

N° d'ordre : 2265

THÈSE

présentée pour obtenir le titre de

DOCTEUR DE L'INSTITUT NATIONAL POLYTECHNIQUE DE TOULOUSE

École doctorale d'Informatique et Télécommunications

Spécialité en Signal, Image et Acoustique

en cotutelle avec l'Institut d'Automatique, Académie des Sciences de Chine

Par

Timothée BAILLOEUL

**Contours actifs et information a priori pour l'analyse de changements :
application à la mise à jour de cartes numériques du bâti urbain à partir
d'images optiques de télédétection haute résolution**

Soutenue le 21 octobre 2005 devant le jury composé de :

Long Quan	Professeur	Université des Sciences et Technologies de Hong Kong, Chine.	<i>Président</i>
Mihai Datcu	Professeur	Ecole Nationale Supérieure des Télécommunications, France.	<i>Rapporteur</i>
ZhanYi Hu	Professeur	Institut d'Automatique, Académie des Sciences de Chine.	<i>Rapporteur</i>
Baogang Hu	Professeur	Institut d'Automatique, Académie des Sciences de Chine.	<i>Co-directeur</i>
Ian Jermyn	Chargé de recherche	Institut National de Recherche en Informatique et en Automatique, France.	<i>Examineur</i>
Philippe Marthon	Professeur	Institut National Polytechnique de Toulouse, France.	<i>Co-directeur</i>
Véronique Prinnet	Professeur associé	Institut d'Automatique, Académie des Sciences de Chine.	<i>Examineur</i>
Bruno Serra	Ingénieur, Docteur	Alcatel Alenia Space, France.	<i>Examineur</i>
PengFei Shi	Professor	Université Jiao Tong de Shanghai, Chine.	<i>Invité</i>
Jixian Zhang	Professeur	Institut National de Relevé Topologique et de Cartographie, Chine.	<i>Invité</i>

Remerciements

Je tiens tout d'abord à remercier Baogang Hu et Marc Jaeger pour m'avoir accueilli au LIAMA et pour leur prévenance à rendre mon séjour au laboratoire le plus agréable possible.

Je suis reconnaissant à Véronique Prinet et Jean-Guy Planès pour m'avoir proposé ce sujet de thèse et pour la confiance qu'ils m'ont accordée. Merci à Véronique pour avoir encadré mes recherches, pour les libertés qu'elle m'a laissées et pour avoir toujours été présente lors des passages délicats de ce long trajet de thèse.

Mes remerciements vont également à Bruno Serra et Philippe Marthon qui ont aussi suivi mes travaux. Merci à Philippe pour ses remarques, ses critiques constructives et pour m'avoir reçu dans son unité de recherche. J'adresse toute ma sympathie à Bruno pour avoir été un correspondant actif, volontaire et efficace à Alcatel Alenia Space. Je remercie son soutien et sa gentillesse.

Je remercie très vivement les membres du jury pour le temps qu'ils ont passé à la lecture de ce document.

Je remercie Josiane Zérubia et toute son équipe pour m'avoir accueilli pendant plusieurs mois au sein du projet Ariana. Je tiens particulièrement à remercier Ian Jermyn et Marie Rochery pour leur assistance et leurs conseils sur les modèles quadratiques. Merci à Ian pour sa patience, nos nombreuses discussions et l'intérêt qu'il a porté à mes travaux.

Je suis également reconnaissant à Pinxiang Chen, Haitao Zhang et Lebing Sun de l'Institut de Relevé Topographique et de Cartographie de Pékin pour nous avoir prêté les cartes et les images aériennes utilisées dans cette étude. Je tiens à remercier très sincèrement Haitao Zhang pour sa disponibilité et son aide diligente.

Merci à Véronique Prinet, Philippe Marthon, Baogang Hu, Fengfeng Tang, Bizhen Hong, Wenli Lou, Li Li, Martine Lacoste et Serge Lefevre pour m'avoir aidé à surmonter les difficultés de la mise en place de la première cotutelle entre l'Académie de Sciences de Chine et l'INPT. Mille mercis à Fengfeng Tang pour avoir été un secrétaire exemplaire, efficace et enthousiaste.

Merci également à Jean-François Barczi pour sa bonne humeur, et pour avoir répondu sans relâche à mes questions d'ordre informatique.

Je remercie chaleureusement tous les étudiants et les stagiaires du LIAMA qui participent à l'ambiance sympathique du laboratoire. J'ai une pensée particulière pour Jinghui Duan et Fengfeng Tang qui ont toujours été prompts à résoudre mes problèmes avec la langue chinoise. J'adresse toute mon amitié à Wanlin Zhu pour nos discussions hebdomadaires sur les contours actifs et nos parties de badminton endiablées. Merci à vous tous pour votre révision efficace du résumé chinois de la thèse.

Enfin, un grand merci à ma famille française et chinoise pour leur soutien inaltérable. Merci à mes parents qui m'ont toujours encouragé de loin (au sens cartographique du terme) : je remercie ma mère pour ses nombreuses relectures du manuscrit et mon père pour sa curiosité inépuisable envers mes activités scientifiques. Merci à Jing, ma femme, pour son optimisme et ses encouragements qui ont été d'un réconfort essentiel durant ces années de thèse.

Résumé

La maintenance de données cartographiques est traditionnellement un processus de photo-interprétation d'images de télédétection. Avec l'émergence commerciale des images satellitaires optiques de résolution sub-métrique (Ikonos, Quickbird) la mise à jour fiable et régulière de cartes numériques à grande échelle devient abordable. Cependant, l'actualisation manuelle des données cartographiques, lente et coûteuse, ne permet plus de répondre au flux grandissant des données de télédétection. De nombreux travaux en vision par ordinateur ont été dédiés à l'automatisation de cette tâche, offrant des résultats satisfaisants sur des scènes rurales ou péri-urbaines. Les milieux urbains restent néanmoins plus complexes à analyser à haute résolution : la présence d'ombres, d'occlusions et la profusion de détails rendent l'interprétation des images difficile.

Cette thèse a pour objectif de proposer une méthodologie visant à analyser les changements entre une carte numérique de bâtiments et des données de télédétection optiques haute résolution plus récentes. L'idée centrale de notre approche est d'utiliser l'information géométrique et spécifique des bâtiments symbolisés dans la carte afin de faciliter leur reconnaissance dans une image satellitaire panchromatique et de s'affranchir des difficultés du milieu urbain. Cette connaissance *a priori* est incorporée comme une contrainte de forme dans un modèle variationnel de recalage fin carte-image par contours actifs. Nous proposons des solutions nouvelles pour l'amélioration de la robustesse des contours actifs contraints : la fusion de l'image avec un Modèle Numérique de Surface (MNS) permet de surmonter le problème des minima locaux ; la variation spatio-temporelle de la contrainte de forme permet aussi d'accroître l'efficacité de la mise en correspondance tout en réduisant les temps de convergence. Les variabilités locales de forme carte-image dues soit à des erreurs de saisie lors de la photo-interprétation, soit aux effets de simplification et de généralisation de la carte, sont aussi considérées. Nous proposons un nouveau critère énergétique de contrainte de forme, alliant énergie linéaire et quadratique, permettant aux contours actifs de s'affranchir localement et parallèlement de la forme *a priori* entachée d'erreur.

L'alignement des données par contours actifs permet ainsi d'accroître la cohérence entre la carte et l'image et améliore la position planimétrique des objets cartographiques retrouvés dans l'image. Ce recalage fin résout le problème des variabilités exogènes carte-image indépendantes d'un changement effectif et rend ainsi l'analyse de changements subséquente plus fiable. Des indices de changement simplement calculés à partir du MNS et d'une image satellitaire multispectrale sont fusionnés pour la détection de changements univoques du bâti urbain. En l'absence de changement univoque, une méthode basée sur le vote de Hough quantifie la ressemblance entre le polygone cartographique recalé et des primitives segments extraites de l'image panchromatique. La fusion du score de Hough avec la variation géométrique enregistrée par le recalage des contours actifs permet la formulation d'une probabilité de non-changement visant à assister un opérateur dans sa tâche de révision de carte. Nous illustrons l'efficacité de l'approche proposée à partir d'expériences menées avec une carte d'échelle 1 :10,000 et une image satellitaire Quickbird (0.6 m/pixel) de la ville de Pékin qui connaît actuellement un fort développement de son tissu urbain.

Abstract

The maintenance of cartographic data usually applies to the photo-interpretation of remote sensing images. With the commercial advent of optical satellite images with sub-meter resolution (Ikonos, Quickbird) a higher frequency of digital map revision and monitoring is affordable and realistic at a large scale. However, manual map updating is slow and costly and it cannot handle the increasing flow of available remote sensing data. Many works have been dedicated to the automation of this task in Image Processing and Computer Vision, yielding remarkable results over rural or suburban areas. Urban areas are however more difficult to analyze at a high resolution : the presence of shadows, occlusions and details profusion make their interpretation difficult.

This thesis aims at proposing a methodology to analyze changes between a digital map of buildings and more recent high resolution optical remote sensing data. The key idea of the proposed approach is to use the specific geometrical information derived from the buildings symbolized in the map in order to ease their recognition in a panchromatic satellite image while overcoming urban difficulties. This prior knowledge is embedded in active contours models as a shape constraint formulated in a variational approach. We propose new solutions to improve the robustness of shape constrained active contours : the fusion of the image with a digital surface model (DSM) allows to overcome the critical issue of local minima ; the spatio-temporal variation of the shape constraint also enables to increase the efficiency of map-to-image matching while decreasing the computational workload. Local shape discrepancies between the cartographic objects and the image are considered (these variabilities are either due to delineation mistakes or simplification/generalization effects in the map). We propose a new shape constraint energetic criterion, made of linear and quadratic energies, to allow local and parallel discrepancies of the active contour from the mistaken cartographic shape prior.

The fine map-to-image matching achieved by the active contours increases the consistency between the map and the image and improves the planimetric localization of cartographic objects found in the image. The fine matching attempts to solve the issue of exogenous variabilities between the map and the image ; it therefore makes the subsequent change detection decision more reliable. Change indicators simply calculated from the DSM and a multispectral satellite image are merged to detect obvious changes of buildings. When no obvious change is detected, a method based on Hough voting quantifies the similarity between the refined cartographic object and segment primitives extracted from the panchromatic image. The fusion of the Hough voting score with the geometric variation due to the active contours matching enables to formulate a “non-change” probability aiming at assisting a photo-interpreter performing map updating. We show the efficiency of the proposed scheme with experiments carried out with a 1 :10,000 scale map and a Quickbird satellite image (0.6 m/pixel) depicting Beijing city, which has nowadays a rapid urban growth.

概述

地图数据的更新经常要采用遥感图像处理技术中的图像理解技术。随着商用一米以下分辨率的光学卫星图像的出现 (Ikonos, Quickbird), 大规模、经常性的地图更新和监测是可负担和实际可行的。然而, 人工地图更新不仅时间长而且费用高, 不能胜任处理越来越多的遥感图像。很多研究试图利用图像处理和计算机视觉领域中的方法来使这一过程自动化, 并产生了很多针对农村和城镇的有价值的结果。城镇地区的高分辨率图像更难分析, 因为阴影的出现、建筑物的遮挡和过多的细节增加了分析的难度。

本论文通过分析数字地图的建筑和高分辨率遥感影像信息之间的差异, 提出了一个能够解决上述难题的方法。该方法主要是利用从地图所标示的建筑物所获得的特定的地理信息, 使从新近获得的全色 (黑白) 卫星图像识别建筑物变得更加容易。先验信息作为形状约束被包含进活动轮廓模型的能量范函中。新的能量约束提高了活动轮廓模型进化过程中克服局部极值的能力。论文中提出了新的解决方法如下: 融合数字表面模型 (DSM) 作为先验知识可以克服局部极值问题; 时空可变的形状约束系数也可以有效增加匹配精度同时减少计算量。论文提出了新的形状限制准则, 其允许图像目标与提取的先验形状之间存在局部形状差异 (这些差异可能是地图绘制过程中或者简化造成的), 这些约束由线性和二次能量形式构成, 作为形状约束被包含进能量范函。

通过活动轮廓得到的精细地图-图像匹配可以提高两者的一致性, 提高地图中建筑物的形状和位置精度。精细匹配试图解决地图和图像间的外围变化, 因此使后续的变化测量决定更可靠。数字表面模型与多谱卫星图像之间的变化用来检测建筑物存在的变化。当没有这种明显的变化时, Hough 变换用来量化地图中物体的边缘和图像中建筑物的边缘之间的相似性, Hough 变换结合活动轮廓模型匹配产生的几何差异允许在辅助地图更新时提供需要更新的概率。为了验证提出的算法, 论文采用了 1: 10, 000 比例的北京地区 (一个快速变化的城市) 地图和 Quickbird 卫星图像 (0.6 米分辨率), 实验验证了本方法是相当有效的。

Table des matières

Notations et conventions	17
1 Introduction	19
1.1 Contexte applicatif et besoins d'une mise à jour de cartes urbaines	19
1.2 Objectif scientifique et approche choisie	20
1.3 Innovations	21
1.4 Plan de la thèse	21
1.5 Cadre de la thèse	23
2 Analyse et formulation du problème	25
2.1 Introduction	25
2.2 État de l'art sur l'analyse de scènes et la mise à jour de cartes à partir d'images de télédétection haute résolution	25
2.2.1 Spécificités et stratégies pour la mise à jour de cartes	25
2.2.1.1 Comparaison d'images	25
2.2.1.2 Comparaison carte-image	26
2.2.2 Quelques travaux sur la fusion de données de télédétection pour l'analyse de scènes à haute résolution	27
2.2.2.1 Fusion d'images optiques	27
2.2.2.2 Fusion d'images optiques et altimétriques	28
2.2.3 Utilisation de contraintes issues de connaissance <i>a priori</i>	29
2.2.3.1 Connaissance générique	29
2.2.3.2 Connaissance spécifique	30
2.2.4 Synthèse	33
2.3 Position du problème de la mise à jour de cartes	33
2.3.1 Cadre conceptuel général	33
2.3.1.1 Définition des univers relatifs à la télédétection	33
2.3.1.2 Définition de $U_{interprétation}$	34
2.3.2 Problématique adressée dans la thèse	35
2.3.2.1 Analyse de changements carte-image considérée dans la thèse	35
2.3.2.2 Difficultés des milieux urbains	38
2.4 Méthodologie proposée	38
2.5 Conclusion	40
3 Génération d'un modèle numérique de surface orthoscopique	41
3.1 Introduction	41
3.2 Stéréoscopie photogrammétrique	41
3.2.1 Stéréoscopie - rappel et principe général	42

3.2.2	Mise en correspondance automatique d'images par masques de corrélation adaptatifs	44
3.2.2.1	Critère de ressemblance par corrélation croisée	45
3.2.2.2	Corrélation par masques adaptatifs	47
3.2.3	Principe de la chaîne de reconstruction utilisée	50
3.3	Optimisation des paramètres de prise de vue	51
3.3.1	Paramètres de prise de vue	51
3.3.1.1	Calcul des paramètres d'affinité : orientation interne	53
3.3.1.2	Évaluation de la qualité des paramètres de prise de vue	53
3.3.2	Optimisation globale	56
3.3.2.1	Fonction de coût	56
3.3.2.2	Référentiel cartésien géocentré	56
3.3.2.3	Résultats	59
3.3.3	Évaluation de la qualité de géocodage des MNSs	61
3.4	Orthorectification d'un MNS épipolaire	63
3.4.1	Orthorectification par méthode directe	63
3.4.1.1	Interpolation triangulaire (interpolation bilinéaire sur les 3 plus proches voisins)	63
3.4.1.2	Algorithme	65
3.4.1.3	Résultats	65
3.4.2	Pré-traitements sur l'image de disparité	66
3.4.2.1	Filtrage médian	67
3.4.2.2	Filtre de réjection	67
3.4.2.3	Synthèse	68
3.5	Fusion de MNSs	68
3.6	Conclusion	71
4	Présentation et choix de contours actifs avec contrainte de forme pour le recalage fin carte-image	73
4.1	Introduction	73
4.2	État de l'art des contours actifs contraints par une forme <i>a priori</i>	74
4.2.1	Les contours actifs	74
4.2.1.1	Représentation d'un contour actif	74
4.2.1.2	Approche géométrique et variationnelle	78
4.2.1.3	Contours actifs basés sur l'information de frontière	79
4.2.1.4	Contours actifs basés sur l'information de région	83
4.2.2	Segmentation sous contrainte de forme spécifique	88
4.2.2.1	Approche non statistique	88
4.2.2.2	Approche statistique	90
4.2.2.3	Invariance de la contrainte de forme par transformation géométrique	91
4.2.3	Segmentation sous contrainte de forme générique : les modèles quadratiques	92
4.3	Contours actifs pour le recalage fin carte-image : choix et implantation	93
4.3.1	Choix de la représentation du contour actif	93
4.3.2	Choix du terme d'attache aux données	93
4.3.2.1	Modèles basés sur l'information de région	94
4.3.2.2	Modèle basé sur l'information de frontière	95
4.3.3	Choix de la contrainte de forme	96
4.3.4	Algorithme et optimisation	97

4.3.4.1	Réduction des temps de calculs : utilisation de la technique de bande étroite rapide	97
4.3.4.2	Diagramme d'évolution du contour actif contraint	99
4.4	Conclusion	101
5	Contours actifs contraints robustes	103
5.1	Introduction	103
5.2	Estimation robuste des paramètres de T_{sim} par algorithme du simplex	103
5.3	Fusion de termes exogènes dans la fonctionnelle d'énergie d'attache aux données	105
5.4	Insertion plus souple de la forme <i>a priori</i> : flexibilité spatio-temporelle de λ	108
5.5	Applications et analyse des résultats expérimentaux	110
5.5.1	Protocole expérimental	111
5.5.2	Comparaison simplex vs. descente de gradient pour l'estimation des paramètres de T_{sim}	111
5.5.3	Contribution de la connaissance de forme <i>a priori</i> pour la segmentation	113
5.5.4	Sensibilité à l'initialisation des contours actifs contraints	115
5.5.4.1	Comparaison des modèles basés sur l'information de région	115
5.5.4.2	Modèle basé sur l'information de frontières	123
5.5.5	Solutions pour une moindre sensibilité à l'initialisation et aux minima locaux	123
5.5.5.1	Résultats avec fusion de données exogènes dans le terme d'attache aux données	126
5.5.5.2	Résultats avec une contrainte de forme spatio-temporelle	129
5.6	Limites du modèle avec contrainte de forme	134
5.7	Conclusion	137
6	Contrainte de forme autorisant les variations parallèles	139
6.1	Introduction	139
6.2	Énergie de contrainte de forme autorisant les variations parallèles	139
6.3	Connaissance <i>a priori</i> générique par modèles quadratiques	140
6.3.1	Énergie quadratique correctrice	141
6.3.1.1	Modèle	141
6.3.1.2	Gradient de l'énergie	142
6.3.2	Algorithme	143
6.3.2.1	Calcul de la force quadratique	144
6.3.2.2	Extension de la force	144
6.3.3	Difficultés d'implantation	144
6.4	Expériences et analyse des résultats	146
6.4.1	Résultats sans correction quadratique	146
6.4.2	Résultats avec correction quadratique	150
6.5	Conclusion	152
7	Analyse de changements carte-image	153
7.1	Introduction	153
7.2	Méthodologie	153
7.3	Calcul et fusion d'indices simples de changement	155
7.3.1	Indice calculé à partir de la donnée altimétrique (MNS)	155
7.3.1.1	Limitations de la comparaison carte-MNS	155
7.3.1.2	Approche	155
7.3.1.3	Résultats expérimentaux	157
7.3.2	Indice calculé à partir d'une image multispectrale	161

7.3.3	Fusion de la prise de décision	163
7.4	Analyse de changements carte-image	164
7.4.1	Mise en correspondance de segments carte-image par vote de Hough	164
7.4.1.1	Principe	164
7.4.1.2	Algorithme	165
7.4.1.3	Énergie associée au vote de Hough	167
7.4.2	Mesure de la variation géométrique due au recalage par contours actifs	167
7.4.3	Probabilité de détection de non-changement	168
7.5	Résultats expérimentaux	169
7.5.1	Protocole expérimental	169
7.5.2	Étude de quelques cas de changements/non-changements	170
7.5.2.1	Détection de changement	170
7.5.2.2	Détection de non-changement	173
7.5.2.3	Ambiguïté et fausse alarme avec une carte entachée d'erreurs locales	179
7.5.2.4	Synthèse des résultats	180
7.5.3	Analyse statistique	181
7.5.3.1	Résultats avec une carte de bonne qualité	181
7.5.3.2	Résultats avec une carte de qualité moyenne	182
7.6	Conclusion	183
8	Conclusion	185
8.1	Synthèse	185
8.2	Discussion des limitations	187
8.2.1	Limites du modèle de contrainte de forme	187
8.2.2	Limites de l'analyse de changements	187
8.2.3	Limites des contours actifs	188
8.3	Perspectives	188
A	Présentation et pré-traitements des données	191
A.1	Présentation des données	191
A.1.1	Données de télédétection	191
A.1.1.1	Images aériennes	191
A.1.1.2	Images satellitaires	193
A.1.2	Données cartographiques	193
A.2	Pré-traitements des données de télédétection	194
A.2.1	Sous-échantillonnage des images aériennes	194
A.2.2	Rectification des images satellitaires	195
B	Algorithme du Simplex	197
C	Modélisation de localisation directe et inverse	199
C.1	Modélisation inverse	199
C.2	Modélisation directe	200
D	Calcul rapide de la moyenne et de la variance à l'intérieur et l'extérieur d'un contour actif représenté par un ensemble de niveaux	203

E	Calcul de variations d'énergies de contrainte de forme et quadratique	205
E.1	Équation d'évolution dérivée de la fonctionnelle de contrainte de forme par différence de normales	205
E.2	Dérivée de l'énergie de correction quadratique	207
E.2.1	Énergie $J_1(C)$	207
E.2.2	Énergie $J_2(C)$	208
E.2.3	Variation totale de l'énergie $J_{correction}(C)$	208
F	Représentation 3D d'une scène urbaine	211
F.1	Objectif et rappel de la méthodologie	211
F.2	Principe de la modélisation	211
F.3	Résultats sur la ville de Pékin	211
	Bibliographie	213
	Liste des publications	221
	Table des figures	223
	Liste des tableaux	229

Notations et conventions

Typographie

- Une variable scalaire est notée en italique maigre : x
- Un vecteur colonne est noté en minuscule, droite, grasse : $\mathbf{x}$
- Un vecteur ligne est un vecteur colonne transposé : $\mathbf{x}^T$
- Une matrice est représentée par une majuscule, droite, grasse : $\mathbf{M}$

Notations

$\langle \mathbf{x}, \mathbf{y} \rangle$	produit scalaire de deux vecteurs
$ x $ ou $ \mathbf{x} $	valeur absolue d'un scalaire ou norme Euclidienne d'un vecteur
$\overline{AB}$	longueur du segment $[AB]$ où A et B sont deux points d'un espace Euclidien
∇f	gradient d'une fonction scalaire
$\operatorname{div}(\mathbf{x})$	divergence d'un vecteur
Δf	Laplacien d'une fonction scalaire
$\min(a, b)$	minimum de deux arguments réels
$\max(a, b)$	maximum de deux arguments réels
$\operatorname{sign}(a)$	fonction caractéristique de l'argument réel a . Elle vaut $+1$ si a est strictement positif ; -1 si a est strictement négatif ; 0 si a est nul
$H(\cdot)$	fonction échelon de Heaviside
$\delta(\cdot)$	distribution de Dirac
$H_a(\cdot), \delta_a(\cdot)$	approximations régulières et dérivables de $H(\cdot)$ et $\delta(\cdot)$
$g(x)$	fonction à valeurs réelles tendant asymptotiquement vers zéro. $g: \mathbb{R}^+ \rightarrow \mathbb{R}$ $g(x) = \frac{1}{1+x^p}$, avec $p \in \mathbb{N}$, typiquement $p = 1, 2$
$I(c, l)$	valeur de l'image I à la colonne c et la ligne l

Acronymes

- **MAP** : **M**aximum **A** *P*osteriori
- **ACP** : **A**nalyse par **C**omposante **P**incipale
- **MDL** : **M**inimum **D**escription **L**ength
- **SIG** : **S**ystème d'**I**nformation **G**éographique
- **BISM** : **B**eijing **I**nstitute of **S**urveying and **M**apping
- **MNS** : **M**odèle **N**umérique de **S**urface
- **LIDAR** : **L**ight **D**etection **A**nd **R**anging
- **UTM** : **U**niversal **T**ransverse **M**ercator

- **CSTB** : **C**entre **S**cientifique et **T**echnique du **B**âtiment
- **GPS** : **G**lobal **P**ositioning **S**ystem
- **RVB** : **R**ouge **V**ert **B**leu
- **HSI** : **H**ue **S**aturation **I**ntensity
- **RANSAC** : **R**ANdom **S**Ample **C**onsensus
- **IGN** : Institut **G**éographique **N**ational (France)
- **BDTopo** : **B**ase de **D**onnées **T**opographiques, produit IGN
- **PPS** : **P**oint **P**rincipal de **S**ymmétrie
- **BNS** : **B**eijing **N**ational **S**ystem
- **SAR** : **S**ynthetic **A**perture **R**adar
- **GVF** : **G**radient **V**ector **F**low
- **NDVI** : **N**ormalized **D**ifference **V**egetation **I**ndex

Chapitre 1

Introduction

1.1 Contexte applicatif et besoins d'une mise à jour de cartes urbaines

Depuis quelques années, les images optiques de télédétection ayant une résolution sub-métrique dans le domaine civil ne sont plus l'apanage des images aériennes. L'imagerie satellitaire, comme en témoignent les programmes Quickbird et Ikonos, atteint dorénavant des résolutions similaires avec les avantages de moindre coût et d'une fréquence d'acquisition plus élevée. C'est un atout sans précédent pour la maintenance régulière des données cartographiques à grande échelle. La mise à disposition de cartes actualisées est indispensable pour de nombreuses applications. A titre d'exemple, bénéficier d'une fréquente mise à jour du tissu urbain est essentiel pour planifier à long terme le développement durable d'une cité. Une représentation fiable et récente du réseau routier permet de solutionner les problèmes de trafic automobile ; connaître avec précision la localisation des zones résidentielles permet d'optimiser les transports dans les zones mal desservies et de savoir où développer les prochaines habitations. La pollution urbaine, de nature atmosphérique ou sonore, peut être mieux comprise et mieux contrôlée grâce à une carte enrichie de l'information d'altitude du bâti. Dans un contexte très différent, celui des catastrophes naturelles (tremblements de terre, inondations) ou moins naturelles (conflits militaires), la production rapide de cartes des dégâts - en termes d'infrastructures et de bâtiments détruits - est un élément clé pour pouvoir apporter un secours rapide à la population et permettre une réhabilitation efficace. On comprend alors qu'avant tout, une carte fréquemment mise à jour représente un outil d'aide à la décision, qu'elle doive être prise dans l'urgence ou pour une planification à long terme.

Alors que le besoin en cartes fiables et actualisées est grandissant et que le flux des images de télédétection est croissant, le goulot d'étranglement de la production de données cartographiques est son traitement manuel. La maintenance de carte est traditionnellement réalisée soit par relevé sur le terrain, soit par des opérateurs détectant des changements entre une ancienne base de données et des images de télédétection plus récentes. Dans les deux cas, ce processus est lent et donc coûteux, et par conséquent ne permet pas de répondre à la demande. Un besoin urgent d'automatiser la photo-interprétation sollicitant des techniques de compréhension d'images par ordinateur s'est alors fait ressentir. Depuis une vingtaine d'années, de nombreux travaux ont ainsi déjà été réalisés dans le but de faciliter la mise à jour de cartes, ou d'essayer de la rendre automatique. Force est de constater que l'utilisation de l'imagerie très haute résolution comme seule source d'information pour la maintenance de données cartographiques à grande échelle est un échec.

1.2 Objectif scientifique et approche choisie

Dans cette thèse, nous adressons le problème de la mise à jour de données cartographiques du bâti urbain à partir d'imagerie satellitaire optique haute résolution (0.6 m/pixel) et de données géospatiales auxiliaires, notamment altimétriques de type MNS (Modèle Numérique de Surface). L'objectif de ce travail de recherche consiste à étudier et à proposer une méthodologie pour l'analyse de changements entre une représentation symbolique des bâtiments d'une scène urbaine (carte numérique) et des données de télédétection plus récentes. Dans cette étude, l'analyse de changements concerne la vérification de la présence ou de la disparition dans l'image d'un bâtiment symbolisé dans une carte numérique plus ancienne.

La comparaison entre une carte et une image optique haute résolution présente deux difficultés majeures. La première est relative à l'interprétation automatique ou semi-automatique de l'image d'un paysage urbain : la profusion de détails, la présence d'ombres, le faible contraste des objets et autres artefacts, sont autant d'obstacles faisant échouer les approches traditionnelles du traitement des images et de reconnaissance de formes. La seconde est due aux erreurs et aux simplifications contenues dans la carte qui introduisent des variabilités que nous appellerons *exogènes* entre la donnée sémantique et l'image. Ces variabilités exogènes sont spécifiques à chaque objet cartographique et sont susceptibles d'être interprétées à tort comme un changement effectif.

En vue de surmonter ces problèmes, nous proposons une méthodologie basée sur la fusion de données et l'utilisation de connaissance *a priori*. La fusion de données géospatiales permet de tirer avantage de la complémentarité de leurs informations dans le but d'assurer une plus grande robustesse des traitements. La connaissance *a priori* de position et de forme - fournie par la carte numérique - permet quant à elle de restreindre l'espace de recherche des objets à retrouver dans l'image pour une comparaison avec la carte.

Nous proposons par conséquent une approche basée sur les principes suivants :

- Détecter les changements évidents entre la carte et les données de télédétection par une méthode simple. Une détection des changements univoques du bâti de la carte tirera parti de la fusion d'une donnée altimétrique (MNS) et d'une image multispectrale de télédétection. Ces changements étant sans ambiguïtés, ils ne nécessitent pas de traitements complexes des images. Nous ferons ainsi appel à des traitements éprouvés et reconnus par les communautés de traitement de l'image et de télédétection.
- Développer une méthode plus sophistiquée pour les cas plus complexes de changement/non-changement. Nous désignons par situations complexes les cas où l'information de la carte est partiellement erronée ou lorsque celle de l'image est altérée par les artefacts des milieux urbains. Dans ces circonstances, la détection de changement ne peut être menée par des traitements simples. Nous avons, pour ce faire, choisi une approche en deux temps :
 - Effectuer la correction des variabilités exogènes entre la carte et une image satellitaire panchromatique par la technique des contours actifs contraints. Les contours actifs offrent un cadre théorique particulièrement approprié à l'intégration de connaissance de forme *a priori* et à la fusion de données. Ils sont par ailleurs parfaitement adaptés à l'intégration de données symboliques vectorielles (carte numérique) et matricielles (image). Cependant, les modèles de contours actifs sous contrainte proposés dans la littérature ne répondent pas entièrement aux problèmes liés à notre application. En particulier, la profusion de détails dans les

images très haute résolution en région urbaine tend à multiplier l'occurrence des minima locaux de la fonctionnelle d'énergie minimisée par les contours actifs. La présence d'erreurs locales dans la forme *a priori* dérivée de la carte limite également le pouvoir de recalage de la technique. L'un des objectifs de cette thèse sera de s'affranchir des problèmes précités en proposant des solutions nouvelles pour une convergence plus robuste, plus rapide du contour avec en même temps une plus grande flexibilité vis-à-vis de la forme connue *a priori*.

- Analyser les changements et les non-changements en comparant l'objet cartographique affiné par les contours actifs avec l'image panchromatique. Cette analyse vise à produire un indicateur de changement dont la fiabilité est accrue par le recalage fin des contours actifs.

1.3 Innovations

Les principales innovations de ce travail de thèse concernent les modèles de contours actifs. La première originalité regarde l'adaptation des contours actifs contraints à la problématique du recalage carte-image et de l'analyse de changements entre une carte et des données optiques de télédétection haute résolution.

Nos autres contributions sont relatives à l'amélioration de la robustesse de ces modèles vis-à-vis des minima locaux. Pour palier ce problème nous proposons deux solutions : l'une concerne la fusion dans le terme d'attache aux données d'un MNS en plus de l'image analysée. La meilleure discrimination du bâti dans le MNS rend le contour actif moins sensible à l'initialisation et diminue les temps de convergence. La seconde solution consiste à relaxer spatialement et dynamiquement la contrainte de forme pendant la convergence du contour actif. Celui-ci atteint alors une solution approximative de l'objet à recaler dans l'image avant que la contrainte de forme soit uniformément rétablie afin de régulariser le contour.

La contribution la plus innovante de ce travail de thèse est la création de nouvelles énergies de contrainte de forme permettant au contour actif de s'affranchir localement et parallèlement de la forme *a priori*. La présence d'erreurs locales de forme des objets cartographiques nous a conduit à la création de tels modèles.

Enfin, nous avons tiré parti de la robustesse de l'algorithme d'optimisation du simplex dans des circonstances où son utilisation fut alors jusqu'ici inexplorée. En particulier, son intégration dans la contrainte de forme des contours actifs se révélera plus performante que la technique de descente de gradient traditionnellement employée.

1.4 Plan de la thèse

Dans un premier temps, le chapitre 2 retracera un bref état de l'art sur la mise à jour de données cartographiques et l'extraction d'objets à partir de données de télédétection haute résolution. Nous constaterons que l'usage de connaissance *a priori* spécifique extraite de la carte et la fusion de données sont les tendances actuelles pour faciliter l'analyse de scènes. Nous exposerons ensuite précisément la problématique adressée dans cette thèse, ses difficultés, et l'approche que nous proposons pour mener à bien l'analyse de changements carte-image. Notre méthodologie s'appuie à la fois sur l'utilisation de connaissance *a priori* et sur la fusion de données auxiliaires. Nous utiliserons une donnée altimétrique comme source d'information additionnelle. Le chapitre suivant décrit comment cette donnée a été générée à partir d'images aériennes stéréoscopiques.

Le chapitre 3 décrit la génération d'un Modèle Numérique de Surface (MNS) à partir de couples d'images stéréoscopiques aériennes. Le MNS est utilisé comme une source additionnelle d'information aux données satellitaires pour aider l'analyse du bâti urbain. Le problème de calibration de caméra a été traité au moyen d'un nouvel algorithme d'optimisation destiné à estimer précisément les paramètres de prise de vue des images aériennes. Cette approche permet d'estimer les paramètres en utilisant la redondance d'information de points de contrôle extraits de n images stéréoscopiques. Une fonction de coût non linéaire construite à partir du modèle de la caméra est minimisée afin d'estimer les paramètres de calibration. Les MNSs générés sont ensuite géocodés, orthorectifiés et fusionnés. L'information contenue dans le MNS sera ultérieurement utilisée lors de la détection de changements univoques (chapitre 7) ainsi qu'avec les modèles de contours actifs (chapitre 5). Le chapitre suivant rappelle les fondements théoriques de ces outils mathématiques.

Le chapitre 4 présente dans une première partie un bref état de l'art des contours actifs. Nous classerons les différents modèles selon leur mode de représentation, leurs termes d'attache aux données et de régularisation, avec une attention spéciale à l'insertion de contrainte de forme exogène connue *a priori*. Enfin, nous motiverons notre choix pour une représentation par ensemble de niveaux du contour actif et exposerons les modèles d'attache aux données et de contrainte de forme qui seront employés dans cette étude.

Le chapitre 5 montre comment accroître la robustesse des contours actifs contraints par la forme et destinés à réaliser une mise en correspondance entre les bâtiments cartographiques et l'image. La première contribution consiste à proposer un algorithme d'optimisation différent pour l'estimation des paramètres de pose rendant la contrainte de forme invariante par transformation globale. La seconde est relative à l'insertion du MNS dans le terme d'attache aux données des contours actifs. Une telle représentation est complémentaire de l'image satellitaire et permet une meilleure discrimination du bâti du reste de la scène. La troisième innovation regarde l'incorporation d'une contrainte de forme variant en fonction de l'espace et du temps. Celle-ci est destinée à conférer plus de flexibilité aux contours actifs afin de les rendre moins sensibles aux minima locaux de l'énergie minimisée. Nous constaterons aussi que le modèle de contrainte de forme est malgré tout inapte à surmonter les erreurs locales de forme contenues dans la carte. Le chapitre suivant adresse ce problème avec la formulation de nouvelles énergies.

Le chapitre 6 décrit la formulation d'un nouveau critère énergétique de contrainte de forme, alliant énergie linéaire et quadratique. Celui-ci est destiné à s'affranchir des erreurs présentes dans la carte afin d'améliorer le recalage avec l'image. Nous proposons une solution consistant à autoriser certains écarts du contour actif par rapport à la forme de référence. La classe des écarts autorisés est celle des déplacements contraints par le parallélisme et correspond aux erreurs les plus souvent rencontrées dans la carte. Ce modèle est complété d'une énergie quadratique imposant au contour actif des contraintes génériques de rectilinéarité et d'orthogonalité. Les travaux présentés aux chapitres 5 et 6 sont destinés à améliorer le recalage carte-image, i.e. la cohérence des deux représentations. Ceci contribue à accroître la fiabilité de l'analyse de changements décrite dans le chapitre suivant.

Le recalage fin par contours actifs nous a permis de réduire les variabilités entre l'objet cartographique et sa représentation dans l'image. Ces variabilités sont signes de différences relativement mineures entre la carte et l'image, et peuvent être issues ou non de réels changements entre les deux données. Nous exposons dans le chapitre 7 une méthodologie pour l'analyse des changements entre une carte de bâtiments et des données de télédétection multi-sources. L'approche que nous proposons se fonde sur la fusion de traitements de la donnée altimétrique, multispectrale et panchromatique et se scinde en deux étapes. La première fusionne des indices de changement destinés

à valider uniquement la disparition univoque d'un bâtiment. Ces indices sont calculés à partir de méthodes simples utilisant l'image multispectrale et la donnée altimétrique. Dans la majorité des cas où un changement n'a pu être identifié de façon évidente, une seconde phase plus complexe est mise en œuvre afin de caractériser le degré de ressemblance entre l'objet cartographique et sa représentation dans l'image satellitaire panchromatique. Ce traitement tire parti du recalage fin par contours actifs et permet la formulation d'une probabilité de non-changement pour chaque élément de la carte.

Chacun des chapitres 3, 5, 6 et 7 est illustré par des résultats expérimentaux qui sont analysés et comparés aux approches traditionnelles.

1.5 Cadre de la thèse

Cette thèse réalisée en cotutelle est le fruit d'une coopération franco-chinoise entre le LIAMA (Académie des Sciences de Chine), Alcatel Alenia Space et le LIMA (ENSEEIH/INPT). La thèse fut co-financée par Alcatel Alenia Space et le LIAMA/NLPR dans le cadre des projets 01.10 (*Digital map updating on urban areas using high resolution images*) et MOST 863 (*Multi-source data fusion and urban development planning decision support*).

Chapitre 2

Analyse et formulation du problème

2.1 Introduction

Ce chapitre donne un aperçu des travaux récents effectués sur l'interprétation de scènes urbaines à partir de données de télédétection à haute résolution. Cet état de l'art n'a pas la prétention d'être exhaustif et sera orienté vers la description de méthodes utilisant la fusion de données de télédétection et cartographiques afin de faciliter l'analyse difficile des paysages urbains. Nous dresserons ensuite la problématique de la mise à jour de cartes dans un cadre conceptuel général. Ceci nous mènera à formuler plus spécifiquement le problème que nous tenterons de résoudre dans cette étude, à savoir la mise à jour de cartes numériques de bâtiments à partir d'imagerie optique haute résolution en milieu urbain. Enfin, nous exposerons une méthodologie destinée à résoudre cette problématique et basée sur l'atténuation des variabilités exogènes entre la carte et l'image.

2.2 État de l'art sur l'analyse de scènes et la mise à jour de cartes à partir d'images de télédétection haute résolution

2.2.1 Spécificités et stratégies pour la mise à jour de cartes

La mise à jour automatique ou semi-automatique de données cartographiques grâce à des données de télédétection est un problème à l'intersection de plusieurs domaines du traitement d'image et de l'analyse de forme. Elle fait appel à la fois à la détection de changement entre deux représentations multi-temporelles d'une même scène [14, 98] et à la reconnaissance de forme pour l'extraction d'objets (routes, bâtiments...) [63, 45, 8, 72, 74].

2.2.1.1 Comparaison d'images

Les approches de détection de changement en télédétection se basent traditionnellement sur une comparaison d'images se faisant soit au niveau du pixel, de primitives extraites des images - typiquement des invariants - ou de l'objet. Une revue complète des techniques de détection de changement inter-images est disponible dans [98]. Quel que soit l'algorithme utilisé, une phase préliminaire d'alignement des données est indispensable pour leur comparaison ultérieure. Cette mise en correspondance est à la fois géométrique et radiométrique.

