



HAL
open science

Extraction de réseaux de rues en milieu urbain à partir d'images satellites à très haute résolution spatiale

Renaud Peteri

► **To cite this version:**

Renaud Peteri. Extraction de réseaux de rues en milieu urbain à partir d'images satellites à très haute résolution spatiale. domain_other. École Nationale Supérieure des Mines de Paris, 2003. Français. NNT: . pastel-00000508

HAL Id: pastel-00000508

<https://pastel.hal.science/pastel-00000508>

Submitted on 9 Feb 2004

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.



Collège doctoral

N° attribué par la bibliothèque

|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|

THESE

pour obtenir le grade de :

DOCTEUR DE L'ÉCOLE DES MINES DE PARIS

Spécialité «*Informatique temps réel - Robotique - Automatique*»

présentée et soutenue publiquement
par

Renaud PÉTERI

le 10 décembre 2003

Titre de la thèse :

**Extraction de réseaux de rues en milieu urbain à partir
d'images satellites à très haute résolution spatiale**

Directeur de thèse : Thierry Ranchin

JURY :

Jean-Paul	MARMORAT	Président
Albert	BIJAOU	Rapporteur
Marc	PIERROT-DESEILLIGNY	Rapporteur
Christiane	WEBER	Examineur
Olivier	GORETTA	Examineur
Thierry	RANCHIN	Examineur

Remerciements

Avant d'adopter le style assez impersonnel (et froid, il faut bien l'avouer) qu'impose l'écriture scientifique, la page de remerciements est traditionnellement le lieu où le thésard, fourbu mais content, se retourne enfin sur son parcours du combattant pour se laisser aller à quelques remarques personnelles.

Ainsi, cette page est certainement la plus agréable à rédiger, même si elle n'est pas forcément la plus facile, tant il ne faut oublier personne...

Je tiendrai tout d'abord à remercier les membres de mon jury qui ont accepté de me délivrer le «Saint Graal». Merci donc à :

Jean-Paul Marmorat, professeur au Centre de Mathématiques Appliquées de l'École des Mines de Paris, pour m'avoir accueilli dans sa formation doctorale et pour avoir accepté de présider mon jury de thèse.

Albert Bijaoui, astronome à l'Observatoire de Nice, qui m'a fait l'honneur d'être rapporteur de ma thèse et dont les conseils sur l'analyse multi-échelle et la transformée en ondelettes m'ont été bénéfiques.

Marc Pierrot-Deseilligny, directeur du laboratoire MATIS de l'IGN, pour avoir accepté le rôle de rapporteur et pour ses remarques très constructives. Je tiens à remercier à travers lui tout le MATIS, et plus particulièrement Sylvain Airault et Hervé Le Men pour les précieux conseils et l'aide qu'ils m'ont fournis pendant ma thèse.

Christiane Weber, directrice de recherche au laboratoire Image & Ville à Strasbourg, pour son avis expert de géographe et pour avoir suivi mon travail avec grande attention.

Cette thèse a été financée par une bourse DGA/CNRS du ministère de la Défense, que je souhaite remercier pour m'avoir permis de mener à bien mes travaux. Je remercie tout particulièrement Olivier Goretta pour ses conseils avisés et pour avoir régulièrement suivi ma thèse depuis le début.

Voici le moment de remercier ceux que j'ai eu la chance de côtoyer pendant plus de trois ans, à savoir tous les membres du Groupe Télédétection & Modélisation de l'École des Mines de Paris. Merci à :

Lucien (*Wald*), pour m'avoir accueilli dans ton laboratoire et pour avoir mis tous les moyens matériels nécessaires pour mener à bien ma thèse. Ceci est une première raison de te remercier. Une seconde est ton impressionnante capacité à concevoir et bonifier les idées, qui font qu'il n'est pas une fois où je ne suis ressorti de ton bureau regonflé!

Thierry (*Ranchin*) mon directeur de thèse, dont j'étais le premier thésard. Merci pour m'avoir fait lever la tête du guidon quand il le fallait, et pour m'avoir enseigné l'art de la

rédaction scientifique. Enfin, je tiens à te remercier pour avoir su me gérer avec une très grande «subtilité» dans l'éprouvante phase de rédaction. Garde longtemps ton côté très humain !

Michel (*Albuisson*) le responsable informatique que tous les centres nous envient - et dont je n'étais pas le chouchou, contrairement à ce que les jaloux disent ! Merci à la fois pour ta décontraction et pour avoir toujours pourvu à mes besoins informatiques.

François (*Cauneau*), en face duquel j'ai beaucoup appris, Lionel (*Ménard*) le guitariste-webmaster si sympathique.

Mes deux compères Nico (*Fichaux*) et Sylvain (*Cros*), qui sont pour beaucoup dans le bon déroulement de cette thèse et des pauses-café. Je vous souhaite de réussir ce que vous désirez.

Merci aussi à Ludo (*Basly*) qui m'a initié à la Polpet', à Thierry (*Poglio*), à Anthony (*Ung*), à la relève si sympathique assurée par Anne-Lise (*Beaulant*) et Claire (*Thomas*). Merci à tous les stagiaires que j'ai rencontrés et plus particulièrement à Julien (*Celle*) pour ta science de la programmation et pour avoir assuré le service après-vente jusqu'au bout. Enfin, à tous les membres de l'École des Mines du site de Sophia-Antipolis que j'ai pu côtoyer et plus particulièrement aux *Funky* membres de Duende.

Terminons par les remerciements les plus personnels.

Un très grand «köszönöm szépen» pour ma douce Dora : ta présence à mes côtés a eu une importance qui ne se mesure pas. Merci pour tes tisanes. Au fait, je te dois 124 468 heures de travaux d'intérêt collectif...

A mes grands-parents et toute ma famille -au sens large- que j'adore. Notamment ma sœur Fanny et mon frère Florent à qui je souhaite tout le bonheur possible dans leur vie.

Enfin, un énorme merci à mes parents Anne-Marie et Dominique pour m'avoir mis au monde et pour l'éducation que j'ai eu la chance de recevoir de votre part. Vous n' imaginez pas combien cette thèse est un peu la vôtre ...

Voilà c'est fini. Ah oui, pour ceux que cela intéresse, il y a juste après plus de 150 pages de grand frisson, peuplées de monstres étranges comme les *DoubleSnakes* ...

Table des matières

Contexte de l'étude et objet de la thèse	3
1 États de la rue et de l'art	13
1.1 La rue dans tous ses états	13
1.1.1 Variabilité intrinsèque et extrinsèque des routes	13
1.1.2 Variabilité typologique par rapport à l'environnement	14
1.1.3 Variabilité due aux conditions et aux moyens d'observation	15
1.1.4 Variabilité due au contexte de l'application	16
1.1.5 Une problématique pour différentes communautés	16
1.2 Avant de commencer l'état de l'art	17
1.2.1 Les différents modèles	17
1.2.2 Les différentes étapes	18
1.2.3 Organisation de l'état de l'art	18
1.3 Travaux en milieu rural ou péri-urbain	19
1.3.1 Travaux sur le linéique	19
1.3.2 Travaux sur le surfacique	25
1.4 Travaux dans le contexte urbain	28
1.4.1 Les travaux sur le linéique	28
1.4.2 Les travaux sur le surfacique	29
1.5 Conclusions sur ce chapitre	31
2 Extraction de bords parallèles par contours actifs et analyse multi-échelle	33
2.1 Objectifs	33
2.2 Les contours actifs	34
2.2.1 Introduction	34
2.2.2 Contours actifs discrets	35
2.2.3 Procédure d'optimisation par algorithme glouton	36
2.3 Transformée en ondelettes et analyse multi-échelle	37
2.3.1 La transformée en ondelettes continue	37
2.3.2 La transformée en ondelettes «à trous»	38
2.4 Contours actifs et énergie multi-échelle par ondelettes	42
2.4.1 Détection multi-échelle de contours	42
2.4.2 Énergie multi-échelle des contours actifs	44
2.5 <i>DoubleSnakes</i> pour l'extraction des contours parallèles	44
2.5.1 Principe	44
2.5.2 Énergie de parallélisme	45

2.5.3	Paramètres des <i>DoubleSnakes</i>	46
2.6	Applications	47
2.6.1	Effet de l'énergie de parallélisme sur un <i>DoubleSnake</i>	47
2.6.2	Effet de l'approche multi-échelle	48
2.6.3	Application à une scène réelle	50
3	La méthode d'extraction globale	53
3.1	Méthode développée	53
3.1.1	Rappel des objectifs	53
3.1.2	Choix effectués	53
3.1.3	Schéma de la méthode d'extraction globale	54
3.1.4	Les entrées de la méthode	54
3.1.5	L'algorithme de traitement	55
3.1.6	Évaluation des résultats	56
3.2	Module du graphe	57
3.2.1	Formalisme utilisé	57
3.2.2	Description	57
3.2.3	Extraction des polygones du graphe	58
3.2.4	Obtention du graphe complet à partir des différentes polygones	61
3.3	Module de reconstruction	64
3.3.1	La méthode d'extraction surfacique du réseau	64
3.3.2	Initialisation des <i>DoubleSnakes</i> pour l'extraction des tronçons de rue à bords parallèles	65
3.3.3	Traitement des intersections	67
3.3.4	Déroulement de l'algorithme de reconstruction	69
4	Référence et évaluation	73
4.1	L'évaluation et le besoin d'une référence	73
4.1.1	Introduction	73
4.1.2	Objectifs	74
4.2	Les différents types de référence	75
4.2.1	Référence basée sur des données de terrain	75
4.2.2	Référence basée sur l'interprétation d'images	75
4.3	Méthode pour l'extraction d'une référence	76
4.3.1	Objet référence	78
4.3.2	Zone de tolérance	79
4.4	Application à une scène réelle	82
4.4.1	Protocole d'expérimentation	82
4.4.2	Extraction d'une référence et de la zone de tolérance	83
4.5	Évaluation quantitative à partir de la référence et de la zone de tolérance	88
4.5.1	Évaluation quantitative à partir de la référence	88
4.5.2	Évaluation quantitative à partir de la zone de tolérance	88
4.6	Conclusion sur ce chapitre	89

5 Applications de la méthode	91
5.1 Introduction	91
5.1.1 Les cas d'étude	91
5.1.2 L'évaluation quantitative	91
5.2 Application à une image synthétique	92
5.3 Image Ikonos de la zone de Hasselt, Belgique	98
5.4 Image Quickbird de la zone de Fredericton, Canada	107
5.5 Image de la caméra numérique de l'IGN de la ville d'Amiens, France	116
5.6 Analyse et perspectives pour l'algorithme de traitement	122
Conclusion et perspectives	124
Bibliographie	133
Publications	140
Annexes	143
A L'analyse multirésolution	145
B Les stades de fonction	149
C «Greedy Algorithm» : Pseudo-Code	151

Contexte de l'étude et objet de la thèse

Contexte et motivations

La cartographie de la voirie en milieu urbain correspond à un besoin social et économique important. Du point de vue des collectivités publiques, il est primordial pour les villes d'avoir une cartographie à jour de leur réseau routier (aménagement urbain, collecte des eaux usées, pollution, ...). Au delà de la gestion de la voirie, les applications connexes sont nombreuses (utilisation des cartes pour les réseaux de télécom, le géo-marketing, les bases de données pour la navigation, les services WAP...).

Il existe ainsi une demande dans la création ou la mise à jour de cartes. Actuellement, la cartographie du réseau de rues est essentiellement effectuée par interprétation d'images avec un apport de mesures de terrain. Cette procédure est un travail long et fastidieux et demande à être ré-effectuée à chaque mise à jour du réseau. A titre d'exemple, l'IGN¹ a entrepris il y a quelques années la constitution de la base de données topographique, la BDTopo[®], dont la saisie est faite à partir de photographies aériennes (principalement au 1/30 000) et dont la fréquence de mise à jour prévue moyenne est de 7 ans. Or une acquisition entièrement manuelle et complète mettant en jeu la centaine d'opérateurs de l'IGN peut être estimée à plusieurs millions d'heures de travail (Ruskoné 1996). On comprend alors l'intérêt d'automatiser ou d'assister la saisie du réseau de rues. La création d'un groupe de travail européen (*EuroSDR*², *ex-OEEPE*) dédié à l'extraction et la mise à jour des réseaux de rues traduit cet enjeu ainsi que le besoin d'interactions entre les chercheurs et les utilisateurs. L'automatisation de cette tâche intéresse aussi la Défense Nationale pour la constitution *ex-nihilo* de bases de données géographiques.

L'imagerie à très haute résolution

Du point de vue du géographe, l'échelle d'intérêt en milieu urbain varie du 1/200 au 1/1 000 000. La disponibilité d'images urbaines à très haute résolution spatiale constitue un apport indéniable pour l'étude du milieu urbain et sa cartographie (Puissant et Weber 2002). En effet, toutes les échelles pour étudier le milieu urbain peuvent ainsi être atteintes par des images de télédétection : les images aériennes pour les échelles les plus grandes et

¹Institut Géographique National

²European Spatial Data Research

<http://www.bauw.unibw-muenchen.de/institute/inst10/euroedr/>

les images satellites pour les échelles plus basses.

Depuis longtemps, l'image aérienne, acquise par chambre de prise de vue à film argentique, a été utilisée pour l'observation de la Terre. Ces images ont été récemment numérisées.

Il existe maintenant des caméras numériques aéroportées offrant de très bons résultats. La caméra numérique de l'IGN est composée de quatre têtes de caméra CCD³ portant des filtres : 3 filtres R, V et B dans le domaine du visible et un filtre infrarouge (Souchon et Thom 2003). Dans le cas de la mission sur la ville d'Amiens de 1999, la résolution spatiale des images acquises atteignait 20 cm par pixel.

Il existe aussi depuis peu sur le marché des images satellites à très haute résolution spatiale (par exemple les satellites SPOT 5, Ikonos, Quickbird ou EROS). Le terme «très haute résolution» est bien évidemment relatif au contexte d'étude. Dans notre cadre d'étude, nous entendons par très haute résolution spatiale des résolutions inférieures à 5 m dans le canal panchromatique⁴ (Puissant et Weber 2002).

Le tableau 0.1 présente les caractéristiques de ces satellites à très haute résolution spatiale.

	Date de lancement	Bande panchromatique	Bandes multispectrales
SPOT 5	2002	2,5 m - 5 m	10 m
Ikonos 2	1999	0,82 m	3,28 m
Quickbird	2000	0,61 m	2,44 m
EROS 1A	2000	1 m - 1,8 m	-

TAB. 0.1: les satellites à très haute résolution spatiale et leurs résolutions

Il convient de préciser que ces résolutions sont données pour des conditions de prise de vue au nadir⁵, correspondant à la meilleure résolution que l'on peut obtenir. Ainsi, pour le satellite Quickbird, la résolution des pixels varie de 61 cm au nadir à 72 cm pour un dépointage maximum de l'optique de 25 ° par rapport au nadir. Dans le cas du satellite Quickbird, les images sont souvent ré-échantillonnées à 70 cm pour avoir une résolution homogène sur toute l'image.

Les images 0.1(a) et 0.1(b) sont des exemples du niveau de détails des capteurs actuels du domaine civil sur des scènes urbaines. La première image 0.1(a) provient de la caméra numérique aéroportée de l'IGN et offre une résolution spatiale de 20 cm. La deuxième image 0.1(b) provient du satellite Quickbird dont la résolution est ici de 70 cm en mode panchromatique. Sur ces deux images, on peut noter le grand nombre de détails observables (véhicules, marquages au sol, ...). Sur de telles images, la rue, dont les deux bords sont visibles, est localisable avec une très grande précision.

³Charge-Coupled Device

⁴bande spectrale très large, couvrant la partie visible du spectre électromagnétique ou plus

⁵point au sol situé à la verticale du satellite ou du capteur



(a) Vue d'Amiens (caméra numérique de l'IGN). © IGN

(b) Vue de Strasbourg par le satellite Quickbird. © DigitalGlobe

FIG. 0.1: exemple d'images à très haute résolution spatiale : l'image 0.1(a) provient de la caméra IGN et a été acquise à la résolution de 20 cm. L'image 0.1(b) provient du satellite Quickbird de résolution spatiale 70 cm en mode panchromatique.

Pour l'extraction de rues, d'autres types d'information peuvent être utilisés conjointement à l'image à haute résolution spatiale. On peut mentionner l'apport des mesures terrain comme amorce pour un processus de détection ou de suivi. L'information d'élévation provenant d'un MNE⁶ peut se révéler aussi très intéressante pour extraire les rues dans le cas où elles sont partiellement occultées (cas des échangeurs) ou pour lever des ambiguïtés (par exemple radiométrie similaire entre une rue et un bâtiment). Ces MNE sont généralement obtenus par pointés manuels sur un couple d'image, par corrélation d'images stéréoscopiques ou par des mesures LIDAR⁷ (De Joinville 2001).

Objectifs

La recherche à l'Ecole des Mines de Paris est principalement à but applicatif. Nous nous sommes donc placés dans le cas typique d'un service technique chargé d'effectuer la cartographie d'une ville, et nous nous sommes penchés tout d'abord sur les données utilisables, puis sur le développement de méthodes adaptées.

L'utilisation du LIDAR demande encore des recherches et il est très peu utilisé par les cabinets de géomètres et les services techniques des villes. De plus, le coût d'une campagne d'acquisition est important.

De même, les campagnes aériennes pour l'acquisition d'images stéréoscopiques sont lourdes

⁶Modèle Numérique d'Elévation : en plus du profil en altitude du sol, un MNE intègre le sursol (bâti, végétation...)

⁷LIght Detection And Ranging. *En français : détection et télémétrie par la lumière*

et coûteuses à monter. Le traitement des données est complexe et nécessite d'être effectué par un expert en photogrammétrie.

Nous pensons que l'image optique visible restera l'instrument courant des cabinets de géomètres, au moins pour la décennie à venir. Nous avons donc choisi de prendre l'image comme donnée unique pour effectuer la cartographie.

Ce genre d'image optique est acquise à l'aide de moyens aéroportés ou spatioportés. Les moyens satellites sont très pratiques dans les cas nombreux où l'accès aux sites d'intérêt est difficile : problèmes administratifs (autorisation de survol), éloignement géographique, catastrophes naturelles, états de conflits ou de guerres ...

Par ailleurs, en milieu urbain, il est préférable d'avoir une haute résolution spatiale et une moins bonne résolution spectrale (mode panchromatique) qu'une basse résolution spatiale et une haute résolution spectrale (canaux multispectraux)(GDTA 1995). Il existe des méthodes de fusion qui permettent d'augmenter la résolution spatiale des canaux multispectraux par injection d'informations à haute résolution spatiale provenant du canal panchromatique (Wald 2002, Ranchin *et al.* 2003). Toutefois, elles ne sont pas pour l'instant assez abouties dans le cas de la résolution métrique et sub-métrique (Laporterie *et al.* 2003).

Nous travaillerons donc à partir d'images panchromatiques à haute résolution spatiale comme seule source d'information. Les images traitées présentent une grande plage de résolution spatiale (de 70 cm à 5 m de résolution). Un effort limité a aussi été fait pour l'application de notre méthode à des images aériennes acquises par la caméra numérique de l'IGN de résolution spatiale de 20 cm.

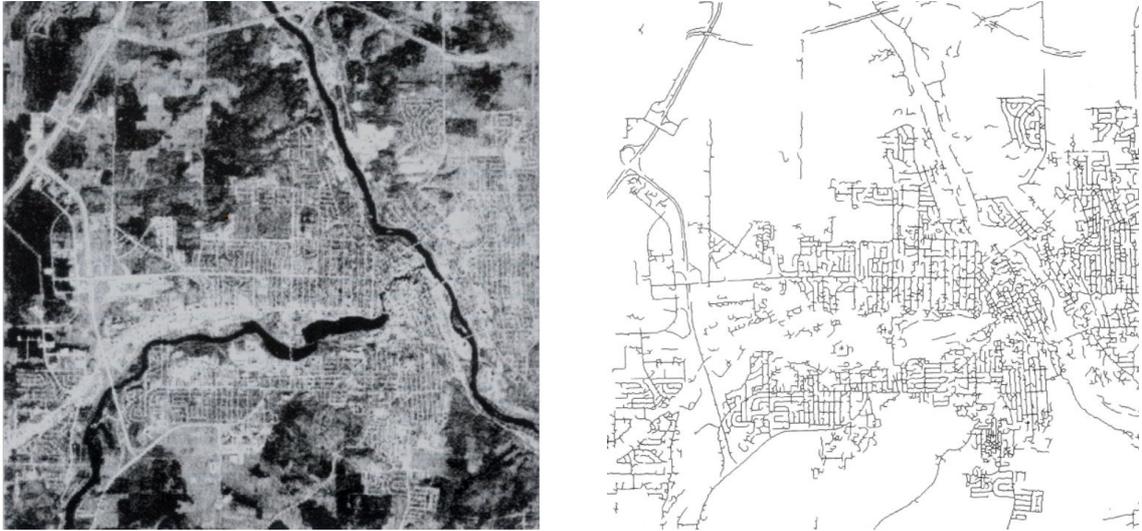
Nous prendrons aussi par la suite des images parfaitement géo-référencées. Cette hypothèse a relativement peu d'influence sur la méthode que nous développons, mais en a sur la qualité géographique de la carte résultante.

Du linéique au surfacique

L'automatisation de l'extraction du réseau routier à partir d'une image satellite n'est pas un problème récent et a fait l'objet de nombreux travaux en traitement d'images et en télédétection depuis plus de vingt ans. Cependant, l'extraction automatique ou semi-automatique⁸ de routes est toujours, à cause de sa complexité, un thème actif de recherche. Avant l'apparition de la très haute résolution spatiale, les méthodes d'extraction développées dans la littérature portaient sur les images satellites dont la résolution est au mieux de 10 m. A ces résolutions, les routes sont représentées par des lignes d'une épaisseur de 1 à 3 pixels, et ces méthodes amènent à une extraction d'un réseau linéique. Les routes visibles sur ce type d'images sont les voies de communication principales du réseau routier telles que les routes nationales, les autoroutes ou les boulevards en ville. Les routes secondaires ne sont pas visibles ou ont un aspect fragmenté dans l'image.

La figure 0.2 est un exemple de résultat d'algorithme d'extraction de réseaux de rues linéiques (d'après Wang *et al.* 1996). L'image traitée provient du capteur HRV du satellite SPOT, qui possède une bande panchromatique à moyenne résolution (10 mètres).

⁸voir annexe B pour la définition des termes



(a) Image panchromatique SPOT HRV de la région de Sherbrooke, Canada (10 m de résolution)

(b) Réseau linéique extrait par la méthode de Wang *et al.* (1996)

FIG. 0.2: exemple d'extraction de réseaux de rues linéiques en milieu urbain sur une image à résolution spatiale moyenne. © Wang *et al.* (1996).

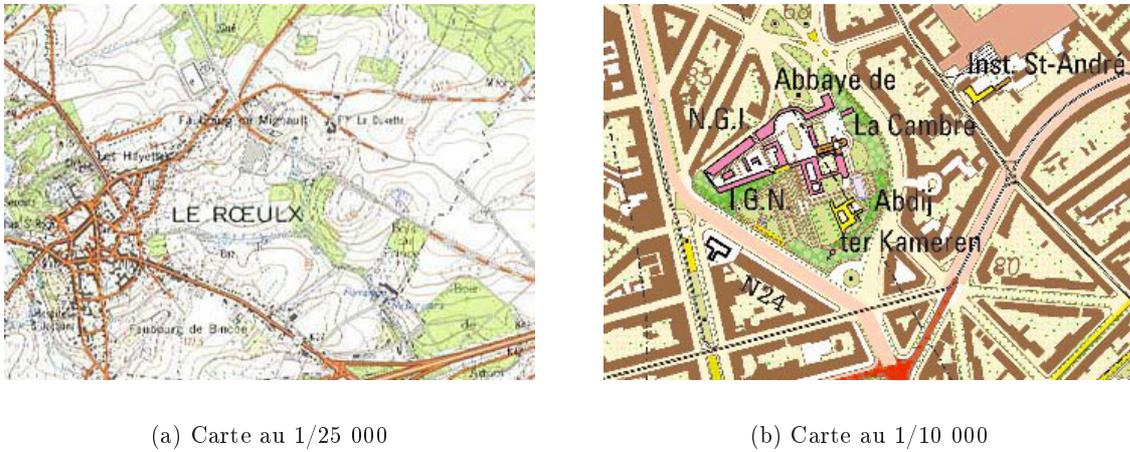
Avec l'apparition de la haute résolution spatiale en milieu urbain, la perception des routes passe du linéique au surfacique (plus de 3 pixels de large). Par surfacique, nous entendons que la route n'est plus représentée par un ligne d'une certaine épaisseur, mais par deux contours parallèles définissant une surface souvent inhomogène en radiométrie (spécialement dans notre cadre d'étude du milieu urbain).

A la résolution spatiale submétrique que permettent d'atteindre les images aériennes, les rues ont une représentation surfacique. Pourtant, très peu de méthodes ont été développées en milieu urbain.

Ceci peut s'expliquer par l'utilisation de telles images pour la cartographie : elles étaient principalement utilisées pour l'établissement de cartes au 1/25 000. Selon les principes de généralisation cartographique, la représentation des rues à ces échelles est telle que l'axe de la rue est bien positionné, mais que sa largeur sur la carte est supérieure à son emprise réelle (figure 0.3(a)). L'IGN belge⁹ dispose depuis peu de cartes au 1/10 000 pour l'ensemble de son territoire. A cette échelle (figure 0.3(b)), la rue est représentée en tant que surface, et sa largeur en représentation cartographique correspond à son emprise réelle au sol.

Le peu de méthodes développées est aussi lié à la très grande difficulté d'une approche tout automatique en milieu urbain, principalement pour l'extraction automatique et fiable des amorces de rues. En effet, le contexte urbain et la haute résolution spatiale de l'image induisent une forte augmentation du bruit et des artefacts.

⁹<http://www.ign.be>



(a) Carte au 1/25 000

(b) Carte au 1/10 000

FIG. 0.3: deux exemples de représentation cartographique des rues © IGN-Belgique

De nouveaux types d'artefacts

L'augmentation de la résolution spatiale va permettre d'extraire la rue avec une meilleure précision géographique, et une meilleure identification des différents types de voies de communication (typologie selon la largeur), ou de caractériser l'emprise du réseau de rues. La contrepartie est une augmentation significative du bruit dans l'image, rendant l'extraction plus difficile.

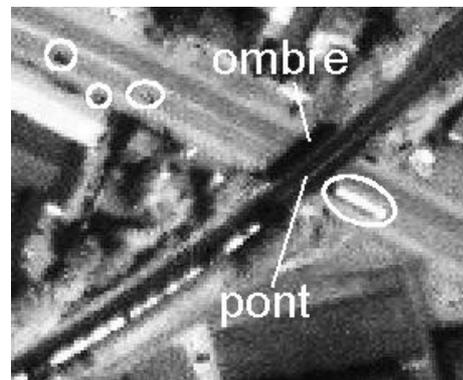


FIG. 0.4: occlusions et artefacts inhérents au contexte urbain. Les véhicules sont entourés.

L'image de la figure 0.4 est un bon exemple pour illustrer quelles sortes d'artefacts engendrent à la fois le contexte urbain et la haute résolution spatiale. Les véhicules (entourés), le pont et son ombre vont perturber le procédé automatique d'extraction. D'autres types d'artefacts rencontrés sont les arbres, les surfaces bitumées (parking, aéroport) ou les bâtiments avec des radiométries similaires à celles des rues et ayant un contraste important avec leur environnement.

Un autre problème sensible en milieu urbain est l'angle de prise de vue (figure

0.5(a)). L'image 0.5(b) illustre les conséquences d'une prise de vue qui n'est pas au nadir : plusieurs bâtiments ainsi que leurs ombres portées occultent une partie des rues, et il peut même devenir difficile pour un opérateur humain de localiser la partie manquante.

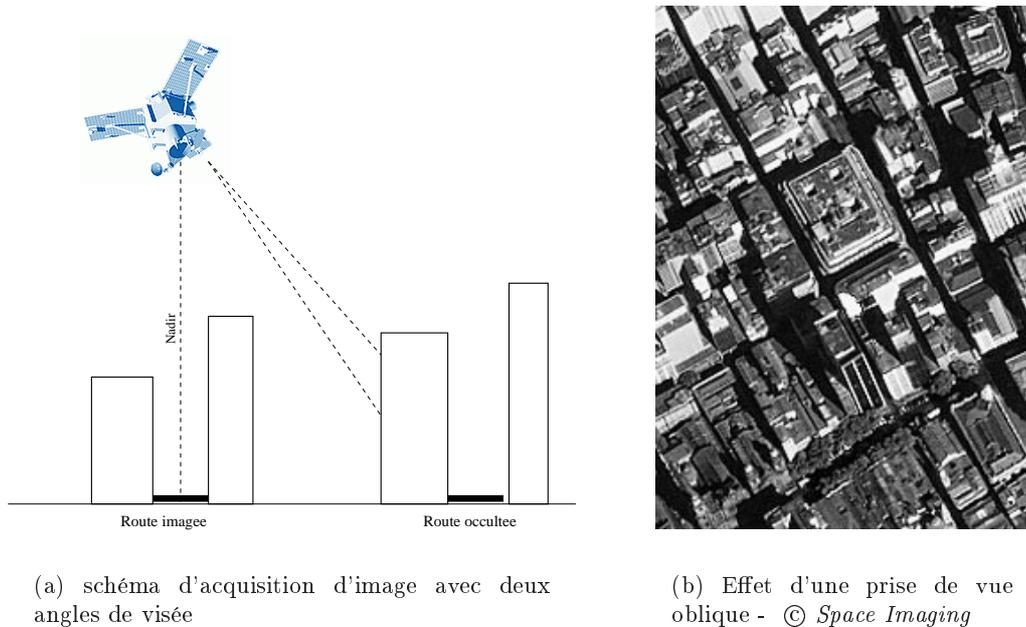


FIG. 0.5: effet de l'angle de prise de vue sur la visibilité des rues

Nous prendrons comme hypothèses une prise de vue proche du nadir (± 15 degrés) et ne considérerons pas des rues à plusieurs niveaux (comme les échangeurs), dont les ambiguïtés ne peuvent être levées qu'avec l'apport de MNE.

Nécessité d'une nouvelle méthode

Ce passage d'une représentation linéique à une représentation surfacique du réseau nécessite le développement de méthodes différentes de celles employées pour l'extraction de réseaux linéiques, les modèles de rues utilisés n'étant plus applicables pour la très haute résolution spatiale (Shi et Zhu 2002). De plus, ces méthodes doivent prendre en compte la particularité du milieu urbain. Le domaine urbain est un milieu très bruité, dans lequel existe une grande variété typologique des rues (rues principales, voies rapides, ruelles, ...). Cette méthode doit être adaptable à tous les types d'images à très haute résolution spatiale (de 70 cm à 5 m).

Organisation du manuscrit

Dans un premier temps, la grande variabilité d'aspects de la rue dans une image télé-détectée sera présentée. La difficulté de pouvoir en définir un modèle générique a conduit à un grand nombre de méthodes d'extraction, dont nous ferons un état de l'art (*chapitre 1*).

Les méthodes appliquées à des milieux ruraux ou péri-urbains seront d'abord présentées. On différenciera les méthodes les plus anciennes ayant porté sur l'extraction de réseaux linéiques, et celles portant sur l'extraction de rues en tant qu'éléments de surface, que ce soit à partir d'images aériennes ou des nouveaux capteurs satellites à très haute résolution spatiale. Les travaux sur le milieu urbain, peu nombreux, feront l'objet d'un traitement à part.

Analysant les causes d'échecs des différentes méthodes mais aussi leurs aspects prometteurs pour nos objectifs, nous en dégagerons un procédé original permettant l'extraction des contours parallèles de rues dans un environnement bruité (*chapitre 2*).

Une nouvelle méthode globale d'extraction du réseau de rues incluant ce modèle sera ensuite présentée (*chapitre 3*). Cette méthode est modulaire et permet une extraction surfacique des réseaux de rues en milieu urbain.

Un accent particulier a été mis sur l'évaluation des résultats (*chapitre 4*). Cette évaluation est essentielle car elle permet de caractériser les résultats produits par un algorithme automatique mais aussi d'en donner une mesure de fiabilité. Une méthode a été développée pour extraire une référence de comparaison lorsque la seule donnée est l'image numérique. Les critères quantitatifs associés seront aussi exposés.

La méthode a été ensuite appliquée à plusieurs images provenant de différents capteurs (Ikonos, Quickbird, SPOT 5 ou la caméra numérique aéroportée de l'IGN...) et sur différentes scènes (différents degrés et types d'urbanisation). L'étude des résultats d'extraction et leur évaluation qualitative et quantitative sera l'objet du chapitre 5.

Le document se termine par une conclusion générale et les perspectives qu'offre la méthode. Si les résultats sont prometteurs dans l'optique d'une aide à la cartographie, nous indiquerons aussi des voies de recherches à explorer.

Chapitre 1

États de la rue et de l'art

Pour tout problème d'extraction d'informations et d'objets à partir d'une image, il convient pour bien répondre aux objectifs visés, de se poser au préalable plusieurs questions.

Quel est l'objet à extraire et quelle est sa variabilité typologique ? Comment apparaît l'objet dans l'image ? Quelles informations propres à l'objet cherche-t-on à extraire ? Pour quelles applications ? Et avec quelle fiabilité ?

Nous nous sommes donc posés cette série de questions dans le cadre de l'extraction de rues (section 1.1).

L'abondance des travaux sur l'extraction de rues traduit cette variété de problèmes posés et donc de solutions retenues. Nous en dressons un état de l'art (sections 1.3 et 1.4) et en tirons les enseignements pour la méthode que nous proposons (section 1.5).

1.1 La rue dans tous ses états ...

La définition et la représentation de la rue pour un système de reconnaissance de formes vont dépendre de plusieurs facteurs.

1.1.1 Variabilité intrinsèque et extrinsèque des routes

Quel que soit son environnement, la radiométrie d'une route est étroitement liée à son type de revêtement et à son degré d'usure. Dans une même image, on peut trouver des portions de route sombres pour lesquelles le revêtement a été refait et d'autres usagées, voire non revêtues, qui apparaissent claires. La poussière le long du bord des routes et surtout des autoroutes (notamment due à la poussière des pneus) a aussi un effet sur la radiométrie.

Les caractéristiques géométriques des routes (largeur, courbure, longueur ...) sont aussi très variables selon la période de construction. Enfin, les intersections représentent une portion à part de la route. De par leur diversité, l'extraction précise des carrefours est complexe et demande un traitement particulier (Boichis 2000).

1.1.2 Variabilité typologique par rapport à l'environnement

La variabilité typologique de la rue ou de la route est aussi très dépendante de leur contexte. La distinction entre les termes «routes» et «rues» tient compte de cette différence contextuelle : la route est une *«voie de communication terrestre aménagée, plus importante que le chemin, située hors d'une agglomération ou reliant une agglomération à une autre (opposé à rue)»*, alors qu'une rue est une *«voie bordée, au moins en partie, de maisons dans une agglomération»* (d'après le dictionnaire *Le Petit Robert*).

Que ce soient des routes ou des rues, leurs aspects sont liés à l'environnement dans lequel elles se situent. On peut ainsi trouver des courbures faibles et variant peu pour des routes nationales ou pour des autoroutes, alors qu'une route en montagne pourra comporter de nombreux virages à fortes courbures (routes en lacet).

Le contexte urbain présente aussi des particularités comparées aux milieux péri-urbains ou ruraux. Considérons par exemple les trois images suivantes du canal panchromatique de 1 m de résolution du satellite Ikonos :

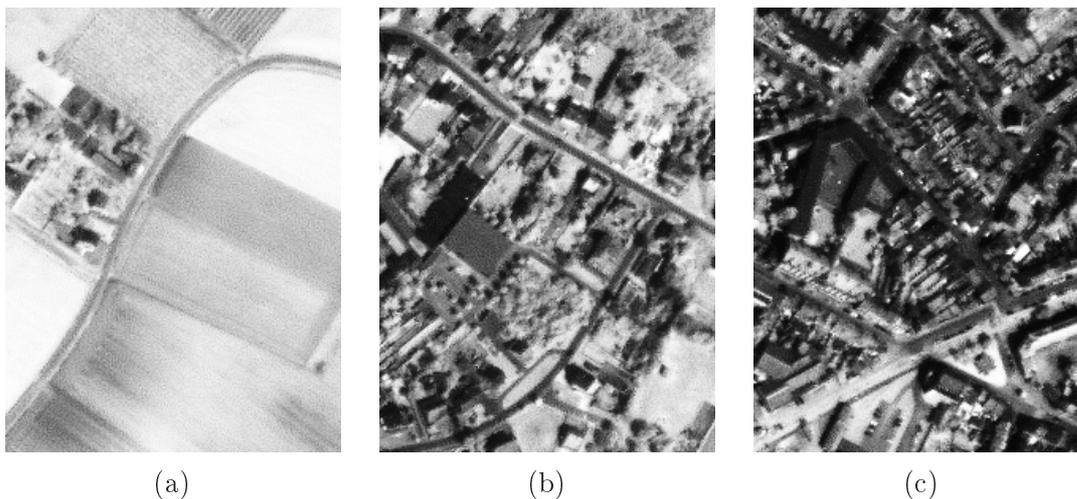


FIG. 1.1: représentation de milieux ruraux (a), péri-urbains (b) et urbains (c) dans la région de Hasselt (Belgique). © 2000 Space Imaging Europe. Image Ikonos aimablement fournie par la société Geographic Information Management

On peut se rendre compte visuellement que la difficulté à extraire le tracé des rues augmente lorsque l'environnement s'urbanise. Hinz *et al.* (2001) décrivent deux principales raisons à cela, illustrées figure 1.1 :

1. *l'hétérogénéité des objets appartenant à la même classe est habituellement plus grande. Par exemple, les immeubles en centre ville sont typiquement plus compliqués que des maisons plus ou moins isolées en zone péri-urbaine.*
2. *l'apparence d'un objet est souvent influencée par les objets de son voisinage et donc, il diverge du modèle d'objet sous-jacent. Par exemple, des immeubles projetant leur ombre sur une route produisent de fortes variations d'intensité dues à l'ombre au milieu d'une route habituellement claire et homogène.*

Le contexte urbain et la très haute résolution spatiale accroissent le nombre d'artefacts rencontrés. Les véhicules ou les ombres portées des arbres sont des perturbations pour tout processus d'extraction de rues. On peut aussi mentionner d'autres types d'artefacts comme les zones goudronnées (parking, aéroport) ou les immeubles qui présentent une radiométrie similaire aux rues et un contraste important avec leur environnement.

Pour ce qui est de l'analyse du milieu urbain, l'infrastructure de la ville étudiée joue également un grand rôle. Ainsi, les villes caractérisées par des structures urbaines espacées et par un réseau urbain quadrangulaire (villes d'Amérique du Nord par exemple) sont relativement faciles à cartographier. En revanche, les villes anciennes européennes ou asiatiques aux trames serrées et aux réseaux routiers tortueux nécessitent une résolution spatiale plus fine (Weber 1995).

Cette dernière remarque nous amène alors à parler des conditions d'acquisition, qui jouent aussi un rôle important pour la perception de la rue dans l'image.

1.1.3 Variabilité due aux conditions et aux moyens d'observation

1.1.3.1 *Les conditions physiques d'acquisition*

Indépendamment de la nature de la rue, les conditions météorologiques (chaussée humide, présence de verglas ou de neige) lors de l'acquisition de l'image vont modifier la radiométrie de la rue. De plus, les ombres portées, artefacts très gênants pour le milieu urbain (notamment les ombres portées des bâtiments), vont différer suivant la date et l'heure d'acquisition de l'image. De même, la présence d'aérosols ou de nuages va influencer sur l'aspect radiométrique de la rue dans l'image.

1.1.3.2 *Les moyens d'observation*

La manière dont la rue va être acquise par le capteur (appareil photographique, caméra numérique ou barrettes de CCD) joue un rôle prépondérant dans sa représentation dans l'image.

L'angle de prise de vue. Tout d'abord une prise de vue qui n'est pas au nadir peut produire des occlusions partielles, voire totales, de la rue dans l'image. Cet aspect est non négligeable en milieu urbain, la rue étant très souvent située entre deux rangées de bâtiments. Une prise de vue oblique induit aussi des distorsions géométriques (pouvant être minimisées par orthorectification).

La résolution spatiale du capteur. L'apparence des rues dans l'image est directement liée à la résolution spatiale du capteur satellite ou aéroporté. Comme nous l'avons présenté dans le chapitre d'introduction, sur des images à haute résolution spatiale, la perception des routes passe du linéique (moins de 3 pixels de large) au surfacique (deux contours anti-parallèles définissant une surface souvent hétérogène en radiométrie).

Cette hétérogénéité radiométrique va elle aussi dépendre de la finesse de résolution spatiale du capteur : l'augmentation de la résolution spatiale engendre en contrepartie une augmentation significative du «bruit» pouvant occulter la rue, comme les véhicules, le marquage

au sol ou les arbres. En milieu urbain, «l'adéquation entre la résolution spatiale et la discrimination des éléments va varier en fonction de la taille, de l'organisation et de la densité des objets urbains (réseaux routiers, bâtiments, ...）」 (Weber 1995, page 98).

La résolution spectrale du capteur. La nature spectrale des données est également importante lors de la détection des rues sur les images. En effet, sur certaines bandes spectrales (SPOT panchromatique ou Landsat TM 1, 2 et 3 par exemple), les rues apparaissent comme des lignes brillantes par rapport à leur environnement, alors que sur d'autres bandes (Landsat TM 4 et 5 par exemple), elles se présentent comme des lignes sombres. Il faut donc tenir compte de la bande spectrale dans laquelle l'image a été acquise pour avoir une idée de l'apparence d'une rue par rapport à son environnement sur cette image. Cet effet s'estompe avec l'augmentation de la résolution spatiale, les objets deviennent identifiables indépendamment de leur réponse spectrale.

Représentation tridimensionnelle de la rue. Les recherches sur la reconstruction automatique de paysages 3D à partir de couples d'images stéréoscopiques ou par levés LIDAR rendent potentiellement possible le calcul de MNE. Un MNE permet d'obtenir l'altitude de la rue et donc de la représenter dans l'espace à trois dimensions. Si pour l'instant cela est encore du domaine de la recherche, il sera possible à long terme de produire des bases de données 3D des villes incluant les rues.

1.1.4 Variabilité due au contexte de l'application

Dans les Systèmes d'Information Géographique (SIG), la représentation d'un même objet peut être différente selon la vision de son utilisateur, ainsi que l'échelle de représentation choisie. C'est notamment le cas d'une rue dans une ville. Le préposé au cadastre la représente sous la forme de deux polygones. Les responsables du revêtement utiliseront quant à eux une surface. Au niveau de la gestion du trafic routier, la rue est représentée par un morceau de graphe. Quant au responsable des réseaux souterrains, il prendra en compte le volume constitué par la projection de la section utile normale le long de la chaussée (Ranchin 2000).

