1(x) = \begin{cases} +\infty, & \text{si } \exists s \in x : |S(s) \cap S(\mathbf{C} \cup \mathbf{x} \setminus s)| > \frac{|S(s)|}{2} \\ U_{11}(n) + \sum_{j=0}^{n-1} U_{13}(\alpha_j, \alpha_{j+1}), & \text{sinon} \end{cases} \quad (5.10)$$

Finalement, l'énergie *a priori* associée au processus \mathbf{X}_c , défini par rapport à $c \in \mathbf{C}$, se définit ainsi :

$$U_p(\mathbf{x}) = \sum_{x \in \mathbf{x}} U_1(x) \quad (5.11)$$

5.3.3 Incorporation des propriétés radiométriques

L'incorporation des propriétés radiométriques se fait par un terme d'attache aux données fondé sur une mesure locale du contraste de la silhouette de la configuration courante \mathbf{x} avec son environnement proche.

La silhouette de la configuration courante se définit comme l'union des silhouettes des segments composant les lignes brisées de \mathbf{x} . La **silhouette d'un segment** est composée :

- du segment discrétisé obtenu par la méthode de Bresenham [Bresenham, 1965], méthode efficace de tracé d'un segment sur une image fournissant une chaîne de pixels entre les deux extrémités du segment telle que le segment passe par tous les pixels de la chaîne.
- de l'ensemble des pixels voisins du segment discrétisé dans la direction normale au segment et tels que la valeur v du niveau de gris du pixel vérifie :

$$g(v|m_0, \sigma_O) > g(v|m_F, \sigma_F) \quad (5.12)$$

où $g(\cdot|m, \sigma)$ est la fonction de vraisemblance gaussienne de moyenne m et de variance σ . m_0 et σ_O correspondent à la moyenne et la variance empiriques du niveau de gris de la silhouette de l'ensemble des objets détectés par l'algorithme d'initialisation. m_F et σ_F correspondent à la moyenne et la variance empiriques du fond de l'image. La condition (5.12) revient à vérifier que le rapport de log-vraisemblance des paramètres m_0 et σ_O associés aux objets par rapport aux paramètres m_F et σ_F associés au fond de l'image est positif :

$$\log\left(\frac{\sigma_F}{\sigma_O}\right) + \frac{(v - m_F)^2}{2\sigma_F^2} - \frac{(v - m_O)^2}{2\sigma_O^2} > 0 \quad (5.13)$$

La largeur des lignes brisées est ainsi prise en compte de façon implicite à partir des observations.

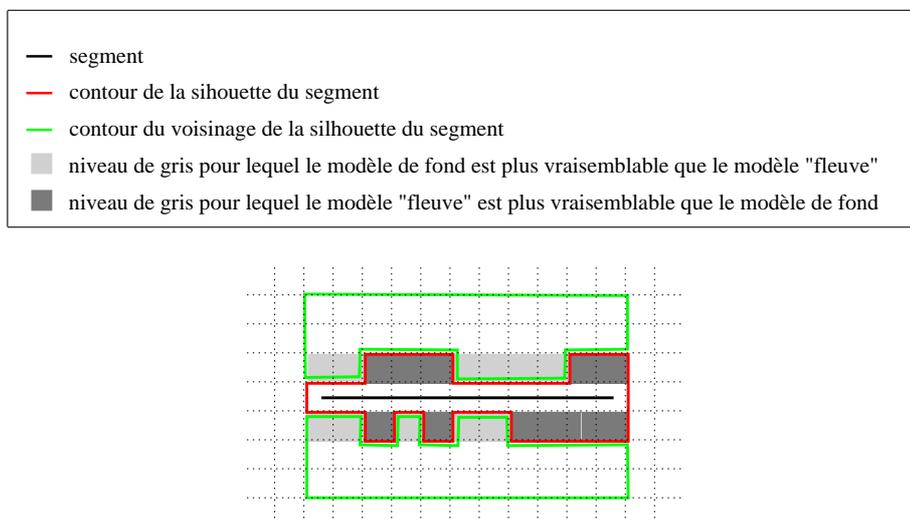


FIG. 5.9 – Masque adaptatif associé à chaque segment composant les lignes brisées de la configuration.

La **mesure de contraste locale** utilisée est fondée sur l'évaluation du contraste entre la silhouette S d'un segment s et l'ensemble de pixels F , représentant le fond proche de S . Cet ensemble F est composé de deux ensembles connexes de pixels, situés de part et d'autre de la silhouette, comme cela est illustré par la figure 5.9. Chacun de ces deux ensembles est composé des pixels voisins du segment discrétisé n'ayant pas vérifié la condition (5.13) et de deux chaînes de pixels successives colinéaires au segment discrétisé. Le contraste entre S et F est évalué par la mesure statistique t , donnée par l'équation (3.40), qui est généralement utilisée pour effectuer un test de Student (test statistique de l'hypothèse de moyennes égales pour deux groupes disjoints). De plus, le fleuve étant censé être de radiométrie inférieure à celle du fond, nous annulerons la valeur du contraste dans le cas où \bar{S} , la moyenne du niveau de gris sur S , est supérieure à \bar{F} , la moyenne du niveau de gris sur F . Finalement, la mesure locale du contraste associé au masque (S, F) est donnée par :

$$v(S, F) = \begin{cases} t(S, F) & \text{si } \bar{S} < \bar{F} \\ 0 & \text{sinon} \end{cases} \quad (5.14)$$

où t est donné par l'équation (3.40).

Nous choisissons d'incorporer les propriétés radiométriques des données en utilisant les données en tant que champ externe pour l'évaluation de mesures de contraste locales, en chaque segment de la configuration. L'approche par champ externe classique consisterait à définir l'énergie de la configuration comme la somme des potentiels fondés sur les mesures de contraste évaluées sur chaque segment de la configuration. Un défaut

de cette définition est que l'ajout d'un segment dont la silhouette est complètement superposée à la silhouette de la configuration courante pourra faire fortement baisser l'énergie s'il est bien placé. Or, un tel ajout est redondant. Nous proposons ici une **approche par champ externe non redondante**. Celle-ci se rapproche d'une approche bayésienne dans laquelle l'attache aux données est directement évaluée sur la silhouette de la configuration : un pixel ne comptera donc qu'une seule fois. Pourquoi alors ne pas utiliser directement une approche bayésienne ? C'est l'approche que nous avons initialement testée avec l'hypothèse de deux modèles gaussiens : un pour le fond, l'autre pour les objets. Cette hypothèse s'est révélée trop simplificatrice. En effet, bien que le fond paraisse homogène à première vue, on peut observer quelques zones homogènes dont l'intensité se rapproche plus de celle du fleuve. Lors des tests, nous avons pu observer une formation de serpentins dans ces zones. De plus, dans le cas où les bords de l'objet sont mal détectés (l'initialisation, faite à basse résolution, ne fournit qu'une détection approximative), une ligne brisée peut être générée le long du fleuve.

L'approche utilisée est la suivante : pour une configuration donnée, nous associons un masque adaptatif M_s à chaque segment s de la configuration, auquel est associée une valeur v_s mesurant le contraste de ce masque. L'ensemble de pixels considéré pour calculer le terme d'attache aux données est alors l'union de tous les masques de pixels ainsi créés. Autrement dit, cet ensemble est constitué de la silhouette $S(\mathbf{x})$ de la configuration \mathbf{x} et du fond proche $F(\mathbf{x})$ de cette silhouette. Chacun des pixels de $S(\mathbf{x}) \cup F(\mathbf{x})$ peut appartenir à un ou plusieurs masques de segments. On assigne alors au pixel p le minimum des valeurs de contraste associées aux masques M_1, \dots, M_N incluant le pixel p :

$$V_c(p) = \min_{i=1, \dots, N} v_i \quad (5.15)$$

où v_i est la valeur de contraste du masque $M_i = (S_i, F_i)$: $v_i = v(S_i, F_i)$, où v est donné par l'équation (5.14). Rappelons que cette valeur est d'autant plus grande que le contraste entre S_i et F_i est élevé. Finalement, l'**énergie d'attache aux données** est donnée par l'équation suivante :

$$U_d(\mathbf{x}) = \sum_{p \in S(\mathbf{x}) \cup F(\mathbf{x})} u_c(p) \quad (5.16)$$

où $U_c(p)$ est un potentiel fondé sur la mesure de contraste $V_c(p)$. Ce potentiel est donné par :

$$U_c(p) = \begin{cases} 2 & \text{si } V_c(p) < \tau_1 \\ 1 - 2 \frac{V_c(p) - \tau_1}{\tau_2 - \tau_1} & \text{si } \tau_1 \leq V_c(p) \leq \tau_2 \\ -1 & \text{si } V_c(p) > \tau_2 \end{cases} \quad (5.17)$$

où τ_1 et τ_2 sont des seuils positifs à fixer selon le contraste entre les fleuves et leur environnement dans l'image. Pour la détection des fleuves dans l'image 5.2, nous avons choisi empiriquement les valeurs suivantes : $\tau_1 = 4$ et $\tau_2 = 8$.

5.4 Extraction du réseau par une modélisation hiérarchique

Nous présentons, dans ce paragraphe, un algorithme d'extraction du réseau hydrographique fondé sur la modélisation hiérarchique proposée dans le paragraphe 5.3. La partie surfacique du réseau (branches d'épaisseur supérieure à trois pixels) est rapidement obtenue par une segmentation par champ de Markov, suivie de post-traitements visant à obtenir une représentation objet de la scène. Cette phase d'initialisation est décrite dans le paragraphe 5.4.1. Ensuite, l'extraction du linéique se fait par un algorithme récursif fondé sur la définition de processus de lignes brisées par rapport aux branches précédemment détectées. Cette phase de génération de nouvelles branches est décrite dans le paragraphe 5.4.2.

5.4.1 Initialisation du réseau

L'initialisation du réseau est fondée sur le résultat de l'extraction du surfacique obtenue par une modélisation par champ de Markov. Nous proposons ici d'extraire de ces données pixéliques le réseau sous forme d'une forêt de lignes brisées, chacun des arbres de la forêt correspondant à une des composantes connexes obtenues après le post-traitement morphologique appliqué au résultat de l'ICM. Cette étape de passage du pixélique à une représentation de la scène par une collection d'objets est nécessaire si l'on veut pouvoir extraire le réseau complet sous forme d'une collection d'objets, chaque objet correspondant à un fleuve observé. La phase d'initialisation est constituée de deux étapes : la première étape concerne le passage de l'information pixélique à la forêt de lignes brisées ; la deuxième étape consiste à prolonger toutes les feuilles des arbres de la forêt.