Dans le cas d'images acquises avec des prises de vue peu différentes, une transformation affine suffit à la superposition géométrique des données ; dans le cas contraire, des transformations non globales sont alors estimées comme le flot optique [10]. Le choix des primitives dérivées de l'image à mettre en correspondance est délicat. En effet, la qualité de l'extraction de celles-ci conditionnera la précision de la fusion des données et ainsi la fiabilité de la méthode de détection de changement.

Dans [14], P. Blanc propose une méthode de mise en correspondance basée sur l'intensité des pixels (corrélation) de précision sub-pixellaire. Dans le cas de prises de vue avec des résolutions sensiblement différentes, Dufournaud *et al.* mettent en correspondance des primitives ponctuelles rendues invariantes par rotation et translation [37]. En remarquant qu'une image à basse résolution est le résultat d'une convolution d'une image haute résolution par un noyau Gaussien de variance $s\sigma$, les auteurs discrétisent l'espace d'échelle s et estiment la meilleure similitude plane directe mettant en correspondance les points extraits de deux images dont la différence en résolution peut atteindre un facteur six.

La correction radiométrique des données consiste à atténuer les différences d'illumination et d'éclairement représentées dans les images afin de rendre leur comparaison ultérieure efficace. La technique la plus usuelle consiste à normaliser les images pour que celles-ci aient la même moyenne et variance radiométrique [67, 118]. D'autres techniques plus sophistiquées basées sur des transformations linéaires de l'éclairement [34, 57] ou la modélisation de l'illumination ont été proposées [76, 127].

2.2.1.2 Comparaison carte-image

La problématique de la comparaison d'une carte avec une image est cependant différente de celle inter-images car la donnée cartographique est une représentation synthétique, de haut niveau sémantique et simplifiée de la réalité. A l'instar de la comparaison d'images pour la détection de changement, la comparaison carte-image comprend une phase préliminaire de fusion des données. Cependant, aucune correction radiométrique ne s'applique puisque la carte n'est pas une image de télédétection. Le problème de superposition d'une carte avec une image satellitaire SPOT-4 a été étudié par M. Roux dans [106]. Une carte de format papier est numérisée puis pré-traitée pour extraire le réseau routier. Celui-ci est ensuite recalé par programmation dynamique sur des primitives linéaires et de type "jonctions" dérivées de l'image. Les cartes actuelles sont généralement issues de Systèmes d'Information Géographique (SIG) et sont de fait numériques. La mise à disposition de telles cartes représente un avantage certain : la donnée cartographique est vectorielle et ne nécessite pas de pré-traitement comme pour les cartes papier numérisées ; la carte est numérique et géocodée, son interfaçage avec des images de télédétection est ainsi facilité. En particulier, si l'image est aussi géocodée, la superposition des données préalable à la détection de changement est directe.

A ce jour, la quantité de travaux menés sur la comparaison de données cartographiques et de télédétection en vue d'une mise à jour de cartes reste bien moindre que celle relative à la détection de changement entre deux ou plusieurs images. Alors que les recherches menées sur la mise à jour de cartes grâce à l'imagerie optique haute résolution restaient marginales dans les années quatre-vingt-dix [54, 110], elles tendent à se multiplier comme en témoignent les programmes d'agences européennes pour la cartographie : projet ATOMI en Suisse [38, 6, 133, 134], révision de carte ATKIS en Allemagne [124, 122, 123, 42] et TOP10DK aux Pays-Bas [62, 83, 61]. Parmi les travaux déjà effectués sur le sujet de la mise à jour de données cartographiques, on peut distinguer ceux tirant parti de la connaissance *a priori* spécifique dérivée de la carte afin de faciliter le traitement, de ceux ignorant cette information de haut niveau. Dans ce dernier cas, des approches ascendantes sont souvent utilisées : extraction préalable des primitives pertinentes de l'image, puis raisonnement à plus haut niveau en émettant des hypothèses sur l'appartenance à la carte de l'objet extrait dans l'image. D'une façon générale, le problème d'une telle approche est la difficulté d'extraire exhaustivement et correctement toutes les primitives recherchées de l'image. Ainsi, celles qui sont manquantes pourraient être interprétées à tort comme un changement lors de la comparaison avec la carte. Il est aussi incertain d'extraire des primitives correspondant à un objet de l'image dont la nature est cohérente

avec ceux symbolisés dans la carte. S. Servigne est confrontée à ces problèmes dans le cadre de la mise à jour de données cadastrales de la ville de Padoue à partir d'images aériennes [110]. Le résultat d'une segmentation (bas-niveau) des images est utilisé pour tenter de le mettre en correspondance avec les données cartographiques. Dans un premier temps, une technique de relaxation est utilisée pour tenter d'associer un élément de la base cadastrale à sa représentation segmentée dans l'image. Les résultats sont peu satisfaisants car la segmentation produit de nombreux contours, souvent fragmentés, pour un même objet. Afin d'améliorer la mise en correspondance, une autre méthode basée sur les régions est appliquée injectant de la connaissance *a priori* relative à la texture des objets. Une base de données de textures est apprise manuellement sur l'image et recense neuf différents types : terrain, culture, macadam, eau, végétation, toit, façade, ombre, piéton. La comparaison des objets cartographiques et de l'image est finalement effectuée au regard de critères de ressemblance de forme, texture et surface. O. Jamet dans [54] justifie le choix de ne pas utiliser l'information de la carte puisqu'elle est non récente et parfois peu fiable et risquerait ainsi de biaiser le processus de détection de changement. Il propose alors une méthode d'appariement de primitives segments à partir d'un couple stéréoscopique d'images aériennes qui sont comparées aux vecteurs d'une carte numérique du bâti urbain (IGN, BDTopo). Cette comparaison formalisée dans le cadre de la théorie des croyances de Dempster-Shafer [112] permet de modéliser l'incertitude associée à la carte.

La connaissance *a priori* dérivée de la carte est cependant un avantage indéniable pour améliorer l'extraction d'objets de scènes urbaines ou pour la mise à jour de données cartographiques [35, 36, 120, 122, 124, 48, 3, 123, 4, 9, 62, 12, 61, 83, 42]. Les scènes urbaines imagées à haute résolution contiennent une multitude de détails et autres *artefacts* qui rend leur analyse automatique ou semi-automatique difficile (ombres, distorsion due à la perspective de la prise de vue, occlusions). L'information spécifique dérivée de la carte renseigne sur l'emplacement de l'objet à retrouver dans l'image, ce qui limite à la fois l'espace de recherche et le taux de fausses alarmes. Outre la connaissance spécifique issue de la carte, la fusion de données auxiliaires de télédétection est aussi la tendance pour l'analyse de scènes urbaines à haute résolution. Ces données supplémentaires se révèlent utiles pour lever les ambiguïtés relatives à l'interprétation de la scène (images multispectrales, infra-rouges, altimétriques, couples stéréoscopiques) [124, 122, 48, 123, 47, 46, 62, 84, 83, 61]. Dans la suite de ce chapitre, nous proposons d'exposer les travaux récents sur l'extraction d'objets et la mise à jour de cartes à partir d'images haute résolution en milieu urbain selon ces deux angles d'approche, à savoir la fusion de données de télédétection et l'utilisation de connaissance *a priori*, en particulier celle spécifique issue de cartes existantes.

2.2.2 Quelques travaux sur la fusion de données de télédétection pour l'analyse de scènes à haute résolution

2.2.2.1 Fusion d'images optiques

“La fusion de données constitue un cadre formel dans lequel s'expriment les moyens et techniques permettant l'alliance des données provenant de sources diverses. Elle vise à l'obtention d'information de plus grande qualité ; la définition exacte de « plus grande qualité » dépendra de l'application” [121]. Dans le cadre de la fusion d'images optiques haute résolution, nous pouvons citer les méthodes visant à améliorer la résolution spatiale d'images multispectrales à partir d'images panchromatiques de meilleure résolution. Une revue complète de ces techniques de fusion est disponible dans [125]. Le gain en qualité est alors une description plus complète de la scène, alliant une haute résolution géométrique avec de nombreuses signatures spectrales, à même de faciliter l'analyse et l'extraction d'objets. Ce gain est néanmoins entaché d'artefacts de reconstruction de l'image fusionnée qui se manifestent par des biais dans le domaine spatial et spectral. Une étude comparant les performances de certaines de ces techniques a été menée par F. Laporterie dans [64]. Dans [42], M.

Gerke fait l'usage d'une image satellitaire Ikonos résultant d'une fusion d'une image panchromatique et multispectrale (*Pan-sharpening* ou *PXS*) pour la mise à jour de cartes du réseau routier ATKIS. Les informations multispectrales permettent de caractériser efficacement la route (grâce à la bande proche infra-rouge) tout en satisfaisant les contraintes de précision topographique de la carte.

Notons que la fusion de données optiques ne concerne pas uniquement les images de résolutions différentes. En effet, des images panchromatiques et stéréoscopiques de résolutions identiques sont aussi utilisées pour faciliter l'analyse de scène, en particulier au niveau du bâti [108, 81, 78, 53]. Dans [78], Nevatia *et al.* proposent une approche ascendante pour la détection de bâtiments à partir d'images aériennes : des primitives linéaires sont premièrement extraites d'une image, et sont ensuite associées pour formuler des hypothèses de toits en forme de parallélogrammes. Les hypothèses sont vérifiées par la détection des ombres et des murs du bâti. Les meilleures hypothèses sont retenues et fusionnées avec celles émanant de l'analyse d'autres images. Une revue complète d'extraction et de reconstruction 3D de bâtiments est disponible dans [63, 45, 72, 8]

Remarquons enfin que des images issues de capteurs optiques de même résolution et ayant des signatures spectrales différentes sont aussi utilisées. La donnée de mesures multispectrales est particulièrement intéressante pour la classification des objets de la scène. Dans les travaux de V. Walter [122], des images multispectrales aériennes (RVB, 2 m/pixel) sont utilisées pour une classification supervisée par maximum de vraisemblance des zones agricoles, urbaines, routières et de plans d'eau. L'information multi-bandes couplée à celle de texture des images permet d'obtenir des résultats encourageants sur des images rurales et péri-urbaines. Nous verrons en section suivante que pour les milieux purement urbains, une donnée altimétrique supplémentaire s'avère nécessaire pour une discrimination efficace du bâti.

2.2.2.2 Fusion d'images optiques et altimétriques

La mise à disposition d'un modèle numérique de surface (MNS) représentant l'altitude des objets de la scène est particulièrement intéressante pour l'analyse des milieux urbains. Le MNS est généralement le fruit d'une reconstruction 3D à partir d'au moins deux images, ou d'une mesure directe de la réflexion d'un signal laser (Lidar). Un MNS permet de distinguer les structures aériennes du sol, facilitant l'extraction du bâti. Cette dernière est néanmoins rendue délicate lors de la présence de végétation qui se confond avec les bâtiments sur l'image altimétrique.

Dans [48], N. Halla et V. Walter filtrent la végétation d'un MNS Lidar par une analyse multi-échos : le signal laser traversant la canopée engendre deux échos. L'écho le plus rapide correspond au feuillage de la végétation alors que le second est relatif au sol. En considérant uniquement le second écho, le MNS Lidar généré contient très peu de végétation aérienne et limite de façon drastique les ambiguïtés entre la canopée et le bâti. Dans une seconde phase, les auteurs poursuivent les travaux de [122] en appliquant une classification par maximum de vraisemblance de la scène : le MNS est alors considéré comme la quatrième bande d'une image multispectrale (RV et proche infra-rouge). L'utilisation conjointe du MNS avec la bande proche infra-rouge permet une classification performante du bâti et de la végétation. La donnée du MNS permet aussi d'estimer les zones d'ombres de la scène et d'éviter ainsi leur mauvaise classification.

Dans [83], des images couleur (RVB, 0.5 m/pixel) et un MNS normalisé LIDAR sont les données d'entrée d'un algorithme de classification supervisée du bâti. Une première étape consiste à scinder la classe apprise des bâtiments en sous-groupes de caractéristiques spectrales homogènes (algorithme ISODATA [5]). Dans un second temps, une classification basée sur la distance de Mahalanobis est

menée sur les données de télédétection. Le résultat de la classification est ensuite comparé à une carte numérique du bâti pour la détection de changement.

Dans [47] des contours actifs de type *snakes* sont utilisés pour segmenter des bâtiments représentés dans une image Ikonos (1 m/pixel) et un MNS LIDAR. Les contours actifs sont initialisés grâce à une détection grossière des bâtiments à partir du MNS, ils évoluent ensuite vers les hauts gradients de l'image satellitaire tout en étant contraints par des forces de gonflage dérivées du MNS [27].

2.2.3 Utilisation de contraintes issues de connaissance *a priori*

Nous opposons dans les sections suivantes les travaux sur l'extraction d'objets ou la mise à jour de cartes selon qu'ils sont basés sur la connaissance générique ou spécifique (carte) d'un paysage étudié. Une revue complète des différents types de connaissance *a priori* utilisés en télédétection est disponible en [9, 35].

2.2.3.1 Connaissance générique

La connaissance générique renseigne sur des propriétés générales relatives à la radiométrie et/ou la géométrie d'un objet à extraire dans l'image. Elles sont issues du raisonnement humain sur la base de connaissances apprises et permettent le plus souvent de définir des modèles d'objets (typiquement les routes et les bâtiments). Des exemples de connaissance générique sont :

- Les routes ont des bords parallèles, une largeur constante et une faible courbure excepté aux intersections.
- Les routes ont une radiométrie homogène et plus claire que leur environnement.
- Les bâtiments ont des bords rectilignes et des coins.
- Les bâtiments ont une altitude supérieure au sol et y projettent des ombres.

Cette connaissance est dérivée d'observations et se veut donc empirique. Péteri et Ranchin utilisent dans [95, 97, 96] des contours actifs de type *snake* ([58], cf. section 4.2.1.3) pour extraire le réseau routier à partir d'une image satellitaire Ikonos représentant un milieu urbain dense avec une résolution de un mètre. L'intérêt du contour actif est double : il permet d'avoir *in fine* une extraction non fragmentée du réseau routier ; il incorpore aussi facilement une information de haut niveau telle que des contraintes géométriques. Notamment dans [96], une contrainte de parallélisme est utilisée afin de segmenter les bords de la route. Cette contrainte géométrique issue de connaissance générique permet une segmentation plus robuste. Une analyse multi-échelle accroît la robustesse de la méthode vis-à-vis du niveau de détail de l'image.

M. Rochery *et al.* utilisent également le parallélisme dans [101, 100, 102] pour l'extraction de routes par contours actifs. La contrainte est formulée grâce à l'introduction d'une énergie quadratique qui permet une interaction à longue distance des points du contour actif. Contrairement aux travaux de Péteri [95, 97, 96], cette approche n'est pas sensible à l'initialisation et incorpore naturellement la notion de réseau. Cependant, les résultats sont réalisés à partir d'images péri-urbaines à moyenne résolution.

Dans [84], H. Oriot utilise des *snakes* statistiques afin de segmenter les bâtiments d'un couple d'images stéréoscopiques haute résolution. Une fonction de coût est définie de façon à favoriser l'intégration de régions de fortes disparités à l'intérieur du contour. L'extraction du bâtiment est affinée par l'information de gradient des images et par une contrainte générique favorisant les angles à 90°

et 180°. Cette approche est semi-automatique dans la mesure où l'initialisation du contour actif est manuelle. L'évolution du contour minimisant une fonction de coût (approche non variationnelle) est basée sur l'insertion, mise à jour et suppression des nœuds du *snake*.

Dans [119], Vinson *et al.* utilisent des modèles déformables rectangulaires afin de segmenter les bâtiments d'une ortho-image. Une pré-détection des bâtiments par segmentation d'un modèle numérique de surface sert à l'initialisation du modèle qui convergera vers les lieux de haut gradient de l'image, i.e. les bords des bâtiments. A la convergence, les paramètres du modèle rectangulaire (barycentre, longueur, largeur, orientation) sont estimés. Les résultats des extractions rectangulaires peuvent être ensuite fusionnés. Le choix d'un modèle rectangulaire reflète la connaissance empirique relative à la forme des bâtiments la plus souvent rencontrée.

Alors que la connaissance *a priori* générique est largement utilisée pour aider l'extraction d'objets, elle est dépendante d'aspects sociétaux et culturels et peut ainsi varier à une échelle nationale ou mondiale (à titre d'exemple le réseau routier des vieux quartiers de Pékin n'a rien à voir avec celui de Manhattan). Cette information n'est donc pas généralisable vis-à-vis de l'endroit géographique représenté par des données de télédétection. Elle se doit d'être complétée par une connaissance plus dédiée et spécifique et moins contrainte par la connaissance empirique humaine.

2.2.3.2 Connaissance spécifique

La connaissance spécifique est typiquement dérivée de données cartographiques. La carte renseigne sur la localisation, la forme et la nature des objets représentés par les données de télédétection. C'est un *a priori* fort qui rend plus robuste la tâche d'extraction ou de mise à jour de cartes. L'utilisation de la carte se fait principalement selon trois aspects. Le premier concerne la réduction de l'espace de recherche d'un objet recherché dans les données de télédétection. Ceci permet de réduire les temps de calcul ainsi que le taux de fausses alarmes. Le second regarde la définition de zones d'apprentissage pour la classification supervisée de scènes. Alors que ces zones sont traditionnellement définies manuellement, la donnée cartographique permet d'en automatiser la saisie. Enfin, la carte peut être utilisée pour enrichir une base de connaissance avec des informations spécifiques de l'image analysée.

Définition de zones d'apprentissage

Les travaux de B.P. Olsen [83], N. Haala [48] et V. Walter [122] sont tous basés sur la classification supervisée d'images multi-capteurs (cf. section 2.2.2). Ces méthodes de classification nécessitent des zones d'apprentissage définies manuellement par des experts. Ces auteurs utilisent une carte existante pour définir ces zones afin de rendre le système automatique. Cependant, l'utilisation d'une telle connaissance dérivée de la carte fait l'hypothèse que le nombre d'objets cartographiques ayant changé est faible devant le nombre total d'objets. En effet, un nombre de changements trop important fausserait la classification subséquente. Pour éviter les erreurs d'apprentissage aux frontières des objets cartographiques, A. Busch réduit l'espace d'apprentissage par filtrages morphologiques [16]. A la différence des travaux de [83, 48, 122], ceux de A. Busch concernent des images optiques de moyenne résolution. Dans le cas d'images très haute résolution telles que celles utilisées par B.P. Olsen, une distorsion due à la perspective est visible au niveau des bâtiments. La carte étant en géométrie orthoscopique (vue au nadir), les données ne sont pas compatibles pour une superposition préalable à l'apprentissage. Néanmoins, la carte TOP10DK utilisée est à trois dimensions : la connaissance de la prise de vue des images de télédétection permet alors de projeter efficacement chaque bâtiment cartographique sur sa représentation homologue dans l'image.

Diminution de l'espace de recherche

Dans [107], M. Roux et H. Maître font usage d'une carte papier numérisée à l'échelle 1 :25,000 de la ville de Paris pour améliorer la reconstruction 3D de bâtiments. La carte, dans laquelle les blocs de bâtiments sont isolés, permet de restreindre l'espace de recherche de points homologues obtenus par corrélation pour le calcul de disparités à partir d'images aériennes stéréoscopiques. L'espace de recherche étant diminué, les erreurs d'appariement sont moins courantes et une fenêtre de corrélation plus petite est utilisée pour l'obtention de résultats plus fins.

Dans [115, 35], les auteurs utilisent une carte routière comme initialisation d'un processus de détection de changement avec une image aérienne haute résolution plus récente. Dans [35], M.E. de Gunst effectue une comparaison entre les données cartographiques et l'image par corrélation entre la carte et un profil radiométrique issu de l'image. Lorsque la corrélation est satisfaisante, il n'y a pas de changement. Lorsqu'elle est faible, une deuxième phase du processus basée sur un raisonnement contextuel est déclenchée : selon l'environnement où un changement a été détecté, une liste d'objets susceptibles d'être observés dans l'image est dressée par ordre de priorité. Chaque type d'objet candidat déclenche un type de segmentation particulier dont le résultat est confronté par tests d'hypothèses à des modèles définis *a priori*. Si tous les tests sont infructueux, la fenêtre de recherche est agrandie, les paramètres des méthodes d'extraction sont adaptés. Dans [115], C. Steger propose une analyse de profil radiométrique de la route guidée par la carte afin de valider le non-changement des parties qui la composent. Lorsque la carte est peu précise, des contraintes génériques de parallélisme et de rectilinéarité de la route sont utilisées pour s'affranchir des imprécisions. Les éléments de la carte peuvent être rejetés, vérifiés ou ambigus. Un second algorithme détecte les nouvelles routes en inspectant le voisinage des parties validées de la carte lors de la phase précédente.

Dans le cadre du projet ATOMI [38], les auteurs utilisent des sources redondantes d'information (carte numérique au 1 :25,000, modèle numérique de surface (MNS), images multispectrales) pour l'extraction de routes [133] et la reconstruction de bâtiments [79]. La carte est utilisée pour restreindre l'espace de recherche. Dans [133], des images aériennes RVB stéréoscopiques ainsi qu'un MNT sont employées pour l'extraction de routes. Dans [79], des données vectorielles de zones résidentielles sont couplées à une pré-détection de bâtiments issue du MNS. Ces données ainsi que les images multispectrales sont simultanément utilisées pour une classification non supervisée du bâti dans l'image. Une seconde phase consiste à recalibrer les objets cartographiques sur les résultats de la classification. Les échelles de la carte et de l'image étant différentes, les bâtiments symbolisés sont souvent éloignés de leur représentation dans l'ortho-image. Une transformation affine estimée par des scores empiriques permet le recalage carte-image et une reconstruction 3D subséquente.

Agouris *et al.* [4] ont développé une méthode permettant de détecter les variations géométriques d'une route symbolisée dans une carte et représentée dans une image de télédétection plus récente grâce aux *snakes*. L'originalité de leur approche est l'introduction d'un nouveau terme d'énergie de contrainte externe qui dépend du degré d'incertitude associé à la précision planimétrique de la carte. Cette incertitude est calculée à partir d'une image de télédétection dont la date d'acquisition correspond avec celle de la réalisation de la carte : le contour actif est initialisé grâce aux objets cartographiques et évolue sans contrainte externe dans l'image. Selon les variations d'énergie totale, et en fonction de sa valeur finale, un degré d'incertitude est alloué à chaque nœud du *snake*. Ce degré est attribué selon des règles de logique floue inspirées de la linguistique. Ainsi, une forte variation d'énergie ainsi qu'une énergie finale élevée pour un nœud considéré conduira à attribuer une forte incertitude au point de l'objet cartographique sur lequel était initialisé ce nœud. Cette incertitude

permet aux auteurs d'exprimer une force de rappel similaire à celle d'un ressort fixé à la position du nœud du contour actif dans son état initial. Finalement, la contrainte externe est imposée aux contours actifs qui évoluent sur l'image la plus récente. Qualitativement, la méthode mise en œuvre rend l'énergie de contrainte externe proportionnelle à l'allongement d'un ressort dans un certain voisinage, puis elle devient constante en dehors de celui-ci. L'étendue du voisinage et la raideur du ressort dépendent de l'incertitude précédemment déterminée. Plus l'incertitude est faible, plus le rayon d'action de l'énergie de contrainte est petit avec une forte raideur associée au ressort : le degré de véracité de la carte étant élevé, s'il y a réellement un changement, les gradients des contours de la nouvelle route devraient être suffisamment intenses pour vaincre la contrainte. Au contraire, si l'incertitude est élevée (grand rayon d'action, faible raideur), il ne semble pas justifié d'attirer le contour actif vers sa position initiale définie par la carte. Cette méthode paraît efficace sur des cas simples d'images optiques de résolution moyenne : la route est déviée manuellement sur des portions de la route, simulant ainsi un faible changement. La méthode ne prend pas en compte l'apparition de nouvelles routes intersectant ou non l'ancienne.

Une approche similaire à [4] est introduite dans [12] par Bentabet *et al.* pour l'extraction de routes à partir d'images SAR. Des *snakes*, initialisés grâce à une carte, évoluent pour segmenter plus fidèlement la représentation de la route dans l'image. Les contours actifs utilisés sont attirés par les zones de gradient élevé de l'image. Ils sont par conséquent sensibles au bruit des données de télédétection. Les auteurs proposent un filtrage du *speckle* qui préserve l'information de contour. Considérant le problème du paramétrage des snakes, les auteurs estiment le coefficient de rigidité du contour actif à partir de la courbure moyenne de la route calculée à partir de la carte.

Dans [3, 2], Agouris *et al.* proposent une méthode de détection de changement au niveau des bâtiments symbolisés dans une carte grâce à la donnée de deux images aériennes diachroniques. L'une des images est acquise à la même date que la réalisation de la carte alors que l'autre est plus récente. Chaque objet cartographique est projeté sur l'image cohérente avec la carte. Le sous-échantillonnage du polygone cartographique considéré fournit une liste de points dans l'image ancienne dont la présence sera vérifiée dans l'image plus récente. La détection de changement est effectuée par l'évaluation d'un score de corrélation originellement développée par Grün dans [44]. Le but de cette méthode est de trouver itérativement les six paramètres d'une transformation affine permettant de minimiser, au sens des moindres carrés, les différences radiométriques des pixels de deux fenêtres appartenant à chacune des images. Cette approche permet une très bonne précision de mise en correspondance, cependant, elle rencontre souvent des problèmes de non-convergence d'estimation des paramètres. Les auteurs de [3] résolvent ces problèmes en donnant plus de poids aux contours de bâtiments qu'au reste de l'image. L'analyse du profil radiométrique des contours dans l'image de référence, guidée par la carte, permet de classer les contours en trois catégories : contour simple (transition entre deux niveaux radiométriques), double et triple. Ces profils sont ensuite mathématiquement modélisés par une combinaison de Gaussiennes dont les maxima sont localisés au niveau des contours ; ils sont finalement injectés dans le modèle d'estimation des paramètres de la transformation affine et améliorent sensiblement la technique. L'avantage de cette méthode est de pouvoir détecter des changements partiels d'un bâtiment considéré. La difficulté est la nécessité de posséder une image acquise à la même date que la carte ; la méthode repose aussi sur une parfaite superposition entre la carte et l'image ce qui n'est souvent pas le cas à cause des erreurs dues aux imprécisions cartographiques.

Dans [65], Leitloff *et al.* proposent d'extraire des files de véhicules à partir d'une image panchromatique Quickbird. Cette méthode est basée sur l'extraction de primitives linéaires dérivées du gradient de l'image. La connaissance d'une carte numérique permet aux auteurs de filtrer de

nombreuses primitives n'ayant aucun lien avec les voitures : en faisant l'hypothèse que celles-ci sont alignées avec la route, ils appliquent un filtrage morphologique ne gardant que les primitives parallèles au réseau routier. Ce pré-traitement diminue la complexité calculatoire et les erreurs de la reconnaissance ultérieure de files basée sur l'analyse de profils radiométriques.

Enrichissement d'une base de connaissance

Dans [133], la connaissance *a priori* est composée de règles concernant l'agencement des routes, de primitives caractéristiques telles que le marquage au sol et de cartes existantes. Cette dernière information spécifique est insérée dans un système à base de connaissance. Cette base se verra enrichie des caractéristiques radiométriques de l'image aux endroits où la cohérence avec la carte est vérifiée. Une classification non supervisée basée sur les connaissances acquises extrait le réseau routier.

2.2.4 Synthèse

A la vue de cette analyse bibliographique, nous constatons que la fusion de données de télédétection améliore sensiblement l'interprétation de paysage, et en particulier celle des scènes urbaines qui sont difficiles à analyser. L'injection de connaissance spécifique et *a priori* permet d'accroître les performances de reconnaissance d'objets et de mise à jour de cartes, évitant ainsi les écueils des méthodes ascendantes traditionnelles. Cette information spécifique est le plus souvent utilisée pour restreindre l'espace de recherche, définir des zones d'apprentissage ou enrichir spécifiquement des bases de connaissance. Cependant, la prise en compte de l'information géométrique dispensée par la carte est plus marginalement employée. La difficulté, en contre partie, tient naturellement au fait que la carte et l'image ne contiennent pas exactement la même information (en raison soit de changements, soit d'erreurs dans la carte). Une comparaison carte-image doit par conséquent savoir s'écarter de cette connaissance *a priori* afin de retrouver l'objet tel qu'il est dans l'image. Dans la section suivante, nous clarifions ces aspects de manière générale en formulant la problématique de mise à jour de cartes que nous tentons de résoudre.

2.3 Position du problème de la mise à jour de cartes

Pour poser et formuler le problème de la mise à jour de cartes à partir de données de télédétection, nous proposons d'employer et étendre le cadre conceptuel exposé dans [55] et repris dans [14] pour analyser la détection de changement entre des données multi-dates et multi-sources. Ce cadre est intéressant dans la mesure où il se prête bien à la problématique que nous adressons, mais nécessite d'être adapté.

2.3.1 Cadre conceptuel général

2.3.1.1 Définition des univers relatifs à la télédétection

Ce formalisme consiste à considérer l'univers objectif U_{objectif} contenant l'ensemble exhaustif des connaissances que l'on peut avoir sur une scène considérée. Cette connaissance objective nous est inaccessible dans sa totalité. Elle se manifeste sous la forme de phénomènes physiques appelés *observables*. La connaissance de toutes les observables ne permet pas d'en déduire exhaustivement U_{objectif} car la fonction de transcription physique de cet univers vers $U_{\text{observable}}$ n'est pas inversible (problème mal posé). Ainsi, la seule façon de connaître partiellement U_{objectif} revient à mesurer une ou plusieurs observables par le biais d'un appareil de mesure. Il est cependant à noter que pour une même scène considérée, les observables varient selon certains facteurs. Dans le cas de l'imagerie

optique ils sont par exemple :

- L'éclairement qui varie en fonction de la position du soleil par rapport à la terre. Selon l'illumination, les ombres portées dues à la présence de bâtiments varieront.
- La géométrie d'observation (angle zénithal de visée). Dans le cas spécifique d'un milieu urbain dense, cette géométrie explique la présence ou non d'occlusions dues aux bâtiments en fonction de l'angle de prise de vue.
- Les conditions météorologiques.
- Les phénomènes parasites de réflexions multiples (par l'atmosphère, le sol ou les objets composant la scène) viendront perturber et biaiser l'observable que l'on désire véritablement mesurer.

Le choix du moyen d'observation dépend bien sûr de l'observable à mesurer. Dans le cas de l'imagerie optique, nous choisissons d'observer les ondes électromagnétiques émises par le soleil et ensuite réfléchies par la scène sur une fenêtre spectrale du domaine visible et infra-rouge proche (capteur passif). La mesure faite de cette observable appelée *observation* est aussi entachée d'erreurs propres à l'instrument de mesure physique. A titre d'exemple, le bruit du capteur de mesure, ou un mauvais calibrage de ce dernier viendra altérer l'information finale qui nous sera disponible et à partir de laquelle nous espérons remonter à la connaissance de $U_{objectif}$.

Nous comprenons dès lors que $U_{observation}$ n'est qu'une représentation partielle et biaisée de $U_{objectif}$. Ces altérations ne sont pas forcément stables dans le temps. Ainsi lors de la comparaison d'images diachroniques pour la détection de changement, ces artefacts introduisent des variabilités exogènes indésirables car elles sont susceptibles d'être interprétées à tort comme un changement. S'affranchir des variabilités exogènes pour détecter les changements effectifs s'étant produits dans $U_{objectif}$ est une tâche difficile qui est au cœur de tout problème de détection de changement entre deux images.

2.3.1.2 Définition de $U_{interprétation}$

Les travaux menés par P. Blanc dans [14] tentent d'adresser cette question de manière générale par un recalage géométrique et radiométrique des données. Cependant, dans [14] les données comparées pour la détection de changement appartiennent toutes à $U_{observation}$. Notre cas a ceci de différent que nous comparons une observation avec une carte. Une carte n'est définitivement pas une observation, elle est le fruit d'une interprétation, schématisation, simplification d'une observation. La carte, qui appartient à $U_{interprétation}$, est aussi susceptible de contenir des erreurs inhérentes à la transcription cartographique de $U_{observation}$ vers $U_{interprétation}$. Ces dernières sont dues aux erreurs d'interprétation de l'opérateur ayant réalisé la carte (ou d'un processus automatique s'il existe). Les erreurs de photo-interprétation sont de deux sortes. L'une concerne une erreur d'interprétation quant à la nature de l'objet manuellement extrait. Par exemple, une parcelle de gazon ou d'espace de stationnement pourrait être interprétée comme le toit d'un bâtiment observé au nadir tant leurs caractéristiques sont susceptibles d'être similaires (zone radiométriquement homogène dont les contours sont caractérisés par de forts gradients). Ces erreurs sont très rares car les cas ambigus de photo-interprétation sont le plus souvent vérifiés sur le terrain. Les erreurs les plus courantes concernent la délimitation des contours des objets. Dans [80], N. Nideröst illustre ces erreurs de délimitation qui varient selon l'interprétation des cartographes (figure 2.1).



FIG. 2.1 – Erreurs de délinéation des contours : différentes segmentations (en noir) du même objet (vérité terrain en blanc) par plusieurs photo-interpréteurs. Le schéma est extrait de [80].

A ceci s'ajoute la phase de simplification des contours des objets insérés dans la carte qui, même si ce n'est pas une erreur d'interprétation, représente un écart avec la réalité des objets observés dans l'image. Dans le cas des bâtiments, certains d'entre eux ont une forme complexe qu'il est inutile de transcrire avec précision dans la carte pour l'échelle considérée. La forme est ainsi simplifiée de façon plus ou moins arbitraire par l'opérateur. Il est également difficile de détourner chaque bâtiment lorsqu'ils sont agglomérés entre eux, nombreux et de petite taille. Une solution retenue consiste à inclure un groupe d'objets dans une seule entité cartographique. La figure 2.4 illustre cet effet de généralisation. Enfin, la cohérence des échelles de la carte et de l'image pour la détection de changement est un facteur essentiel. En effet, une carte de faible échelle superposée à une image de haute résolution exhibera des effets significatifs de généralisation et d'imprécision de la localisation, qui rendra *in fine* la détection de changement ardue voire incohérente (figure 2.2).

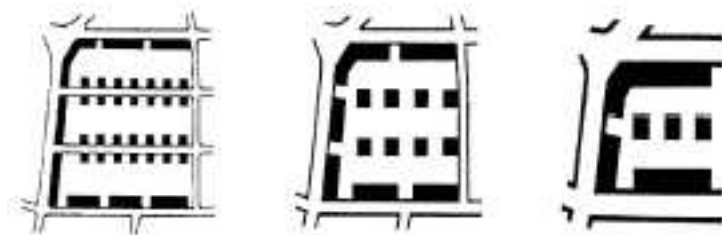


FIG. 2.2 – Effet de généralisation dû à l'échelle. La même scène est représentée à différentes échelles : 1 :25,000 / 1 :50,000 / 1 :100,000 (de gauche à droite). Le schéma est extrait de [80].

Une méthode fiable de détection de changement carte-image se doit donc de palier les variabilités exogènes présentes dans les univers $U_{observable}$, $U_{observation}$ et $U_{interprétation}$ afin de ne pas les interpréter comme un changement effectif (table 2.1).

2.3.2 Problématique adressée dans la thèse

2.3.2.1 Analyse de changements carte-image considérée dans la thèse

Dans ce travail de thèse, nous nous intéressons à la mise à jour des bâtiments de cartes numériques à deux dimensions à partir d'une image satellitaire panchromatique haute résolution plus récente. La maintenance de données cartographiques revêt principalement quatre aspects :

1. La validation/suppression des objets cartographiques qui sont détectés comme présents/absents dans l'image satellitaire plus actuelle.
2. L'insertion dans la carte de nouveaux objets détectés dans l'image satellitaire plus actuelle.
3. L'amélioration de la précision planimétrique des objets cartographiques existants dans l'image satellitaire plus actuelle.

Univers		Artefacts sources de variabilités exogènes
$U_{observable}$	Géométrie d'éclairement	Ombres dues aux objets de la scène en fonction de la hauteur du soleil
	Géométrie d'observation	Occlusions différentes en fonction de l'angle de prise de vue
$U_{observation}$		Bruit du capteur, erreurs de quantification (numérisation), erreurs de calibrage.
$U_{interprétation}$		Erreurs d'interprétation sur la nature de l'objet cartographique
		Inexactitudes de l'extraction de l'objet
		Simplifications, généralisations

TAB. 2.1 – Classification des artefacts urbains en fonction des différents univers.

4. L'enrichissement de la carte par insertion de nouveaux attributs. L'ajout d'information 3D à une carte 2D en est un exemple.

Dans cette étude nous explorerons les points 1,3 et 4 de cette liste. Ainsi nous ne détecterons pas les nouveaux bâtiments pour les ajouter à la carte. Nous nous limiterons à une détection de changement *carte-image* consistant à considérer individuellement chaque bâtiment symbolisé dans la carte et à vérifier s'il existe toujours dans les données de télédétection plus récentes. Si le bâtiment de la carte existe toujours, sa position et sa forme sont susceptibles d'être affinées grâce aux données de télédétection (nous considérons que l'information délivrée par l'image est plus fiable que celle des données cartographiques). Le tableau 2.2 récapitule les différents types de changements susceptibles d'être rencontrés. Le changement de type *d* est marginal et ne sera pas considéré, celui de type *h* est

Nature du changement	Nouvelle situation
a. Bâtiment totalement détruit	Sol nu
b. Bâtiment totalement détruit	Bâtiment quelconque
c. Bâtiment totalement détruit	Bâtiments plus petits intersectant l'ancienne empreinte au sol
d. Bâtiment totalement détruit	Bâtiment avec la même empreinte au sol
e. Bâtiment totalement détruit	Objets $\neq$ bâti (routes, arbres, mobilier urbain, ...)
f. Bâtiment partiellement détruit	Bâtiment partiellement détruit
g. Bâtiment avec de nouvelles extensions	Bâtiment partiellement étendu
h. Bâtiment en cours de construction	Bâtiment avec la même empreinte au sol mais avec une altitude plus élevée

TAB. 2.2 – Changements possibles d'un bâtiment dans $U_{objectif}$

difficilement détectable puisque que la carte qui sera utilisée contient une information uniquement à deux dimensions. Il sera par conséquent écarté.

Remarquons que dans la mesure où nous considérons la présence potentielle d'erreurs locales de forme dans la carte, les cas *f* et *g* représentent des problèmes mal posés de la détection de changement. En effet, il est impossible de savoir si un changement partiel carte-image de faible

amplitude est dû à une erreur de délinéation de la carte ou à un changement effectif dans le monde réel (extension, destruction partielle du bâti). Ainsi, nous ne considérerons pas ces changements partiels. Néanmoins, nous développerons une méthodologie qui tentera de corriger les incohérences carte-image, qu'elles soient dues à des artefacts de la carte ou à des changements réels.

Nous supposerons que les échelles de la carte et des données de télédétection utilisées correspondent, ainsi nous nous affranchissons d'effets de généralisation dus à l'échelle. Cependant les erreurs de délinéation et les effets de généralisation de la carte inhérents à $U_{interprétation}$ seront à prendre en compte (figures 2.3-2.4).



FIG. 2.3 – Superposition d'une carte du bâti (polygones jaunes) avec une image satellitaire. Les flèches blanches pointent des inexactitudes locales de forme de la carte.



FIG. 2.4 – Superposition d'une carte du bâti (polygones jaunes) avec une image satellitaire illustrant l'effet de généralisation : un seul polygone cartographique englobe de nombreux bâtiments contigus.