Pour l'établissement de cartes, il existe une véritable typologie associée à un symbolisme particulier. Les routes et les rues sont représentées par des symboles linéaires dont la largeur, la texture et la couleur dépendent de la viabilité mais aussi de la série de cartes (et du fabricant). Il existe aussi tout un code hiérarchique (la généralisation cartographique) indiquant suivant l'échelle quel type d'objet doit être représenté prioritairement à un autre. La connaissance du réseau de rues est donc en partie subjective et liée à son application.

1.1.5 Une problématique pour différentes communautés

L'extraction de rues et de routes est depuis presque trente ans un thème actif de recherche - les premiers travaux à notre connaissance remontent à 1976 (Bajcsy et Tavakoli). Cela traduit la difficulté du problème, et aussi son importance pour plusieurs communautés scientifiques. Le thème de l'extraction de rues à partir d'images satellites ou aériennes intéresse la communauté de traitement d'images pour des problèmes théoriques ou celle des cartographes et des géographes dans un cadre de la télédétection plus appliqué.

L'abondance des travaux illustre cette diversité d'applications et d'efforts qui ont été menés. Dans la section suivante, nous en proposons un état de l'art.

1.2 Avant de commencer l'état de l'art

Avant d'effectuer un état de l'art sur l'extraction de rues, nous allons préciser les modèles de rues les plus souvent employés dans la littérature, ainsi que les différentes étapes possibles composant les algorithmes.

1.2.1 Les différents modèles

Nous l'avons vu (section 1.1), il existe une grande variabilité d'aspects de la rue dans l'image. Il est donc difficile de trouver un modèle générique de la rue. Toutefois, quelques caractéristiques majeures peuvent être dégagées. La plupart des méthodes d'extraction de rues se basent sur des modèles simples de rues. Les caractéristiques principales de ces modèles sont communes à beaucoup d'auteurs, à savoir que la rue a une radiométrie relativement homogène le long de son axe, et présente un contraste par rapport à son environnement. De plus, la largeur de la rue et sa courbure sont supposées varier lentement.

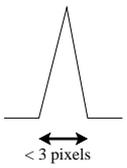
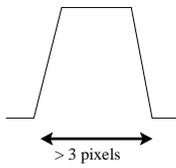
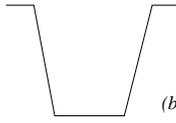
Aspect par rapport l'environnement	Modèle linéique	Modèle surfacique
Routes brillantes		
Routes sombres		
	(a)	(b)

FIG. 1.2: les deux modèles de profils radiométriques «idéaux» de la rue

Sur les images satellites à basse ou moyenne résolution spatiale (plus de 10 m), la rue apparaît comme un élément linéaire de 1 à 3 pixels de large (figure 1.2.a). Sur les images aériennes ou satellites à haute ou très haute résolution spatiale, la rue apparaît comme un ruban de radiométrie assez homogène (sauf en milieu urbain) ayant des bords contrastés et parallèles (figure 1.2.b). La figure 1.2 présente ces profils radiométriques «idéaux» des rues, servant de modèle à la plupart des algorithmes.

D'un point de vue du contexte, plusieurs objets peuvent être alignés sur les bords de la rue, comme les bâtiments et les arbres. De plus, par la fonctionnalité même des rues, des véhicules s'y trouvent.

Enfin, le réseau formé par les rues est supposé connexe et présente peu d'extrémités libres, mis à part celles liées au bord de l'image.

1.2.2 Les différentes étapes

Les algorithmes d'extraction de rues peuvent être composés d'une ou plusieurs étapes : la détection, l'extraction, le suivi ou la validation. Ces étapes peuvent elles-mêmes être décomposées en plusieurs fonctions (voir annexe B).

La détection consiste à établir dans une zone d'intérêt, la présence d'une route ou non.

L'extraction consiste à extraire le réseau dans l'image globale.

Le suivi consiste à «suivre» la rue à partir de points d'amorce en fonction d'un modèle théorique.

La validation est une étape d'auto-évaluation des résultats. Elle peut occasionner des rebouclages dans l'algorithme de type résolution de conflit.

Ces étapes peuvent être combinées : ainsi les amorces d'un algorithme de suivi peuvent provenir d'un algorithme extrayant un réseau partiel et non forcément connexe.

1.2.3 Organisation de l'état de l'art

Notre cadre d'étude est le milieu urbain et la très haute résolution spatiale. Nous cherchons à extraire les rues de l'image en tant qu'éléments de surface. La plupart des travaux menés jusqu'à présent concernent le milieu rural ou péri-urbain. Il s'agit donc de méthodes d'extraction de routes. Toutefois, la similitude entre route et rue est suffisamment grande pour que l'étude de ces travaux puissent nous être profitable.

Nous traiterons d'abord les travaux développés pour le contexte rural ou péri-urbain (section 1.3). Nous ferons la distinction entre les méthodes d'extraction linéique des routes (section 1.3.1) et celles d'extraction surfacique (section 1.3.2).

Puis nous nous intéresserons aux travaux, peu nombreux, développés pour le milieu urbain (1.4). Nous étudierons là aussi les méthodes d'extraction linéique des rues (section 1.4.1) pour finalement finir sur notre cadre d'étude, c'est-à-dire les méthodes d'extraction surfacique en milieu urbain (section 1.4.2).

Enfin, nous concluons cet état de l'art par les approches les plus prometteuses pour notre contexte (section 1.5).

1.3 Travaux en milieu rural ou péri-urbain

1.3.1 Travaux sur le linéique

A la fois pour des raisons historiques (la résolution spatiale des capteurs satellites) et parce que le sujet d'étude est ardu, beaucoup des travaux sur l'extraction de routes ont porté sur l'extraction de réseaux linéiques. Ces recherches ont donné un grand nombre d'approches et de résultats différents.

Seront aussi présentées dans cette section les méthodes sur les images à haute résolution spatiales ou aériennes dont l'objectif est d'extraire l'axe de la rue, souvent pour être inclus dans un SIG¹.

1.3.1.1 Extraction de lignes

L'extraction de lignes consiste à rechercher dans toute l'image les pixels ou les zones présentant les principales caractéristiques d'une rue. Ces méthodes n'assurent pas toujours l'exhaustivité ou l'absence de fausses détections. Certains de ces algorithmes peuvent être utilisés comme un «pré-traitement» permettant d'extraire des pixels ou des zones ayant une probabilité plus ou moins forte d'appartenir à des routes. Ils peuvent ensuite être couplés avec un algorithme d'interprétation de plus haut niveau qui pourra reconstruire la géométrie ainsi que la topologie du réseau (Pierrot-Deseilligny *et al.* 1993).

Le filtrage linéaire. Wang et Howarth (1987) appliquent des masques directionnels à leurs images afin de détecter et extraire les routes représentées par des lignes (minces ou épaisses), sur des images à 30 mètres de résolution spatiale. Les masques construits représentent les 8 directions possibles à suivre pour une ligne (en voisinage 8-connexité). La convolution de ces masques avec une image Landsat à 30 m de résolution spatiale et des critères de sélection (la ligne représente une route ou non) permettent d'extraire les amorces (pixels et direction à suivre) de l'algorithme de suivi de routes, qui fait suite à cette première étape de leur algorithme d'extraction du réseau routier.

La morphologie mathématique. Destival (1987) et Serendero (1989) utilisent des filtres non linéaires de morphologie mathématique pour extraire les lignes, brillantes par rapport à leur environnement, supposées correspondre à une route, à partir d'images pan-chromatiques (SPOT-HRV). La transformation dite du «chapeau haut de forme», puis un seuillage, sont appliqués afin d'obtenir une image binaire de pixels appartenant au réseau routier. Cette transformation permet de ne conserver de l'image originale que les pics ou les crêtes qui «entrent» (au moins dans une des directions) dans le chapeau supposé de hauteur infinie. Le résultat dépend de la largeur et/ou de la hauteur des sommets à conserver.

Les réseaux de neurones. Doucette *et al.* (2001) présente une approche inspirée des réseaux de Kohonen, SORM (*Self-Organising Road Map*). Le but est d'extraire l'axe central des rues sur des images haute résolution (moins de 2 m). La méthode a l'avantage de s'affranchir de la définition classique du bord des rues et d'exploiter l'information des canaux multispectraux (lors d'une phase de classification). Toutefois, l'algorithme montre

¹Système d'Information Géographique

ses limites lors qu'il rencontre des objets de radiométrie similaire à celle des routes (toits de bâtiments). Cette méthode est semi-automatique en ce sens où la classification qui sert d'entrée aux réseaux de neurones est supervisée par un opérateur humain (en plus de quelques paramètres de seuil à fixer).

Bhattacharya et Parui (1997) conçoivent l'extraction de routes sur les images de télédétection comme un problème de classification, où les pixels doivent être classés en «route» et «non route». La méthode a été appliquée sur des images acquises en bande proche infrarouge par le satellite IRS-1A (résolution spatiale 36,25 m) et en bande panchromatique par SPOT 4 (résolution spatiale 10 m).

Processus ponctuels de Markov. Stoica (2000) effectue une détection des réseaux linéiques dans les images SPOT en utilisant la géométrie stochastique. Afin de se soustraire à l'approche pixellique, le comportement des objets dans l'image est directement étudié. Le réseau linéique dans une image est ainsi modélisé par un réseau de segments. Les segments évoluent jusqu'au moment où ils trouvent le réseau dans l'image. L'évolution se fait en minimisant une fonction d'énergie à l'aide d'un algorithme de type recuit simulé utilisant une dynamique MCMC² à sauts réversibles. L'aspect géométrique du réseau (alignement, connectivité, dimension) est géré par le terme *a priori* (Candy model). Le terme permettant de trouver l'emplacement du réseau dans l'image, l'attache aux données, est désigné à partir de tests d'hypothèses : un segment de route est un ruban homogène le long d'une direction. La méthode est appliquée sur des images aériennes, SPOT et radar (RSO), et des milieux principalement ruraux.

1.3.1.2 Suivi de lignes

Les méthodes de suivi de lignes consistent à «suivre» la route en comparant ses caractéristiques avec un modèle théorique (souvent par corrélation des profils radiométriques de la route le long de son axe principal). Leur rôle est principalement d'assister en quasi temps-réel l'opérateur humain lors de la saisie du réseau routier. Les algorithmes de suivi nécessitent des amorces (souvent un point et la direction de la rue). Ces amorces peuvent soit être fournies par un opérateur, soit provenir d'un algorithme d'extraction en amont. Même si les objectifs sont identiques, les méthodes de suivi peuvent différer.

La programmation dynamique. Plusieurs méthodes de suivi de routes utilisent la programmation dynamique afin d'optimiser une fonction de coût, ce qui revient à la recherche d'un chemin optimal dans un graphe.

Merlet et Zérubia (1996) établissent leur détecteur de lignes sur une méthode de segmentation basée sur l'algorithme de Fishler qui utilise les champs Markoviens et une programmation dynamique. L'algorithme de Fischler permet d'obtenir, par construction, des lignes minces et continues. Cet algorithme est étendu à des cliques d'ordre supérieur pour tenir compte du contraste route/environnement, et à des voisinages d'ordre supérieur pour introduire la notion de courbure. De plus, un modèle dynamique permet d'utiliser cette information de courbure de manière plus globale et naturelle. Ce détecteur de lignes est appliqué à une image panchromatique à 10 m de résolution spatiale. C'est un traitement

²Monte Carlo Markov Chain

semi-automatique, où les points de départ de l'algorithme sont définis par l'utilisateur.

Jedynak (1995) adapte le jeu des «vingt questions» à la reconnaissance de réseaux routiers épais à partir d'images SPOT. Cet algorithme utilise un arbre de décision où chaque branche représente une direction possible pour un segment de la route recherché, en fonction du segment détecté précédent. Ici, le problème consiste à déterminer la stratégie optimale pour minimiser le nombre moyen de questions (branches) nécessaires à la recherche de la solution vraie (direction suivie par la route).

Gruen et Li (1995) utilisent eux aussi la programmation dynamique pour la recherche du chemin optimal. Les paramètres de contraste, de courbure ainsi que les points déjà détectés comme appartenant à la route sont utilisés pour trouver le point suivant de la route. Le suivi de route est précédé d'un rehaussement local des contrastes, effectué à l'aide d'une transformée en ondelettes adaptée aux caractéristiques radiométriques de la route. La méthode est semi-automatique au niveau de l'amorce.

Suivi par homogénéité directionnelle. L'algorithme utilisé par Airault et Jamet (1995) est une aide à la saisie du réseau routier. Il est appliqué à des images aériennes de 60 cm de résolution spatiale de paysages ruraux. Cet algorithme cherche à extraire précisément l'axe des routes. Sa stratégie comporte deux étapes.

Une étape de détection : elle est basée sur l'optimisation d'un critère d'homogénéité directionnelle de la variance locale. Cette étape est assurée par un algorithme de suivi de routes. Celui-ci recherche le chemin optimal parmi un arbre de chemins possibles (chaque nouveau pas a une liberté angulaire de $\pm 45^\circ$). Cette étape assure seulement que le tracé de l'axe se situe entre deux bords anti-parallèles. Les paramètres tels que la radiométrie, la largeur ou l'orientation des routes sont mis à jour dynamiquement durant le suivi.

Une étape de restitution : elle permet de recalculer le résultat de la précédente étape avec précision par rapport à l'axe de la route.

Cet algorithme est semi-automatique au niveau de l'amorce, un opérateur fournissant un point de départ et une direction de propagation à l'initialisation.

Filtrage de Kalman. Véran (1993) et Vosselman et De Knech (1995) basent leur algorithme de suivi de route sur le filtrage de Kalman.

L'approche de Véran (1993) est originale : il considère le suivi de route comme la poursuite d'un véhicule supposé circuler au milieu de la chaussée à une vitesse constante. Sa méthode est basée sur un filtrage de Kalman étendu et tient compte d'un (cas des carrefours) ou des deux bords de la route pour son suivi. Ces bords de routes sont détectés en utilisant un filtre de Canny-Deriche (Deriche 1987). La direction de la route est estimée par la direction du maximum d'homogénéité dans le voisinage du pixel concerné. Cet algorithme est semi-automatique : l'opérateur initialise le vecteur d'état initial et fixe certains paramètres, comme l'orientation initiale. L'information apportée par une carte topographique au 1/25 000 recalculée permet le passage des carrefours. L'auteur applique son algorithme sur des images aéroportées de 1 m de résolution spatiale.

Vosselman et De Knech (1995) utilisent un filtrage récursif de Kalman pour leur algorithme de suivi de routes. Celui-ci estime les paramètres (comme la largeur) décrivant la position et l'état de la route suivie à un instant donné. Ces paramètres sont estimés en comparant par moindres carrés le profil d'intensité de la route suivie à l'instant k avec un profil d'intensité moyen de segment de route. Le modèle de route prend en compte la radiométrie

(luminosité et contraste) et la géométrie (largeur et variation de largeur des routes). Cet algorithme est initialisé par un opérateur qui positionne un segment appartenant à la route recherchée. Cet algorithme est appliqué à des images aéroportées de 1,6 m de résolution spatiale.

Coopération d'algorithmes. Mc Keown et Denlinger (1988) utilise une méthode de coopération de plusieurs algorithmes de suivi de routes. Pour une même route, plusieurs algorithmes sont mis en œuvre : seul le meilleur résultat est conservé. Les deux algorithmes les plus utilisés par leur méthode sont le suivi par corrélation de profil et le suivi des bords de routes. L'algorithme de suivi par corrélation de profil utilise comme points de départ un pixel appartenant à la route, la direction et la largeur de la route considérée. Il est basé sur la corrélation qui existe entre le profil radiométrique à examiner et un profil de référence de la route. Si la correspondance entre les deux sections est pauvre, la recherche se poursuit dans une autre direction à partir du même point de départ jusqu'à trouver une bonne correspondance. Un des modèles de trajectoires possibles de la route donne les points permettant de créer la ou les sections à analyser. L'algorithme de suivi des bords de route permet de lier les points possédant une forte amplitude et une forte orientation du gradient dans la direction supposée de la route. Il s'inspire de la méthode de Nevatia et Babu (1980) mais utilise un opérateur gradient de Sobel pour une meilleure précision angulaire. La direction de la route recherchée est prévue par une extrapolation parabolique des directions prises par les profils analysés précédemment.

Ces deux algorithmes permettent d'établir, indépendamment l'un de l'autre, un modèle de l'axe de la route, de sa largeur et d'autres propriétés locales de la route. Un module intermédiaire permet alors d'établir le taux de réussite de chaque algorithme de suivi et d'extraire les différentes propriétés de la route considérée. Ce module permet également de suspendre un algorithme de suivi en cas d'échec et de le relancer à partir d'un modèle de route trouvé par l'autre algorithme.

Enfin, un module, dit de haut niveau, permet une description symbolique de la route en terme de position de l'axe, de la largeur, du revêtement de la route considérée, de la présence ou non de véhicules sur la chaussée, de ponts...

Cette méthode a été ensuite appliquée à des images aéroportées de 1 à 3,5 m de résolution spatiale.

1.3.1.3 Approches multi-sources.

L'approche multi-source en interprétation d'image a pour but l'utilisation de plusieurs sources d'information complémentaires pour parvenir à une extraction fiable des objets d'intérêt. Les informations peuvent être symboliques : une base de données cartographiques ou une carte topographique de la scène par exemple, ou bien numériques : différentes images d'un même site acquises par différents capteurs par exemple.

Guidée par une carte topographique. Pierrot-Deseilligny *et al.* (1993) utilise une carte scannée pour reconstruire le graphe du réseau de rues. Le système utilise l'information portée par le réseau partiellement extrait de l'image, ainsi que des *connaissances a priori* sur la topologie du réseau : connectivité globale et à l'échelle locale du réseau, non-redondance, absence d'extrémités libres, faible courbure, continuité sur le type de la route

(une autoroute ne se prolonge pas par un sentier). Une approche algorithmique de résolution est ensuite choisie. Elle comprend plusieurs étapes, ordonnées séquentiellement selon le sous-ensemble des règles sur les *connaissances a priori* qu'elles traitent.

Rellier *et al.* (2000) propose une méthode de mise en correspondance locale des cartes avec des images SPOT, en utilisant un champ markovien sur un graphe. La carte est utilisée pour initialiser le graphe. Ensuite, le graphe est modifié en minimisant une fonction d'énergie. L'énergie comporte des termes d'attache aux données relatifs à la carte et l'image, et un terme de structure imposant des contraintes topologiques. La méthode donne de bons résultats pour le recalage mais elle ne détecte pas les nouvelles routes par rapport à la carte et ne peut donc pas être utilisée pour la mise à jour.

Guidée par une base de donnée. Bordes (1997) propose un système d'interprétation d'images aériennes pour l'extraction des routes. Le système utilise des données externes fournies par une base de données cartographiques (la BD Carto[®]). La démarche générale d'interprétation proposée est une démarche descendante. La BD Carto[®] est utilisée pour générer des hypothèses de route. L'interprétation de la scène consiste alors à extraire dans l'image les routes correspondant à ces différentes hypothèses, c'est-à-dire à trouver la position exacte de la route dans l'image. Pour chaque hypothèse, l'extraction de route consiste à choisir l'algorithme le plus adapté et à régler ses paramètres en fonction des caractéristiques intrinsèques et contextuelles de l'hypothèse de route.

Utilisation de la troisième dimension spatiale. Zhang *et al.* (2000) présente le système d'interprétation développé dans le projet ATOMI (voir Eidenbenz *et al.* (2000)), dont le but est d'extraire l'axe central des routes à partir de cartes topographiques au 1/25 000 scannées. Le système utilise de nombreuses sources d'information qui sont complémentaires mais aussi redondantes. Notamment, des images aériennes stéréoscopiques avec plusieurs canaux sont utilisées. Les segments de route initialisés grossièrement dans l'image à partir de la carte vectorisée sont placés dans l'espace 3D, et les éléments linéaires dans leur voisinage sont alors extraits. Les segments parallèles sont ensuite déterminés en utilisant les interactions dans les espaces 2D (les images) et 3D. Les segments parallèles avec la même orientation dans l'espace 3D sont ensuite groupés. Enfin, la méthode utilise l'information sur la hauteur des objets de la scène provenant de MNT³ et MNE afin de lever les ambiguïtés et les occlusions sur les routes. La méthode est appliquée à des images aériennes et donne des résultats convaincants. Il est notamment possible d'extraire l'axe central par le marquage au sol lorsque l'extraction des bords de rues échoue. Les zones de tests sont exclusivement en environnements ruraux ou péri-urbains.

Gruen et Li (1997) combine l'ajustement aux moindres carrés avec la technique des contours actifs, ou *snakes*, de Kass *et al.* (1987, voir chapitre 2). Ils définissent ainsi un objet, le *LSB-Snake*, pour l'extraction de rues à partir d'images aériennes ou satellites. Les *LSB-Snakes* approximent les éléments linéaires par des courbes B-splines, et trois types d'observations sont utilisés pour décrire leurs propriétés. Le premier est un modèle décrivant un sommet radiométrique dans l'image. La relation entre ce modèle (*template* en anglais) et l'image est fixée par six paramètres modélisant la transformation géométrique. Les points d'amorce fournis par un opérateur représentent le deuxième type d'observations. Les contraintes

³Modèle Numériques de Terrain : décrit le profil en altitude du sol

de régularités du *LSB-Snake* sont le troisième type d'observations. La position finale du contour actif est calculée en minimisant selon la norme L_2 le résidu pour les trois types d'observation. Enfin, une représentation 3D des éléments linéaires est effectuée à partir de deux images ou plus de la scène, ainsi que d'un modèle numérique de terrain. Cette méthode a l'avantage de pouvoir déterminer elle-même la précision de l'extraction et, grâce aux *LSB-Snakes*, de s'affranchir des occlusions locales.

Utilisation des canaux multispectraux. Xiaoying et Davis (2003) extraient automatiquement le réseau routier à partir des canaux du satellite Ikonos (1 m de résolution spatiale en mode panchromatique, 4 m pour les canaux rouge, vert, bleu et proche infrarouge). La résolution des canaux multispectraux est amenée à 1 m par un procédé de fusion d'images.

L'algorithme se compose de quatre phases séquentielles. Une phase de pré-traitement pendant laquelle les zones avec de la végétation sont «masquées» afin de réduire la région d'intérêt pour l'extraction du réseau routier. Le masquage s'effectue en calculant l'indice de végétation (le NDVI⁴) et en éliminant les pixels dont l'indice est supérieur à 0,3. Le même procédé est appliqué pour masquer les étendues d'eau (avec le canal vert). Ensuite des amorces de routes sont sélectionnées à partir d'une segmentation de l'image (par K-mean) et d'un groupement perceptuel. Les objets linéaires de l'image sont alors détectés à partir du canal panchromatique et du canal NDVI en utilisant les propriétés géométriques définies dans (Steger 1998). Enfin, un suivi de route par l'algorithme de Vosselman et De Knech (1995, voir section (1.3.1.2)) est effectué, toujours en utilisant le canal panchromatique et le NDVI.

Le procédé est tout-automatique et atteint des taux de complétude entre 70% et 90%. Toutefois, la méthode est appliquée sur des scènes où il y a peu de bruit (ombres, véhicules. . .) et en milieu rural boisé où le calcul du NDVI s'avère discriminant.

1.3.1.4 Les systèmes d'interprétation

Dans le cadre de la télédétection, les systèmes d'interprétation (nommés aussi systèmes d'exploitation et de valorisation d'images) ne cherchent pas directement à extraire les objets d'une image, mais plutôt à interpréter la scène qu'elle contient. Ils se basent sur le principe de pensée d'un interprète d'images face à une recherche d'objets particuliers dans une image.

Dans son algorithme d'extraction linéique du réseau routier, Khatir (1988) assimile les amorces de lignes détectées à des composantes connexes. La recherche des composantes qui peuvent représenter des routes s'effectue selon certains critères basés sur une vision générale des routes. L'algorithme utilise les trois critères suivants :

- les routes mènent à des zones urbaines ou des villages,
- les routes ont une courbure régulière,
- les routes ont une radiométrie en général plus claire que leur environnement.

A ces trois critères sont associées des probabilités de validité. Grâce à une phase d'apprentissage utilisant comme vérité-terrain la carte correspondant à la zone étudiée, une

⁴Normalized Difference Vegetation Index.

Il est défini comme : $NDVI = \frac{NIR - R}{NIR + R}$, NIR étant le canal proche infrarouge et R le canal rouge du visible.

composante connexe peut être définie comme un certain type de route avec une certaine probabilité. Le seuillage de ces différentes probabilités permet de ne conserver que les composantes connexes représentant des routes de manière sûre. L'algorithme est appliqué sur une image SPOT (10 m de résolution spatiale).

Garnesson (1991) utilise un système-expert, appelé MESSIE, de type tableau noir pour extraire les routes sur des images aéroportées de 40 et 86 cm de résolution spatiale. Le système utilise le principe suivant :

- détection des contours basée sur une segmentation sélective de l'image (recherche des régions claires par exemple),
- chaînage sur l'image binaire des contours et élimination des chaînes trop courtes,
- approximation des chaînes par des segments de droite,
- pour finir, un groupement perceptuel pour extraire les routes en se basant sur des attributs de proximité, de similarité, de continuation et de symétrie.

Nous pouvons aussi inclure les systèmes d'interprétation de Mc Keown et Denlinger (1988), Jedynak (1995), Eidenbenz *et al.* (2000) et de Bordes (1997) décrits précédemment.

1.3.2 Travaux sur le surfacique

Les approches qui ont été évoquées précédemment ont en commun d'extraire le réseau routier de manière linéique. Sur les images satellites basse ou moyenne résolution (de type SPOT 4, Landsat TM), cette approche provient directement de l'aspect de la route dans l'image. Sur les images à haute résolution spatiale (images aériennes ou des nouveaux capteurs satellites à très haute résolution spatiale), les méthodes avaient pour but d'extraire l'axe de la route, souvent pour être intégré dans un SIG. Extraire l'axe de la route n'implique pas systématiquement l'estimation de sa largeur. Par exemple dans Zhang *et al.* (2000), l'extraction de l'axe central peut se faire en détectant le marquage au sol. De plus, ces approches ne prennent pas en compte la forme des intersections.

Nous présentons dans cette section les méthodes qui extraient les réseaux routiers de manière surfacique, donc à partir d'images aériennes ou satellites à haute résolution spatiale.

Approches par contours actifs. La technique des contours actifs (Kass *et al.* 1987, voir chapitre 2.1) a plusieurs avantages dans le cadre de l'extraction de rues. Un modèle géométrique des routes à extraire peut être intégré dans la fonctionnelle d'énergie guidant les contours actifs. De plus, les contraintes globales sur toute la longueur du contour actif permettent de ne pas prendre en compte les points s'écartant trop de la forme à extraire dans le cas d'occlusions locales (comme l'ombre portée des arbres). Les contours actifs doivent cependant être initialisés près de la route à extraire.

Fua et Leclerc (1990) ont utilisé les contours actifs pour la détection des routes et des bâtiments à partir d'images aériennes. Les auteurs montrent que les courbes ayant minimisé leur fonctionnelle d'énergie, c'est-à-dire les solutions pour lesquelles la courbure est faible ou l'amplitude du gradient est constante, sont de bonnes approximations des contours. De plus, ils donnent une méthode adaptative pour déterminer les paramètres du modèle d'énergie. Les courbes sont modélisées comme des rubans dont les bords sont des courbes parallèles. Un terme traduisant cette contrainte est ajouté à l'énergie d'attraction.

Neuenschwander *et al.* (1997) présentent une stratégie d'optimisation, les *Ziplock Snakes*,

afin de réduire la sensibilité à l'initialisation des contours actifs. L'utilisateur positionne manuellement les extrémités du *Ziplock Snake*, qui sont ensuite rapprochées des contours selon l'intensité de l'image et son gradient. Le principe de la méthode est alors de diviser le contour actif en deux parties : une active et une passive. Pendant l'optimisation, les forces «images» sont appliquées uniquement sur les parties actives du contour actif, alors que la partie passive n'est soumise qu'à des contraintes géométriques. Les forces images sont ensuite propagées de proche en proche par analogie avec une fermeture éclair (d'où le nom *Ziplock*). La méthode assure ainsi que les parties actives restent proches du contour à extraire et que les structures non pertinentes ne viennent pas perturber l'extraction. La méthode est ensuite appliquée pour extraire les bords de rues sur des images aériennes à partir de points d'initialisation très distants. L'algorithme est toutefois mis en défaut dans le cas des intersections. En pratique, l'utilisateur peut rectifier l'extraction en rajoutant un nouveau point de contrôle au *Ziplock Snake*.

Simulation de propagation de rayons. Guigues et Viglino (2000) présentent une approche originale d'extraction automatique de réseaux routiers guidée par un réseau routier imprécis (comme la BD Carto[®]). La méthode s'appuie sur deux éléments principaux : un détecteur des structures parallèles d'une image, fonctionnant par simulation de propagations et réflexions de rayons lumineux, et un algorithme fonctionnant par programmation dynamique, qui construit à partir du parallélisme précédemment détecté un réseau de routes hypothétiques. Un vote impliquant de nombreux critères est alors utilisé pour sélectionner le parcours le plus probablement emprunté par la route. Cette méthode est avant tout un procédé bas-niveau qui nécessite par la suite une interprétation par un système plus haut-niveau, notamment pour le traitement des intersections. Toutefois, le système fonctionne correctement en zones rurales, et se révèle prometteur dans les cas plus complexes comme pour les routes à plusieurs voies ou les autoroutes.

Utilisation de l'information 3D. Wang et Trinder (1998, 2003) utilisent l'information tridimensionnelle avec une approche similaire à celle de Gruen et Li (1997). La méthode est aussi basée sur la détection dans l'espace 2D (ici deux images) des paires de segments parallèles. Ces paires sont positionnées dans l'espace 3D par des contours actifs (similaires aux *LSB-SNakes* de Gruen et Li). Un groupement complexe des segments, prenant en compte à la fois des contraintes spatiales, géométriques et radiométriques, permet de lever les ambiguïtés et les occlusions locales comme les véhicules sur la route. La méthode est appliquée sur des images à 50 cm de résolution spatiale et sur des milieux principalement ruraux.

Shaoh *et al.* (2003) se focalisent sur l'extraction des carrefours (et aussi des bâtiments) à partir de paires stéréoscopiques d'images aériennes. Ils définissent quatre niveaux d'abstraction : le point, le segment, la surface et enfin l'intersection, reconstruite dans l'espace 3D. L'extraction a lieu de manière ascendante hiérarchique et est guidée par les relations topologiques à tous les niveaux d'abstraction.

Approches multi-échelles. Nous l'avons vu (section 1.1.3.2), la représentation de la rue dans l'image dépend de la résolution spatiale de l'image : sur les images à très haute résolution spatiale (meilleure que de 5 m par pixel), les rues apparaissent

comme des surfaces allongées et assez homogènes. Sur des images avec une moins bonne résolution spatiale, les rues sont des éléments linéaires de l'image. De plus, en fonction de la résolution, certaines caractéristiques de la route, comme le marquage au sol ou le terre-plein central, sont visibles ou non.

Les travaux de Baumgartner *et al.* (1996), Baumgartner (1998), Baumgartner *et al.* (1999), ont été les premiers à prendre en compte les caractéristiques multi-échelles de la rue. Ils définissent des modèles de rues afin de contraindre l'extraction sur les régions qui ont le plus de chances de contenir des rues de manière fiable. Ces modèles prennent en compte le contexte (rural, péri-urbain ou urbain) et les interactions avec les autres objets (bâtiments, arbres, véhicules). A partir de l'image de résolution dégradée, les lignes les plus fiables représentant les rues sont extraites par l'algorithme de Steger (1998) et sont utilisées pour extraire les segments candidats de l'image à pleine résolution. A partir de critères géométriques, radiométriques et du contexte, les segments parallèles extraits sont ensuite appariés, les segments proches sont alors groupés et les discontinuités sont comblées de manière à rendre le réseau connexe. Enfin, les intersections sont reconstruites, en utilisant la technique des *Ziplock Snakes* (Neuenschwander *et al.* 1997). Le procédé est tout automatique et donne de très bons résultats en milieu rural. Pour des milieux plus bruités comme en forêt ou en milieu urbain, l'extraction automatique est incomplète et nécessite l'intervention d'un opérateur humain ou l'utilisation d'algorithmes d'extraction ou de suivi utilisant le contexte comme le marquage au sol ou les véhicules.

Laptev *et al.* (2000) reprennent l'approche multi-échelle des travaux précédents et se concentrent sur le traitement des occlusions (en utilisant les contours actifs définis par Fua et Leclerc 1990) et des intersections (en utilisant les *Ziplock Snakes* de Neuenschwander *et al.* 1997). Le procédé se révèle efficace pour combler les «trous» laissés par l'étape de détection des segments parallèles, ainsi que pour extraire les intersections avec précision. La méthode est évaluée quantitativement avec les critères définis par Heipke *et al.* (1997). Les exemples d'extraction présentés ont une précision meilleure ou plus faible que le pixel par comparaison à une référence manuellement extraite. Le domaine d'application de l'algorithme est principalement le domaine rural. Les échecs de la méthode apparaissent pour les rues n'ayant pas une largeur constante (pour des résolutions fines), que l'on rencontre plus souvent en environnement urbanisé.

1.4 Travaux dans le contexte urbain

La plupart des méthodes qui ont été décrites dans les sections précédentes sont applicables au contexte rural, voire péri-urbain, mais montrent leurs limites en environnement urbain. Le milieu urbain est en effet un environnement complexe (voir section 1.1.2) où il est difficile de développer des algorithmes robustes et fiables, en particulier pour le cas tout-automatique.

Dans cette section nous présentons les travaux portant spécifiquement sur l'extraction des réseaux de rues en milieu urbain, que ce soit pour l'extraction linéique ou surfacique. Ces travaux sont récents et encore peu nombreux comparés aux méthodes développés sur les milieux ruraux ou péri-urbains. De plus, au vu de la difficulté du contexte, ils se restreignent pour la plupart au cas de réseaux quadrangulaires (en grille).

1.4.1 Les travaux sur le linéique

Wang *et al.* (1996) appliquent des masques directionnels à leurs images afin de détecter et extraire les rues représentées par des lignes (minces ou épaisses), sur des images à 10 m de résolution spatiale. L'application de filtres directionnels permet de détecter les rues quelles que soient leurs directions (horizontales, verticales ou diagonales) ou leurs courbures (faibles ou prononcées). La taille de ces filtres dépend de l'épaisseur de la ligne recherchée. Douze filtres de convolution sont utilisés au total, ce qui permet d'extraire à partir d'une image panchromatique à 10 m de résolution spatiale les rues de 2 à 3 pixels de large.

Haverkamp (2002) propose une méthode pour extraire de manière presque automatique (la largeur des rues est à fixer par l'utilisateur) les réseaux en grille à partir d'images du satellite Ikonos. L'algorithme comprend deux étapes. La première utilise les canaux multispectraux de manière similaire à Xiaoying et Davis (2003) : les zones avec de la végétation sont «masquées» en calculant l'indice de végétation NDVI et en éliminant les pixels au-dessus d'un certain seuil. La deuxième étape est basée sur le calcul à partir de l'image panchromatique de la texture autour de chaque pixel : s'il présente une homogénéité directionnelle de texture suffisante dans deux directions ou plus, il est retenu comme faisant partie du réseau. Ensuite, suivant qu'un pixel appartient à une rue ou à une intersection, la structure du réseau peut être extraite. La méthode permet d'identifier 82,9% des rues avec un écart-type d'un pixel en terme de localisation. Toutefois, elle est spécifique aux réseaux en grille rencontrés dans les villes américaines, et est mise en échec en présence de bruit comme les véhicules.

Shi et Zhu (2002) partent du principe qu'une rue peut être considérée comme un ensemble de segments contenu à l'intérieur d'un long ruban étroit. Un seuillage binaire de l'image est tout d'abord effectué pour simplifier la détection. Pour une orientation donnée, les segments parallèles sont détectés (chaîne de plus de 30 pixels à «1») puis agglomérés en fonction de critères de longueur, de parallélisme et de proximité. Cette opération est répétée dans toutes les directions (3×180 directions au total). Des opérations de morphologie mathématique (fermeture et squeletisation) permettent ensuite d'obtenir un réseau linéique et connexe. La méthode ne nécessite pas de points d'amorce mais plusieurs

seuils doivent être fixés de manière empirique par l'utilisateur.

Une méthode semi-automatique de détection de la voirie urbaine (Couloigner *et al.* 1997) a été appliquée à des images multispectrales issues des canaux multibandes HRV des données simulées de SPOT 1-3 (20 m de résolution spatiale) et du canal simulé de SPOT 5 (10 m de résolution spatiale). Cette méthode est basée sur une classification ascendante hiérarchique des canaux multibandes des données. Le dépouillement visuel de l'arborescence hiérarchique des classes présentes dans les images de tests permet de ne conserver que les classes estimées appartenir au réseau routier urbain. En final de cette détection, un squelette de la voirie est créé, toutefois assez incomplet. Les résultats de cette méthode ont été améliorés par l'utilisation d'une méthode de fusion de données (Couloigner et Ranchin 1998). La méthode de fusion de données permet d'augmenter la résolution spatiale des canaux multispectraux par injection d'informations à haute résolution spatiale provenant du canal panchromatique (Wald 2002, Ranchin *et al.* 2003). Les images multibandes synthétisées par cette méthode ont une résolution spatiale de 10 m pour SPOT 1-3 et de 5 m pour SPOT 5.

1.4.2 Les travaux sur le surfacique

Price (2000) se concentre sur l'extraction surfacique des réseaux en grille à partir d'images de résolution variant entre 20 et 80 cm. Le réseau de rues est modélisé comme étant une combinaison de «grilles» avec une taille de maille régulière (la taille d'un bloc de bâtiments). Les intersections sont les nœuds de la grille formée par des segments de rue. Après une initialisation manuelle de trois points définissant la première maille, la grille est étendue par itérations successives en ajoutant de nouvelles mailles et en ajustant leurs segments sur les sommets radiométriques de l'image. Pendant la phase finale de vérification, l'information d'altitude fournie par un modèle numérique d'élévation est incorporée de manière à recalculer précisément le réseau et à éliminer les mauvais segments. Cette phase utilise l'information contextuelle qui est qu'en ville, les rues représentent les minima du MNE.

Couloigner et Ranchin (2000), Péteri *et al.* (2001) présentent une méthode semi-automatique pour extraire les réseaux routiers urbains par hiérarchie de classes de rues. L'algorithme est appliqué sur des images entre 0,85 et 10 m de résolution spatiale et se focalise aussi sur les réseaux en grille. Il comprend deux étapes. La première étape est l'extraction des deux bords de la rue en utilisant les approximations de l'image à différentes résolutions spatiales (par application de l'algorithme «à trous», voir section 2.3.2). Le procédé consiste à rechercher les points d'intersection des profils transversaux de la rue sur l'image à la résolution originale et les deux premières approximations. L'utilisateur doit fournir les points d'initialisation et la classe de la rue à extraire. L'algorithme suit ensuite la rue dans une procédure «estimation-recalage» utilisant pour chaque point supposé appartenir à la rue les profils radiométriques provenant de la décomposition à différentes résolutions. La deuxième étape permet ensuite d'extraire les caractéristiques de la rue (terre-plein central, terre-plein secondaire). Ce procédé itératif utilise à cette étape les images de coefficients d'ondelettes pour localiser, s'ils sont présents, le terre-plein central ou les terre-pleins secondaires.

La précision géométrique des résultats est en moyenne de deux pixels par rapport à une

référence définie manuellement, quelle que soit la résolution des images originales. L'intérêt de la méthode réside dans l'utilisation de plusieurs résolutions de l'image pour réduire les effets du bruit ainsi que dans la définition de plusieurs classes génériques de rues. La méthode n'est toutefois applicable que sur les réseaux de rues droites et la topologie du réseau n'est pas gérée.

Nous concluons cet état de l'art par les travaux de Hinz *et al.* (2001), Hinz et Baumgartner (2003) dont le système d'interprétation développé pour l'extraction en milieu urbain est actuellement l'un des plus complets. Les auteurs constatent eux aussi le peu de travaux jusqu'à présent consacrés en contexte urbain. Leur méthode synthétise leur longue expérience dans la recherche sur l'extraction de routes (Heipke *et al.* 1995; Baumgartner *et al.* 1996; Baumgartner 1998; Wiedemann *et al.* 1998; Steger 1998; Baumgartner *et al.* 1999; Laptev *et al.* 2000). Elle incorpore aussi des innovations prenant en compte le cadre complexe qu'est le milieu urbain.

Les entrées du système sont plusieurs vues aériennes de la scène se superposant (au dessus de 20 cm de résolution spatiale), ainsi qu'un MNT assez précis (entre 1-2 m). La méthode est automatique et présente trois caractéristiques principales.

- Les stratégies automatiques d'extraction et les modèles de rues sont adaptés en fonction de leur contexte. Les stratégies d'extraction vont dépendre du contexte global de la rue. Trois types de contexte sont utilisés : urbain, forêt et rural (Baumgartner *et al.* 1999). Les modèles de rues prennent en compte leur contexte local, comme la présence de bâtiments ou de véhicules ou le marquage au sol. Par exemple, un contexte «occlusion» entraînera l'utilisation d'une autre vue de la scène où la rue sera visible. Des modèles multi-échelles des rues sont aussi utilisés (Heipke *et al.* 1995).

- Plusieurs vues de la même scène sont utilisées. Cela permet d'utiliser les redondances d'information, de prédire et de résoudre les problèmes d'occlusion, ainsi que d'avoir une représentation des rues dans l'espace 3D.

- La méthode d'extraction comporte des modèles internes qui évaluent la pertinence des différentes hypothèses pour la suite du traitement. Par exemple, si une hypothèse de rue est identifiée sur plusieurs vues se recouvrant, une valeur importante de confiance lui sera affectée.

Sur les exemples présentés, la méthode arrive à une détection de 75% des rues et de 5 % de fausses alarmes. Pour une approche tout automatique et un contexte très bruité, ces résultats sont bons, et la méthode est notamment robuste vis-à-vis des occlusions.

Plusieurs points sont encore en cours d'amélioration comme l'extraction des intersections ou la détection des véhicules. De plus, les auteurs précisent que les paramètres ont été optimisés empiriquement. Enfin, ils reconnaissent aussi ce thème est encore à l'étape de la recherche fondamentale et qu'il reste encore beaucoup à faire afin d'atteindre un stade pleinement opérationnel.

1.5 Conclusions sur ce chapitre

Au cours de ce chapitre, la grande variabilité typologique des rues dans l'image a été présentée. Plusieurs raisons ont été avancées, notamment le changement de résolution des nouveaux capteurs à très haute résolution spatiale. Cette grande variabilité de la rue a conduit à un grand nombre de méthodes d'extraction, utilisant un large champ des techniques de reconnaissance de formes et de traitement d'images. Plusieurs points sont ressortis de cet état de l'art.

Caractérisation des performances. Avant tout, la manière de caractériser les résultats d'une méthode d'extraction de rues va dépendre du contexte de la méthode et de ses objectifs. Ainsi, pour un algorithme tout-automatique d'extraction, les critères de taux d'omission et de fausses alarmes sont les critères principaux. Pour un algorithme de suivi dont l'amorce provient d'un opérateur, et qui est souvent à but opérationnel, on privilégiera plutôt la précision géométrique du résultat.