Étape 1 : du pixel à l'objet

Pour passer des pixels aux objets, nous proposons un algorithme en deux sous-étapes :

1. Détection du squelette de chaque composante connexe :

Le squelette correspond aux lignes de crêtes de l'hypersurface D définie par la distance de Hausdorff aux bords, évaluée en chaque pixel de l'objet considéré (*i.e.* de la composante connexe). Cette distance est calculée après détection des bords et propagation de la distance vers l'intérieur de l'objet en quatre connexité. Le squelette est obtenu par amincissements successifs préservant la connexité du squelette (8-connexité). On procède aux amincissements par retrait des pixels à chaque niveau de distance $D = 0, 1, 2, \dots$ si ceux-ci n'appartiennent pas à une ligne de crête. On entend par retrait d'un pixel, un assignation de la valeur 0 à ce pixel. En outre, la valeur 0 est attribuée au fond de l'image. Si, au contraire, il appartient à une ligne de crête, la valeur 1 lui est assignée. Pour tester si un pixel p appartient à une ligne de crête, nous procédons à une analyse sur une fenêtre

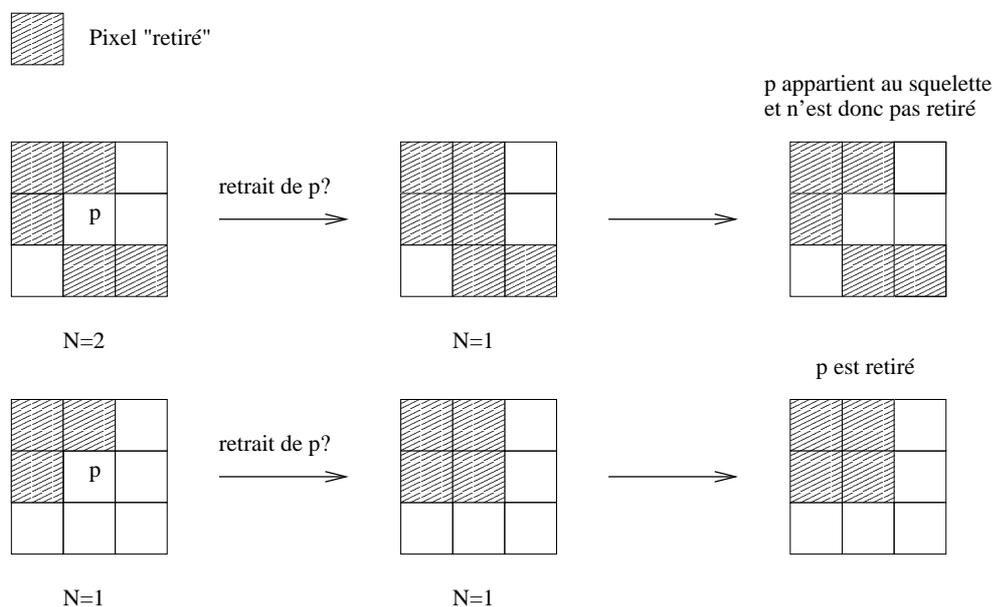


FIG. 5.10 – Retrait du pixel p en fonction de l'évolution du nombre de composantes connexes constituées de pixels "retirés".

de taille 3×3 centrée en p . Pour cela, il suffit de compter le nombre de composantes connexes (en 4-connexité) constituées de pixels ayant la valeur 0 dans cette fenêtre. Si le retrait de p induit la diminution du nombre de composantes connexes, alors p fait partie de la ligne de crête et celui-ci n'est pas retiré. Ceci est illustré dans la figure 5.10. Remarquons que cette procédure ne marche pas pour les pixels terminaux (feuilles de l'arbre) car il n'y a alors qu'une seule composante connexe, constituée de pixels ayant la valeur 0, présente dans la fenêtre. Avant la procédure d'amincissement, nous incorporons donc au squelette les pixels appartenant aux lignes terminales d'épaisseur 1. Le résultat de l'algorithme est donné dans la figure 5.11.

2. Passage du squelette à un arbre de lignes brisées :

Nous procédons, premièrement, au passage du squelette à un arbre binaire composé de chaînes de pixels. La racine de l'arbre est initialisée au point culminant de l'hypersurface D . Nous procédons, ensuite, au chaînage des pixels jusqu'au premier point d'embranchement détecté. A partir de ce point, la racine correspond à cette première chaîne et deux descendants de cette racine sont initialisés par les deux directions possibles. Une mise en oeuvre récursive nous permet d'obtenir la totalité de l'arbre.

Chaque chaîne de pixel est ensuite vectorisée sous forme d'une ligne brisée. Une largeur, obtenue par la valeur de la distance au bord D , est assignée à chaque

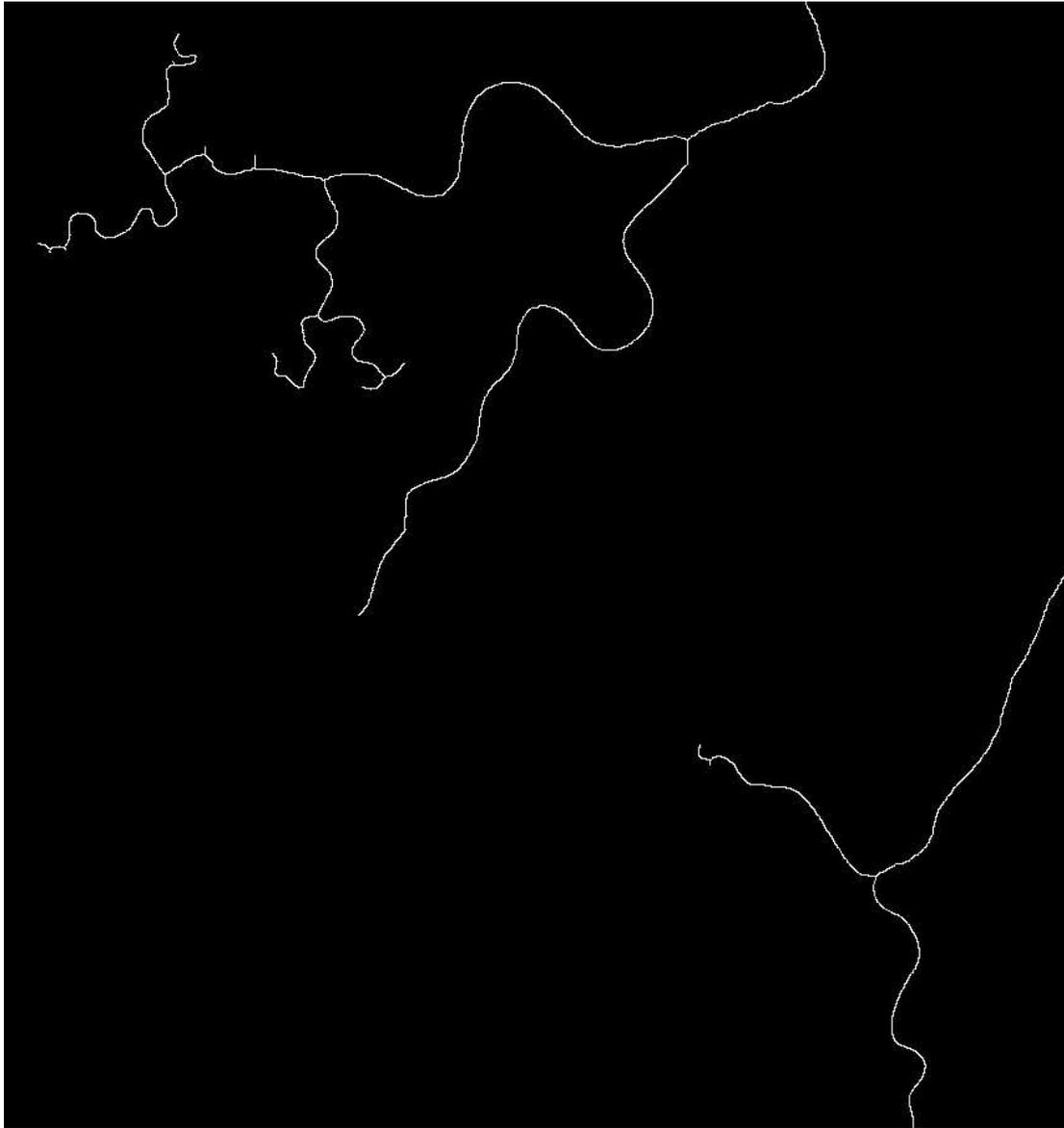


FIG. 5.11 – Squelette des composantes connexes.

point de contrôle.

Nous procédons, enfin, au passage de cet arbre binaire à un arbre n-aire de lignes brisées. La branche principale est initialisée par la vectorisation de la racine à laquelle est concaténée :

- F_1 , son descendant le plus proche au sens de la largeur et de la différence d'orientation entre le vecteur final de l'ascendant et le vecteur initial du descendant ;
- puis F_2 , le descendant de F_1 la plus proche de F_1 ;
- et ainsi de suite jusqu'à l'une des feuilles de l'arbre.

Les descendants mis de côté correspondent à autant de début de branches descendantes de la racine qui sont prolongées de la même façon que la racine.

Étape 2 : prolongement des branches par recuit simulé

Le passage du pixel aux objets nous fournit les lignes brisées correspondant à la partie surfacique des fleuves présents dans la scène observée. La représentation des objets détectés n'est alors que partielle. En effet, comme on peut le voir sur la figure 5.12, les terminaisons des fleuves (*i.e.* à proximité de la source) ne sont pas détectées. Ceci est dû au fait qu'elles correspondent à des éléments linéiques dans l'image. Nous proposons alors de réaliser le prolongement de ces objets par recuit simulé sur un algorithme de Monte Carlo par chaîne de Markov [Robert, 1996].

Nous parcourons l'arbre de façon récursive et pour chaque branche $c = (p, v_1, \dots, v_n)$, où v_i correspond aux paramètres décrivant le segment i , nous proposons d'estimer les paramètres finals $(v_{n+1}, \dots, v_{n'})$ qui maximisent la densité de la configuration à laquelle appartient x conditionnellement à tout ce qui a été détecté. En outre, nous considérons les paramètres initiaux comme fixés. Nous cherchons donc les paramètres finals $\hat{v} = (\hat{v}_{n+1}, \dots, \hat{v}_{n'})$ qui minimisent l'énergie associée à la nouvelle branche $c_{\hat{v}} = (p, v_1, \dots, v_n, \hat{v})$ obtenue par l'ajout de ces paramètres :

$$\hat{v} = \arg \min_v \left[U_1(c_v) + \sum_{p \in (M_{n+1}, \dots, M_{n'})} U_c(p) \right] \quad (5.18)$$

où U_1 , donné par l'équation (5.10), correspond à l'énergie *a priori* sur la forme d'une ligne brisée et son agencement par rapport au reste du réseau ; U_c , donné par l'équation (5.17), au potentiel d'attache aux données fondé sur la mesure du contraste en p ; et M_i est le masque de pixels associé au segment i .