2.3.2.2 Difficultés des milieux urbains

Nous focalisons notre étude sur les milieux urbains denses. Les difficultés intrinsèques à $U_{observable}$ et $U_{observation}$ sont nombreuses pour ce type de paysage pour lequel l'utilisation d'images de télédétection à très haute résolution fait apparaître des artefacts qui n'existent pas à des échelles moyennes et petites. Ces artefacts peuvent être les ombres dues aux bâtiments et aux voies routières aériennes, une distorsion géométrique importante due aux effets de perspective au niveau du bâti, ou encore les effets d'occlusion. Les distorsions géométriques de l'image inhérentes aux effets de perspective et à l'angle de prise de vue ont pour effet de rendre la localisation planimétrique des objets vus dans l'image peu précise tout en modifiant leur forme. Un moyen de corriger cet artefact est d'orthorectifier l'image mais ce traitement n'a pu être effectué avec nos données. Les ombres sont un problème certain car elles occultent de l'information et créent des contours artificiels indésirables pour tout traitement de photo-interprétation par ordinateur. Dans les travaux de D. Boldo et H. Le Men [15], une méthode est proposée afin d'atténuer les effets d'ombres et d'occlusions du ciel. En milieu urbain, les occlusions sont dues aux structures de haute altitude (bâtiments), cependant il faut faire la distinction entre occlusion observée sur l'image et l'effet d'une occlusion qui n'est pas observée. Une occlusion observée sur l'image engendre l'occultation d'une information (un bâtiment en cache un autre). L'effet d'une occlusion qui n'est pas physiquement observée dégrade l'information en l'occultant partiellement : une rue peut être ombragée à cause d'occlusion du ciel (la zone ombragée n'est pas directement éclairée par le soleil, mais par le ciel environnant ce qui a pour effet d'atténuer la réflexion de l'objet de la scène) tout en étant visible sur l'image. Un autre effet typiquement urbain est la rétro-illumination des façades qui éclairent les objets au sol, introduisant ainsi un biais dans la mesure de leur luminance [15]. Enfin, d'autres difficultés sont liées à l'activité humaine qu'il est possible de déceler sur des images à très haute résolution. Ainsi, une route embouteillée, où la densité de véhicules est très élevée, aura une radiométrie très différente d'une route avec un trafic fluide. Il est toutefois à noter que si ces artefacts sont des problèmes additionnels, ils peuvent aussi apporter une information non négligeable. Ainsi, la détection d'une ombre signifie qu'un bâtiment existe dans son voisinage, la détection de véhicules confirme l'existence d'une route (connaissance générique *a priori*).

2.4 Méthodologie proposée

Notre méthodologie suppose la mise à disposition de deux images satellitaires haute résolution panchromatique et multispectrale (RVB et proche infra-rouge) de la même scène et une carte numérique de bâtiments plus ancienne à deux dimensions. Nous supposons que la carte et les images satellitaires sont initialement superposables (soit par la donnée de points de contrôle soit par un algorithme de recalage global ou encore l'information de géocodage). Ce recalage global n'est pas suffisant pour résoudre les variabilités exogènes propres à chaque bâtiment de la carte.

Nous ferons l'usage d'un modèle numérique de surface (MNS) auxiliaire visant à aider la discrimination du bâti du reste de la scène. Dans notre étude, ce MNS sera généré à partir d'images aériennes stéréoscopiques plus récentes que la carte à mettre à jour.

La méthode que nous proposons pour mener à bien l'analyse de changement avec la carte est la suivante :

1. Optimisation des paramètres de prise de vue des images aériennes et génération d'un MNS par corrélation des images aériennes stéréopaires (chapitre 3).
2. Pour chaque bâtiment symbolisé dans la carte :

- (a) Détection de changements univoques par fusion de données multispectrales et altimétriques. Cette détection de changement est limitée au cas *a* du tableau 2.2. Elle ne peut en aucun cas confirmer un non-changement. Elle est destinée à détecter les cas sans ambiguïtés de changement entre la carte et les données de télédétection (chapitre 7).
- (b) Atténuation des variabilités exogènes survenant dans $U_{interprétation}$ grâce à l'image satellitaire panchromatique. Sont étudiés uniquement les bâtiments qui n'ont pas été détectés comme ayant changé à l'étape (a). Nous utilisons des contours actifs destinés à mettre en correspondance finement chaque bâtiment symbolisé dans la carte avec sa représentation homologue dans l'image (chapitres 4,5,6). Dans le cas d'un bâtiment cartographique absent de l'image (changement), la mise en correspondance échoue et le changement sera ultérieurement détecté en étape (c). Pour effectuer le recalage fin carte-image par contours actifs, nous tirons parti de la connaissance *a priori* et spécifique dérivée de la carte afin de surmonter les difficultés des milieux urbains denses. Cette connaissance est à la fois :
 - i. **Géographique.** Les données cartographiques et de télédétection étant initialement superposées, nous savons où chercher dans l'image pour effectuer le recalage fin. Ainsi, le contour actif sera initialisé près du toit du bâtiment à recaler grâce à l'information géographique contenue dans la carte. Il est important de remarquer que lorsque l'image n'est pas orthorectifiée, la superposition carte-image n'est efficace que pour les bâtiments peu à moyennement élevés (la distorsion due à la perspective visible au niveau des bâtiments élevés éloigne le toit de l'objet cartographique orthoscopique). La technique des contours actifs étant sensible à l'initialisation, notre méthode ne peut donc prendre en compte les bâtiments d'altitude très élevée. Lorsque les données de télédétection sont orthorectifiées, cette limitation disparaît.
 - ii. **Géométrique.** Nous connaissons à l'avance la forme du bâtiment à retrouver dans l'image si aucun changement n'est survenu. Cette information sera insérée comme une contrainte de forme dans le modèle des contours actifs et garantira un meilleur succès de mise en correspondance avec une moindre sensibilité aux artefacts urbains. La fusion d'un MNS dans le modèle des contours actifs ainsi qu'une variation de la contrainte de forme montreront une plus grande capacité du recalage fin vis-à-vis de l'incertitude géographique des bâtiments de la carte. Alors que les variabilités exogènes sont atténuées, la précision des objets cartographiques est susceptible d'être améliorée.
- (c) Détection de changement carte-image panchromatique. Un score de cohérence est calculé entre les segments de l'objet cartographique recalé et ceux extraits de l'image panchromatique. Ce score, combiné à la variation géométrique enregistrée par le recalage des contours actifs, permettra le calcul d'une probabilité de non-changement indiquant la présomption de changement/non-changement pour le bâtiment de la carte considéré (chapitre 7).

Le résultat de la méthode est une probabilité de changement/non-changement pour chaque bâtiment de la carte. Optionnellement, l'altitude dérivée du MNS peut venir enrichir la carte originellement à deux dimensions (phase 2.b.ii, cf. annexe F). Si l'image satellitaire est orthorectifiée, alors la localisation planimétrique des bâtiments est améliorée par recalage fin (phase 2.b).

Les images satellitaires, aériennes et les cartes numériques que nous utiliserons, afin de vérifier expérimentalement l'efficacité de la méthode, sont présentées en annexe A. Les images de télédétection proviennent du capteur embarqué à bord du satellite Quickbird. Les résolutions des images

panchromatiques et multispectrales sont respectivement 0.6 m/pixel et 2.4 m/pixel. L'angle de prise de vue par rapport au nadir varie entre 5° et 9° . La distorsion due à la perspective est donc relativement faible mais correspond néanmoins à la configuration sous-optimale des données non orthorectifiées évoquée en 2.b.i. Ces images satellitaires ainsi que les données aériennes (0.2 m/pixel) représentent toutes la ville de Pékin qui connaît actuellement une forte expansion et restructuration de son tissu urbain.

2.5 Conclusion

Ce chapitre a brossé un état des travaux récents menés sur l'extraction d'objets (routes, bâtiments) et la mise à jour de cartes à partir de données de télédétection haute résolution. Il en ressort que la fusion de données auxiliaires est la tendance actuelle pour résoudre les ambiguïtés des scènes urbaines. Plus récemment, l'utilisation de cartes dans le processus de fusion a montré son efficacité. La connaissance *a priori* apportée par la carte est spécifique et permet soit de restreindre l'espace de recherche, soit de définir des zones d'apprentissage, soit d'enrichir une base de connaissances avec des caractéristiques propres à l'image. Après avoir exposé le problème de la mise à jour de cartes sous l'angle général des univers propres aux observables, observations et interprétations, nous avons formulé plus spécifiquement la problématique de la thèse. Enfin, nous avons proposé notre méthodologie pour l'analyse de changements entre une carte numérique de bâtiments et des images de télédétection plus récentes. Celle-ci est basée sur la fusion de données de télédétection et propose de résoudre le problème des variabilités exogènes entre la carte et les images par la technique des contours actifs afin d'achever une détection de changement carte-image fiable. L'approche proposée se limite à l'analyse des changements *a,b,c,e* du tableau 2.2 et est restreinte aux bâtiments *peu à moyennement élevés* lorsque les données de télédétection ne sont pas orthorectifiées.

Chapitre 3

Génération d'un modèle numérique de surface orthoscopique

3.1 Introduction

Ce chapitre propose de décrire le processus de génération d'un modèle numérique de surface (MNS) orthoscopique à partir de couples d'images aériennes stéréoscopiques. Rappelons que dans la méthodologie générale, ce MNS est une source d'information supplémentaire et discriminante au niveau du bâti qui visera à faciliter la mise à jour de cartes à partir d'une seule image satellitaire. Dans un premier temps, les principes fondamentaux de la stéréovision seront brièvement rappelés, ainsi que l'algorithme de reconstruction 3D ayant servi à la génération d'un MNS non orthoscopique. Cet algorithme est basé sur une technique de corrélation d'images stéréoscopiques qui préserve le contour des objets du paysage [86, 87]. Ce logiciel impose d'avoir une connaissance précise des paramètres de prise de vue des images impliquées dans la génération du MNS. La seconde section de ce chapitre propose une méthode permettant d'affiner la précision de tels paramètres connus *a priori*. Cette approche, qui est la principale contribution exposée dans ce chapitre, fait appel à l'optimisation de fonctionnelles non-linéaires par l'algorithme du *simplex*. Des résultats quantitatifs démontreront l'efficacité de la méthode proposée. Enfin, nous terminerons ce chapitre avec le géocodage, l'orthorectification et la fusion des MNSs générés par la chaîne de reconstruction 3D.

3.2 Stéréoscopie photogrammétrique

La stéréoscopie est relative aux méthodes qui permettent d'obtenir une mesure de relief d'une scène à partir de deux photographies prises sous deux angles de vue distincts. Inspirée du système visuel humain, la stéréoscopie est couramment utilisée en photogrammétrie lors de la constitution de cartes en relief, mais aussi en microscopie optique et électronique.

Dans ce chapitre, nous nous intéressons uniquement à la restitution automatique du relief d'une scène observée par des moyens de télédétection (figure 3.1). Le résultat d'une telle reconstruction est la localisation dans l'espace à trois dimensions de chaque point de la scène vu dans au moins deux images. Dans le cas particulier de la télédétection, les points reconstruits du paysage sont souvent référencés par rapport au terrain, i.e. par rapport à un référentiel cartographique. L'estimation des dénivelés d'un paysage est appelée Modèle Numérique de Terrain (MNT) ou Modèle Numérique de Surface (MNS) selon qu'on représente le relief du sol uniquement ou le sol et ses superstructures (bâtiments, végétation...). A la différence d'un MNS obtenu par imagerie LIDAR, un MNS réalisé par stéréo-photogrammétrie est le résultat d'un traitement complexe de reconstruction, ce qui le rend sous-optimal. En effet, cette reconstruction est sujette aux erreurs présentes dans la connaissance

de la prise de vue (calibration) ainsi qu'à la performance des modèles mathématiques employés. Les auteurs de [7] ont cependant montré que leur reconstruction 3D à partir d'images satellitaires stéréoscopiques Ikonos s'est révélée meilleure que la donnée LIDAR sur une faible étendue du MNS généré.

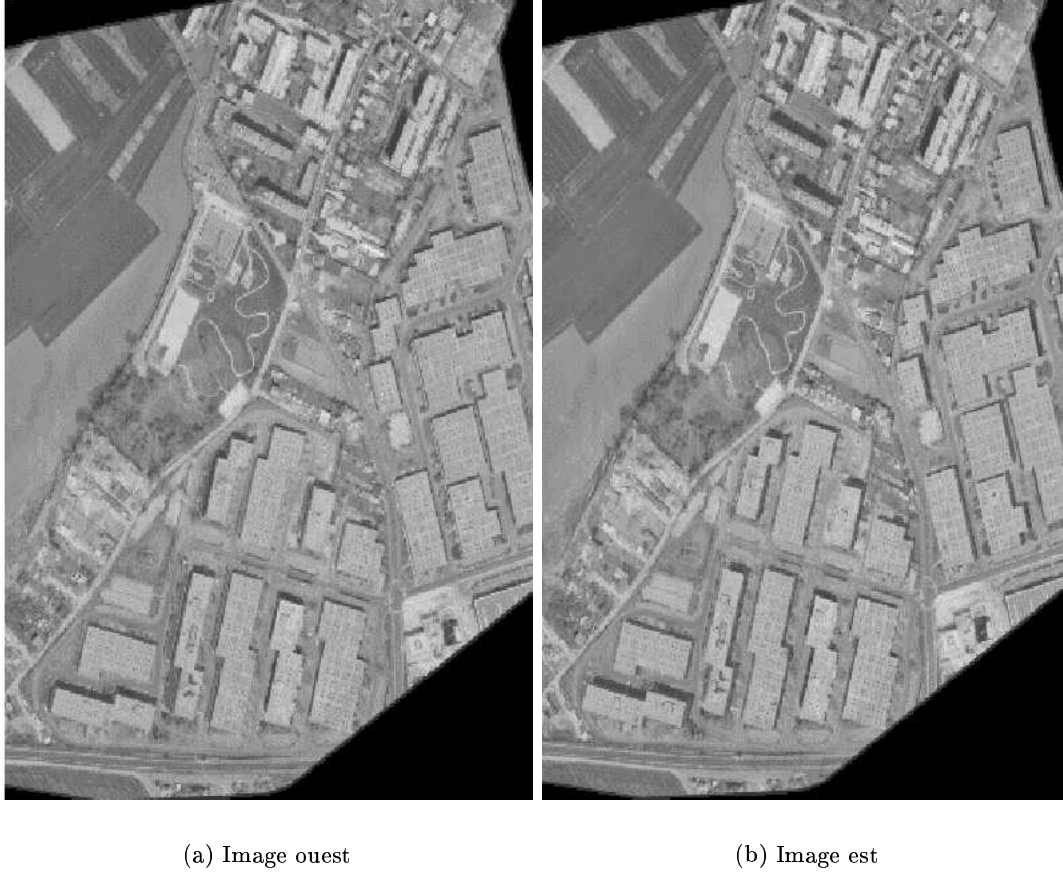


FIG. 3.1 – Exemple d'un couple d'images stéréoscopiques. Les images sont des données satellitaires simulées à partir d'images aériennes.

3.2.1 Stéréoscopie - rappel et principe général

Soit un point P de l'espace appartenant à un paysage représenté par deux images de télédétection différentes I_A et I_B . Ces images ont été prises à partir de positions différentes de la caméra. C_A et C_B dénotent les positions des centres optiques pour chacune des prises de vue. Le but de la stéréoscopie est de déterminer la localisation du point P de la scène par le calcul de ses coordonnées tridimensionnelles à partir de ses projections P_A et P_B dans les images (figure 3.2).

La différence relative entre P_A et P_B est appelée parallaxe et est inversement proportionnelle à la distance entre le point P et les plans focaux des caméras. La parallaxe représentant la distance en mètres dans le plan de mesure entre les deux représentations d'un même point, peut s'exprimer en pixels dans l'image, on l'appelle alors disparité. Par définition, la disparité d est le rapport entre la parallaxe et la résolution ρ de l'image en m/pixel :

$$d = \frac{\text{parallaxe}}{\rho}$$

L'intersection du plan (P, C_A, C_B) avec les plans focaux forme les droites épipolaires conjuguées. La distance $\overline{C_A C_B}$ est appelée la base que l'on notera B . Les centres optiques sont supposés être à une altitude H du sol.

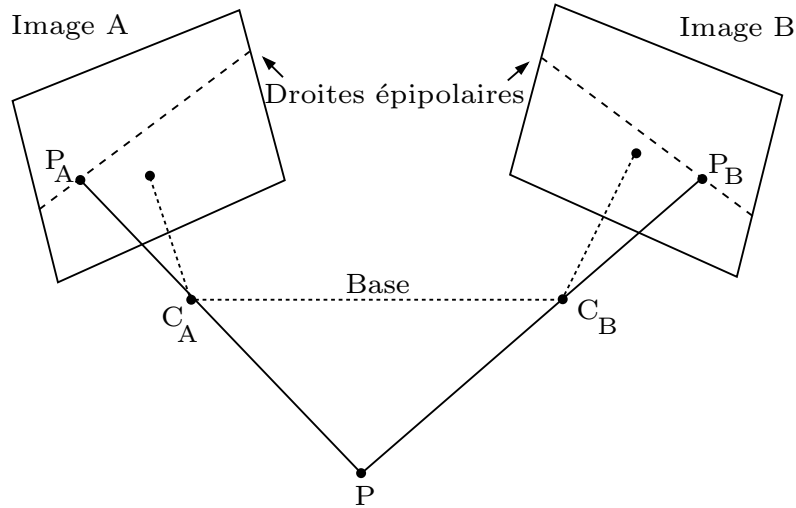


FIG. 3.2 – Pour la restitution stéréoscopique, deux images sont acquises de deux points de vue différents. Un point P de la scène se projette sur chaque image à l'intersection du plan image et de la droite joignant le point P et le centre optique C_A ou C_B .

La connaissance des conditions de prise de vue (orientation et paramètres intrinsèques de la caméra) permet de déterminer les directions de visée des pixels P_A et P_B , i.e. les droites $(P_A C_A)$ et $(P_B C_B)$. On peut alors calculer la position tridimensionnelle de P en considérant l'intersection des directions de visée des deux points homologues P_A et P_B (phase appelée aéro-triangulation). La reconstruction 3D par stéréoscopie revient donc à résoudre deux problèmes :

1. **Calibration.** Une connaissance précise des paramètres de prise de vue est requise pour un calcul fiable des directions de visée. Ces informations peuvent être estimées à partir des images et de points d'appui au sol ou connues *a priori*. En section 3.3, nous proposerons une méthode permettant d'affiner la précision de ces paramètres, dont seule une estimation imprécise est connue *a priori*.
2. **Mise en correspondance.** Pour un pixel P_A donné, il faut retrouver son homologue P_B dans l'image conjuguée afin de calculer l'intersection des visées. Ce problème appartenant à la reconnaissance de formes est le plus délicat. Il est à noter que l'espace de recherche de P_B pour un P_A donné peut être réduit de façon drastique par la connaissance des droites épipolaires conjuguées. Nous exposerons en section 3.2.2 la méthode de mise en correspondance surfacique implantée dans la chaîne de reconstruction ayant servi à la génération du MNS.

Les paramètres importants de la prise de vue stéréoscopique sont la base B et la hauteur H , ou plutôt leur rapport B/H . Dans le cas d'un satellite, H est fixée par l'orbite de ce dernier et B est le seul paramètre sur lequel il est possible d'influer. En imagerie aérienne, le réglage de ces paramètres est plus souple.

- Un rapport B/H élevé suppose deux prises de vue avec des angles importants par rapport au nadir de la scène observée. L'avantage d'une telle configuration est une bonne localisation altimétrique des objets. En effet, plus l'angle des directions de visée inter-images se rapproche

de 90° , moins l'incertitude sur les directions de visée induit de grandes erreurs sur l'altimétrie calculée. Cependant un rapport B/H important accroît significativement les risques d'occlusion (point vu dans une image mais caché dans l'autre à cause du relief de la scène) ainsi que le taux d'échec de reconnaissance des points homologues. Un rapport B/H important ne conviendra donc pas à un relief urbain car ce dernier varie fortement et rapidement induisant ainsi de nombreuses occlusions.

- A contrario, un rapport B/H peu élevé sous-entend deux prises de vues quasi semblables avec un angle faible par rapport au nadir de la scène. Le nombre d'occlusions est réduit, les points à apparier sont plus aisément reconnaissables. En revanche la précision altimétrique est dégradée car déterminée à partir de visées quasi parallèles engendrant de fortes erreurs s'il existe des imprécisions relatives à la calibration du capteur de télédétection. Il est donc nécessaire de choisir un rapport B/H adapté à la nature du relief de la scène observée. Outre les deux cas limites que nous venons d'exposer, B/H varie typiquement entre 0.1 et 1.2.

Finalement, remarquons que nous avons implicitement supposé jusqu'ici que les directions de visée de chaque couple de points homologues s'intersectent. Cela est évidemment faux à cause des erreurs et imprécisions des instruments de mesure permettant de calculer les directions de visée. Dans la pratique, le point d'intersection P est généralement défini comme le centre du plus petit segment joignant les deux directions de visée (figure 3.3).

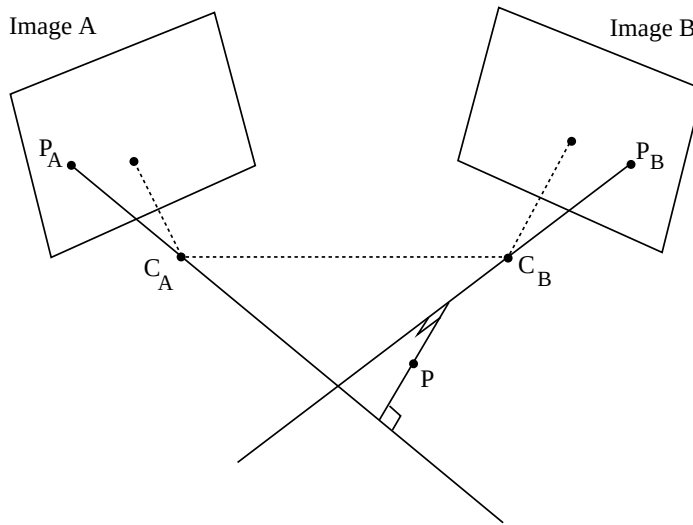


FIG. 3.3 – Pour deux points conjugués P_A et P_B , l'orientation relative permet de calculer la position du point P dans le système de coordonnées lié au couple d'images. P est défini comme le centre du plus petit segment reliant les droites (P_A, C_A) et (P_B, C_B) .

3.2.2 Mise en correspondance automatique d'images par masques de corrélation adaptatifs

Le but de la mise en correspondance est de trouver pour chaque pixel de l'image I_A de référence, son homologue dans l'image I_B . La relation géométrique exacte permettant de passer du point (i_A, j_A) de I_A à (i_B, j_B) de I_B est la suivante :

$$(i_B, j_B) = (i_A + d_c, j_A + d_l)$$

où : (d_c, d_l) est la disparité vraie au point (i_A, j_A) selon les colonnes et les lignes respectivement.

La mise en correspondance automatique d'images nécessite la maximisation d'une fonction de ressemblance entre deux points ou deux régions appartenant à chacune des images. L'approche surfacique consiste à chercher les points homologues à partir de la ressemblance de leurs *voisinages* respectifs. Soit un pixel de l'image de référence, on considère une vignette centrée en ce point, puis on cherche son homologue en balayant une fenêtre de même taille dans l'autre image. Il suffit de formuler un critère de ressemblance afin de quantifier la similarité des fenêtres.

3.2.2.1 Critère de ressemblance par corrélation croisée

Les techniques fondées sur la corrélation des intensités des images ont été utilisées dans de nombreuses applications commerciales de la stéréo-photogrammétrie, elles sont également l'une des plus anciennes méthodes développées en vision par ordinateur. Le principe de la corrélation croisée pour mesurer la similarité entre deux points est de calculer le coefficient d'inter-corrélation à partir de deux vignettes de taille $K \times L$. Plus le score de corrélation est élevé, plus la présomption de similarité sera forte. Le coefficient de corrélation centrée normalisée au pixel (i_A, j_A) de l'image de référence I_A et au pixel $\mathbf{s} = (s_c, s_l)$ dans l'image I_B s'écrit :

$$c_{i_A, j_A, \mathbf{s}} = \frac{\gamma_{i_A, j_A, \mathbf{s}}(I_A, I_B)}{\sigma_{i_A, j_A}(I_A) \sigma_{i_A, j_A, \mathbf{s}}(I_B)}$$

avec la fonction d'inter-corrélation :

$$\gamma_{i_A, j_A, \mathbf{s}}(I_A, I_B) = \sum_{m=i_A-K}^{i_A+K} \sum_{n=j_A-L}^{j_A+L} (I_A(m, n) - \bar{I}_A) (I_B(m + s_c, n + s_l) - \bar{I}_B)$$

et les variances :

$$\sigma_{i_A, j_A}^2(I_A) = \sum_{m=i_A-K}^{i_A+K} \sum_{n=j_A-L}^{j_A+L} (I_A(m, n) - \bar{I}_A)^2$$

et

$$\sigma_{i_A, j_A, \mathbf{s}}^2(I_B) = \sum_{m=i_A-K}^{i_A+K} \sum_{n=j_A-L}^{j_A+L} (I_B(m + s_c, n + s_l) - \bar{I}_B)^2$$

où :

- $\bar{I}_X$ représente la moyenne d'intensité des pixels de l'image X à l'intérieur de la fenêtre de taille $K \times L$ considérée.
- La fenêtre de corrélation dans l'image conjuguée est centrée sur le pixel $\mathbf{s} = (s_c, s_l)$. s_c et s_l appartiennent à l'intervalle $[-W, W]$ qui définit l'aire de recherche dans l'image conjuguée ($W > K, W > L$).

La normalisation du coefficient de corrélation par les écarts-type des vignettes, ainsi que le fait qu'il soit centré, permet de s'affranchir d'un biais radiométrique linéaire présent entre les deux images. Le processus de corrélation est réitéré pour chaque pixel de l'image conjuguée au sein de l'aire de recherche, ce qui permet de construire une surface de corrélation comprise entre 0 et 1 pour chaque pixel de l'image de référence (figure 3.4).

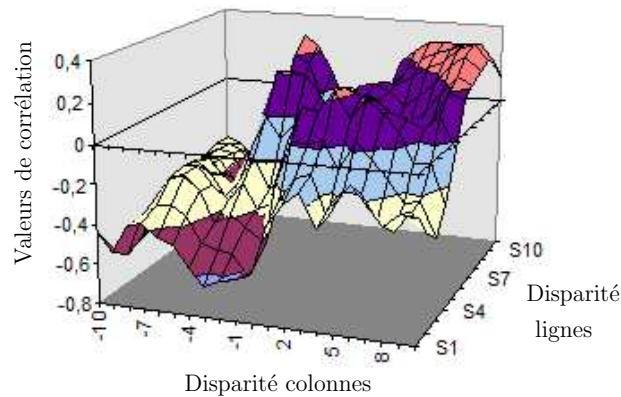


FIG. 3.4 – Exemple de surface de corrélation.

La surface de corrélation illustrée en figure 3.4 exhibe plusieurs modes. C'est ici une surface complexe. La surface de corrélation idéale correspond au cas où l'homologue est identifié de manière unique et non ambiguë, elle ne comprend alors qu'un seul pic de fort niveau. La position du maximum correspond à celle de l'homologue du pixel traité. Ce maximum est extrait et permet de calculer la disparité associée à ce pixel. Une interpolation de cette surface de corrélation au voisinage du maximum discret permet d'obtenir une précision sub-pixellaire de la position du véritable maximum, et donc de la disparité. La recherche de la disparité sub-pixellaire est réitérée pour chaque pixel de l'image de référence afin de réaliser une carte de disparité.

L'évaluation de la qualité de la mise en correspondance se fait selon les critères suivants :

- hauteur du pic de corrélation : la valeur de la corrélation doit être supérieure à un seuil ϵ_{corr} .
- largeur du pic de corrélation : l'étalement du pic au sommet se doit d'être restreint.
- unicité : présence limitée de maxima locaux (ambiguïtés) au voisinage du pic de corrélation.
- divergence de la fonction d'interpolation : le maximum sub-pixellaire ne doit pas être détecté trop loin du maximum discret.

Ces critères de qualité permettent de filtrer les erreurs et les ambiguïtés d'appariement afin d'obtenir une image de disparité la plus fiable possible. Si pour un pixel de l'image de référence, aucun pixel de l'image conjuguée ne satisfait les conditions précitées, la disparité n'est pas renseignée. Ce cas survient lorsque la ressemblance entre les deux pixels n'est pas assez franche, ou lorsqu'un point de la scène n'est visible que sur une seule image (occlusion).

D'autres contraintes de corrélation peuvent être prises en compte pour fiabiliser et accélérer la mise en correspondance :

- **Contrainte de réciprocité** : si X est l'homologue conjugué de Y, Y est l'homologue de X, d'où l'intérêt de réaliser une corrélation croisée (avec une image puis l'autre comme référence) puis un filtrage pour éliminer les points de disparité incohérente.
- **Contraintes externes** : elles limitent la combinatoire lors de la recherche de l'homologue afin d'augmenter la vitesse de traitement. Ainsi, rectifier les deux images en géométrie épipolaire permettra de chercher l'homologue d'un pixel dans l'image de référence sur une droite de l'autre image, et non plus une zone du plan. De plus, si on connaît *a priori* les altitudes maximales et minimales de la scène, on limite la recherche à une partie seulement de la droite

épipolaire correspondante. Cela permet de limiter les temps calculs puisque le problème de recherche de pixels homologues est alors réduit à une dimension. De plus, la restriction de l'espace de recherche limite l'occurrence d'erreurs de corrélation. La rectification en géométrie épipolaire nécessite néanmoins une connaissance *précise* des paramètres de prise de vue. Nous ne détaillons pas ce processus de rééchantillonnage des images qui est illustré en figures 3.5-3.6 et détaillé en [129].

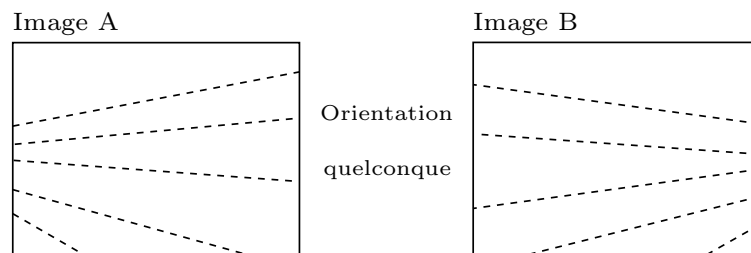


FIG. 3.5 – Paire d'images en géométrie quelconque

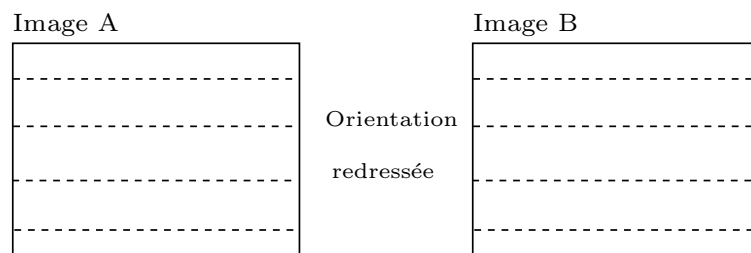


FIG. 3.6 – Paire d'images en géométrie épipolaire

3.2.2.2 Corrélation par masques adaptatifs

La méthode de corrélation avec des vignettes de dimensions fixes présente l'avantage d'obtenir des cartes de disparités denses. Cette technique est bien adaptée à la restitution de MNT, cependant elle présente une limite pour la génération de MNS urbains. En effet, elle a tendance à lisser les transitions altimétriques du fait des disparités non-homogènes contenues dans les vignettes. Ceci est particulièrement critique pour la reconstruction du relief du bâti.

La méthode de corrélation par masques adaptatifs a été conçue pour palier ce problème et restituer des MNS urbains présentant de fortes discontinuités altimétriques. Cette technique a été développée par N. Paparoditis dans [86, 87]. Cette approche permet d'adapter la forme de la vignette aux contours des objets présents dans l'image de façon à réduire les variations de disparité. Le principe de corrélation par masques adaptatifs est illustré en figure 3.7. On notera que cette approche est basée sur l'hypothèse que les discontinuités altimétriques et plus généralement les ruptures de pente de la scène se caractérisent par de forts gradients d'intensité dans l'espace image.

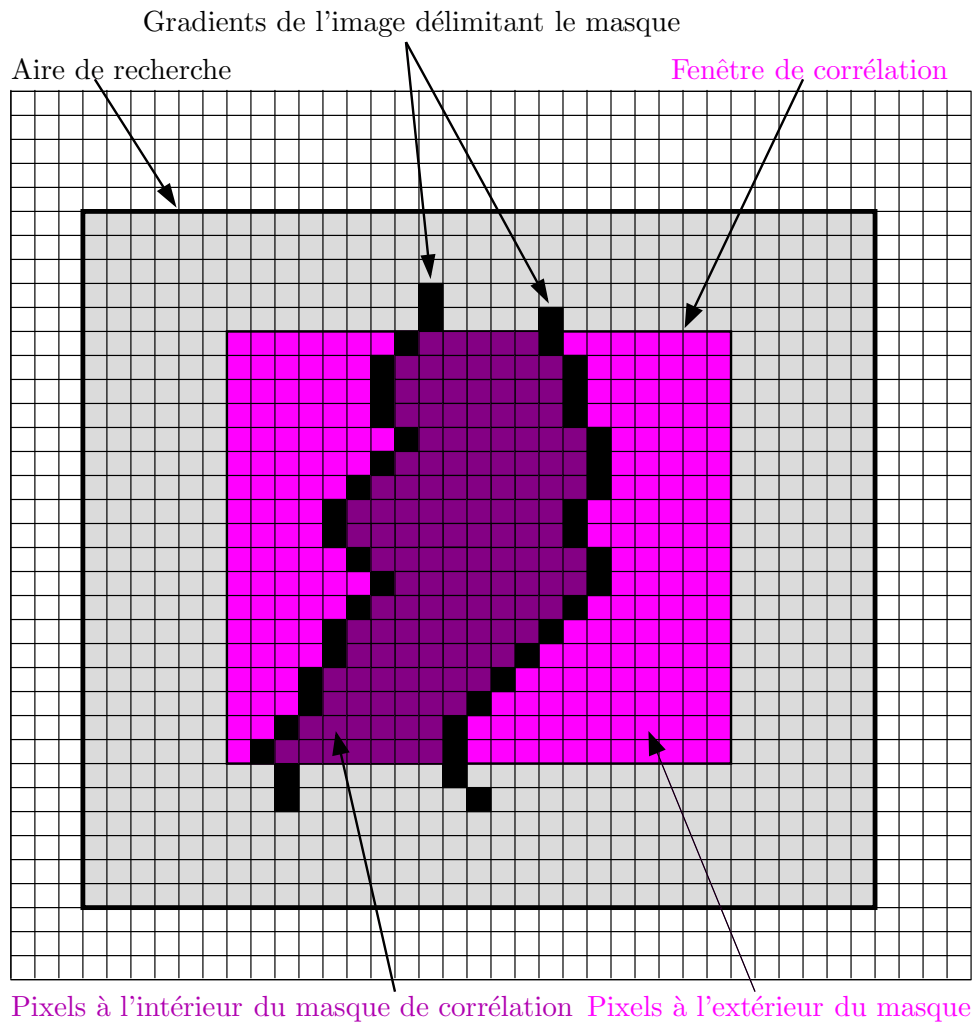


FIG. 3.7 – Masque adaptatif de corrélation.

Les figures 3.8 et 3.9 montrent l'apport de la corrélation par masques adaptatifs : la prise en compte des contours des objets permet d'éviter les erreurs d'appariement au voisinage du bâti.

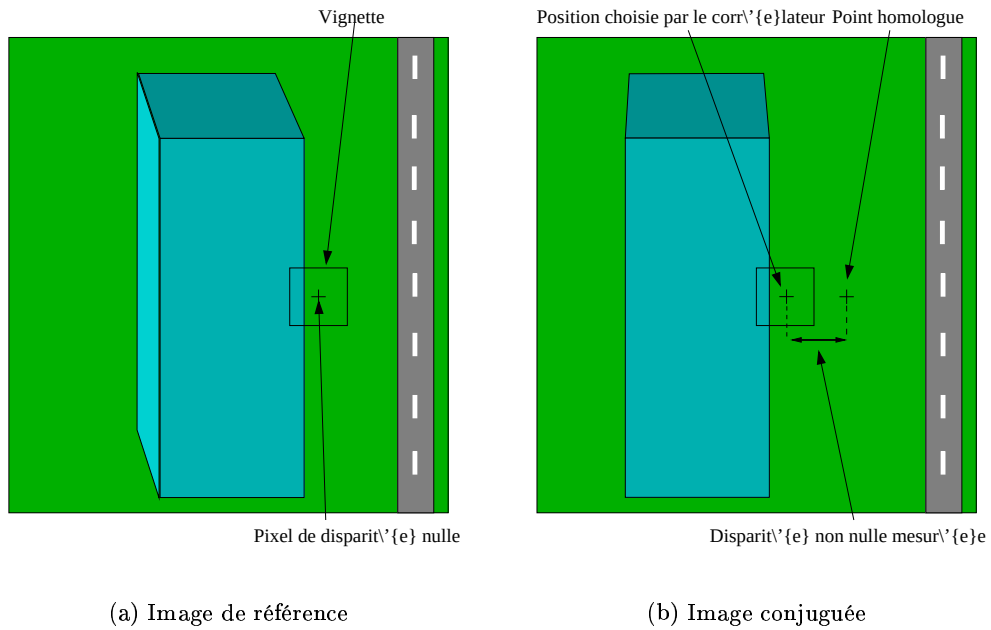


FIG. 3.8 – Erreur dans le cas de zones occultées avec une fenêtre de corrélation fixe. (a) un pixel de disparité nulle est considéré. (b) son homologue dans l'image conjuguée n'est pas retrouvé à cause de la prise en compte d'une partie du bâtiment dans la fenêtre de corrélation.

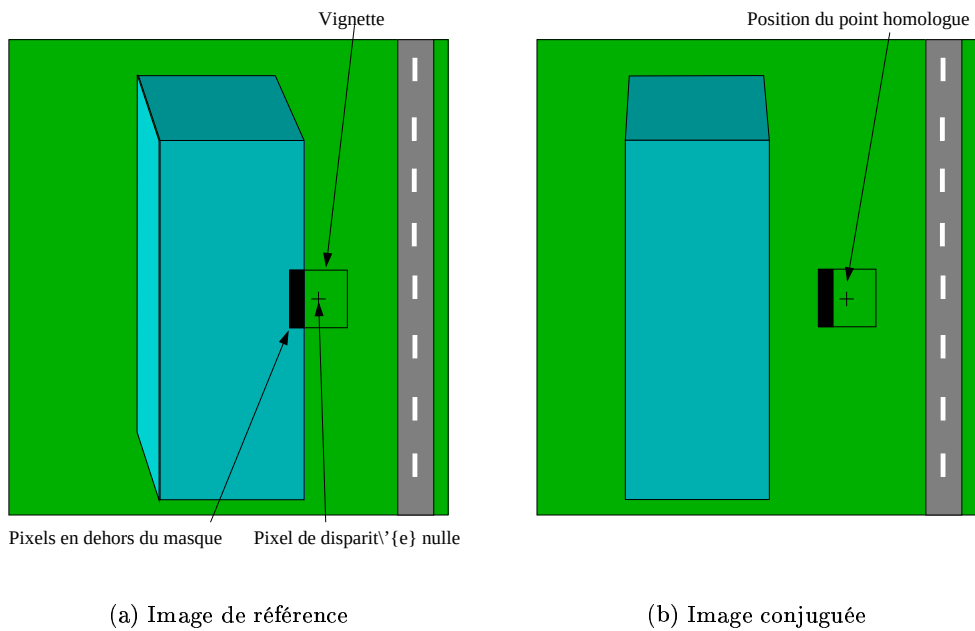


FIG. 3.9 – Application du masque sur la vignette de contexte. La partie du bâtiment est retirée de la fenêtre de corrélation grâce à la géométrie adaptative du masque. Le pixel homologue est retrouvé.

La taille du masque adaptatif de corrélation est ajustée localement et automatiquement. Cette adaptation locale est effectuée en considérant une taille minimale pour laquelle le rapport signal sur bruit soit suffisamment important pour parvenir à une corrélation de qualité. La qualité est évaluée par la mesure de la borne statistique dite de Cramer-Rao qui est un minorant de l'écart-type de l'erreur d'estimation de la disparité. La figure 3.10 illustre les meilleures performances des masques adaptatifs par rapport aux vignettes fixes.