Suivant la difficulté de la scène, un taux de bonne détection de 85% peut être considéré comme un très bon résultat pour un algorithme tout automatique, alors que cela sera jugé insuffisant dans un cadre opérationnel, où l'on requiert plus de 95% de bonnes détections. De plus, l'apport de données externes (MNE, base de données), et à un degré moindre l'utilisation de plusieurs canaux spectraux, peuvent grandement améliorer les résultats d'extraction. Enfin, il n'existe actuellement pas de protocole standard ni de banques d'images de référence pour pouvoir comparer les performances des différents algorithmes d'extraction de rues (même si des organismes comme EuroSDR ou l'IGN y travaillent actuellement). Il est donc difficile de comparer quantitativement les méthodes et de les classer par ordre de performance. Par contre, on peut analyser leurs causes d'échec et identifier comment certaines méthodes ont réussi à les surmonter.

Les causes d'échec. Parmi les causes d'échec, on trouve en premier lieu les échecs dus à l'écart des données par rapport au modèle théorique choisi. Il peut s'agir d'écarts par rapport à un modèle radiométrique (la rue n'apparaît pas dans l'image comme une bande homogène), des écarts par rapport à un modèle géométrique (non constance du parallélisme des bords ou de largeur de la rue), ou, plus rarement, des écarts par rapport au modèle topologique (réseau non connexe). Ces écarts aux modèles sont principalement dus au «bruit» apparaissant sous formes multiples : principalement le bruit d'occlusion (véhicules sur la chaussée, ombres portées des arbres et des bâtiments), mais aussi le bruit d'acquisition.

Plus le bruit et le niveau de détails augmentent, plus il devient difficile de définir un modèle assez générique pouvant s'adapter à la grande variabilité d'aspects de la rue dans l'image. Ainsi, un grand nombre des méthodes présentées montrent leurs limites lorsque le milieu s'urbanise et donc lorsque le bruit s'accroît. Notre contexte d'étude étant le milieu urbain et la très haute résolution spatiale (autre facteur d'augmentation du bruit), il est malgré tout possible de dégager de l'état de l'art des approches prometteuses pour ce contexte.

Les approches prometteuses pour le milieu urbain et la haute résolution spatiale. Les causes d'échecs sont donc dues en partie à des écarts par rapport au modèle théorique. Il faut ainsi distinguer la représentation de la rue dans l'image et sa

réalité sous-jacente. Il est donc important que l'algorithme d'extraction puisse inclure toutes les représentations dans l'image d'un même objet rue et donc gérer les occlusions partielles de la rue.

L'utilisation de données externes, notamment des MNE (Zhang *et al.* 2000; Hinz et Baumgartner 2003) ou l'utilisation de plusieurs images (Wang et Trinder 1998; Wang et Trinder 2003), constituent un apport pour la cartographie, notamment pour lever les ambiguïtés créées par les artefacts. Toutefois, comme nous l'avons mentionné dans le chapitre d'introduction, ces données ne sont pas toujours disponibles pour les cabinets de topographes ou dans les pays en voie de développement. De plus, leur coût d'obtention est élevé. Dans notre cadre opérationnel, nous ne travaillons qu'à partir d'une seule image. Nous considérerons donc les méthodes basées sur l'image comme unique donnée permettant d'extraire la rue en tant qu'élément de surface dans un environnement bruité.

Parmi celles-ci, on peut dégager un certain nombre d'approches prometteuses pour notre contexte.

- L'utilisation du contexte (présence de files de voitures sur la rue, marquage au sol ...) peut permettre de lever certaines ambiguïtés ou servir d'indicateurs de confiance pour l'algorithme. Cependant, de par son aspect aléatoire dans l'image, il doit être utilisé comme une aide pour l'algorithme et non comme une condition nécessaire à la présence d'une rue.
- Permettant de se soustraire en partie du bruit, le traitement multi-échelle de l'image peut être un grand apport pour le milieu urbain (Couloigner et Ranchin 2000; Hinz *et al.* 2001).
- Il est préférable en milieu bruité de développer un modèle de rue basé sur des critères géométriques forts (comme le parallélisme des bords de la rue) plutôt que sur des critères purement radiométriques (comme le profil de la rue).
- Il est important d'introduire des connaissances *a priori* sur les rues pour contraindre le modèle. C'est ce que permettent les contours actifs de manière souple. Plusieurs travaux ont montré leur intérêt dans le cadre de l'extraction de routes (Wang et Trinder 1998; Laptev *et al.* 2000; Wang et Trinder 2003), et donc potentiellement pour le milieu urbain.
- Enfin, il semble qu'un seul type d'approche ne soit pas suffisante. Seule l'association de plusieurs méthodes permettra d'accroître la robustesse de l'algorithme vis-à-vis du bruit, notamment, la coopération de plusieurs modèles du réseau de rues (linéique, surfacique) ou de plusieurs approches de traitement d'images (approche contour ou approche région).

Partant de ces remarques, le prochain chapitre propose un procédé pour extraire les tronçons de rues à bords parallèles (chapitre 2). Son utilisation dans la méthode globale d'extraction de rues à partir d'images à très haute résolution spatiale sera ensuite l'objet du chapitre 3.

Chapitre 2

Extraction de bords parallèles par contours actifs et analyse multi-échelle

Dans ce chapitre, nous définissons une nouvelle classe de contours actifs intégrant l'analyse multi-échelle, qui va permettre d'extraire les sections de rue à bords parallèles. Leur utilisation dans le cadre de la méthode globale d'extraction de rues développée sera ensuite l'objet du chapitre 3.

2.1 Objectifs

Les intersections mises à part, le réseau routier est composé de tronçons de rues délimités par des contours parallèles (au moins localement). Dans un environnement bruité, le parallélisme des bords de rues représente une caractéristique géométrique plus fiable que l'homogénéité radiométrique. Afin d'extraire de manière globale tout le réseau routier, nous avons tout d'abord cherché à extraire les sections de rue à bords parallèles. Notre but est de fournir une aide à l'extraction de rue à partir d'une image numérique à très haute résolution spatiale d'une scène urbaine. Le modèle d'extraction de bords parallèles retenu devait remplir un certain nombre de critères :

Être générique. L'aspect des rues dans l'image présente une grande variabilité (section 1.1). Différentes classes de rues peuvent être rencontrées en milieu urbain (avenue, autoroute, rue, ruelle, ...). Le modèle doit aussi être générique vis-à-vis de la résolution spatiale de l'image, pouvant varier entre 70 cm et 5 m.

Gérer les artefacts. Nous nous situons dans un environnement, le milieu urbain, générant beaucoup d'artefacts. Ces artefacts (véhicules sur la chaussée, ombres portées des arbres et des bâtiments, marquage au sol ...) provoquent des occlusions locales de la rue à extraire. Il convient donc de pouvoir s'en affranchir ou du moins de les minimiser.

Inclure des connaissances *a priori*. Afin de guider un modèle d'extraction de formes, il est important de pouvoir lui incorporer les connaissances *a priori* sur l'objet à

extraire. Dans le cas des rues, la courbure est régulière et relativement faible, et les variations sur la largeur et le parallélisme des bords sont lentes.

Être évolutif. Dans l'optique de développements futurs, le modèle doit permettre d'inclure de manière souple de nouvelles contraintes ou de nouveaux modèles.

Être rapide. Enfin, la solution retenue doit permettre un gain de temps par rapport à une interprétation manuelle de la scène.

En tenant compte de ces critères, un modèle d'extraction de bords parallèles a été développé, faisant intervenir des contours actifs combinés à une analyse multi-échelle.

Les contours actifs constituent un outil bien approprié au problème de l'extraction de rues (Fua et Leclerc 1990, Gruen et Li 1997, Laptev *et al.* 2000, Agouris *et al.* 2001). Ils peuvent s'adapter à la forme à extraire ainsi que franchir des occlusions locales. De plus, par leur structure, les contours actifs permettent d'introduire de manière souple des contraintes, en définissant des termes spécifiques. Enfin, ils permettent d'adopter une approche hiérarchique descendante et donc d'intégrer les connaissances *a priori* sur les rues et le réseau.

Toutefois, les contours actifs sont très dépendants de leur initialisation et peuvent être piégés par des minima locaux d'énergie. L'utilisation de l'analyse multi-échelle permet d'accroître la robustesse vis-à-vis du bruit sur la rue (Baumgartner *et al.* 1999, Couloigner et Ranchin 2000) et d'accélérer la convergence de l'algorithme.

Dans les sections suivantes, une présentation générale des contours actifs et de la variante utilisée dans la méthode seront exposées (section 2.2). La transformée en ondelettes et son lien avec l'analyse multi-échelle seront ensuite présentés (section 2.3).

Ce cadre mathématique rigoureux permettra de définir le terme d'attache aux données des contours actifs (section 2.4). Un nouveau type de contour actif, le *DoubleSnake*, sera introduit, permettant l'extraction des contours à bords parallèles dans un environnement bruité (section 2.5).

Enfin, quelques exemples sur des scènes synthétiques et des scènes réelles seront présentés afin d'illustrer les propriétés des *DoubleSnakes* (section 2.6).

2.2 Les contours actifs

2.2.1 Introduction

Les contours actifs (appelés couramment *snakes*) ont été introduits par Kass, Witkin, et Terzopoulos (1987). Leur principe théorique est d'approximer les objets à extraire par des modèles élasto-dynamiques. Ces derniers vont se déformer en étant attirés par des énergies images telles que les contours et en étant forcés à suivre des contraintes de régularité sur leur forme. Les contours actifs sont des courbes, fermées ou non, qui doivent être initialisées à proximité de la forme à détecter, par un opérateur ou un algorithme de pré-traitement.

Le procédé d'extraction revient ensuite à un problème d'optimisation d'une fonctionnelle d'énergie sous contraintes.

On distingue généralement deux catégories de contours actifs : les modèles continus et les modèles discrets. Les modèles continus représentent le contour de manière analytique. Les points d'initialisation vont servir de points de contrôle à la courbe analytique générée par approximation de type *spline*. Dans ce cas, la minimisation de l'énergie est équivalente à la résolution de l'équation d'Euler (voir Kass *et al.* 1987) et pourra être traitée selon une approche variationnelle.

Dans le cas du modèle discret, le contour actif est représenté par une ligne polygonale définie par des noeuds et des segments connectant ces noeuds. Les relations géométriques et radiométriques entre ces noeuds sont traduites en termes de fonctions d'énergie, et l'extraction de l'objet revient à un problème d'optimisation de la fonctionnelle d'énergie sur le support discret de l'image.

Notre choix s'est porté pour le modèle des contours actifs sur une approche purement discrète, plus simple et plus rapide en terme d'implémentation. De plus, l'implémentation numérique des modèles continus induit naturellement un processus de discrétisation faisant intervenir un échantillonnage régulier du contour, alors que le modèle discret n'impose nullement cette contrainte.

La prochaine section portera donc sur les contours actifs discrets et le modèle développé. Pour une présentation détaillée des contours actifs continus et de leur modèles évolutifs, on pourra se référer à l'ouvrage collectif de Cocquerez et Philipp (1995).

2.2.2 Contours actifs discrets

Le modèle de contour actif discret que nous avons choisi est celui décrit par Williams et Shah (1992).

Il reprend le modèle originel (Kass *et al.* 1987) en redéfinissant les termes posant problème, notamment l'énergie de continuité.

2.2.2.1 Les termes d'énergie

L'évolution du contour actif est régie par la fonctionnelle suivante :

$$E = \sum_i [\alpha(i)E_{cont}^i + \beta(i)E_{curv}^i + \gamma(i)E_{image}^i] \quad (2.1)$$

Les deux premiers termes de l'équation (2.1) correspondent à l'énergie interne E_{int} de régularisation. Ils maintiennent la continuité entre les points du contour actif en contrôlant la longueur de chaque segment ainsi que la courbure en chaque point.

Le terme de continuité \mathcal{C}^0 s'écrit :

$$E_{cont} = \bar{d} - |v_i - v_{i-1}| \quad (2.2)$$

où \bar{d} est la distance moyenne entre tous les points du contour, recalculée à chaque itération, et i représente l'indice du point v_i où est évaluée cette énergie.

Le second terme de l'équation (2.1) contrôle la courbure du contour. L'énergie correspondante est calculée selon :

$$E_{curv} = |v_{i-1} - 2v_i + v_{i+1}|^2 \quad (2.3)$$

Cette fonction donne une bonne approximation de la courbure et est rapide en temps de calcul.

Le dernier terme d'énergie E_{image} de l'équation (2.1) correspond à l'énergie image. C'est le terme d'attache aux données, qui guide le contour actif vers l'objet à extraire. Son rôle est donc très important et son expression sera détaillé à la section 2.4.

Enfin, les termes $\alpha(i)$, $\beta(i)$ et $\gamma(i)$ sont des coefficients de pondération affectés à chacune des énergies au point considéré. Ils permettent de privilégier un terme plutôt qu'un autre. Ces paramètres sont la plupart du temps pris constants sur toute la courbe.

2.2.3 Procédure d'optimisation par algorithme glouton

Dans le modèle proposé par Williams et Shah (1992), la minimisation de la fonctionnelle d'énergie des contours actifs va être effectuée par un algorithme dit *glouton*. Les algorithmes gloutons (en anglais : *greedy algorithms*) construisent une solution de façon incrémentale, en choisissant à chaque étape la direction qui est localement la plus prometteuse.

Les auteurs montrent que leur approche permet d'obtenir des résultats comparables à l'algorithme de minimisation exacte par programmation dynamique de Amini *et al.* (1988), tout en ayant un temps d'exécution beaucoup plus rapide. La complexité par programmation dynamique est en $O(nm^3)$ (avec n : le nombre de points du contour et m : la taille du voisinage de recherche) contre $O(nm)$ pour l'algorithme glouton.

Nous avons donc choisi cette minimisation par algorithme glouton, car dans notre contexte d'assistance à l'interprétation, le temps d'exécution est un paramètre important.

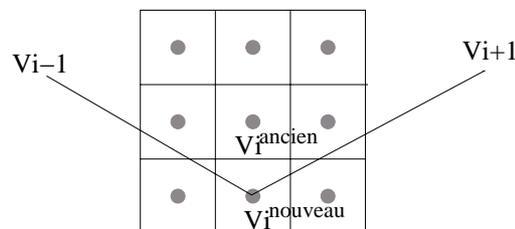


FIG. 2.1: procédure d'optimisation

D'un point de vue algorithmique, l'énergie est calculée en chaque point du voisinage du nœud traité (sur la figure 2.1, le voisinage est 3×3). Le nœud v_{i-1} a déjà été optimisé et s'est déplacé vers sa nouvelle position. Le nœud suivant v_{i+1} ne s'est pas encore déplacé. Les positions de ces deux points sont utilisées avec les différents emplacements possibles pour v_i afin d'évaluer les différents termes d'énergie. Le point de la fenêtre de recherche minimisant l'énergie est alors retenu pour v_i . La procédure est répétée tout le long du contour actif, ceci définissant une boucle. A la fin de chaque boucle, les critères d'arrêt sont évalués (comme le nombre de points à bouger) et l'énergie totale est mise à jour. Lorsque les critères d'arrêt sont satisfaits, le contour actif est alors considéré dans sa position d'équilibre et la boucle s'arrête. Les nœuds finaux du contour actif sont alors utilisés pour créer les segments du contour final. Le pseudo-code de l'algorithme de Williams et Shah (1992) se trouve en annexe (C).

2.3 Transformée en ondelettes et analyse multi-échelle

Les contours actifs sont très dépendants de leur initialisation et peuvent être piégés par des minima locaux d'énergie. L'utilisation de la transformation en ondelettes et de l'analyse multi-échelle va permettre d'accroître la robustesse vis-à-vis des coefficients de pondération et du bruit, ainsi que d'accélérer la convergence de l'algorithme.

2.3.1 La transformée en ondelettes continue

Comme la transformée de Fourier, la transformée en ondelettes permet de décomposer une fonction $f(x) \in L^2(\mathbb{R})^1$ sur une base de fonctions élémentaires : *les ondelettes*. Cette base, ou famille, est générée par translations et dilatations d'une unique fonction Ψ appelée *ondelette-mère*.

Dans le cas unidimensionnel, cette famille est de la forme :

$$\left\{ \Psi_{a,b}(x) = \frac{1}{\sqrt{|a|}} \Psi\left(\frac{x-b}{a}\right), (a, b) \in \mathbb{R}^* \times \mathbb{R} \right\} \quad (2.4)$$

a est appelé *pas de dilatation*, et b *pas de translation*.

Les ondelettes sont des fonctions oscillantes bien définies à la fois en temps et en fréquence. Elles possèdent trois propriétés fondamentales communes : la localisation, l'oscillation et la régularité et satisfont toutes à une condition d'admissibilité. Dans le cas unidimensionnel, la transformée en ondelettes d'une fonction $f(x) \in L^2(\mathbb{R})$ s'écrit :

$$WT_f(a, b) = \int_{-\infty}^{+\infty} f(x) \overline{\Psi_{a,b}(x)} dx = \langle f(x), \Psi_{a,b}(x) \rangle \quad (2.5)$$

avec $\overline{\Psi}$ le complexe conjugué de l'ondelette analysante Ψ , et \langle, \rangle désignant le produit

¹ ensemble des fonctions de carré intégrable

scalaire de $L^2(\mathbb{R})$.

Pour une échelle a et une position b , $WT_f(a, b)$ représente l'information contenue par $f(x)$ pour cette échelle et à cette position. Elle fournit donc la décomposition locale de l'information d'une fonction de carré intégrable.

Inversement, on peut reconstruire exactement $f(x)$ en intégrant sur toutes les échelles et à toutes les positions de la manière suivante :

$$f(x) = \frac{1}{C_\Psi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} WT_f(a, b) \Psi_{a,b}(x) \frac{da db}{a^2} \quad (2.6)$$

avec la condition d'admissibilité de l'ondelette mère :

$$C_\Psi = \int_0^{+\infty} |\hat{\Psi}(\nu)| \frac{d\nu}{\nu} < +\infty, \hat{\Psi} \text{ étant la transformée de Fourier de } \Psi.$$

L'équation 2.6 peut être interprétée de deux manières différentes :

- on peut reconstruire $f(x)$ quand on connaît sa transformée en ondelettes,
- on peut écrire $f(x)$ comme une superposition d'ondelettes.

Ces deux points de vue conduisent à des applications différentes qui sont, d'une part le traitement, et d'autre part l'analyse des signaux.

Pour les applications à l'image numérique, la transformée en ondelettes est étendue à deux variables d'espace, et on utilise sa version discrète. La discrétisation passe par celle des paramètres de dilatation et de translation a et b . Un cas très utilisé est lorsque les paramètres valent $a = 2^j$ et $b = k2^j$, $(k, j) \in \mathbb{Z}^2$, on parle alors transformée en ondelettes *dyadique*.

2.3.2 La transformée en ondelettes «à trous»

L'algorithme «à trous» (Holschneider *et al.* 1989) est une transformée en ondelettes dyadique rapide qui s'implémente par des bancs de filtres. Shensa (1992) a décrit son lien avec l'analyse multirésolution, présentée en annexe A. Cette implémentation est très proche de la transformée en ondelettes (bi)orthogonales rapide, mais sans sous-échantillonnage.

Cas monodimensionnel Étant donné un filtre F de coefficients $F[n]$, on note $F_j[n]$ le filtre obtenu en insérant 2^{j-1} zéros entre chaque coefficient de F (d'où le nom d'algorithme à trous), et on pose : $\bar{F}_j[n] = F_j[-n]$.

Les données échantillonnées $a_0[n]$ d'un signal monodimensionnel f peuvent être considérées comme le produit scalaire au point n de f par une fonction d'approximation, ou fonction d'échelle ϕ , jouant le rôle d'un filtre passe-bas.

On a alors :

$$a_0[n] = \langle f(x), \phi(x - n) \rangle \quad (2.7)$$

On définit aussi une ondelette mère ψ par sa projection dans le sous-espace engendré par la fonction ϕ , et vérifiant les conditions d'admissibilité.

Dans la version monodimensionnelle de la transformée en ondelettes à trous, ϕ étant la fonction d'échelle et ψ l'ondelette mère, nous avons $\forall j \geq 0$:

$$a_j[n] = \langle f(x), \phi_{2^j}(x - n) \rangle \text{ et } d_j[n] = \langle f(x), \psi_{2^j}(x - n) \rangle \quad (2.8)$$

$$\text{avec } \phi_{2^j}(x) = \frac{1}{2^j} \phi\left(\frac{x}{2^j}\right) \text{ et } \psi_{2^j}(x) = \frac{1}{2^j} \psi\left(\frac{x}{2^j}\right).$$

a_j est l'approximation de a à l'échelle j .

d_j représente les valeurs discrètes de la transformée en ondelettes. Ils portent l'information sur les détails perdus lors du passage de l'échelle $j - 1$ à l'échelle j .

La transformée en ondelettes à trous peut être implémentée par un banc de filtres. Elle est alors définie selon les relations de récurrence :

$$a_{j+1}[n] = a_j * \bar{h}_j[n] \quad (2.9)$$

$$d_{j+1}[n] = d_j * \bar{g}_j[n] \quad (2.10)$$

où le filtre \bar{h} est un filtre passe-bas associé à la fonction d'échelle ϕ et \bar{g} est un filtre passe-haut associé à la fonction ondelette mère ψ .

Enfin, il est possible de reconstruire le signal à l'échelle j à partir des approximations et des plans de détails à l'échelle $j + 1$:

$$a_j[n] = \frac{1}{2} (a_{j+1} * \tilde{h}_j[n] + d_{j+1} * \tilde{g}_j[n]) \quad (2.11)$$

Les filtres \tilde{h} et \tilde{g} sont les filtres duaux permettant de reconstruire le signal f (voir Shensa (1992), Starck (1992), Mallat (1997) pour plus de détails sur la reconstruction).

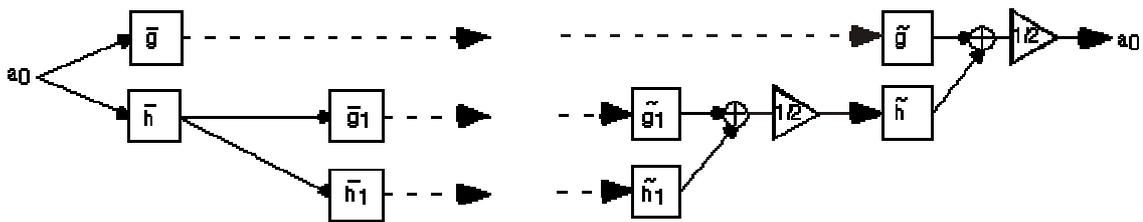


FIG. 2.2: schéma des bancs de filtres pour l'algorithme à trous

La figure 2.2 représente la décomposition et la reconstruction de l'algorithme à trous par bancs de filtres.

Cas bidimensionnel Pour l'application de l'algorithme «à trous» aux images numériques, on choisit en pratique une fonction d'échelle à variables séparées :

$$\phi_{2^j}(x, y) = \phi_{2^j}(x) * \phi_{2^j}(y) \quad (2.12)$$

Pour tout niveau d'échelle $j \geq 0$, on pose :

$$a_j[n, m] = \langle f(x, y), \phi_{2^j}(x - n)\phi_{2^j}(y - m) \rangle \quad (2.13)$$

et on assimile les données discrètes de l'image aux $a_0[n, m]$.

Les coefficients d'ondelettes discrets sont alors :

$$d_j^1[n, m] = W_{2^j}^1 f(n, m) \quad (2.14)$$

$$d_j^2[n, m] = W_{2^j}^2 f(n, m) \quad (2.15)$$

L'algorithme à trous permet alors de calculer la transformée dyadique rapide de la manière suivante :

$$a_{j+1}[n, m] = a_j * \bar{h}_j \bar{h}_j[n, m], \quad (2.16)$$

$$d_{j+1}^1[n, m] = a_j * \bar{g}_j \delta_j[n, m], \quad (2.17)$$

$$d_{j+1}^2[n, m] = a_j * \delta_j \bar{g}_j[n, m]. \quad (2.18)$$

δ_j désigne la fonction de Dirac discrète dont le $j^{ième}$ coefficient vaut 1 et les autres 0.

Les approximations de l'image originale (les plans de contextes) sont donc obtenues par convolutions successives par les filtres \bar{h} de lissage.

La famille des fonctions B-Splines constituent une base d'interpolation à support compact que l'on peut choisir aussi régulière que l'on veut (Bijaoui *et al.* 1994, Brigger et Unser 1998). Les fonctions d'échelle et d'ondelette que nous utiliserons sont des splines d'ordre 2 (fig. 2.3), qui permettent une analyse quasi-isotrope, tout en restant compactes et régulières à l'ordre 1 inclus.

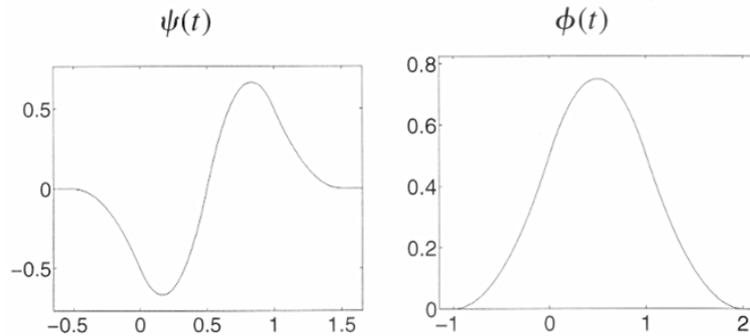


FIG. 2.3: ondelette et fonction d'échelle spline quadratique (d'après Mallat 1997)

Le tableau 2.1 rassemble les coefficients pour les filtres de décomposition et de reconstruction correspondant.

n	$h[n]/\sqrt{2}$	$h[n]/\sqrt{2}$	$g[n]/\sqrt{2}$	$\tilde{g}[n]/\sqrt{2}$
-2				-0.03125
-1	0.125	0.125		-0.21875
0	0.375	0.375	-0.5	-0.6875
1	0.375	0.375	0.5	0.6875
2	0.125	0.125		0.21875
3				0.03125

TAB. 2.1: coefficients des filtres pour les fonctions d'échelle et d'ondelette spline quadratique (Mallat 1997)

Pour produire une dilatation de la fonction d'échelle, un zéro est ajouté entre chaque coefficient des filtres \bar{h} et \bar{g} . Ainsi, il n'y a aucun sous-échantillonnage de l'image originale et l'application de l'algorithme à trous produit donc à chaque itération trois images de même taille, ce qui facilite l'analyse des images.

Remarque : dans la version la plus simple et la plus couramment utilisée, les coefficients d'ondelettes sont obtenus par soustraction pixel à pixel de deux plans de contexte successifs. Dans ce cas, l'algorithme produit à chaque itération un plan de contexte et une image non-directionnelle de coefficients d'ondelettes.

Contrairement à une analyse multirésolution (voir annexe A), la représentation par l'algorithme à trous est redondante à travers les échelles (base d'ondelettes non orthogonale) et est sur-échantillonnée au sens de Shannon. Elle accroît donc la quantité d'information. Toutefois, ce critère n'est pas déterminant car nous utilisons la transformée en ondelettes pour analyser le signal et non pour le compresser (où dans ce cas la décorrélation entre les échelles est essentielle). De plus, l'algorithme à trous constitue un algorithme rapide (le nombre de convolutions est constant pour toutes les échelles) et, du fait de la non-décimation, est invariant par translation.

2.4 Contours actifs et énergie multi-échelle par ondelettes

De manière à accroître la robustesse vis-à-vis du bruit, l'énergie image de la fonctionnelle des contours actifs (formule 2.1) va être évaluée à différentes échelles. Pour cela, on effectue une représentation multi-échelle des contours de l'image associée à une transformation en ondelettes (Mallat et Zhong 1992; Mallat 1997).

2.4.1 Détection multi-échelle de contours

Une approche pour détecter les contours d'une image consiste à filtrer l'image par la dérivée première d'un noyau de lissage θ d'une certaine largeur de bande. Si ce noyau de lissage est une gaussienne, il s'agit du détecteur de Canny (1986).

Il existe un lien très fort entre les coefficients fournis par un filtre de Canny, et la transformée en ondelettes à une échelle donnée. Ainsi, comme présenté par Mériaudeau *et al.* (1995), une famille d'ondelettes peut être générée à partir du filtre généralisé de Canny-Deriche (Bourennane *et al.* 1993).

L'analyse par ondelettes permet de caractériser et d'extraire les singularités d'une image à différentes résolutions spatiales (Mallat et Zhong 1992; Mahmoodi *et al.* 1997; Mallat 1997). Son application aux images de télédétection permet de conserver le lien physique existant entre la résolution spatiale de l'image et la taille des objets d'intérêt (Bijaoui *et al.* 1994; Li et Shao 1994; Ranchin 1997). Ceci nous amènera à définir des classes de rues, qui seront présentées au chapitre 3. Nous avons choisi de plus de nous placer dans le cadre mathématique bien défini de l'analyse multi-échelle par ondelettes, car elle permet de faire évoluer la méthode en définissant de nouvelles approches (caractérisation des singularités de l'image, indices de texture). Elle est en ce point plus générique que la convolution par une série de gaussiennes dont on ferait varier le paramètre d'échelle.

Une version multi-échelle du détecteur de Canny (1986) est ainsi implémentée dans le cas d'une analyse dyadique (Mallat et Zhong 1992). Pour cela, on considère une ondelette en dimension deux, définie par dérivation partielle d'un noyau de lissage θ . La fonction de lissage est choisie comme convergeant vers 0 en l'infini avec une intégrale égale à 1. Ces conditions imposent à l'ondelette d'être bien définie (leur intégrale est nulle).

Le noyau de lissage va être dilaté itérativement et joue le rôle de la fonction d'échelle dans l'analyse multi-échelle.

Les deux composantes en x et en y de la transformée dyadique sont alors proportionnelles au gradient de la convolution du signal avec le noyau θ :

$$\begin{pmatrix} W_{2^j}^1 f(x, y) \\ W_{2^j}^2 f(x, y) \end{pmatrix} = 2^j \begin{pmatrix} \frac{\partial}{\partial x} (f * \theta_{2^j})(x, y) \\ \frac{\partial}{\partial y} (f * \theta_{2^j})(x, y) \end{pmatrix}$$

Soit :

$$\begin{pmatrix} W_{2^j}^1 f(x, y) \\ W_{2^j}^2 f(x, y) \end{pmatrix} = 2^j \vec{\nabla} (f * \theta_{2^j})(x, y)$$

Le module du gradient est donc proportionnel au module de la transformée en ondelette :

$$\|\mathbf{W}_{2^j} \mathbf{f}(\mathbf{x}, \mathbf{y})\| = \sqrt{|W_{2^j}^1 f(x, y)|^2 + |W_{2^j}^2 f(x, y)|^2}$$

Les points de contours multi-échelles sont les points où le module de la transformée en ondelettes dyadique est maximal dans la direction pointée par le vecteur précédent.

La figure 2.4 illustre cette extraction multi-échelle des contours. A partir de l'image originale, les composantes de la transformée en ondelettes $W_{2^j}^1 f(x, y)$ et $W_{2^j}^2 f(x, y)$ ainsi que son module $\|\mathbf{W}_{2^j} \mathbf{f}(\mathbf{x}, \mathbf{y})\|$ sont calculés pour différentes échelles. Les petits détails (correspondant aux hautes fréquences spatiales, disparaissent lorsque la résolution de l'image diminue sous l'effet du lissage.

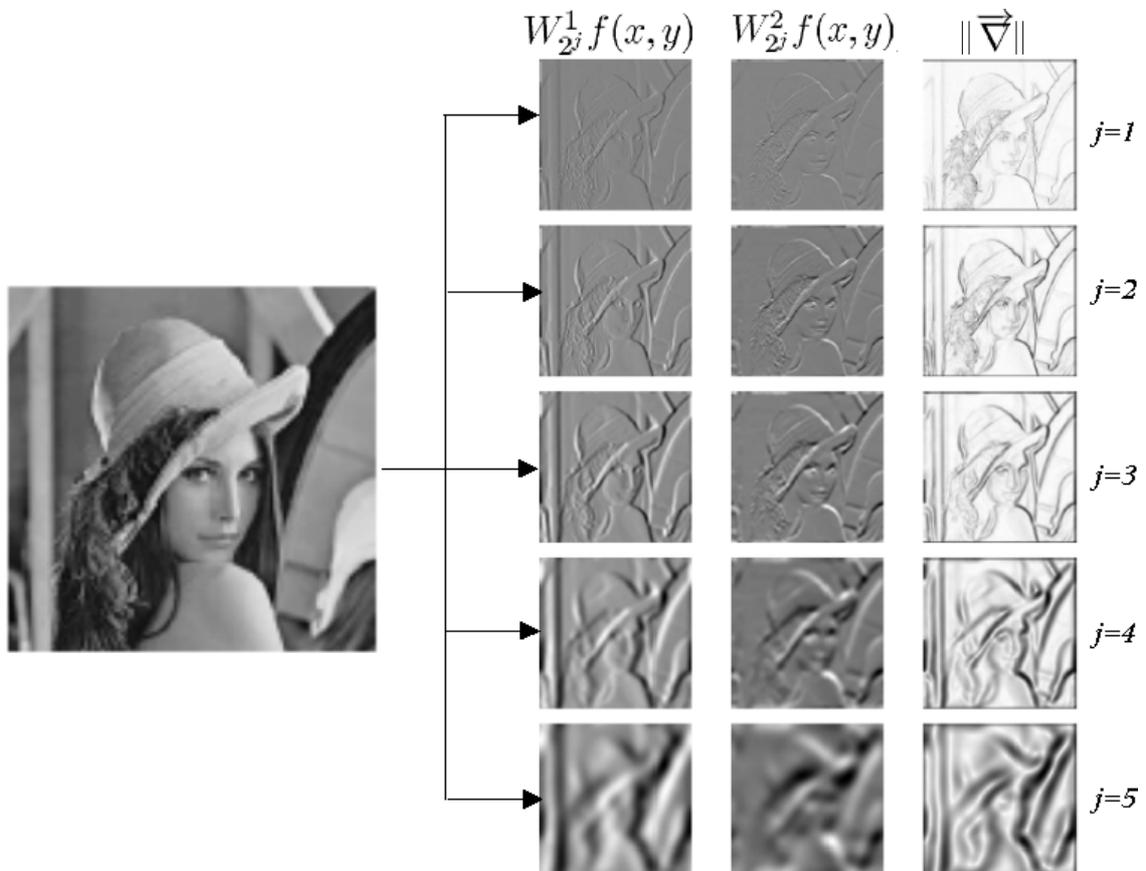


FIG. 2.4: composantes et module de la transformée en ondelettes à différentes échelles

2.4.2 Énergie multi-échelle des contours actifs

Pour les filtres que nous avons présentés (tableau 2.1), les composantes en x et en y de la transformée en ondelettes représentent les approximations par des différences finies des dérivées partielles suivant x et y du noyau de lissage défini par $\theta_{2^j}(x, y) = 4\phi_{2^j}(2x)\phi_{2^j}(2y)$. Plus précisément, le filtre $\bar{g}_j[n]$ associé à la transformée en ondelettes approxime une dérivée translatée de 2^j . Il faut donc compenser cette translation lorsque l'on évalue le module du gradient correspondant à la transformée en ondelettes (comme décrit dans la section 2.4.1).

Le module du gradient est donc défini pour l'échelle j comme :

$$\|\mathbf{W}_{2^j}\mathbf{f}(\mathbf{x}, \mathbf{y})\|^2 = |d_j^1[n - 2^{j-1}, m]|^2 + |d_j^2[n, m - 2^{j-1}]|^2 \quad (2.19)$$

On définit alors, pour le contour actif, une énergie image relative à chaque plan de contexte. Elle est calculée en utilisant les coefficients d'ondelettes à différentes échelles :

$$E_{2^j}^i{}_{image} = -\frac{1}{2^j} \sqrt{|W_{2^j}^1 f(i)|^2 + |W_{2^j}^2 f(i)|^2} \quad (2.20)$$

avec $W_{2^j}^{1,2} f(i)$ les coordonnées de la transformée en ondelettes, i l'indice du point considéré et j le niveau de l'échelle.

On a donc finalement :

$$E_{2^j}^i{}_{image} = -\frac{1}{2^j} \sqrt{|d_j^1[n - 2^{j-1}, m]|^2 + |d_j^2[n, m - 2^{j-1}]|^2} \quad (2.21)$$

2.5 *DoubleSnakes* pour l'extraction des contours parallèles

2.5.1 Principe

L'extraction des segments de rue à contours parallèles présente des spécificités. En effet, la rue est un objet dont les principales particularités sont le contraste par rapport à son environnement et le parallélisme de ses bords.

Le modèle de contour actif développé, le *DoubleSnake*, prend en compte ces connaissances *a priori* (Péteri *et al.* 2003; Péteri et Ranchin 2003). Le modèle est composé de deux contours actifs multi-échelles ouverts (ici appelés brins du *DoubleSnake*), possédant les contraintes d'évolution des contours actifs simples, tout en interagissant afin de garder un parallélisme entre eux. Cette condition sur le parallélisme est assortie d'une condition sur la distance entre les deux brins.

La base de l'algorithme repose sur l'algorithme glouton (Williams et Shah 1992) décrit section 2.2.3. Le calcul de l'énergie du *DoubleSnake* comprend un nouveau terme $E_{//}$, correspondant à une énergie forçant le *DoubleSnake* à maintenir ses deux brins localement parallèles.

La fonctionnelle d'énergie d'un des deux brins du *DoubleSnake* est alors définie pour l'échelle j comme :

$$E_j = \sum_i \left[\alpha^i E_{cont}^i + \beta^i E_{curv}^i + \gamma^i E_{2^j image}^i + \delta^i E_{//}^i \right] \quad (2.22)$$

où i représente le point i d'un des brins.

E_{cont}^i et E_{curv}^i sont les énergies internes de régularisation sur la forme du *DoubleSnake*.

L'énergie image est calculée (voir section 2.4) en utilisant les coefficients d'ondelettes $W_{2^j}^{1,2} f(i)$ à différentes échelles j :

$$E_{2^j image}^i = -\frac{1}{2^j} \sqrt{|W_{2^j}^1 f(i)|^2 + |W_{2^j}^2 f(i)|^2}. \quad (2.23)$$

Le terme $E_{//}^i$ correspond à l'énergie de parallélisme, et est décrit dans la section suivante.

2.5.2 Énergie de parallélisme

Le terme $E_{//}^i$ vise à prendre en compte la notion de parallélisme entre les deux brins du *DoubleSnake* et à assurer la régularité de la forme de la rue.

Ainsi, dans le principe, si un des deux brins est piégé par le bruit sur la rue dans un minimum local de son énergie image, il pourra en être «sorti» par l'autre brin dont l'énergie image est plus stable et correspond réellement au bord de la rue.

En milieu urbain, il arrive que la largeur de la rue varie, par exemple lorsqu'il y a un élargissement de la rue aux abords d'un carrefour. Le parallélisme des bords n'est alors plus respecté sur l'ensemble de la rue. Pour cette raison, et pour avoir aussi plus de souplesse sur la forme des rues à extraire, nous avons privilégié un parallélisme local, dont la taille du voisinage peut varier (pouvant à la limite être global). On pourra ainsi «propager» les variations locales sur la largeur de la rue, le long du contour actif.

Une hypothèse est faite pour pouvoir évaluer l'énergie de parallélisme $E_{//}$: les deux brins qui composent le *DoubleSnakes* doivent avoir le même nombre de noeuds. Cette contrainte n'est pas trop forte car cette condition est remplie à l'initialisation des *DoubleSnakes* (le procédé d'initialisation des *DoubleSnakes* sera exposé au chapitre 3) et de plus, les rues ont généralement des bords de même longueur.

Cette condition permet d'apparier les points deux à deux, sans avoir à calculer pour chaque point son homologue le plus proche sur le brin symétrique.

La contrainte locale de parallélisme est illustrée figure 2.5. Ici, la taille du voisinage est fixée aux deux plus proches voisins du point considéré P_i . On cherche à minimiser la différence de direction et d'amplitude entre le vecteur \mathbf{V}_i et les vecteurs \mathbf{V}_{i-1} et \mathbf{V}_{i+1} .

Lorsque le *DoubleSnake* évoluera, il essayera de minimiser cette différence, gardant ainsi un parallélisme local entre ses deux brins ainsi qu'une distance similaire.

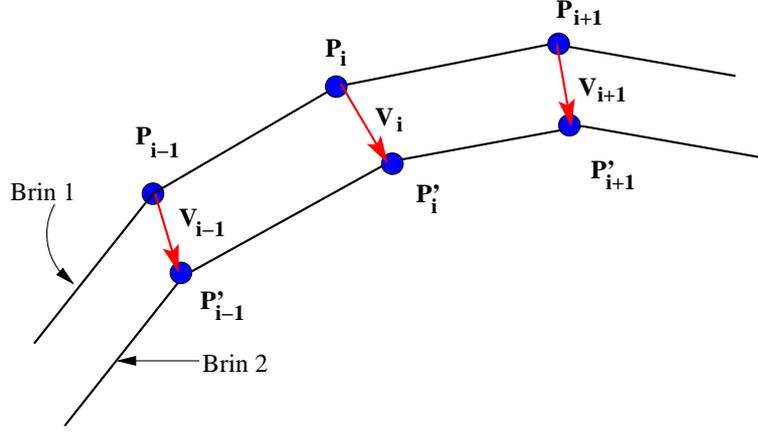


FIG. 2.5: évaluation du parallélisme des deux brins lors de l'extraction des contours parallèles

On cherche donc à avoir pour le cas général d'un voisinage $[-M, +M]$ au point i (ici pour le brin 1) :

$$\sum_{j=i-M}^{j=i+M} (\mathbf{V}_i^1 - \mathbf{V}_j^1) \rightarrow \mathbf{0}$$

On définit alors l'énergie de parallélisme comme :

$$E_{//}^i = \frac{1}{2M} \sum_{j=i-M}^{j=i+M} (|V_{i_x}^1 - V_{j_x}^1| + |V_{i_y}^1 - V_{j_y}^1|) \quad (2.24)$$

Pour les points extrêmes des brins pour lesquelles la formule 2.24 ne s'applique pas, on restreint alors le voisinage.

L'énergie de parallélisme est ensuite multipliée par un coefficient $\delta(i)$ afin de pondérer son influence dans la fonctionnelle d'énergie.

2.5.3 Paramètres des *DoubleSnakes*

Les paramètres pondérant les différents termes d'énergie des *DoubleSnakes* dans l'équation 2.22 sont fixés de manière empirique sur toute l'image. De manière à réduire leur nombre, ils sont les mêmes pour tous les *DoubleSnakes* intervenant dans une scène, et sont choisis constants sur tout le contour actif.

Leur adaptation à la résolution de l'image ou au type de tissu urbain pourrait se faire au travers d'une étude de sensibilité des résultats à ces paramètres. Leur détermination automatisée paraît cependant difficile. On pourrait alors penser à effectuer un grand nombre de tests sur différents cas (différentes typologies de réseau, différents capteurs satellites) pour extraire des paramètres statistiquement représentatifs.