Remarque 18 *En pratique, nous ne cherchons pas à créer des masques associés aux objets surfaciques (obtenus par la phase d'initialisation) pour le calcul de $V_c(p)$. Pour le calcul de $V_c(p)$, nous ne considérons que les masques des nouveaux segments contenant le pixel p . Nous imposons, lors de la construction de ces masques, qu'ils ne comprennent*

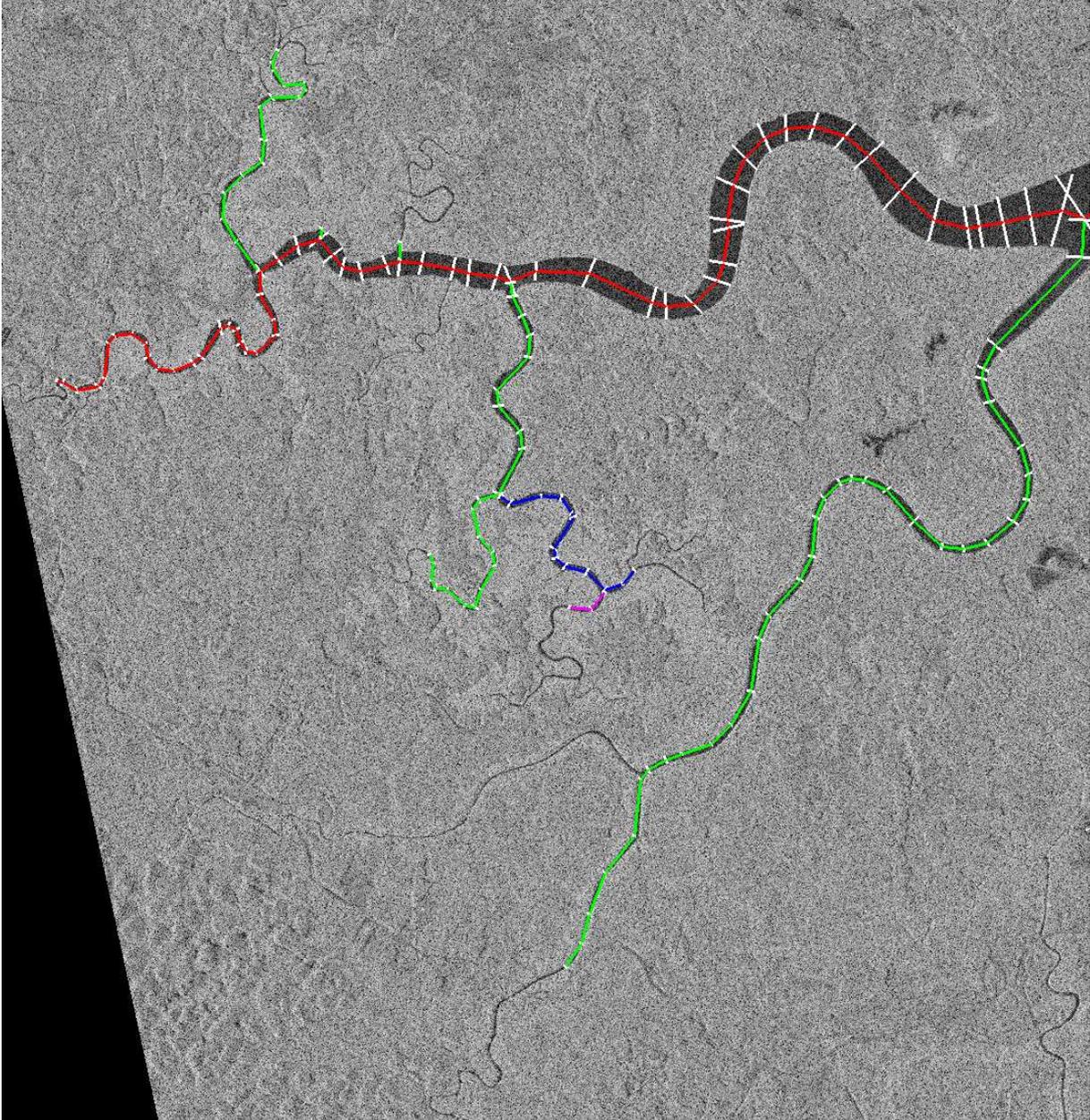


FIG. 5.12 – Arbre n-aire représentant la première composante connexe (en blanc dans la figure 5.5). La ligne rouge correspond à la racine de l'arbre; les lignes vertes correspondent aux filles de la racine; les lignes bleues aux filles des lignes vertes; les lignes violettes aux filles des lignes bleues. Les lignes blanches représentent les largeurs des lignes brisées au niveau de chaque points de contrôle.

pas de pixel appartenant à la silhouette des objets surfaciques, i.e. les pixels appartenant aux composantes connexes détectées dans la phase d'initialisation.

L'optimisation est réalisée sur chaque branche par un **recuit simulé** sur un algorithme d'échantillonnage avec une décroissance adaptative (voir paragraphe 2.6.3). L'algorithme d'échantillonnage est un algorithme de type **Metropolis-Hastings** comprenant uniquement des propositions de perturbations de la ligne brisée à prolonger qui ne modifient pas les paramètres initiaux de cette ligne. Contrairement au cas des processus de lignes brisées décrits dans le chapitre 4, la proposition de ces perturbations n'induit pas un calcul de l'attache aux données sur toute la ligne brisée, mais uniquement sur les nouveaux segments proposés par la perturbation. Nous gagnons ainsi en temps de calcul. Nous utilisons deux types de mouvements réversibles. Le premier concerne l'ajout et le retrait d'un segment à la ligne brisée. Le deuxième est une translation d'un point de contrôle de la ligne brisée, où le vecteur de translation est tiré uniformément dans un carré centré autour de l'origine $(0, 0)$.

Une fois le prolongement effectué, nous retirons de l'arbre les branches de longueur réduite, considérées comme des fausses alarmes.

Les figures 5.13 et 5.14 montrent respectivement le résultat de l'algorithme sur chacun des deux arbres obtenus par l'étape 1 de la phase d'initialisation. Le résultat du prolongement des huit branches du premier arbre a été obtenu en trois heures et dix sept minutes (avec un processeur $2GHz$), soit une moyenne de 25 minutes par branche. Une branche composée d'un seul segment n'a pas été prolongée. Elle est par conséquent retirée de l'arbre final. Le résultat du prolongement des trois branches du deuxième arbre a été obtenu en une heure et vingt minutes. La branche racine (en rouge) ayant été complètement détectée dans la première étape, aucun segment n'est rajouté à la fin de l'optimisation. Pour cette branche, une seconde a suffi pour atteindre la convergence. Les deux autres branches n'ayant pratiquement pas été détectées dans la phase d'initialisation sont obtenues en plus d'une demi-heure. Mis à part le temps de calcul, ces résultats sont très satisfaisants étant donné le bruit de chatoiement présent dans l'image d'entrée et le faible contraste des fins de branches comme le montre la figure 5.15. On notera cependant qu'une des lignes du premier arbre n'a pu être prolongée correctement (la plus grande ligne verte sur la figure 5.13). Ceci est dû à un contraste qui s'atténue fortement lorsque l'on se rapproche de la source du fleuve. Peut-être serait-il judicieux d'incorporer au terme d'attache aux données d'autres informations telles que l'homogénéité d'une branche, une diminution du contraste en fin de branche, ou encore des informations provenant d'autres capteurs.

5.4.2 Génération de nouvelles branches

La phase d'initialisation nous permet d'avoir une représentation objet du réseau surfacique. Plus exactement, le réseau est constitué de plusieurs arbres d'objets. Chaque objet est décrit par une ligne brisée et une silhouette sur l'image issue de la segmenta-