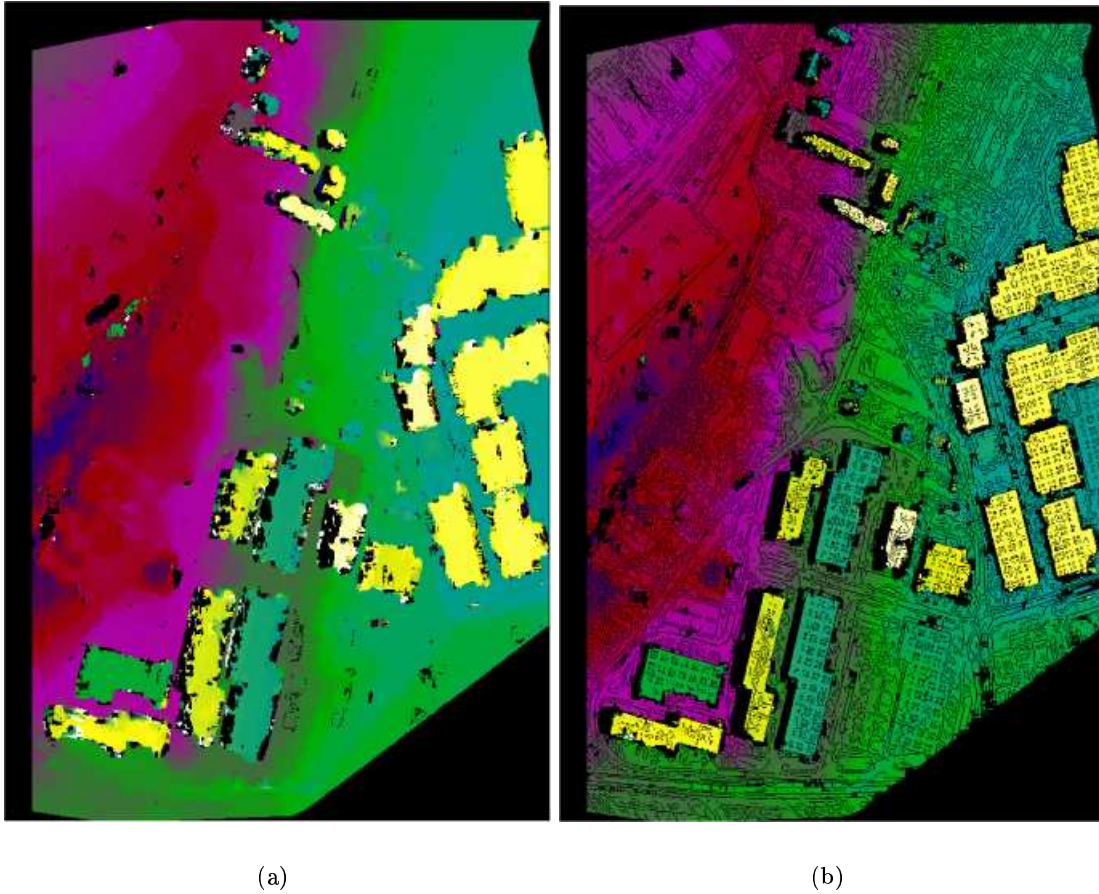


FIG. 3.10 – Image de disparité générée à partir du couple de la figure 3.1 : (a) corrélation classique (b) corrélation par masques adaptatifs.

3.2.3 Principe de la chaîne de reconstruction utilisée

La chaîne de reconstruction 3D utilisée permet de traiter aussi bien des couples d'images aériennes que satellitaires. Afin de réduire les temps de calcul et les erreurs de corrélation, les images sont premièrement rectifiées en géométrie épipolaire grâce à la connaissance des paramètres de prises de vue. Les images sont ensuite mises en correspondance par la technique des masques adaptatifs détaillée en section 3.2.2.2. Le résultat de la mise en correspondance est une image de disparité pour chaque paire d'images corrélées. Les rôles de l'image de référence et de l'image conjuguée sont ensuite échangés pour générer une seconde image de disparité équivalente. Les deux images de disparité sont ensuite fusionnées et l'aéro-triangulation est finalement mise en œuvre pour produire un MNS géocodé en géométrie épipolaire. Chaque pixel de ce MNS contient une information de localisation spatiale codée par des coordonnées géodétiques lorsqu'elle a pu être reconstruite (longitude, latitude et élévation par rapport à un ellipsoïde de référence), sinon, le pixel est non renseigné

(occlusion, faible score de corrélation, etc.). Le détail du processus de la chaîne de reconstruction est illustré en figure 3.11.

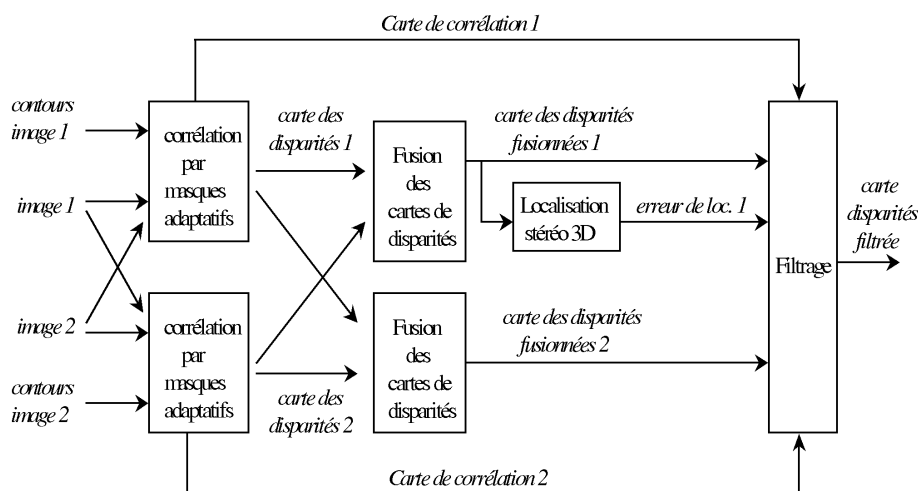


FIG. 3.11 – Synthèse du processus de reconstruction 3D.

3.3 Optimisation des paramètres de prise de vue

Nous considérons dans la suite du document la donnée de trois images aériennes représentant la ville de Pékin. Ces images sont le résultat de la numérisation de films analogiques (cf. annexe A). La précision des paramètres de prise de vue délivrés avec les trois images est cruciale pour la rectification en géométrie épipolaire précédant la phase de corrélation permettant de réaliser le MNS. Elle est aussi déterminante quant à la précision de l'information spatiale qui sera associée à chaque pixel du MNS par aéro-triangulation.

3.3.1 Paramètres de prise de vue

Les paramètres de prise de vue en imagerie optique aérienne sont les suivants :

1. Position du centre optique de la caméra exprimée en coordonnées tridimensionnelles par rapport à un référentiel $\mathfrak{R}$.
2. Angles d'attitude de la caméra par rapport à $\mathfrak{R}$.
3. Distance focale f de la caméra.
4. Point Principal de Symétrie (PPS) du plan focal.
5. Distorsion du système optique (négligeable pour les images de Pékin)
6. Paramètres d'affinité : six paramètres correspondant à la rotation, translation et facteur d'échelle modélisant la numérisation de l'image. Il est nécessaire d'introduire une rotation et une translation pour rendre compte du non alignement de la photographie avec le scanner ainsi que du décalage entre le premier pixel numérisé et celui correspondant au premier pixel imageant la photographie. Le facteur d'échelle correspond au grossissement (zoom) du scanner.

Les paramètres délivrés par le Beijing Institute of Surveying and Mapping (BISM) sont référencés par rapport au Beijing National System 1954 (BNS) qui est un référentiel de projection cartographique de type Gauss-Krüger (type UTM avec un facteur d'échelle du méridien central égal à 1) associée à l'ellipsoïde de Krasovsky. Les angles d'attitude de la caméra sont les paramètres $(\varphi, \omega, \kappa)$

couramment utilisés en photogrammétrie (et non les angles de tangage, roulis et lacet (ϕ, θ, ψ) utilisés en navigation aérienne [11]). Les angles (ϕ, θ, ψ) illustrés en figure 3.12 sont référencés par rapport à un plan tangent à l'ellipsoïde terrestre (horizon) ainsi qu'aux directions Nord et Est alors que $(\varphi, \omega, \kappa)$ expriment la rotation entre le repère caméra et un repère terrestre associé à une *projection cartographique* (figure 3.13). La relation de passage entre ces deux triplés d'angles n'est donc pas triviale et fait intervenir des équations non linéaires émanant de la projection cartographique.

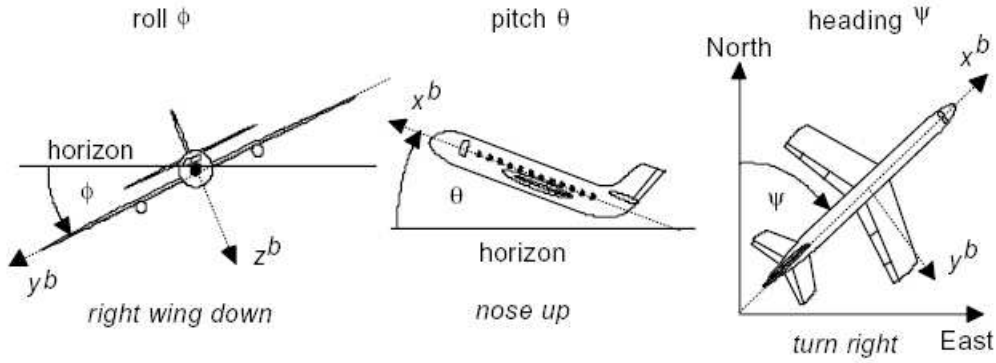


FIG. 3.12 – Définition des angles d'attitude (roulis, tangage, lacet) d'une caméra embarquée sur une plate-forme aérienne (en supposant que les axes de rotation de l'avion et de la caméra sont identiques). Le schéma est extrait de [11].

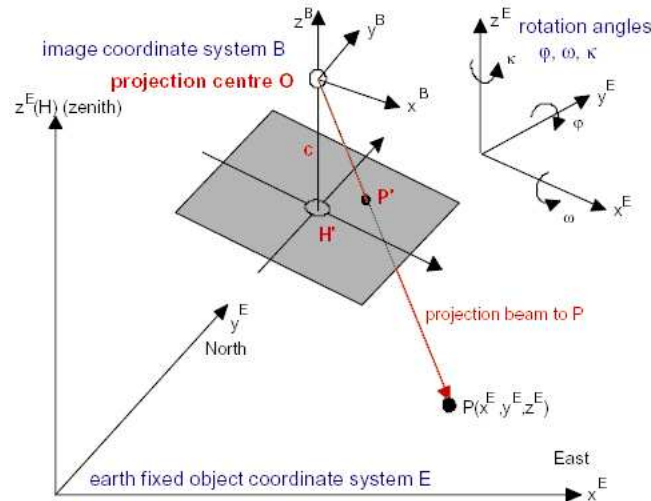


FIG. 3.13 – Illustration des angles $(\varphi, \omega, \kappa)$ utilisés en photogrammétrie. Ces angles expriment la rotation entre le référentiel caméra (x^B, y^B, z^B) et le référentiel terrain (x^E, y^E, z^E) associé à une projection cartographique. Le schéma est extrait de [11].

Les paramètres d'affinité ne sont pas directement fournis mais doivent être calculés grâce à huit marques fiduciaires présentes aux extrémités de l'image (figures A.1-3.14). Ces marques fiduciaires sont exprimées en millimètres dans le plan focal du système optique, et peuvent être repérées manuellement ou automatiquement dans l'image en coordonnées pixellaires. L'estimation des paramètres d'affinité est alors aisément réalisée par résolution d'un système linéaire surdéterminé par la technique des moindres carrés.

3.3.1.1 Calcul des paramètres d'affinité : orientation interne

Huit marques fiduciales sont exprimées en coordonnées millimétriques dans le plan focal de la caméra.

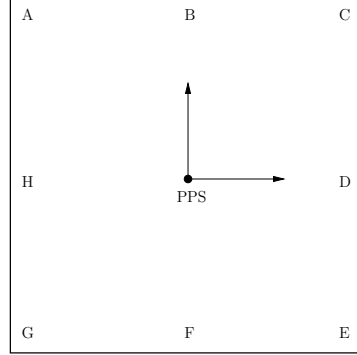


FIG. 3.14 – Huit marques fiduciales représentées dans le plan focal de la caméra. Le repère du plan focal est centré sur le point principal de symétrie (PPS) du système optique.

Les coordonnées en pixels de ces mêmes marques peuvent aussi être collectées manuellement ou automatiquement sur les images numérisées. Il est alors possible de résoudre le système suivant en possédant au moins trois marques :

$$\begin{cases} lig = T_{lig} + a_{00}X_{camera} + a_{01}Y_{camera} \\ col = T_{col} + a_{10}X_{camera} + a_{11}Y_{camera} \end{cases}$$

où :

- (col, lig) sont les coordonnées en pixels d'une marque fiduciale considérée Φ dans l'image. Ces données sont mesurées sur l'image
- (X_{camera}, Y_{camera}) sont les coordonnées de la marque Φ en millimètres dans le plan focal. Ces coordonnées sont fournies par les spécifications du constructeur de la caméra.
- (T_{lig}, T_{col}) sont les paramètres d'affinité en translation à estimer.
- $(a_{ij})_{i,j \in \{0,1\}}$ sont les paramètres d'affinité de rotation/facteur d'échelle à estimer.

La résolution d'un tel système linéaire se fait par la technique des moindres carrés. Nous avons choisi d'utiliser les huit marques simultanément pour une meilleure précision des estimations. Les résidus sur les huit marques varient entre 0,33 et 0,36 pixels sur les trois images. L'estimation de ces paramètres est effectuée sur les images aériennes non sous-échantillonnées afin de préserver la précision des marques fiduciales.

3.3.1.2 Évaluation de la qualité des paramètres de prise de vue

Comme nous l'avons souligné en introduction de cette section, la qualité des paramètres de prise de vue est déterminante pour l'utilisation des images avec la chaîne de reconstruction 3D. Cette qualité doit être examinée selon deux aspects. L'un concerne la précision de l'information de géo-référencement des pixels de l'image dans la scène 3D (précision absolue), l'autre regarde la précision de la localisation d'un point de la scène présent dans les deux images d'un couple stéréoscopique (précision relative) en vue d'une mise en géométrie épipolaire ultérieure. L'évaluation de ces deux critères de qualité des paramètres de prise de vue nécessite la donnée d'informations supplémentaires. En effet, des points de contrôle (GCPs¹) exprimés dans un référentiel $\mathfrak{R}$ ainsi

¹De l'anglais *Ground Control Points*.

qu'en coordonnées pixels dans les images sont indispensables pour évaluer la précision absolue. Des points conjugués repérés dans deux images d'un couple stéréoscopique (TPs²) permettent d'évaluer la précision relative. Pour les mêmes raisons qu'au paragraphe 3.3.1.1, l'évaluation de la qualité est effectuée sur les images non sous-échantillonnées.

▷ Précision absolue

Formalisation

L'évaluation de la précision absolue d'une image consiste à projeter tous ses GCPs (exprimés dans le repère $\mathfrak{R}$) dans l'image par modélisation inverse (annexe C.1). La modélisation inverse permet, grâce aux paramètres de prise de vue, de calculer la position d'un point 3D de la scène dans l'image. Il est alors possible de calculer la distance moyenne entre les points projetés dans l'image et les GCPs imagés correspondants. Environ dix GCPs par image ont été fournis par le BISM.

Notons :

- $GCP_{k,\mathfrak{R}}^l$ le $k^{ième}$ GCP de l'image l exprimé en coordonnées tridimensionnelles par rapport au référentiel terrain $\mathfrak{R}$.
- $GCP_{k,Im}^l$ le $k^{ième}$ GCP de l'image l exprimé en coordonnées pixels.
- $GCP_{k,proj}^l$ est le $k^{ième}$ GCP de l'image l projeté dans l'image par modélisation inverse à partir de $GCP_{k,\mathfrak{R}}^l$.

Une image l contient N_l GCPs, dans notre cas $l = \{1, 2, 3\}$ et $N_l \simeq 10$. Le critère d'évaluation de la précision absolue conférée par les paramètres de prise de vue pour l'image l devient :

$$d_{abs,l} = \frac{\sum_{k=1}^{N_l} \overline{GCP_{k,proj}^l GCP_{k,Im}^l}}{N_l}$$

Résultats

Les résultats de la précision absolue avec les paramètres de prise de vue fournis sont de l'ordre de deux à trois pixels pour les images 1 et 3 (table 3.1). Les paramètres sont moins bons sur la deuxième image avec presque huit pixels de différence en moyenne et un écart type important. Une optimisation absolue semble nécessaire au regard de ces résultats.

l	$d_{abs,l}$ en pixels	$\max \left(\overline{GCP_{k,proj}^l GCP_{k,Im}^l} \right)$	$\sigma \left(\overline{GCP_{k,proj}^l GCP_{k,Im}^l} \right)$
1	2.3	4.8	1
2	7.5	14.1	4.7
3	2.7	6.4	1.5

TAB. 3.1 – Résultats en pixels de la quantification de la précision absolue des paramètres de prise de vue pour les trois images aériennes.

²De l'anglais *Tie Points*.

▷ Précision relative

Formalisation

L'évaluation de la précision relative consiste à projeter tous les TPs d'une image sous forme de droites épipolaires dans l'image stéréoscopique conjuguée. Le critère de précision relative est la moyenne des distances entre les droites épipolaires calculées et les TPs de la seconde image. Pour déterminer une droite épipolaire dans la seconde image d'un couple stéréoscopique, un TP de la première image est projeté dans le repère terrain $\mathfrak{R}$ par modélisation directe à deux altitudes différentes : alt_{min} et alt_{max} (annexe C.2). Ces deux points 3D sont ensuite projetés dans la seconde image par modélisation inverse. La projection dans le repère $\mathfrak{R}$ fait intervenir les paramètres de prise de vue de la première image alors que ceux de la seconde sont sollicités pour la modélisation inverse.

Notons :

- $TP_{k,1}^m$ le $k^{ième}$ TP de la première image du $m^{ième}$ couple stéréoscopique. Il est exprimé en coordonnées pixels.
- $TP_{k,2}^m$ le $k^{ième}$ TP de la seconde image du $m^{ième}$ couple stéréoscopique. Il est exprimé en coordonnées pixels.
- $EPI_{k,proj}^m$ est la $k^{ième}$ droite épipolaire projetée dans la seconde image du $m^{ième}$ couple stéréoscopique. Elle est calculée à partir de $TP_{k,1}^m$.

Les deux images du $m^{ième}$ couple stéréoscopique contiennent chacune N_m TPs, dans notre cas $m = \{1, 2\}$ et $N_m = 10$. Le critère d'évaluation de la précision relative pour le couple m devient :

$$d_{rel,m} = \frac{\sum_{k=1}^{N_m} dist \left(EPI_{k,proj}^m, TP_{k,2}^m \right)}{N_m} \quad (3.1)$$

La fonction distance $dist(.)$ exprimée dans l'équation (3.1) est la distance d'un point à une droite, i.e. la distance entre le point et l'intersection entre la droite et la normale à cette dernière passant par le point. Les TPs ont été extraits manuellement, puis affinés par corrélation sub-pixellaire [14]. Vingt TPs sont ainsi sélectionnés pour finalement garder les dix meilleurs.

Résultats

Les résultats de précision relative obtenus ne sont pas assez bons pour être utilisés par la chaîne de reconstruction 3D (table 3.2). En effet, une précision relative inférieure à trois pixels est requise pour la phase préalable de mise en géométrie épipolaire. La précision des paramètres de prise de vue fournis par le BISM n'est pas suffisante pour assurer un géo-référencement de qualité ainsi que la mise en géométrie épipolaire des images. Une optimisation de ces paramètres s'avère donc nécessaire.

m	$d_{rel,m}$ en pixels	$\max \left(dist \left(EPI_{k,proj}^m, TP_{k,\in}^m \right) \right)$	$\sigma \left(dist \left(EPI_{k,proj}^m, TP_{k,\in}^m \right) \right)$
1	1.7	3	0.6
2	3.9	7.7	1.8

TAB. 3.2 – Résultats en pixels de la quantification de la précision relative des paramètres de prise de vue pour chacune des paires d'images aériennes.

3.3.2 Optimisation globale

3.3.2.1 Fonction de coût

L'optimisation des paramètres de prise de vue des trois images se doit de tenir compte à la fois de la précision absolue et relative. Les paramètres optimisés permettront ainsi de mettre les images en géométrie épipolaire (pour assurer la réussite de la corrélation 1D du logiciel de reconstruction) tout en assurant une information de localisation géographique de qualité (chaque pixel du MNS résultant possède des coordonnées de longitude, latitude et élévation calculées par aéro-triangulation). L'optimisation est menée sur les images non sous-échantillonnées. Il suffira de diviser les paramètres d'affinité par le facteur de sous-échantillonnage pour les rendre compatibles avec les images de taille réduite en entrée du logiciel de reconstruction. Le fait de posséder plus de deux images est un avantage pour la qualité de l'optimisation. En effet, la redondance d'information permettra de mieux minimiser chaque distance $d_{abs,l}$ et $d_{rel,m}$ intervenant dans l'optimisation. La fonction de coût J_{param} à minimiser est alors :

$$J_{param} = \sum_{l=1}^n d_{abs,l} + \sum_{m=1}^{n-1} d_{rel,m} \quad (3.2)$$

Remarque : si une zone de la scène est imagée par plus de deux photographies, nous ne considérerons que le couple ayant le recouvrement le plus élevé. Ainsi pour n images, nous supposerons qu'il existe $n - 1$ couples stéréo-graphiques.

La fonction de coût exposée en (3.2) n'est pas linéaire en fonction des paramètres de prise de vue des images. L'approche couramment utilisée pour minimiser J_{param} consiste à linéariser la fonction de coût afin d'effectuer l'optimisation. Nous avons choisi de garder J_{param} dans sa forme non linéaire originale et d'utiliser l'algorithme du simplex pour mener à bien l'estimation des paramètres de vue optimaux. Le simplex est un algorithme robuste d'optimisation adapté aux fonctions non linéaires ; le détail de cette méthode est exposé en annexe B. Nos motivations à choisir l'algorithme du simplex sont les suivantes :

- Le simplex optimisant directement des fonctions non linéaires, il n'est pas nécessaire de linéariser la fonction de coût J_{param} .
- J_{param} n'étant pas linéarisée, une estimation plus précise des paramètres de prise de vue est susceptible d'être obtenue.
- Le simplex n'a pas de paramètres de réglage, ce qui rend son utilisation aisée.

Les degrés de liberté de l'optimisation sont tous les paramètres de prise de vue des images. A chaque image sont associés 15 paramètres : 3 pour le centre optique, 3 pour les angles d'attitude, 1 pour la distance focale, 2 pour le PPS et 6 pour les paramètres d'affinité. Cependant la distance focale ainsi que le PPS sont supposés identiques quelle que soit l'image considérée. Pour n images il y aura ainsi $12n + 3$ paramètres (soit 39 dans cette étude).

3.3.2.2 Référentiel cartésien géocentré

Les paramètres de prise de vue en entrée du logiciel de reconstruction doivent être référencés par rapport à un référentiel cartésien direct géocentré (cela concerne l'expression de la position et de l'attitude de la caméra). L'origine d'un tel repère est confondu avec le centre de gravité de la terre. L'axe (OZ) est l'axe de rotation de la terre et pointe vers le nord. L'axe (OX) est tel que le méridien de Greenwich soit inclus dans le plan (OXZ). (OY) est tel que le repère ($OXYZ$) soit

direct. Désormais, $\mathfrak{R}$ désigne ce référentiel (voir [39] pour d'autres référentiels couramment utilisés).

La position du centre optique de la caméra pour chaque prise de vue doit désormais être exprimée dans ce repère $\mathfrak{R}$. Puisque nous disposons de la position dans le repère de projection BNS, il a fallu convertir cette dernière grâce aux équations de [39, pp. 100-101]. Quant aux angles d'attitude du système optique, il est délicat d'essayer de les convertir du système BNS vers $\mathfrak{R}$. La principale difficulté est la différente nature des référentiels utilisés : l'un est à trois dimensions ($\mathfrak{R}$), l'autre est le résultat d'une projection cylindrique transverse (BNS). Il a été plus simple d'estimer ces angles par optimisation absolue sur chaque image en utilisant les GCPs préalablement convertis dans le système $\mathfrak{R}$. Il est légitime de se demander si ces angles, initialisés à 0, n'auraient pas pu être estimés lors de l'optimisation globale. L'expérience a montré que le calcul des angles échouait dans ce cas, d'où la nécessité de le faire préalablement par optimisation absolue sur chaque image. Le diagramme de la figure 3.15 synthétise le fonctionnement global du programme d'optimisation des paramètres de prise de vue.

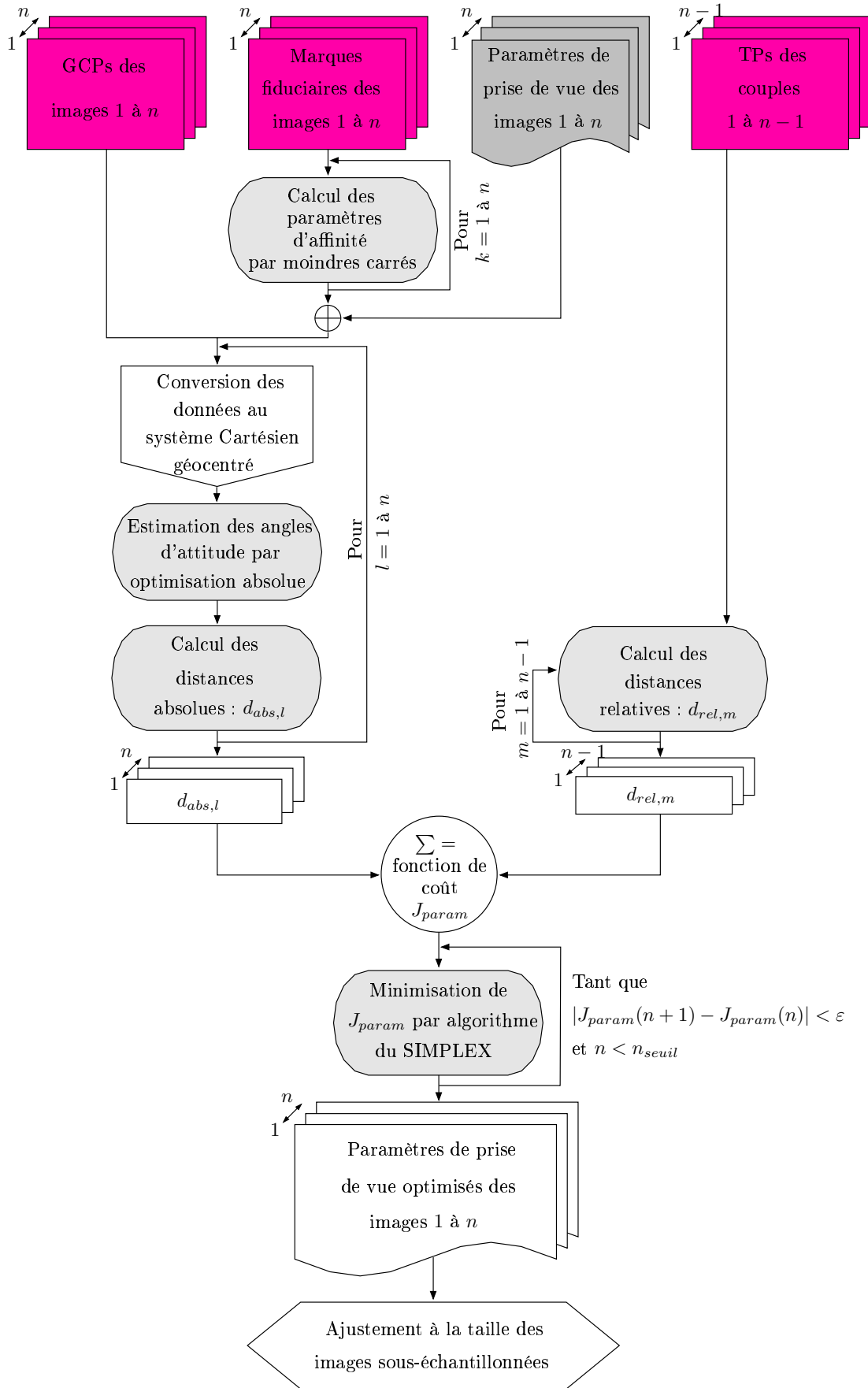


FIG. 3.15 – Synthèse algorithmique de l'optimisation globale des paramètres de prise de vue de n images aériennes.

3.3.2.3 Résultats

Les résultats obtenus avec les paramètres de prise de vue des trois images de Pékin sont donnés en tableau 3.3.

l	$d_{abs,l}$ en pixels	$\max \left(dist \left(GCP_{k,proj}^l, GCP_{k,Im}^l \right) \right)$	$\sigma \left(dist \left(GCP_{k,proj}^l, GCP_{k,Im}^l \right) \right)$
1	1.5	4.5	1.5
2	1.8	4.1	1.2
3	1.8	5.1	1.5
m	$d_{rel,m}$ en pixels	$\max \left(dist \left(EPI_{k,proj}^m, TP_{k,\in}^m \right) \right)$	$\sigma \left(dist \left(EPI_{k,proj}^m, TP_{k,\in}^m \right) \right)$
1	0.3	1	0.3
2	0.3	0.65	0.25

TAB. 3.3 – Résultats de l’optimisation des paramètres de prise de vue avec les images à 0.21 m/pixel de résolution.

Si l’on ramène ces résultats aux images sous-échantillonnées (par un facteur 3) utilisées par la chaîne de reconstruction 3D, on a :

l	$d_{abs,l}$ en pixels
1	0.5
2	0.6
3	0.6
m	$d_{rel,m}$ en pixels
1	0.1
2	0.1

TAB. 3.4 – Résultats de l’optimisation des paramètres de prise de vue avec les images sous-échantillonnées par un facteur trois.

On constate que l’optimisation a permis d’améliorer la précision des paramètres de prise de vue, et ceci de façon significative pour la précision relative. Il est néanmoins important de mentionner que ces résultats sont sensibles à la répartition et au nombre des points de contrôle dans les images. Environ dix points uniformément répartis sur chaque image sont nécessaires pour obtenir des résultats satisfaisants. La figure 3.17 illustre le MNS généré par la chaîne de reconstruction 3D avec les paramètres optimisés et le couple stéréoscopique de la figure 3.16.

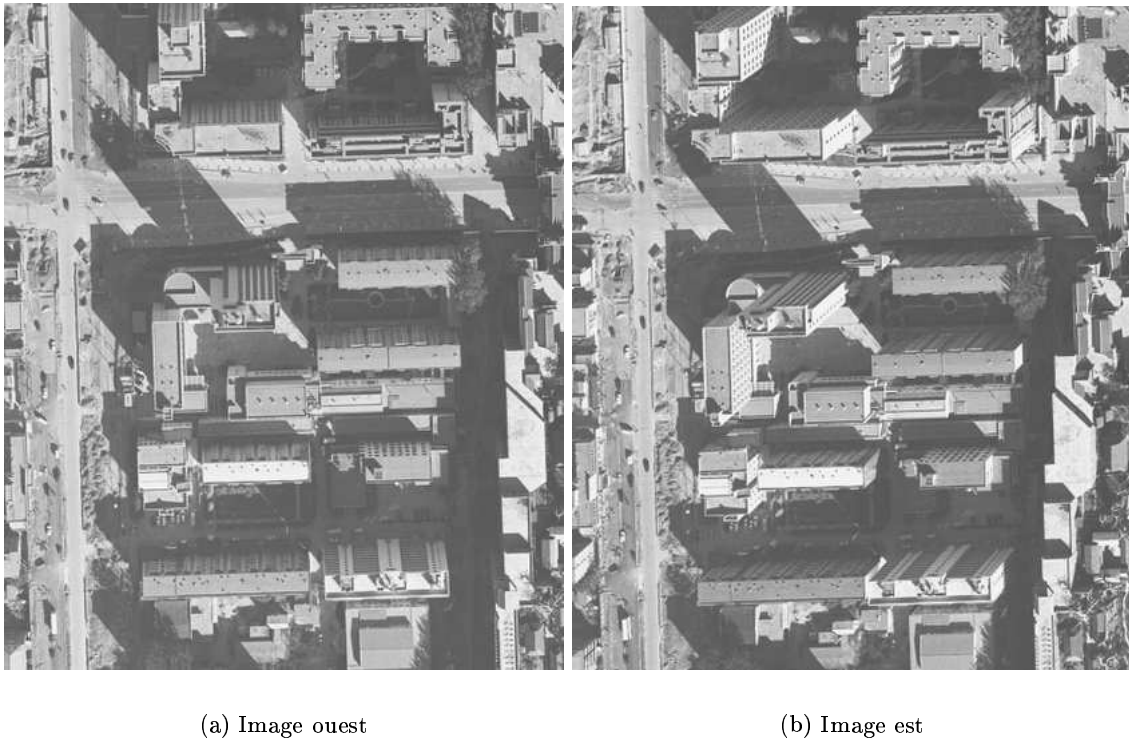


FIG. 3.16 – Couple d'images stéréoscopiques rectifiées en géométrie épipolaire.

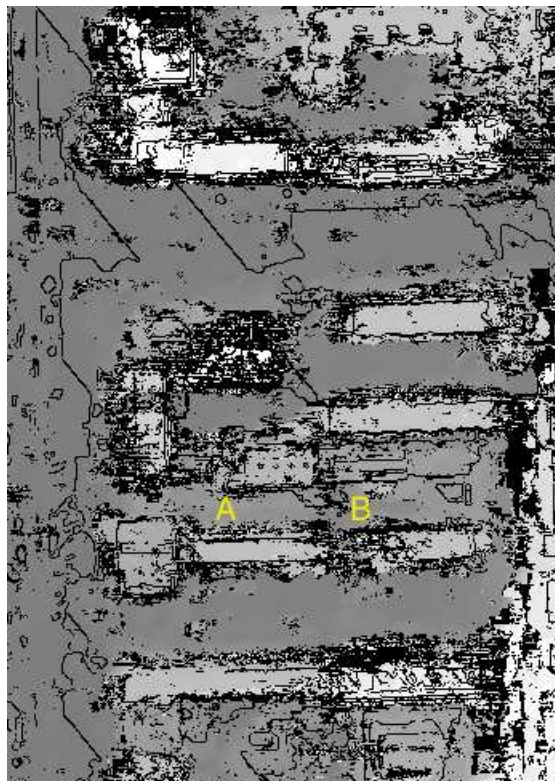


FIG. 3.17 – MNS en géométrie épipolaire. Les zones non renseignées sont en noir ; la luminance est proportionnelle à l'altitude.

Le MNS généré possède de nombreux pixels non renseignés (en noir sur l'image de la figure 3.17) au niveau du bâti. Ce sont en effet des zones où il existe des occlusions et où les façades de bâtiments sont difficiles à corréler. Les contours artificiels dus aux ombres et les objets présents sur une seule des images (voitures ...) provoquent des zones non renseignées de faible dimension sur la route. Cependant, les contours de bâtiments sont bien préservés, et la plupart d'entre eux ont une altitude assez homogène. Les ombres portées sur le sol sont aussi bien gérées (pas de détection de faux bâtiments). Même si cette image est visuellement imparfaite, elle contient néanmoins une information pertinente permettant de localiser le bâti et de le distinguer du sol dans la majorité des cas. Les limites du corrélateur sont illustrées aux endroits annotés A et B sur l'image de la figure 3.17 : ces zones semblent appartenir au sol (région de faible intensité, par conséquent de faible élévation) alors qu'en réalité elles correspondent à deux bâtiments (figure 3.16). Les bâtiments A et B représentent un cas extrême du milieu urbain dense. Ils sont en effet enclavés et dans une zone d'ombre qui permet difficilement d'observer leurs disparités et leurs contours.

3.3.3 Évaluation de la qualité de géocodage des MNSs

L'évaluation de la qualité du géocodage du MNS a été réalisée par la donnée de points de contrôle (GCPs). Chaque point de contrôle est exprimé en coordonnées pixellaires dans les images aériennes ainsi qu'en coordonnées 3D dans le système cartographique de Pékin : Est (E), Nord (N) et élévation (Z). Les coordonnées cartographiques sont exprimées en mètres, l'altitude Z est référencée par rapport à l'ellipsoïde de Krasovsky. La qualité du MNS pourra être quantifiée lorsque les coordonnées de ces GCPs seront mesurées dans le MNS puis comparées à leur vérité terrain. La première étape consiste à repérer les points de contrôle dans les images de référence en géométrie épipolaire sous-échantillonnées ayant servi à la génération du MNS. Puisqu'il existe une exacte correspondance entre les images de référence et les MNS épipolaires générés à partir d'elles, il est alors possible de lire les valeurs (E, N, Z) dans les MNS pour les GCPs considérés. Nous avons généré deux MNS épipolaires à partir de 3 images stéréo-paires deux à deux. Pour chaque MNS, nous possédons une dizaine de GCPs parmi lesquels quelques-uns n'ont pas été impliqués dans l'optimisation des paramètres de prise de vue. Les tableaux 3.5 et 3.6 rassemblent les résultats.

	<i>Erreur Absolue</i>			
<i>GCP</i>	ΔN	ΔE	ΔZ	$\Delta Plan$
1	1.0	0.1	0.8	1.0
2	-0.5	-0.1	0.4	0.5
3	1.5	0.6	2.9	1.6
4	-0.5	0.8	1.5	0.9
5	0.0	0.6	2.4	0.6
6	0.0	-1.5	0.2	1.5
7	0.0	0.0	-0.5	0.0
8	-0.5	-0.5	0.5	0.7
9	0.0	-0.2	0.8	0.2
10	-1.0	-0.1	0.2	1.0
11	-0.5	0.5	-0.3	0.7
12	0.0	-0.3	0.2	0.3
Moyenne	0.0	0.0	0.8	0.7
<i>Écart type</i>	0.7	0.6	1.0	0.5

TAB. 3.5 – Erreur absolue entre les GCPs et la reconstruction 3D du MNS_1 . ΔN : erreur dans la direction nord, ΔE : erreur dans la direction est, ΔZ erreur altimétrique, erreur dans le plan : $\Delta Plan = \sqrt{\Delta N^2 + \Delta E^2}$.

	<i>Erreur Absolue</i>			
<i>GCP</i>	ΔN	ΔE	ΔZ	$\Delta Plan$
1	0.5	-0.9	1.0	1.0
2	0.0	-0.2	-0.2	0.2
3	-0.5	0.6	2.7	0.8
4	-0.5	0.1	1.6	0.5
5	0.5	0.0	1.4	0.5
6	2.5	-3.4	-0.8	4.2
7	-1.5	0.5	0.4	1.5
8	0.0	0.2	1.1	0.2
9	0.0	-0.2	0.5	0.2
10	-0.5	0.2	2.2	0.5
11	-1.5	0.4	-0.2	1.5
12	0.0	0.8	-1.3	0.8
13	0.0	-0.2	0.1	0.2
14	0.0	-0.5	-0.3	0.5
Moyenne	-0.1	-0.2	0.6	0.9
<i>Écart type</i>	1.0	1.0	1.1	1.1

TAB. 3.6 – Erreur absolue entre les GCPs et la reconstruction 3D du MNS_2 . ΔN : erreur dans la direction nord, ΔE : erreur dans la direction est, ΔZ erreur altimétrique, erreur dans le plan : $\Delta Plan = \sqrt{\Delta N^2 + \Delta E^2}$.

Les quantités pertinentes de ces tableaux sont les écarts types qui représentent la précision du géocodage dans la direction Nord, Est, Z ou dans le plan. On remarque que la précision planimétrique varie entre 0.5 m et 1 m, la précision altimétrique est de 1 m. Ces résultats semblent satisfaisants et valident la phase préliminaire d'optimisation des paramètres de prise de vue de la caméra, ainsi que le calcul d'aéro-triangulation de la chaîne 3D.

3.4 Orthorectification d'un MNS épipolaire

L'orthorectification consiste à corriger la distorsion due à la perspective présente dans l'image. Chaque pixel d'une image en géométrie orthoscopique est tel qu'il serait vu au nadir. La nécessité d'orthorectifier des images de scènes urbaines survient uniquement pour des images à très haute résolution représentant un relief ou des bâtiments élevés. Dès lors que l'image est orthorectifiée, elle est superposable avec des données cartographiques qui sont aussi dans cette géométrie. Le MNS issu de la chaîne de reconstruction 3D est en géométrie épipolaire. L'information de géo-référencement associée à chaque pixel du MNS permettra d'orthorectifier ce dernier.

3.4.1 Orthorectification par méthode directe

L'orthorectification par méthode directe consiste à considérer chaque pixel du MNS en géométrie épipolaire et à le projeter dans une grille terrain géocodée [82]. Pour une même localisation planimétrique dans la grille orthoscopique, plusieurs pixels du MNS épipolaire peuvent y être projetés. Seul le pixel de plus haute altitude sera gardé (technique du *z-buffer*). La grille orthoscopique peut être codée en longitude/latitude ou selon une projection cartographique. Dans notre cas, nous désirons superposer le MNS avec les données cartographiques. Par conséquent, l'information de longitude/latitude présente dans le MNS épipolaire généré a été convertie en coordonnées (E, N, Z) de la projection cartographique BNS [39, pp. 99]. Chaque pixel est ensuite projeté dans la grille orthoscopique elle-même associée au référentiel BNS. La relation entre les pixels de la grille orthoscopique et les coordonnées terrain BNS est du type :

$$\begin{cases} E = E_{min} + iR \\ N = N_{max} - jR \end{cases} \quad (3.3)$$

où :

- R est le pas de la grille en mètres.
- (E_{min}, N_{max}) sont les coordonnées terrain BNS de l'origine contenue dans le MNS épipolaire.
- (i, j) sont les coordonnées pixels de la grille orthoscopique, respectivement colonne et ligne.
- (E, N) sont les coordonnées terrain BNS associées au pixel (i, j) .