2.6 Applications

D'un point de vue de l'évolution des *DoubleSnakes* à une échelle j fixée, tout revient à considérer pour une itération une seule fonctionnelle d'énergie, comme si on avait mis les deux brins bout-à-bout. On optimise ensuite la fonctionnelle point à point suivant le principe de l'algorithme glouton (section 2.2.3).

Sur les exemples qui suivent, l'énergie de parallélisme est évaluée à partir des deux voisins les plus proches du point considéré.

2.6.1 Effet de l'énergie de parallélisme sur un *DoubleSnake*

La figure (2.6) illustre l'effet seul de l'énergie de parallélisme sur un *DoubleSnake*, tous les autres termes d'énergie étant mis à zéro. Depuis leur position initiale (figure 2.6.a) les deux brins ont évolué en s'influçant mutuellement pour arriver à une position finale d'équilibre (figure 2.6.b). Les deux brins sont parallèles : le vecteur reliant deux points appariés d'un brin à l'autre présente toujours la même norme et la même direction.

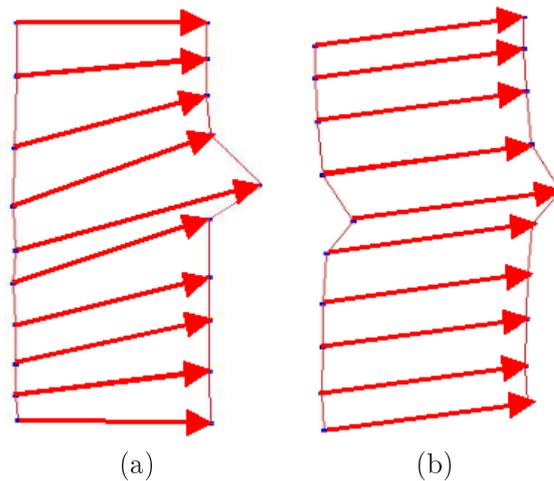


FIG. 2.6: effet de l'énergie de parallélisme sur un *DoubleSnake*

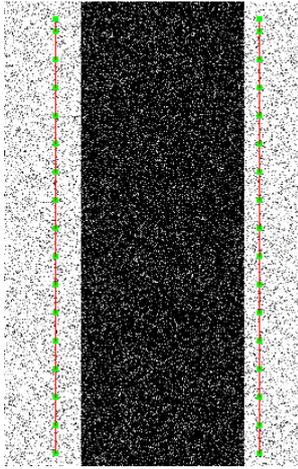
Les valeurs numériques confirment l'état d'équilibre atteint (tableau 2.2) : les résidus (0,01 pour le coefficient directeur et 1,16 pour la norme) pouvant s'expliquer par la discrétisation spatiale imposée par la grille pixel.

Ecart-type σ	Position initiale	Position finale
Ecart-type sur la norme	18,79	1,16
Ecart-type sur les coefficients directeurs	0,21	0,01

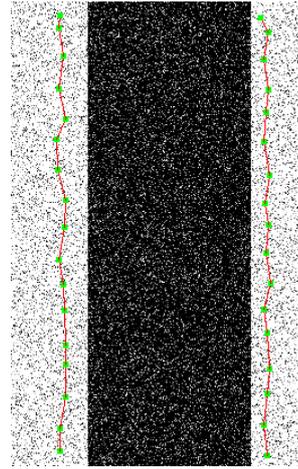
TAB. 2.2: écart type sur les vecteurs liés à l'énergie de parallélisme

2.6.2 Effet de l'approche multi-échelle

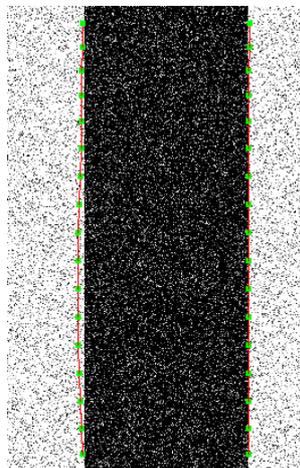
L'exemple suivant illustre l'efficacité de l'approche multi-échelle par rapport un calcul direct du gradient sur l'image pour le calcul de l'énergie image des contours actifs. Sur la figure 2.7 sont présentés deux cas avec la même initialisation (figure 2.7(a)). Il s'agit d'une image synthétique bruitée composée d'une bande noire à détecter.



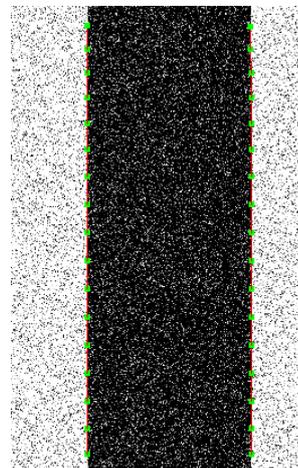
(a) Initialisation



(b) Résultat *sans* l'approche multi-échelle



(c) Résultat intermédiaire (échelle 2)



(d) Résultat final *avec* l'approche multi-échelle

FIG. 2.7: influence de l'approche multi-échelle sur les résultats

Le premier cas (figure 2.7(b)) présente le résultat d'extraction sans l'utilisation de l'approche multi-échelle pour le calcul de l'énergie image. A la fin de leur évolution, les brins du contour actif sont restés éloignés de la bordure à détecter : les points du contour actif

se sont fait piéger par le bruit et n'ont pas été influencés par l'objet à détecter, trop distant.

Dans le second cas (figures 2.7(c) et 2.7(d)), l'approche multi-échelle a été appliquée pour le calcul de l'énergie image. Le traitement multi-échelle a été effectué jusqu'à l'échelle $j = 4$.

La figure 2.7(c) présente le résultat intermédiaire d'extraction, après l'évolution des contours actifs jusqu'à l'échelle 2. Jusqu'à cette étape, l'énergie image a été prédominante, ce qui a permis aux contours actifs de converger rapidement vers l'objet à extraire. L'évolution des contours actifs après l'échelle 2, est ensuite principalement guidée par les termes de régularisation sur la forme. Le résultat final d'extraction est présenté figure 2.7(d) : le contour actif a été capable de détecter précisément les contours de la bande noire, malgré la présence du bruit.

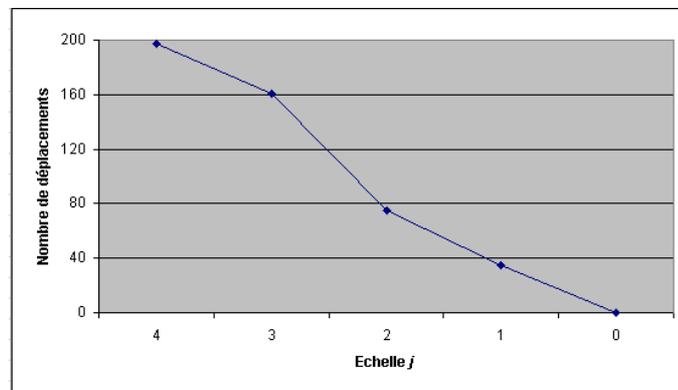
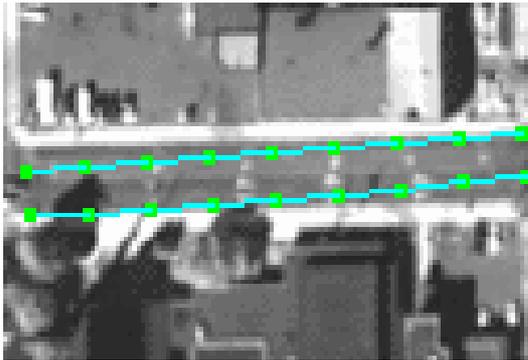


FIG. 2.8: nombre de déplacements de points en fonction de l'échelle. $j = 0$ correspond à l'image à la résolution initiale.

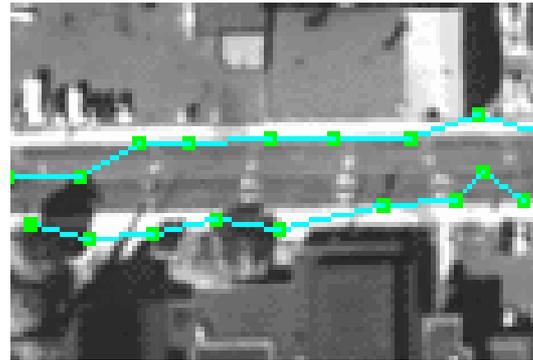
Le graphique 2.8 indique, pour chaque échelle, le nombre de points déplacés lors de la minimisation de la fonctionnelle d'énergie des contours actifs. Ce nombre est tout d'abord très élevé dans les hautes échelles où l'énergie image prédomine et tend à faire converger rapidement les contours actifs. Le nombre de points déplacés décroît ensuite rapidement, lorsque les contours actifs subissent principalement l'influence de l'énergie de régularisation, jusqu'à arriver à une position d'équilibre à l'échelle 0, pour laquelle plus aucun point ne bouge.

2.6.3 Application à une scène réelle

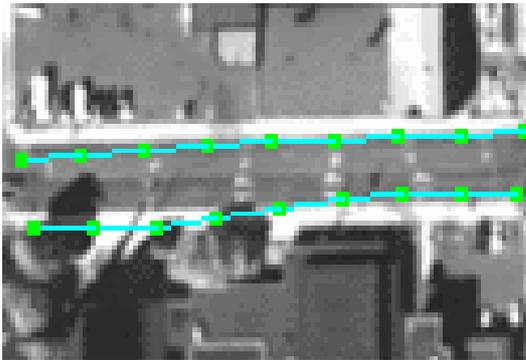
Dans cette section, l'application des *DoubleSnakes* à une scène réelle est présentée. Il s'agit d'une image provenant du satellite Quickbird d'une résolution spatiale dans le mode panchromatique de 0,7 mètre (figure 2.9).



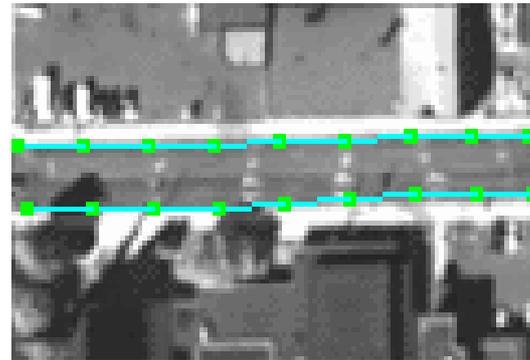
(a) Image originale avec l'initialisation



(b) Résultat *sans* l'approche multi-échelle et *avec* l'énergie de parallélisme



(c) Résultat *sans* force de parallélisme et *avec* l'approche multi-échelle



(d) Résultat *avec* l'approche multi-échelle et *avec* l'énergie de parallélisme

FIG. 2.9: test sur une image Quickbird (0,7 m de résolution spatiale) © DigitalGlobe

Cette image représente une rue droite en milieu urbain et comporte du bruit (ombres et marquage au sol).

L'initialisation des *DoubleSnakes* a été superposée sur la figure 5.8(a) et sera présentée au chapitre 3.

Différents cas ont été envisagés :

- L'énergie image des *DoubleSnakes* est calculée avec le gradient de l'image à pleine résolution (figure 2.9(b)). Ici, la nature du bruit étant plus variée que pour l'exemple 2.7.b (bruit d'occlusion dû à l'ombre portée de l'arbre en bas à droite de l'image, marquage au

sol), le phénomène de divergence est encore plus accentué.

- L'énergie image est calculée à partir des coefficients d'ondelettes jusqu'à l'échelle 5, mais l'énergie de parallélisme des *DoubleSnakes* est mise à 0. Le résultat final (figure 2.9(c)) présente un aspect plus régulier que pour l'exemple précédent et est moins dépendant de variations locales du gradient. Toutefois, un des brins a été attiré par le gradient créé par l'arbre, sans que l'autre, correctement positionné au final, n'ait réussi à l'en empêcher.

- Les *DoubleSnakes* évoluent suivant la fonctionnelle d'énergie (2.22), avec l'énergie image calculée à partir des coefficients d'ondelettes et le terme d'énergie de parallélisme. La figure 2.9(d) présente le résultat de l'extraction après l'évolution des *DoubleSnakes* à toutes les échelles : l'utilisation de l'approche multi-échelle et de contraintes géométriques fortes comme le parallélisme ont permis aux contours actifs de ne pas être piégés par le bruit.

Dans ce chapitre, nous avons défini les *DoubleSnakes*, contours actifs basés sur l'analyse multi-échelle permettant d'extraire les sections de rues avec bords parallèles dans un environnement bruité.

Leur initialisation, dont nous n'avons pas parlé, ainsi que leur utilisation dans le cadre de la méthode globale d'extraction de rues, est l'objet du chapitre 3 suivant.

Chapitre 3

La méthode d'extraction globale

Dans ce chapitre nous proposons une méthode globale permettant d'extraire les réseaux de rues de manière surfacique en milieu urbain. La méthode est modulaire, incluant les DoubleSnakes présentés au chapitre 2.

3.1 Méthode développée

3.1.1 Rappel des objectifs

Comme présenté dans l'introduction, notre étude se situe dans le contexte urbain et l'imagerie satellite à très haute résolution spatiale. Nous travaillons à partir de l'unique information portée par l'image. L'objectif est d'extraire les réseaux de rues et d'évaluer quantitativement le résultat de l'extraction. Le contexte urbain implique d'une part que la méthode développée soit robuste vis-à-vis du bruit, et d'autre part qu'elle soit assez générique pour s'adapter aux divers types de réseaux de rues urbains (rectilignes ou avec une courbure, de différentes largeurs, . . .).

3.1.2 Choix effectués

Notre travail se situe dans un contexte opérationnel, avec une contrainte opératoire qui est la fiabilité du résultat : il faut minimiser le plus possible le nombre de fausses détections ou de sur-détections. Comme évoqué au chapitre 1, le choix de l'approche, semi-automatique ou automatique, va dépendre des objectifs visés. Ainsi, un taux de bonnes détections de 85% peut être considéré comme un bon résultat pour un algorithme tout automatique alors que cela sera jugé insuffisant dans un cadre opérationnel, où l'on requiert plus de 95% de bonnes détections. Le contexte bruité du milieu urbain accroît encore la difficulté pour les algorithmes automatiques d'atteindre ce taux de performance.

Dans sa proposition de projet, le groupe de travail européen EuroSDR notait qu'à l'heure actuelle, une automatisation totale n'est pas envisageable (*a fortiori* en milieu urbanisé), d'une part en terme de performances et d'autre part en terme d'évaluation de leur qualité. Il préconise un effort de recherche sur l'extraction d'objets via les méthodes semi-automatiques (Mayer et Baltsavias 2002). Il existe en effet de nombreux aspects lors de l'extraction (que l'on distingue de la phase de détection) qui mettent les méthodes semi-automatiques en échec, notamment les artefacts sur la chaussée (véhicules, marquage au

sol) ou les occlusions partielles de la route. De plus, leur présence s'accroît en milieu urbain.

Renforcés par des discussions avec des spécialistes de l'IGN et la DGA¹, nous avons alors effectué le choix d'une méthode semi-automatique au niveau de l'amorce pour l'étape du graphe (présentée section 3.2). Nous avons donc développé une méthode basée sur l'image seule avec une amorce. Cette dernière peut être fournie sous formes diverses : points, tronçon, élément du tronçon ou base de données.

3.1.3 Schéma de la méthode d'extraction globale

La figure 3.1 suivante présente les différentes étapes constituant la méthode d'extraction globale.

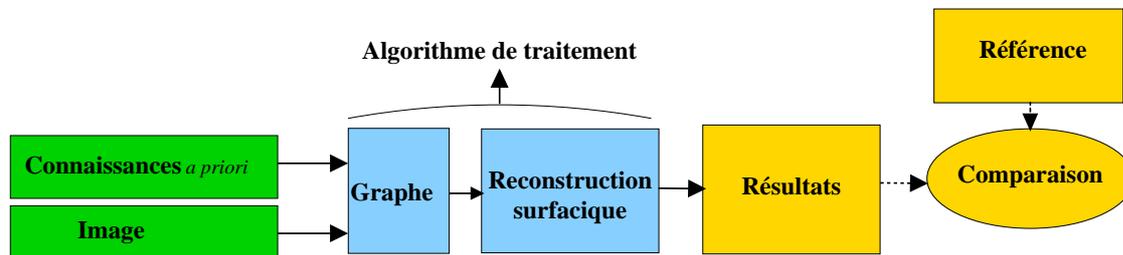


FIG. 3.1: méthode d'extraction globale

Cette méthode se compose : des entrées, de l'algorithme de traitement, puis de l'étape d'évaluation des résultats avec comparaison à une référence.

Ces différents éléments sont développés dans les sections suivantes.

3.1.4 Les entrées de la méthode

Les images. La première entrée est l'image : nous travaillons sur des images satellites en bande panchromatique, ou plus largement sur une image mono-spectrale à haute résolution spatiale.

La zone géographique d'intérêt est le domaine urbain, incluant l'agglomération péri-urbaine. La résolution spatiale de ces images peut varier entre 70 cm et 5 mètres.

Les connaissances a priori sur les rues et le réseau. Les principales causes d'échec des méthodes d'extraction dans un environnement bruité sont pour l'essentiel dues à l'écart par rapport au modèle théorique (voir section 1.5). Le modèle de rue utilisé doit donc être assez souple pour correspondre aux diverses représentations de la rue dans l'image.

Ainsi, les principales connaissances *a priori* introduites et qui composent le modèle de rue sont :

¹Délégation Générale pour l'Armement

- la rue est représentée dans l'image en tant qu'élément de surface (la distance entre les deux bords de la rue est supérieure à 3 pixels),
- les règles de construction de la voirie garantissent une continuité sur les variations de la courbure de la rue,
- l'homogénéité radiométrique le long de la direction principale de la rue ne constitue pas un critère fiable lorsque l'environnement est bruité. Toutefois, elle sera utilisée pour une première estimation de la localisation de la rue (approche «région»),
- le parallélisme local des bords de rues et la constance locale de la largeur de la rue seront eux utilisés pour la localisation précise de la rue (approche «contour»).

Concernant le réseau de rues, les principales caractéristiques retenues sont sa connexité, la connectivité locale, et l'absence d'extrémités libres (à l'exception des segments de rues situés aux bords de l'image).

3.1.5 L'algorithme de traitement

Principe. Il s'agit de la partie centrale de la méthode. L'algorithme est basé sur la coopération entre une représentation linéique et une représentation surfacique du réseau de rues. Ainsi, l'algorithme se décompose en deux parties séquentielles : une étape d'extraction du graphe du réseau de rues et une étape de reconstruction des rues dans l'image en tant qu'éléments de surface.

Extraction du graphe du réseau. Le réseau routier peut être représenté par une structure de graphe composé d'arcs et de nœuds. Il s'agit dans cette phase d'extraire un graphe topologiquement correct du réseau de rues. Cette étape vise à donner correctement les connexions spatiales entre les différents arcs représentant les rues, ainsi qu'une information sur leur localisation dans l'image.

L'extraction du graphe du réseau de rues intègre une fonction de suivi de lignes mais peut aussi provenir d'une base de données géographiques (comme la base de données *BDTopo*[®] de l'IGN) ou être réalisée par l'opérateur.

Dans tous les cas, on obtient une initialisation du graphe représentant le réseau de rues, dont les sommets sont les intersections des rues.

En sortie de cette étape, le réseau de rues sera représenté par un graphe topologiquement correct. Toutefois à cette étape, le réseau de rues est représenté de manière linéique, et les polygones formant le réseau ne sont pas nécessairement recalés sur l'axe des rues.

Étape de reconstruction surfacique. L'étape suivante est la reconstruction surfacique des rues dans l'image. Cette étape peut être considérée comme l'étape de segmentation de l'algorithme.

Cette deuxième étape fait intervenir les contours actifs multi-échelles, les *DoubleSnakes*, définis au chapitre 2. L'initialisation des *DoubleSnakes* se fait à partir du résultat de l'extraction du graphe du réseau : l'estimation de la localisation spatiale de la rue est fournie par le graphe extrait précédemment. En retour, l'étape de reconstruction permettra de recalculer précisément le graphe du réseau sur l'axe des rues.

Enfin, les intersections bénéficieront d'un traitement particulier. En effet, elles ne correspondent plus au modèle de l'objet rue, défini par les *DoubleSnakes*, notamment pour les

bords parallèles.

Propriétés de cette approche. L'algorithme de traitement a la particularité d'utiliser deux modes de représentation des rues : la représentation linéique et la représentation surfacique. La méthode tire parti d'une coopération entre ces deux modes de représentation de la rue : l'étape de représentation linéique permet d'initialiser l'étape de représentation surfacique, et l'étape de représentation surfacique permet de recalibrer précisément le graphe sur l'axe de la rue.

Un autre avantage de cette séparation du traitement est de permettre à l'utilisateur d'effectuer une vérification intermédiaire, et si nécessaire une correction, entre la phase d'extraction du graphe du réseau et la phase de reconstruction. On vise ainsi un contrôle continu de la qualité d'extraction.

Si nous proposons dans cette thèse des algorithmes pour ces deux modules, on peut aussi envisager l'utilisation d'autres méthodes pour chacune des étapes. Par exemple, la première étape utilise une représentation linéique du réseau par son graphe. On peut donc inclure à cette étape toute méthode permettant une extraction linéique du réseau, quitte à diminuer artificiellement la résolution de l'image (voir la section 1.3.1 dans le chapitre 1 sur l'état de l'art). La méthode est donc ouverte aux développements futurs.

3.1.6 Évaluation des résultats

Un accent particulier a été mis sur l'évaluation des résultats par rapport à l'interprétation humaine (chapitre 4). Cette évaluation est essentielle car elle permet de caractériser les résultats produits par un algorithme automatique mais aussi d'en donner une mesure de fiabilité. Une référence est nécessaire pour comparer les résultats obtenus et ainsi obtenir des critères quantitatifs d'évaluation. Une méthode pour extraire une référence à partir de l'image numérique, et des critères quantitatifs de qualité ont été établis (section 4.3).

Les deux modules de l'algorithme de traitement sont détaillés individuellement dans les sections suivantes. La partie d'évaluation des résultats de la méthode (figure 3.1) fera l'objet du chapitre 4.

3.2 Module du graphe

3.2.1 Formalisme utilisé

Le réseau routier peut être représenté par une structure de graphe composé d'arcs et de nœuds. Les arcs sont des polygones définies par la liste de leurs sommets. Les extrémités de ces arcs sont constituées par des nœuds dont l'attribut «degré» indique le nombre d'arcs qui s'y connectent (figure 3.2).

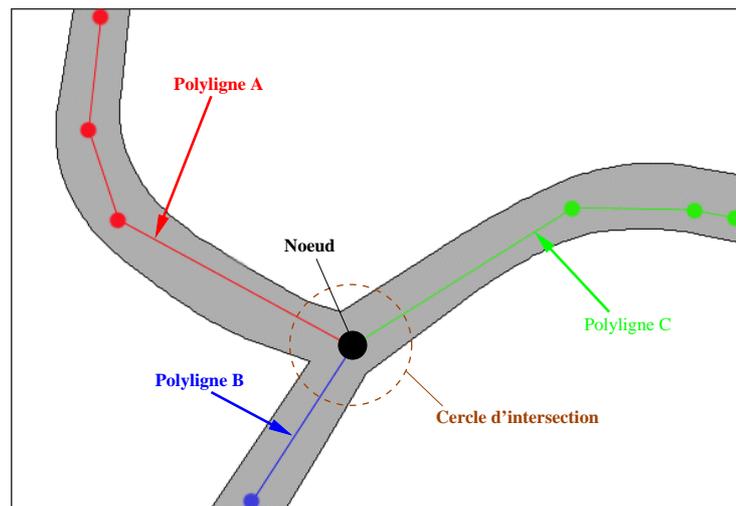


FIG. 3.2: polygones constituant le graphe du réseau

Dans un souci de cohérence et de compatibilité, nous avons choisi le même formalisme que pour les bases de données (notamment la *BDT^{opo}*[®] de l'IGN) : un tronçon de rue est défini d'une intersection à une autre et est représenté par une polyligne.

Dans l'image, cette distinction implique un traitement différent entre les tronçons à bords parallèles caractérisant la rue, et les intersections. En effet, ces deux «objets» géographiques n'ont pas des topologies similaires et présentent aussi des différences de forme. Un cercle, appelé cercle d'intersection, va délimiter les deux types de traitement à effectuer : traitement des contours parallèles et des intersections. Ce dernier traitement sera détaillé dans la section décrivant le module de reconstruction (section 3.3).

3.2.2 Description

Il s'agit dans cette phase d'extraire un graphe topologiquement correct du réseau de rues. Comme nous l'avons mentionné dans la section 3.1.2 sur les choix effectués, la contrainte principale est la fiabilité du résultat : les fausses détections ou les sur-détections sont à proscrire.

Pour réaliser cet objectif et augmenter la robustesse de la méthode, une approche semi-automatique au niveau de l'amorce du module du graphe a été retenue. Dans le cas

de l'algorithme de suivi présenté dans la section (3.2.3) suivante, l'approche est semi-automatique, l'opérateur fournissant un point d'amorce et une direction de propagation. Les polygones peuvent aussi provenir de bases de données géographiques ou être saisies manuellement par l'opérateur.

Le module du graphe comprend deux phases :

- extraction des polygones du graphe (section 3.2.3),
- obtention du graphe complet par raccordement des polygones extraites (section 3.2.4).

3.2.3 Extraction des polygones du graphe

L'algorithme de suivi de lignes (Airault et Jamet 1995). Cette partie est basée sur l'algorithme de suivi de lignes de Airault et Jamet (1995), dont le code nous a été aimablement fourni par le laboratoire MATIS de l'IGN. Cet algorithme est semi-automatique au niveau de l'amorce du suivi, l'opérateur fournissant un point et une direction de propagation.

Son intégration nous a paru profitable pour plusieurs raisons. Tout d'abord, cet algorithme de suivi a montré son efficacité notamment pour le passage d'occlusions locales, le changement de revêtement de la rue ou le franchissement des intersections. Ses limites et ses causes d'arrêts sont connues (Ruskoné 1996). Enfin, la structure modulaire de la méthode (figure 3.1), permettait son intégration de manière souple vis-à-vis du module de reconstruction surfacique.

Cet algorithme de suivi s'applique principalement pour des environnements ruraux ou péri-urbains. Il rencontre en effet des difficultés dans les environnements fortement urbanisés induisant beaucoup de bruit (comme un fort marquage au sol ou de nombreux véhicules). Cependant, même des imprécisions lors du suivi sont acceptables dans notre cas, car on ne cherche pas à cette étape une localisation précise, ni même à ce que la polygône soit entièrement incluse entre les deux bords de rue.

Le principe de l'algorithme de suivi est de générer un arbre de recherche de chemins possibles, puis de sélectionner le meilleur chemin, correspondant à la rue à détecter, qui minimise une fonction de coût. Cette fonction de coût évalue l'homogénéité de la texture locale dans les directions de propagation possibles. Cette homogénéité de texture se traduit par une variance minimale dans la direction correspondant à l'orientation de la rue. La robustesse de l'algorithme est assurée par une optimisation du critère d'homogénéité directionnelle sur une distance assez grande (la profondeur de l'arbre de recherche).

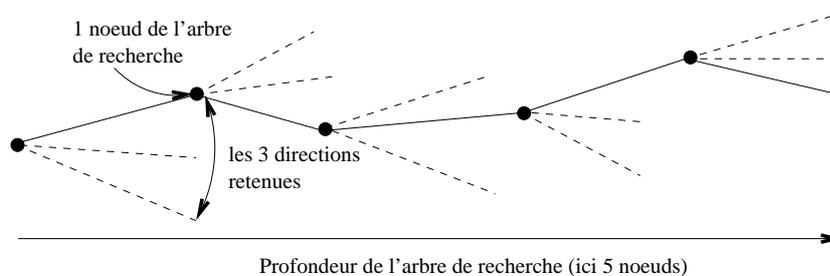


FIG. 3.3: arbre de recherche pour l'algorithme de suivi

Chaque nœud de l'arbre de recherche est composé des 3 meilleures directions retenues (les directions angulaires étant quantifiées de 0 à 31 dans le sens trigonométrique), et nous avons fixé la profondeur de l'arbre à 5 (figure 3.3). Enfin, les calculs de variance sont effectués sur des voisinages allongés de longueur variable (entre 10 et 30 pixels).

Le choix du meilleur chemin consiste donc à déterminer le chemin minimisant la fonction de coût sur un ensemble de $3^5 = 243$ combinaisons possibles.

L'expression de la fonction de coût pour un chemin est :

$$COUT_{chemin} = \frac{\sum_{i=1,5} (variance_i * (\delta_i * Coef_{rigid}))}{\sum_{i=1,5} (longueur_i)} \quad (3.1)$$

avec

$variance_i$: la variance sur la radiométrie calculée sur tout le segment situé à la profondeur i de l'arbre de recherche.

δ_i : l'écart entre la direction du segment i et son segment précédent.

$Coef_{rigid}$: coefficient de rigidité pondérant les changements de direction.

$longueur_i$: la longueur d'un segment situé à la profondeur i de l'arbre de recherche.

Dans l'équation (3.1), on privilégie dans l'arbre de recherche les segments les plus longs. La variance dans ce cas a en effet plus de chance de bien caractériser la texture de la rue.

L'algorithme originel s'intègre dans la chaîne de production de l'IGN et comprend une interface utilisateur élaborée ainsi qu'une gestion des intersections. Dans notre version, l'algorithme de suivi de ligne (noté *Suivi_ligne*) renvoie, pour un point et une direction fournis en entrée, une série de 5 points correspondant à la polyligne extraite.

Son intégration a été réalisée selon le schéma 3.4.

Ce chaînage semi-automatique permet à tout instant du suivi automatique de corriger les possibles dérives du suivi de ligne. Il s'effectue successivement pour chacune des polygones. L'algorithme de suivi s'arrête lorsqu'il rencontre un bord de l'image ou sur décision de l'opérateur. On passe alors à la polyligne suivante.

Exemple d'extraction de polygones. La figure 3.5 présente le résultat du suivi sur une image urbaine du satellite Quickbird (70 cm de résolution spatiale). Depuis l'amorce fournie par un opérateur (un point et une direction), l'algorithme a suivi la rue, sans intervention de l'opérateur et ne s'est arrêté que lorsqu'il a rencontré le bord de l'image (cas d'arrêt normal). Au niveau de l'intersection, c'est la rue avec la texture la plus homogène (compte de tenu de la position et de la direction au moment du calcul de la variance) qui a été choisie par l'algorithme. L'optimisation du critère d'homogénéité directionnelle sur une distance assez grande a permis de passer l'occlusion importante causée par les arbres (entourée sur la figure 3.5). Enfin, on peut observer que la taille des segments varie en fonction de l'homogénéité locale de la variance.

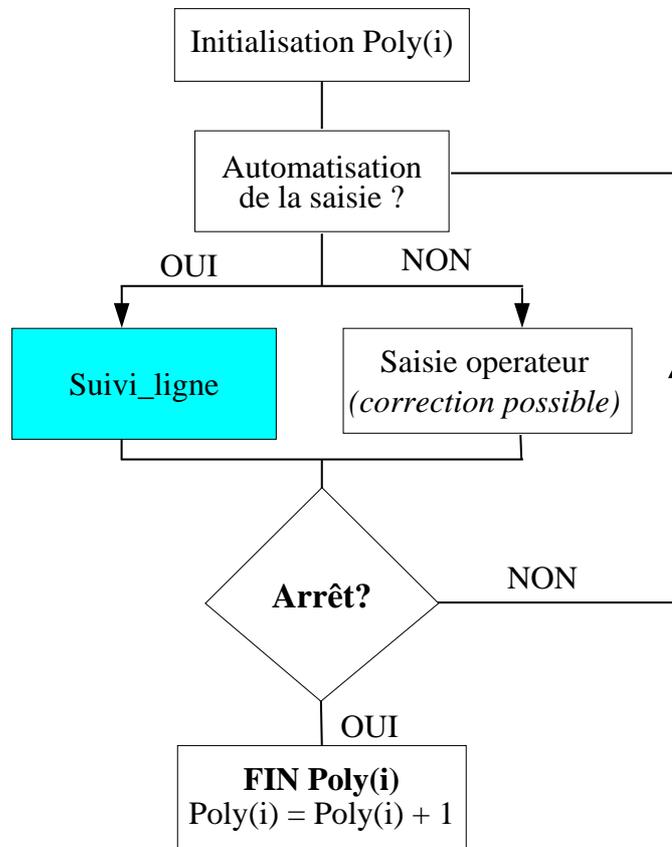


FIG. 3.4: diagramme d'extraction des polygones

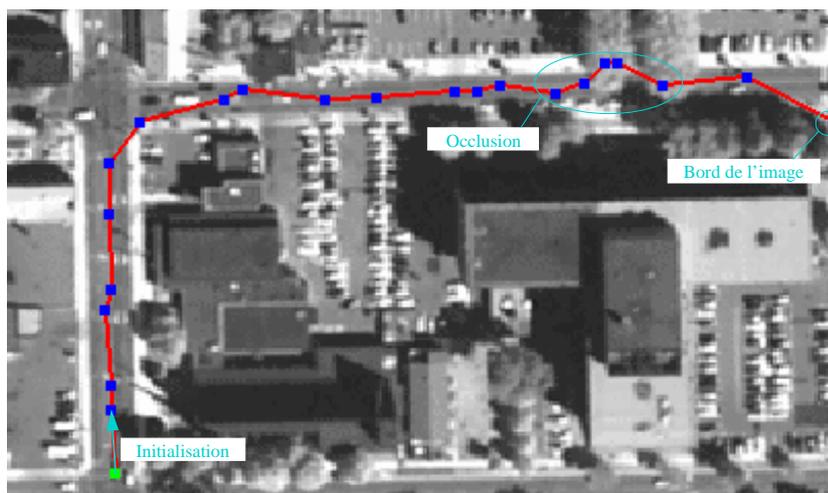


FIG. 3.5: résultat d'un suivi de rue

3.2.4 Obtention du graphe complet à partir des différentes polygones

Objectif. Dans cet section, nous allons chercher à reconstruire la topologie du graphe du réseau à partir des différentes polygones (non connectées) obtenues par l'algorithme de suivi.

Une rue est définie d'une intersection à une autre et est représentée dans le graphe du réseau par une polygone. Les rues se connectent entre elles aux intersections, et définissent les nœuds du graphe. De par leur différence de forme et de topologie, les rues et les intersections doivent être traitées séparément, tant pour leur détection que pour leur extraction. Le cercle d'intersection, représenté sur la figure 3.2, va délimiter les deux types de traitement à effectuer : traitement des tronçons à contours parallèles et traitement des intersections. Ce cercle est déterminé par l'opérateur. Il nous a en effet semblé que le gain de temps d'une extraction automatique des intersections ne serait pas compensé par l'étape obligatoire de vérification des hypothèses d'intersection par l'opérateur. Si la tâche est simple et rapide pour un opérateur (placer le centre du cercle et son rayon), elle est très difficile pour un procédé automatique à cause de la grande variabilité de forme et de topologie des intersections et des carrefours. Toutefois, on peut mentionner les travaux de Boichis (2000) dont la thèse porte exclusivement sur l'extraction de carrefours routiers à partir d'une base de données. Le cercle a été choisi car c'est une forme isotrope qui ne privilégie aucune direction. Son diamètre est tel qu'il englobe la totalité de l'intersection. Son centre sera le nœud du graphe définissant une intersection.



FIG. 3.6: polygones extraites par l'algorithme de suivi de rues. Les points verts correspondent aux amorces fournies par l'opérateur, les points bleus ont été extraits par l'algorithme de suivi.

La figure 3.6 présente le résultat de l'extraction des polygones par l'algorithme de suivi (scène étendue de la figure 3.5). À ce stade, nous avons un ensemble de polygones, non connectées entre elles. De plus, certaines d'entre elles «traversent» plusieurs intersections

et sont donc composées de plusieurs tronçons de rue au sens des bases de données (cas de la polyligne extraite sur la figure 3.5).

Exemple d'extraction du graphe complet. Il convient donc d'extraire la topologie du graphe à partir de cet ensemble de polygones. Pour cela, les polygones dont les points sont localisés dans un cercle d'intersection vont être tronqués. On va ensuite raccorder le premier point de la polygone situé hors du cercle d'intersection avec le centre de ce cercle (figure 3.7(a)). Pour le cas où une polygone traverse une (ou plusieurs) intersections (cas de la polygone extraite figure 3.5), la polygone est «découpée» en plusieurs tronçons qui vont être raccordés comme précédemment avec le centre du cercle d'intersection le plus proche.

Une fois le graphe complet obtenu, un sous-échantillonnage des polygones de ce graphe va être effectué en ne retenant qu'un pixel sur trois entre les extrémités (figure 3.7(b)). L'intérêt de ce sous-échantillonnage est de simplifier le graphe avant son utilisation comme entrée dans le module de reconstruction surfacique. En effet, dans les zones de forte variance radiométrique, l'algorithme de suivi n'a progressé qu'avec des segments de petite longueur, créant ainsi un grand nombre de points proches. Si ce nombre élevé de points a été nécessaire au moment du suivi, il n'est plus justifié après.

C'est ce graphe simplifié qui va définir l'entrée de l'algorithme de reconstruction surfacique, présenté dans la section suivante.



(a) obtention du graphe complet à partir des polygones extraits par l'algorithme de suivi



(b) obtention du graphe complet sous-échantillonné à partir des polygones extraits par l'algorithme de suivi

FIG. 3.7: obtention du graphe complet du réseau

3.3 Module de reconstruction

3.3.1 La méthode d'extraction surfacique du réseau

Description. En sortie du module du graphe, le réseau est topologiquement correct mais représenté de manière linéique. De plus, les polygones du graphe ne sont pas nécessairement bien localisées par rapport à l'axe de la rue.

Le rôle de ce module de reconstruction est d'extraire le réseau de rues de manière surfacique en utilisant le graphe topologique comme initialisation. Autrement dit, il s'agit d'extraire le contour des rues et des intersections dans l'image, tout en gardant les informations sur la topologie du réseau provenant du module précédent.

Lien avec les contours actifs multi-échelles. Compte tenu de la variété de typologie des rues en ville, il nous faut disposer d'un modèle pouvant s'adapter à cette diversité de forme. Il doit aussi pouvoir s'adapter à l'environnement générant beaucoup d'artefacts qu'est le milieu urbain.

Comme nous l'avons vu au chapitre 2, les contours actifs constituent un outil bien approprié au problème puisqu'ils permettent de s'adapter à la forme à extraire ainsi que de passer les occlusions locales. De plus, les contours actifs permettent d'intégrer des connaissances *a priori* sur les rues et le réseau en définissant des termes d'énergie spécifiques. Nous avons montré, dans la section 2.4, que leur association avec l'analyse multi-échelle permet d'accroître leur robustesse vis-à-vis du bruit. Dans cette optique, ont été développés les contours actifs multi-échelles puis les *DoubleSnakes* permettant l'extraction des tronçons de rue à bords parallèles (section 2.5).

Le module de reconstruction comprend deux étapes séquentielles différenciant le traitement des rues et celui des intersections. Le cercle d'intersection (section 3.2.4) situé au nœud du graphe du réseau va différencier le traitement à effectuer. Son rayon est assez grand pour englober l'intégralité de l'intersection (voir figure 3.2).

A l'extérieur du cercle d'intersection, les rues sont considérées comme ayant des bords localement parallèles et sont extraites par les *DoubleSnakes* présentés à la section 2.5. Les *DoubleSnakes* ont besoin d'être initialisés à proximité de la forme à extraire : cette initialisation est fournie par le module précédent du graphe. Elle sera décrite dans la section suivante.

A l'intérieur de ce cercle, les intersections sont ensuite extraites par des contours actifs multi-échelles simples, initialisés par la position finale des *DoubleSnakes* après évolution (section 3.3.3).

La section qui suit présente comment les *DoubleSnakes* vont être initialisés pour l'extraction des tronçons de rue à bords parallèles (traitement hors du cercle d'intersection).

3.3.2 Initialisation des *DoubleSnakes* pour l'extraction des tronçons de rue à bords parallèles

Du graphe aux *DoubleSnakes*. En plus de l'information sur la topologie du réseau, le graphe va permettre l'initialisation du module de reconstruction. Cette initialisation est celle des *DoubleSnakes*, qui vont extraire dans un premier temps les tronçons de rue à bords parallèles.

Les contours actifs doivent d'abord être initialisés, et si possible le plus près de la forme à extraire. Dans le cas présent, l'initialisation des *DoubleSnakes* est faite à partir des polygones du graphe (figure 3.8.a) qui vont être échantillonnées (figure 3.8.b). Nous avons fixé le pas d'échantillonnage constant pour toutes les polygones. On peut toutefois envisager de le faire varier en fonction de la courbure locale de la polygone ou en fonction de la résolution de l'image.

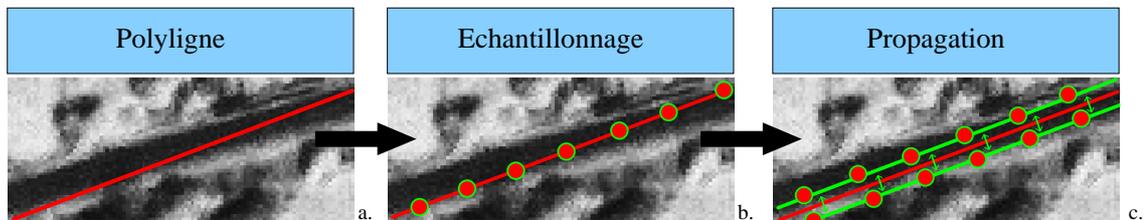


FIG. 3.8: procédé d'initialisation des *DoubleSnakes* par les polygones

Comme la rue à extraire est un élément surfacique déterminé par ses deux bords, cette polygone va être dédoublée et «propagée» dans les deux directions orthogonales à la polygone (figure 3.8.c). Les deux polygones échantillonnés qui en résultent vont constituer les deux brins initiaux d'un *DoubleSnake*.

La distance de propagation va dépendre de la résolution de l'image ainsi que du type de rue à extraire (autoroute, ruelle, ...). On a défini 4 classes de rues correspondant aux catégories de ruelle, rue, boulevard et autoroute urbaine. Couloigner (1998) proposait une définition de classes de rues selon leur profil radiométrique à différentes résolutions spatiales. Considérant le bruit important apparaissant dans notre contexte (image à très haute résolution spatiale et milieu urbain), il était préférable de différencier les rues suivant leur classe de largeur.

La distance de propagation $Distance_{prop}$ pour une $Classe_i$ ($i \in [1, 4]$) est définie comme :

$$Distance_{prop}(Classe_i) = \frac{LARG_{Classe_i}}{2 \times RES_{image}}$$

avec RES_{image} la résolution spatiale (en mètre) de l'image et $LARG_{Classe_i}$ représentant la largeur de rue (en mètre) d'une classe. La distance de propagation est en pixel. Le facteur $\frac{1}{2}$ provient du fait que la distance de propagation correspond à une demi-largeur (voir figure 3.8.c) .