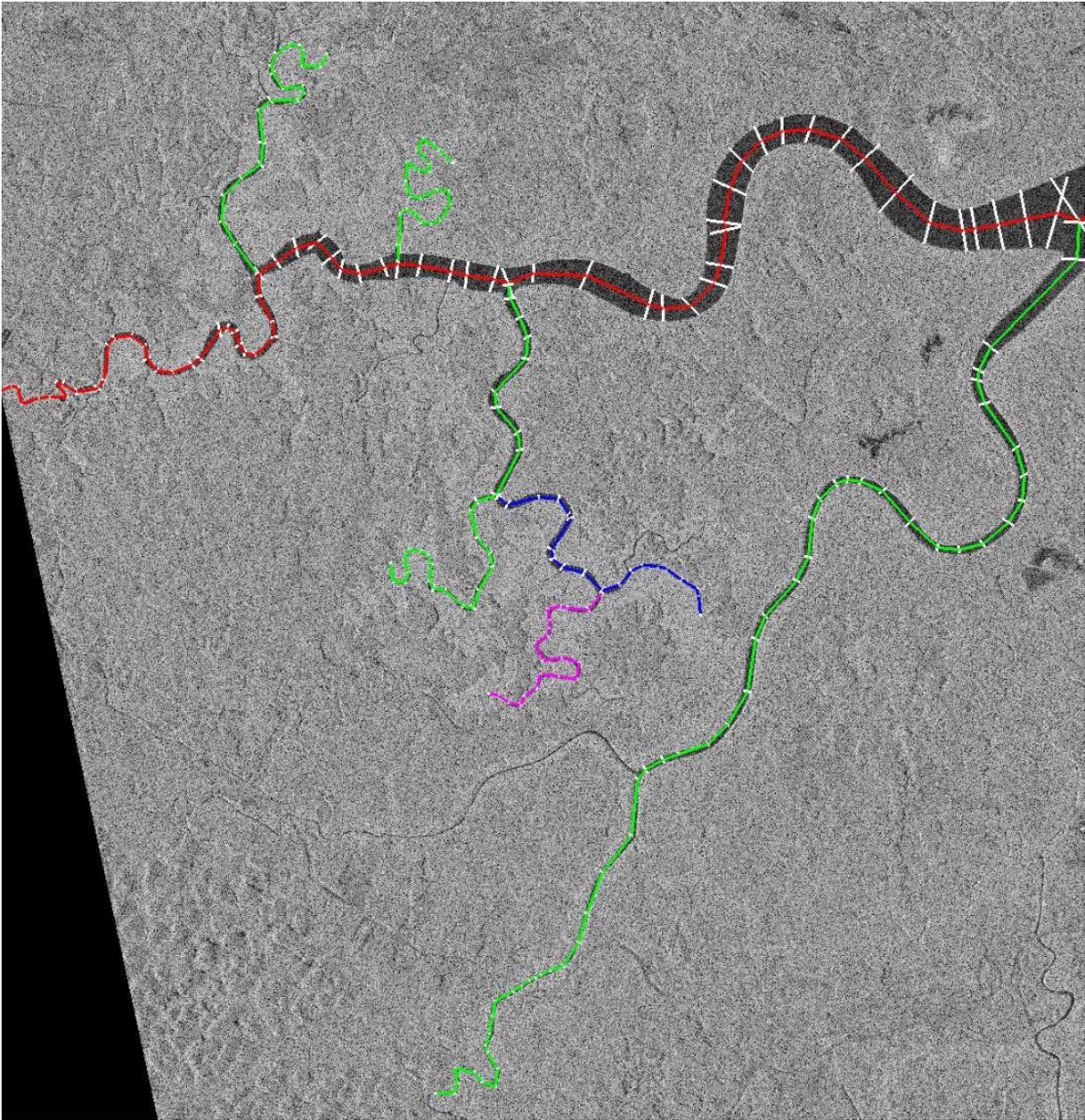


FIG. 5.13 – Prolongement du premier arbre par recuit sur un algorithme de type Metropolis-Hastings.

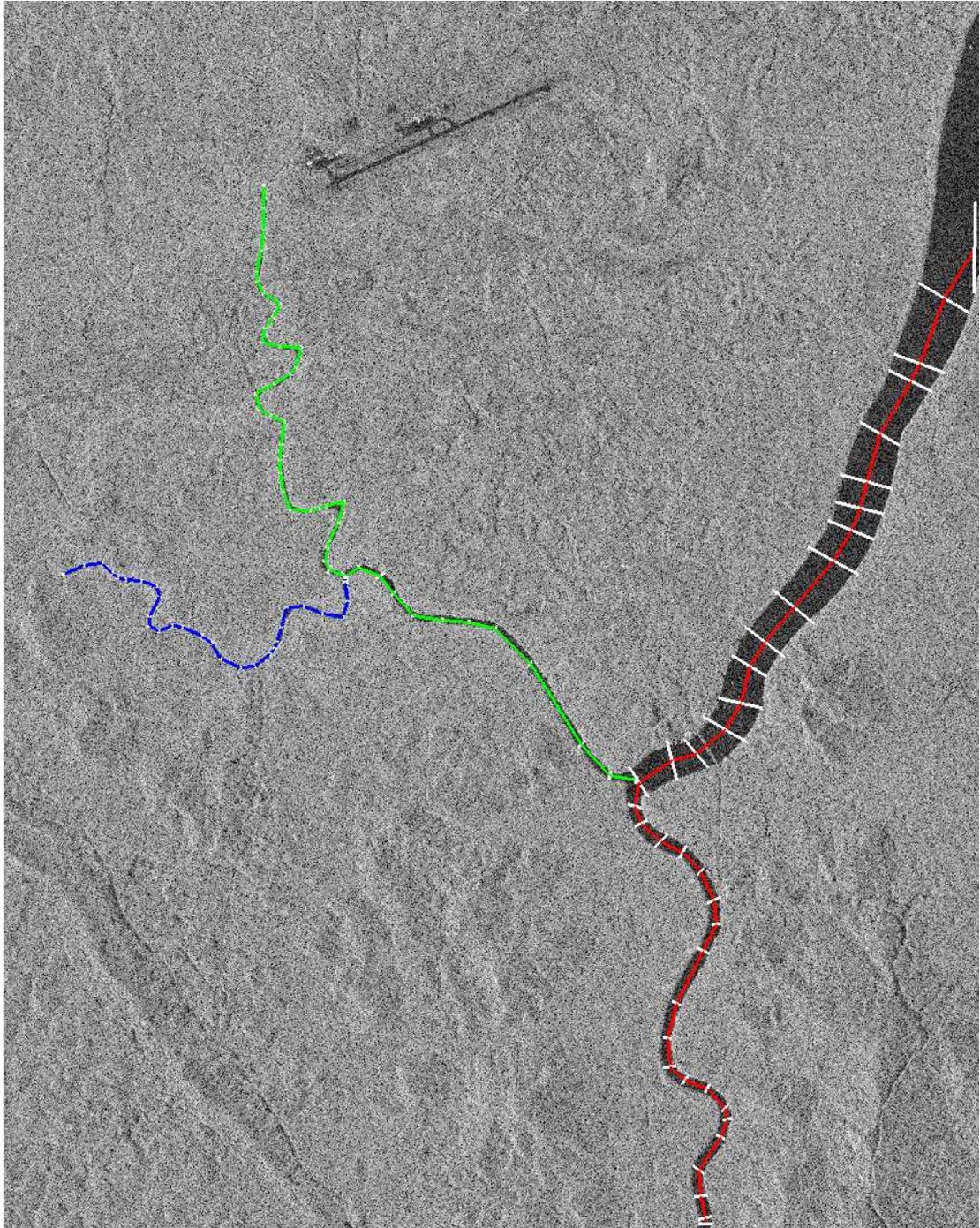


FIG. 5.14 – Prolongement du deuxième arbre par recuit sur un algorithme de type Metropolis-Hastings.

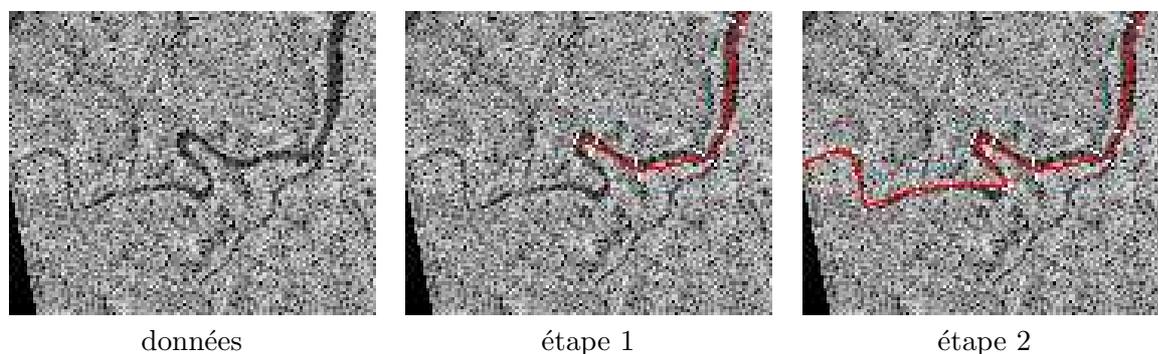


FIG. 5.15 – Zoom permettant de visualiser les résultats de l'étape 1 (passage des composantes connexes aux lignes brisées) et de l'étape 2 (prolongement des lignes brisées) sur la fin d'une branche dont l'extraction est délicate : un faible contraste s'ajoute au bruit de chatoiement de l'image radar ERS utilisée.

tion par champ de Markov. Plus précisément, chaque pixel p des composantes connexes, obtenues après post-traitement du résultat de l'ICM, est assigné à la silhouette de l'objet le plus "proche". L'objet le plus proche d'un pixel p est l'objet o dont la distance entre p et un des segments décrivant o , moins la largeur de o au niveau de la projection orthogonale de p sur ce segment, est minimale.

La modélisation hiérarchique du réseau nous permet de compléter le réseau partiel, obtenu grâce à la phase d'initialisation, par un algorithme récursif permettant la génération de nouvelles branches à partir des branches précédemment obtenues. Cet algorithme s'applique à chacun des arbres initiaux. Il est résumé dans la table 5.3.

<p>Initialisation : c est la racine de l'arbre.</p> <ol style="list-style-type: none"> 1. Génération des nouveaux ascendants de c par recuit simulé. On obtient la configuration \mathbf{x} constituée des anciens et nouveaux descendants de c. 2. Pour chaque $x_i \in \mathbf{x}$, poser $c = x_i$ et aller en 1.
--

TAB. 5.3 – Algorithme de génération de nouvelles branches.

Pour générer les nouvelles branches, il nous faut tout d'abord définir une zone d'influence pour chaque objet dans laquelle nous pourrions générer de nouvelles branches via la définition d'un processus dans cette zone. Pour plus de simplicité, nous optons pour une construction pixélique de cette zone d'influence. Pour un objet de niveau 0

(racine), la zone d'influence est donnée par l'ensemble des pixels voisins de la silhouette de l'objet considéré. Pour un objet de niveau N , la zone d'influence est donnée par les pixels voisins de la silhouette de l'objet considéré qui n'appartiennent pas à la silhouette des objets de niveau inférieur.