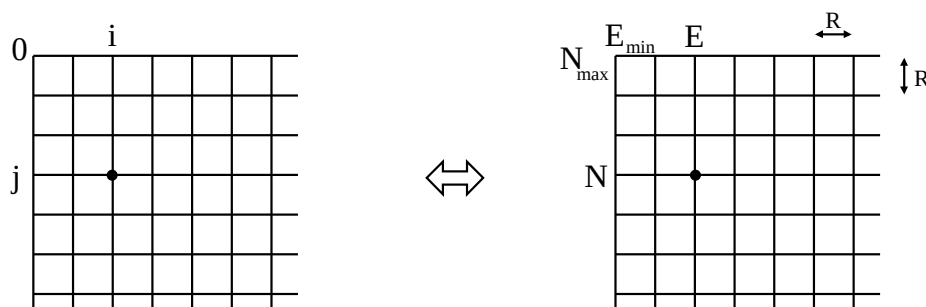


FIG. 3.18 – Grille orthoscopique associée à la projection cartographique.

3.4.1.1 Interpolation triangulaire (interpolation bilinéaire sur les 3 plus proches voisins)

Projeter individuellement chaque pixel du MNS épipolaire dans la grille orthoscopique ne peut aboutir à un ortho-MNS de qualité. En effet, un pixel projeté dans la grille orthoscopique aura des coordonnées non entières. La contribution du pixel projeté envers son voisinage devra faire l'objet d'une interpolation. La principale question est de décider sur quel voisinage contribuera le pixel. Si le voisinage est trop petit, certains points de la grille orthoscopique seront non renseignés et le

résultat sera peu dense. Si le voisinage d'influence est trop élevé, il se peut que des points de la grille se voient affecter une valeur d'altitude incorrecte. Ce problème sera d'autant plus sensible au niveau des toits des bâtiments.

Une manière de trancher cette question est d'adopter une interpolation triangulaire. Au lieu de considérer individuellement chaque pixel du MNS épipolaire, considérons un triangle isocèle dont les côtés égaux font un pixel. Ce triangle (ABC) du MNS épipolaire est illustré en figure 3.19 (à gauche). Ce triangle est ensuite projeté dans la grille orthoscopique grâce aux informations de géocodage associées à chaque pixel A , B et C et grâce au système d'équations (3.3).

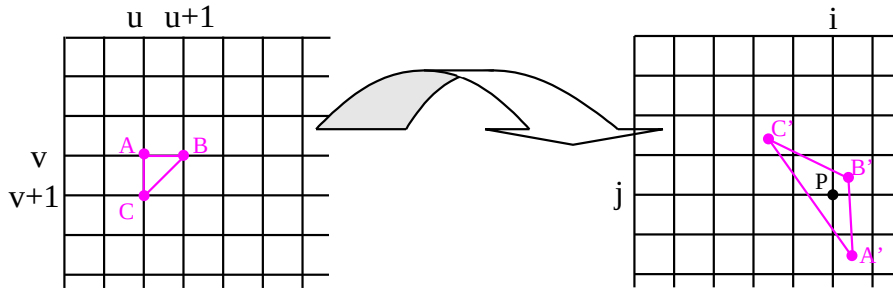


FIG. 3.19 – Projection d'un triangle du MNS épipolaire vers la grille orthoscopique.

Les points de la grille orthoscopique inclus dans le triangle projeté se verront affecter une altitude et une radiométrie selon les formules d'interpolation suivantes :

$$z_{ij} = \frac{z_A \overline{PB} \overline{PC} + z_B \overline{PA} \overline{PC} + z_C \overline{PA} \overline{PB}}{\overline{PA} \overline{PB} + \overline{PB} \overline{PC} + \overline{PC} \overline{PA}} \quad (3.4)$$

où :

- z_{ij} est l'altitude du point P de la grille et de coordonnées (i, j) . P est inclus dans le triangle projeté (A, B, C) .
- z_X est l'altitude du point X ($X = A, B$ ou C).

Si la distorsion due à la perspective est importante, rares seront les cas où un pixel de la grille sera inclus dans un seul triangle projeté. Pour illustrer ce propos, prenons l'exemple d'une arête d'un bâtiment parallélépipédique soumise à une distorsion de perspective dans l'image. Tous les pixels de cette arête ont une même position planimétrique (E_0, N_0) avec des altitudes différentes. Dans la grille orthoscopique, ce point se verra ainsi inclus dans plusieurs triangles (figure 3.20).

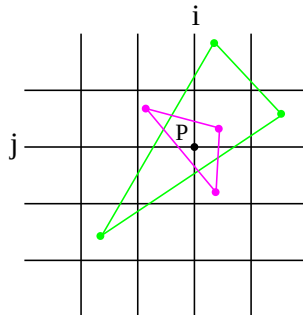


FIG. 3.20 – Multiples triangles projetés dans la grille orthoscopique et intersectant le même pixel.

Notre objectif est d'orthorectifier le MNS épipolaire, i.e. de réaliser une vue verticale pour chaque pixel de la grille orthoscopique. Ainsi, pour un pixel donné de la grille, nous affecterons l'altitude maximale parmi celles calculées par les différents triangles incluant ce point.

3.4.1.2 Algorithme

L'algorithme d'orthorectification directe avec interpolation triangulaire est le suivant :

1. Initialisation de l'image du MNS orthoscopique " MNS_{ortho} " à -1 (-1 est la valeur symbolisant l'indétermination : pas de valeur d'altitude renseignée) ; fixer la valeur du pas de la grille orthoscopique R en mètres.
2. Convertir les coordonnées géodétiques du MNS épipolaire " MNS_{epi} " en coordonnées cartographiques BNS selon [39, pp. 99]. Déterminer les valeurs E_{min} et N_{max} du MNS_{epi} converti.
3. Pour chaque pixel (u, v) de MNS_{epi} , on pose $B = (u + 1, v)$ et $C = (u, v + 1)$:
 - (a) Pour $A = (u, v)$ et $A = (u + 1, v + 1)$
 - i. Si les pixels A, B et C sont tous renseignés :
 - Projection du triangle (ABC) sur la grille orthoscopique selon l'équation (3.3) qui devient $(A'B'C')$.
 - Pour chaque pixel (i, j) de la grille orthoscopique inclus dans $(A'B'C')$:
 - Calculer l'altitude z_{ij} selon l'équation (3.4).
 - $MNS_{ortho}(i, j) = \max(MNS_{ortho}(i, j), z_{ij})$

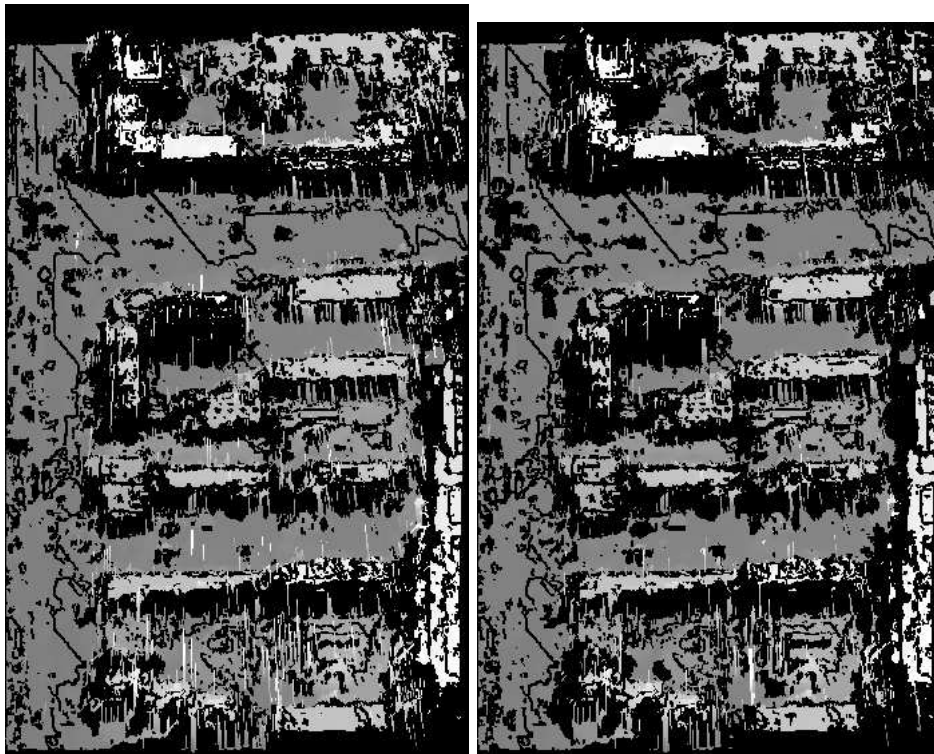
3.4.1.3 Résultats

La figure 3.21 illustre un MNS orthorectifié et projeté dans le système BNS avec un pas $R = 0.65$ m.



FIG. 3.21 – MNS orthorectifié avec un pas $R = 0.65$ m et $\varepsilon_{corr} = 0.4$.

Les commentaires relatifs au nombre important de pixels non renseignés en figure 3.21 ainsi que de leur localisation au niveau du bâti sont identiques à ceux formulés pour le MNS en géométrie épipolaire (figure 3.17). Cependant on constate l'apparition d'artefacts de type rayures claires. Ces artefacts sont inhérents à la technique d'interpolation triangulaire développée en section 3.4.1.1 ainsi qu'aux erreurs de corrélation contenues dans le MNS épipolaire. En effet, certains pixels du MNS épipolaire sont le résultat d'une corrélation erronée engendrant le calcul d'une altitude anormalement élevée par rapport à l'environnement du pixel. Les coordonnées planimétriques de ce point sont aussi fausses. Un triangle impliquant un, ou plusieurs de ces pixels corrompus inclura un nombre anormalement élevé de pixels lors de sa projection dans la grille orthoscopique. De plus, ces pixels recevront une altitude trop élevée ce qui explique l'aspect clair des artefacts. Il existe plusieurs moyens d'atténuer ce problème. Le premier consiste à modifier le seuil de réjection du score d'intercorrélation ε_{corr} intervenant lors de l'appariement des images stéréoscopiques. Le résultat présenté en figure 3.21 a été généré avec un seuil $\varepsilon_{corr} = 0.4$. Le danger d'augmenter ε_{corr} est de rendre la décision de renseigner un pixel plus sélective, et ainsi de voir s'agrandir les zones indéterminées.



(a) MNS ortho généré avec $\varepsilon_{corr} = 0.5$

(b) MNS ortho généré avec $\varepsilon_{corr} = 0.6$

FIG. 3.22 – MNS orthorectifiés avec différents seuils de corrélation.

La comparaison des figures 3.21, 3.22.a et 3.22.b montre la propagation des zones indéterminées en fonction du seuil de corrélation ε_{corr} . Les artefacts sont atténués lorsque ε_{corr} augmente.

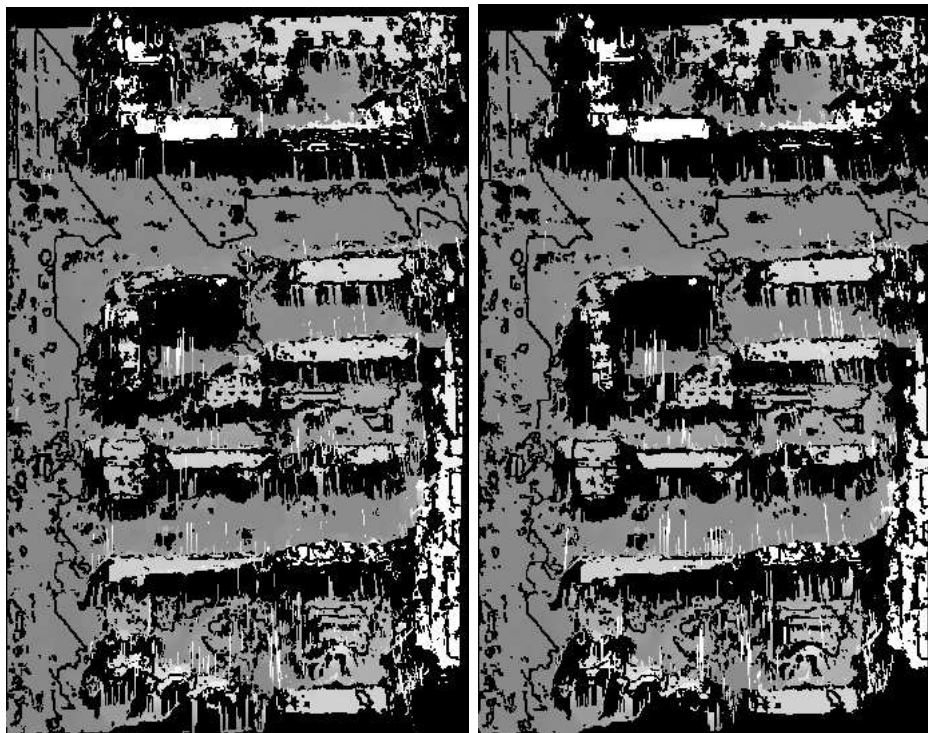
3.4.2 Pré-traitements sur l'image de disparité

Un autre moyen de réduire ces artefacts consiste à filtrer l'image de disparité qui sert à générer le MNS épipolaire. Deux filtres ont été essayés pour tenter de supprimer ces pixels erronés. Un filtrage médian de taille $N \times N$ ($N = 3$ ou 7) et un filtre de réjection de taille $N \times N$ ($N = 7$). Le filtre de réjection considère le pixel courant non déterminé s'il existe plus de x% de pixels non renseignés

dans la fenêtre l'entourant. Ce filtre a été conçu suite à l'observation que les pixels erronés sont souvent isolés et entourés par des pixels non renseignés.

3.4.2.1 Filtrage médian

Le filtrage médian dont les résultats sont illustrés en figure 3.23 a été mené uniquement sur les pixels renseignés de l'image de disparité. On constate que ce filtrage permet d'atténuer les artefacts sans les supprimer totalement. Appliquer une fenêtre plus grande ne réduit pas mieux les artefacts. Les artefacts sont aussi moins bien atténués que par l'augmentation du seuil de corrélation.



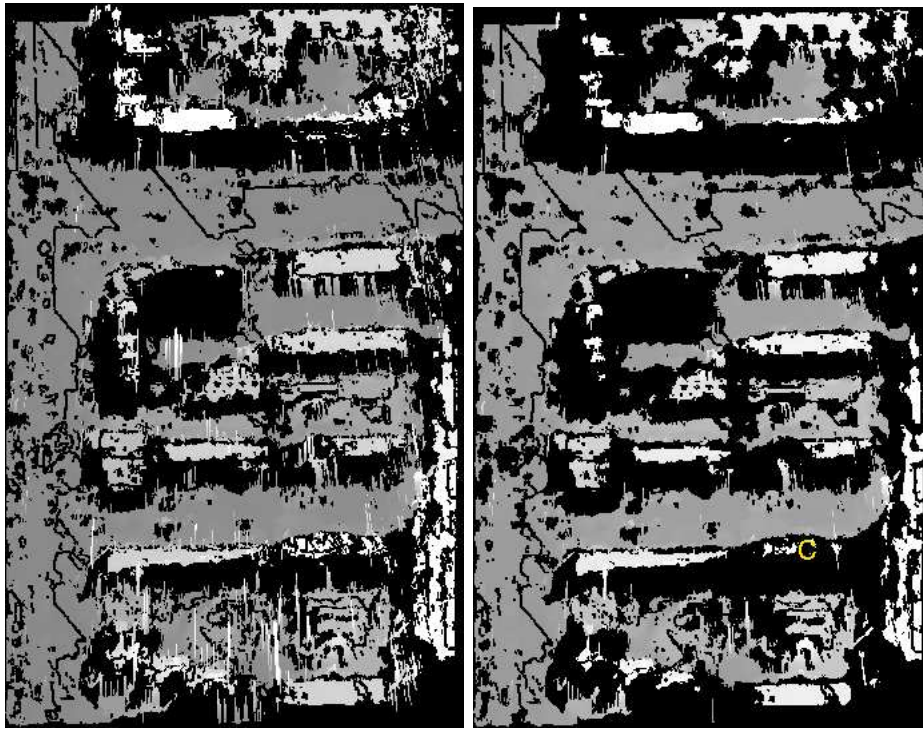
(a) Fenêtre 3×3

(b) Fenêtre 7×7

FIG. 3.23 – MNS ortho réalisé avec une image de disparité filtrée avec un filtre médian.

3.4.2.2 Filtre de réjection

La comparaison des figures 3.21, 3.24.a et 3.24.b montre clairement l'influence du seuil de réjection pour une fenêtre de taille fixée. Lorsque le seuil décroît, les artefacts sont sensiblement éliminés mais les zones non renseignées se répandent au point d'occulter de l'information pertinente (e.g. bâtiment rogné et annoté en figure 3.24.b par la lettre C).



(a) Réjection à 50 pour-cent

(b) Réjection à 30 pour-cent

FIG. 3.24 – MNS ortho réalisé avec une image de disparité filtrée avec un filtre de réjection de taille 7×7 .

Les bords de toits de bâtiments sont aussi affectés par le seuil de réjection du filtre. Alors que l'algorithme de corrélation implanté dans la chaîne de reconstruction 3D préserve les bords de bâtiments, ce filtre a tendance à les rogner ainsi qu'à fragmenter l'information d'altitude des toits.

3.4.2.3 Synthèse

Plusieurs méthodes ont été testées afin de réduire les artefacts dus aux pixels erronés de l'image de disparité. Aucune n'est parfaite dans l'absolu puisque chacune d'elles détériore sensiblement l'information contenue dans le MNS. Expérimentalement, nous avons observé que les meilleurs résultats étaient obtenus en utilisant consécutivement toutes les méthodes exposées et avec des paramètres ne détériorant pas trop l'image traitée. Ainsi nous pouvons augmenter légèrement le seuil de corrélation lors de la génération du MNS épipolaire ($\varepsilon_{corr} = 0.5$) afin d'éviter de générer trop de pixels indéterminés tout en supprimant quelques pixels erronés. L'image de disparité peut ensuite être filtrée par un filtre médian (3×3), puis par le filtre de réjection 7×7 avec un seuil de 60% de pixels non renseignés. La figure 3.26.b montre le résultat de l'orthorectification avec un tel pré-traitement de l'image de disparité sur une autre zone du MNS.

3.5 Fusion de MNSs

L'information précise de géocodage contenue dans chaque MNS orthorectifié va nous permettre d'effectuer la fusion des deux MNS géocodés. L'algorithme développé à cet effet s'inspire largement de la fonction *correlar* utilisée par la chaîne de reconstruction 3D. A la différence de la fonction *correlar* qui fusionne des images de disparités, nous fusionnons des MNS orthorectifiés et géocodés.

Cet algorithme tente de fusionner au mieux ces MNS en tenant compte de la cohérence inter-MNS (validation de l'information si les valeurs émanant d'autres MNS sont proches) et intra-MNS (validation de l'information de fusion vis-à-vis de sa cohérence par rapport au voisinage). Dans notre cas, la fusion s'apparente plutôt à de la "concaténation" ou encore "mosaïquage" puisque l'aire de recouvrement des MNS représente seulement un quart de leur surface. L'algorithme de fusion se veut le plus général possible en considérant la fusion de N et non seulement deux MNS orthorectifiés. Cet algorithme s'exprime de la façon suivante :

I) Détermination de la grille englobante. Chaque MNS a une étendue parfaitement connue dans le système cartographique dans lequel il est projeté. La première étape consiste à construire la grille géocodée du MNS final, résultat de la fusion. Cette grille est la plus petite grille contenant les N MNS à fusionner.

II) Boucle sur chaque pixel de la grille englobante :

Pour chaque pixel pix de la grille :

1. Test d'appartenance aux N MNSS

- (a) Pour $i \in \{1, \dots, N\}$ teste si $pix \in MNS_i$
- (b) Si pix n'appartient à aucun MNS : **rejet**
- (c) Si pix appartient à un seul MNS : **on assigne la valeur du MNS à pix** (que la valeur soit renseignée ou non)
- (d) Si pix appartient à m MNS dont m_r sont renseignés ($m \in \{2, \dots, N\}$)
 - i. Si $m_r = 0$, **rejet**
 - ii. Si $m_r = 1$, **on assigne la valeur de l'unique MNS renseigné en ce point à pix**
 - iii. Si $m_r \in \{2, \dots, m\}$
 - A. Si les m_r points des m_r MNS sont cohérents entre eux :
 - * **On assigne à pix la valeur moyenne des points cohérents par rapport à leur voisinage³**
 - * Si aucun point n'est cohérent avec le voisinage : **rejet**
 - B. Si les $m_r - k$ points des m_r MNS sont cohérents entre eux ($k \in \{1, \dots, m_r - 2\}$), on choisit les deux points les plus cohérents entre eux :
 - * Si ce deux points sont cohérents par rapport au voisinage : **on assigne à pix leur valeur moyenne.**
 - * Si un seul est cohérent par rapport au voisinage : **on assigne à pix la valeur de ce point.**
 - * Sinon : **rejet**
 - C. Si les $m_r - k$ points des m_r MNS sont incohérents entre eux :
 - * Si la cohérence par rapport au voisinage est calculable : **on assigne à pix la valeur du point le plus cohérent par rapport au voisinage.**
 - * Sinon : **rejet**

III) Fin.

³On effectue la même opération s'il est impossible de calculer la cohérence par rapport au voisinage des m_r points.

Dans cet algorithme, les notions de cohérence sont définies de la façon suivante :

- Cohérence entre deux points p_i et p_j émanant de deux MNS_i , MNS_j : p_i et p_j sont cohérents entre eux $\iff |altitude(p_i) - altitude(p_j)| < \varepsilon_{inter-coherence}$.
- Cohérence par rapport au voisinage d'un point p_i émanant de MNS_i : p_i est cohérent par rapport au voisinage $\iff |altitude(p_i) - altitude(voisinage)| < \varepsilon_{intra-coherence}$.

L'altitude du voisinage est la moyenne des points du MNS fusionné (résultat) localisés dans une fenêtre causale (le pixel d'intérêt se trouve en bas à droite de la fenêtre). Typiquement, les calculs ont été menés avec une fenêtre de 5×5 pixels. Puisque que les MNS possèdent des pixels non renseignés pouvant entraver le calcul de la moyenne de la cohérence par rapport au voisinage (s'il existe trop peu de points renseignés dans la fenêtre, ce calcul n'a pas de sens), on considère que cette cohérence est "incalculable" si plus de 50% des pixels de la fenêtre sont non renseignés.

Le résultat de la fusion de MNSs illustré en figures 3.25-3.26 a été réalisé avec $\varepsilon_{inter-coherence} = \varepsilon_{intra-coherence} = 5m$. La figure 3.25 illustre la fusion des deux MNS. La figure 3.26 montre plus précisément la zone encadrée en rouge.

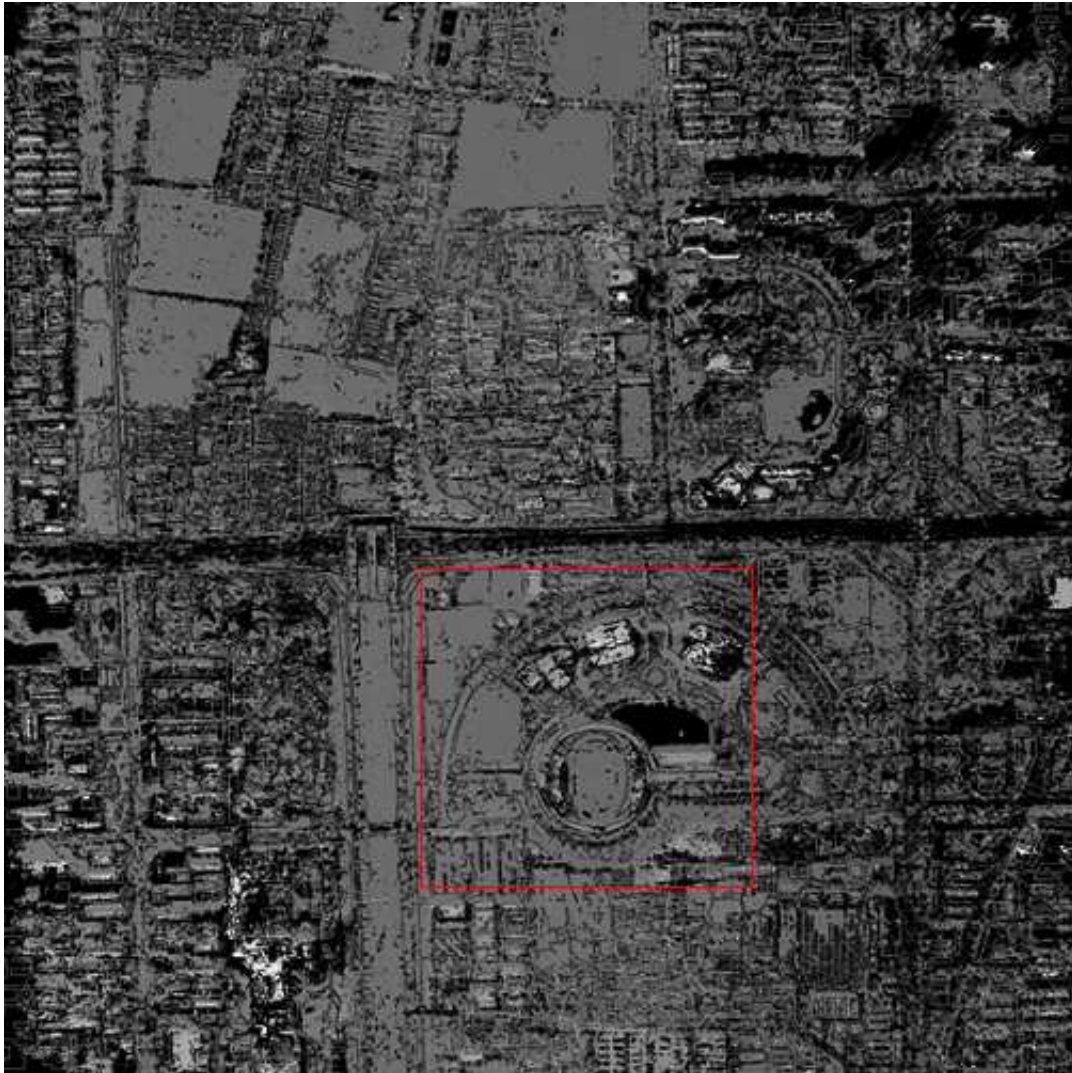


FIG. 3.25 – MNS fusionné $MNS_1 + MNS_2$.



(a) Extrait de l'image aérienne

(b) MNS fusionné correspondant

FIG. 3.26 – Agrandissement de la zone rectangulaire rouge de la figure 3.25.

3.6 Conclusion

Ce chapitre a permis de présenter les travaux effectués sur la génération de MNSs orthoscopiques. La principale contribution de cette étude est l'optimisation de paramètres de prise de vue d'images aériennes. Ces paramètres connus *a priori* avec une précision insuffisante ont été optimisés par l'algorithme du simplex grâce à la donnée de GCPs et de TPs. Une attention particulière a été portée sur l'évaluation quantitative de la précision de géocodage des MNS générés par la chaîne de reconstruction avec les paramètres optimisés. Les MNSs reconstruits atteignent une précision planimétrique et altimétrique de l'ordre du mètre, ce qui valide la méthode d'optimisation des paramètres de prise de vue. Il est néanmoins important de remarquer que la qualité de l'optimisation dépend du nombre et de la répartition uniforme des GCPs utilisés. Nous avons également détaillé l'orthorectification et la fusion de plusieurs ortho-MNSs géocodés. Le MNS fusionné et orthoscopique final peut dès lors être superposé avec les données cartographiques et l'image satellitaire grâce à l'information de géocodage. Cette superposition sera d'autant plus cohérente avec les données cartographiques qui sont aussi en géométrie orthoscopique.

Enfin, nous avons pu remarquer, grâce aux résultats illustrés, que la reconstruction 3D produit des MNSs ayant de nombreux pixels non renseignés. Dans le MNS de la figure 3.25, ils représentent 33% de l'image. Ces zones indéterminées sont principalement localisées autour des bâtiments en raison des occlusions. En effet, deux images ne sont pas suffisantes pour renseigner l'altitude de l'intégralité de la scène. La tendance actuelle est d'utiliser des triplés ou des quadruplés d'images pour restituer le relief d'un paysage [136]. Les MNS générés dans cette étude ont ainsi un aspect grossier et semblent difficilement utilisables. Les erreurs de corrélation y contribuent aussi. Celles-ci sont inhérentes à la difficulté des milieux urbains qui présentent de nombreux détails mettant en échec le corrélateur. Il est aussi à noter que les conditions de rapport signal à bruit n'étaient pas optimales : les images étant analogiques, elles ont été scannées. La numérisation introduit inévitablement du bruit. Nous avons remarqué que la bande spectrale bleue était particulièrement bruitée, ce qui est dû à un défaut soit de la caméra, soit du scanner. Néanmoins, nous verrons au chapitre suivant que l'aspect imparfait de l'ortho-MNS produit ne sera pas un obstacle à son utilisation comme source d'information supplémentaire complétant l'image satellitaire.

Chapitre 4

Présentation et choix de contours actifs avec contrainte de forme pour le recalage fin carte-image

4.1 Introduction

Ce chapitre présente les modèles déformables appelés contours actifs. Ils seront destinés à mettre en correspondance les bâtiments symbolisés dans la carte avec leur représentation homologue dans une image satellitaire panchromatique haute résolution. Nous exploitons la flexibilité des contours actifs ainsi que leur potentiel à incorporer de l'information de haut niveau pour mener à bien ce recalage fin dont l'objectif est double : i) les objets cartographiques auront une localisation spatiale affinée si l'image est de meilleure précision que la donnée cartographique, ce qui correspond à l'une des composantes de la *mise à jour de cartes*. ii) la mise en correspondance permet d'amoindrir les *variabilités exogènes* carte-image qui sont sources d'erreurs pour toute détection de changement subséquente. Les scènes urbaines représentées dans les images satellitaires Quickbird utilisées comportent un niveau de détail élevé ainsi que certains *artefacts urbains* uniquement remarquables à très haute résolution (occlusions, ombres, faible contraste des objets,...). Afin de surmonter ces difficultés nous proposons de tirer avantage de la connaissance *a priori* contenue dans la carte. De façon générale, la connaissance fournie par la carte revêt trois aspects. Elle permet de renseigner la nature de l'objet à recaler dans l'image (bâtiment, route, ...), sa localisation (les données cartographiques et de télédétection sont supposées globalement superposées), et sa forme. L'information de localisation permettra d'initialiser le contour actif proche du bâtiment représenté dans l'image, et sa forme sera *contrainte* par la silhouette du bâtiment symbolisé dans la carte afin de surmonter les difficultés des images urbaines. Dans une première partie, ce chapitre décrira un état de l'art des contours actifs. Nous classerons les différents modèles selon leur mode de représentation, d'attache aux données et de régularisation, avec une attention spéciale pour l'insertion de contrainte de forme exogène connue *a priori*. Enfin, nous motiverons notre choix pour une représentation par ensemble de niveaux du contour actif et exposerons les modèles d'attache aux données et de contrainte de forme qui seront employés dans cette étude.

4.2 État de l’art des contours actifs contraints par une forme *a priori*

4.2.1 Les contours actifs

Les contours actifs sont des techniques de segmentation permettant d’extraire un objet d’intérêt d’une image. Cette segmentation n’est pas immédiate, elle requiert une phase dynamique du contour (d’où la dénomination “actif”) qui évoluera itérativement au cours du temps artificiel t , de sa position initiale vers les bords de l’objet à extraire. Une telle évolution temporelle peut se formaliser mathématiquement sous la forme d’une *équation d’évolution* exprimant explicitement ou implicitement la vitesse du contour actif. Il existe plusieurs manières d’obtenir une équation d’évolution. L’une consiste à dériver cette équation de la minimisation d’une fonctionnelle d’énergie, on parle alors d’approche variationnelle. Une alternative consiste à construire une équation d’évolution par analogie avec d’autres disciplines scientifiques comme la Physique, c’est l’approche géométrique. Nous nous attacherons à décrire, puis utiliser l’approche variationnelle dans ce travail de thèse. Avec une telle approche, la fonctionnelle d’énergie peut être schématisée comme la somme de deux classes de termes énergétiques. La première classe concerne l’énergie interne du contour visant à contrôler des contraintes intrinsèques à celui-ci, comme par exemple sa régularité. Nous verrons en section 4.2.2 qu’il est possible d’insérer des contraintes géométriques plus spécifiques dérivées de la connaissance *a priori* que l’on a sur l’objet à segmenter dans l’image. La deuxième classe est relative au *terme d’attache aux données* qui fera interagir le contour actif avec des caractéristiques extraites de l’image. Dans les travaux de Foulonneau [40] et Jehan-Besson [56], ce terme d’énergie externe est appelé *critère* et est construit à partir de *descripteurs*. Un descripteur est une mesure faite sur l’image permettant de caractériser une frontière ou une région. Un descripteur de frontière serait par exemple la carte des gradients de l’image, un descripteur d’une région R pourrait être la moyenne des pixels de l’image inclus dans R . Selon le choix du descripteur adopté dans la fonctionnelle d’énergie, on dérivera différents types de contours actifs plus ou moins performants en fonction de la nature de l’image à analyser.

On voit donc se dessiner différentes catégories de contours actifs. Celles-ci se différencient par la façon avec laquelle on déduit l’équation d’évolution, par le mode de représentation du contour actif, et finalement le terme d’attache aux données qui peut être soit basé sur les frontières, les régions ou bien les deux. L’objectif de cette section est de détailler et d’expliquer ces différents moyens de classer les contours actifs. Dans le reste de ce document, nous nous restreignons au cas des contours actifs bi-dimensionnels évoluant dans le plan. Nous conseillons au lecteur les travaux de [13, 111, 88, 109, 56, 103, 40] pour un exhaustif état de l’art des contours actifs.

4.2.1.1 Représentation d’un contour actif

Représentation explicite

Historiquement, c’est la représentation explicite du contour actif qui a émergé en premier dans la communauté de Vision par Ordinateur grâce aux travaux pionniers de Kass, Witkin et Terzopoulos en 1987 [58]. Une telle représentation consiste à paramétrer le contour actif¹ $\mathbf{C}$ par un paramètre arbitraire p et le temps t : $\mathbf{C}(p, t)$.

$$\mathbf{C} = \{ \mathbf{C}(p, t) \in \mathbb{R}^2, C^2 \mid p \in [a, b] \subset \mathbb{R}, t \in \mathbb{R}^+ \} \quad (4.1)$$

Il est important de remarquer qu’il existe d’autres moyens de représenter un contour de façon

¹Nous considérerons dans ce document que $\mathbf{C}$ est un contour orienté dans le sens trigonométrique.

explicite (B-Splines, base de Fourier,...). Nous nous attacherons cependant à expliciter celle décrite en équation (4.1) par souci de simplicité. L'équation d'évolution de $\mathbf{C}$ peut alors être formalisée de façon générale, en décomposant la vitesse de $\mathbf{C}(p, t)$ selon sa composante normale $\mathbf{N}$ et tangentielle $\mathbf{T}$ en (p, t) dans un repère de Frenet $(\mathbf{T}, \mathbf{N})$:

$$\frac{\partial \mathbf{C}(p, t)}{\partial t} = v_N(p, t) \mathbf{N}(p, t) + v_T(p, t) \mathbf{T}(p, t) \quad (4.2)$$

où : $\mathbf{N}(p, t)$ est le vecteur unitaire normal au contour $\mathbf{C}$ en $\mathbf{C}(p, t)$; $\mathbf{T}(p, t)$ est le vecteur unitaire tangent au contour $\mathbf{C}$ en $\mathbf{C}(p, t)$. Dans [109], Sapiro montre que si la vitesse normale ne dépend pas du paramétrage, alors la composante tangentielle de la vitesse n'influence pas la déformation du contour actif, mais uniquement son paramétrage. On peut donc, sans perte de généralité, supposer dans le reste de ce document que le contour actif évolue selon sa normale :

$$\frac{\partial \mathbf{C}(p, t)}{\partial t} = v_N(p, t) \mathbf{N}(p, t) \quad (4.3)$$

La définition présentée en équation (4.1) correspond à une courbe $\mathbf{C}$ continue du plan. Puisqu'une image est physiquement représentée de façon discrète par une grille régulière pixellaire, le contour $\mathbf{C}$ se doit d'être aussi discrétisé. En pratique, la discrétisation du contour est un sous-échantillonnage de ce dernier en fonction de la variable p (figure 4.1).

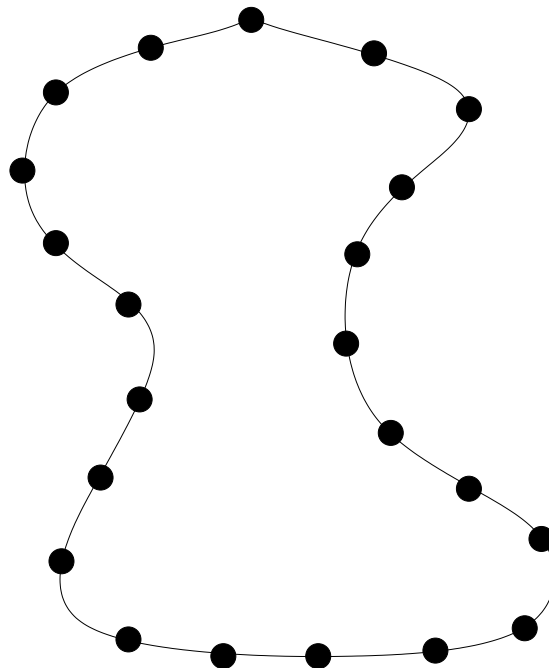


FIG. 4.1 – Sous-échantillonnage d'un contour représenté de façon explicite et paramétré par son abscisse curviligne.

Le contour actif se résume alors à une liste de n œuds dont la position est mise à jour au fil de l'évolution de $\mathbf{C}$ selon l'équation (4.3) qui devient :

$$\mathbf{C}(m_i, t + 1) = \mathbf{C}(m_i, t) + \Delta t v_N(m_i, t) \mathbf{N}(m_i, t) \quad (4.4)$$

avec m_i le $i^{\text{ème}}$ œud parmi N . Si l'on suppose que le nombre de n œuds N est fixe au cours du temps, on entrevoit les problèmes suivants : si la longueur de $\mathbf{C}$ croît au cours du temps, l'approxi-

mation polygonale de l'objet à segmenter sera grossière ; si la longueur diminue, on risque d'avoir des nœuds extrêmement proches les uns des autres ce qui provoquera des instabilités numériques si des dérivées aux différences finies sont calculées le long de $\mathbf{C}$ pour le calcul de v_N (son dénominateur tend alors vers 0). Diminuer le pas temporel Δt de l'équation (4.4) permet d'éviter ces instabilités mais mène à l'allongement des temps de calcul qui deviennent rapidement impraticables. Il existe trois façons de résoudre ce problème : i) le rééchantillonnage des nœuds de $\mathbf{C}$ afin de garder une distance inter-nœuds constante, ii) ajouter des termes de diffusion dans l'expression de la vitesse visant à adoucir la courbure du contour actif (la distance inter-nœuds est alors minorée et non-nulle), iii) filtrer les oscillations dues aux instabilités. Ces solutions non intrinsèques sont sous-optimales, créent d'autres problèmes indépendants de la problématique de segmentation, et influencent fortement le résultat d'extraction d'objet. Deux autres problèmes propres à la représentation explicite subsistent : le contour $\mathbf{C}$ ne peut changer de topologie, de plus, des artefacts de croisement du contour peuvent apparaître. Le premier problème est résolu par un re-paramétrage du contour actif [73]. Le second inconvénient peut être traité par des techniques de débouclage [51].

Représentation implicite

En 1988, les travaux de Osher et Sethian [85] présentent une façon moins naturelle de représenter un contour. C'est une représentation implicite venant de la Physique des interfaces, appelée "ensemble de niveaux" ou encore *level sets* en anglais. Le principe est le suivant : on représente un contour fermé $\mathbf{C}(t)$ comme le zéro d'une fonction à $n + 1$ dimensions (dans le plan, $n = 2$, cf. figure 4.2). Originellement, c'est une fonction de distance Euclidienne signée qui a été proposée pour représenter la $(n + 1)^{\text{ième}}$ dimension de l'ensemble de niveaux. Cependant, d'autres fonctions peuvent être choisies, ainsi Haker *et al.* représentent une surface 3D comme le zéro d'une fonction u solution de l'équation harmonique de Laplace $\Delta u = 0$ [49].