Le tableau 3.1 présente les classes de largeur correspondant aux catégories de ruelle, rue, boulevard et autoroute urbaine. Elles ont été déterminés en prenant en compte les spécifications de l'aménagement de la voirie urbaine (guide du CETUR : Centre d'Études des Transports URbains 1988).

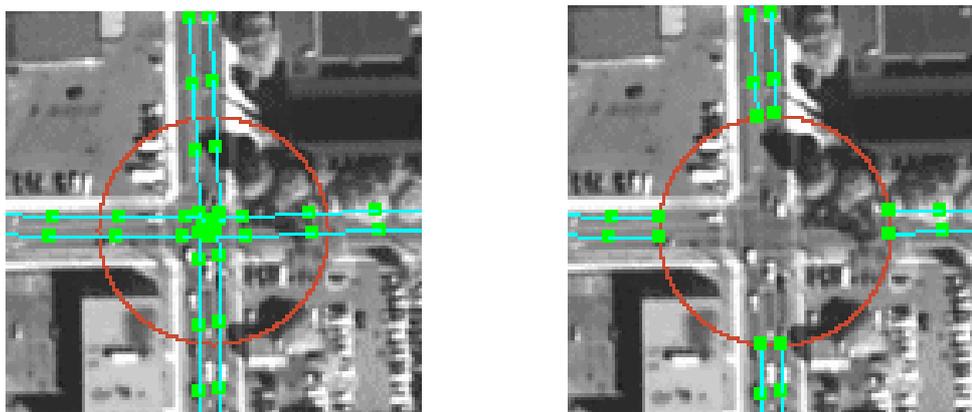
	<i>Classe 1</i> Ruelle	<i>Classe 2</i> Rue	<i>Classe 3</i> Boulevard	<i>Classe 4</i> Autoroute urbaine
Largeur de classe (en mètre)	3	6	12	20

TAB. 3.1: tableau des classes de largeur

A titre d'exemple, la distance de propagation correspondant à un boulevard (rue de classe 3) sera de 8,57 pixels sur une image du satellite Quickbird (résolution spatiale de 0,7 m), tandis qu'elle sera de 6 pixels pour une image du satellite Ikonos (la résolution spatiale du capteur est de 1 mètre).

Initialisation des *DoubleSnakes* au niveau des intersections. La frontière entre l'extraction des sections de rues à bords parallèles par les *DoubleSnakes* et l'extraction des intersections se situe au niveau du cercle d'intersection. Après la phase de propagation (figure 3.8), les *DoubleSnakes*, dérivant des polygones définies d'un nœud du graphe à l'autre, pénètrent dans le cercle d'intersection (figure 3.9(a)).

Comme l'extraction des contours parallèles ne doit pas déborder sur la zone d'extraction des intersections, un pré-traitement est effectué à l'initialisation des *DoubleSnakes*, consistant à éliminer les points des brins se trouvant à l'intérieur du cercle. Les extrémités des *DoubleSnakes* sont placées alors sur le cercle d'intersection (figure 3.9(b)). Lors de l'évolution des *DoubleSnakes*, ces points auront leurs déplacements contraints à rester sur le cercle.



(a) Avant la troncature

(b) Après la troncature

FIG. 3.9: initialisation des *DoubleSnakes* aux intersections

L'énergie de parallélisme des *DoubleSnakes* telle qu'elle a été définie section 2.5 impose que les deux brins d'un même *DoubleSnake* soient composés du même nombre de points. Cette hypothèse, qui est vérifiée à l'initialisation (figure 3.8) peut ne plus être vraie après l'élimination des points à l'intérieur du cercle. Une vérification est donc effectuée et le brin qui comporte plus de points que son symétrique est tronqué.

3.3.3 Traitement des intersections

Description. Comme évoqué précédemment, l'extraction des intersections réclame un traitement spécifique par rapport à l'extraction des contours parallèles. En effet, les propriétés des rues utilisées lors de l'extraction des contours parallèles (bords parallèles, homogénéité radiométrique dans la direction de la rue, contraste par rapport à l'environnement de la rue) ne sont plus toutes valides. De plus, les formes possibles des intersections sont très variées (rond-point, fourche, ...). Nous ne traiterons pas l'extraction des ronds-points à sens giratoire, cas complexe car non convexe. On pourra se référer à la thèse de Boichis (2000), sur l'extraction des carrefours et des ronds-points à partir d'une base de données géographiques.

Initialisation des intersections. Le traitement des intersections intervient une fois que la phase d'extraction des contours parallèles par les *DoubleSnakes* est terminée. A cette étape, les extrémités des *DoubleSnakes* ont minimisé leur énergie en étant contraint à se mouvoir sur le cercle d'intersection. Ces points vont servir à initialiser l'extraction de l'intersection. La figure 3.10 décrit cette initialisation.

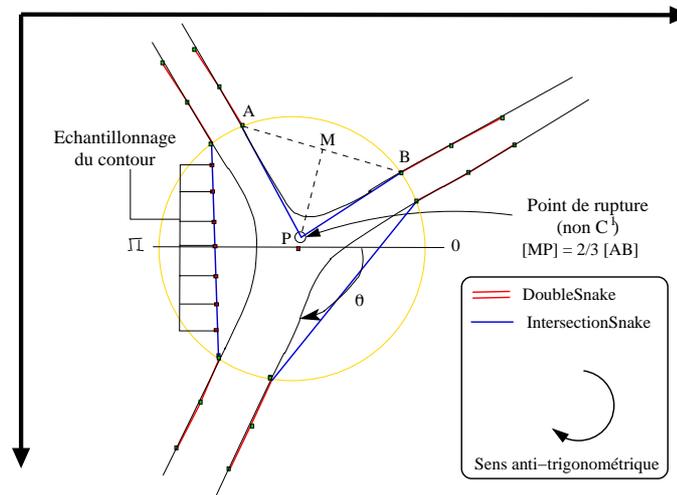


FIG. 3.10: initialisation des contours actifs d'intersection

Les extrémités des *DoubleSnakes* se trouvant sur le cercle sont ordonnées suivant l'angle θ qu'elles forment avec l'axe des abscisses ($\theta \in [-2\pi, 0]$). Chaque point est alors apparié avec le point le plus proche (en terme d'angle) et n'appartenant pas au même *DoubleSnake*. Ils

définissent ainsi les extrémités de nouveaux contours actifs, qui sont ensuite échantillonnés (figure 3.10).

Si l'angle formé par les brins de deux points appariés est en dessous d'un certain seuil ($3\Pi/5$), un point de rupture (pas de continuité \mathcal{C}^1) est introduit de manière à améliorer l'initialisation du contour actif d'intersection. La figure 3.10 illustre la manière dont est déterminé le point de rupture P : il se situe au $2/3$ du segment reliant le milieu M du segment [AB] et le centre du cercle d'intersection ($[MP]=\frac{2}{3}[AB]$). Ceci implique qu'il se situera à l'intérieur du cercle. Le facteur $2/3$ ainsi que l'angle de $3\Pi/5$ pour l'introduction du point de rupture ont été fixés de manière empirique sur un ensemble de tests. Le point de rupture évoluera par la suite avec un paramètre de continuité \mathcal{C}^1 mis à zéro (soit $\beta = 0$ dans l'équation 2.1). Ce point correspond en effet aux coins des rues, pour lesquelles la continuité \mathcal{C}^1 n'est pas respectée. Ces deux nouveaux segments vont ensuite être échantillonnés.

Les *IntersectionSnakes*. La structure de contour actif utilisée au cours de cette étape (*IntersectionSnake*) consiste en un contour multi-échelle unique. Il est «ouvert» au sens où le premier point n'est pas rattaché au dernier, mais ses deux extrémités sont fixes, déterminées par la position finale des extrémités des *DoubleSnakes* (figure 3.10). Les *IntersectionSnakes* ont des contraintes spécifiques lors de leur évolution : ils doivent maintenir une continuité d'ordre \mathcal{C}^1 à chacune de leurs extrémités avec le brin du contour parallèle qu'ils prolongent. De manière à garantir une continuité entre les intersections et les contours parallèles, une pondération importante dans la fonctionnelle d'énergie des *IntersectionSnakes* est donnée au terme de continuité \mathcal{C}^1 (sauf aux points de rupture). Enfin, un *IntersectionSnake* est contraint à ne jamais sortir du cercle d'intersection à l'intérieur duquel il a été défini. En effet, le cercle d'intersection englobant la totalité de l'intersection, il ne peut exister de structures en dehors.

Les *IntersectionSnakes*, tout comme les *DoubleSnakes* sont des contours actifs dont la fonctionnelle est définie pour une résolution spatiale donnée (section 2.4.2). Leur évolution sur plusieurs niveaux de résolution va permettre d'augmenter leur convergence et d'accroître leur robustesse vis-à-vis du bruit.

3.3.4 Déroulement de l'algorithme de reconstruction

Description. La figure (3.11) représente le déroulement de l'algorithme d'extraction surfacique.

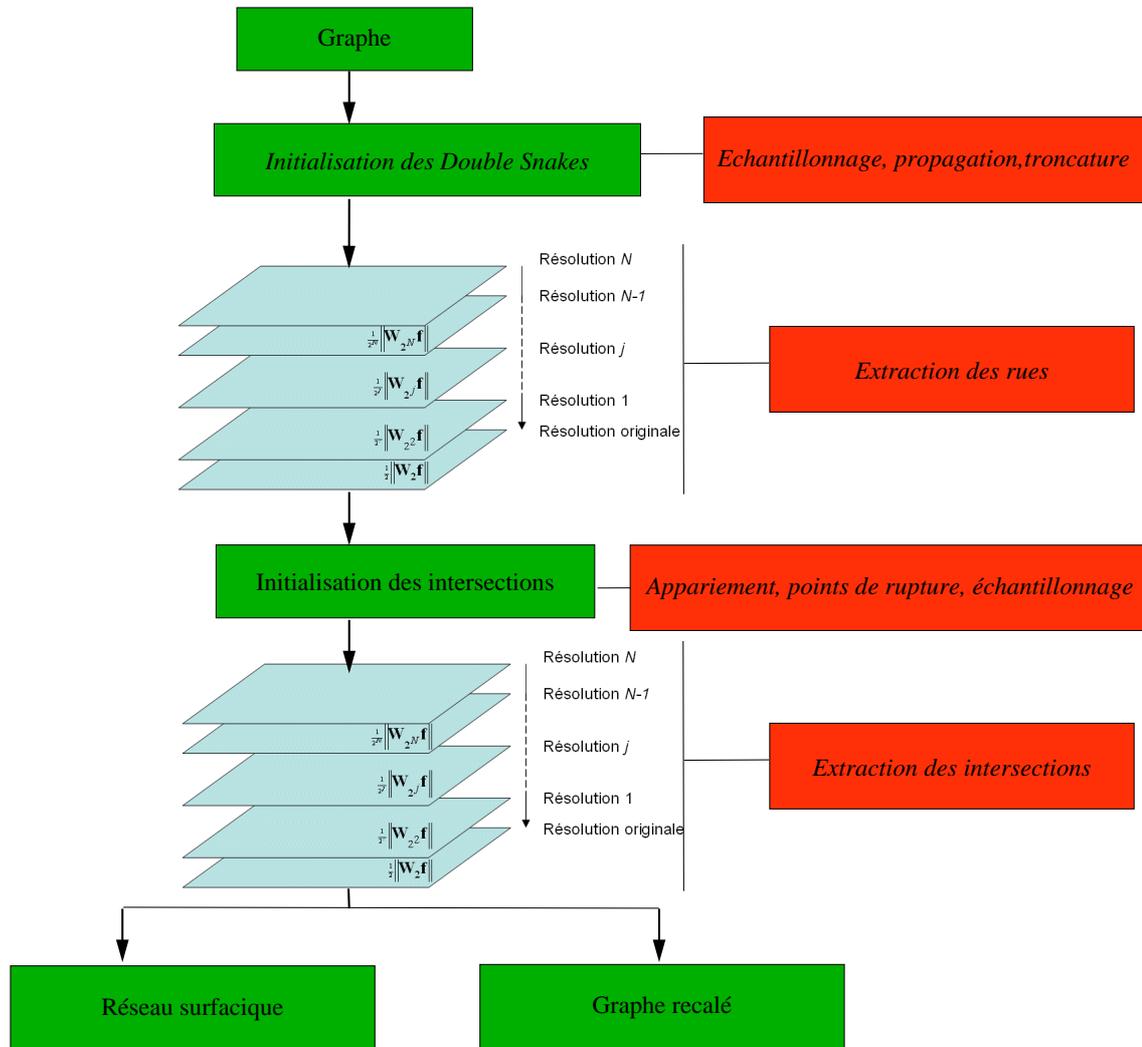


FIG. 3.11: algorithme de reconstruction surfacique

A partir du graphe du réseau, l'initialisation des *DoubleSnakes* est effectuée suivant les 3 étapes définies section 3.3.2 : échantillonnage, propagation dans la direction normale à la polyligne et à une distance dépendant de la classe du *DoubleSnake*, puis troncature des *DoubleSnakes* au niveau du cercle d'intersection.

Ensuite, une analyse multi-échelle est effectuée à partir de l'image à la résolution originale, donnant une série de plans de contexte et une série de plans d'ondelettes.

Pour chaque plan de contexte, une énergie image est calculée à partir du plan d'ondelette associé. Cette énergie va attirer les contours actifs vers les détails présents à la résolution du plan de contexte considéré.

Le niveau de profondeur de décomposition de l'analyse multi-échelle est un paramètre important. Il s'agit de trouver la résolution minimale à laquelle disparaissent les artefacts de taille inférieure à la largeur de la rue (voitures, marquage au sol, ...) tout en conservant ses contours. Ce seuil va dépendre de la largeur de la rue à extraire et donc de sa classe d'appartenance et de la résolution de l'image (voir tableau 3.1).

En terme d'échelle, il s'agit pour une classe i de déterminer l'échelle j minimale telle que l'on ait : $2^j \geq 2 \times Distance_{prop}(Classe_i)$

Le tableau 3.2 présente, pour différents satellites, à partir de quelle échelle de décomposition chaque classe de rue va démarrer la minimisation de sa fonctionnelle d'énergie.

	<i>Classe 1</i>	<i>Classe 2</i>	<i>Classe 3</i>	<i>Classe 4</i>
Quickbird	3	4	5	5
Ikonos	2	3	4	5
SPOT 5	1	2	3	3

TAB. 3.2: niveau de décomposition en fonction de la classe de rue considéré pour différents satellites

Une fois ces étapes d'initialisation effectuées, l'extraction des contours parallèles peut commencer. Les *DoubleSnakes* évoluent sur tous les plans de contexte successifs, de la résolution la plus basse fixée par leur classe jusqu'à l'image de résolution initiale. La pondération choisie pour l'énergie image décroît avec la résolution. On privilégie d'abord l'énergie image, puis ensuite les termes de régularisation, afin d'éviter aux *DoubleSnakes* d'être piégés par des maxima locaux d'énergie image.

Les contours actifs d'intersection sont ensuite initialisés (section 3.3.3) : appariement des points sur le cercle d'intersection, ajout de points de rupture éventuels puis échantillonnage.

Les *IntersectionSnakes* vont évoluer eux aussi de l'image à la résolution la plus basse (que l'on a fixée empiriquement à celle des rues de classe 2) jusqu'à la résolution initiale de l'image. Deux autres contraintes apparaissent lors de la minimisation de leur fonctionnelle d'énergie : les extrémités sont fixes et le contour actif doit maintenir une continuité d'ordre \mathcal{C}^1 à chacune de ses extrémités avec le brin du contour parallèle qu'il prolonge.

A la fin des différentes itérations, tous les contours actifs se trouvent alors dans un état d'équilibre. Il est à noter que la chaîne d'implémentation de l'algorithme de reconstruction (transformée en ondelettes par l'algorithme à trous et algorithme glouton pour l'implémentation des contours actifs), permet un traitement rapide.

Recalage du graphe du réseau. Il est, après cette étape, possible et simple de recalculer le graphe topologique d'entrée : les polygones sont recalés sur l'axe de la rue en évaluant l'axe médian des *DoubleSnakes*. Les intersections seront, elles, toujours raccordées

au centre du cercle d'intersection.

Par rapport au graphe d'entrée, le graphe de sortie est donc recalé sur l'axe central mais est aussi suréchantillonné (taux d'échantillonnage des *DoubleSnakes*), et apparait donc plus «lisse».

Le chapitre 4 suivant présente le procédé qui a été défini pour l'évaluation quantitative des résultats. Des exemples d'application de la méthode globale sur différentes scènes et différents capteurs feront l'objet du chapitre 5.

Chapitre 4

Référence et évaluation

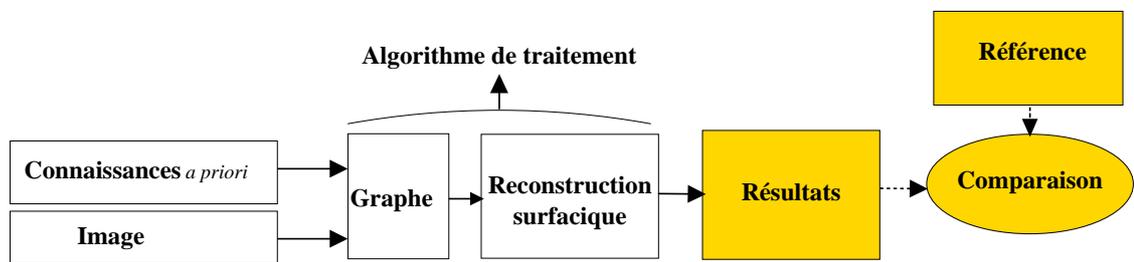


FIG. 4.1: référence et évaluation de la méthode globale

Ce chapitre est consacré à l'évaluation quantitative des performances de la méthode d'extraction. Dans ce but, il convient de définir au préalable une référence qui sera alors comparée au résultat de l'extraction automatique. Les différentes manières de définir un objet référence dans le cadre de méthodes d'extraction de rues à partir d'images satellites seront tout d'abord présentées. Une méthode pour extraire une référence relative, lorsque la seule donnée à disposition est l'image numérique, est ensuite proposée. Des critères quantitatifs d'évaluation appropriés sont évoqués. Cette méthode est enfin illustrée sur une image satellite à très haute résolution spatiale.

4.1 L'évaluation et le besoin d'une référence

4.1.1 Introduction

L'évaluation des résultats est essentielle pour tout système automatique de traitement, et *a fortiori* pour les méthodes d'extraction d'objets à partir d'images de télédétection. Elle permet de caractériser les résultats produits mais aussi d'en donner une mesure de fiabilité.

On peut distinguer deux types d'évaluations. Une évaluation par la méthode elle-même, que l'on peut qualifier d'évaluation «interne». Il s'agit dans ce cas de diagnostics pouvant être réalisés au niveau des différents modules composant la méthode, et indiquant si le résultat est jugé correct, incorrect ou à préciser.

Ce type d'évaluation doit être complété par une évaluation «externe», faite par rapport à

une référence indépendante avec laquelle s'effectuera la comparaison.

Nous n'aborderons ici que l'évaluation «externe», obligatoire pour une validation rigoureuse ne dépendant pas de la méthode de traitement.

Pour l'extraction de routes à partir d'images satellitales ou aéroportées, les travaux portant sur l'évaluation quantitative des résultats sont récents et assez peu nombreux (Heipke *et al.* 1997; Couloigner et Ranchin 1998; Harvey 1999).

Ils aboutissent à des critères quantitatifs comme la complétude ou la redondance du réseau routier. Tous utilisent une référence servant à la comparaison. Il s'agit dans la grande majorité des cas d'une référence tracée manuellement sur l'image et considérée comme la réalité de terrain. Cependant, il y a peu de discussions sur le choix et l'influence de cette référence sur les résultats d'évaluation. De plus, il n'existe actuellement pas de séries d'images de référence permettant de comparer les résultats de différents algorithmes d'extraction de rues. Les sections suivantes proposent une réflexion et une méthodologie pour l'établissement d'une référence relative de comparaison dans le cas où la seule donnée est l'image numérique. Les critères numériques attachés sont aussi définis.

4.1.2 Objectifs

Pour pouvoir extraire une référence, le cas favorable est de disposer d'une image géo-référencée et d'une base topographique de précision adaptée aux objectifs visés. Cependant, il est fréquent de ne pas disposer de relevé topographique, et qui plus est à jour, notamment dans les pays en voie de développement de superficie étendue. C'est le cas que nous considérerons ici. De la même manière que nous avons choisi pour la méthode d'extraction de ne pas utiliser de données externes de type MNE (section 3.1.2), nous considérerons pour l'évaluation quantitative que nous avons seulement l'image numérique comme source d'information. Nous nous situons donc dans le cadre d'une estimation «relative» de la qualité des résultats, où nous cherchons à obtenir des performances similaires à l'interprétation humaine.

Constatant la grande variabilité d'interprétation possible lors de la saisie d'une référence manuelle à partir de l'image, le problème de la fiabilité de cette référence se pose. En effet, pour les images à très haute résolution spatiale, la représentation en tant que surface des rues ainsi que l'augmentation des détails visibles ont accru cette variabilité. Sur les images de résolution décimétrique, les rues sont des éléments linéiques engendrant moins de différences d'interprétation.

Dans le cas général, si l'algorithme permet d'atteindre une précision subpixel, la référence choisie doit avoir cette précision. En effet, une référence image basée pixel peut estimer un algorithme non fiable alors que ses résultats ont une précision subpixel (par exemple le détecteur subpixel de lignes présenté par Steger (1998)).

D'autre part, nous cherchons une référence qui soit représentative de la réalité et non pas une représentation symbolique utilisant des procédés de généralisation cartographique.

Dans ce chapitre, nous allons tout d'abord présenter les différents types de référence possibles pour un procédé d'extraction de rues à partir d'images numériques. Puis nous

proposons une méthode pour extraire une référence à partir de l'image numérique. Enfin, la dernière section consiste en une application de cette méthode dans le cadre de l'extraction de rue sur des images satellites.

4.2 Les différents types de référence

4.2.1 Référence basée sur des données de terrain

Les données provenant de relevé de terrain (comme des mesures par GPS¹) sont une estimation directe de la réalité physique de la rue à détecter et permettent de la localiser dans l'espace géographique avec une bonne précision. Cependant, le lancement d'une campagne de mesures *in situ* est un procédé long et coûteux, difficilement réalisable à grande échelle. De plus, utiliser des données de terrain comme référence demande des mises à jour régulières, et n'est pas applicable aux images d'archives. D'autre part, les données issues d'une enquête terrain sont exprimées en unité physique (m ou m^2) et ne sont pas comparables directement au résultat de l'extraction (dans le domaine image). En effet, les effets de l'échantillonnage spatial de la scène et de la quantification par les capteurs CCD dépendent de plusieurs facteurs (conditions d'acquisition, humidité de la scène, conditions d'éclairage, trouble atmosphérique, ...). Il convient donc de trouver une méthode d'appariement automatique entre les deux types de données, étape complexe qui induit ses propres erreurs.

4.2.2 Référence basée sur l'interprétation d'images

Par sa capacité à pouvoir tirer parti des informations contextuelles, l'interprète humain constitue un « système » très efficace en reconnaissance de formes. La possibilité d'utiliser l'interprétation humaine pour l'évaluation d'algorithmes automatiques de traitement d'images a été notamment l'objet de travaux conjoints avec des psychologues. Afin de pouvoir effectuer la comparaison avec des méthodes automatiques, ces études ont porté sur la recherche de critères quantitatifs des interprétations humaines : orientation du détecteur de contour (Fram et Deutsch 1975) ou mesure de similarité pour l'indexation d'images (Squire et Pun 1997). Ces recherches ont montré la bonne capacité de l'interprète à identifier un objet avec un contraste faible par rapport à son environnement, tout en soulignant l'importance à prendre comme interprète des « experts » dans le domaine traité.

Dans le problème de la détection de rues, l'interprète humain fonctionne comme un algorithme « intelligent » capable de prendre en compte le contexte de l'image, à la différence des algorithmes locaux de bas-niveau. Il peut s'affranchir des obstacles tels que les occlusions (ombres sur la rue, ponts, ...) et des artefacts gênants pour les méthodes automatiques.

4.2.2.1 Évaluation visuelle *a posteriori*

La manière la plus simple et la plus répandue pour évaluer un résultat d'extraction de formes est l'estimation visuelle. L'opérateur compare la forme extraite automatiquement avec sa vision de la forme à détecter.

¹Global Positioning System

Cette référence image, que l'on peut considérer comme *a posteriori*, a l'inconvénient d'être variable, au sens où l'estimation de l'objet par l'interprète d'images est différente à chaque réalisation de l'algorithme. De plus, elle ne permet de développer que des critères qualitatifs et non quantitatifs.

4.2.2.2 Évaluation visuelle a priori

Demander à un « expert » de localiser à l'avance sur l'image les zones d'intérêt est une pratique courante en traitement d'images. En imagerie cérébrale par exemple, les points caractéristiques du cortex (comme les commissures antérieures et postérieures) sont estimés par un médecin spécialiste et pourront ainsi servir de points d'appui pour un recalage. Ainsi pour la détection de rues, certains auteurs (Couloigner 1998; Boichis 2000; Doucette *et al.* 2001) établissent une référence visuelle *a priori*, un interprète d'images définissant à l'avance la forme à détecter. Pour Ruskoné (1996) et Bordes (1997), la référence est la base de données *BDTopo*[®] de l'IGN. Elle est obtenue par restitution photogramétrique et interprétation d'images stéréoscopiques suivi de vérifications sur le terrain. La référence est donc fixe par rapport à chaque résultat de l'algorithme.

4.2.2.3 Choix effectué

Dans notre cas, la disponibilité, la fréquence d'obtention des données *in situ* ainsi que leur coût apparaissent comme des obstacles à leur utilisation comme référence pour un procédé d'extraction. En accord avec nos objectifs, nous avons privilégié une référence basée sur l'interprétation de l'image numérique. La référence est, dans ce cas, obtenue sur le même support, l'image, que l'algorithme automatique.

Il est alors préférable de définir la référence *a priori* et donc de la fixer par rapport à chaque résultat de l'algorithme. Cependant, de par la subjectivité de la saisie d'un contour, il existe une variabilité d'interprétation par l'ensemble des interprètes.

Dans le paragraphe suivant, nous présentons une méthode permettant de minimiser cette variabilité et de définir un objet référence ainsi qu'une zone de tolérance. Cette dernière permet la définition d'un intervalle de confiance pour l'objet référence ainsi qu'une zone caractérisant le résultat d'une extraction automatique.

4.3 Méthode pour l'extraction d'une référence

Nous établissons notre référence dans le cas d'un objet cartographique à extraire d'une scène de télédétection, puis nous l'appliquerons dans le cas de l'extraction du réseau routier. Nous ne considérons que le cas d'objets *connexes*.

Dans la suite, nous partirons donc du principe que le tracé planimétrique des deux bords de la rue est acquis par interprétation de l'image. Or l'estimation de la référence est liée à la subjectivité de l'interprète d'images. Pour réduire cette variabilité, notre choix est d'établir une référence à partir de mesures statistiques sur plusieurs interprétations d'images. On choisit un groupe d'interprètes ayant une habitude de l'analyse de scène à partir d'images numériques.

Un modèle vecteur pour la saisie est utilisé, ce qui permet une représentation des objets dans un espace géométrique continu. Il permet une localisation précise et la description

fine des caractéristiques géométriques des objets saisis.

Ce choix du modèle vecteur pour la saisie se justifie par le fait que l'interprète d'images ne perçoit pas l'image seulement à la résolution du pixel : le contexte et les connaissances *a priori* (l'objet à reconnaître par exemple) lui donnent une vision de la scène qui se traduit par une précision subpixel. Pour l'imagerie numérique, une estimation empirique indique que l'interprète d'images peut avoir une précision moyenne au niveau du pointé de 0,3 pixel (source DGA). De plus, comme nous l'avons mentionné dans les objectifs, nous voulons obtenir une référence de précision subpixel.

Les caractéristiques de cette référence sont donc :

- référence *a priori* (invariante pour chaque réalisation de l'algorithme),
- la référence est un objet vecteur dans l'espace analytique,
- la référence est déduite de mesures statistiques sur un groupe de plusieurs interprètes d'images.

Le schéma (4.2) ci-dessous illustre l'extraction d'une référence à partir d'un ensemble d'interprétations.

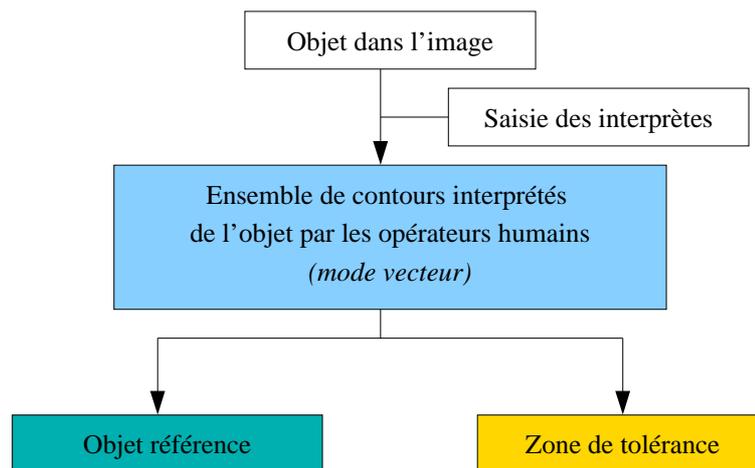


FIG. 4.2: extraction de la référence et de la zone de tolérance

L'analyse et le traitement des différentes saisies vont permettre d'extraire un objet référence et une zone de tolérance. Les paragraphes suivants proposent une définition de ces objets ainsi que des critères quantitatifs qui peuvent être utilisés avec ces objets.

4.3.1 Objet référence

4.3.1.1 Définition

Il s'agit ici de définir un objet de référence, qui sera représentatif de l'ensemble des interprétations. La saisie des interprètes se fait en mode vectoriel, par des fonctions linéaires par morceaux.

4.3.1.2 Détermination

La figure (4.3) représente le procédé d'extraction du contour qui servira de référence à partir des N saisies des interprètes. Par souci de clarté, nous n'avons représenté que deux saisies vectorielles ainsi qu'un seul bord de l'objet (ici la rue) à extraire.

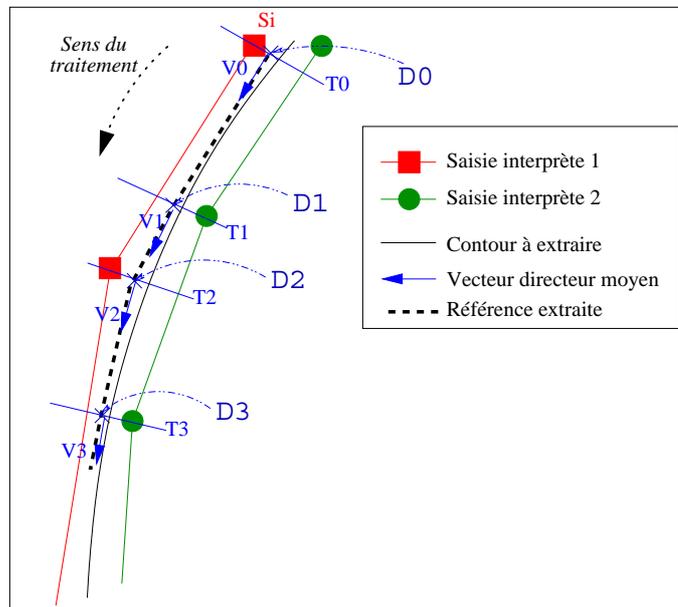


FIG. 4.3: extraction de la référence à partir des interprétations d'un bord de l'objet

L'extraction de la polyligne de référence se fait comme suit (le sens de traitement est défini sur la figure 4.3, mais a peu d'influence sur le résultat final) :

- le point de départ du traitement est choisi comme étant le sommet de polyligne rencontré en dernier dans le sens du traitement (sommet S_i sur la figure 4.3). En ce point, on calcule le vecteur directeur \mathbf{V}_0 , moyenne des $N-1$ vecteurs directeurs des segments de polyligne situés avant le sommet S_i ,
- on considère ensuite le transect (T_0) orthogonal au vecteur \mathbf{V}_0 . Une moyenne est effectuée entre les points situés à l'intersection entre (T_0) et les différentes polygones (en rejetant les modes extrêmes). On obtient le point D_0 qui sera un sommet de la polyligne de référence,
- on réévalue le vecteur directeur moyen \mathbf{V}_1 , à partir des N segments de polyligne situés après le transect (T_0). On obtient ainsi le transect (T_1) orthogonal à \mathbf{V}_1 que

l'on « propage » dans la direction de \mathbf{V}_1 jusqu'au premier sommet rencontré. On recalcule alors le point D_1 ,

- cette procédure est répétée jusqu'à ce l'on rencontre le sommet final d'une des polygones.

Par rapport à une polygone saisie par un interprète, le contour référence obtenu a un nombre de sommets plus élevé et ses variations de pente sont atténuées. En effet, la variation de pente entre chaque sommet de la référence correspond à la contribution d'une seule polygone. Du fait du moyennage et de l'augmentation du nombre de sommets, la polygone de référence extraite apparaît donc plus lisse que les polygones saisies par chacun des interprètes.

4.3.2 Zone de tolérance

4.3.2.1 Définition

L'intervalle de tolérance est représentatif de l'écart d'interprétation entre les différents interprètes.

L'utilité de la zone de tolérance est double.

- Elle permet de caractériser l'incertitude sur la référence extraite. En effet, une zone de tolérance peu étendue traduit une faible variabilité des interprétations et donc une référence extraite plus fiable.
- Elle permet aussi d'évaluer la qualité de l'extraction par des méthodes automatiques en établissant des critères quantitatifs.

Dans les deux cas, l'utilisation de l'interprétation humaine peut engendrer l'introduction d'un biais sur la référence et sur la zone de tolérance (possibilité d'une erreur collective d'interprétation).

4.3.2.2 Détermination

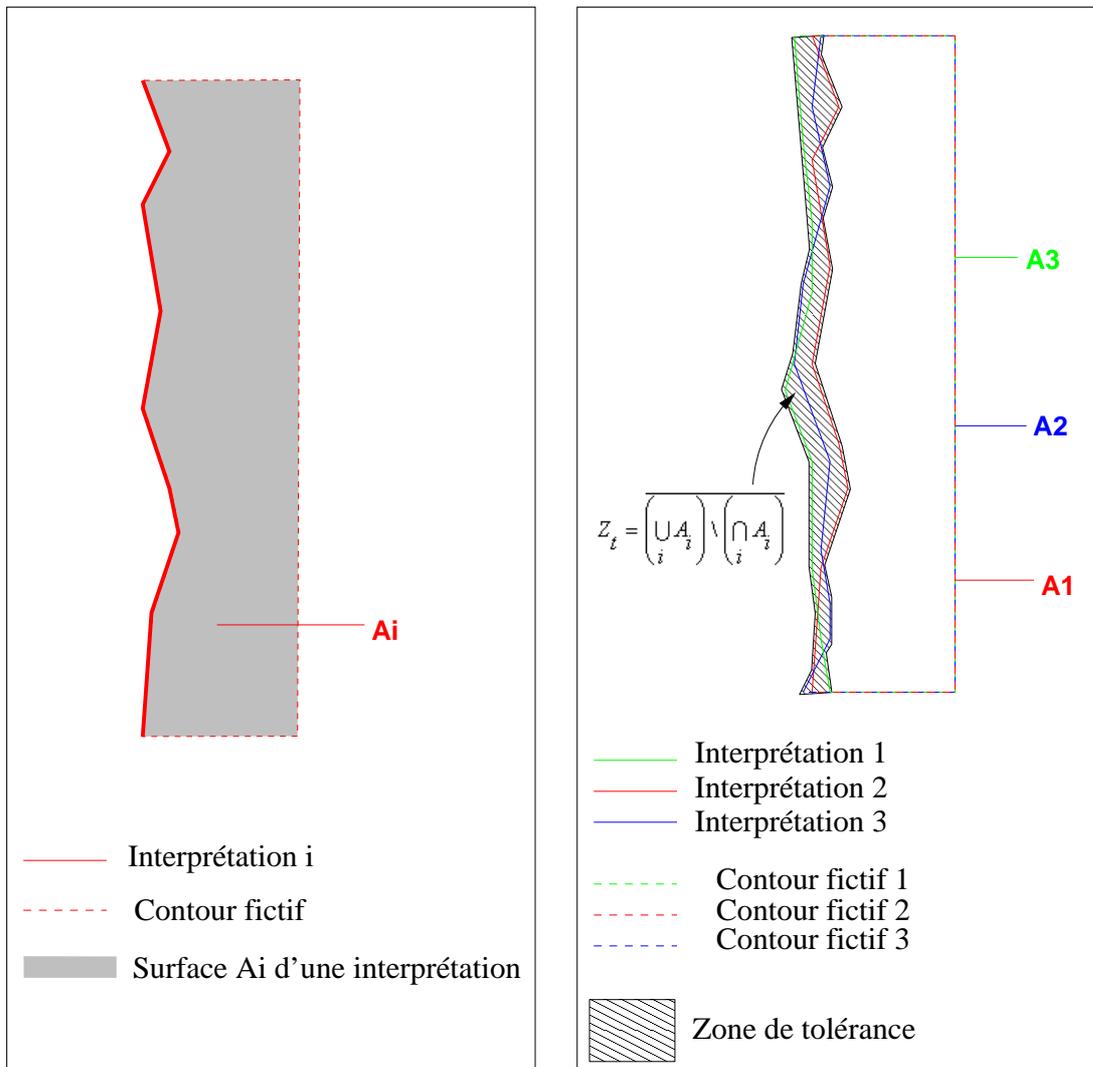
On considère la surface A_i définie par le contour de l'objet saisi par l'interprète i . Dans le cas de courbe non fermée, cas des bords de rue, on fermera la forme par un contour fictif (voir figure 4.4(a)). Le contour fictif est un objet intermédiaire qui doit être commun à toutes les interprétations. Il permet de déterminer les surfaces A_i et la zone de tolérance Z_t et n'intervient plus dans la suite du traitement.

La zone de tolérance Z_t est alors définie comme l'union des A_i privée de l'intersection des A_i :

$$Z_t = \overline{\left(\bigcup_i A_i \right)} \setminus \overline{\left(\bigcap_i A_i \right)} \quad (4.1)$$

avec A_i la surface d'une interprétation i , la barre située au dessus désignant *la fermeture* de l'ensemble.

La zone de tolérance Z_t est représentée sur la figure 4.4(b).



(a) Surface définie par une interprétation et un contour fictif

(b) Zone de tolérance

FIG. 4.4: détermination de la zone de tolérance à partir des interprétations

4.3.2.3 Procédé algorithmique

L'implémentation de la formule (4.1) peut se révéler délicate de par l'apparition de nouveaux points lorsque deux polygones se croisent.

La solution algorithmique choisie, est d'effectuer une triangulation de *Delaunay* contrainte par les différentes polygones et les contours fictifs. La frontière de l'ensemble défini est alors composée des segments n'appartenant qu'à un seul triangle (figure 4.5).

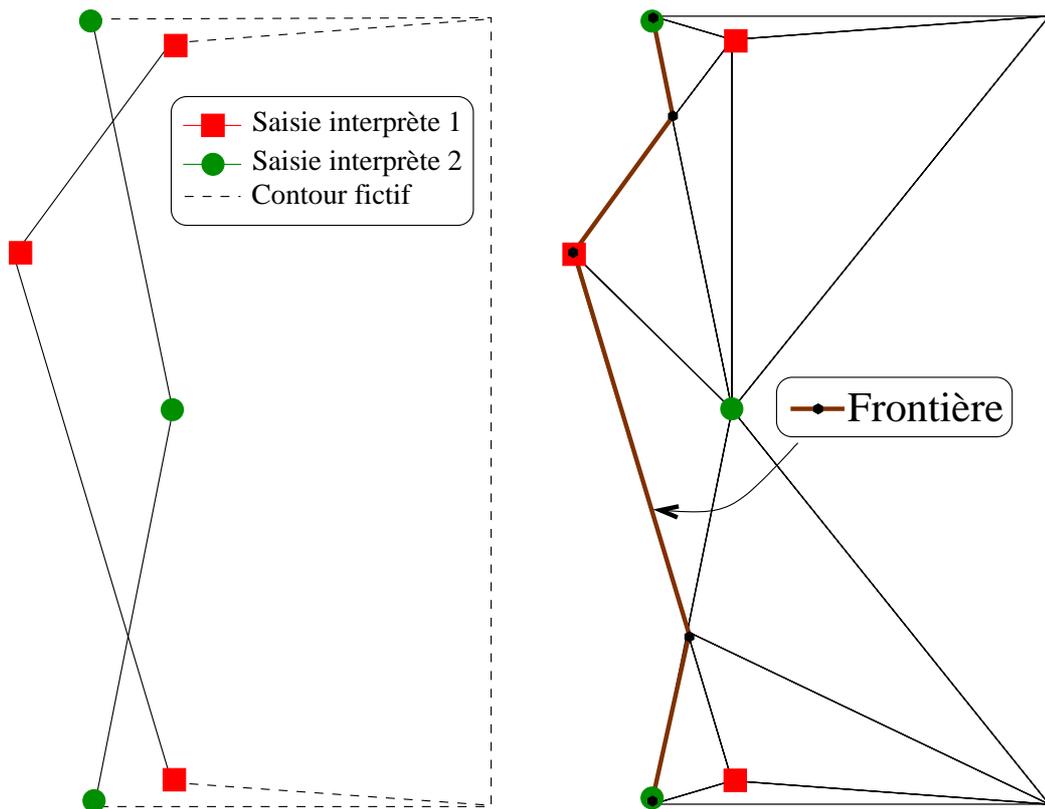


FIG. 4.5: frontière «gauche» de la zone de tolérance obtenue par triangulation de Delaunay contrainte.

4.4 Application à une scène réelle

4.4.1 Protocole d'expérimentation

L'image sur laquelle la campagne de saisie a été effectuée est une image du satellite Ikonos. Cette image nous a été gracieusement fournie par la société GIM² à des fins de test.

L'image a une résolution spatiale de 1 mètre et est extraite de la région de Hasselt, Belgique. La scène représente un milieu péri-urbain traversé par une rue principale sur laquelle les méthodes d'extraction de la référence et de la zone de tolérance vont être appliquées (figure 4.6).



FIG. 4.6: image originale de la rue à acquérir
© 2000 Space Imaging Europe

Il a été demandé à un panel de 11 interprètes d'images de numériser leur vision de cette rue sur l'image en plaçant les bords (saisie vectorielle). Ces personnes travaillent dans le domaine de l'image satellitale et sont donc familières des images numériques. La diversité entre les saisies sera considérée comme la variabilité d'interprétation. La saisie vectorielle

²Geographic Information Management

des bords de rue s'est effectuée avec le logiciel ArcView, qui comprend plusieurs aides à l'interprétation, dont un zoom pour que l'interprète ait à la fois une vision locale et une vision du contexte. L'opérateur a aussi la possibilité de régler le contraste de l'image pour faire ressortir les contours.