Une fois la zone d'influence définie pour un objet donné, la génération des nouveaux descendants se fait conditionnellement à tout ce qui a été détecté. Nous cherchons donc à estimer la configuration \mathbf{x} , contenant les lignes brisées pré-détectées, qui minimise l'énergie U :

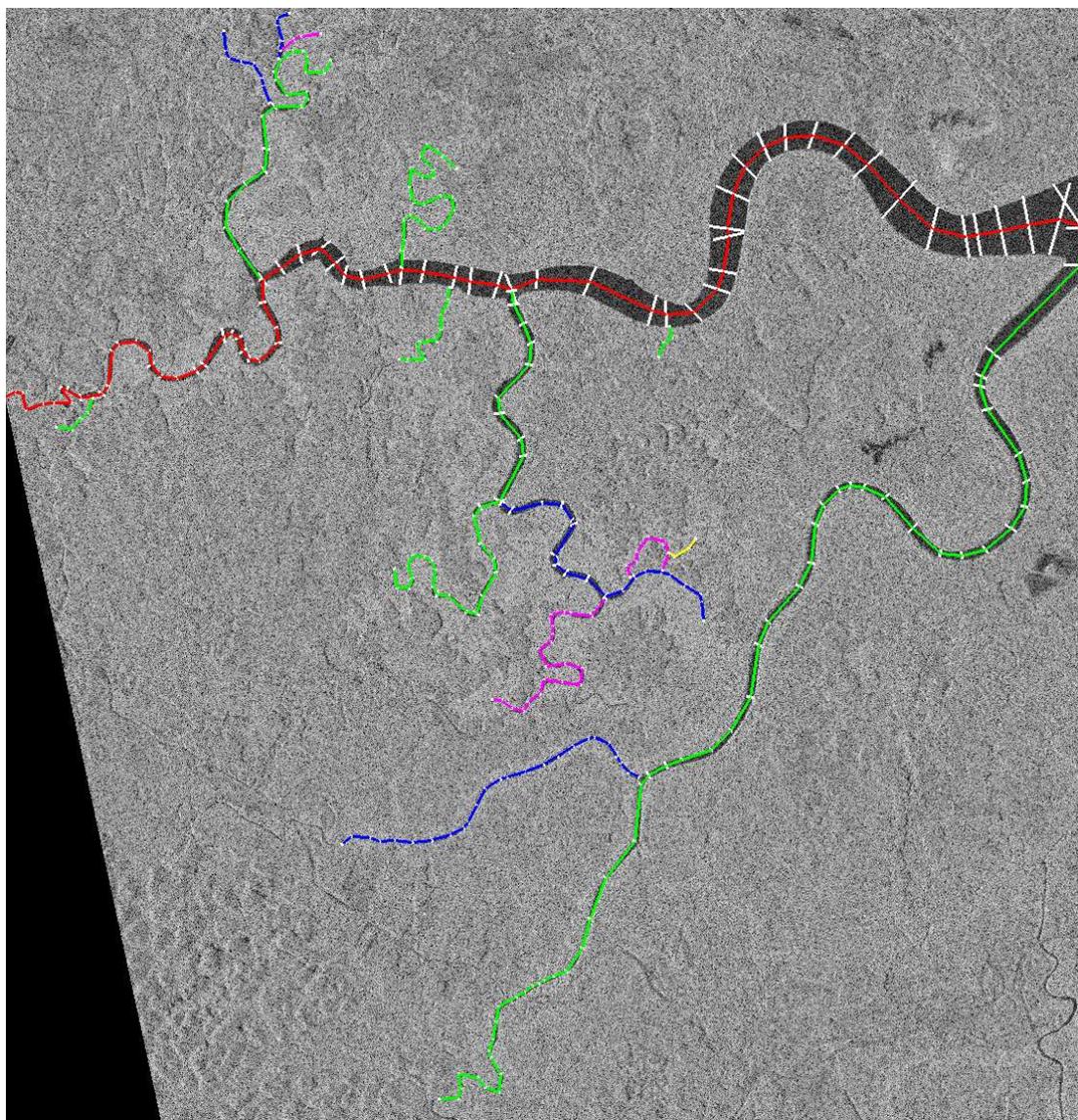
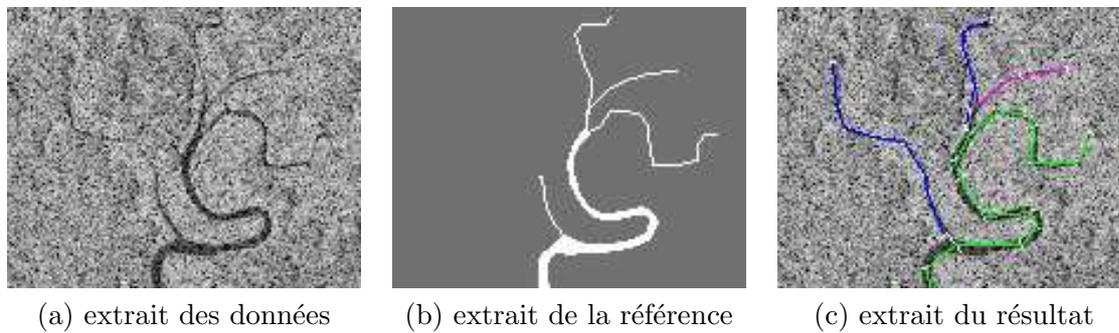
$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \supseteq \mathbf{c}_i} [U_p(\mathbf{x}) + U_d(\mathbf{x})] \quad (5.19)$$

où \mathbf{c}_i est la configuration initiale composée des lignes brisées pré-détectées, U_p est l'énergie *a priori* donnée par l'équation (5.11) et U_d est l'énergie d'attache aux données donnée par l'équation (5.16). Pour cela, nous utilisons un **recuit simulé** sur un algorithme MCMC à sauts réversibles [Geyer et Møller, 1994, Geyer, 1999, Green, 1995]. La décroissance de la température est adaptative et l'arrêt de l'algorithme se fait lorsque l'énergie n'est plus modifiée pendant un nombre fixé d'itérations. L'algorithme d'échantillonnage est un algorithme de type **Metropolis-Hastings** dont le noyau de proposition est composé d'une naissance et mort de lignes brisées contenant un seul segment et des mêmes perturbations que celles utilisées pour le prolongement des branches dans la phase d'initialisation : mouvement d'un point de contrôle ; ajout/retrait d'un segment à la ligne brisée.

L'algorithme de génération des nouvelles branches a été testé sur les deux arbres obtenus par la phase d'initialisation.

Le résultat de cet algorithme appliqué à l'arbre initial donné dans la figure 5.13 est présenté dans la figure 5.16. Il a été obtenu en moins de 20 minutes avec un processeur 3 GHz. Le résultat est satisfaisant dans le sens où une seule branche extraite manuellement n'a pu être détectée par notre algorithme. De plus, il ne présente pratiquement pas de fausses alarmes : une seule des branches détectées par l'algorithme n'est pas présente dans l'extraction manuelle. Enfin, on ne constate pratiquement pas de surdétectations du à un prolongement excessif d'une branche : seules deux branches sont trop prolongées, dont une des deux nous semble plus correcte que le tracé manuel comme le zoom présenté dans la figure 5.16. Outre la branche qui n'avait pu être prolongée dans l'étape 2 de la phase d'initialisation, une seule branche n'a pu être prolongée correctement. Là encore, cela est dû à un affaiblissement du contraste lorsque l'on se rapproche de la source du fleuve.

Aucun nouveau descendant n'a été généré pour l'arbre initial donné dans 5.14. Il n'y avait effectivement pas de nouveaux affluents à extraire.



(d) résultat sur l'arbre 1

FIG. 5.16 – Génération de nouvelles branches à partir du premier arbre. Les images (a), (b) et (c) correspondent à un zoom sur une partie du réseau pour laquelle le réseau extrait automatiquement (b) semble plus exhaustif que celui extrait manuellement (c).

Conclusion

Nous avons abordé des problèmes de modélisation du réseau linéique (en particulier, réseaux routiers ou hydrographiques) dans les images satellitaires et aériennes. Nous avons développé un ensemble de méthodes originales pour l'extraction non supervisée du réseau dans un cadre de géométrie stochastique. Dans cette conclusion, nous présentons une synthèse des travaux effectués, puis nous esquissons des perspectives dans le prolongement de cette étude.

Synthèse des travaux effectués

Les principales contributions de nos travaux concernent l'exploitation des propriétés radiométriques des données et des caractéristiques géométriques et topologiques des réseaux d'intérêt pour une extraction non supervisée des réseaux routiers et hydrographiques. Pour ce faire, nous avons opté pour une modélisation du réseau par processus objet. Plus précisément, trois types de modélisation ont été proposés :

- une modélisation des routes et des rivières de largeur constante par processus de segments (la largeur étant supposée constante sur tout le réseau) ;
- une modélisation des routes et des rivières de largeur constante par processus de lignes brisées (la largeur pouvant varier d'une branche à l'autre du réseau) ;
- une modélisation hiérarchique des fleuves et de leurs affluents fondée sur la modélisation des affluents d'un fleuve par un processus de lignes brisées.

Nous avons pu montrer, au travers de la construction des différents modèles, que les processus objets constituent un outil puissant pour la définition d'un modèle *a priori* incorporant de fortes contraintes géométriques et topologiques. La pertinence de cette modélisation a pu être vérifiée sur de nombreux exemples. Dans leur ensemble, les résultats expérimentaux sont très encourageants. En effet, le réseau obtenu est généralement constitué de longues branches continues et de faible courbure, et présente relativement peu de surdétections et d'omissions, étant donné le caractère complètement automatique de la méthode.

Une contribution non négligeable de ces travaux se situe au niveau de la construction de l'algorithme d'échantillonnage. Si la structure générale de l'algorithme d'échantillonnage reste identique pour chacun des modèles, il n'en est pas de même pour les différents sous-noyaux de proposition de perturbation. Il est, en effet, essentiel de définir des perturbations adaptées au modèle utilisé. L'ajout de perturbations pertinentes nous a

ainsi permis d'obtenir des résultats en un temps raisonnable. Néanmoins, cette approche reste lourde en temps de calcul, notamment si les objets manipulés sont complexes.

Contribution de la méthode d'extraction par processus de segments

La modélisation du réseau par processus de segments nous a permis de développer une méthode d'extraction du réseau linéique sous forme de vecteurs, complètement automatique, à partir d'images aériennes ou satellitaires. Le modèle *a priori* "Quality Candy" s'est avéré particulièrement approprié au cas de l'extraction du réseau routier. En effet, l'utilisation de coefficients relatifs à la qualité pour la relation de connexion conduit à l'obtention d'un réseau de faible courbure et continu. Ce modèle s'adapte de façon encourageante au cas des réseaux plus sinueux comme cela a été montré par les résultats obtenus sur les forêts galeries. De plus, le terme d'attache aux données proposé peut être utilisé pour différents types de données : moyenne et haute résolution, optique et radar. Les résultats ont montré l'intérêt d'utiliser un pré-calcul de l'attache aux données en terme d'efficacité, notamment lorsqu'il est combiné avec une naissance dépendant de ce pré-calcul. Une légère perte en qualité est cependant observée lors de l'utilisation d'un pré-calcul. L'approche bayésienne s'est avérée plus performante que l'approche par champ externe sur des images de résolution moyenne, non perturbées par un bruit géométrique. Par contre, du fait d'un respect plus important des données et d'une modélisation très simple du bruit, elle est plus sensible au bruit géométrique que les approches par champ externe.

Contribution de l'extension aux lignes brisées

Les résultats expérimentaux ont montré que la principale contribution d'une modélisation par lignes brisées est la possibilité de détecter de façon très précise les réseaux sinueux. De plus, les intersections en Y sont mieux détectées grâce à une modélisation des jonctions du réseau au travers d'une interaction de connexion. En terme d'omissions et de surdétectations, les résultats sont équivalents à l'approche par processus de segment. Le principal inconvénient de cette modélisation est le temps de calcul nécessaire à l'extraction du réseau.