Dans le cadre des travaux exposés en [85], on appelle ϕ la fonction d'ensemble de niveaux vérifiant les propriétés suivantes :

- ϕ est une fonction de Lipschitz à valeurs réelles, $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$
- quel que soit le temps t , le zéro de $\phi(\mathbf{x}, t)$ est le contour fermé $\mathbf{C}(t)$:

$$\forall \mathbf{x} \in \mathbb{R}, \mathbf{C}(t) = \{\mathbf{x} \in \mathbb{R}^2 \mid \phi(\mathbf{x}, t) = 0\} \quad (4.5)$$

- $\phi(\mathbf{x})$ est la distance signée entre le point $\mathbf{x}$ et le contour $\mathbf{C}(t)$. Le signe de $\phi(\mathbf{x})$ dépend de l'appartenance de $\mathbf{x}$ à la région interne ou externe au contour que nous noterons respectivement $\Omega_{in}(t)$ et $\Omega_{out}(t)$. Par convention, nous choisirons $\phi(\mathbf{x})$ positive à l'intérieur de $\mathbf{C}(t)$ et négative à l'extérieur :

$$\begin{cases} \phi(\mathbf{x}, t) = 0 & \text{si } \mathbf{x} \in \mathbf{C}(t) \\ \phi(\mathbf{x}, t) = d(\mathbf{x}, \mathbf{C}(t)) & \text{si } \mathbf{x} \in \Omega_{in}(t) \\ \phi(\mathbf{x}, t) = -d(\mathbf{x}, \mathbf{C}(t)) & \text{si } \mathbf{x} \in \Omega_{out}(t) \end{cases} \quad (4.6)$$

avec $d(\mathbf{x}, \mathbf{C}(t))$ la plus petite distance Euclidienne du point $\mathbf{x}$ au contour $\mathbf{C}(t)$:

$$d(\mathbf{x}, \mathbf{C}(t)) = \min_{\mathbf{x}_c \in \mathbf{C}(t)} |\mathbf{x} - \mathbf{x}_c| \quad (4.7)$$

Une telle représentation possède certains avantages sur la représentation explicite. Premièrement, un contour représenté par un ensemble de niveaux peut changer de topologie au cours du temps, il n'est donc pas nécessaire de la connaître *a priori*. L'équation d'évolution est plus stable numériquement et évite le remaillage des nœuds de la représentation explicite. Le problème des boucles du contour est naturellement résolu par la topologie flexible de $\mathbf{C}(t)$. Enfin, l'ensemble de niveaux

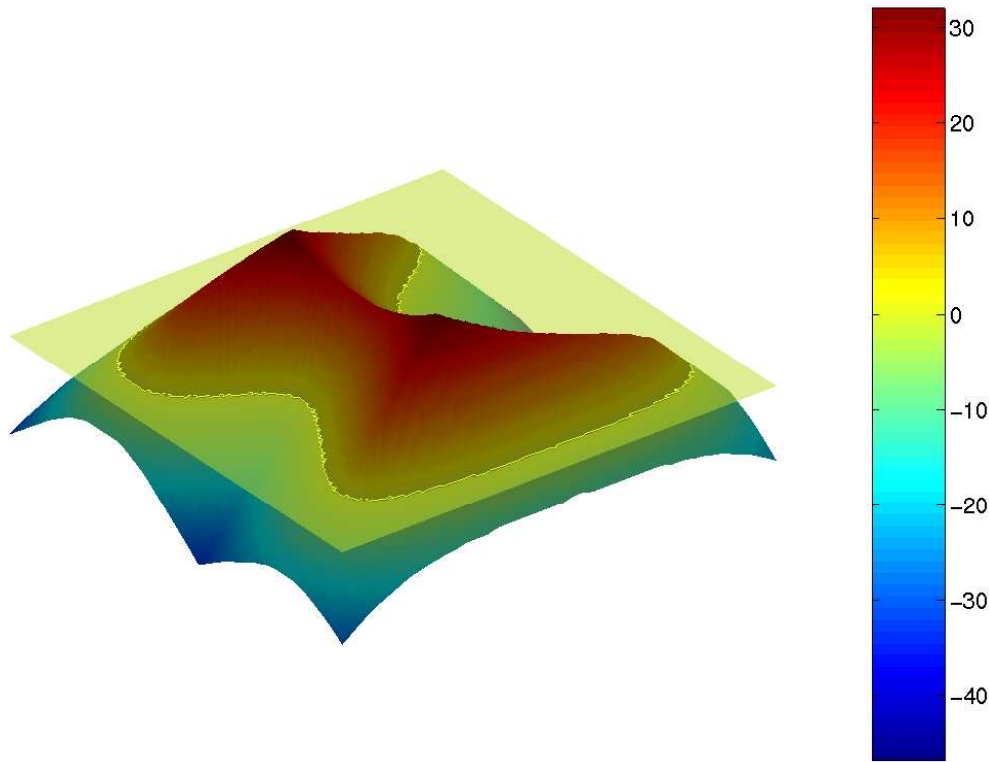


FIG. 4.2 – Représentation à trois dimensions de l'intersection d'un ensemble de niveaux avec le plan de l'image (niveau zéro). Le contour représenté implicitement est celui de la figure 4.1. La légende représente la distance Euclidienne signée au contour.

permet d'accéder à de nombreuses propriétés géométriques intéressantes et intrinsèques au contour, ou de n'importe quelle ligne de niveau. Ainsi on peut déterminer la normale à ϕ en n'importe quel point de l'image, et donc *a fortiori* en n'importe quel point du contour pour le peu qu'on connaisse l'ensemble des points tels que $\phi(\mathbf{x}, t) = 0$. La normale intérieure au contour est donnée par :

$$\mathbf{N}(\mathbf{x}) = \frac{\nabla \phi(\mathbf{x})}{|\nabla \phi(\mathbf{x})|} \quad (4.8)$$

La courbure κ vient aussi naturellement :

$$\kappa = \text{div}(\mathbf{N}) = \nabla \cdot \left(\frac{\nabla \phi(\mathbf{x})}{|\nabla \phi(\mathbf{x})|} \right) \quad (4.9)$$

$$= \frac{\phi_{xx}(\mathbf{x}) \phi_y^2(\mathbf{x}) - 2\phi_x(\mathbf{x}) \phi_y(\mathbf{x}) \phi_{xy}(\mathbf{x}) + \phi_{yy}(\mathbf{x}) \phi_x^2(\mathbf{x})}{(\phi_x^2(\mathbf{x}) + \phi_y^2(\mathbf{x}))^{3/2}} \quad (4.10)$$

Les ensembles de niveaux permettent aussi d'effectuer des opérations logiques sur des ensembles telles que le calcul d'union ou d'intersection.

L'équation d'évolution ne concerne plus explicitement le contour $\mathbf{C}(t)$ comme en équation (4.2), mais l'ensemble de niveaux. On déduit l'équation d'évolution de $\phi(\mathbf{x}, t)$ en différenciant l'équation (4.5) par rapport au temps t :

$$\forall \mathbf{x}(t) \in \mathbf{C}(t), \quad \frac{\partial \phi(\mathbf{x}(t), t)}{\partial t} = 0 \quad (4.11)$$

Il vient alors :

$$\frac{\partial \phi(\mathbf{x}, t)}{\partial t} + \left\langle \nabla \phi(\mathbf{x}, t), \frac{\partial(\mathbf{x}(t))}{\partial t} \right\rangle = 0 \quad (4.12)$$

$$\phi_t(\mathbf{x}, t) + \langle \nabla \phi(\mathbf{x}, t), \mathbf{x}_t(t) \rangle = 0 \quad (4.13)$$

En décomposant l'expression de $\mathbf{x}_t(t)$ de façon similaire à l'équation (4.2), on a finalement :

$$\phi_t(\mathbf{x}, t) = - \langle \nabla \phi(\mathbf{x}, t), v_N(\mathbf{x}, t) \mathbf{N}(\mathbf{x}, t) \rangle \quad (4.14)$$

Il est important de remarquer que la formulation implicite s'affranchit du paramétrage p du contour actif $\mathbf{C}(t)$, ce qui posait un problème de stabilité numérique et de topologie en représentation explicite. La discrétisation de la fonction $\phi(\mathbf{x}, t)$ peut donc se faire sur une grille régulière de pas Δx et Δy (approche Eulérienne), contrairement à la discrétisation Lagrangienne explicite selon p plus délicate (figure 4.1). Pour un contour donné $\mathbf{C}(t)$, l'ensemble de niveaux associé revient au calcul de la distance signée à ce contour (équation (4.6)). Pour un ensemble de niveaux donné, si l'on veut retrouver le contour $\mathbf{C}(t)$, il suffit de détecter les passages à zéro de $\phi(\mathbf{x}, t)$ (équation (4.5)). Le contour extrait est alors une approximation linéaire du véritable contour représenté par $\phi(\mathbf{x}, t)$ (figure 4.5).

Remarques :

- Il est important de remarquer qu'un ensemble de niveaux n'est pas juste un artifice de calcul permettant de mettre en œuvre différemment l'équation d'évolution d'un contour actif. Un ensemble de niveaux est une représentation intrinsèquement différente de ses homologues explicites, qui par conséquent aboutit à une résolution numérique alternative.
- La représentation implicite présente une perte d'information spatiale par rapport à la représentation par n œuds (*snakes*). Alors qu'il est possible de savoir si l'on est situé sur le contour (niveau zéro), il est en revanche impossible de déterminer *où* on se situe sur ce dernier. La relation d'ordre entre les points du contour est perdue. Le problème des n œuds est dual : ils permettent la localisation sur le contour puisqu'ils sont organisés sous forme d'une séquence ordonnée. Cependant la connaissance de la distance de n'importe quel pixel de l'image au contour n'est pas directe.
- Remarquons qu'une représentation explicite par n œuds permet de modéliser des contours fermés ou ouverts. Il est plus difficile de traiter les contours ouverts dans le cas des ensembles de niveaux.
- Enfin, la complexité calculatoire des ensembles de niveaux est plus grande que les méthodes explicites. Ceci est inhérent à l'encodage sous forme d'une fonction de distance signée de dimension $n + 1$ (n étant la dimension de l'image), alors que la façon explicite gère un nombre de n œuds beaucoup plus réduit.

4.2.1.2 Approche géométrique et variationnelle

Approche géométrique

Lorsqu'un mode de représentation implicite ou explicite a été choisi, il convient de déterminer la vitesse normale $v_N(\mathbf{x}, t)$ qui permettra de déformer le contour actif. La détermination de ce champ vectoriel de vitesse est le point crucial des contours actifs puisque leur habilité à segmenter l'objet d'intérêt représenté dans une image en dépend.

L'approche géométrique, par analogie avec la Physique, ou par des considérations purement mathématiques, établit directement l'équation d'évolution du contour actif. Dans [17] et [69, 70], les

auteurs décrivent un contour actif soumis au flux Euclidien de la chaleur ou de la courbure moyenne κ (analogie avec l'équation de propagation de la chaleur) : $v_N = \kappa$. Ce flux a des propriétés diffusantes et lissantes et évite au contour actif de présenter des singularités [111]. Ceci n'est qu'une contrainte de régularité interne au contour actif. Pour faire évoluer le contour, les auteurs introduisent une vitesse normale constante : $v_N = \kappa + c$. Cette constante est analogue à l'effet d'une force de gonflage ou de rétraction introduite par Cohen *et al* [27]. Le terme d'attache aux données de l'image est une fonction g pondérant la vitesse normale. Dans leurs expériences, les auteurs utilisent un terme basé sur les frontières de l'objet à segmenter : lorsque le contour atteint des zones de gradient élevé de l'image, la fonction de pondération tend vers zéro. L'expression de la vitesse normale est finalement :

$$v_N = g(|\nabla I|)(\kappa + c) \quad (4.15)$$

Le principal problème de cette approche est la sensibilité à l'initialisation du contour actif. Ceci est inhérent au terme d'attache aux données basé sur les frontières ainsi qu'à la force de gonflage/rétraction univoque durant l'évolution du contour. L'initialisation requiert une forte connaissance *a priori* de la localisation de l'objet à extraire. Cette technique s'est avérée performante en approche semi-automatique sur des images médicales.

Approche variationnelle

L'approche variationnelle consiste à formuler une fonctionnelle d'énergie J dont la minimisation par calcul des variations fournira l'équation d'évolution du contour actif $\mathbf{C}(t)$. La fonctionnelle est composée de termes d'énergie, ou *critères*, qui sont intrinsèques (contraintes internes sur $\mathbf{C}(t)$) ou extrinsèques (attache aux données). Alors que les critères intrinsèques sont souvent basés sur la frontière définie par $\mathbf{C}(t)$, les critères extrinsèques sont soit basés sur les frontières ou les régions. Nous utiliserons cette dichotomie opérée sur les critères extrinsèques pour différencier les contours actifs basés sur l'information de frontière de ceux basés sur l'information de région. La technique de descente de gradient est traditionnellement utilisée pour déduire l'équation d'évolution du contour $\mathbf{C}$ à partir du calcul des variations :

$$\frac{\partial \mathbf{C}}{\partial t} = -\frac{\partial J}{\partial \mathbf{C}} \quad (4.16)$$

Un contour actif évoluant selon l'équation (4.16) tendra à minimiser l'énergie J dont le minimum correspond à la segmentation des objets recherchés dans l'image.

4.2.1.3 Contours actifs basés sur l'information de frontière

Un contour actif basé sur l'information de frontière tient uniquement compte de l'information présente dans un voisinage de $\mathbf{C}(t)$. Ainsi, seule l'information présente au niveau des bords de l'objet à segmenter sera utilisée. Pour formuler une fonctionnelle, il convient premièrement de choisir un *descripteur* de frontière que l'on appellera $k_b(\mathbf{x})$. Un tel descripteur peut être la carte du module du gradient de l'image, ou le champ de vecteurs de gradient en chaque pixel de l'image I . Le critère construit à partir du descripteur est alors l'intégrale le long du contour $\mathbf{C}(t)$:

$$J_b(\mathbf{C}(t)) = \int_{\mathbf{C}(t)} k_b(\mathbf{x}) d\mathbf{a}(\mathbf{x}) \quad (4.17)$$

où $d\mathbf{a}(\mathbf{x})$ est un élément d'aire. Nous choisissons de présenter deux types de contours actifs basés frontière : les *snakes* et les *contours géodésiques*.

- Les snakes

Dans [58], Kass, Witkin et Terzopoulos en 1987 représentent le contour actif de façon explicite en formulant la fonctionnelle comme la somme de trois termes basés frontière :

$$J_b(\mathbf{C}(p, t)) = \alpha \int_a^b \left| \frac{\partial \mathbf{C}(p, t)}{\partial p} \right|^2 dp + \beta \int_a^b \left| \frac{\partial^2 \mathbf{C}(p, t)}{\partial p^2} \right|^2 dp - \gamma \int_a^b |\nabla I(\mathbf{C}(p, t))|^2 dp \quad (4.18)$$

avec $\{\alpha, \beta, \gamma\} \in \mathbb{R}^+$. Les deux premiers termes sont des critères intrinsèques, cas particulier de l'opérateur de Tikhonov développé à l'ordre deux et avec des poids w_k constants :

$$J_{b,Tikhonov}(\mathbf{C}(p, t)) = \sum_{k=0}^n \int_{\mathbf{C}(t)} w_k(p) \left| \frac{\partial^k \mathbf{C}(p, t)}{\partial p^k} \right|^2$$

Le premier critère intrinsèque est assimilable à une contrainte élastique sur les n œuds de discrétisation de $\mathbf{C}(p, t)$, entraînant les n œuds voisins à se rapprocher. Le second critère intrinsèque est une contrainte de rigidité sur le contour actif, régulant sa courbure (modèle des plaques minces). Le dernier terme de l'équation (4.18) est le critère extrinsèque d'attache aux données basé sur le module du gradient de l'image. On note que plus $\mathbf{C}(p, t)$ sera placé sur des lieux de gradient élevé de l'image, plus la fonctionnelle globale $J_b(\mathbf{C}(p, t))$ sera faible. Dans une approche variationnelle, la seconde étape consiste à dériver l'équation d'évolution du contour actif dont la solution est un minimum de la fonctionnelle. Dans [58], les auteurs utilisent les équations d'Euler-Lagrange afin de minimiser $J_b(\mathbf{C}(p, t))$. Dans le formalisme de Lagrange, S désigne "l'action" et est définie par :

$$S = \int \mathcal{L}(t, q_i, \dot{q}_i) dt \quad (4.19)$$

où $\mathcal{L}$ est le Lagrangien, différence entre l'énergie cinétique et potentielle généralisée d'un système mécanique à n particules. $\mathcal{L}$ dépend des positions et des vitesses des particules ainsi que du temps. Minimiser "l'action" (principe de moindre action) consiste à annuler une variation infinitésimale $\tilde{S}$ de S engendrée par une variation infinitésimale $\tilde{\mathcal{L}}$ du Lagrangien. La minimisation de S revient à résoudre les n équations différentielles suivantes :

$$\frac{\partial \mathcal{L}}{\partial q_i} = \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right), \quad i \in \{1, \dots, n\} \quad (4.20)$$

Dans l'hypothèse d'un contour actif fermé (conditions cycliques aux limites), l'équation (4.20) revient à résoudre l'équation aux dérivées partielles suivante :

$$\alpha \frac{\partial^2 \mathbf{C}(p, t)}{\partial p^2} + \beta \frac{\partial^4 \mathbf{C}(p, t)}{\partial p^4} + \gamma \frac{\partial \int_{\mathbf{C}(t)} |\nabla I(\mathbf{C}(p, t))|^2 dp}{\partial \mathbf{C}(p, t)} = 0 \quad (4.21)$$

En discrétisant le contour actif en n œuds m_i , les auteurs de [58] résolvent l'équation (4.21) selon les différences finies et aboutissent *in fine* à l'équation matricielle suivante :

$$\mathbf{A}\mathbf{V} + \mathbf{f}_v = 0$$

avec $\mathbf{A}$ une matrice de Toeplitz pentadiagonale, $\mathbf{V} = (\mathbf{C}(m_0), \dots, \mathbf{C}(m_n))$ et

$$\mathbf{f}_v = \left(\frac{\partial \int_{\mathbf{C}(t)} |\nabla I(\mathbf{C}(p, t))|^2 dp}{\partial \mathbf{C}(m_i)} \right)_{i \in \{1, \dots, n\}}$$

La solution recherchée est $\mathbf{V}$, cependant, $\mathbf{A}$ n'est pas inversible (0 est une valeur propre de $\mathbf{A}$). On cherche alors la solution de manière itérative. La méthode proposée dans [58] rencontre des instabilités numériques, en particulier à cause du calcul de la dérivée d'ordre quatre de l'équation (4.21). De plus, le contour actif a tendance à se contracter sur la partie du contour qui a le plus fort gradient. Dans [126], les auteurs proposent un algorithme rapide évitant les écueils de [58]. Le principal apport de cette méthode est de proposer une alternative aux estimées des dérivées premières et secondes par différences finies dans la fonctionnelle, accroissant ainsi la rapidité de calcul et la précision des résultats. Les coefficients de pondération (α, β) de ces dérivées sont aussi variables au cours des itérations : selon un critère relatif à la courbure (seuillage) et à la valeur du plus fort gradient à proximité du noyau courant, leurs valeurs peuvent diminuer afin de modéliser des points anguleux. Cette catégorie de contours actifs est néanmoins limitée :

- La fonctionnelle exprimée en équation (4.18) dépend du paramétrage p du contour actif, ce qui la rend non intrinsèque.
- Le résultat de la segmentation est sensible vis-à-vis des paramètres α , β , et γ qu'il faut régler et ajuster en fonction de l'image à analyser. Il n'existe pas cadre théorique pour fixer leurs valeurs, mais seulement des choix empiriques.
- De façon inhérente au mode de représentation, la topologie du contour ne peut évoluer. De plus celui-ci ne peut segmenter que des formes convexes.
- Comme nous l'avons déjà évoqué, ce type de contour est instable numériquement à cause de sa représentation explicite.
- Enfin, l'initialisation du contour doit être très proche de l'objet à segmenter dans l'image, ce qui de manière générale présente peu d'intérêt.

- Les contours géodésiques

Dans [18], Caselles *et al.* montrent que minimiser la fonctionnelle de l'équation (4.18) avec $\beta = 0$ revient à trouver une courbe géodésique dans un espace de Riemann dont la métrique dépend de l'image analysée. La fonctionnelle proposée est alors :

$$J_b(\mathbf{C}(p, t)) = \alpha \int_a^b \left| \frac{\partial \mathbf{C}(p, t)}{\partial p} \right|^2 dp + \gamma \int_a^b g(|\nabla I(\mathbf{C}(p, t))|) dp \quad (4.22)$$

La fonctionnelle exprimée en équation (4.22) a les mêmes inconvénients que celle proposée en équation (4.18) : elle est non intrinsèque et dépend de paramètres arbitraires α et γ . Grâce au principe de Maupertuis [18], Caselles *et al.* montrent que minimiser la fonctionnelle de l'équation (4.22) revient à minimiser la longueur du contour actif L_R (d'où la dénomination *géodésique*) selon une métrique propre à l'image :

$$L_R = \int_0^{L(\mathbf{C}(t))} g(|\nabla I(\mathbf{C}(s, t))|) ds \quad (4.23)$$

Comparée à la longueur Euclidienne de $\mathbf{C}(t)$ qui est par définition $L_E = \oint ds$, on constate que la nouvelle définition de la longueur comprend une pondération en $g(|\nabla I(\mathbf{C}(s, t))|)$ qui est un terme attaché aux données. Ainsi, plus le contour actif traversera des zones de gradient élevé de l'image, plus sa longueur diminuera. Le problème de segmentation revient ainsi à trouver un contour géodésique dont la métrique dépend de l'image. En minimisant la fonctionnelle désormais intrinsèque de l'équation (4.23) par calcul des variations, les auteurs établissent l'équation d'évolution du contour

par descente de gradient. En représentant le contour par un ensemble de niveaux ils obtiennent :

$$\phi_t(\mathbf{x}, t) = g(|\nabla I(\mathbf{C}(p, t))|) \kappa |\nabla \phi(\mathbf{x}, t)| - \langle \nabla g(|\nabla I(\mathbf{C}(p, t))|), \nabla \phi(\mathbf{x}, t) \rangle \quad (4.24)$$

Le premier terme de l'équation (4.24) est relatif à la courbure du contour actif et a un effet régularisateur. Ce terme apparaît naturellement et justifie la suppression du terme contrôlant la rigidité du contour dans la fonctionnelle donnée en équation (4.18). Ce terme garantissant la régularité de la courbe est pondéré par la fonction g , arrêtant ainsi la progression du contour sur les zones de haut gradient. Le second terme est la véritable nouveauté des contours géodésiques. En effet, les termes basés sur la courbure ont déjà été introduits auparavant avec les approches géométriques. En revanche, le second terme de l'équation (4.24) permet de stabiliser le contour actif sur des bords caractérisés par une valeur variable du gradient. Habituellement, une progression du contour dont la vitesse est pondérée par g est efficace si la valeur du gradient est suffisamment élevée pour faire tendre g vers zéro. Dans la pratique la valeur du gradient n'est pas nécessairement très élevée, engendrant un phénomène de "fuite" du contour actif au travers des zones où $|\nabla I|$ n'est pas significatif. Ce type de phénomène apparaît aussi lorsque la valeur du gradient n'est pas constante le long du bord d'un objet à segmenter. Dans ce dernier cas, le deuxième terme de l'équation permet de palier ce problème en exerçant une force de rappel, stabilisant le contour, et évitant ainsi les écueils des approches purement géométriques.

Les contours géodésiques surpassent les "snakes" sur plusieurs plans. Premièrement, la fonctionnelle des contours géodésiques est intrinsèque. Ces derniers prennent en compte le cas d'une valeur variable du gradient le long d'un bord d'objet à extraire. De plus, ils permettent de segmenter des objets non convexes, et sont facilement transposables en ensembles de niveaux. Finalement, on peut saluer et souligner la contribution des contours géodésiques à unifier les approches variationnelles et géométriques.

Les inconvénients des contours actifs basés sur l'information de frontière

Bien qu'ils soient abondamment utilisés pour la segmentation d'image, les contours actifs basés frontière revêtent de nombreux problèmes limitant leur domaine d'application à des images relativement simples. Le principal inconvénient de cette approche est le caractère purement *local* de l'information utilisée pour mouvoir le contour actif. Cela a pour effet de devoir initialiser le contour actif très proche de l'objet que l'on désire segmenter. Le descripteur couramment choisi pour ce genre de contours actifs est le module du gradient de l'image. Ainsi, si tout ou partie du contour actif se retrouve dans des zones de luminance homogène, dépourvues de gradient élevé, le contour ne sera mû par aucune force externe. Enfin, l'autre problème du caractère local de l'information de frontière est la sensibilité au bruit de l'image. Les descripteurs de frontière étant calculés localement, les voisinages considérés sont petits devant la taille de l'image et sont ainsi largement influencés par le bruit. Pré-filtrer l'image est toujours une solution pour réduire ce problème, cependant, le filtrage a pour effet de délocaliser les bords de l'objet d'intérêt, rendant la segmentation imprécise.

Pour palier le problème de la sensibilité à l'initialisation des contours actifs basés sur l'information de frontière, Cohen [27] propose d'ajouter une force artificielle de gonflage ou rétraction visant à mener le contour actif sur les bords de l'objet à segmenter. Ceci est équivalent à ajouter une constante c dans l'équation d'évolution des contours géométriques ou géodésiques de la façon suivante :

$$\phi_t(\mathbf{x}, t) = [\kappa + c] g(|\nabla I(\mathbf{C}(p, t))|) |\nabla \phi(\mathbf{x}, t)| - \langle \nabla g(|\nabla I(\mathbf{C}(p, t))|), \nabla \phi(\mathbf{x}, t) \rangle \quad (4.25)$$

En fonction du signe de c , le contour actif s'étend ou se rétracte selon sa normale. Le choix du

signe se révèle ainsi critique car il nécessite la connaissance *a priori* de la localisation de l'objet à extraire par rapport à l'état initial du contour $\mathbf{C}(t = t_0)$. Le cas où le contour actif est partiellement à l'intérieur et l'extérieur de l'objet à segmenter est aussi problématique.

- *Diffusion de l'information de gradient par Gradient Vector Flow (GVF)*

Afin de rendre l'information de gradient de l'image non locale, Xu et Prince [128] proposent de diffuser le gradient d'une image sous la forme d'un champ de vecteurs de composantes $(\mathbf{u}, \mathbf{v})$. Soit f le module normalisé du gradient d'une image I . La diffusion de l'information de gradient se fait par la minimisation de la fonctionnelle :

$$J_{GVF}(u, v) = \int_{\Omega} \mu_{GVF} (\mathbf{u}_x^2 + \mathbf{u}_y^2 + \mathbf{v}_x^2 + \mathbf{v}_y^2) + |\nabla f|^2 |(\mathbf{u}, \mathbf{v}) - \nabla f|^2 d\mathbf{x} \quad (4.26)$$

où μ_{GVF} est une constante positive.

Le premier terme de l'équation (4.26) est relatif à l'allure du champ $(\mathbf{u}, \mathbf{v})$. Pour les lieux de gradient faible de l'image, ce terme est prédominant et régularise $(\mathbf{u}, \mathbf{v})$. A contrario, pour les gradient élevés, le second terme d'attache aux données impose à $(\mathbf{u}, \mathbf{v})$ à ressembler au gradient de l'image ∇f . Dans [128], les auteurs proposent de déduire le GVF $(\mathbf{u}, \mathbf{v})$ à partir d'une descente de gradient minimisant J_{GVF} :

$$\begin{cases} \mathbf{u}_t = \mu_{GVF} \Delta \mathbf{u} - (\mathbf{u} - f_x) |\nabla f|^2 \\ \mathbf{v}_t = \mu_{GVF} \Delta \mathbf{v} - (\mathbf{v} - f_y) |\nabla f|^2 \end{cases} \quad (4.27)$$

Le GVF ainsi obtenu peut être soit incorporé avec des contours actifs représentés explicitement [128] ou implicitement [92]. Dans [92], N. Paragios utilise ce champ de vecteurs diffusé avec des contours actifs représentés par des ensembles de niveaux. Voici l'un des modèles présenté dans [92] :

$$\phi_t(\mathbf{x}, t) = g(|\nabla I(\mathbf{C}(p, t))|) (\kappa |\nabla \phi(\mathbf{x}, t)| - \langle (\mathbf{u}, \mathbf{v}), \nabla \phi(\mathbf{x}, t) \rangle) \quad (4.28)$$

Le premier terme de l'équation (4.28) régularise la courbe alors que le second aligne la normale du contour avec le GVF et tend à diriger $\mathbf{C}$ vers les lieux de gradient élevé. Cette technique de *gradient vector flow* permet de rendre moins locale l'information de frontière tout en autorisant le contour actif à segmenter des objets non convexes.

4.2.1.4 Contours actifs basés sur l'information de région

Une alternative efficace aux contours actifs basés sur l'information de frontière est l'utilisation des contours actifs basés sur l'information de région, qui intrinsèquement font usage d'une information *globale* de l'image. Les descripteurs d'une région R sont souvent des quantités statistiques telles que la moyenne, la variance, la texture ou l'histogramme de cette même région. La fonctionnelle d'énergie construite à partir de tels descripteurs $k(\mathbf{x}, R)$ est alors une intégrale double sur la région $R(t)$ délimitée par $\mathbf{C}(t)$. Optionnellement, il est possible d'ajouter une composante basée frontière de descripteur $k_b(\mathbf{x})$ soit pour introduire un terme de régularisation, soit pour insérer un terme d'attache aux données complémentaire de ceux apportés par l'information de région :

$$J_r(R(t)) = \int_{R(t)} k(\mathbf{x}, R(t)) d\mathbf{x} + \int_{\mathbf{C}(t)} k_b(\mathbf{x}) d\mathbf{a} \quad (4.29)$$

Les contours actifs basés sur l'information de région font partie d'un domaine de recherche très actif depuis les années 90 en Vision par Ordinateur. Originellement, les travaux de Zhu *et al.* [137] pré-

sentent une méthode de compétition de régions dans un cadre mixte à la fois Bayésien et minimisant le critère de *description de longueur minimale* (MDL). Les auteurs considèrent les pixels de l'image comme une réalisation d'une variable aléatoire dont la densité de probabilité Gaussienne a des paramètres différents selon l'appartenance ou non à la région délimitée par le contour actif. C'est sur la base de ces travaux et de ceux de [18] que N. Paragios introduira (dans une approche purement Bayésienne) les *régions actives géodésiques* qui intègrent le terme basé contour des contours géodésiques pour la segmentation supervisée d'images scalaires, et le suivi d'objets en mouvement dans des séquences vidéos [88, 91]. Les travaux de M. Rousson [104, 103] prolongent ceux de N. Paragios en les étendant aux images en couleurs et à la segmentation de zones texturées ne répondant plus aux critères Gaussiens.

Parallèlement, les travaux de D. Mumford et J. Shah réalisés en 1985 ont permis d'ouvrir une autre branche, plus géométrique, dans le domaine de la segmentation basée sur les régions [75]. Ces travaux ont récemment été rendus populaires dans la communauté de Vision par Ordinateur grâce à T. Chan et L. Vese [19]. Rousson montre dans [103] que la limite de partition minimale des travaux de Mumford et Shah est un cas particulier de l'approche Bayésienne. Les travaux récents sur la segmentation basée région foisonnent [19, 20, 21, 25, 40, 75, 89, 88, 93, 91, 104, 103, 117, 137]. Ils traitent la segmentation ou la classification supervisée/non supervisée de N partitions de l'image, l'utilisation d'information de texture, et l'analyse de séquences d'images. Dans le cadre de notre étude, nous nous restreignons à la segmentation bi-modale d'une image scalaire. Si le nombre de partitions est connu à l'avance ($N = 2$), leurs propriétés statistiques dans l'image ne le sont pas (à l'inverse des approches supervisées). Nous présentons les deux modèles principaux de la segmentation basée région qui ont été implantés pour notre application : le modèle de Mumford-Shah (et ses dérivés) ainsi que le celui basé sur une approche Bayésienne.

Fonctionnelle de Mumford-Shah

En 1985, D. Mumford et J. Shah proposent dans [75] une méthode de segmentation, applicable à n'importe quel type de signal. Dans le cadre du traitement des images, l'idée est la suivante : pour une image I donnée, on cherche une image idéale U homogène par morceaux qui est une approximation de I . L'image U est une collection de régions homogènes dont les intersections forment un ensemble B composé de frontières régulières. Les auteurs ont formalisé cette idée sous la forme de la fonctionnelle suivante :

$$J_{MS} = \mu^2 \int \int_R (U - I)^2 + \mu \int_{R-B} |\nabla U|^2 + \nu L_B \quad (4.30)$$

où : $(\mu, \nu) \in \mathbb{R}^+$ et L_B est la longueur des frontières composant l'ensemble B . Le premier terme de cette fonctionnelle impose à U de ressembler à l'image I . Le second force U à être homogène à l'intérieur de chaque région (les discontinuités inter-régions sont préservées). Enfin, le dernier terme impose aux frontières partitionnant l'image U d'avoir une longueur L_B minimale. La minimisation de cette fonctionnelle revient donc à trouver l'ensemble des frontières de B segmentant l'image I en parties homogènes. Cette technique de segmentation permet ainsi de débruiter une image dans le même temps.

Plus récemment, Chan et Vese ont appliqué cette technique de segmentation au cas limite de deux régions, c'est la limite de partition minimale [19]. Les auteurs réduisent U à la moyenne de I à l'intérieur et à l'extérieur de la région R . Cette simplification appelée *cartoon limit* en anglais, ou plus poétiquement *univers de Mondrian* [40] restreint U à être constante (et non homogène) par

morceaux. L'expression de la fonctionnelle sous forme d'ensemble de niveaux revient alors à :

$$J_{CV}(\phi, t) = \int_{\Omega} (I(\mathbf{x}) - c_{in}(\phi(\mathbf{x}, t)))^2 H_a(\phi(\mathbf{x}, t)) d\mathbf{x} \\ + \int_{\Omega} (I(\mathbf{x}) - c_{out}(\phi(\mathbf{x}, t)))^2 (1 - H_a(\phi(\mathbf{x}, t))) d\mathbf{x} + \nu \int_{\Omega} |\nabla H_a(\phi(\mathbf{x}, t))| d\mathbf{x} \quad (4.31)$$

Les quantités c_{in} et c_{out} étant par définition la moyenne des pixels de l'image I à l'intérieur et l'extérieur du contour actif au temps t :

$$c_{in}(\phi(\mathbf{x}, t)) = \frac{\int_{\Omega} I(\mathbf{x}) H_a(\phi(\mathbf{x}, t)) d\mathbf{x}}{\int_{\Omega} H_a(\phi(\mathbf{x}, t)) d\mathbf{x}} \quad (4.32)$$

$$c_{out}(\phi(\mathbf{x}, t)) = \frac{\int_{\Omega} I(\mathbf{x}) (1 - H_a(\phi(\mathbf{x}, t))) d\mathbf{x}}{\int_{\Omega} (1 - H_a(\phi(\mathbf{x}, t))) d\mathbf{x}} \quad (4.33)$$

où H_a est une approximation régulière de la fonction de Heaviside. La minimisation de la fonctionnelle de l'équation (4.31) par la technique de descente de gradient produit l'équation d'évolution suivante :

$$\phi(\mathbf{x}, t)_t = - \left\{ (I(\mathbf{x}) - c_{in}(\phi(\mathbf{x}, t)))^2 - (I(\mathbf{x}) - c_{out}(\phi(\mathbf{x}, t)))^2 + \kappa\nu \right\} \delta_a(\phi(\mathbf{x}, t)) \quad (4.34)$$

où δ_a est une approximation régulière de la distribution de Dirac. Cette approche permet une segmentation moins sensible à l'initialisation du contour actif comparé à ceux basés frontière. C'est aussi une méthode facile à implanter et relativement peu coûteuse en temps de calcul. Cependant, ce modèle restreint la segmentation aux objets constants par morceaux. Il est néanmoins possible de s'éloigner de cette hypothèse si l'objet d'intérêt se démarque *sensiblement* du fond de l'image. C'est le cas de certains résultats de [19] illustrant la segmentation de galaxies (constellation d'étoiles) qui ne sont pas homogènes mais qui se distinguent singulièrement de l'espace inter-sidéral. Chan et Vese ont par la suite introduit dans [20] une simplification moins drastique de l'image U en l'autorisant à être homogène par morceaux. Les valeurs de U dépendent désormais de la position pixellaire $\mathbf{x}$ et de l'appartenance à l'intérieur ou l'extérieur de la région délimitée par le contour actif :

$$U(\mathbf{x}) = \begin{cases} U_{in}(\mathbf{x}), & \text{si } \mathbf{x} \in \Omega_{in}(t) \\ U_{out}(\mathbf{x}), & \text{si } \mathbf{x} \in \Omega_{out}(t) \end{cases}$$

La fonctionnelle de Mumford-Shah devient alors :

$$J_{CV}(\phi, t) = \int_{\Omega} (I(\mathbf{x}) - U_{in}(\mathbf{x}))^2 H_a(\phi(\mathbf{x}, t)) d\mathbf{x} \\ + \int_{\Omega} (I(\mathbf{x}) - U_{out}(\mathbf{x}))^2 (1 - H_a(\phi(\mathbf{x}, t))) d\mathbf{x} \\ + \mu \int_{\Omega} |\nabla U_{in}(\mathbf{x})| H_a(\phi(\mathbf{x}, t)) d\mathbf{x} \\ + \mu \int_{\Omega} |\nabla U_{out}(\mathbf{x})| (1 - H_a(\phi(\mathbf{x}, t))) d\mathbf{x} \\ + \nu \int_{\Omega} |\nabla H_a(\phi(\mathbf{x}, t))| d\mathbf{x} \quad (4.35)$$

En dérivant l'équation (4.35) respectivement par rapport à $U_{in}(\mathbf{x})$, $U_{out}(\mathbf{x})$ et $\phi(\mathbf{x})$, les auteurs

déduisent les équations d'évolution de l'image idéale U et du contour actif implicitement représenté par ϕ . Dans une approche différente, Tsai *et al* estiment $U_{in}(\mathbf{x})$ et $U_{out}(\mathbf{x})$ grâce à une théorie stochastique de l'optimisation [117].

Approche Bayésienne

Nous rappelons la formulation du *maximum a posteriori* (MAP) exprimée dans [88, 91, 103] et ses liens avec des travaux antérieurs aussi basés sur une approche Bayésienne.

Soit $p(\mathcal{P}(\Omega)|I)$ la probabilité d'obtenir une partition $\mathcal{P}(\Omega)$ de l'image pour une image I donnée. Une segmentation optimale de l'image est réalisée lorsque cette probabilité est maximisée. Selon la formule de Bayes, cette probabilité s'exprime sous la forme :

$$p(\mathcal{P}(\Omega)|I) = \frac{p(I|\mathcal{P}(\Omega))}{p(I)} p(\mathcal{P}(\Omega)) \quad (4.36)$$

avec $p(I)$: probabilité d'observer une réalisation I de l'image. $p(\mathcal{P}(\Omega))$ est la probabilité d'obtenir une partition $\mathcal{P}(\Omega)$ de l'image parmi toutes les partitions possibles. $p(I|\mathcal{P}(\Omega))$ est la probabilité d'obtenir une image I pour une partition $\mathcal{P}(\Omega)$ connue *a priori*. Pour le problème de segmentation à partir d'une seule image, la probabilité d'observation d'une réalisation I de l'image est constante. Ainsi, l'équation (4.36) revient à :

$$p(\mathcal{P}(\Omega)|I) \sim p(I|\mathcal{P}(\Omega)) p(\mathcal{P}(\Omega)) \quad (4.37)$$

Dans [105, 103], Rousson injecte dans $p(\mathcal{P}(\Omega))$ une connaissance *a priori* sur la forme des partitions optimales à obtenir. D'une autre façon, ce terme permet de régulariser le contour actif s'il dépend de la longueur $L(\mathbf{C})$ de ce dernier :

$$p(\mathcal{P}(\Omega)) \sim \exp^{-\nu L(\mathbf{C})} \quad (4.38)$$

avec ν une constante appartenant à $\mathbb{R}^+$. Maximiser la probabilité ci-dessus revient bien à trouver une partition de longueur minimale. Ce terme est équivalent à celui introduit dans la fonctionnelle de Mumford-Shah en équation (4.30). Dans [88, 91], N. Paragios considère $p(\mathcal{P}(\Omega))$ constante. Il ajoute à la fonctionnelle basée région dérivée du MAP le terme basé frontière des contours géodésiques qui comporte un terme diffuseur et régularisateur.