Comme le montre la figure 4.7 où deux profils radiométriques de la rue sont représentés, le profil radiométrique de la rue n'est pas constant. C'est pour cette raison que dans notre méthode d'extraction (section 3.1.4) nous avons privilégié les contraintes géométriques comme le parallélisme des bords de la rue plutôt que des contraintes sur la radiométrie.

Le transect (4.7.b) correspond à un profil radiométrique «classique» de rue : homogène le long de l'axe de la rue et contrasté dans la direction transverse à cet axe.

Le transect (4.7.a) a par contre un profil radiométrique altéré, dû à la présence d'occlusions (ici des arbres). Dans ce cas, le bord de la rue ne représente pas un extremum radiométrique idéal. Il est donc nécessaire d'avoir une vision de la rue dans son ensemble pour une bonne extraction, contribuant en conséquence à la variabilité d'interprétation.

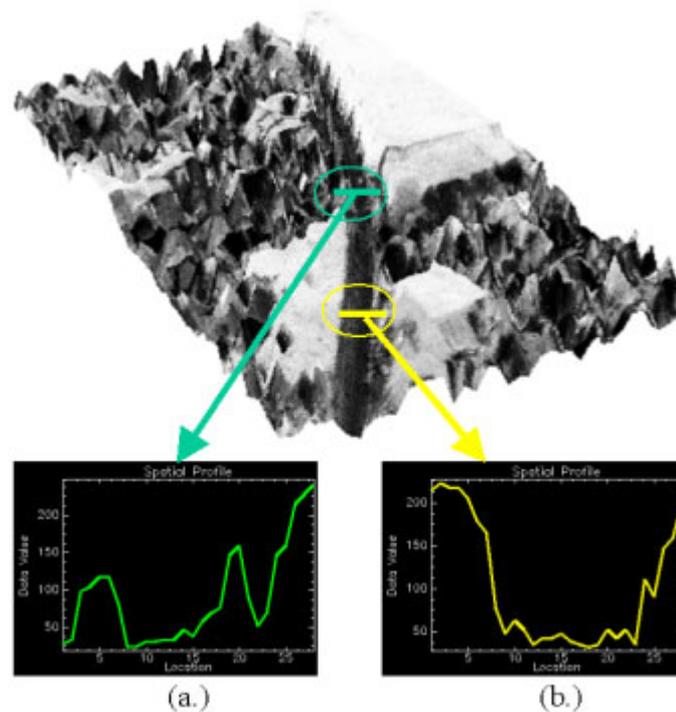


FIG. 4.7: représentation en intensité radiométrique de la figure 4.6 et exemple de transects de la rue

4.4.2 Extraction d'une référence et de la zone de tolérance

4.4.2.1 Analyse de la saisie

Sur la figure 4.8, la superposition des 11 acquisitions est représentée. Prises séparément, chaque acquisition semble correspondre au bord de la rue dans l'image. Cependant, la variabilité d'interprétation est apparente, et se traduit par un épaissement des 2 bords de la rue sur cette image.



FIG. 4.8: superposition des 11 acquisitions

Afin d'effectuer une analyse plus précise de cette variabilité, nous avons choisi de nous focaliser sur la région comprenant une partie du bord droit de la rue (cadre sur la figure 4.8). Cette région possède une courbure relativement importante ainsi que des occlusions dues aux arbres, et génère donc une grande variabilité d'interprétation.

La figure 4.9 représente le cadre de la figure 4.8 (dimensions $15 \text{ m} \times 20 \text{ m}$), ainsi que les 11 acquisitions. Deux polygones (6 et 10) ont des sommets compris dans cette zone (S_6 et S_{10} sur la figure 4.9). En accord avec la méthode décrite dans la section 4.3, les sommets du contour vont être calculés en S_6 puis S_{10} .

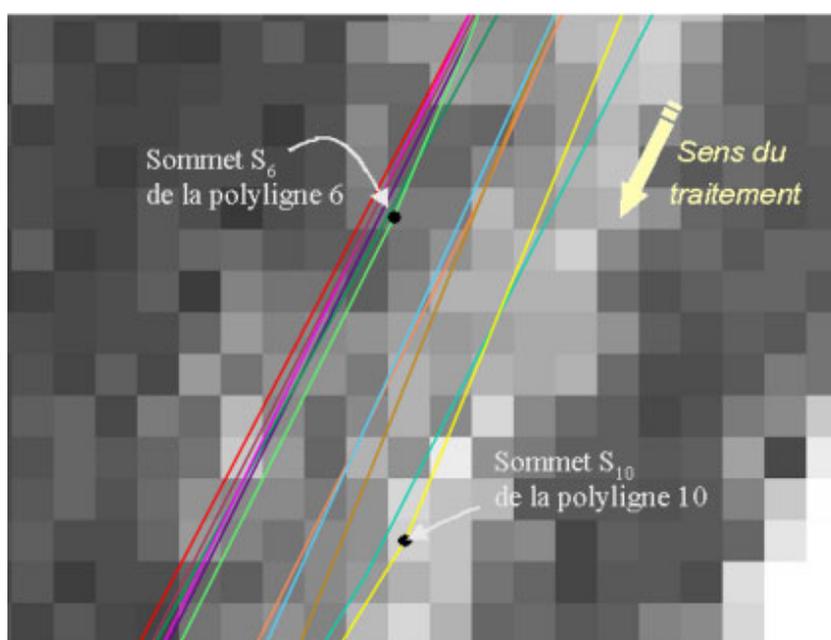


FIG. 4.9: superposition des 11 acquisitions (zoom sur le bord droit de la rue 4.8)

4.4.2.2 Contour de référence

On suppose que le vecteur directeur moyen \mathbf{V}_6 a été estimé auparavant en moyennant les coefficients directeurs des polygones situées avant le sommet S_6 (relativement au sens du traitement). On considère ensuite le transect (T_6) au sommet S_6 orthogonal à \mathbf{V}_6 (figure 4.10).

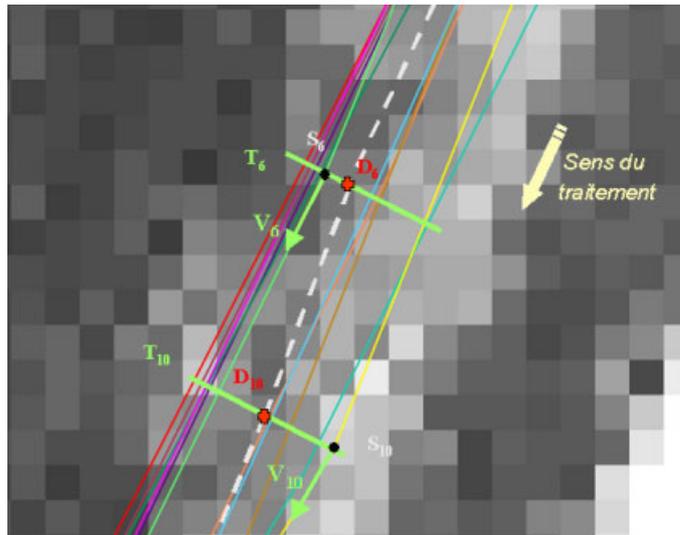


FIG. 4.10: extraction du contour de référence (bord droit de la rue). La ligne en tirets correspond au contour de référence obtenu.

Une moyenne est effectuée entre les points situés à l'intersection entre (T_6) et les différentes polygones (en rejetant les points extrêmes). Le point D_6 obtenu correspond à un sommet pour la polygone de référence (figure 4.3).

Sur la figure 4.11, le profil radiométrique du transect T_6 est représenté.

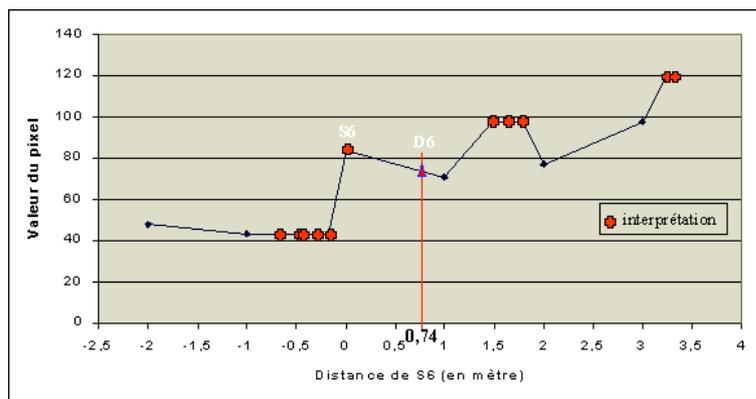


FIG. 4.11: profil radiométrique du transect T_6 (bord droit)

Les différents points correspondant aux différentes interprétations sont reportés sur cette courbe, avec en abscisse la distance de ces points par rapport au point S_6 . Le point D_6 correspond à la moyenne de ces distances (en rejetant les 2 points extrêmes).

L'écart-type par rapport à D_6 des différentes interprétations au transect T_6 est de 1,33 m (soit 1,33 pixel pour l'image Ikonos à 1 mètre de résolution).

Le même procédé a été appliqué pour le transect T_{10} , donnant un écart-type de 1,65 m. Même en ne retenant que la précision du pixel, un tel écart montre la variabilité d'une référence basée sur une seule interprétation.

Par rapport aux autres polygones saisis par les interprètes, le contour référence obtenu a un nombre de sommets plus élevé et des variations de pentes atténuées, et apparaît donc plus lisse.

4.4.2.3 Zone de tolérance

La figure 4.12 représente la zone de tolérance pour le bord droit de la rue, calculée selon la méthode définie par triangulation de Delaunay contrainte.

La frontière de la zone est composée pour sa limite gauche de la polyligne 1 et pour sa limite droite d'une partie de la polyligne 10 et d'une partie de la polyligne 11. La polyligne de référence extraite précédemment est aussi tracée dans la zone de tolérance, ainsi que ses deux sommets D_6 et D_{10} .

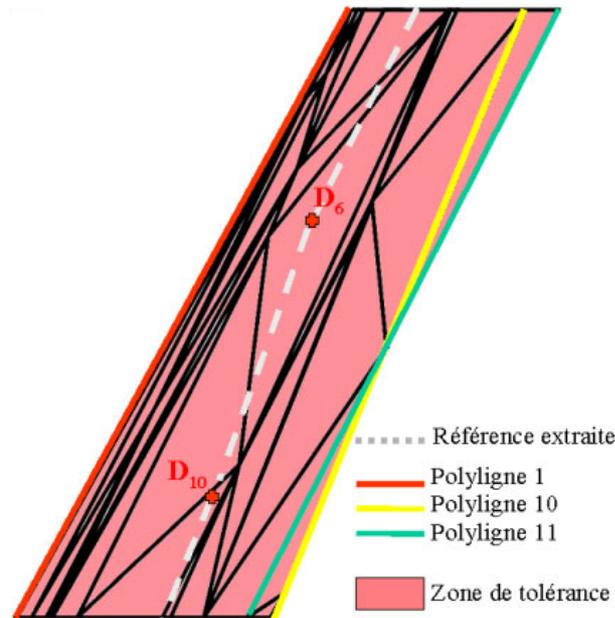


FIG. 4.12: zone de tolérance pour le bord droit de la rue calculée à partir d'une triangulation de Delaunay contrainte

La référence se situe bien dans la zone de tolérance, et dans ce cas presque dans l'axe de cette zone. La référence extraite, ainsi que la zone de tolérance, sont des objets analytiques. Suivant le type d'utilisation souhaitée ou de la précision recherchée, elles peuvent être utilisées comme données vectorielles et donc intégrables dans un système d'information géographique, ou être converties sur une grille pixel.

4.5 Évaluation quantitative à partir de la référence et de la zone de tolérance

4.5.1 Évaluation quantitative à partir de la référence

Afin de comparer un contour extrait et la référence obtenue, trois critères de similarité géométrique vont être utilisés.

Distance de Hausdorff. Le premier critère de similarité retenu est la distance de Hausdorff. Elle permet de bien caractériser la ressemblance et la proximité de deux objets vectoriels (Abbas 1994).

Soit les deux polygones :

$P_{ref} = (p1_{ref}, \dots, pm_{ref})$, la polygone de référence, et $P_{ext} = (p1_{ext}, \dots, pm_{ext})$ la polygone extraite automatiquement.

La distance de Hausdorff est définie comme :

$$H(P_{ext}, P_{ref}) = \max(h(P_{ext}, P_{ref}), h(P_{ref}, P_{ext}))$$

avec :

$$h(A, B) = \max_{a \in A} (\min_{b \in B} (d(a, b)))$$

$d(a, b)$ étant la distance euclidienne entre les deux points a et b .

La distance de Hausdorff vérifie les propriétés d'identité, de symétrie et d'inégalité.

Critère de longueur. Soit L_{ext} est la longueur totale du contour extrait et L_{ref} la longueur totale du contour référence, on définit l'erreur relative sur la longueur comme :

$$Err_{longueur} = \frac{|L_{ext} - L_{ref}|}{L_{ref}} \quad (4.2)$$

Critère de surface. Si l'objet extrait et l'objet référence sont des contours fermés, on peut, de plus, définir un critère de similarité de surface.

Soit A_{ext} l'aire totale du contour extrait (fermé) et A_{ref} l'aire totale du contour référence, on définit l'erreur relative sur la surface comme :

$$Err_{surface} = \frac{|A_{ext} - A_{ref}|}{A_{ref}} \quad (4.3)$$

4.5.2 Évaluation quantitative à partir de la zone de tolérance

La zone de tolérance traduit la variabilité des interprétations humaines. Calculer le pourcentage de la longueur du contour extrait par un algorithme inclus dans cette zone de tolérance va traduire le degré de similarité du résultat avec l'interprétation humaine. Ce critère est moins discriminant que la comparaison à une référence, mais il permet de déterminer les résultats trop éloignés de l'interprétation visuelle.

4.6 Conclusion sur ce chapitre

Ce chapitre présente une méthode pour établir une référence de comparaison lorsque la seule source d'information à disposition est l'image satellite ou aérienne. Basée sur l'interprétation d'image, cette référence permet une évaluation relative de la qualité des résultats. Dans le but de prendre en compte la variabilité d'interprétation entre différents opérateurs, il a été proposé une méthode pour rendre la référence moins subjective. C'est en effet la comparaison à une référence objective qui va permettre d'évaluer et de quantifier les performances des méthodes d'extraction automatique.

Une référence *a priori* est définie, en représentation vectorielle et basée sur des statistiques calculées sur plusieurs interprètes d'images. Une acquisition moyenne est estimée qui servira à l'évaluation quantitative en utilisant la distance de Hausdorff.

Une zone de tolérance est aussi définie, représentative de la variabilité d'interprétation. L'évaluation quantitative se fera en calculant le pourcentage du contour extrait situé dans cette zone. Le cas réel sur lequel a été appliquée cette méthode confirme cette importante variabilité.

Plus généralement, la référence et la zone de tolérance permettent de développer toutes sortes de critères quantitatifs (rectitude de la rue, emprise au sol, ...). Plusieurs critères quantitatifs sur les rues ainsi que des critères sur le réseau sont présentés dans Péteri *et al.* (2003). Il est à noter que les critères seront à adapter suivant le type des objets à extraire (cas des bâtiments par exemple).

Dans le cas où l'on dispose, sur une zone donnée, d'images à plusieurs résolutions différentes, la référence peut être constituée avec les images de meilleures résolutions, qui constituent la «vérité terrain» vis-à-vis des autres images.

La variabilité d'interprétation visuelle augmente d'autant plus que la scène est complexe (cas du domaine urbain). Elle augmente également avec l'amélioration de la résolution spatiale des capteurs. Pour ces raisons, il nous était important de bien définir la référence puis les critères de comparaison avant d'estimer les performances de l'algorithme.

Le chapitre 5 suivant propose des applications de la méthode d'extraction sur différentes scènes et différents capteurs satellites. L'évaluation quantitative des résultats d'extraction sera effectuée par la méthode présentée dans ce chapitre.

Chapitre 5

Applications de la méthode

Dans ce chapitre, nous présentons différentes applications de la méthode d'extraction de rues définie au chapitre 3. La méthode est appliquée sur une image synthétique puis sur des images de scènes réelles. Ces images proviennent des satellites à très haute résolution spatiale comme Ikonos ou Quickbird, ainsi que de la caméra numérique de l'IGN. Les résultats sont analysés et évalués quantitativement selon la méthode définie au chapitre 4.

5.1 Introduction

5.1.1 Les cas d'étude

Les exemples présentés dans ce chapitre rassemblent différents cas permettant de tester les performances de la méthode : quartier péri-urbain, zone urbaine, rues avec différentes courbures, réseaux quadrangulaires, différents types d'intersections et plus ou moins d'artefacts sur les rues. Nous avons aussi considéré des images de différents capteurs et à différentes résolutions spatiales.

Les deux premières applications (une scène synthétique, section 5.2, et une scène réelle du capteur Ikonos, section 5.3) vont illustrer les différentes étapes composant la phase de reconstruction définie section 3.3. Pour ces deux cas, le graphe initial est donné manuellement. Ensuite, deux exemples complets (extraction du graphe topologique puis reconstruction surfacique du réseau) sont donnés, concernant des scènes urbaines observées par le satellite Quickbird et la caméra numérique de l'IGN.

5.1.2 L'évaluation quantitative

De manière intuitive pour un interprète humain, la rue est définie par ses deux bords d'une intersection à l'autre. Ainsi, pour le protocole de saisie défini au chapitre 4, il a été demandé aux interprètes d'images de saisir successivement les tronçons de rue à bords parallèles puis les intersections (et non de suivre un bord sur toute l'image). On peut alors comparer les performances de l'algorithme selon deux classes : les tronçons de rue à bords parallèles et les intersections. On en dérivera ensuite une caractérisation globale du résultat d'extraction.

5.2 Application à une image synthétique

Description de la scène. Ce test permet d'illustrer les différentes étapes composant l'étape de reconstruction. La scène synthétisée représente trois rues se croisant en une intersection en Y (figure 5.1). Deux classes de rues différentes ont été définies dans cette scène : la rue 1 et la rue 2 sont des rues de classe 3 et la rue 3 est de classe 2. Ceci va déterminer la distance de propagation mais aussi l'échelle à partir de laquelle les *DoubleSnakes* vont commencer à minimiser leur énergie ($j = 4$ pour les rues 1 et 2 et $j = 3$ pour la rue 3). Du bruit simulant des petits artefacts ponctuels a été ajouté afin de tester la robustesse de l'algorithme de reconstruction vis-à-vis de ce bruit.

Les différentes étapes. Sur la figure 5.1, le graphe du réseau et le cercle d'intersection sont reportés. Le graphe a été saisi manuellement mais n'a pas été placé exactement sur l'axe central, notamment pour la rue 3 (figure 5.1).

Sur la figure 5.2, les contours parallèles ont été initialisés par la propagation et l'échantillonnage des polygones du graphe. Les portions de contours contenus dans le cercle d'intersection ont été tronquées, et les points extrêmes ont été placés sur le cercle d'intersection.

Sur la figure 5.3, les tronçons à bords parallèles ont été extraits par les trois *DoubleSnakes* après avoir évolué à différentes échelles. Pendant l'évolution des contours actifs, les points situés sur le cercle d'intersection ont été contraints d'y rester. L'énergie image calculée à partir des différentes échelles, associée à l'énergie de parallélisme a permis un bon comportement des *DoubleSnakes* vis-à-vis des pics de bruit. La rue 3, la moins bien initialisée, a notamment été détectée correctement.

Sur la figure 5.4, les contours actifs d'intersection ont été initialisés par la position finale des *DoubleSnakes*. On peut noter que le type d'initialisation diffère selon l'angle entre deux rues adjacentes : ainsi pour l'intersection reliant la rue 1 et la rue 3, un point de rupture a été introduit par l'algorithme (l'angle valant moins de $\frac{3\pi}{5}$).

La figure 5.5 présente le résultat final de l'extraction, qui s'est montré robuste vis-à-vis du bruit introduit. Les *IntersectionSnakes* ont commencé la minimisation de leur fonctionnelle à partir de l'échelle 3. Ils ont conservé la continuité \mathcal{C}^1 au niveau de cercle d'intersection.

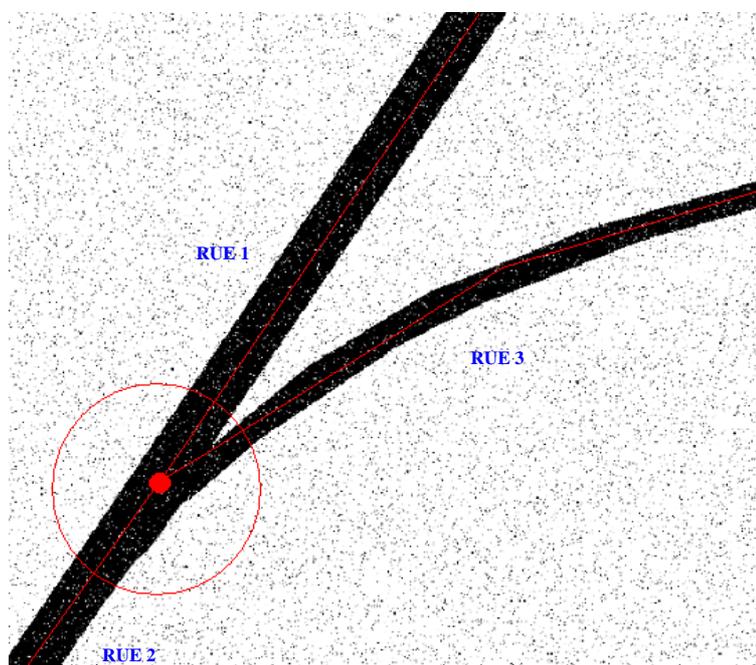


FIG. 5.1: graphe du réseau et cercle d'intersection

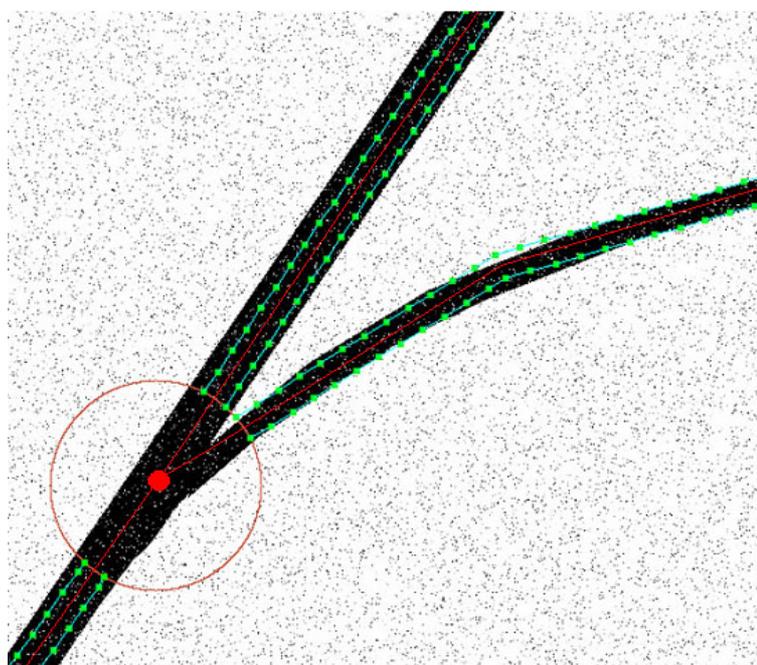


FIG. 5.2: après propagation des contours parallèles

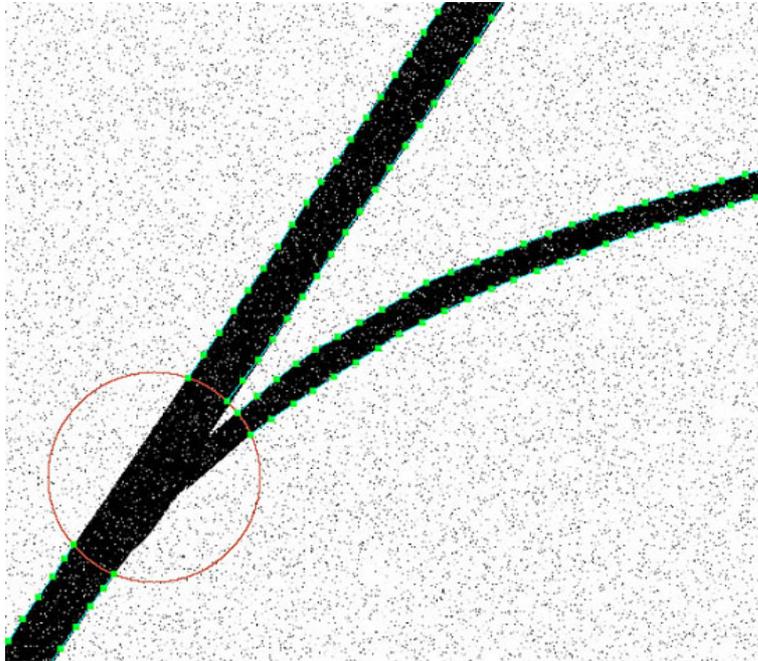


FIG. 5.3: après extraction des contours parallèles

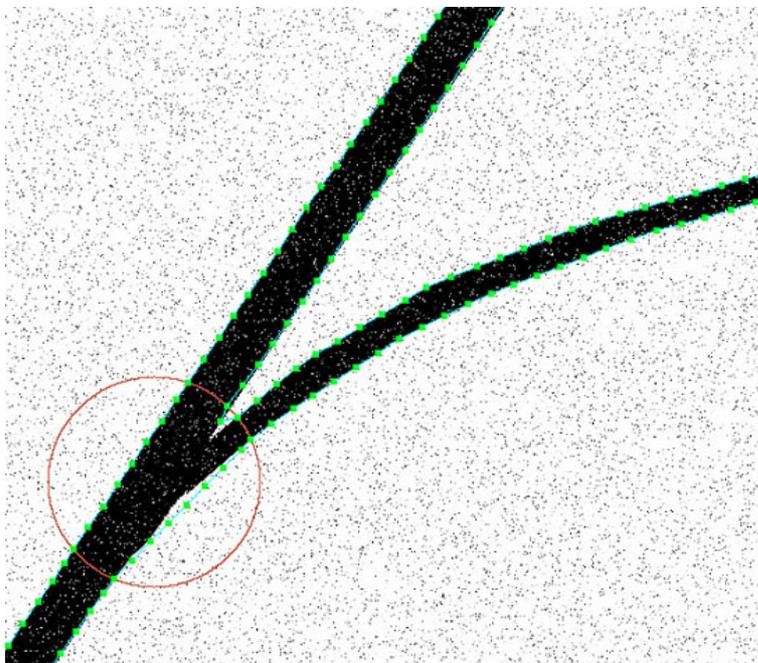


FIG. 5.4: initialisation de l'intersection

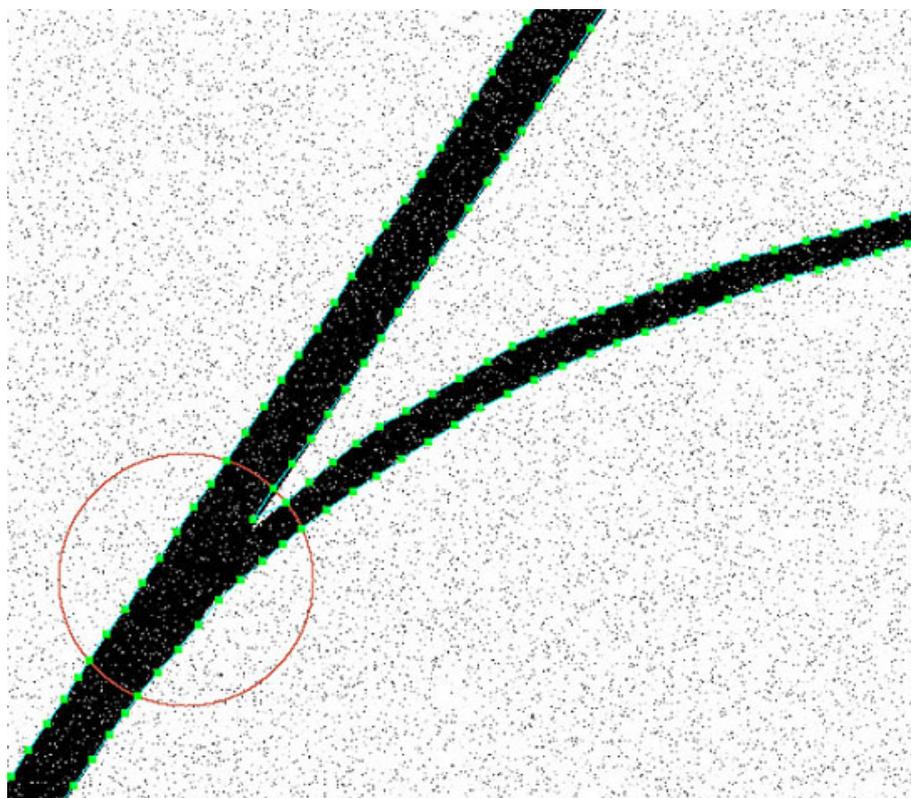


FIG. 5.5: après extraction de l'intersection

Recalage de l'axe central de la rue. Même si cela ne constitue pas notre objectif premier, il peut être intéressant d'extraire l'axe central de la rue en sortie de l'algorithme de reconstruction. On peut ensuite l'inclure dans une base de données géographiques. Il est estimé tout simplement à partir de l'extraction surfacique en calculant le milieu du segment reliant deux points appariés.

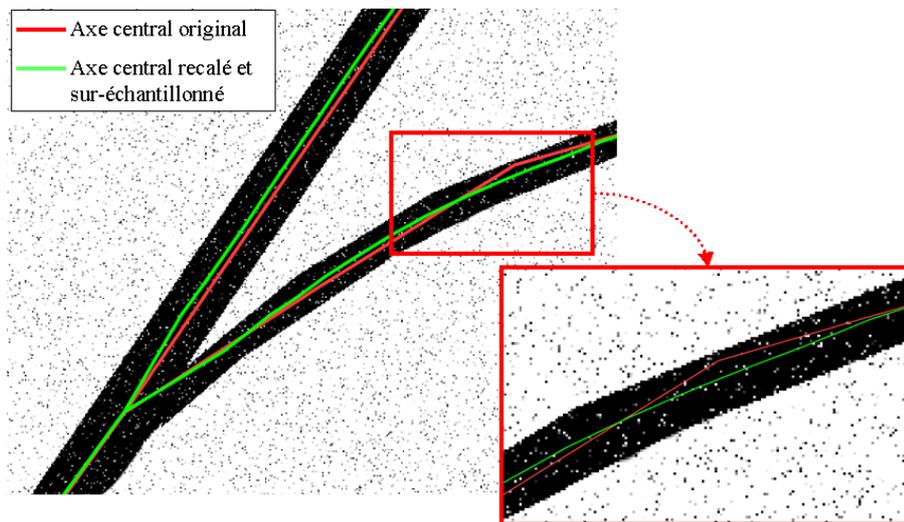


FIG. 5.6: recalage du graphe du réseau après l'étape de reconstruction

La figure 5.6 représente le graphe initial et le graphe recalé, ainsi qu'un zoom sur une portion de rue où sont tracés la polyligne initiale du graphe ainsi que l'axe central extrait par l'algorithme. Ce dernier est mieux recalé que la polyligne de départ et est aussi plus échantillonné (le taux d'échantillonnage est celui des *DoubleSnakes*).

Évaluation des résultats. Ce cas est une image synthétique dont on connaît la localisation de la rue avec précision (avant l'ajout du bruit). Il n'a pas été nécessaire dans ce cas de faire saisir la référence de comparaison par plusieurs opérateurs humains.

Les critères d'évaluation retenus sont la distance de Hausdorff et la similarité entre les longueurs et les surfaces pour la référence et la rue extraite par l'algorithme. Les tableaux 5.8, 5.9 et 5.10 présentent ces différents critères quantitatifs.

La distance de Hausdorff traduit l'écart maximal entre deux polygones (tableau 5.8). Dans le cas des tronçons de rue, et compte tenu des effets de quantification spatiale, elle peut être considérée comme très faible (de l'ordre du pixel). Sa valeur de 2,2 pour les intersections est principalement due au léger décalage sur la pointe de l'intersection entre la rue 1 et 3 (les contraintes de rigidité ont prévalu sur le terme d'attache aux données).

Les critères de surface et de longueur indiquent une très bonne adéquation entre la référence et les rues extraites : l'erreur relative totale n'est que de 0,31 % pour le critère de longueur et de 0,3 % pour celui de surface.

	Moyenne sur les tronçons de rue	Moyenne sur les intersections
Distance de Hausdorff	1,06	2,2

TAB. 5.1: distance de Hausdorff rue extraite - rue référence (en pixel)

Longueur	Rue extraite	Rue référence	Différence	Différence relative
Tronçons de rue	2350,41	2341,77	8,64	0,36%
Intersections	341,5	342,6	1,1	0,32%
Cumulée	2691,9	2683,6	8,3	0,31%

TAB. 5.2: critère de longueur (en pixel)

Surface	Rue extraite	Rue référence	Différence	Différence relative
Tronçons de rue	36650	36740	90	0,2%
Intersections	7082	7130	48	0,6%
Cumulée	43732	43870	138	0,3%

TAB. 5.3: critère de surface (en pixel²)

Les sections suivantes présentent plusieurs applications de la méthode sur des scènes réelles vues à partir de différents capteurs.

5.3 Image Ikonos de la zone de Hasselt, Belgique

Description de la scène. L'image 5.7 provient du satellite Ikonos de la société SpaceImaging. Ce satellite acquiert des images en mode panchromatique à la résolution de 1 mètre. Elle nous a été aimablement fournie par la société GIM (Geographic Information Management) à des fins de tests. La scène est un quartier résidentiel en périphérie de la ville d'Hasselt en Belgique. L'image comprend une intersection de deux rues principales que nous avons définies comme étant des rues de classe 2. Les *DoubleSnakes* vont donc évoluer à partir de l'échelle 3 (voir tableau 3.2).

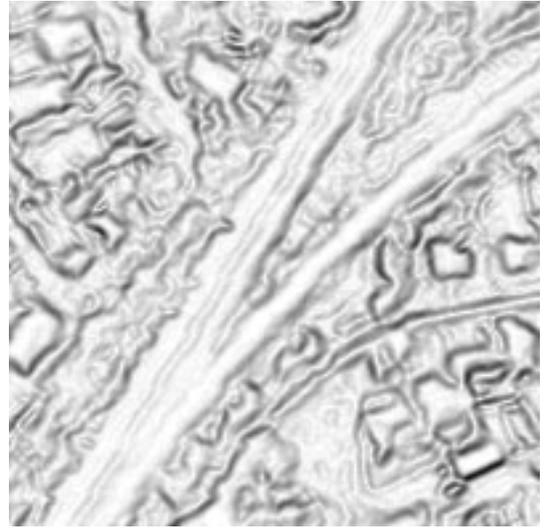


FIG. 5.7: image Ikonos (région de Hasselt, Belgique) © *SpaceImaging*

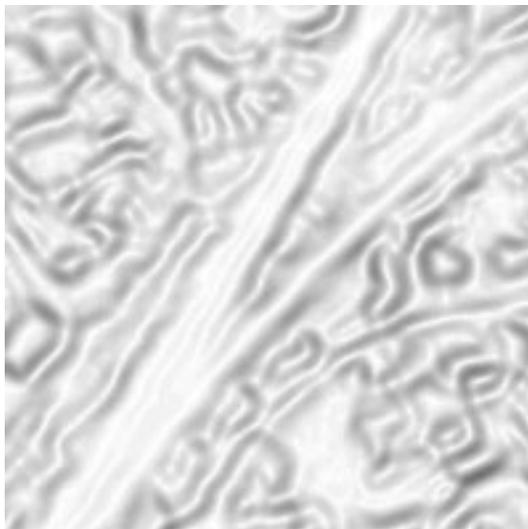
La figure 5.8 présente une zone de cette image où le module de la transformée en ondelettes est représenté pour les échelles 1, 2 et 3. On peut remarquer que les détails fins, comme le marquage au sol, disparaissent à l'échelle 3 sous l'effet de lissage. Le module de la transformée en ondelettes étant proportionnel à l'énergie image, les *DoubleSnakes* ne seront pas piégés par ces maxima locaux d'énergie lorsqu'ils évolueront à l'échelle 3.



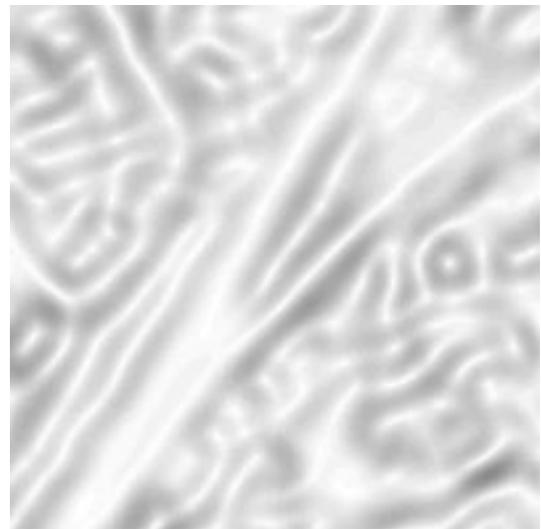
(a) extrait de l'image 5.7 à la résolution originale © *SpaceImaging*



(b) module de la transformée en ondelettes à l'échelle 1



(c) module de la transformée en ondelettes à l'échelle 2



(d) module de la transformée en ondelettes à l'échelle 3

FIG. 5.8: module de la transformée en ondelettes à différentes échelles. Le marquage au sol central a disparu à l'échelle 3 sous l'effet de lissage.

Extraction des rues. Sur la figure 5.9, le graphe d'entrée, saisi ici manuellement de manière peu précise par rapport à l'axe central, a été superposé. Le cercle d'intersection est aussi reporté.



FIG. 5.9: image originale avec le graphe

Les images 5.10, 5.11 et 5.12 représentent les différentes phases du déroulement de l'algorithme, similairement à l'exemple synthétique précédent.

Sur l'image 5.10, les polygones constituant le graphe d'entrée ont été échantillonnés puis propagés.

Sur l'image 5.11, les sections de rues à bords parallèles ont été extraites par les *DoubleSnakes* après avoir évolué à différentes échelles. L'utilisation de l'approche multi-échelle a permis aux *DoubleSnakes* de ne pas être piégés par le bruit (principalement le marquage au sol). De plus, les contraintes de régularisation et l'énergie de parallélisme ont permis d'extraire correctement les deux bords de la rue malgré une initialisation de mauvaise qualité (cas du milieu de la rue 3).

Le résultat final d'extraction après l'initialisation et le traitement des intersections est représenté sur l'image 5.12. Visuellement, tous les contours ont été extraits correctement. Seule la pointe de l'intersection (entre la rue 2 et la rue 3) a été tronquée à cause de sa faible énergie image et de son initialisation trop éloignée.

Conformément à notre démarche et en s'appuyant sur le protocole de comparaison défini au chapitre 4, il est nécessaire d'établir la référence et sa zone de tolérance. Ainsi, une saisie de la rue a été réalisée par un ensemble d'interprètes d'images.



FIG. 5.10: après propagation des polygones



FIG. 5.11: après traitement des contours parallèles



FIG. 5.12: après traitement des intersections

Évaluation quantitative Il a été demandé à 8 experts de saisir vectoriellement leur interprétation de la scène avec le logiciel ArcView. Le résultat de cette saisie est présenté sur la figure 5.13. Un zoom au niveau de l'intersection est représenté figure 5.14. On peut visuellement remarquer un certain épaississement des contours, traduisant l'écart entre les différentes interprétations.

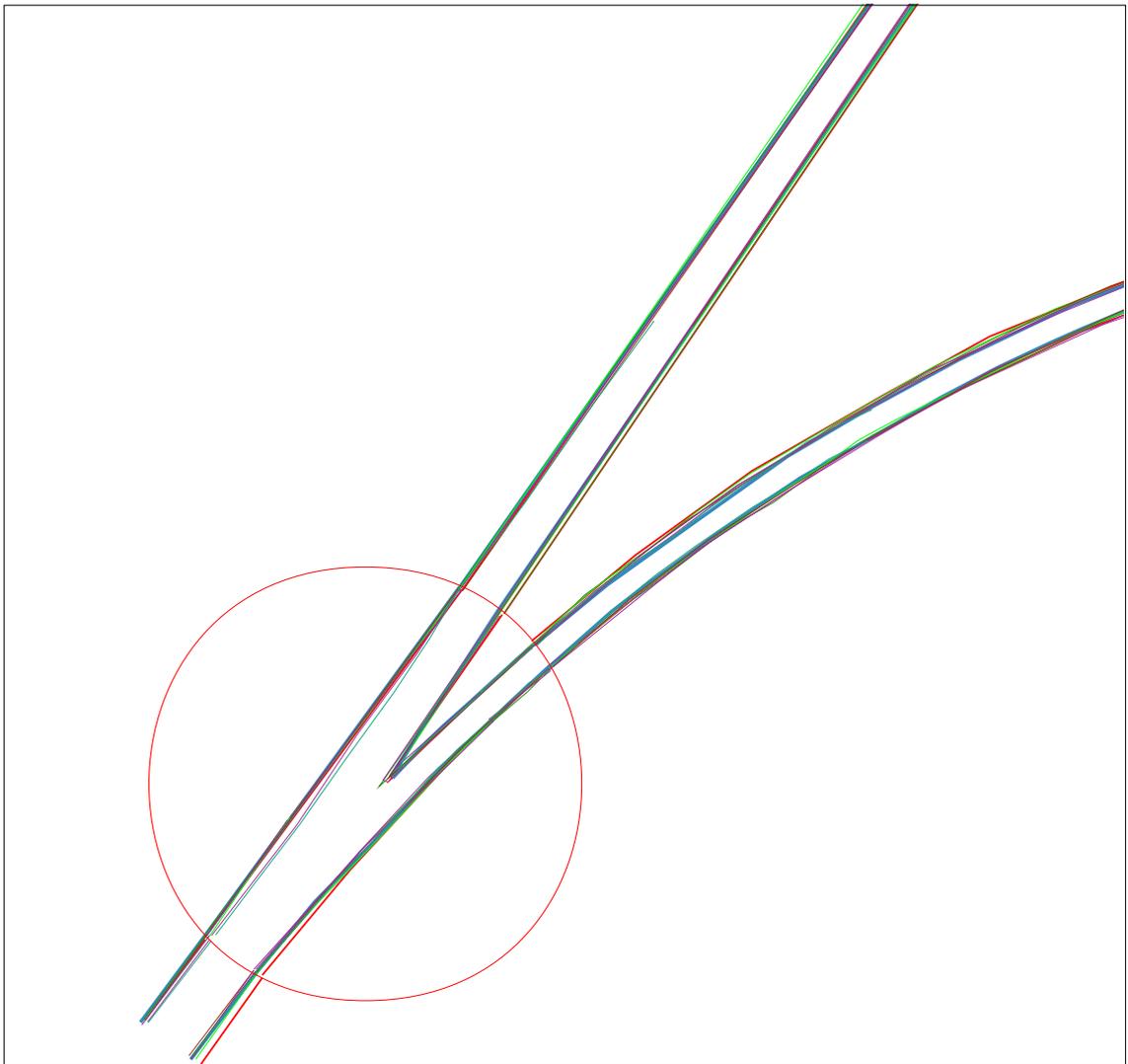


FIG. 5.13: superposition des 8 interprétations de la scène

A partir de ces interprétations, la zone de tolérance et la référence ont été extraites comme décrit au chapitre 4.