Contribution de la méthode d'extraction d'un réseau hiérarchique

La modélisation par champ de Markov s'est révélée très efficace en terme de temps de calcul et pour la détection complète de la partie surfacique du réseau hydrographique sur une image ERS en Guyane. Néanmoins, la modélisation par champ de Markov ne permet pas d'extraire facilement les branches fines du réseau. Nous avons montré sur cet exemple que les processus ponctuels marqués apporte une solution lorsque les limites des approches markoviennes sont atteintes. En effet, cette modélisation objet nous permet de détecter tous les fleuves présents dans l'images et de prolonger complètement la plupart des affluents. Ceci est réalisé de façon efficace grâce à l'utilisation du résultat de la segmentation obtenue par champ de Markov et grâce à l'exploitation de la structure arborescente du réseau hydrographique.

Perspectives

Les perspectives que nous envisageons dans le prolongement de ces travaux de thèse s'articulent autour de trois axes de recherche. En premier lieu, il nous semble intéressant de poursuivre dans le domaine de la modélisation de la scène observée par processus ponctuels marqués. D'autre part, un travail important reste à entreprendre au niveau de l'optimisation, le temps de calcul étant le principal inconvénient de la méthode. Enfin, ce travail ne constitue pas une fin en soi pour la mise à jour ou la production de données cartographiques : il reste à développer des outils permettant l'exploitation de ces résultats par un cartographe ou leur intégration dans des systèmes d'interprétation.

Modélisation de la scène observée par processus ponctuels

Tous les processus de référence utilisés lors de cette étude sont des processus de Poisson homogènes. Il serait intéressant de tester l'approche en utilisant un processus de référence non homogène, dont la mesure d'intensité est fondée sur les données. Ainsi, l'acceptation des objets bien placés pourra se faire de façon naturelle au début de l'algorithme. De même, la loi de probabilité associée aux marques des objets a été choisie uniforme. Une piste à explorer serait de proposer des lois plus proches des connaissances *a priori* sur la forme des objets. Par exemple, pour les lignes brisées, il serait sans doute plus pertinent de définir la densité de probabilité de l'orientation d'un segment en fonction de l'orientation du segment précédent (par exemple, une densité gaussienne centrée sur la valeur de l'orientation précédente).

Les processus de segments ne permettent pas, pour l'instant, de modéliser correctement les réseaux à largeur variable, comme, par exemple, un réseau routier composé de routes et d'autoroutes. Pour l'extraction d'un tel réseau, il serait intéressant, comme cela est déjà fait pour les lignes brisées, de rajouter une marque correspondant à l'épaisseur d'un segment dans le modèle. L'épaisseur d'une branche étant supposée varier de façon progressive (rivière), voire pas du tout (route), une idée serait alors de prendre en compte la différence d'épaisseur de deux segments connectés dans le potentiel de connexion. Pour ce qui est des lignes brisées, une largeur variable au sein d'une même ligne brisée serait sans doute plus adaptée à la modélisation des rivières et fleuves qui sont de largeur plus fine à la source qu'à l'embouchure.

Concernant l'exploitation des propriétés radiométriques des données, nous n'avons, pour l'instant, utilisé qu'une seule source d'information. Dans le cas d'une approche par champ externe sans pré-calculs, les informations utilisées sont les données radiométriques issues d'une seule image. Dans le cas d'une approche par champ externe avec pré-calculs et dans le cas de l'approche bayésienne, les informations utilisées proviennent d'un détecteur de structures linéaires appliqué à une seule image. Afin d'améliorer la qualité du réseau extrait par l'algorithme, nous envisageons dans un futur proche de travailler dans un cadre de fusion de données et donc de bénéficier de l'apport de plusieurs sources : données multi-capteurs, multi-bandes, multi-dates,

multi-résolutions. De plus, il serait sûrement pertinent de combiner plusieurs types de détecteurs (détecteurs de lignes, d'intensité, etc.) pour améliorer la qualité des extractions fondées sur des pré-calculs.

A plus long terme, nous prévoyons une modélisation complète de la scène observée, via l'introduction de différents types d'objets dans le modèle, correspondant par exemple aux routes, rivières, maisons et arbres présents dans l'image.

Enfin, cette modélisation du réseau linéique pour l'extraction du réseau routier et hydrographique pourrait être adaptée à d'autres applications telles que :

- la détection de contours, en modifiant le terme d'attache aux données ;
- la détection de réseaux subsurfaciques à partir d'images radar ERS ;
- la reconnaissance des discontinuités physiques de l'espace géologique ;
- l'extraction des vaisseaux sanguins à partir de données volumiques.

Optimisation

Nous avons vu que le principal défaut de l'approche proposée est le temps de calcul nécessaire à la convergence du recuit simulé. En effet, pour espérer obtenir un réseau complet, il est important de ne pas décroître trop vite en température au niveau de certaines températures critiques. Le choix d'une décroissance adaptative se fondant sur le comportement de l'algorithme, semble donc pertinent pour ne passer lentement qu'au niveau des températures critiques. Le schéma adaptatif que nous avons utilisé est un des plus simples proposés dans la littérature et une étude plus poussée doit être menée dans ce domaine pour une accélération importante de la convergence, sans perte de qualité au niveau du résultat de l'extraction.

Mis à part pour l'extraction du réseau hiérarchique, tous les algorithmes d'extraction ont été initialisés avec une configuration vide. Ce choix a été fait pour montrer que l'extraction n'était pas sensible à l'initialisation. Mais il est évident que pour gagner en efficacité, une initialisation pertinente pourrait être judicieuse. Nous espérons ainsi pouvoir partir à température basse et obtenir un résultat plus rapidement.

Afin de gagner en efficacité, il pourrait également être intéressant de proposer une optimisation déterministe à partir d'une certaine température.

Production et mise à jour cartographique

L'application finale de ces travaux pourrait être la production ou la mise à jour de cartes. Pour en arriver là, de nombreux aspects restent à explorer.

Pour faciliter l'exploitation de ces résultats, il est essentiel de développer des outils d'auto-évaluation du réseau extrait permettant la réalisation d'un post-traitement

et permettant à un utilisateur de se concentrer sur les parties ambiguës du réseau. Ceci pourra être fait par une analyse des propriétés radiométriques et de texture des différentes zones de chaque partie du réseau, conjointement à une analyse des caractéristiques géométriques et topologiques des différentes branches obtenues. Certains critères nous permettront ainsi de proposer des fusions de branches qui n'ont pu être réalisées lors de l'extraction automatique et de retirer certaines fausses alarmes. Ces critères pourront également être utilisés pour indiquer le degré de confiance accordé à telle ou telle partie du réseau.

Pour la mise à jour cartographique, il faut tout d'abord se poser la question de la mise en correspondance du réseau avec la carte géographique existante. Ce problème est loin d'être évident, étant donné les distorsions introduites par les cartographes dans un souci de généralisation. Il sera important de définir des mesures de performance. Elle pourront, d'une part, être utilisées pour améliorer itérativement la qualité de la mise en correspondance et, d'autre part, fournir des informations utiles pour la mise à jour cartographique.

Enfin, il reste à étudier la possibilité d'incorporer des informations cartographiques dans l'algorithme d'extraction. On pourra, par exemple, initialiser l'algorithme par le réseau issu de la carte et proposer une optimisation en prenant en compte une mesure de distance entre le réseau courant et le réseau issu de la carte.

Publications

Rapports de Recherche

1. C. Lacoste, X. Descombes, et J. Zerubia. A comparative study of point processes for line network extraction in remote sensing. Rapport de Recherche 4516, INRIA, Sophia Antipolis, France, juillet 2002.

Journaux français

2. C. Lacoste, X. Descombes, J. Zerubia, et N. Baghdadi. Extraction automatique des réseaux linéiques à partir d'images satellitaires et aériennes par processus markov objet. *Bulletin de la S.F.P.T.*, 170:13–22, 2003.

Journaux internationaux

3. C. Lacoste, X. Descombes, et J. Zerubia. Point Processes for Unsupervised Line Network Extraction in Remote Sensing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. **Soumis en 2003**, révisions envoyées fin août 2004.

Conférences francophones avec actes

4. C. Lacoste, X. Descombes, J. Zerubia, et N. Baghdadi. Extraction de réseaux linéiques à partir d'images satellitaires par processus Markov objet. *Actes du colloque GRETSI*, Paris, France, 2003.

Conférences internationales avec actes

5. C. Lacoste, X. Descombes, et J. Zerubia. Road network extraction in remote sensing by a Markov object process. *Proceedings of IEEE ICIP*, Barcelona, Spain, Septembre 2003.
6. X. Descombes, F. Kruggel, C. Lacoste, M. Ortner, G. Perrin, et J. Zerubia. Marked point process in image analysis : from context to geometry. *Proceedings of*

International Conference on Spatial Point Process Modelling and its Application (SPPA), Castellon, Espagne, Avril 2004.

7. C. Lacoste, X. Descombes, J. Zerubia, et N. Baghdadi. A Bayesian geometric model for line network extraction from satellite images. *Proceedings of ICASSP*, Montréal, Québec, Canada, mai 2004.
8. C. Lacoste, X. Descombes, J. Zerubia, et N. Baghdadi. Unsupervised line network extraction from remotely sensed images by polyline process. *Proceedings of EUSIPCO*, Vienne, Autriche, septembre 2004.

Séminaires

9. C. Lacoste, X. Descombes, et J. Zerubia. A comparative study of point processes for line network extraction from images. *Mini-Symposium in Stochastic Geometry*, CWI, Amsterdam, Pays-Bas, mai 2002.
10. C. Lacoste, X. Descombes, et J. Zerubia. Extraction du réseau linéique en télédétection par Processus Markov Objet. BRGM, Orléan, France, octobre 2002.
11. C. Lacoste, X. Descombes, et J. Zerubia. Extraction du Réseau Linéique en Télédétection par Processus Markov Objet, *Séminaire Ariana - Mistral: "Stratégies stochastiques d'exploration d'état appliquées au traitement d'image et à la modélisation de réseaux*, INRIA Sophia-Antipolis, France, mai 2003.
12. C. Lacoste, X. Descombes, et J. Zerubia. Line network extraction in remote sensing by spatial processes. *Symposium DLR-Ariana*, DLR, Oberpfaffenoffen, Allemagne, novembre 2003.
13. C. Lacoste, X. Descombes, et J. Zerubia. CAROLINE: a CARTographic Oriented Line Network Extraction model. *Spatial Point Process Modelling and its Applications*, Castellon, Espagne, avril 2004.