Le calcul de $p(I|\mathcal{P}(\Omega))$ requiert quelques hypothèses pour être mené à bien :

- Les régions de la partition optimale ne sont pas corrélées :

$$p(I|\mathcal{P}(\Omega)) = p(I|\mathcal{P}(\Omega_{in}, \Omega_{out})) = p(I|\mathcal{P}(\Omega_{in})) p(I|\mathcal{P}(\Omega_{out})) \quad (4.39)$$

Cette hypothèse est raisonnable puisque le but de la segmentation est de séparer des régions de l'image dont les propriétés sont différentes.

- En revanche, l'hypothèse suivante est plus forte et restrictive : *les pixels d'une même région sont indépendants et ont la même probabilité d'être observés*. Cette hypothèse est mise en défaut dans des zones texturées, ou à motifs périodiques où il existe une interaction locale entre les pixels. L'équation (4.37) devient alors :

$$p(I|\mathcal{P}(\Omega)) = \prod_{\mathbf{x} \in \Omega_{in}} p(I(\mathbf{x})|\mathcal{P}(\Omega_{in})) \prod_{\mathbf{x} \in \Omega_{out}} p(I(\mathbf{x})|\mathcal{P}(\Omega_{out})) \quad (4.40)$$

La maximisation peut être reformulée sous la forme de la minimisation de l'énergie suivante :

$$J(\Omega_{in}, \Omega_{out}) = -\ln \{p(\mathcal{P}(\Omega) | I)\} \quad (4.41)$$

En remplaçant les expressions des équations (4.38) et (4.40) dans la formulation de l'équation (4.41), on a finalement la fonctionnelle d'énergie :

$$J(\Omega_{in}, \Omega_{out}) \sim -\int_{\Omega_{in}} \ln(p(I(\mathbf{x}) | \mathcal{P}(\Omega_{in}))) - \int_{\Omega_{out}} \ln(p(I(\mathbf{x}) | \mathcal{P}(\Omega_{out}))) + \nu L(\mathbf{C}) \quad (4.42)$$

Dans [88, 91], [103] et [137], les auteurs supposent que les densités de probabilité suivent des lois normales dont les seuls paramètres sont la moyenne c et la variance σ calculées sur Ω :

$$p(I(\mathbf{x}) | c, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp^{-\frac{(I(\mathbf{x})-c)^2}{2\sigma^2}} \quad (4.43)$$

Dans [88, 91], les paramètres des distributions Gaussiennes sont connus à l'avance par un processus supervisé. Dans [103] et [137], ces paramètres sont calculés dynamiquement au cours des itérations du contour actif. Dans [137], l'expression de la fonctionnelle ne provient pas du MAP mais de la technique dite *MDL* (Minimum Description Length).

Avec l'hypothèse d'une distribution Gaussienne, et en représentant le contour actif par un ensemble de niveaux, la fonctionnelle d'énergie devient :

$$\begin{aligned} J_{Bayes}(\phi, t) = & \int_{\Omega} \left(\frac{(I(\mathbf{x}) - c_{in}(\phi(\mathbf{x}, t)))^2}{2\sigma_{in}^2} + \ln \left(\sqrt{2\pi\sigma_{in}^2} \right) \right) H_a(\phi(\mathbf{x}, t)) d\mathbf{x} \\ & + \int_{\Omega} \left(\frac{(I(\mathbf{x}) - c_{out}(\phi(\mathbf{x}, t)))^2}{2\sigma_{out}^2} + \ln \left(\sqrt{2\pi\sigma_{out}^2} \right) \right) (1 - H_a(\phi(\mathbf{x}, t))) d\mathbf{x} + \nu L(\mathbf{C}) \end{aligned} \quad (4.44)$$

où les variances à l'intérieur et l'extérieur du contour actif sont par définition :

$$\sigma_{in}^2(\phi(\mathbf{x}, t)) = \frac{\int_{\Omega} (I(\mathbf{x}) - c_{in}(\phi(\mathbf{x}, t)))^2 H_a(\phi(\mathbf{x}, t)) d\mathbf{x}}{\int_{\Omega} H_a(\phi(\mathbf{x}, t)) d\mathbf{x}} \quad (4.45)$$

$$\sigma_{out}^2(\phi(\mathbf{x}, t)) = \frac{\int_{\Omega} (I(\mathbf{x}) - c_{out}(\phi(\mathbf{x}, t)))^2 (1 - H_a(\phi(\mathbf{x}, t))) d\mathbf{x}}{\int_{\Omega} (1 - H_a(\phi(\mathbf{x}, t))) d\mathbf{x}} \quad (4.46)$$

L'équation d'évolution du contour actif dérivée de la fonctionnelle (4.42) par descente de gradient est alors :

$$\phi(\mathbf{x}, t)_t = \left\{ -\frac{(I(\mathbf{x}) - c_{in}(\phi(\mathbf{x}, t)))^2}{2\sigma_{in}^2(t)} + \frac{(I(\mathbf{x}) - c_{out}(\phi(\mathbf{x}, t)))^2}{2\sigma_{out}^2(t)} + \ln \left(\frac{\sigma_{out}^2(t)}{\sigma_{in}^2(t)} \right) + \kappa\nu \right\} \delta_a(\phi(\mathbf{x}, t)) \quad (4.47)$$

Les paramètres des distributions Gaussiennes sont évalués à chaque itération par la technique de maximum de vraisemblance qui revient à la simple estimation de la moyenne et de la variance dans ce cas. Dans [137], Zhu *et al* aboutissent à une fonctionnelle similaire à celle de l'équation (4.47) sans le terme de régularisation. Leur démarche basée sur le MDL occulte cependant les hypothèses clairement énoncées dans l'approche Bayésienne pour aboutir à ce résultat. On remarque que pour $\sigma_{in}^2(t) = \sigma_{out}^2(t) = cst$, on retrouve le modèle de Mumford-Shah exprimé en équation (4.34) par Chan et Vese.

4.2.2 Segmentation sous contrainte de forme spécifique

Qu'ils soient basés frontière ou région, les modèles présentés dans la section précédente sont fonctions de l'intensité des pixels de l'image. Ainsi, si l'objet à segmenter est altéré par du bruit, des occlusions ou un faible contraste au niveau de ses bords, le résultat sera largement influencé par ces artefacts qui dégraderont la qualité de l'extraction. Ceci est particulièrement critique pour le cas de certaines images médicales (ultra-sonores par exemple) qui ont un faible rapport signal à bruit et présentent des organes dont certaines parties du contour sont très ténues ou occultées. De ce fait, c'est la communauté de Vision par Ordinateur en Imagerie Médicale qui a été la première et la plus active à résoudre ce problème par l'insertion de connaissance de forme *a priori* dans le processus de segmentation.

La connaissance *a priori* relative à un objet à segmenter dans une image est multiple. Elle concerne à la fois la radiométrie de l'objet (couleur, luminance), ses propriétés de texture ou sa forme. Dans ce chapitre, nous nous intéressons uniquement à la connaissance *a priori* de forme. Celle-ci permet de contraindre le contour actif à ressembler à une forme (ou une famille de formes) de référence qui régularisera spécifiquement le contour sur les lieux de l'image où des artefacts sont présents.

Depuis le début des années 90, de nombreux travaux ont montré combien la donnée d'une forme spécifique de référence pouvait améliorer les résultats de segmentation [33, 29, 31, 32, 30, 41, 40, 66, 21, 24, 23, 25, 22, 93, 105, 103, 114, 131, 130]. Nous proposons une brève revue des travaux les plus marquants sur l'insertion de forme *a priori* en distinguant les approches statistiques (qui utilisent un jeu de formes de référence) de celles déterministes. De façon générale, la contrainte de forme est introduite par une métrique permettant de comparer le contour actif au temps t avec la forme *a priori*. Dans le cadre des approches variationnelles, cette métrique est utilisée pour la formulation d'une énergie de contrainte de forme J_{shape} qui est alors ajoutée à celle relative à l'attache aux données :

$$J = J_{image} + \lambda J_{shape} \quad (4.48)$$

où J est la fonctionnelle d'énergie globale, J_{image} est l'énergie d'attache aux données et $\lambda \in \mathbb{R}^+$. Nous proposons de décrire ces distances en fonction de la façon dont est représenté le contour actif (explicite ou implicite). Cette dichotomie n'est pas parfaite puisque certaines approches peuvent être formulées dans l'un ou l'autre des systèmes de représentation. Nous le préciserons le cas échéant.

4.2.2.1 Approche non statistique

Représentation explicite

Dans [41] A. Foulonneau propose d'utiliser les moments géométriques afin de caractériser la forme du contour actif et celle de la forme de référence. Une distance portant sur ces moments est ensuite créée pour mesurer l'écart du contour actif à la forme *a priori*. Afin d'assurer une représentation plus compacte et moins redondante, les moments géométriques sont projetés sur une base orthogonale de polynômes de Legendre. Les paramètres de cette représentation explicite ont l'avantage d'être intrinsèquement invariant par translation et transformation par facteur d'échelle. Le détail de leur extension à l'invariance par rotation, et plus généralement par transformation affine est disponible dans [40]. Comme toute représentation fondée sur l'utilisation d'une base (de Legendre, ou Fourier), la troncature de l'ordre de cette dernière est une décision délicate. En effet, pour représenter des formes comportant des singularités (coins), il faut utiliser un ordre élevé pour assurer une approximation de qualité. Ceci est un problème qui augmente la complexité calculatoire

et restreint le pouvoir de représentation de ces méthodes explicites.

Représentation implicite

Y. Chen est la première à introduire un terme de contrainte de forme grâce aux ensembles de niveaux [24, 23]. Cependant, il convient de remarquer que cette formulation peut aussi être représentée explicitement. L'auteur propose un terme énergétique de contrainte de forme invariant par similitude plane directe T_{sim} qui dans le cadre d'une approche variationnelle prend la forme suivante :

$$J_{shape}(\mathbf{C}, T_{sim}) = \int_0^1 d^2(T_{sim} \mathbf{C}(p)) |\mathbf{C}_p(p)| dp \quad (4.49)$$

où $d(\mathbf{x}) \equiv d(\mathbf{C}^*, \mathbf{x})$ est la distance entre le point de coordonnées $\mathbf{x}$ et la forme de référence $\mathbf{C}^*$. Ce terme est pondéré par un poids constant λ et ajouté au terme d'énergie d'attache aux données des contours géodésiques. Les résultats de cette approche montrent une amélioration notable de la segmentation d'images médicales. Cependant, elle reste très sensible à l'initialisation à cause du terme basé frontière de l'attache aux données. L'invariance par similitude nécessite l'optimisation des paramètres de T_{sim} au cours de l'évolution du contour actif. L'auteur estime leurs valeurs par descente de gradient. Le réglage du poids λ n'est pas une question triviale. Un poids trop faible risque une mauvaise segmentation de l'objet alors qu'un poids trop élevé empêchera le contour actif d'être fidèle à l'information dérivée de l'image.

Dans [32], D. Cremers propose le terme d'énergie de contrainte suivant :

$$J_{shape}(\phi, \psi_0) = \int_{\Omega} (\phi(\mathbf{x}) - \psi_0(\mathbf{x}))^2 d\mathbf{x} \quad (4.50)$$

ce terme quadratique est la différence entre l'ensemble de niveaux ϕ représentant le contour actif et celui encodant la forme de référence ψ_0 . Cette formulation a la qualité de préserver la caractéristique intrinsèque de topologie flexible conférée par les ensembles de niveaux. En revanche, cette énergie dépend de la taille du domaine d'intégration et n'est invariante par aucune transformation. D. Cremers remarque qu'une telle distance ne permet que d'extraire l'objet connu *a priori* dans l'image, opérant tel un filtrage dans l'espace des formes du contour actif. Il propose alors une nouvelle énergie capable d'appliquer localement la contrainte de forme dans l'image :

$$J_{shape}(\phi, \psi_0, L) = \int_{\Omega} (\phi(\mathbf{x}) - \psi_0(\mathbf{x}))^2 (L(\mathbf{x}) + 1)^2 d\mathbf{x} + \int_{\Omega} \lambda^2 (L(\mathbf{x}) - 1)^2 d\mathbf{x} + \gamma \int_{\Omega} |\nabla H(L(\mathbf{x}))| d\mathbf{x} \quad (4.51)$$

La fonction L dite de label dynamique prend uniquement les valeurs $+1$ ou -1 . La fonction L renforce la contrainte de forme lorsqu'elle tend vers 1 , et l'annule en convergeant vers -1 . Ce dernier comportement est favorisé lorsque le contour actif est éloigné de la forme *a priori*, statique dans le cas présent. Le dernier terme de l'équation (4.51) assure la régularité de la fonction L . Cette énergie permet à la fois de segmenter un objet corrompu grâce à l'*a priori* ainsi que les objets environnants de même radiométrie sans connaître à l'avance l'endroit où appliquer la contrainte.

Dans [105] les auteurs proposent l'énergie de contrainte suivante :

$$J_{shape}(\phi, \psi_0) = \int_{\Omega} (\phi(\mathbf{x}) - \psi_0(T_{sim}\mathbf{x}))^2 H(\phi(\mathbf{x})) d\mathbf{x} \quad (4.52)$$

Cette formulation a l'avantage de ne plus dépendre de l'espace d'intégration Ω puisque l'intégrale porte désormais sur l'intérieur du contour actif. Elle est aussi invariante par similitude plane directe

comme en équation (4.49). Mais comme le remarque D. Cremers dans [31], cette distance quadratique n'est pas symétrique et est incapable de prendre en compte des objets à plusieurs composantes. Ce dernier propose alors une pseudo-distance vérifiant les propriétés de positivité et de symétrie sans pour autant satisfaire l'inégalité triangulaire :

$$J_{shape}(\phi, \psi_0) = \int_{\Omega} (\phi(\mathbf{x}) - \psi_0(\mathbf{x}))^2 \frac{h(\phi(\mathbf{x})) + h(\psi_0(T_{sim}\mathbf{x}))}{2} d\mathbf{x} \quad (4.53)$$

où : $h(\phi(\mathbf{x})) = \frac{H(\phi(\mathbf{x}))}{\int_{\Omega} H(\phi(\mathbf{x})) d\mathbf{x}}$. Cette énergie permet ainsi de segmenter avec succès un objet multi-composantes partiellement effacé et bruité.

Dans [21], une autre énergie de contrainte symétrique et indépendante du domaine d'intégration est formulée :

$$J_{shape}(\phi, \psi_0) = \int_{\Omega} (H(\phi(\mathbf{x})) - H(\psi_0(T_{sim}\mathbf{x})))^2 d\mathbf{x} \quad (4.54)$$

Ce terme compare les aires intérieures au contour actif et à la forme de référence au sens de la norme L_2 .

4.2.2.2 Approche statistique

L'objet à segmenter dans l'image peut présenter une certaine variabilité par rapport à la forme de référence. Ceci est par exemple le cas des images médicales où la forme d'un organe varie d'un patient à l'autre, selon le mode d'acquisition, l'orientation de la caméra ou le temps dans le cas de séquences vidéos (échographie cardiaque). Puisque ces variations par rapport à la référence ne sont pas modélisables par une transformation linéaire de type affine, de nombreux travaux ont focalisé leurs efforts sur l'apprentissage de formes diverses afin de conférer plus de souplesse à la contrainte de forme. Ainsi, le contour actif est autorisé à se déformer dans un sous-espace des formes défini par l'apprentissage. Puisque les formes apprises sont redondantes, elles sont souvent projetées dans un sous-espace orthogonal grâce à l'analyse en composantes principales (ACP) afin de déterminer les modes principaux de variabilité (i.e. de déformation). Un souci inhérent au traitement de formes d'apprentissage est leur alignement. En effet, il convient de les placer dans un même système de référence afin de pouvoir les comparer sans biais. L'approche communément adoptée est d'estimer la meilleure similitude plane directe permettant un alignement optimal (méthode Procrustes). Lorsque certaines formes sont trop différentes, elles sont classées en sous-familles par *clustering* [24].

Représentation explicite

Les travaux pionniers sur l'insertion de contrainte statistique de forme est à attribuer à Staib et Duncan [114]. Les auteurs proposent de représenter explicitement le contour actif sur une base de Fourier. La donnée d'échantillons d'apprentissage leur permet d'évaluer une densité de probabilité sur les paramètres de la représentation (qui sont les coefficients de chacun des éléments de la base). En supposant que la densité de probabilité de forme *a priori* est Gaussienne, ils formulent leur critère de segmentation par MAP. La segmentation optimale est ainsi obtenue lorsque le contour actif satisfait l'attache aux données avec un maximum de vraisemblance avec les formes de référence de l'apprentissage.

Une approche similaire est reprise par Cootes *et al* dans [28]. La différence provient du mode de représentation par *snakes*. Ces derniers prélèvent des "points de contrôle" à des endroits choisis sur les formes d'apprentissage. Ce processus est manuel ou assisté et représente une contrainte forte : i) le choix des points de contrôle peut avoir une influence sur le résultat final. ii) c'est une étape

lente qui restreint le flux de données à traiter. Nous verrons que ce problème a été résolu dans [33, 105]. Après avoir aligné les formes d'apprentissage, ces dernières sont projetées dans une sous-base orthogonale par ACP. Il n'existe pas à proprement parler de critère énergétique permettant de comparer le contour actif avec la référence. A chaque itération, la variation spatiale de chacun des nœuds attirés par les zones de haut gradient est projetée dans l'espace des formes, ce qui contraint le déplacement, et *in fine* la forme du contour actif. Ces travaux sont traduits dans un cadre Bayésien et généralisés à l'invariance par transformation affine dans [131].

Dans [33], D. Cremers introduit les *Diffusion Snakes* basés sur une représentation par B-splines. L'avantage de la base de B-splines est une meilleure compacité de la représentation des formes d'apprentissage ainsi qu'un caractère local (une variation d'un point de contrôle modifie localement la courbe contrairement à une base de Fourier). De plus, le pouvoir de généralisation est étendu aux formes singulières comportant des coins. Une analyse statistique par ACP permet de déterminer les variations principales des points de contrôle. La contrainte de forme est formulée comme une distance de Mahalanobis entre le contour actif et la forme moyenne de référence. Ce terme à l'avantage d'être intrinsèquement invariant par similitude plane directe.

Représentation implicite

Leventon *et al* proposent dans [66] de représenter les formes d'apprentissage sous forme d'ensembles de niveaux. A l'instar des méthodes présentées, l'ACP est utilisée pour déterminer les modes principaux de déformation. La contrainte de forme est aussi dérivée d'une approche de type MAP. Cependant, elle apparaît dans l'équation d'évolution comme un terme correcteur égal à la différence entre l'ensemble de niveaux du contour actif et celui de la forme la plus probable (combinaison linéaire des modes principaux de l'apprentissage). Cette façon peu naturelle d'introduire la contrainte de forme est formalisée plus généralement dans la fonctionnelle d'énergie d'une approche variationnelle dans [105, 103]. Dans ces travaux, les formes d'apprentissage sont aussi représentées par leurs ensembles de niveaux. Ceux-ci sont premièrement alignés par l'estimation d'une similitude plane directe optimale selon l'équation (4.50). La forme moyenne est représentée comme la moyenne des ensembles de niveaux $\phi_{\mathcal{M}}(\mathbf{x}, t)$, la variance $\sigma_{\mathcal{M}}(\mathbf{x})$ est calculée pour chaque pixel à partir des formes de l'apprentissage. Le critère énergétique de contrainte s'inscrivant dans une démarche MAP est alors :

$$J_{shape}(\phi, T) = \int_{\Omega} \left(\frac{(s\phi(\mathbf{x}, t) - \phi_{\mathcal{M}}(T\mathbf{x}, t))^2}{s\sigma_{\mathcal{M}}(T\mathbf{x})} + \ln(\sigma_{\mathcal{M}}(T\mathbf{x})) \right) \delta_a(\phi(\mathbf{x}, t)) d\mathbf{x} \quad (4.55)$$

Le premier terme de cette énergie représente la contrainte de forme qui est normalisée par la variance moyenne mesurant la confiance du modèle. Ainsi, les zones à forte répétitivité parmi les échantillons d'apprentissage (faible variance) verront une contrainte de forme particulièrement renforcée.

4.2.2.3 Invariance de la contrainte de forme par transformation géométrique

Comme nous venons de le constater, l'invariance de la contrainte de forme par rapport à certaines transformations géométriques est cruciale. En effet, sans cette invariance, le pouvoir de segmentation serait réduit à un objet de l'image correspondant exactement à la référence. L'invariance par transformation affine est un problème récurrent en Vision par Ordinateur. Certaines représentations permettent une invariance naturelle à une sous-classe des transformations affines, comme par exemple les similitudes planes [33, 114]. D'autres ont besoin de l'ajout de paramètres extrinsèques permettant de réaliser l'invariance [105, 21, 23, 24].

Dans le cas d'un critère énergétique de contrainte de forme représentée par un ensemble de niveaux (équation (4.54)), il est possible de calculer analytiquement l'ensemble de niveaux ϕ_2 d'un contour représenté par ϕ_1 et transformé par une similitude plane directe T_{sim} :

$$\phi_2(\mathbf{x}) = s\phi_1(T_{sim}^{-1}\mathbf{x}) \quad (4.56)$$

où s est le facteur d'échelle de T_{sim} . L'équation (4.56) peut être explicitée sous la forme :

$$\phi_2(x, y) = s\phi_1\left(\frac{(x - \mu_x)\cos(\theta) + (y - \mu_y)\sin(\theta)}{s}, \frac{-(x - \mu_x)\sin(\theta) + (y - \mu_y)\cos(\theta)}{s}\right) \quad (4.57)$$

où θ est l'angle de rotation, et (μ_x, μ_y) la translation de T_{sim} . L'extension de la relation analytique entre ensembles de niveaux à une transformation avec un facteur d'échelle anisotrope ou une transformation affine est plus délicate et reste une question ouverte. Dans [30], D. Cremers propose une formulation de contrainte intrinsèque sans l'estimation parallèle des paramètres de pose. Cependant, elle reste limitée à l'invariance par translation et agrandissement/réduction par facteur d'échelle. Enfin, les travaux intéressants réalisés par Riklin-Raviv *et al.* [99] permettent de modéliser une projection de type perspective en intersectant la fonction de distance signée par un plan incliné (et non le plan horizontal du niveau 0).

4.2.3 Segmentation sous contrainte de forme générique : les modèles quadratiques

Une nouvelle classe de contours actifs d'ordre supérieur a été proposée par M. Rochery et I. Jermyn dans [102]. La finalité de ces récents modèles est d'incorporer des contraintes géométriques globales au sein du contour actif. Alors que cette contrainte est plus spécifique que celles génériques de régularisation du contour, elle l'est cependant moins que l'incorporation de forme *a priori* décrite aux paragraphes précédents. Les contours actifs proposés sont fondés sur une approche variationnelle et minimisent l'énergie *quadratique* suivante définie sur l'espace des 1-chaînes :

$$J_{quad}(C) = \oint \oint \left\langle \vec{\mathbf{t}}(p), F(\mathbf{C}(p), \mathbf{C}(p')) \vec{\mathbf{t}}(p') \right\rangle dp dp' \quad (4.58)$$

où $\vec{\mathbf{t}}(p)$, $\vec{\mathbf{t}}(p')$ sont les vecteurs tangents au contour actif $\mathbf{C}$ aux points p et p' . F est un tenseur covariant pondérant l'interaction entre les deux points du contour $\mathbf{C}(p)$ et $\mathbf{C}(p')$. La dérivation de J_{quad} par rapport à la courbe $\mathbf{C}$ permet de déduire une force de contrainte de forme non locale pour chaque point $\mathbf{C}(p)$ s'exprimant comme une intégrale simple sur le contour.

Dans [100], les auteurs utilisent les modèles quadratiques pour faciliter l'extraction du réseau routier à partir d'images de télédétection. Ils s'intéressent au cas particulier où le tenseur F est la matrice identité, et formulent une énergie visant à favoriser la segmentation de structures aux bords parallèles (routes). L'énergie quadratique alors proposée est :

$$J_{quad}(C) = - \oint \oint \left\langle \vec{\mathbf{t}}(p), \vec{\mathbf{t}}(p') \right\rangle \Psi(|\mathbf{C}(p) - \mathbf{C}(p')|) dp dp' \quad (4.59)$$

où Ψ est une fonction décroissante représentant le potentiel d'interaction entre les points du contour d'abscisse p et p' :

$$\Psi(x) = \begin{cases} 1 & \text{si } x < d_{min} - \varepsilon \\ 0 & \text{si } x > d_{min} + \varepsilon \\ \frac{1}{2} \left(1 - \frac{x - d_{min}}{\varepsilon} - \frac{1}{\pi} \sin\left(\pi \frac{x - d_{min}}{\varepsilon}\right) \right) & \text{sinon} \end{cases} \quad (4.60)$$

On remarque que l'énergie formulée en équation (4.59) est minimale lorsque deux points voisins du contour ($|\mathbf{C}(p) - \mathbf{C}(p')| < d_{min}$) ont des tangentes colinéaires. Ceci favorise l'extraction de structures rectilignes. Cette énergie est en revanche élevée pour deux points voisins avec des tangentes anti-parallèles. Ainsi, les deux bords parallèles d'une route extraite auront une largeur minimale d_{min} au delà de laquelle la répulsion due à l'énergie tend vers zéro en même temps que le profil Ψ . D'autres variantes d'énergies quadratiques impliquant des termes d'attache aux données ont été proposées. Les résultats de tels contours actifs représentés par leurs ensembles de niveaux sont prometteurs. L'insertion de telles contraintes géométriques fortes et non locales se révèle efficace, et la représentation implicite permet d'inclure naturellement la notion de réseau.

4.3 Contours actifs pour le recalage fin carte-image : choix et implantation

4.3.1 Choix de la représentation du contour actif

Dans le cadre du recalage fin carte-image que nous proposons d'étudier dans cette thèse, nous optons pour une *représentation implicite* du contour actif par *ensemble de niveaux* pour les raisons suivantes :

1. La représentation implicite permet une modélisation naturelle des coins. Ce point est particulièrement déterminant pour notre application puisque nous considérons des objets (bâtiments) présentant ces singularités géométriques.
2. La représentation par ensemble de niveaux est dépourvue de paramètres de réglage contrairement à la représentation explicite. Son utilisation est donc plus générale et aisée.
3. La représentation implicite permet de modéliser naturellement des objets à topologie complexe. Par exemple, des bâtiments comportant des cours intérieures pourront être modélisés.
4. La topologie d'un contour actif représenté par un ensemble de niveaux peut changer durant la convergence. Cette propriété semble peu pertinente pour notre application puisque la contrainte de forme est supposée déterminer la topologie du contour actif. Néanmoins nous verrons que la flexible topologie des ensembles de niveaux sera mise à profit pour résoudre le problème de minima locaux (cf. section 5.4).
5. Les ensembles de niveaux sont plus stables numériquement que les contours explicitement représentés, en particulier pour la modélisation de singularités géométriques.
6. La comparaison de deux formes représentées par leurs ensembles de niveaux est directe et efficace. La mise en œuvre de la contrainte de forme *a priori* s'en trouve facilitée.

Néanmoins, les inconvénients de la représentation implicite sont :

1. Une complexité calculatoire accrue. Cela est inhérent à la dimension d'ordre $n+1$ de l'ensemble de niveaux représentant le contour actif. Nous verrons en section 4.3.4.1 comment la restriction de l'ensemble de niveaux à une bande étroite permet de réduire les temps de calcul.
2. La fonction distance de l'ensemble de niveaux n'est pas préservée par les équations d'évolution de type Hamilton-Jacobi. Ceci implique une réinitialisation périodique de l'ensemble de niveaux au cours de l'évolution du contour actif. Le coût calculatoire s'en trouve ainsi augmenté.

4.3.2 Choix du terme d'attache aux données

Dans cette étude, nous nous servirons de deux modèles d'attache aux données basés sur *l'information de région* : celui de Chan et Vese [19] et le modèle Bayésien [91]. Le modèle de Chan et

Vese est d'autant plus efficace que l'image à analyser est constante par morceaux. Ceci restreint le champ d'application de notre méthode aux bâtiments dont la radiométrie du toit est uniforme. Nous verrons comment le modèle Bayésien permet de s'affranchir de cette hypothèse avec la possibilité de traiter les bâtiments homogènes. Notre motivation à choisir des contours actifs basés région repose sur leur moindre sensibilité à l'initialisation et au bruit comparés à ceux basés sur les frontières. Néanmoins, nous explorerons également le potentiel des contours actifs basés sur *l'information de frontière* pour leur capacité à segmenter des bâtiments non homogènes. Afin de rendre le contour moins sensible à l'initialisation, nous utiliserons la technique de Gradient Vector Flow (GVF) qui permet de diffuser l'information très locale des gradients de l'image. Les équations d'évolution des ensembles de niveaux relatives aux modèles de contours actifs que nous utiliserons sont :

4.3.2.1 Modèles basés sur l'information de région

– *Modèle de Chan et Vese*

$$\phi_t(\mathbf{x}, t) = -\delta_a(\phi(\mathbf{x}, t)) \left(-[I(\mathbf{x}) - c_{out}(\phi(\mathbf{x}, t))]^2 + [I(\mathbf{x}) - c_{in}(\phi(\mathbf{x}, t))]^2 \right) \quad (4.61)$$

avec ϕ l'ensemble de niveaux représentant le contour actif, δ_a est une approximation régularisée de la distribution de Dirac, c_{in} et c_{out} sont les moyennes de l'image I à l'intérieur et l'extérieur du contour respectivement.

– *Modèle Bayésien*

$$\phi_t(\mathbf{x}, t) = \left\{ -\frac{(I(\mathbf{x}) - c_{in}(\phi(\mathbf{x}, t)))^2}{\sigma_{in}^2(t)} + \frac{(I(\mathbf{x}) - c_{out}(\phi(\mathbf{x}, t)))^2}{\sigma_{out}^2(t)} + \ln \left(\frac{\sigma_{out}^2(t)}{\sigma_{in}^2(t)} \right) \right\} \delta_a(\phi(\mathbf{x}, t)) \quad (4.62)$$

où σ_{in}^2 et σ_{out}^2 sont les variances de l'image I à l'intérieur et l'extérieur du contour respectivement. Dans les équations (4.61-4.62), le terme de régularisation a été ignoré pour deux raisons : il a tendance à arrondir les coins que pourrait présenter le contour actif, nous cherchons à éviter cet effet puisque nous voulons segmenter des bâtiments dont les contours présentent souvent ces singularités. De plus, le terme de contrainte de forme aura un effet régularisateur.

– *Domaine d'application*

- Bâtiment peu élevé (distorsion due à la perspective négligeable). Cette contrainte provient de l'incompatibilité des géométries dans lesquelles les données cartographiques et de télé-détection sont projetées. Alors que la carte est orthoscopique, l'image satellitaire n'est pas rectifiée. Dans le cas des bâtiments très élevés, l'initialisation du contour actif provenant de la carte se situera au niveau de l'empreinte au sol dans l'image. Le contour actif sera ainsi éloigné du toit à segmenter, et rendra le résultat du recalage fin non robuste puisque les contours actifs sont sensibles à l'initialisation.
- Bâtiment sans effet de généralisation dans la carte, avec un toit de forme quelconque. La généralisation cartographique consistant à regrouper plusieurs bâtiments dans un unique polygone rendra la radiométrie de l'objet hétérogène, ce qui est incompatible avec les modèles basés région proposés.
- La représentation cartographique du bâtiment n'est pas entachée d'erreur. Nous discuterons à la fin du chapitre suivant de la présence d'erreurs dans la carte qui illustrera les limites du modèle proposé avec contrainte de forme (cf. section 5.6).

4.3.2.2 Modèle basé sur l'information de frontière

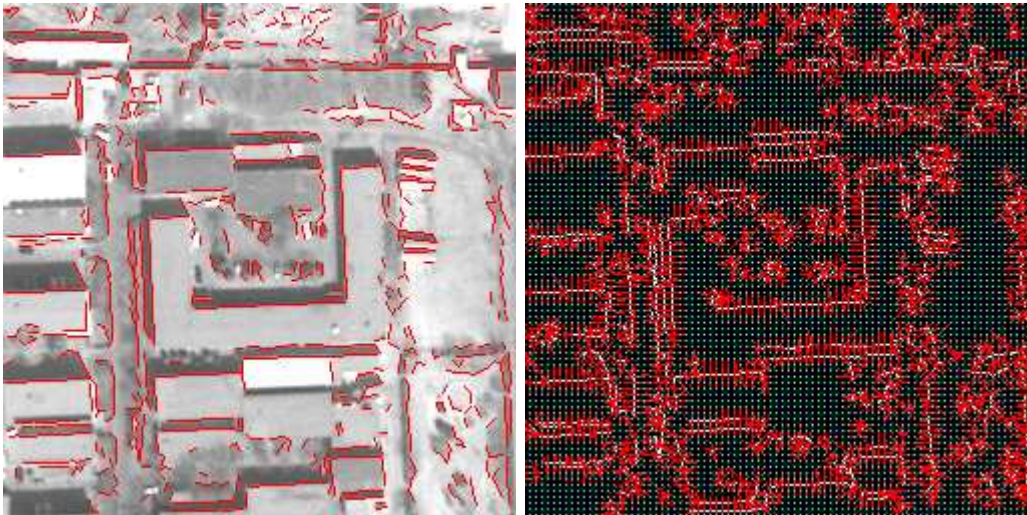
– *Modèle géodésique avec Gradient Vector Flow*

Nous adaptons le modèle exposé en équation (4.28) issu de [92] de la façon suivante :

$$\phi(\mathbf{x}, t)_t = - \langle (\mathbf{u}, \mathbf{v}), \nabla \phi(\mathbf{x}, t) \rangle \quad (4.63)$$

où $(\mathbf{u}, \mathbf{v})$ est le champ de vecteurs GVF résultant de la diffusion de l'information de gradient de l'image. Par rapport à l'équation (4.28), le terme de régularisation a été omis ainsi que le terme de pondération inversement proportionnel au module du gradient de l'image. Nous justifions la première omission par l'insertion ultérieure d'une contrainte de forme qui aura un effet de régularisation. La seconde est justifiée par le fait que le champ $(\mathbf{u}, \mathbf{v})$ est nul sur les lieux de haut gradient ce qui est redondant avec la fonction $g(|\nabla I|)$ de l'équation (4.28).

Enfin, la dernière adaptation au modèle que nous apportons est l'information de gradient qui sera diffusée pour le calcul du champ $(\mathbf{u}, \mathbf{v})$. Nous diffuserons des primitives segments extraites à partir des lieux de gradient élevé de l'image. Ces primitives ne sont pas seulement issues d'une détection de contours, mais d'un chaînage subséquent qui permet de trouver les segments de l'image. Les primitives segments sont d'un plus haut niveau et sont plus à même de caractériser un bâtiment dont l'empreinte au sol et le toit sont le plus souvent composés de bords rectilignes. Le calcul des segments opère ainsi tel un filtrage des gradients de l'image, évitant la prise en compte d'objets périphériques tels les arbres ou le mobilier urbain qui ne répondent pas au critère de rectilignité. La figure 4.3 illustre l'extraction de ces primitives.



(a) Primitives segments extraites par chaînage des points de gradient élevé (rouge)

(b) Vecteur de gradient des segments diffusé par la méthode de "Gradient Vector Flow"

FIG. 4.3 – Exemple d'extraction de segments (a) et du gradient vector flow associé (b).

– *Domaine d'application*

- Bâtiment peu élevé.
- Bâtiment avec ou sans effet de généralisation dans la carte, avec un toit de forme quelconque.

L'hétérogénéité des bâtiments regroupés en un seul objet peut être gérée par le modèle basé uniquement sur les gradients de l'image.

- La représentation cartographique du bâtiment n'est pas entachée d'erreur.

4.3.3 Choix de la contrainte de forme

L'information *a priori* dérivée de la carte revêt trois aspects. Premièrement, c'est une information sur la nature de l'objet que l'on cherchera à recaler dans l'image, nous traitons uniquement la couche du bâti. Deuxièmement, c'est une information de localisation qui nous permet de rechercher un bâtiment au bon endroit dans l'image, nous nous servons de cette connaissance pour initialiser le contour actif. Enfin, la carte fournit la forme du bâtiment que l'on est susceptible de retrouver dans l'image. Nous contraignons le contour actif à ressembler à cette forme dérivée de la carte en utilisant le critère énergétique décrit en équation (4.54). La dérivée de cette énergie par rapport à l'ensemble de niveaux ϕ donne l'équation suivante :

$$\phi(\mathbf{x}, t)_t = -\frac{\partial J_{image}}{\partial \phi} + 2\lambda (H(\phi(\mathbf{x})) - H(\psi(\mathbf{x}))) \delta_a(\phi(\mathbf{x}, t)) \quad (4.64)$$

avec

$$\psi = \psi_0 \circ T_{sim}$$

où : J_{image} est un critère énergétique quelconque d'attache aux données, ϕ est l'ensemble de niveaux représentant le contour actif ; ψ_0 est l'ensemble de niveaux représentant la forme *a priori* dérivée de la carte, $\lambda \in \mathbb{R}^+$.

Le critère de l'équation (4.54) a l'avantage d'être indépendant du domaine d'intégration Ω , et fournit une mesure discriminante de deux formes représentées implicitement par ϕ et ψ_0 [103]. La contrainte de forme est invariante par une transformation globale T_{sim} qui permet d'introduire des degrés de liberté entre le contour actif représenté par ϕ et la forme *a priori* statique ψ_0 . T_{sim} est une similitude plane directe qui permet de rendre l'insertion de forme *a priori* invariante par rotation d'angle θ , de translation μ et dilatation/diminution par facteur d'échelle s identique dans les directions (Ox) et (Oy) de l'image :

$$T_{sim}\mathbf{x} = s \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \mathbf{x} + \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix} \quad (4.65)$$

Selon l'expression de l'équation (4.54), il n'est pas possible de déduire analytiquement les paramètres $\xi = (s, \theta, \mu_x, \mu_y)$ de T_{sim} à partir du système :

$$\left(\frac{\partial J_{shape}}{\partial \xi_i} \right)_{i=1, \dots, 4} = 0 \quad (4.66)$$

A défaut, des techniques d'optimisation numérique sont employées pour estimer les paramètres $\hat{\xi}$ au cours de l'évolution du contour actif. L'optimisation par descente de gradient est couramment employée pour résoudre cette question. Nous discuterons d'une alternative à la descente de gradient en section 5.2.

4.3.4 Algorithme et optimisation

4.3.4.1 Réduction des temps de calculs : utilisation de la technique de bande étroite rapide

Si nous implantions au sens strict l'équation d'évolution (4.64), il faudrait passer en revue tous les pixels $\mathbf{x}$ de l'image pour renseigner complètement et mettre à jour la fonction ϕ . Ceci coûterait considérablement en temps de calcul, et serait inutile. En effet, seules les valeurs de ϕ proches de zéro nous intéressent car c'est le passage à zéro de l'ensemble de niveaux qui représente le contour actif. En pratique, on effectue la mise à jour dans une bande étroite autour du niveau zéro de l'ensemble ϕ . Lorsque le contour actif est sur le point de sortir de la bande étroite, cette dernière est reconstruite à partir de la dernière position du contour.

La fonction distance de l'ensemble de niveaux n'est pas préservée par les équations d'évolution de type Hamilton-Jacobi qui sont habituellement dérivées des approches variationnelles. Ainsi, la condition $|\nabla\phi| = 1$ n'est plus vérifiée au cours du temps. Deux options sont possibles pour résoudre ce problème. La plus satisfaisante consiste à dériver des équations d'évolution qui ne sont pas de type Hamilton-Jacobi, c'est ce que propose José Gomes dans son travail de thèse, ou plus synthétiquement dans [43]. La complexité des calculs est accrue mais résout le problème de façon élégante. La deuxième façon que nous avons choisie consiste à réinitialiser l'ensemble de niveaux toutes les N itérations comme lorsque le contour est sur le point de sortir de la bande étroite. Le concept de réinitialisation dans une bande étroite a été originellement implanté dans un article de D.L. Chopp traitant de la physique des interfaces [26]. Dans [116] Sussman propose une méthode de reconstruction de bande étroite selon l'équation différentielle :

$$\phi_\tau(\mathbf{x}, t) = \text{sign}(\phi(\mathbf{x}, \tau)) (1 - |\nabla\phi(\mathbf{x}, \tau)|) \quad (4.67)$$

Une telle approche propose une approximation de la fonction distance, qui est d'autant plus fine que le nombre d'itérations est grand. Le problème d'une telle méthode est la nécessité de calculer le gradient de l'ensemble de niveaux ϕ , ce qui peut être délicat au voisinage d'une singularité géométrique. Une mauvaise estimation du gradient en ces lieux entraînera et propagera des erreurs dans l'estimation de la fonction distance. Dans [60], R. Keriven discute d'algorithmes de réinitialisation plus anciens. R. Keck discute de façon très détaillée de l'influence de la réinitialisation sur la précision des calculs dans le cadre de la mécanique des fluides [59]. Dans [90], N. Paragios et R. Deriche proposent une méthode locale plus rapide que [116] baptisée Hermès, dans [111] Sethian décrit une technique basée sur le *Fast Marching*. Nous avons opté pour la méthode proposée par Yui en 2002, appelée "bande étroite rapide" (*Fast Narrow Band*), et qui est plus rapide que les approches précitées [132]. Cette méthode consiste en deux passes :

1. Pour chaque pixel de coordonnées $\mathbf{x}$ appartenant au niveau zéro de l'ensemble de niveaux, on assigne une approximation de la fonction distance aux pixels voisins selon les voisinages circulaires illustrés en figure 4.4 : on réitère ce processus pour des voisinages circulaires progressivement plus grands jusqu'à atteindre la largeur de bande étroite désirée. Lorsqu'un pixel se voit assigner plusieurs fois une valeur différente, la valeur minimale est gardée. A la fin de cette phase, une bande étroite contenant des approximations grossières de la fonction distance est construite.
2. Pour chaque pixel de la bande étroite, on recherche les pixels appartenant au contour (niveau 0 de l'ensemble de niveaux) dans un voisinage circulaire de rayon égal à la valeur préalablement assignée. Parmi tous les pixels de contour trouvés, leurs distances Euclidiennes au pixel considéré sont calculées ; la valeur minimale est finalement attribuée au pixel considéré de la bande étroite.