La figure 5.15 représente un zoom au niveau de l'intersection, où sont représentés la zone de tolérance (zone bleue), la référence obtenue (trait pointillé rouge), ainsi que le contour extrait par l'algorithme (trait plein vert). La zone de tolérance est assez étroite,

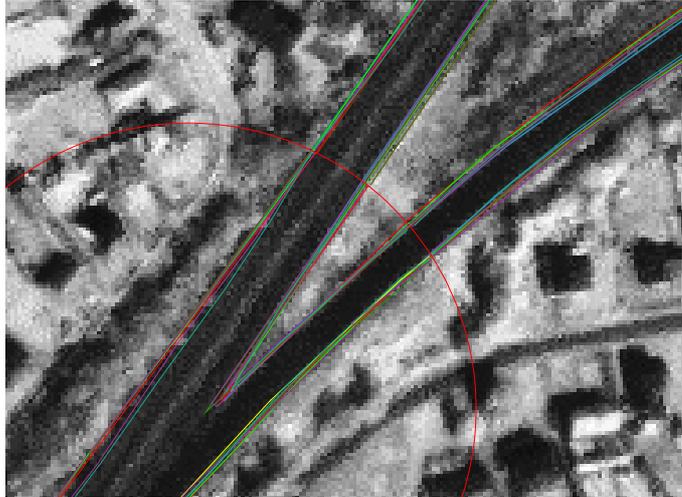


FIG. 5.14: superposition des 8 interprétations de la scène (zoom)

signe d'une bonne cohérence entre les différents interprètes d'images. La référence est bien incluse dans la zone de tolérance. Enfin, on peut observer la bonne localisation du contour extrait par l'algorithme, qui mise à part la pointe de l'intersection, se situe bien dans la zone de tolérance et est proche de la référence. La distance de Hausdorff et l'étude des longueurs et des surfaces extraites quantifient la similarité géométrique entre la référence et la rue extraite. Le pourcentage du contour extrait par l'algorithme automatique inclus dans la zone de tolérance traduit son adéquation avec l'interprétation humaine.

Les différents critères quantitatifs sont rassemblés dans les tableaux suivants :

	Moyenne sur les tronçons de rue	Moyenne sur les intersections
Distance de Hausdorff	2,2 m	17,6 m

TAB. 5.4: distance de Hausdorff entre la rue extraite et la rue référence

Longueur	Rue extraite	Rue référence	Différence	Différence relative
Tronçons de rue	1198 m	1206 m	8 m	0,6 %
Intersections	360 m	447 m	87 m	19,4 %
Cumulée	1558 m	1653 m	95 m	5,7 %

TAB. 5.5: critère de longueur

Surface	Rue extraite	Rue référence	Différence	Différence relative
Tronçons de rue	8400 m^2	8224 m^2	176 m^2	2,1 %
Intersections	5344 m^2	4282 m^2	1062 m^2	20,7 %
Cumulée	13744 m^2	12506 m^2	1238 m^2	9,9 %

TAB. 5.6: critère de surface

	Rue extraite
Tronçons de rue	91%
Intersections	81 %

TAB. 5.7: pourcentage de la rue inclus dans la zone de tolérance

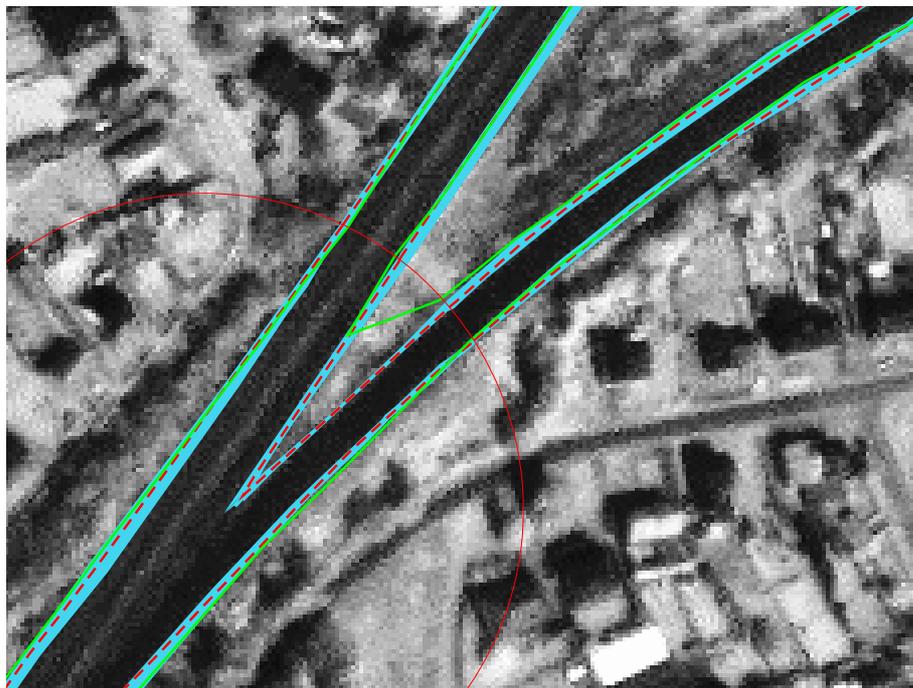


FIG. 5.15: rues extraites (trait plein vert), référence (trait pointillé rouge) et zone de tolérance (bleu)

Au niveau des tronçons de rue à bords parallèles, la moyenne sur les distances de Hausdorff est faible (2,2 mètres, 1 pixel faisant 1 mètre). Elle est de 17,6 m pour les intersections, due à la pointe mal détectée de l'intersection.

Le pourcentage d'erreur sur la longueur des tronçons de rue est proche de 0, ce qui confirme l'inspection visuelle. Elle s'élève à 19,4 % pour l'intersection, correspondant à l'intersection tronquée. Pour la comparaison des surfaces, le pourcentage d'erreur est faible pour les tronçons (2,1%), mais est de 20,7% pour l'intersection (comme pour le critère de longueur, la pointe de l'intersection est la principale contribution à cette erreur). Globalement, les critères quantitatifs confirment l'impression visuelle : il existe une bonne correspondance géométrique entre la référence et la rue extraite. Cela se traduit par un taux cumulé de 94,3% sur les longueurs et de 90,1% sur les surfaces, avec de meilleures performances pour les bords parallèles que pour les intersections.

Du point de vue de la ressemblance avec l'interprétation humaine, les pourcentages sont aussi élevés pour les tronçons à bords parallèles : plus de 90 % des contours sont inclus dans la zone de tolérance. Ce taux est de 81 % pour les intersections, la pointe faisant baisser le pourcentage. D'autre part, que cela soit pour les tronçons ou les intersections, lorsque le contour n'est pas dans la zone de tolérance, il n'en est pas très éloigné (voir l'intersection sur la figure 5.15).

Temps des saisies En moyenne, les 8 interprètes ont mis 4 minutes et demi pour saisir toute la scène. L'algorithme extrait les mêmes informations en 45 secondes (sur un PC de 2,4 GHz sous Linux), en comptant la phase manuelle d'extraction du graphe et la phase de reconstruction automatique. Le gain de temps est donc d'un facteur 6 pour l'opérateur.

5.4 Image Quickbird de la zone de Fredericton, Canada

Description de la scène La scène suivante provient du satellite Quickbird de la société DigitalGlobe. Ce satellite acquiert des images en mode panchromatique à 0,61 m. L'image présentée a été rééchantillonnée à 70 cm. Cette image a été acquise le 31 août 2003 au dessus de la ville de Fredericton au Canada (figure 5.16). Elle a été gracieusement mise à notre disposition dans le cadre du programme Géoide¹ et du projet AutoMap. La ville de Fredericton est typique des villes nord-américaines, où le réseau de rues est quadrangulaire.

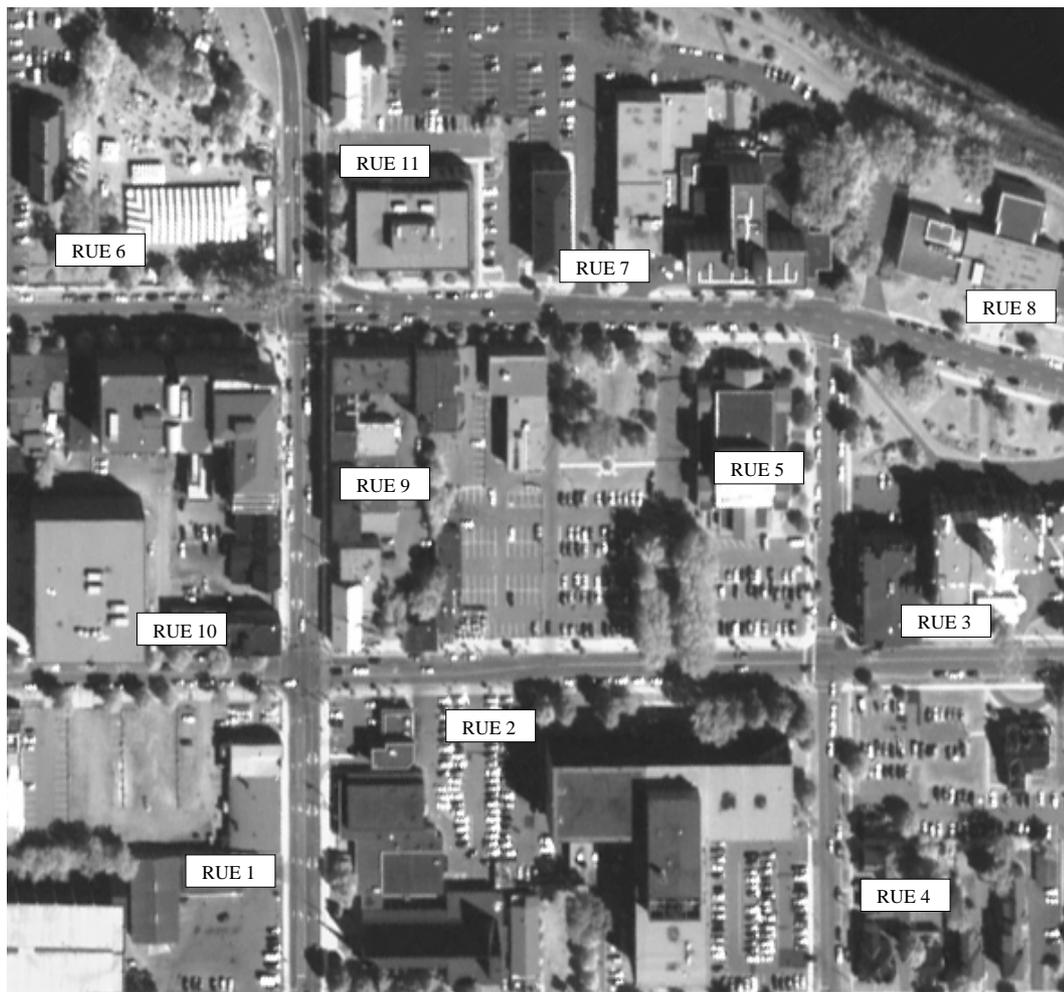


FIG. 5.16: image Quickbird du centre ville de Fredericton, Canada © DigitalGlobe

Cette scène représente un cas plus complexe que l'exemple précédent sur la zone de Hasselt. En effet, le milieu est plus urbanisé et présente un bruit plus important sur la chaussée, comme des véhicules ou le marquage au sol (cas des rues 6 et 9 notamment). Il y a aussi un grand nombre d'occlusions, principalement les arbres et les ombres portées des

¹<http://www.geoide.ulaval.ca>

bâtiments (rue 2, 3, 5 et 10).

Enfin, il y a 4 intersections, 3 en «X» et une en «Y» (entre les rues 5, 7 et 8).

Les sections suivantes présentent les différentes étapes d'extraction de ce réseau de rues. Le graphe est ici extrait par l'algorithme de suivi présenté section 3.2.3. La phase de reconstruction surfacique est ensuite présentée, les étapes intermédiaires étant moins détaillées que pour les exemples montrés précédemment.

Extraction du graphe La figure 5.17 présente le résultat du suivi de rues. Cinq points d'amorce (en vert sur la figure 5.17) ont été fournis par l'opérateur, ainsi que les directions initiales de propagation. Il résulte 5 polygones de suivi, dont certaines traversent plusieurs intersections. Sur tout le suivi, il y a eu trois cas où l'opérateur a dû corriger la trajectoire (points en jaune). Ces erreurs sont dues à l'occlusion par les ombres portées sur les rues 6 et 11, et à l'homogénéité radiométrique du bâtiment sur la rue 10.



FIG. 5.17: après l'extraction de polygones par l'algorithme de suivi

L'étape suivante va consister en l'extraction de la topologie du graphe à partir de cet ensemble de polygones obtenu et des cercles d'intersection. Comme décrit à la section 3.2.4, les polygones dont les points extrêmes sont localisés dans un cercle d'intersection vont être tronquées, et vont être raccordées au centre de ce cercle (figure 5.18). Lorsque les polygones traversent plusieurs intersections, ce qui est la majorité des cas, elles sont «découpées» en plusieurs tronçons.

Le graphe est ensuite sous-échantillonné en ne retenant qu'un point des polygones sur trois (figure 5.19).

À ce stade, le graphe du réseau est complet et topologiquement correct mais les polygones du graphe ne sont pas parfaitement centrées sur l'axe de la rue (par exemple pour les rues 2, 4 ou 11). Ce graphe va ensuite permettre d'initialiser l'étape de reconstruction surfacique.



FIG. 5.18: obtention du graphe topologique



FIG. 5.19: sous-échantillonnage du graphe topologique

Etape de reconstruction surfacique. A partir du graphe obtenu, les polygones du graphe sont échantillonnés et propagés à la distance définie par la classe des rues. Ici, les rues sont les artères principales de Fredericton et sont considérées comme de classe 3. Les *DoubleSnakes* vont donc évoluer à partir de l'échelle 5.

La figure 5.20 présente le résultat après l'extraction des tronçons de rue à bords parallèles et l'initialisation des intersections.



FIG. 5.20: initialisation des intersections après l'extraction des tronçons à bords parallèles

Enfin, la figure 5.21 présente le résultat final d'extraction.

On s'aperçoit visuellement que l'algorithme a réussi à extraire des rues mal initialisées (comme la rue 4). Ceci est dû aux termes de régularisation et à l'énergie de parallélisme. Grâce à l'utilisation des différents plans d'ondelettes, l'algorithme montre une bonne robustesse vis-à-vis du bruit sur la chaussée comme les véhicules (cas des rues 7 ou 9) ou le marquage au sol (rue 9). Il a aussi réussi à passer plusieurs occlusions causées par les arbres (rue 2) ou les ombres portées (rues 3, 6, et 7).

Certains tronçons ont été moins bien extraits : aux abords de l'intersection avec les rues

1, 10 et 9, les bords de la rue 2 s'écartent progressivement. Ceci est dû à la présence du parking et de nombreuses occlusions. Pour la rue 11, l'initialisation loin de l'axe de la rue associée avec le grand nombre d'ombres portées ont fait «dériver» le *DoubleSnake*, qui a toutefois gardé une largeur correcte.

Pour les intersections, certaines présentent des erreurs d'extraction (elles sont un peu «coupées») alors que d'autres ont été extraites correctement.



FIG. 5.21: résultat final après l'extraction des intersections

La figure 5.22 présente un zoom sur l'intersection en «X» entre les rues 5, 7 et 8. Cet exemple illustre l'introduction ou non des points de ruptures (voir l'initialisation des intersections sur la figure 5.20). Visuellement sur la figure 5.22, les intersections ont été bien extraites, et les termes de régularisation ont permis de passer les occlusions ou les zones de faible contraste. De plus, l'introduction du point de rupture a permis une bonne extraction de l'angle droit de l'intersection.



FIG. 5.22: résultat final d'extraction : zoom sur l'intersection entre les rues 5, 7 et 8

Evaluation quantitative Il a été demandé aux 8 interprètes de saisir sur cette scène le réseau de rues. Le résultat est présenté figure 5.23. La variabilité d'interprétation apparaît plus importante que dans le cas précédent pour l'image Ikonos. La scène est en effet plus complexe, et les intersections sont moins bien discernables. Certains tronçons ont engendré une grande variabilité (cas des rues 4, 6, 8 et 10) ainsi que presque toutes les intersections.

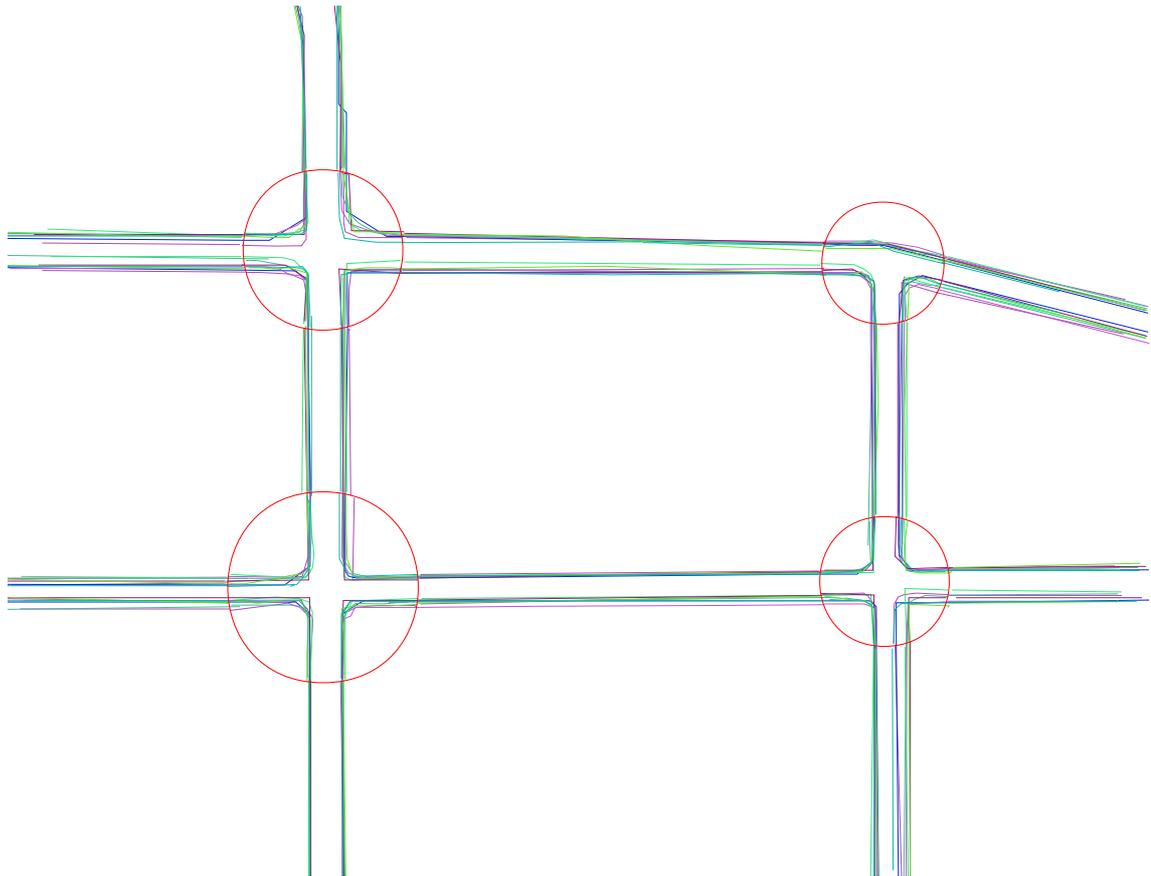


FIG. 5.23: superposition des 8 interprétations de la scène

Sur la figure 5.24 est représenté un zoom sur l'intersection entre les rues 6, 7, 9 et 11. La zone de tolérance (en bleu) et la référence (trait pointillé rouge) sont extraites à partir de ces interprétations. La zone de tolérance est bien plus large que sur le cas précédent, traduisant la différence importante de jugement entre les différents interprètes d'images. Le contour extrait par l'algorithme est représenté en trait plein vert. Cet exemple est représentatif des différents résultats que l'on obtient : à cause d'une initialisation trop éloignée et de la présence d'ombres portées, la rue 11 est décalée par rapport à la référence et est entièrement hors de la zone de tolérance. À l'inverse les rues 6 et 7 ont été très exactement extraites, sont proches de la référence et entièrement incluses dans la zone de tolérance. La rue 9 conjugue les 2 cas : le bord droit n'a pas été extrait correctement à cause de l'ombre portée trop étendue, engendrant une mauvaise initialisation de l'intersection. Le bord gauche a été extrait très précisément, et a permis une bonne initialisation de

l'intersection. On peut aussi observer figure 5.24 que certaines intersections sont un peu coupées alors que d'autres ont été extraites correctement.

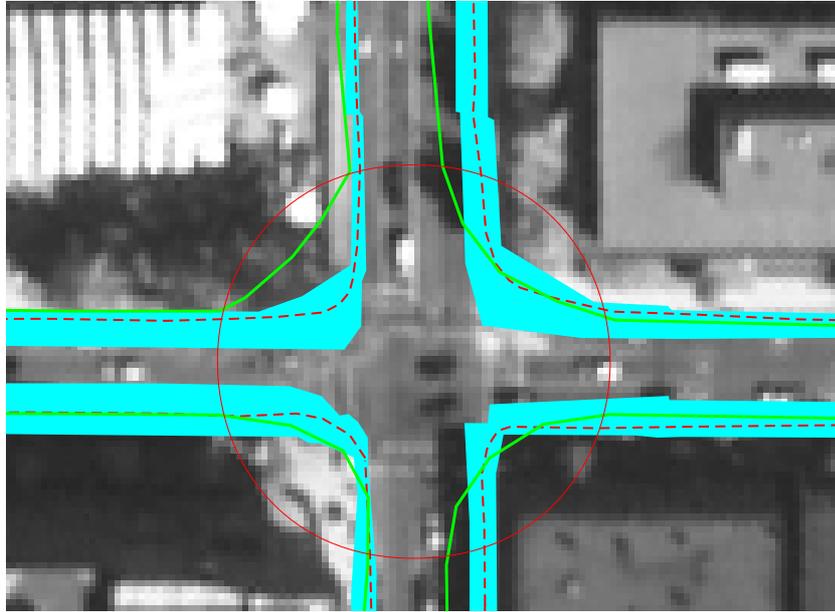


FIG. 5.24: rues extraites (trait plein vert), référence (trait pointillé) et zone de tolérance (bleue)

Les impressions visuelles sont confirmées par les valeurs numériques issues de la référence et de la zone de tolérance. Globalement les critères géométriques montrent une bonne correspondance entre le réseau extrait et la référence. La distance de Hausdorff, traduisant le maximum sur les écarts, est en moyenne de moins de 3 mètres pour les tronçons et de moins de 5 mètres pour les intersections. Les longueurs du réseau extrait et de celui de référence sont similaires (de l'ordre de 4% de différence relative). La différence sur les surfaces est faible pour les tronçons de rue (2,2%). Elle est de 16,9% sur les intersections qui ont été plus difficiles à extraire.

	Moyenne sur les tronçons de rue	Moyenne sur les intersections
Distance de Hausdorff	2,8 m	4,9 m

TAB. 5.8: distance de Hausdorff rue extraite - rue référence

Du point de vue de la zone de tolérance, plus de 70 % des bords de rue se trouvent à l'intérieur, ce qui est un bon résultat au vu de la complexité de la scène et du bruit présent. Le taux de 52,5 % pour les intersections traduit la dépendance des *IntersectionSnakes* par rapport à l'initialisation fournies par les *DoubleSnakes* et encore plus en amont par

Longueur	Rue extraite	Rue référence	Différence	Différence relative
Tronçons de rue	1906 m	1942 m	36 m	1,9 %
Intersections	581 m	650 m	69 m	11 %
Cumulée	2487 m	2592 m	105 m	4,1 %

TAB. 5.9: critère de longueur

Surface	Rue extraite	Rue référence	Différence	Différence relative
Tronçons de rue	13645 m^2	13351 m^2	294 m^2	2,2 %
Intersections	7980 m^2	6824 m^2	1156 m^2	16,9 %
Cumulée	21625 m^2	20215 m^2	1410 m^2	7,0 %

TAB. 5.10: critère de surface

l'algorithme de suivi. En effet, un décalage au niveau des points des *DoubleSnakes* situés sur le cercle (voir figure 5.24) va engendrer un décalage final sur le résultat d'extraction des intersections.

	Rue extraite
Tronçons de rue	71,3%
Intersections	52,5%

TAB. 5.11: pourcentage de la rue inclus dans la zone de tolérance

Temps de saisie. Les 8 interprètes ont mis en moyenne 22 minutes pour saisir toute la scène. Il leur a fallu sur certaines rues très occultées plusieurs ajustements. L'algorithme met lui, en comptant la phase d'extraction du graphe (semi-automatique) et la phase de reconstruction, 1 minute et 15 secondes. Cela fait plus d'un facteur 17 par rapport à un opérateur, ce qui représente un gain de temps très important.

Si le taux d'inclusion dans la zone de tolérance est plus important sur le cas de Hasselt, le gain de temps est dans ce cas complexe plus important, et les similarités géométriques entre la référence et la rue extraite restent élevées.

Enfin il est intéressant de noter que lors de la saisie, une manière naturelle pour l'opérateur est de placer une première estimation du bord de la rue à basse résolution et de le recalculer précisément à l'aide du zoom. Il s'agit là d'un principe similaire à celui que nous avons inclus dans notre méthode automatique. Les contours actifs initialisés et recalés en première estimation à la résolution la plus basse évoluent ensuite et affinent leurs positions jusqu'à la résolution initiale.

5.5 Image de la caméra numérique de l'IGN de la ville d'Amiens, France

Description de la scène La scène suivante nous a été aimablement fournie à des fins de tests par l'IGN. Elle a été prise par la caméra numérique et a une résolution spatiale de 20 cm. Il s'agit d'une vue sur la zone d'Amiens en France, comprenant une intersection en X et une intersection en Y. Toutes les rues ont été considérées de classe 3 (rues principales en ville), à l'exception de la rue 6, moins importante et donc de classe 2. Les rues de classes 3 vont démarrer la minimisation de leur fonctionnelle à l'échelle 9, tandis que la rue de classe 2 évoluera à partir de l'échelle 8.



FIG. 5.25: image de la caméra numérique de l'IGN de la ville d'Amiens, France © IGN

Ce cas est très complexe car il cumule un grand nombre de difficultés : la prise de vue est oblique (on peut voir la façade de certains bâtiments), les ombres portées des bâtiments sont très importantes et créent des changements radiométriques brutaux, et la taille des

véhicules devient importante (une trentaine de pixels).

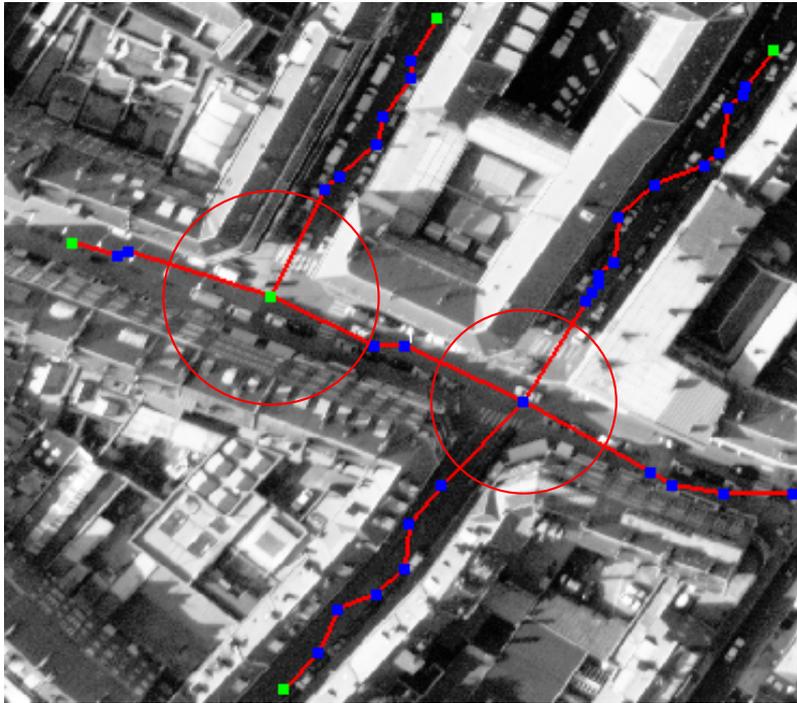
Ce cas se situe hors de nos limites d'application (résolution meilleure que 70 cm et prise de vue oblique). La campagne d'interprétation a été réalisée, mais les résultats n'étaient pas représentatifs de ce que l'on peut extraire dans l'image. En effet, la variabilité d'interprétation est dans ce cas trop extrême, car les interprètes ne sont pas familiers avec ce type d'images. Toutefois, les résultats visuels d'extraction sont présentés, car ils permettent de mettre en valeur dans ce cas les causes de réussites et les limites de la méthode. Ils permettent aussi d'entrevoir les développements futurs possibles.

Extraction du réseau. La figure 5.26 présente les polygones extraites par l'algorithme de suivi. De par la complexité de cette scène et sa résolution, l'algorithme rencontre plus de difficultés pour extraire les polygones. Il n'est en effet pas paramétré pour des rues aussi larges, ce qui occasionne des effets de «zigzags». De plus, les ombres portées ou les toits de bâtiments mettent parfois en défaut le critère d'homogénéité locale de la variance radiométrique. Toutefois, il réussit à suivre globalement bien les rues et notamment à se faufiler entre les files de voitures en stationnement (cas des rues 2 et 5).



FIG. 5.26: après l'extraction de polygones par l'algorithme de suivi

Sur la figure 5.27(a) est représenté le graphe topologique obtenu. La figure 5.27(b) représente le graphe après sous-échantillonnage. Le sous-échantillonnage a diminué l'effet de zigzag. Toutefois, les polygones ne sont pas recalés à cette étape sur l'axe central des rues.



(a) graphe topologique à partir des polygones extraites du suivi



(b) après sous-échantillonnage du graphe topologique

FIG. 5.27: obtention du graphe complet du réseau

Analyse du résultat d'extraction La figure 5.28, représente l'étape intermédiaire à laquelle les tronçons de rue à bords parallèles ont été extraits et les intersections viennent d'être initialisées.



FIG. 5.28: initialisation des intersections après l'extraction des tronçons à bords parallèles

Le résultat final de l'extraction surfacique est représenté figure 5.29. L'impression globale est que l'algorithme a réussi à extraire le réseau de rues avec une bonne cohérence sur la forme, mais que ce sont majoritairement les hauts des façades des bâtiments qui ont été extraits, plus que le rebord des trottoirs.

La première raison de ce mauvais positionnement tient aux conditions d'illumination de la scène : la lumière est rasante, ce qui fait ressortir nettement les bords des toitures des bâtiments et ombre les rues en partie ou même totalement (rues 5 et 6).

La deuxième raison provient de l'angle de prise de vue oblique. Sous cet angle, l'effet des bâtiments est double. Ils occultent le bord des rues (rues 1 et 2), ou leurs façades sont observables (rues 1, 2 et 3) et créent ainsi de nombreux artefacts parallèles à l'axe des rues.

Les deux premières raisons dérivent des conditions d'acquisition de la scène et rendent difficilement discernable le bord des rues. La troisième difficulté est due aux nombreuses files de véhicules en stationnement, qui occultent parfois tout le bord de la chaussée (par exemple sur les rues 2 et 5).



FIG. 5.29: résultat final après l'extraction des intersections

Dans ces conditions très difficiles, on peut cependant considérer que l'algorithme s'est

bien comporté. Certains bords de rue ont été bien détectés (bords inférieurs des rues 1 et 3), et l'algorithme a montré une très bonne robustesse vis-à-vis des occlusions partielles de la chaussée : véhicules au milieu de la rue (rues 1, 3 et 5), marquage au sol (au niveau des intersections) ou éclairage public (rue 1).

En privilégiant les critères géométriques (parallélisme des bords) à un modèle de profil radiométrique, l'algorithme a réussi à s'affranchir des nombreux changements de radiométrie sur la chaussée. Enfin, les contours actifs d'intersection se sont pour la plupart bien adaptés à l'angle formé par les bâtiments.

Il faut noter que le résultat va dépendre de la définition que l'on donne à une rue : si l'on considère que les trottoirs font partie de la rue, l'algorithme produit des résultats intéressants : le réseau de rues apparaît comme le complémentaire des blocs de bâtiments et on peut alors détourner les bâtiments et extraire la structure de la ville. Si la rue est considérée sans les trottoirs, ce qui est la définition commune, il reste alors des pistes de recherche à explorer. La première est la détection de files de voitures (Ruskoné 1996; Hinz et Baumgartner 2003). Elles peuvent être un bon indicateur de la présence d'un bord de rue. Toutefois, il conviendra de distinguer les voitures en stationnement près du bord de la rue (rues 2 et 5), de celles circulant ou attendant au feu rouge et donc au milieu de la chaussée.

Ce cas montre les limites des connaissances *a priori* injectées dans notre méthode. Même si l'utilisation d'informations externes ne rentrait pas dans nos objectifs, une piste prometteuse de recherche serait l'utilisation de l'information portée par les MNE. Bien que la localisation précise des bords de façades soient actuellement un point critique lors de l'établissement de MNE par stéréoscopie ou par mesures LIDAR, l'information d'élévation permettrait de faire la distinction entre un sommet de bâtiment et la chaussée voisine. On pourrait alors envisager son utilisation dans notre méthode en le considérant comme une fonction de potentiel : une fois tous les contours actifs ayant atteint un état d'équilibre, on pourrait «perturber» leurs positions en fonction de leurs altitudes. S'ils ont bien détecté le bord de la rue, ce qui correspond à un minimum local du MNE, alors leurs positions resteront inchangées. S'ils ont détecté un bord de bâtiment (maximum local du MNE), alors ils tomberont dans le minimum local du MNE correspondant au bord de la rue à proximité.

5.6 Analyse et perspectives pour l'algorithme de traitement

A travers cette série d'exemples, plusieurs remarques peuvent se dégager. Tout d'abord, les différentes phases séquentielles de l'algorithme s'enchaînent correctement. Ensuite, l'algorithme a montré une bonne robustesse vis-à-vis du bruit et des occlusions locales, que ce soit pour l'image de Hasselt ou pour des milieux plus urbains comme Fredericton ou Amiens. L'utilisation de contraintes géométrique fortes (termes régularisants, parallélisme des bords) permet de ne pas dépendre d'un modèle de profil radiométrique. L'approche multi-échelle pour l'évolution des contours actifs permet de s'affranchir du bruit et des artefacts à basse résolution et d'effectuer un recalage précis à haute résolution.

Du point de vue de l'extraction des tronçons de rue, les résultats présentent une grande similarité géométrique de longueur et de surface avec la référence issue d'interprétations de l'image. La différence relative n'est que de 0,6% pour la longueur et de 2,1% pour la surface dans le cas péri-urbain de Hasselt et de 1,9% (longueur) et 2,2% (surface) dans le cas urbain de Fredericton. De plus, les tronçons de rue extraits dans le cas péri-urbain sont très proches de l'interprétation humaine (91% dans le cas d'Hasselt). Dans un cas urbain plus bruyé sur la ville de Fredericton, 71% des contours se trouvent dans la zone de tolérance. De plus, les résultats hors zone ne sont pas trop divergents et peuvent être corrigés interactivement par l'opérateur.

D'une manière générale, il existe une meilleure performance d'extraction pour les tronçons à bords parallèles que pour les intersections. En effet, un objet spécifique, le *DoubleSnake*, a été développé permettant de s'adapter à la forme des bords parallèles, alors que la grande variabilité de forme des intersections rend difficile l'introduction de connaissances *a priori* fortes sur leur géométrie. Le modèle d'intersection que nous avons introduit est donc assez simple. Toutefois, pour améliorer les résultats d'extraction des intersections sans contraintes trop strictes, l'utilisation de *Ziplock snakes* (Neuenschwander *et al.* 1997) serait envisageable. Dans ce cas, on ne spécifie que les deux extrémités du contour actif qui optimisera ensuite sa fonctionnelle d'énergie en partant de ses deux extrémités jusqu'à son centre. Son intégration, dans notre cas, se ferait aisément car les intersections sont initialisées par deux points fixes provenant des *DoubleSnakes*. L'intégration des travaux de Boichis (2000), qui propose une modélisation des différents types d'intersection, rond-points compris, pourrait aussi être bénéfique.

Du point de vue du gain de temps par l'interprétation humaine, l'application de l'algorithme sur les différentes scènes a permis un gain en temps pouvant aller jusqu'à un facteur 17 suivant le type de scène rencontrée.

Parmi les améliorations possibles, il serait profitable d'accroître les contraintes sur les points extrêmes des *DoubleSnakes* qui ne sont liés à aucun cercle d'intersection. Ils ont en effet tendance à s'écartier. Il conviendrait de les lier aux bords de l'image et d'étendre le voisinage à plusieurs voisins pour le calcul de l'énergie de parallélisme.

Nous avons fixé le pas d'échantillonnage constant pour toutes les polygones. On peut toutefois envisager de le faire varier en fonction de la courbure locale de la polygone et surtout en fonction de la résolution de l'image : on pourrait ainsi accélérer l'algorithme en sous-échantillonnant les contours actifs à basse résolution et en augmentant leur nombre

de points à haute résolution.

Il peut être intéressant d'extraire des caractéristiques des rues comme leur terre-plein central ou les terre-pleins secondaires. On peut envisager dans ce cas l'établissement de modèles ne comportant plus deux brins comme pour les *DoubleSnakes*, mais 3 ou 5 : tous les brins évolueront eux aussi de manière liée pour garder un parallélisme local entre eux, tout en se recalant sur la caractéristique de la rue qu'ils doivent extraire.

Enfin, des tests sur un grand ensemble de scènes et de résolutions spatiales différentes permettraient de donner plus d'informations sur les limites réelles de la méthode.

Conclusion et perspectives

La disponibilité d'images satellites à très haute résolution spatiale au dessus de zones urbaines est récente. Elle constitue potentiellement un très grand apport dans la perspective de cartographier les zones urbaines à des échelles de l'ordre du 1/10 000. A ces échelles, la rue passe d'une représentation linéique où seul l'axe de la rue est localisé avec précision, à une représentation surfacique où sa largeur en représentation cartographique correspond à son emprise réelle au sol. La très haute résolution spatiale permet donc une localisation plus précise des bords de la rue, mais engendre en contrepartie une augmentation significative du bruit.

Dans cette thèse, nous avons proposé une méthode globale d'extraction des réseaux de rues en milieu urbain à partir des images à très haute résolution spatiale. Son objectif est de répondre à la forte demande dans la création et/ou la mise à jour automatisée de cartes.

Conscient des difficultés liées au coût et au traitement de sources d'information externes comme les modèles numériques d'élévation ou l'utilisation de plusieurs canaux spectraux dans un contexte opérationnel, nous avons choisi de prendre l'image panchromatique comme unique donnée.

L'état de l'art a montré le peu de travaux effectués en milieu urbain pour une extraction surfacique de la rue. Il a confirmé qu'une méthode semi-automatique au niveau de l'amorce devait être retenue dans notre contexte bruité afin de garantir la fiabilité des résultats. Il a aussi permis de mettre en valeur le problème du bruit et de dégager les techniques permettant de le réduire.

Dans la méthode que nous proposons, l'approche multi-échelle et des contraintes géométriques fortes ont été utilisées pour définir des contours actifs multi-échelles permettant d'extraire les bords parallèles des rues. L'utilisation des coefficients de la transformée en ondelettes de l'image à différentes échelles, ainsi que des contraintes géométriques comme le parallélisme, permet au modèle développé, le *DoubleSnake*, d'extraire les tronçons de rues à bords parallèles dans un environnement bruité.

Nous avons aussi exploité la coopération entre la représentation linéique de la rue et sa représentation surfacique. Le graphe topologique du réseau est d'abord extrait et est utilisé pour initialiser l'étape de reconstruction surfacique. Le résultat d'extraction peut alors servir à recaler le graphe précisément sur l'axe des rues.

Enfin, la méthode est modulaire, et permet donc d'incorporer de manière souple de nouvelles techniques à différents niveaux.

La méthode a été appliquée à différentes scènes. Elle a été évaluée quantitativement par comparaison à une référence et à une zone de tolérance, définies à partir de plusieurs interprétations humaines. Dans la limite des cas observés, la méthode donne de bons résultats. L'algorithme montre une meilleure performance pour l'extraction des sections de rues à bords parallèles que pour celle des intersections.

Dans le cas des rues à bords parallèles, on observe une grande similarité géométrique entre les longueurs et les surfaces du réseau extrait et de la référence. Sur les cas évalués quantitativement, entre 70% et 90% du réseau extrait est situé dans la zone de tolérance. Les valeurs sont moins élevées pour les intersections (par exemple le taux d'inclusion dans la tolérance varie de 50% et 80%). Ceci s'explique à la fois par la dépendance des contours actifs d'intersection par rapport à leur initialisation mais aussi par le fait que, contrairement aux sections à bords parallèles, nous n'avons pas développé pour eux de modèle spécifique.

Plusieurs plages de résolutions ont été considérées ainsi que différents types de scènes (péri-urbaine, urbaine, différentes typologies de réseaux). Ces exemples ont permis de mettre en évidence le bon comportement de l'algorithme vis-à-vis du bruit, quels que soient la résolution de l'image et le degré d'urbanisation de la scène étudiée, ce qui tend à confirmer le caractère générique de la méthode.

La méthode a été implémentée par des algorithmes rapides de transformée en ondelettes et de contours actifs, ce qui permet d'augmenter la rapidité globale de l'algorithme. Cette méthode semble donc répondre aux objectifs fixés : comparé à une extraction manuelle, l'algorithme fait baisser le temps d'extraction jusqu'à un facteur 17 suivant la complexité de la scène. Toutefois, ce gain doit être diminué du temps de rectification de certains mauvais positionnements des rues. Le gain de temps peut être ré-estimé comme un facteur 10 approximativement, ce qui reste très intéressant dans une optique opérationnelle.

Enfin, certaines limites de la méthode sont apparues lors de son application sur une scène à 20 cm de résolution, avec des conditions de prise de vue difficiles (vue et illumination obliques). Ces conditions constituent en effet les limites des connaissances *a priori* introduites dans la méthode. Notamment, l'algorithme a tendance à extraire les hauts de bâtiments plutôt que la chaussée des rues. Plus généralement, plusieurs points permettront d'augmenter dans le futur les performances et l'automatisation de la méthode. Ces améliorations ne doivent cependant pas nuire à la fiabilité de l'algorithme.