Bibliographie

1. [Airault et Jamet, 1994] S. Airault et O. Jamet. Détection automatique du réseau routier sur des images aériennes. *Actes du congrès RFIA*, volume 1, pages 519–531, Paris, France, janvier 1994.
2. [Azencott, 1992] Azencott. Markov field approach : parameter estimation by qualitative boxes. Cours : Les Houches, 1992.
3. [Baddeley et van Lieshout, 1993] A. Baddeley et M. N. M. van Lieshout. Stochastic geometry models in high-level vision. *Statistics and Images*, 1:233–258, 1993.
4. [Barzohar et Cooper, 1996] M. Barzohar et D. B. Cooper. Automatic finding of main roads in aerial images by using geometric-stochastic models and estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:707–721, juillet 1996.
5. [Baumgartner et al., 1996] A. Baumgartner, C. Steger, C. Wiedmann, H. Mayer, W. Eckstein et H. Ebner. Update of roads in GIS from aerial imagery: verification and multiresolution extraction. *International Archive of Photogrammetry and Remote Sensing*, 31(3):53–58, 1996.
6. [Baumgartner et al., 1999] A. Baumgartner, C. Steger, H. Mayer, W. Eckstein et H. Ebner. Automatic road extraction based on multi-scale, grouping, and context. *Photogrammetric Engineering and Remote Sensing*, 65(7):777–785, juillet 1999.
7. [Besag, 1986] J. Besag. On the statistical analysis of dirty pictures. *Journal of Royal Statistic Society*, B(68):259–302, 1986.
8. [Bhattacharya et Parui, 1997] U. Bhattacharya et S.K. Parui. An improved back-propagation neural network for detection of road-like features in satellite imagery. *International Journal of Remote Sensing*, 18:3379–3394, avril 1997.
9. [Bicego et al., 2003] M. Bicego, S. Dalfini, G. Vernazza et P. Murino. Automatic road extraction from aerial images by probabilistic contour tracking. *Proceedings of IEEE Int. Conf. on Image Processing*, volume III, pages 585–588, Barcelone, Espagne, septembre 2003.
10. [Bobillet et al., 2003] W. Bobillet, J.P. Da Costa, C. Germain, O. Lavialle et G. Grenier. Contours actifs : application à la détection de rangs de cultures en télédétection haute résolution. *Actes du colloque GRETSI*, Paris, France, septembre 2003.

11. [Bresenham, 1965] J. E. Bresenham. Algorithm for computer control of a digital plotter. *IBM Systems Journal*, 4(1):25–30, 1965.
12. [Canny, 1986] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, novembre 1986.
13. [Clifford et Nicholls, 1994] P. Clifford et G. Nicholls. Comparison of birth-and-death and Metropolis-Hastings Markov chain Monte Carlo for the Strauss process. Rapport de recherche, Statistics Department, Oxford University, juin 1994.
14. [Couloigner et Ranchin, 2000] I. Couloigner et T. Ranchin. Mapping of urban areas: A multiresolution modeling approach for semi-automatic extraction of streets. *Photogrammetric Engineering and Remote Sensing*, 66(7):867–874, juillet 2000.
15. [Deriche, 1987] R. Deriche. Using Canny’s criteria to derive a recursively implemented optimal edge detector. *International Journal of Computer Vision*, 1(2):167–187, 1987.
16. [Descombes *et al.*, 1995] X. Descombes, J.F. Mangin, E. Pechersky et M. Sigelle. Fine structure preserving Markov model for image processing. *9th Scandinavian Conference on Image Analysis*, pages 349–356, Uppsala, Suède, juin 1995.
17. [Destival, 1987] I. Destival. Recherche automatique des réseaux linéaires sur les images SPOT. *Bulletin de la S.F.P.T.*, 105:5–16, 1987.
18. [Dhéréte et Desachy, 1999] P. Dhéréte et J. Desachy. Data fusion for linear geographic feature matching on SPOT images. *Bulletin de la S.F.P.T.*, 153:88–90, avril 1999.
19. [Doucette *et al.*, 2001] P. Doucette, P. Agouris, A. Stefanidis et M. Musavi. Self-organized clustering for road extraction in classified imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 55:347–358, 2001.
20. [Duda et Hart, 1973] R.O. Duda et P.E Hart. *Pattern Classification and Scene Analysis*. John Wiley & Sons, NY, USA, 1973.
21. [Eberly *et al.*, 1994] D. Eberly, R. Gardner, B. Morse, S. Pizer et C. Scharlach. Ridges for image analysis. *Journal of Mathematical Imaging and Vision*, 4(4):353–373, décembre 1994.
22. [Fischler *et al.*, 1981] M. A. Fischler, J. M. Tenenbaum et H. C. Wolf. Detection of roads and linear structures in low-resolution aerial imagery using a multi-source knowledge integration technique. *Computer Graphics and Image Processing*, 15:201–223, 1981.
23. [Fua et Leclerc, 1990] P. Fua et Y. G. Leclerc. Model driven edge detection. *Machine Vision and Applications*, 3:45–56, 1990.
24. [Garcin *et al.*, 2001] L. Garcin, X. Descombes, J. Zerubia et H. Le Men. Building detection by Markov object processes and a MCMC algorithm. Rapport de Recherche 4206, INRIA, Sophia Antipolis, France, juin 2001.

25. [Geman et Geman, 1984] S. Geman et D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.
26. [Geman et Jedynak, 1996] D. Geman et B. Jedynak. An active testing model for tracking roads in satellite images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:1–14, 1996.
27. [Géraud, 2003] T. Géraud. Fast road network extraction in satellite images using mathematical morphology and Markov random fields. *IEEE - EURASIP Workshop on Nonlinear Signal and Image Processing*, juin 2003.
28. [Geyer et Møller, 1994] C. J. Geyer et J. Møller. Simulation and likelihood inference for spatial point process. *Scandinavian Journal of Statistics, Series B*, 21:359–373, 1994.
29. [Geyer, 1999] C.J. Geyer. *Stochastic Geometry: Likelihood and Computation*, chapitre 3, pages 79–140. Chapman & Hall/CRC, 1999.
30. [Green, 1995] P.J. Green. Reversible jump Markov chain Monte-Carlo computation and Bayesian model determination. *Biometrika*, 57:97–109, 1995.
31. [Grün et Li, 1995] A. Grün et H. Li. Road extraction from aerial and satellite images by dynamic programming. *ISPRS Journal of Photogrammetry and Remote Sensing*, 50(4):11–20, août 1995.
32. [Guigues et Vilgino, 2000] L. Guigues et J.-M. Vilgino. Automatic road extraction through light propagation simulation. *International Archive of Photogrammetry and Remote Sensing*, volume XXXIII, Amsterdam, Pays-Bas, 2000.
33. [Gurney, 1980] C. M. Gurney. Threshold selection for line detection algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, 18:04–211, 1980.
34. [Harvey, 1999] W. A. Harvey. Performance evaluation for road extraction. *Bulletin de la S.F.P.T.*, 153:79–87, avril 1999.
35. [Hastings, 1970] W. K. Hastings. Monte Carlo sampling using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.
36. [Haverkamp, 2002] D. Haverkamp. Extracting straight road structure in urban environments using IKONOS satellite imagery. *Optical Engineering*, 41(09):2107–2110, septembre 2002.
37. [Heipke et al., 1995] C. Heipke, C. Steger et R. Multhammer. A hierarchical approach to automatic road extraction from aerial imagery. David M. McKeown Jr. et Ian J. Dowman, éditeurs, *Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision II, Proc. SPIE*, volume 2486, pages 222–231, 1995.
38. [Hivernat et al., 1999] C. Hivernat, S. Randriamasy, X. Descombes et J. Zerubia. Qualification automatique des résultats d’une mise en correspondance de réseaux routiers en vue de la mise à jour cartographique. *Bulletin de la S.F.P.T.*, 153:91–93, 1999.

39. [Hoffmann *et al.*, 1991] K.H. Hoffmann, D. Würtz, C. Groot et M. Hanf. Concepts in optimizing simulated annealing schedules: an adaptive approach for parallel and vector machines. M. Grauer et D.B. Pressmar, éditeurs, *Parallel and Distributed Optimization*. Springer Verlag, 1991.
40. [Huber et Lang, 2001] R. Huber et K. Lang. Road extraction from high-resolution airborne SAR using operator fusion. *Proceedings of International Geoscience and Remote Sensing Symposium*, Sydney, Australie, juillet 2001.
41. [Imberty et Descombes, 2000] M. Imberty et X. Descombes. Simulation de processus objets : Etude de faisabilité pour une application à la segmentation d'images. Rapport de Recherche 3881, INRIA, Sophia Antipolis, France, 2000.
42. [Kendall et Møller, 2000] W. S. Kendall et J. Møller. Perfect Metropolis-Hastings simulation of locally stable spatial point processes. *Advances in Applied Probability*, 32:844–865, 2000.
43. [Kerstan *et al.*, 1978] J. Kerstan, K. Matthes et J. Mecke. *Infinitely divisible point processes*. Wiley, Chichester, 1978.
44. [Koller *et al.*, 1995] T.M. Koller, G. Gerig, G. Székely et D. Dettwiler. Multiscale detection of curvilinear structures in 2-D and 3-D image data. *Proceedings of IEEE Int. Conf. on Computer Vision*, pages 864–869, Cambridge, Massachusetts, juin 1995.
45. [Lacoste *et al.*, 2002] C. Lacoste, X. Descombes et J. Zerubia. A comparative study of point processes for line network extraction in remote sensing. Rapport de Recherche 4516, INRIA, Sophia Antipolis, France, juillet 2002.
46. [Laptev *et al.*, 2000] I. Laptev, T. Lindeberg, W. Eckstein, C. Steger et A. Baumgartner. Automatic extraction of roads from aerial images based on scale space and snakes. *Machine Vision and Applications*, 12:23–31, 2000.
47. [Lindeberg, 1998] T. Lindeberg. Edge detection and ridge detection with automatic scale selection. *International Journal of Computer Vision*, 30(2):77–116, 1998.
48. [Merlet et Zerubia, 1996] N. Merlet et J. Zerubia. New prospects in line detection by dynamic programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(4):426–431, 1996.
49. [Metropolis *et al.*, 1953] M. Metropolis, A. W. Rosenbluth, A. H. Teller et E. Teller. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21:1087–1092, 1953.
50. [Neuenschwander *et al.*, 1997] W. M. Neuenschwander, P. Fua, L. Iverson, G. Székely et O. Kubler. Ziplock snakes. *International Journal of Computer Vision*, 25(3):191–201, 1997.
51. [Ortner, 2001] M. Ortner. Extraction de caricatures de bâtiments sur des modèles numériques d'élévation. Rapport de DEA, INRIA, Sophia Antipolis, France, août 2001.