En premier lieu il conviendrait de faire une étude de la sensibilité aux paramètres de la méthode (pondération des énergies des contours actifs, choix de la classe des rues). Cette étude pourrait permettre, en considérant un grand nombre de cas (différentes typologies de réseau ou différentes résolutions d'image), d'extraire des paramètres statistiquement représentatifs. Même si l'automatisation totale paraît difficile, une piste prometteuse serait de calculer un indice de texture représentatif du tissu urbain. Cet indice permettrait d'adapter les paramètres d'extraction en fonction de la typologie du tissu urbain (quartiers résidentiels périphériques, centre ville).

Un des avantages de notre méthode est de n'utiliser que très peu de sources d'information. Cependant, sa modularité permet l'incorporation de données externes. Par exemple, l'information d'élévation permettrait de lever certaines incertitudes. On pourrait envisager son utilisation à la fin de l'extraction pour distinguer le haut des bâtiments du bord de la chaussée. L'utilisation de plusieurs canaux spectraux pourrait aussi améliorer les résultats d'extraction. Ainsi, l'indice de végétation peut permettre de masquer les arbres le long des rues et restreindre ainsi la zone d'extraction. Les images thermiques peuvent fournir des indications sur les véhicules en mouvement circulant sur la chaussée, et donc sur la présence de la rue. Il convient de veiller toutefois à ce que l'introduction de ces données externes n'augmente pas la complexité et le temps de traitement de manière trop importante, et que le gain de précision et de fiabilité soit en accord avec les besoins opérationnels.

L'utilisation du contexte local de la rue n'a, pour l'instant, pas été exploitée. Les files de voitures, si elles sont bien interprétées comme étant en stationnement et non en circulation, peuvent être de bons indices quant à la présence du bord de la rue. Il en est de même pour les alignements de bâtiments qui traduisent souvent la proximité d'une rue. Le contexte doit être considéré comme un ensemble d'indices sur la présence d'une rue, avec une certaine probabilité d'apparition, et non comme étant nécessaire à l'extraction.

Cette notion de contexte pourrait permettre à l'algorithme de s'auto-estimer et ainsi de fournir des indicateurs de confiance pour l'opérateur. De manière à ne pas créer de biais sur les indicateurs, ils doivent se servir d'informations qui n'ont pas été utilisées par l'algorithme. Ces indicateurs peuvent alors servir à focaliser l'attention de l'opérateur dans les zones où l'extraction est estimée par l'algorithme comme non fiable, et ainsi aider l'opérateur à rectifier le résultat de l'extraction.

Enfin, la méthode proposée dans cette thèse est semi-automatique, comme la majorité des approches opérationnelles présentées dans la littérature scientifique.

Bien entendu, le souhait serait que les systèmes de cartographie soient tout-automatiques. Toutefois, les obstacles sont nombreux, principalement au stade de la détection des rues. L'état actuel des techniques ne permet pas d'envisager cette automatisation totale avant de nombreuses années.

Peut-être qu'une meilleure connaissance et une meilleure modélisation de cet extraordinaire système de perception qu'est la vision humaine permettront d'atteindre une interprétation tout-automatique et fiable. Comprendre les mécanismes de l'interprétation humaine et pouvoir les reproduire dans un système automatisé impliquent l'interaction entre de nombreux domaines de la science, et représentent des défis majeurs pour l'avenir.

Table des figures

0.1	exemple d'images à très haute résolution spatiale : l'image 0.1(a) provient de la caméra IGN et a été acquise à la résolution de 20 cm. L'image 0.1(b) provient du satellite Quickbird de résolution spatiale 70 cm en mode pan-chromatique.	7
0.2	exemple d'extraction de réseaux de rues linéiques en milieu urbain sur une image à résolution spatiale moyenne. © Wang et al. (1996).	9
0.3	deux exemples de représentation cartographie des rues © IGN-Belgique . .	10
0.4	occlusions et artefacts inhérents au contexte urbain. Les véhicules sont entourés.	10
0.5	effet de l'angle de prise de vue sur la visibilité des rues	11
1.1	représentation de milieux ruraux (a), péri-urbains (b) et urbains (c) dans la région de Hasselt (Belgique). © 2000 Space Imaging Europe. Image Ikonos aimablement fournie par la société Geographic Information Management . .	14
1.2	les deux modèles de profils radiométriques «idéaux» de la rue	17
2.1	procédure d'optimisation	36
2.2	schéma des bancs de filtres pour l'algorithme à trous	39
2.3	ondelette et fonction d'échelle spline quadratique (d'après Mallat 1997) . . .	40
2.4	composantes et module de la transformée en ondelettes à différentes échelles	43
2.5	évaluation du parallélisme des deux brins lors de l'extraction des contours parallèles	46
2.6	effet de l'énergie de parallélisme sur un <i>DoubleSnake</i>	47
2.7	influence de l'approche multi-échelle sur les résultats	48
2.8	nombre de déplacements de points en fonction de l'échelle. $j = 0$ correspond à l'image à la résolution initiale.	49
2.9	test sur une image Quickbird (0,7 m de résolution spatiale) © DigitalGlobe	50
3.1	méthode d'extraction globale	54
3.2	polylignes constituant le graphe du réseau	57
3.3	arbre de recherche pour l'algorithme de suivi.	58
3.4	diagramme d'extraction des polylignes	60
3.5	résultat d'un suivi de rue	60
3.6	polylignes extraites par l'algorithme de suivi de rues. Les points verts correspondent aux amorces fournies par l'opérateur, les points bleus ont été extraits par l'algorithme de suivi.	61
3.7	obtention du graphe complet du réseau	63

3.8	procédé d'initialisation des <i>DoubleSnakes</i> par les polygones	65
3.9	initialisation des <i>DoubleSnakes</i> aux intersections	66
3.10	initialisation des contours actifs d'intersection	67
3.11	algorithme de reconstruction surfacique	69
4.1	référence et évaluation de la méthode globale	73
4.2	extraction de la référence et de la zone de tolérance	77
4.3	extraction de la référence à partir des interprétations d'un bord de l'objet	78
4.4	détermination de la zone de tolérance à partir des interprétations	80
4.5	frontière «gauche» de la zone de tolérance obtenue par triangulation de Delaunay contrainte.	81
4.6	image originale de la rue à acquérir	82
4.7	représentation en intensité radiométrique de la figure 4.6 et exemple de transects de la rue	83
4.8	superposition des 11 acquisitions	84
4.9	superposition des 11 acquisitions (zoom sur le bord droit de la rue 4.8)	85
4.10	extraction du contour de référence (bord droit de la rue). La ligne en tirets correspond au contour de référence obtenu.	86
4.11	profil radiométrique du transect T_6 (bord droit)	86
4.12	zone de tolérance pour le bord droit de la rue calculée à partir d'une trian- gulation de Delaunay contrainte	87
5.1	graphe du réseau et cercle d'intersection	93
5.2	après propagation des contours parallèles	93
5.3	après extraction des contours parallèles	94
5.4	initialisation de l'intersection	94
5.5	après extraction de l'intersection	95
5.6	recalage du graphe du réseau après l'étape de reconstruction	96
5.7	image Ikonos (région de Hasselt, Belgique) © <i>SpaceImaging</i>	98
5.8	module de la transformée en ondelettes à différentes échelles. Le marquage au sol central a disparu à l'échelle 3 sous l'effet de lissage.	99
5.9	image originale avec le graphe	100
5.10	après propagation des polygones	101
5.11	après traitement des contours parallèles	101
5.12	après traitement des intersections	102
5.13	superposition des 8 interprétations de la scène	103
5.14	superposition des 8 interprétations de la scène (zoom)	104
5.15	rues extraites (trait plein vert), référence (trait pointillé rouge) et zone de tolérance (bleu)	105
5.16	image Quickbird du centre ville de Fredericton, Canada © <i>DigitalGlobe</i>	107
5.17	après l'extraction de polygones par l'algorithme de suivi	108
5.18	obtention du graphe topologique	109
5.19	sous-échantillonnage du graphe topologique	109
5.20	initialisation des intersections après l'extraction des tronçons à bords parallèles	110
5.21	résultat final après l'extraction des intersections	111
5.22	résultat final d'extraction : zoom sur l'intersection entre les rues 5, 7 et 8	112

5.23	superposition des 8 interprétations de la scène	113
5.24	rues extraites (trait plein vert), référence (trait pointillé) et zone de tolérance (bleue)	114
5.25	image de la caméra numérique de l'IGN de la ville d'Amiens, France © <i>IGN</i>	116
5.26	après l'extraction de polygones par l'algorithme de suivi	117
5.27	obtention du graphe complet du réseau	118
5.28	initialisation des intersections après l'extraction des tronçons à bords parallèles	119
5.29	résultat final après l'extraction des intersections	120
A.1	représentation des approximations successives d'une même image à l'aide d'un algorithme pyramidal	145
C.1	Greedy Algorithm pseudo-code	151

Liste des tableaux

0.1	les satellites à très haute résolution spatiale et leurs résolutions	6
2.1	coefficients des filtres pour les fonctions d'échelle et d'ondelette spline quadratique (Mallat 1997)	41
2.2	écart type sur les vecteurs liés à l'énergie de parallélisme	47
3.1	tableau des classes de largeur	66
3.2	niveau de décomposition en fonction de la classe de rue considéré pour différents satellites	70
5.1	distance de Hausdorff rue extraite - rue référence (en pixel)	97
5.2	critère de longueur (en pixel)	97
5.3	critère de surface (en pixel ²)	97
5.4	distance de Hausdorff entre la rue extraite et la rue référence	104
5.5	critère de longueur	104
5.6	critère de surface	105
5.7	pourcentage de la rue inclus dans la zone de tolérance	105
5.8	distance de Hausdorff rue extraite - rue référence	114
5.9	critère de longueur	115
5.10	critère de surface	115
5.11	pourcentage de la rue inclus dans la zone de tolérance	115

Références

- Abbas, I. (1994). *Bases de données vectorielles et erreur cartographique : Problème posé par le contrôle ponctuel. Une méthode alternative fondée sur la distance de Hausdorff : le contrôle linéaire*. Thèse de doctorat, Université Paris 7, Ecole Nationale des Sciences Géographiques, Paris, France.
- Agouris, P., S. Gyftakis, et A. Stefanidis (2001, June). Dynamic node distribution in adaptive snakes for road extraction. In *Vision Interface 2001*, Ottawa, Canada, pp. pp. 134 – 140.
- Airault, S. et O. Jamet (1995). Détection et restitution automatique du réseau routier sur des images aériennes. *Traitement du Signal* 12(2), 189–200.
- Amini, A., S. Tehrani, et T. Weymouth (1988). Using dynamic programming for minimizing the energy of active contours in the presence of hard constraints. In I. C. S. Press (Ed.), *Second International Conference on Computer Vision*, pp. 95–99. IEEE Computer Society.
- Bajcsy, R. et M. Tavakoli (1976, September). Computer recognition of roads from satellite pictures. *Systems, Man and Cybernetics* 6(9), 623–637.
- Baumgartner, A. (1998, July). Extraction of roads from aerial imagery based on grouping and local context. In *Proc. ISPRS, Commission III, Working Group 3*, Columbus, Ohio, USA. <http://wwwphoto.eng.ohio-state.edu/isprs3/FinalPro.htm>.
- Baumgartner, A., C. Steger, H. Mayer, W. Eckstein, et E. Heinrich (1999). Automatic road extraction based on multi-scale, grouping, and context. *Photogrammetric Engineering and Remote Sensing* 65(7), 777–785.
- Baumgartner, A., C. Steger, C. Wiedermann, H. Mayer, W. Eckstein, et H. Hebner (1996). Verification and multiresolution extraction. *Photogrammetric Engineering and Remote Sensing* 31(3), 53–58.
- Bhattacharya, U. et S. Parui (1997). An improved backpropagation neural network for detection of road-like features in satellite imagery. *International Journal of Remote Sensing* 18(16), 3379–3394.
- Bijaoui, A., J. Starck, et F. Murtagh (1994). Restauration des images multi-échelles par l’algorithme à trous. *Traitement du Signal* 11(3), 229–243.
- Boichis, N. (2000). *Extraction automatique des carrefours routiers par interprétation d’images aérienne guidée par une base de données cartographiques*. Thèse de doctorat, Université de Cergy-Pontoise, Cergy-Pontoise, France. 142 pages.
- Bordes, G. (1997). *Interprétation d’images aériennes guidée par une base de données cartographiques : application à l’extraction automatique des routes*. Thèse de doctorat, Université de Marne-la-Vallée, Marne-la-Vallée, France. 271 pages.

- Bourennane, E., M. Paindavoine, et F. Truchetet (1993). Amélioration du filtre de Canny-Deriche pour la détection des contours sous forme de rampe. *Traitement du Signal et des Images* 10(4), pp. 297–310.
- Brigger, P. et M. Unser (1998, July 19-24,). Multi-scale B-spline snakes for general contour detection. In *Proceedings of the SPIE Conference on Mathematical Imaging : Wavelet Applications in Signal and Image Processing VI*, Volume 3458, San Diego CA, USA, pp. 92–102.
- Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8(6), 679–698.
- Centre d'Études des Transports URbains (1988). *Guide général de la voirie urbaine : Conception, Aménagement, Exploitation*. 8 avenue Aristide Briand 92220 Bagneux, France : Centre d'Études des Transports URbains.
- Cocquerez, J. et S. Philipp (1995). *Analyse d'images : filtrage et segmentation*. Enseignement de la physique. Paris, France : éditions Masson. 457 pages.
- Couloigner, I. (1998). *Reconnaissance de formes dans des images de télédétection du milieu urbain*. Thèse de doctorat, Université de Nice Sophia-Antipolis, Nice, France. 104 pages.
- Couloigner, I. et T. Ranchin (1998, May). Extraction of urban network from high spatial resolution satellite imagery : assessment of the quality of an automated method. In Balkema (Ed.), *Proceedings of the EARSeL Symposium "Operational Remote Sensing for Sustainable Development"*, Enschede, The Netherlands, pp. 309–314.
- Couloigner, I. et T. Ranchin (2000). Mapping of urban areas : A multiresolution modeling approach for semi-automatic extraction of streets. *Photogrammetric Engineering and Remote Sensing* 66(7), 867–874.
- Couloigner, I., T. Ranchin, V. P. Valtonen, et L. Wald (1997). Apport de spot 5 à la cartographie urbaine. *Bulletin de la Société Française de Photogrammétrie et de Télédétection* 145, 33–38.
- De Joinville, O. (2001). *Évaluation de la qualité d'une cartographie urbaine à l'aide d'images aériennes à haute résolution*. Thèse de doctorat, Ecole Nationale Supérieure des Télécommunications, Paris, France. 185 pages.
- Deriche, R. (1987). Using Canny's criteria to derive a recursively implemented optimal edge detector. *The International Journal of Computer Vision* 1(2), 167–187.
- Destival, I. (1987). Recherche automatique de réseaux linéaires sur des images spot. *Bulletin de la Société Française de Photogrammétrie et de Télédétection* 66, 5–16.
- Doucette, P., P. Agouris, A. Stefanidis, et M. Musavi (2001). Selforganized clustering for road extraction in classified imagery. *ISPRS Journal of Photogrammetry & Remote Sensing* 55(5-6), pp. 347–358.
- Eidenbenz, C., C. Käser, et E. P. Baltsavias (2000). Atomi - automated reconstruction of topographic objects from aerial images using vectorised map information. In IAPRS (Ed.), *19th ISPRS Congress*, Volume 33, Amsterdam, The Netherlands, pp. 462–471.
- Fram, J. R. et E. S. Deutsch (1975). On the quantitative evaluation of edge detection schemes and their comparison with human performance. *IEEE Transactions on Computer* 24(6), 616–628.
- Fua, P. et Y. G. Leclerc (1990). Model driven edge detection. *Machine Vision and Applications* 3, 45–56.

- Garnesson, Ph. (1991). *MESSIE : un système d'analyse de scène - application à l'imagerie aérienne*. Thèse de doctorat, Université de Nice Sophia-Antipolis, Nice, France. 225 pages.
- GDTA (1995). Télédétection, urbanisme et aménagement. Cahier A4. Groupement pour le Développement de la Télédétection Aérospatiale, Toulouse, France. 151 pages.
- Gruen, A. et H. Li (1995). Road extraction from aerial and satellite images by dynamic programming. *ISPRS Journal of Photogrammetry and Remote Sensing* 50(4), 11–20.
- Gruen, A. et H. Li (1997). Semi-automatic linear feature extraction by dynamic programming and lsb-snakes. *Photogrammetric Engineering and Remote Sensing* 63(8), 985–995.
- Guigues, L. et J.-M. Viglino (2000). Automatic road extraction through light propagation simulation. In *International Archive of Photogrammetry and Remote Sensing*, Volume XXXIII, Amsterdam, Holland.
- Harvey, W. (1999, Avril). Performance evaluation for road extraction. In *Proceedings of the International Workshop on 3D Geospatial Data Production*, Paris, France. ISPRS.
- Haverkamp, D. (2002). Extracting straight road structure in urban environments using IKONOS satellite imagery. *Optical Engineering* 41(9), 2107–2110.
- Heipke, C., H. Mayer, C. Wiedemann, et O. Jamet (1997). Evaluation of automatic road extraction. In *In International Archives of Photogrammetry and Remote Sensing, volume (32) 3-2W3, pages 47–56, 1997*.
- Heipke, C., C. Steger, et R. Multhammer (1995). A hierarchical approach to automatic road extraction from aerial imagery. *SPIE* 2486, 222–231.
- Hinz, S. et A. Baumgartner (2003). Automatic extraction of urban road networks from multi-view aerial imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* 58, pp. 83–98.
- Hinz, S., A. Baumgartner, et E. Heinrich (2001, November). Modeling contextual knowledge for controlling road extraction in urban areas. In *Proceedings of the IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*, Rome, Italy.
- Holschneider, M., R. Kronland-Martinet, J. Morlet, et P. Tchamitchian (1989). A real time algorithm for signal analysis with the help of the wavelet transform. In *Wavelets, Time-Frequency Methods and Phase Space*, pp. pages 286–297. Springer-Verlag. Berlin, Allemagne.
- Jedynak, B. (1995). *Modèles stochastiques et méthodes déterministes pour extraire les routes des images de la terre vue du ciel*. Thèse de doctorat, Université de Paris Sud, Paris, France. 186 pages.
- Kass, M., A. Witkin, et D. Terzopoulos (1987). Snakes : Active contour models. In *Proc. of IEEE Conference on Computer Vision*, London, England, pp. 259–268.
- Khatir, L. (1988). *Recherches d'algorithmes de localisation de routes dans les images satellites à haute résolution - Application à l'imagerie SPOT*. Thèse de doctorat, Université Paul Sabatier, Toulouse, France. 103 pages.
- Laporterie, F., C. Latry, H. De Boissezon, et M.-J. Lefèvre (2003, 22-23 May). Evaluation of the quality of panchromatic/multispectral fusion algorithms performed on images simulating the future pleiades satellites. In *Proceedings of the 2nd IEEE/ISPRS Joint*

- Workshop on Remote Sensing and Data Fusion over Urban Areas*, Berlin, Germany, pp. 23–31.
- Laptev, I., H. Mayer, T. Lindeberg, W. Eckstein, C. Steger, et A. Baumgartner (2000). Automatic extraction of roads from aerial images based on scale space and snakes. *Machine Vision and Applications* 12, 23–31.
- Li, D. et J. Shao (1994). The wavelet and its application in image edge detection. *ISPRS Journal of Photogrammetry and Remote Sensing* 49(3), pages 4–11.
- Mahmoodi, S., B. Sharif, et E. Chester (1997). Contour detection using multi-scale active shape models. In *Proceedings of the IEEE International Conference on Image Processing*, Volume 2, Santa Barbara, CA, USA, pp. 708–711.
- Mallat, S. (1989). A theory for multiresolution signal decomposition : the wavelet representation. *IEEE Trans. Pattern Anal. Machine Intell.* 11(7), 674–693.
- Mallat, S. (1997). *A Wavelet Tour of Signal Processing*. London, UK. : AP Professional. 577 pages.
- Mallat, S. et S. Zhong (1992, July). Characterization of Signals from Multiscale Edges. *IEEE Trans. Pattern Anal. Machine Intell.* 40(7), 2464–2482.
- Mayer, H. et E. Baltsavias (2002). Proposed OEEPE project "Automated extraction, refinement, and update of road databases from imagery and other data". OEEPE Newsletter. Volume 2, pp 22–24.
- Mc Keown, D. et J. Denlinger (1988). Cooperative methods for road tracking in aerial imagery. In *Proceedings of IEEE Computer Society Conference - Computer vision and pattern recognition*, Ann Harbor, Michigan, USA, pp. 662–672.
- Mériaudeau, F., O. Laligant, et F. Truchetet (1995, September). Etude comparative de differents detecteurs de contours et segmentation au sens contours par frames multiechelles. In *Actes du colloque GRETSI*, Juan les pins, France, pp. pp. 617–620. to appear.
- Merlet, N. et J. Zérubia (1996). New prospects in line detection by dynamic programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18(4), 426–430.
- Meyer, Y. (1992). *Wavelets. Algorithms and Applications*. Philadelphia, PA, USA : Society for Industrial and Applied Mathematics (S.I.A.M).
- Neuenschwander, W., P. Fua, G. Szekely, et O. Kubler (1997, December). Ziplock snakes. *International Journal of Computer Vision* 26(3), 191–201.
- Nevatia, R. et K. Babu (1980). Linear feature extraction and description. *Computer Graphics and Image Processing* 13, pp. 257–269.
- Péteri, R., J. Celle, et T. Ranchin (2003, Mai). Contours actifs et analyse multirésolution pour l'extraction de réseaux de rues à partir d'images satellites. In *Congrès Francophone de Vision par Ordinateur ORASIS'03*, Gérardmer, France, pp. 393–401.
- Péteri, R., I. Couloigner, et T. Ranchin (2001). A multiresolution modelling approach for semi-automatic extraction of streets : application to high resolution images from the IKONOS satellite. In G. Bégni (Ed.), *Proceedings of the EARSeL/SFPT Symposium «New Solutions for a New Millenium»*, Marne-la-Vallée, France, pp. pp. 327–332. Swets & Zeitlinger, Lisse, the Netherlands. ISBN 90-5809-254-2.

- Péteri, R., I. Couloigner, et T. Ranchin (2003). How to assess quantitatively road extracted from high resolution imagery? *Photogrammetric Engineering & Remote Sensing*. to appear.
- Péteri, R. et T. Ranchin (2003, 2-5 June). Multiresolution Snakes for urban road extraction from IKONOS and Quickbird images. In *23rd EARSeL Annual Symposium "Remote Sensing in Transition"*, Ghent, Belgium. to appear.
- Pierrot-Deseilligny, M., H. Le Men, et G. Stamon (1993). Map understanding for GIS data capture : algorithms for road network graph reconstruction. In *ICDAR*, Tsukuba, Japon, pp. 676–679.
- Price, K. (2000). Urban street grid description and verification. In *Workshop on the Application of Computer Vision (WACV2000)*, Palm Springs, California, USA, pp. 148–154.
- Puissant, A. et Ch. Weber (2002). The utility of very high spatial resolution images to identify urban objects. *Geocarto International* 17(1), 31–41.
- Ranchin, T. (1997). Wavelets, remote sensing and environmental modelling. In A. Sydow (Ed.), *Proceedings of the 15th IMACS World Congress on Scientific Computation, Modelling, and Applied Mathematics*, Volume 6, Berlin, Germany, pp. 27–34.
- Ranchin, T. (2000). *Système d'Information Géographique*. Support de cours de l'Institut Géographique du Mali. Ecole des Mines de Paris, Sophia Antipolis, France.
- Ranchin, T., B. Aiazzi, L. Alparone, S. Baronti, et L. Wald (2003). Image fusion. The ARSIS concept and some successful implementation schemes. *ISPRS Journal of Photogrammetry & Remote Sensing* 58, 4–18.
- Rellier, G., X. Descombes, et J. Zerubia (2000). Local registration and deformation of a road cartographic database on a SPOT satellite image. Rapport technique 3939, INRIA, Sophia Antipolis, France. 23 pages.
- Ruskoné, R. (1996). *Extraction automatique du réseau routier par interprétation locale du contexte : application à la production de données cartographiques*. Thèse de doctorat, Université de Marne-la-Vallée, Marne-la-Vallée, France. 157 pages.
- Serendero, M. (1989). *Extraction d'informations symboliques en imagerie SPOT : réseaux de communications et agglomérations*. Thèse de doctorat, Université de Nice Sophia-Antipolis, Nice, France. 160 pages.
- Shaoh, Z., D. Li, et Q. Cheng (2003, July). A topological 3D reconstruction of complicated buildings and crossroads. In *Proc. IEEE International Geoscience and Remote Sensing Symposium*, Volume I, Toulouse, France, pp. 56–58.
- Shensa, M. (1992). Discrete wavelet transforms : Wedding the à trous and mallat algorithms. *IEEE Trans. on Signal Process.* 40(100), pp 2464–2482.
- Shi, W. et C. Zhu (2002, February). The line segment match method for extracting road network from high-resolution satellite images. *IEEE. Trans. Geoscience and Remote Sensing* 40(2), 511–514.
- Souchon, J.-P. et C. Thom (2003). Un canal proche infrarouge pour la caméra numérique de l'IGN. In *Bulletin d'information de l'IGN*, Volume 74.
- Squire, D. et T. Pun (1997, June). A comparison of human and machine assessments of image similarity for the organization of image databases. In *Proceedings of the 10th Scandinavian Conference on Image Analysis*, Lappeenranta, Finland, pp. 51–58.

- Starck, J. (1992). *Analyse en Ondelettes et Imagerie à Haute Résolution Angulaire*. Thèse de doctorat, Université de Nice Sophia-Antipolis, Nice, France. 194 pages.
- Steger, C. (1998, Feb.). An unbiased detector of curvilinear structures. *IEEE Trans. Pattern Anal. Machine Intell.* 20, 113–125.
- Stoica, R. (2000). *Processus ponctuels pour l'extraction de réseaux linéiques dans les images satellitaires et aériennes*. Thèse de doctorat, Université de Nice Sophia-Antipolis, Nice, France. 154 pages.
- Véran, J. (1993). Suivi de routes dans une images aérienne par filtrage de Kalman. Rapport technique, Ecole Nationale Supérieure des Télécommunications, Paris, France. 30 pages.
- Vosselman, G. et J. De Knech (1995). Road tracking by profile matching and Kalman filter. In B. Verlag (Ed.), *Proceedings of Automatic extraction of man-made objects from aerial and space images*, pp. 265–274.
- Wald, L. (2002). *Data Fusion : Definitions and Architectures - Fusion of images of different spatial resolutions*. Les Presses de l'Ecole des Mines, Paris, France. 198 pages.
- Wang, D., D. He, L. Wang, et D. Morin (1996). Extraction du réseau routier à partir d'images SPOT HRV. *International Journal of Remote Sensing* 17(4), 827–833.
- Wang, J. et P. Howarth (1987). Automated road network extraction from Landsat TM imagery. In *Proceedings of the annual ASPRS/ACSM Convention1*, Volume 1, Baltimore, MD, USA, pp. 429–438.
- Wang, Y. et J. Trinder (1998, July). Use of topology in automatic road extraction. In *Proc. ISPRS, Commission III, Working Group 4*, Columbus, Ohio, USA.
- Wang, Y. et J. Trinder (2003). Knowledge-based road extraction from aerial images. Rapport technique, School of geomatic engineering, University of New South Wales, Sydney, Australia. 178 pages.
- Weber, Ch. (1995). *Images satellitaires et milieu urbain*. Collection géomatique, Paris : Hermès. 185 pages.
- Wiedemann, C., C. Heipke, H. Mayer, et S. Hinz (1998, July). Automatic extraction and evaluation of road networks from moms-2p imagery. In *Proc. ISPRS, Commission III, Working Group 3*, Columbus, Ohio, USA. <http://wwwphoto.eng.ohio-state.edu/isprs3/FinalPro.htm>.
- Williams, D. et M. Shah (1992, January). A fast algorithm for active contours and curvature estimation. In *CVIP : Image Understanding*, Volume 55, pp. 14–26.
- Xiaoying, J. et C. Davis (2003, July). Automatic road extraction from high-resolution multispectral IKONOS imagery. In *Proc. IEEE International Geoscience and Remote Sensing Symposium*, Volume III, Toulouse, France, pp. 1730–1732.
- Zhang, C., E. Baltsavias, et A. Gruen (2000, December). Knowledge-based image analysis for 3D road reconstruction. In *Proceedings of the 21st Asian Conference on Remote Sensing*, Taipei, Taiwan, pp. 100–105.

Publications

Revues internationales

- [1] R. Péteri, I. Couloigner et T. Ranchin, 2002– How to assess quantitatively road extracted from high resolution imagery?– Accepté pour publication dans *Photogrammetric Engineering and Remote Sensing*.

Revues nationales

- [2] R. Péteri et T. Ranchin, 2003 – Evaluation quantitative de méthodes d'extraction de rues : choix d'une référence de comparaison.– *Bulletin de la Société Française de Photogrammétrie et de Télédétection* n ° 170 (2003-2), pp. 23-29.

Congrès internationaux

- [3] R. Péteri, J. Celle et T. Ranchin. – Detection and extraction of road networks from high resolution satellite images. – Dans *Proceedings of the IEEE International Conference on Image Processing ICIP'03*, Volume I, pages 301-304, Barcelona, Spain, 14-17 September 2003.
- [4] R. Péteri et T. Ranchin. – Urban street mapping using Quickbird and Ikonos images. – Dans *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium IGARSS 2003*, Toulouse, France, 21-25 July 2003, Volume 3, pp. 1721-1723.
- [5] R. Péteri et T. Ranchin. – Multiresolution snakes for urban road extraction from Ikonos and Quickbird images. – Dans *Proceedings of the 23rd EARSeL Symposium*, Ghent, Belgium, 2-4 June 2003.
- [6] R. Péteri et T. Ranchin. – Extraction and update of street networks in urban areas from high resolution satellite images. – Dans *Proceedings of ISPRS Commission IV symposium, Geospatial Theory, Processing and Applications*, Ottawa, Canada, 8-12 July 2002, C. Armenakis, Y.C. Lee editors, ISSN 1682-1750, IAPRS & SIS, vol. XXIV, Part 4, Commission IV, pp. 768-773.
- [7] R. Péteri et T. Ranchin. – Assessment of road extraction methods from satellite images : reflections and case study on the definition of a reference. – Dans *Proceedings of 22nd EARSeL Annual Symposium "Geoinformation for European-wide integration"*, 4-6 June 2002, Prague, Czech Republic. Tomas Benes editor, Millpress, Rotterdam, Netherlands, pp. 141-147.
- [8] 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. – Dans *In Proceedings of the 21st EARSeL Symposium, Observing our environment from space : new solutions for a new millenium*, Paris, France, 14-16 may 2001, Gérard Begni editor, A. A. Balkema, Lisse, Abingdon, Exton (PA), Tokyo, pp. 327-332.

Congrès et workshops nationaux

- [9] R. Péteri et T. Ranchin. – Extraction de réseaux de rues à partir d'images satellites à haute résolution spatiale – Dans *19^{eme} Colloque Gretsi sur le traitement du signal et des images*, Volume III, pages 264-267, Paris, Septembre 2003.
- [10] R. Péteri, J. Celle et T. Ranchin. – Contours acifs et analyse multirésolution pour l'extraction de réseaux de rues à partir d'images satellites – Dans *ORASIS'03, Actes des journées francophones des jeunes chercheurs en analyse d'images et perception visuelle*, pages 393-401, Gérardmer, France, Mai 2003.

Rapports de recherche

- [11] R. Péteri. – Détection de rues dans de images satellites à haute résolution spatiale. – *Rapport final de Contrat Jeune Chercheur DGA n°0060073/A000*, Décembre 2001.

Annexes

Annexe A

L'analyse multirésolution

Principes Le concept d'analyse multirésolution tel que l'a introduit Mallat (1989) est issu des algorithmes pyramidaux. Dans cette approche, la taille du pixel est définie comme la résolution de référence qui permet de mesurer les variations locales de l'image. Précisons que la résolution définie ici est l'inverse de l'échelle définie pour la cartographie. Ainsi, quand la résolution augmente, la taille du plus petit objet visible diminue.

Pour décrire les algorithmes pyramidaux, Meyer (1992) utilise l'exemple de la cartographie :

«Nous voyons, sur cet exemple, apparaître l'idée fondamentale d'une représentation de l'image par un arbre. Dans le cas de la cartographie, le tronc serait la mappemonde et en cheminant vers les branches, les ramifications, les feuilles, on accéderait à des cartes couvrant des régions de plus en plus petites et fournissant des détails de plus en plus fins, détails qui ne figureraient pas au niveau inférieur.»

Pour interpréter cette représentation de la cartographie à l'aide d'un algorithme pyramidal, il conviendra de renverser les rôles du bas et du haut. . . »

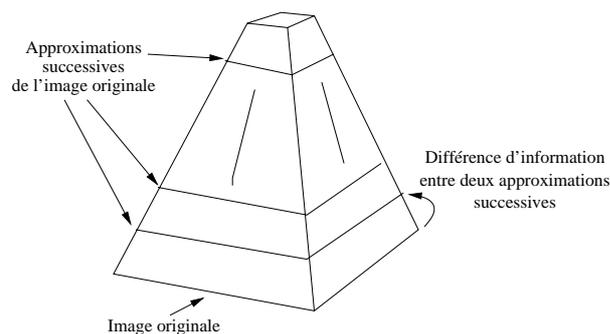


FIG. A.1: représentation des approximations successives d'une même image à l'aide d'un algorithme pyramidal

La figure A.1 est une représentation de l'analyse multirésolution et, plus généralement, des algorithmes pyramidaux. L'analyse multirésolution permet de construire des

approximations successives (les plans de contexte) et de plus en plus grossières de l'image originale (la base de la pyramide). Quand on «monte» dans la pyramide, les différents étages représentent les approximations successives de l'image originale. La limite théorique de ces approximations est le sommet de la pyramide qui correspond à un pixel unique.

Construction d'une analyse multirésolution Une analyse multirésolution est une suite de sous-espaces vectoriels $V_j (j \in \mathbb{Z})$ de $L^2(\mathbb{R}^n)$, où n est la dimension de l'espace (2 dans le cas d'une image), ayant les propriétés suivantes :

$$V_j \subset V_{j+1}$$

$$V_j \cap_j V_j = \{0\}$$

$$\bigcup_j V_j \text{ est dense dans } L^2(\mathbb{R}^n) \text{ soit } \overline{\bigcup_j V_j} = L^2(\mathbb{R}^n)$$

Invariance par dilatation :

$$\forall f(x) \in L^2(\mathbb{R}^n), \forall j \in \mathbb{Z} \text{ on a : } f(x) \in V_0 \Leftrightarrow f(2^j x) \in V_j$$

Invariance par translation :

$$\forall f_j(x) \in L^2(\mathbb{R}^n), \forall k \in \mathbb{Z} \text{ on a : } f_j(x) \in V_j \Leftrightarrow f(x - k) \in V_j$$

Il existe une fonction ϕ telle que $\{\phi(x - k), k \in \mathbb{Z}\}$ forme une base orthonormée de V_0

La fonction ϕ est appelée *fonction d'échelle*.

Dans le cas dyadique et avec des fonctions ϕ et ψ réelles qui est le cas dans lequel nous nous plaçons, on définit la base de V_j comme : $\{\sqrt{2^{-j}}\phi_j(x - 2^{-j}k), k \in \mathbb{Z}\}$

L'association de l'analyse multirésolution et de la transformée en ondelettes permet une analyse complète et un traitement efficace des images. En effet, la différence d'information qui existe entre deux approximations successives d'une même image construites par l'analyse multirésolution est décrite dans les coefficients d'ondelettes $d_{j,k}(f(x))$:

$$d_{j,k}(f(x)) = \langle f(x), \Psi_{j,k}(x) \rangle \quad (\text{A.1})$$

où la base des $\Psi_{j,k}$ forme une base orthonormée.

A partir d'une approximation $a_{j,k}(f(x))$ et des coefficients d'ondelettes, $d_{j,k}(f(x))$, on pourra reconstruire $a_{j-1,k}(f(x))$ selon :

$$a_{j-1,k}(f(x)) = a_{j,k}(f(x)) + \sum_{k \in \mathbb{Z}} d_{j,k}(f(x)) \Psi_{j,k} \quad (\text{A.2})$$

Ainsi, l'analyse multirésolution à l'aide de la transformée en ondelettes permet de reconstruire de manière exacte (sans perte d'information) le signal échantillonné original.

D'un point de vue pratique, l'implémentation de l'analyse multirésolution est réalisé par le biais d'une succession de filtres passe-bas h et filtres passe-haut g . Cette cascade de filtre est appelée un *banc de filtres*.

Annexe B

Les stades de fonction

On peut distinguer cinq stades de fonction dans une chaîne d'*EVI*¹ (*source : DGA/SPOTI*) :

Fonction tout-automatique : l'algorithme évolue de manière autonome.

Fonction semi-automatique : l'interprète fournit à l'initialisation les paramètres adaptés. Il peut faire une correction interactive, effectuée à l'issue du résultat obtenu automatiquement après paramétrage.

Fonction interactive : c'est une saisie manuelle assistée. L'algorithme suit la rue et s'arrête lors qu'il bute sur un obstacle. L'interprète saisit alors le bon point.

Fonction manuelle : l'interprète saisit point à point.

Fonction non disponible : l'intérêt de la fonction a été identifié sans avoir encore été implémenté.

Ces définitions sont adaptées au cas particulier de notre contexte de travail, mais peuvent être généralisées à toute fonction d'*EVI*.

¹*EVI : Exploitation et Valorisation des Images.*

Une fonction d'*EVI* est une chaîne de traitement visant à l'élaboration de produits finis à partir d'images (à partir des caractéristiques géométriques et radiométrique des images)

Annexe C

«Greedy Algorithm» : Pseudo-Code

Le pseudo-code fourni ci-dessous a été repris de l'article de D.J. Williams & M. Shah (Williams et Shah 1992) :

Index arithmetic is modulo n

Initialize α_i , β_i and γ_i to 1 for all i .

```
do
  for  $i = 0$  to  $n$ 
     $E_{min} = \text{BIG}$ 
    for  $j = 0$  to  $m-1$  /*  $m$  est la taille du voisinage */
       $E_j = \alpha_i \cdot E_{cont,j} + \beta_i \cdot E_{curv,j} + \gamma_i \cdot E_{image,j}$ 
      if  $E_j < E_{min}$  then
         $E_{min} = E_j$ 
         $jmin = j$ 
    Move point  $\mathbf{v}_i$  to location  $jmin$ 
    if  $jmin$  not current location,  $ptsmoved += 1$ 
    for  $i = 0$  to  $n-1$ 
       $c_i = | \vec{u}_i / | \vec{u}_i | - u_{i+1} / | u_{i+1} | |^2$ 
    for  $i = 0$  to  $n-1$ 
      if ( $c_i > c_{i-1}$  and  $c_i > c_{i+1}$ 
        and  $c_i > \text{threshold1}$ 
        and  $\text{mag}(\mathbf{v}_i) > \text{threshold2}$ 
        then  $\beta_i = 0$ 
  until  $ptsmoved < \text{threshold3}$ 
```

FIG. C.1: Greedy Algorithm pseudo-code

EXTRACTION DE RESEAUX DE RUES EN MILIEU URBAIN A PARTIR D'IMAGES SATELLITES A TRES HAUTE RESOLUTION SPATIALE

Résumé

La disponibilité d'images satellites à très haute résolution spatiale au dessus de zones urbaines est récente. Elle constitue potentiellement un très grand apport pour la cartographie des villes à des échelles de l'ordre du 1:10 000. La très haute résolution spatiale permet une représentation réelle des rues sur une carte, mais engendre une augmentation significative du bruit.

Dans cette thèse, nous proposons une méthode d'extraction des réseaux de rues en milieu urbain à partir des images à très haute résolution spatiale. Son objectif est de répondre à la forte demande dans la création automatisée de cartes.

La méthode proposée n'utilise que l'image numérique comme source d'information. Elle est semi-automatique au niveau de la détection et exploite la coopération entre la représentation linéique de la rue et sa représentation surfacique. Le graphe topologique du réseau est d'abord extrait et est utilisé pour initialiser l'étape de reconstruction surfacique. Le résultat d'extraction peut alors servir à recalculer le graphe précisément sur l'axe des rues. La méthode utilise des contraintes géométriques fortes afin de ne pas dépendre d'un modèle de profil radiométrique de la rue, trop variable en milieu urbain. Dans cette optique, un modèle de contour actif associé à la transformée en ondelettes, le *DoubleSnake*, a été développé. Son évolution dans un cadre multi-échelle permet d'extraire les sections de rues à bords parallèles dans un environnement bruité. Les positions finales des *DoubleSnakes* permettent ensuite l'extraction des intersections.

La méthode a été appliquée à des images de différents capteurs et avec différents types d'urbanisation. Un protocole innovant d'évaluation quantitative des résultats par comparaison à l'interprétation humaine a permis de montrer le caractère générique de la méthode, ainsi que sa bonne robustesse vis-à-vis du bruit.

Cette méthode constitue un pas vers une cartographie automatisée du réseau de rues urbain.

Mots clefs : télédétection, cartographie urbaine, réseaux de rues, très haute résolution spatiale, extraction de formes, contours actifs, analyse multi-échelle, transformée en ondelettes, évaluation quantitative.

EXTRACTION OF STREET NETWORKS IN URBAN AREAS FROM VERY HIGH SPATIAL RESOLUTION SATELLITE IMAGES

Abstract

The availability of very high spatial resolution satellite images over urban areas is recent. It represents potentially an important contribution to urban mapping at scales of 1:10 000. The very high spatial resolution enables a real representation of streets on a map, but generates a significant increase of noise. This report proposes a method for extracting urban street networks from very high spatial resolution images. Its goal is to meet the need for an automatized building of maps.

The proposed method uses digital image as only input data. It is semi-automatic at the detection step, and takes advantage of cooperation between linear representation of streets and their representation as surface elements. A topological graph of the street network is first extracted, and used for initializing the surface reconstruction step. The extraction result can then be used in order to precisely register the street centerline. This method favors strong geometrical constraints in order to avoid a radiometric profile model of the street, too variable in urban areas. To that aim, a model of active contour associated with the wavelet transform, called *DoubleSnake*, has been developed. Its evolution in a multi-scale framework enables the extraction of parallel street sides in a noisy environment. Then, final positions of *DoublesSnakes* permit the extraction of intersections.

The method has been applied on images from different sensors and with different urban types. An innovative protocol for a quantitative assessment of the results compared to human interpretation has shown its generic aspect, as well as its robustness with respect to noise.

This method is a step toward a fully automatized cartography of the street network.

Keywords : remote sensing, urban cartography, street networks, very high spatial resolution, pattern recognition, active contours, multi-scale analysis, wavelet transform, quantitative assessment.

Laboratoire d'accueil : Centre d'Energétique - Groupe Télédétection & Modélisation
Ecole des Mines de Paris
Rue Claude Daunesse - B.P. 207 -F-06904 Sophia Antipolis Cedex

Thèse présentée par : PÉTERI Renaud **le :** 10 décembre 2003

Discipline : «Informatique temps réel - Robotique - Automatique»
Ecole des Mines de Paris