52. [Pérez *et al.*, 2001] P. Pérez, A. Blake et M. Gangnet. Jetstream: probabilistic contour extraction with particles. *Proceedings of IEEE Int. Conf. on Computer Vision*, volume II, pages 524–531, 2001.
53. [Perrin *et al.*, 2004] G. Perrin, X. Descombes et J. Zerubia. Tree crown extraction using marked point processes. *Proceedings of EUSIPCO*, Vienne, Autriche, septembre 2004.
54. [Peskun, 1973] P. H. Peskun. Optimum Monte Carlo sampling using Markov chains. *Biometrika*, 60:607–612, 1973.
55. [Péteri *et al.*, 2001] R. Péteri, I. Couloigner et T. Ranchin. A multiresolution modelling approach for semi-automatic extraction of streets: application to high resolution images from the Ikonos satellite. *Proceedings of the EARSeL/SFPT Symposium, "Observing our environment from space: new solutions for a new millenium"*, Marne-la-Vallée, France, mai 2001.
56. [Poli et Valli, 1996] R. Poli et G. Valli. An algorithm for real-time vessel enhancement and detection. *Computer Methods and Programs in Biomedicine*, 52:1–22, 1996.
57. [Preston, 1976] C. Preston. Spatial birth and death processes. *Bulletin of the International Statistical Institute*, 46(2):371–391, 1976.
58. [Rellier *et al.*, 2002] G. Rellier, X. Descombes et J. Zerubia. Local registration and deformation of a road cartographic database on a SPOT satellite image. *Pattern Recognition*, 35(10):2213–2222, octobre 2002.
59. [Ripley, 1977] B. D. Ripley. Modelling spatial patterns. *Journal of the Royal Statistical Institute, Series B*, 39:172–212, 1977.
60. [Robert et Casella, 1999] C. Robert et G. Casella. *Monte Carlo Statistical Methods*, chapitre 8. Springer-Verlag, 1999.
61. [Robert, 1996] C. Robert. *Méthodes de Monte Carlo par chaînes de Markov*. Statistique mathématique et probabilité. Economica, 1996.
62. [Rochery *et al.*, 2003] M. Rochery, I. H. Jermyn et J. Zerubia. Contours actifs d'ordre supérieur appliqués à la détection de linéiques dans des images de télédétection. Rapport de Recherche 5063, INRIA, France, décembre 2003.
63. [Roux, 1992] L. Roux. *Recalage d'images multi-sources. Application au recalage d'une image SPOT et d'une carte*. Thèse de Doctorat, ENST, Paris, France, 1992.
64. [Rue et Hurn, 1999] H. Rue et M. Hurn. Bayesian object identification. *Biometrika*, 3:649–660, 1999.
65. [Ruelle, 1970] D. Ruelle. Superstable interactions in classical statistical mechanics. *Communication in Mathematical Physics*, 18:127–159, 1970.
66. [Ruskoné, 1996] R. Ruskoné. *Extraction automatique du réseau routier par interprétation locale du contexte*. Thèse de Doctorat, Université de Marne la Vallée, 1996.
67. [Sampère, 2001] J-P. Sampère. Evaluation de l'utilisation de données supermode SPOT5 pour la mise à jour de BD IGN. *Bulletin de la S.F.P.T.*, 164:96–105, 2001.

68. [Serendero, 1989] M. A. Serendero. *Extraction d'Informations Symboliques en Imagerie SPOT : Réseaux de Communication et Agglomérations*. Thèse de Doctorat, Université de Nice - Sophia Antipolis, 1989.
69. [S.Hinz et Baumgartner, 2003] S.Hinz et A. Baumgartner. Automatic extraction of urban road networks from multi-view aerial imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(1-2):83–98, juin 2003.
70. [Steger, 1996] C.T. Steger. Extracting curvilinear structures: A differential geometric approach. *Proceedings of European Conf. on Computer Vision*, volume 1064 of *Lecture Notes in Computer Science*, pages 630–641. Springer-Verlag, 1996.
71. [Steger, 1998] C.T. Steger. An unbiased detector of curvilinear structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(2):113–125, février 1998.
72. [Stoica et al., 2004] R. Stoica, X. Descombes et J. Zerubia. A Gibbs point process for road extraction in remotely sensed images. *International Journal of Computer Vision*, 57(2):121–136, 2004.
73. [Stoica, 2001] R. Stoica. Processus ponctuels pour l'extraction des réseaux linéiques dans les images satellitaires et aériennes. Thèse de Doctorat, Université de Nice - Sophia Antipolis, février 2001.
74. [Stoyan et al., 1987] D. Stoyan, W.S. Kendall et J. Mecke. *Stochastic Geometry and its Applications*. John Wiley and Sons, Chichester, 1987.
75. [Tupin et al., 1998] F. Tupin, H. Maitre, J-F. Mangin, J-M. Nicolas et E. Pechersky. Detection of linear features in SAR images: Application to road network extraction. *IEEE Transactions on Geoscience and Remote Sensing*, 36(2):434–453, 1998.
76. [Urago et al., 1994] S. Urago, J. Zerubia et M. Berthod. A Markovian model for contour grouping. Rapport de Recherche 2122, INRIA, Sophia Antipolis, France, février 1994.
77. [van Lieshout et Stoica, 2001] M.N.M. van Lieshout et R.S. Stoica. The Candy model revisited: Markov properties and inference. Rapport de Recherche PNA-R0115, CWI, Amsterdam, Pays-Bas, 2001.
78. [van Lieshout, 1993] M.N.M. van Lieshout. Stochastic annealing for nearest-neighbour point processes with application to object recognition. Rapport de Recherche BS-R9306, CWI, Amsterdam, Pays-Bas, 1993.
79. [van Lieshout, 2000] M.N.M. van Lieshout. *Markov Point Processes and their Applications*. Imperial College Press, 2000.
80. [Vermaak et al., 2003] J. Vermaak, N. Lawrence et P. Pérez. Variational inference for visual tracking. *Proceedings of IEEE Int. Conf. on Computer Vision and Pattern Recognition*, Madison, Wisconsin, juin 2003.
81. [Vosselman et de Knecht, 1995] G. Vosselman et J. de Knecht. Road tracing by profile matching and Kalman filtering. *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, pages 265–274, Ascona, Suisse, avril 1995.

82. [Wang *et al.*, 1996] D. Wang, D.C. He, L. Wang et D. Morin. Extraction du réseau routier urbain à l'aide d'images SPOT HRV. *International Journal of Remote Sensing*, 17(4):827–833, 1996.
83. [Wang et Pavlidis, 1993] L. Wang et T. Pavlidis. Detection of curved and straight segments from gray scale topography. *CVGIP : Image Understanding*, 58(3):352–365, novembre 1993.
84. [Winkler, 2003] G. Winkler. *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods: a Mathematical Introduction*. seconde édition, Springer-Verlag, 2003.
85. [Zhang *et al.*, 1999] C. Zhang, S. Murai et E. P. Baltsavias. Road network detection by mathematical morphology. *Bulletin de la S.F.P.T.*, 153:94–96, 1999.
86. [Zhang, 2004] C. Zhang. Towards an operational system for automated updating of road databases by integration of imagery and geodata. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(3-4):127–258, janvier 2004.
87. [Zlotnick et Carnine, 1993] A. Zlotnick et P. Carnine. Finding road seeds in aerial images. *Computer Vision, Graphics, and Image Processing*, 57:243–260, 1993.

Résumé

Cette thèse aborde le problème de l'extraction non supervisée des réseaux linéiques (routes, rivières, etc.) à partir d'images satellitaires et aériennes. Nous utilisons des processus objet, ou processus ponctuels marqués, comme modèles *a priori*. Ces modèles permettent de bénéficier de l'apport d'un cadre stochastique (robustesse au bruit, corpus algorithmique, etc.) tout en manipulant des contraintes géométriques fortes. Un recuit simulé sur un algorithme de type Monte Carlo par Chaîne de Markov (MCMC) permet une optimisation globale sur l'espace des configurations d'objets, indépendamment de l'initialisation.

Nous proposons tout d'abord une modélisation du réseau linéique par un processus dont les objets sont des segments interagissant entre eux. Le modèle *a priori* est construit de façon à exploiter au mieux la topologie du réseau recherché au travers de potentiels fondés sur la qualité de chaque interaction. Les propriétés radiométriques sont prises en compte dans un terme d'attache aux données fondé sur des mesures statistiques.

Nous étendons ensuite cette modélisation à des objets plus complexes. La manipulation de lignes brisées permet une extraction plus précise du réseau et améliore la détection des bifurcations.

Enfin, nous proposons une modélisation hiérarchique des réseaux hydrographiques dans laquelle les affluents d'un fleuve sont modélisés par un processus de lignes brisées dans le voisinage de ce fleuve.

Pour chacun des modèles, nous accélérons la convergence de l'algorithme MCMC par l'ajout de perturbations adaptées.

La pertinence de cette modélisation par processus objet est vérifiée sur des images satellitaires et aériennes, optiques et radar.

Mots-clés : Géométrie stochastique, processus ponctuels marqués, recuit simulé, MCMC à sauts réversibles, extraction de réseaux linéiques, images satellitaires et aériennes.

Abstract

This thesis addresses the problem of the unsupervised extraction of line networks (roads, rivers, etc.) from remotely sensed images. We use object processes, or marked point processes, as prior models. These models benefit from a stochastic framework (robustness w.r.t. noise, algorithms, etc.) while incorporating strong geometric constraints. Optimization is done via simulated annealing using a Reversible Jump Markov Chain Monte Carlo (RJMCMC) algorithm, without any specific initialization.

We first propose to model line networks by a process whose objects are interacting line segments. The prior model is designed to exploit as fully as possible the topological properties of the network under consideration through potentials based on the quality of each interaction. The radiometric properties of the network are modeled using a data term based on statistical measures.

We then extend this model to more complex objects. The use of broken lines improves the detection of network junctions and increases the accuracy of the extracted network.

Finally, we propose a hierarchical model of hydrographic networks in which the tributaries of a given river are modeled by a process of broken lines in the neighborhood of this river.

For each model, we accelerate convergence of the RJMCMC algorithm by using appropriate perturbations.

We show experimental results on aerial and satellite images (optical and radar data) to verify the relevance of the object process models.

Key-words: Stochastic geometry, marked point processes, simulated annealing, RJMCMC, line network extraction, aerial and satellite images.