Cette méthode astucieuse permet de manipuler des voisinages circulaires ce qui limite considérablement l'espace de recherche des pixels du contour. Néanmoins, on peut se montrer critique quant à la nécessité du contour d'avoir une précision pixellaire et non sub-pixellaire. Cette perte de précision peut se révéler pénalisante lors de calculs ultérieurs de courbure. Les auteurs proposent de corriger ce problème par diffusion à l'issue de la deuxième phase de l'algorithme. Nous proposons d'adapter l'algorithme de la bande étroite rapide pour calculer une fonction distance de précision sub-pixellaire :

1. Pour chaque point de coordonnées sub-pixellaires $\mathbf{x}_{contour,sp}$ appartenant au niveau zéro de l'ensemble de niveaux, on arrondit les coordonnées sur la grille pixellaire en mémorisant le lien $\mathbf{x}_{contour} \rightarrow \mathbf{x}_{contour,sp}$, puis on assigne une approximation de la fonction distance aux pixels voisins selon les voisinages circulaires illustrés en figure 4.4 : on réitère ce processus pour des voisinages circulaires progressivement plus grands jusqu'à atteindre la largeur de bande étroite désirée. Lorsqu'un pixel se voit assigner plusieurs fois une valeur différente, la valeur minimale est gardée. A la fin de cette phase, une bande étroite contenant des approximations grossières de la fonction distance est construite.
2. Pour chaque pixel de la bande étroite, on recherche les pixels appartenant au contour dans un voisinage circulaire de rayon égal à la valeur préalablement assignée en ce pixel. Pour chaque pixel de contour trouvé $\mathbf{x}_{contour,k}$, on évalue la distance Euclidienne aux points sub-pixellaires $\mathbf{x}_{contour,sp,k,i}$ reliés à $\mathbf{x}_{contour,k}$. La valeur minimale parmi tous les $\mathbf{x}_{contour,sp,k,i}$ est finalement attribuée au pixel considéré de la bande étroite.

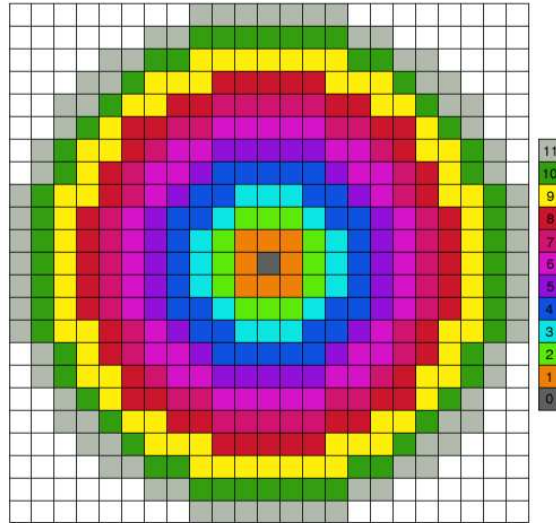


FIG. 4.4 – Voisinages circulaires utilisés pour la construction de la bande étroite rapide. La légende fournit une approximation de la distance pixellaire au centre.

La construction d'une bande étroite intervient donc lors de trois types d'évènements : pour la construction de $\phi(\mathbf{x}, t = 0)$, pour la reconstruction de la bande lorsque le niveau zéro de l'ensemble de niveaux est sur le point d'en sortir, et pour réinitialiser l'ensemble de niveaux lorsque celui-ci dégénère au cours du temps. En pratique, nous avons préféré travailler dans une bande très étroite de deux ou trois pixels, en réinitialisant à chaque itération de l'équation d'évolution du contour actif.

Si la construction de bande étroite s'avère indispensable, son opération duale est aussi nécessaire. Retrouver le contour à partir d'un ensemble de niveaux par détection des zéros de ce dernier intervient pour deux raisons : l'affichage du contour actif, et la réinitialisation de la bande étroite.

Le contour extrait de l'ensemble de niveaux est une approximation linéaire du contour réel (figure 4.5). On le détecte par changement de signe de la fonction $\phi(\mathbf{x}, t)$, et on calcule ses coordonnées sub-pixellaires par interpolation linéaire.

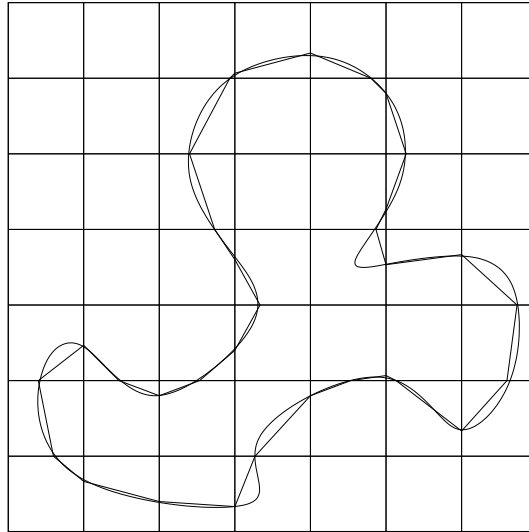


FIG. 4.5 – Approximation linéaire d'un contour.

4.3.4.2 Diagramme d'évolution du contour actif contraint

Le diagramme illustré en figure 4.6 décrit l'algorithme de mise à jour du contour actif. Sur cette figure, la fonction $\mathcal{F}$ est définie par l'équation d'évolution. Dans toutes les expériences menées sur des images réelles, la forme *a priori* polygonale dérivée de la carte et le contour actif initial sont identiques.

- En pratique, l'espace image considéré pour faire évoluer le contour actif est un sous-ensemble de Ω . Typiquement, une sous-image de 256 par 256 pixels est extraite de l'image satellitaire Quickbird (qui en fait 30,000 par 30,000!) à l'endroit du polygone cartographique que l'on désire recalcr et réviser. Si la taille de la boîte englobante du polygone est supérieure à 256 par 256 pixels, on extrait une sous-image plus grande ayant une marge de 100 pixels par rapport à la boîte englobante.
- Dans le cas où les modèles basés région sont utilisés, la caractéristique d'homogénéité des bâtiments représentés dans l'image peut être forcée en la pré-traitant par une diffusion anisotropique (préservant les contours).
- La génération de l'ensemble de niveaux ψ_0 représentant la forme *a priori* est effectuée à l'initialisation et couvre la totalité de la sous-image étudiée (contrairement à l'ensemble ϕ uniquement calculé dans une bande étroite). Pour réduire les temps de calcul, nous utilisons l'algorithme de Maurer [71] qui permet de calculer la distance Euclidienne sur des images binaires en $O(N)$ où N est le nombre de pixels de l'image. Le polygone cartographique vectoriel est ainsi préalablement converti en coordonnées pixellaires avant l'application de cet algorithme.
- Nous avons décidé d'estimer les paramètres de T_{sim} après chaque mise à jour du contour actif.

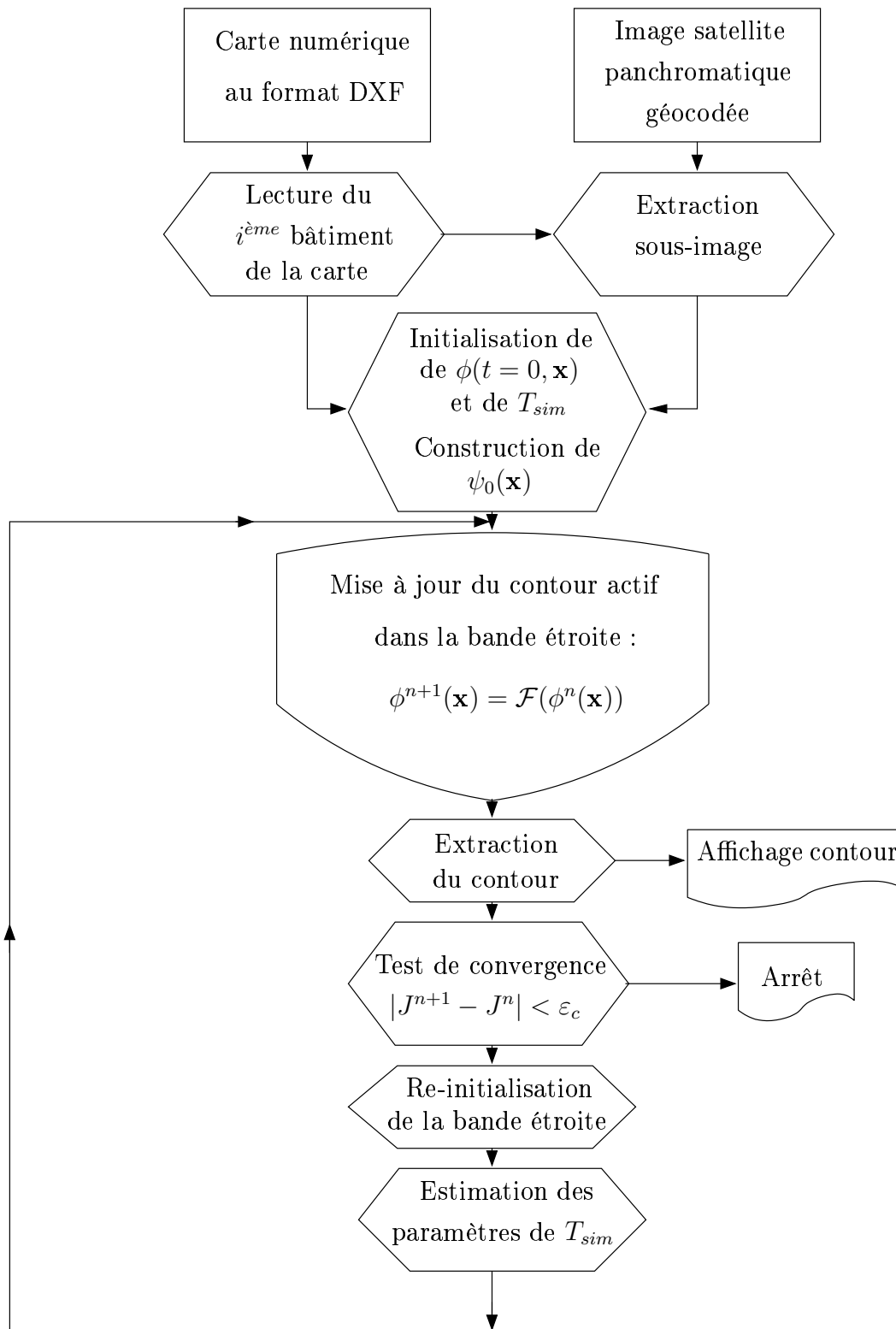


FIG. 4.6 – Diagramme de l'algorithme d'évolution du contour actif

4.4 Conclusion

Ce chapitre a retracé un état de l'art des contours actifs avec une attention particulière pour l'incorporation de contrainte de forme *a priori*. Nous avons exposé notre choix d'un contour actif représenté par ensemble de niveaux pour mener à bien le recalage fin des bâtiments de la carte sur l'image. Différents termes d'attache aux données ont été retenus pour conduire les contours actifs vers les bâtiments de l'image. Deux modèles basés sur l'information de région ont été choisis pour leur simplicité d'implantation et leur moindre sensibilité à l'initialisation. Ils seront néanmoins limités à segmenter les bâtiments homogènes de l'image. Un modèle basé sur l'information de frontière, intrinsèquement plus sensible à l'initialisation, sera utilisé pour le recalage avec les bâtiments non homogènes. Enfin, les détails relatifs à l'implantation et la réduction de la complexité calculatoire des ensembles de niveaux ont été décrits. Le chapitre suivant exposera nos contributions pour améliorer la robustesse des modèles avec contrainte de forme choisis.

Chapitre 5

Contours actifs contraints robustes

5.1 Introduction

Ce chapitre étudie l'amélioration de la robustesse des contours actifs contraints par la forme et destinés à réaliser une mise en correspondance entre les bâtiments cartographiques et l'image. Le but d'un tel recalage est d'accroître la cohérence carte-image afin de permettre une détection de non-changement subséquente fiable. Nous proposons d'améliorer les modèles de contours actifs proposés au chapitre précédent afin d'accroître leur pouvoir de recalage. La première contribution consiste à proposer un algorithme d'optimisation différent pour l'estimation des paramètres de pose rendant la contrainte de forme invariante par transformation globale. La seconde est relative à l'insertion du modèle numérique de surface (MNS) dans le terme d'attache aux données des contours actifs. Une telle représentation est complémentaire de l'image satellitaire et permet une meilleure discrimination du bâti du reste de la scène. La troisième innovation regarde l'incorporation d'une contrainte de forme variant en fonction de l'espace et du temps. Celle-ci est destinée à conférer plus de flexibilité aux contours actifs afin de les rendre moins sensibles aux minima locaux de l'énergie minimisée. Des résultats expérimentaux proposés en fin de chapitre permettent d'évaluer la contribution de ces trois solutions à accroître la robustesse et l'efficacité, en terme de temps de calcul, du recalage carte-image.

5.2 Estimation robuste des paramètres de T_{sim} par algorithme du simplexe

Comme nous l'évoquions en section 4.3.3, le minimum de l'énergie de contrainte de forme décrite de façon générale en équation (5.1) peut ne pas avoir de solution analytique des paramètres $\xi_{i \in (1, \dots, N)}$ de la transformation globale T :

$$J_{shape}(\phi, \psi) = \int_{\Omega} \mathcal{D}(\phi(\mathbf{x}), \psi_0(T\mathbf{x})) d\mathbf{x} \quad (5.1)$$

où $\mathcal{D}$ représente une distance mesurant la similarité entre la forme du contour actif ϕ et la forme de référence ψ_0 . C'est le cas de l'énergie formulée en équation (4.54) où $\mathcal{D}(\phi(\mathbf{x}), \psi_0(T\mathbf{x})) = (H(\phi(\mathbf{x})) - H(\psi_0(T\mathbf{x})))^2$ et $T = T_{sim}$ ($N = 4$). A défaut d'avoir une expression formelle des paramètres qui aurait rendu l'approche intrinsèque, il est alors nécessaire de les estimer par optimisation numérique au cours du processus d'évolution du contour actif. Dans [21, 23, 105, 24], les auteurs proposent de calculer numériquement les paramètres d'une similitude plane directe minimisant l'énergie de l'équation (4.54) par descente de gradient. On a alors quatre équations relatives à

chacun des paramètres de T_{sim} , à savoir l'échelle s , la rotation θ et la translation μ :

$$\begin{cases} s^{n+1} = s^n + \Delta t_s \mathcal{F}(\phi, \psi_0, s^n, \theta^n, \mu_x^n, \mu_y^n) \\ \theta^{n+1} = \theta^n + \Delta t_\theta \mathcal{G}(\phi, \psi_0, s^n, \theta^n, \mu_x^n, \mu_y^n) \\ \mu_x^{n+1} = \mu_x^n + \Delta t_{\mu_x} \mathcal{M}(\phi, \psi_0, s^n, \theta^n, \mu_x^n, \mu_y^n) \\ \mu_y^{n+1} = \mu_y^n + \Delta t_{\mu_y} \mathcal{N}(\phi, \psi_0, s^n, \theta^n, \mu_x^n, \mu_y^n) \end{cases} \quad (5.2)$$

où n est le rang de l'itération, ϕ et ψ_0 sont les ensembles de niveaux représentant respectivement le contour actif et la forme *a priori*. Δt_s , Δt_θ , Δt_{μ_x} et Δt_{μ_y} sont les pas temporels associés à chacune des descentes de gradient estimant les paramètres de T_{sim} . Le détail des expressions des fonctions $\mathcal{F}$, $\mathcal{G}$, $\mathcal{M}$ et $\mathcal{N}$ est disponible dans [21].

L'estimation des paramètres grâce aux quatre descentes du système (5.2) peut se révéler extrêmement périlleuse. Le point délicat est le réglage des pas temporels des quatre descentes relatives à chacun des paramètres. Alors que chaque descente estime un paramètre avec un sens géométrique différent (translation, rotation, échelle), il est fort probable que la vitesse de chaque descente ne soit pas la même. A titre d'exemple, la descente estimant la rotation devrait être bien moins rapide que celles de la translation. Cependant, il n'existe aucun cadre théorique permettant de clarifier cette question ce qui donne lieu à des réglages empiriques difficiles, instables et influençant l'estimation des paramètres. Enfin, l'inter-dépendance des quatre descentes de gradient fait qu'une mauvaise estimation d'un paramètre dans une descente va biaiser les autres au point de faire diverger rapidement l'optimisation. Daniel Cremers mentionne le problème du réglage des descentes de gradients dans [30].

Une alternative plus heureuse que nous avons choisie est l'utilisation de l'algorithme du *simplex* (le détail de cet algorithme est rappelé et illustré en annexe B). On minimise alors l'énergie de contrainte de forme exposée en équation (5.1) en trouvant un jeu de N paramètres optimaux estimés simultanément. L'optimisation se fait par réflexion, expansion et contraction d'un polyèdre de $N + 1$ sommets dans l'espace des paramètres. L'avantage du simplex est l'absence de paramètres de réglage influençant fortement le résultat d'optimisation. Cette approche, moins sensible aux minima locaux de l'énergie, est plus robuste que la descente de gradient. C'est une méthode d'ordre zéro (ne nécessitant pas le calcul du gradient de la fonction de coût) plus à même de trouver un minimum d'une fonctionnelle non strictement convexe. Enfin, le simplex est capable de minimiser des formes plus complexes d'énergie de contrainte de forme, avec plus de paramètres à estimer, sans nécessiter le calcul théorique des dérivées partielles $\frac{\partial J_{shape}}{\partial \xi_i}$ inhérent aux descentes de gradient.

Cependant, le simplex a une complexité calculatoire accrue par rapport à la descente de gradient. En effet, la construction itérative des sommets du simplex demande l'évaluation de la fonction de coût J_{shape} qui est ainsi effectuée de nombreuses fois durant le processus de convergence. J_{shape} est une intégrale sur tout le domaine de l'image et implique une complexité calculatoire élevée, ce qui augmentera significativement les temps de calcul par rapport à la descente de gradient. On peut néanmoins limiter le calcul intégral de J_{shape} au sein d'une bande étroite pour diminuer la charge calculatoire. L'expression de l'énergie de l'équation (4.54) devient alors :

$$J_{shape}(\phi, \psi) = \int_{\Omega} N(\phi(\mathbf{x}), \varepsilon_b) (H(\phi(\mathbf{x})) - H(\psi(\mathbf{x})))^2 d\mathbf{x} \quad (5.3)$$

où la fonction $N(\phi(\mathbf{x}), \varepsilon_b)$ est égale à 1 si $|\phi(\mathbf{x})| \leq \varepsilon_b$, sinon elle est égale à 0. La valeur de ε_b qui est choisie arbitrairement est un compromis entre la complexité calculatoire et la précision souhaitée de l'estimation des paramètres. En effet, une bande large permettra une bonne différenciation

des formes ϕ et ψ , ce qui garantira une évaluation précise et robuste des paramètres, au prix d'un coût calculatoire élevé. A contrario, une bande trop étroite induira une convergence rapide vers une estimation imprécise, voire même un minimum local de la fonctionnelle d'énergie. Une demi-bande de trois pixels $\varepsilon_b = 3$ s'est révélée un compromis satisfaisant. Le seul paramétrage de l'algorithme du simplex est relatif à la construction du simplex initial. Le premier sommet du simplex est calculé à partir du jeu des paramètres initiaux ξ_{ini} , le reste des sommets est construit en ajoutant une variation individuelle $\delta_{\xi(i)}$ à chacun des paramètres de ξ_{ini} . Dans le cas de T_{sim} , quatre paramètres sont à estimer. Le simplex aura ainsi cinq sommets et son état initial sera :

- Sommet initial 1 : calculé à partir de $(s_{ini}, \theta_{ini}, \mu_{x,ini}, \mu_{y,ini})$
- Sommet initial 2 : calculé à partir de $(s_{ini} + \delta_s, \theta_{ini}, \mu_{x,ini}, \mu_{y,ini})$
- Sommet initial 3 : calculé à partir de $(s_{ini}, \theta_{ini} + \delta_\theta, \mu_{x,ini}, \mu_{y,ini})$
- Sommet initial 4 : calculé à partir de $(s_{ini}, \theta_{ini}, \mu_{x,ini} + \delta_{\mu_x}, \mu_{y,ini})$
- Sommet initial 5 : calculé à partir de $(s_{ini}, \theta_{ini}, \mu_{x,ini}, \mu_{y,ini} + \delta_{\mu_y})$

Le choix des valeurs de $\delta_{\xi(i)}$ conditionne la taille du simplex initial et ainsi son aptitude à estimer des paramètres éloignés de ξ_{ini} . Le choix le plus sûr est un simplex initial relativement grand, ce qui garantira une meilleure robustesse de l'optimisation au prix d'un coût calculatoire plus élevé. La comparaison des deux algorithmes à partir de résultats expérimentaux est exposée en section 5.5.2.

5.3 Fusion de termes exogènes dans la fonctionnelle d'énergie d'attache aux données

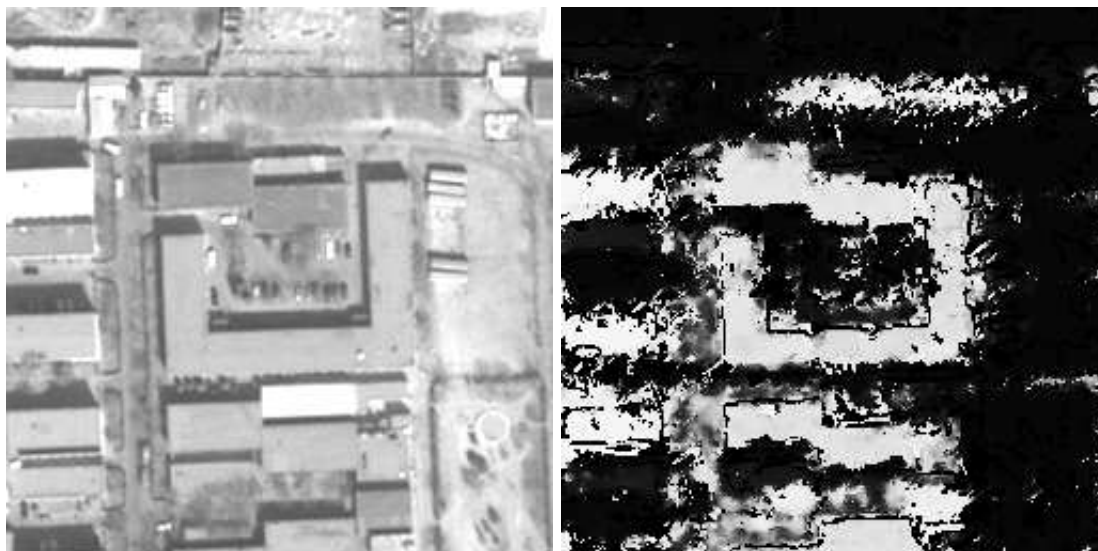
L'examen de la figure 5.1 montre que les bâtiments représentés dans l'image satellitaire panchromatique sont localement discriminés du reste de l'image, mais pas globalement. Ainsi, la radiométrie d'un bâtiment peut être différente de son voisinage proche sans être pour autant discriminée du reste de l'image au sens de quantités statistiques telles que la moyenne. Ces quantités sont cependant utilisées dans les modèles basés sur l'information de région que nous avons choisis. Cette difficulté inhérente aux milieux urbains aura pour effet de rendre la convergence du contour actif basé région sensible à l'initialisation. Celui-ci aura alors plus de risque d'atteindre un minimum local de la fonctionnelle d'énergie à minimiser.

Cependant, le Modèle Numérique de Surface (MNS) que nous avons auparavant généré (cf. chapitre 3) semble mieux répondre aux caractéristiques des modèles basés sur l'information de région. En effet, l'intensité des pixels du MNS est proportionnelle à l'altitude des objets reconstruits, garantissant aux bâtiments de se détacher de leur environnement proche, ou du reste de l'image si la densité du bâti n'est pas trop élevée. On peut forcer la discrimination d'un bâtiment représenté dans le MNS en modifiant l'histogramme de ce dernier. Pour notre application, nous avons adopté le pré-traitement suivant :

$$I_{MNS}^{clipped}(\mathbf{x}) = \frac{1}{2} \max_{\mathbf{x} \in I_{MNS}} (I_{MNS}(\mathbf{x})) \left[1 + \frac{2}{\pi} \arctan \left(\frac{I_{MNS}(\mathbf{x}) - m + s}{\zeta} \right) \right] \quad (5.4)$$

avec m : moyenne d'altitude du MNS à l'intérieur du polygone cartographique superposé au MNS ; $s = \zeta \tan \left(\frac{\pi(2ratio-1)}{2} \right)$ garantit que les pixels du MNS de valeur m se verront assigner une valeur égale à $\max(I_{MNS}(\mathbf{x})) \times ratio$. La valeur de $ratio$ a été fixée à 0.95 dans nos expériences. La variable ζ permet de régler la pente de la fonction arctan, plus ζ sera de valeur faible devant 1, plus la pente sera élevée lorsque $I_{MNS}(\mathbf{x}) = m - s$, accroissant ainsi le contraste entre le bâtiment considéré et les objets moins élevés. A contrario, le bâtiment aura une intensité très peu différente des objets

plus élevés. Nous avons choisi $\zeta = 0.5$. La figure 5.1 présente un bâtiment dans une image satellite et dans le MNS pré-traité selon l'équation (5.4). On constate que le MNS pré-traité se rapproche plus d'une image bi-phasique au sens de la moyenne, et aura plus de chance de fonctionner avec le terme d'attache aux données du modèle de Chan et Vese ou du modèle Bayésien. Néanmoins, la reconstruction 3D du MNS est très "bruitée", les pixels non renseignés y sont nombreux. Ainsi, il semble difficile d'obtenir une segmentation précise du bâtiment en utilisant uniquement le MNS.



(a) Image satellitaire

(b) MNS (intensité de l'image proportionnelle à la hauteur)

FIG. 5.1 – Représentations différentes de la même scène : image satellitaire et reconstruction 3D.

Nous proposons de fusionner l'information radiométrique de l'image satellitaire et d'altitude du MNS pour mener à bien la segmentation de bâtiment. Cette fusion de données exogènes est motivée par les représentations complémentaires de la scène que proposent l'image de télédétection et le MNS :

- Dans l'image satellitaire, les bords du bâtiment sont le plus souvent faciles à caractériser et l'image a un rapport signal à bruit très élevé. Cependant, la présence d'ombres, d'objets périphériques au sol ou occultant partiellement le bâtiment rend la segmentation difficile. L'hétérogénéité de la radiométrie du toit ou de l'environnement du bâtiment peut mettre en échec les modèles d'attache aux données basés sur les régions homogènes.
- A l'inverse, le MNS représente mal les bords du toit du bâtiment car la reconstruction en trois dimensions est délicate sur les lieux de transition d'altitude. Cependant, le problème des ombres et des objets périphériques situés au niveau du sol est écarté. Le bâtiment représenté dans le MNS est mieux discriminé du reste de l'image.

Ces deux représentations se complètent de façon satisfaisante. Cependant si on désire les fusionner, elles doivent remplir certaines conditions de cohérence :

1. **Les données doivent représenter le même objet d'intérêt** : ceci pose le problème de

la date d'acquisition des données. Étant exogènes, les données sont susceptibles d'avoir été acquises à des instants différents, instants entre lesquels des changements auraient pu survenir. Ainsi, un bâtiment représenté dans le MNS généré à partir d'images aériennes de 1999 peut très bien avoir disparu sur l'image Quickbird de 2002. Cette condition de cohérence nous impose de tester la fusion de données exogènes sur des zones où l'objet d'intérêt de la scène est inchangé entre les dates d'acquisition.

2. **Les données doivent être superposables** : chaque pixel des données à fusionner doit représenter la même partie d'un objet de la scène. Cette clause soulève deux questions. La première est relative à la précision du recalage des données. Dans notre cas, l'image et le MNS sont géocodés dans le même système cartographique. L'image a une précision de géocodage inférieure au pixel alors que celle du MNS est égale à deux pixels. Ceci semble suffisant, sachant qu'à la fin de la convergence du contour actif, il est possible de réduire l'influence du MNS et ainsi d'accorder plus de poids à l'image qui est la plus fiable. La deuxième question concerne la géométrie des données à fusionner. Le MNS est en géométrie orthoscopique alors que l'image satellitaire Quickbird ne l'est pas. Ceci causera assurément des problèmes de superposition pour les bâtiments élevés qui présentent une distorsion due à la perspective dans l'image : le toit et l'empreinte au sol ne sont alors plus superposés, contrairement au MNS. Cela nous contraindra à traiter uniquement les bâtiments peu élevés, à défaut de posséder une image satellitaire orthoscopique.

En plus de ces deux conditions de cohérence, une hypothèse implicite a été faite sur l'indépendance des données à fusionner. Celle-ci est réalisée par le caractère exogène des images :

$$p(I_{sat}, I_{MNS} | \phi) = p(I_{sat} | \phi) p(I_{MNS} | \phi)$$

où I_{sat} est l'image satellite, I_{MNS} est l'image du MNS. En supposant que les deux conditions de cohérence ainsi que celle d'indépendance sont vérifiées, nous fusionnons l'image satellitaire et le MNS dans la fonctionnelle d'énergie d'attache aux données basée région de Chan et Vese de la façon suivante :

$$\begin{aligned} J_{\text{fusion,CV}}(\phi) &= J_{\text{CV}}(\phi) \\ &+ \lambda_{MNS} \left[\int_{\Omega} |I_{MNS}(x, y) - c_{MNS, in}|^2 H(\phi(x, y)) dx dy \right. \\ &\left. + \int_{\Omega} |I_{MNS}(x, y) - c_{MNS, out}|^2 (1 - H(\phi(x, y))) dx dy \right] \end{aligned} \quad (5.5)$$

où J_{CV} est définie en équation (4.31), I_{MNS} est l'image du MNS. Les quantités statistiques indexées par MNS sont définies comme en équations (4.32-4.33) mais réfèrent au MNS. Le terme que nous venons d'ajouter à la fonctionnelle J_{CV} est le terme énergétique d'attache aux données de Chan et Vese relatif à la donnée exogène du MNS.

L'extension du modèle Bayésien de l'équation (4.44) est similaire :

$$\begin{aligned} J_{\text{fusion,Bayes}}(\phi) &= J_{\text{Bayes}}(\phi) \\ &+ \lambda_{MNS} \left[\int_{\Omega} \left(\frac{|I_{MNS}(x, y) - c_{MNS, in}|^2}{2\sigma_{MNS, in}^2} + \ln(2\pi\sigma_{MNS, in}^2) \right) H(\phi(x, y)) dx dy \right. \end{aligned}$$

$$+ \int_{\Omega} \left(\frac{|I_{MNS}(x, y) - c_{MNS, out}|^2}{2\sigma_{MNS, out}^2} + \ln(2\pi\sigma_{MNS, out}^2) \right) (1 - H(\phi(x, y))) dx dy \quad (5.6)$$

Il est aussi concevable de fusionner la donnée du MNS avec le modèle d'attache aux données basé sur l'information de frontière (GVF) (section 4.3.2.2). Nous choisissons d'intégrer la donnée exogène du MNS avec un modèle basé sur l'information de région. En effet, le MNS répond mieux à la caractéristique d'une image constante ou homogène par morceaux. Il paraît donc plus justifié d'utiliser un modèle basé sur l'information de région plutôt que sur les frontières. L'intégration du MNS selon le modèle de Chan et Vese revient à l'équation suivante :

$$\phi(\mathbf{x}, t)_t = - \langle (\mathbf{u}, \mathbf{v}), \nabla \phi(\mathbf{x}, t) \rangle$$

$$- \lambda_{MNS} \delta_a(\phi(\mathbf{x}, t)) \left(- [I_{MNS}(\mathbf{x}) - c_{MNS, out}(\phi(\mathbf{x}, t))]^2 + [I_{MNS}(\mathbf{x}) - c_{MNS, in}(\phi(\mathbf{x}, t))]^2 \right) \quad (5.7)$$

Selon les modèles des équations (5.5) à (5.7), le contour actif évoluera parallèlement sur deux images : le MNS et l'image satellitaire, et sera un compromis entre les informations dérivées de ces deux représentations ainsi que celle de la contrainte de forme *a priori*. Il est à noter que la précision de géocodage du MNS ($\simeq 1m$) est inférieure à celle de l'image rectifiée Quickbird qui a pour vocation d'être la source d'information la plus fiable ($\simeq 0.2m$). Ainsi, nous annulons le poids λ_{MNS} à la fin du processus de convergence pour que le contour actif soit plus fidèle à l'image. La phase préliminaire consistant à accorder plus de crédit au MNS est destinée à se rapprocher rapidement de la solution finale sans les problèmes des minima locaux.

Dans [47], T. Guo fait évoluer un contour actif de type *snake* sur une image de télédétection orthoscopique et un MNS réalisé par LIDAR. Une différence notable avec notre approche est l'utilisation d'un terme d'attache aux données basé sur les frontières dans l'image altimétrique et satellitaire, ce qui impose l'emploi de forces artificielles de gonflage. De plus, aucune contrainte de forme n'était employée et un MNS de meilleure qualité, issu d'une mesure et non d'une reconstruction, était utilisé.

La contribution du MNS pour une moindre sensibilité aux minima locaux est présentée en section 5.5.5.1.

5.4 Insertion plus souple de la forme *a priori* : flexibilité spatio-temporelle de λ

Toujours dans l'optique de rendre la mise en correspondance carte-image plus robuste, nous discutons dans cette section de l'influence du poids λ qui met en balance l'information de l'image et la contrainte de forme dans les équations (4.48) et (4.64). Le réglage de cette constante pondérant l'influence de la contrainte de forme n'est pas trivial pour les raisons suivantes :

1. Une contrainte trop faible ne permet pas de régulariser suffisamment le contour actif qui sera ainsi sensible aux altérations de l'image. Ceci mènera à une segmentation erronée de l'objet qui correspond à un minimum local de la fonctionnelle J de l'équation (4.48). Dans ce cas, les minima locaux, que nous nommerons de première espèce, sont nombreux.
2. A l'opposé, une prépondérance de la contrainte de forme limite les possibilités du contour actif à s'échapper des minima locaux de la fonctionnelle à minimiser J . Dans ce genre de situation, il en coûtera plus au contour actif de se déformer pour atteindre la cible dans l'image que de rester immobile et fidèle à la forme imposée par la contrainte. La solution choisie est ainsi un

minimum local (de seconde espèce) qui correspond aussi à une mauvaise segmentation/recalage de l'objet dans l'image.

Déterminer *a priori* une valeur optimale λ_{opt} du poids de contrainte paraît une tâche délicate puisque celle-ci dépendrait à la fois de l'image et du degré d'altération qu'elle comporte ; ces caractéristiques ne sont néanmoins pas connues à l'avance. Dans le contexte de l'imagerie médicale, Y. Chen propose d'essayer plusieurs valeurs du poids de contrainte et de retenir la valeur correspondant à la meilleure solution évaluée grâce à des profils radiométriques [22]. Cette solution paraît sous-optimale puisque λ_{opt} n'est pas assurément estimé, de plus, le processus de segmentation doit être répété plusieurs fois. Enfin, le réglage délicat de λ est d'autant plus difficile que les objets des images traitées sont peu discriminés du fond. Dans ce cas, l'information relative à l'image est affaiblie comparée à la contrainte de forme qui devient prédominante, cela augmente l'occurrence de minima locaux de seconde espèce. Cet effet est d'autant plus significatif lorsque le contour actif est initialisé trop loin de l'objet à retrouver dans l'image.

Pour résoudre le problème des minima locaux inhérents à une sous ou sur-estimation de la contrainte de forme, nous proposons de convertir le poids constant λ des équations (4.48) et (4.64) en une fonction de l'espace. Le but d'une telle fonction est de relaxer l'influence de la contrainte dans un voisinage du niveau 0 de l'ensemble de niveaux ψ , tout en préservant une contrainte forte et uniforme loin de la forme *a priori*. Ceci autorise une plus grande liberté et flexibilité du contour actif qui restent néanmoins spatialement bornées. Par conséquent, le contour actif sera à même de ressembler globalement à la forme *a priori* tout en autorisant des variations locales. Cette flexibilité est une propriété déterminante pour éviter les minima locaux de seconde espèce. Nous proposons de formaliser le profil spatial de relaxation comme une fonction symétrique de la distance à la forme *a priori* :

$$\lambda_{space}(\psi(\mathbf{x})) = \lambda \left(1 - e^{-\left(\frac{\psi(\mathbf{x})}{d}\right)^2} \right) \quad (5.8)$$

La fonction de relaxation spatiale de l'équation (5.8) est nulle sur les lieux de la forme *a priori* représentée par ψ et tend asymptotiquement vers la constante λ . Le paramètre d contrôle la taille de l'espace de relaxation. Cette formulation n'a pas de justification théorique et d'autres expressions plus simples pourraient satisfaire les mêmes critères de relaxation. Nous vérifierons expérimentalement en section 5.5.5.2 que la dérivabilité et le point stationnaire de λ_{space} en $\psi(\mathbf{x}) = 0$ améliorent l'insensibilité aux minima locaux.

L'existence de l'espace de relaxation paramétré par d est destiné à empêcher le contour actif d'être sujet aux minima locaux de seconde espèce. Cependant, la liberté accordée au contour actif dans l'espace de relaxation ne permettra pas d'éviter les minima de première espèce. Pour résoudre ce problème, nous proposons de réduire dans un second temps (à partir de t_1) l'espace de relaxation. Alors que cet espace diminue, l'efficacité de la contrainte de forme est rétablie uniformément sur le domaine de l'image. L'idée est donc d'atteindre une segmentation grossière de l'objet de l'image avec relaxation spatiale ($t < t_1$), puis de rétablir la contrainte de forme *a priori* uniforme dans un second temps pour vaincre les altérations de l'image ($t \geq t_1$). Le poids de contrainte de forme devient ainsi une fonction de l'espace et du temps artificiel du processus itératif. La décroissance de l'espace de relaxation se fait par le remplacement de la constante d par une fonction linéaire par morceaux et décroissante en fonction du temps :

$$d(t) = \begin{cases} d_0 & \text{si } t < t_1 \\ (\varepsilon - d_0) \frac{t-t_1}{t_2-t_1} + d_0 & \text{si } t_1 \leq t < t_2 \\ \varepsilon & \text{si } t_2 \leq t \end{cases} \quad (5.9)$$

où $(d_0, \varepsilon, t_1, t_2) \in \mathbb{R}^{++}$ avec $d_0 > \varepsilon$ et $\varepsilon \ll 1$. d_0 est une constante déterminant la taille initiale de l'espace de relaxation et t_2 règle la vitesse de décroissance. Nous verrons en section 5.5.5.2 qu'une décroissance temporelle plus simple de type échelon de Heaviside est moins efficace à cause de l'inertie due à l'estimation des paramètres de T_{sim} .

Finalement, nous renforçons la relaxation spatiale et le rétablissement de la contrainte uniforme en remplaçant le facteur λ de l'équation (5.8) par une fonction croissante du temps $\lambda_a(t)$ qui contrôlera l'amplitude globale de la fonction spatio-temporelle du poids de contrainte. Durant la phase de relaxation ($t < t_1$), la valeur est minimale pour accorder plus de flexibilité au contour actif. A partir de t_1 , l'amplitude augmente pour atteindre un maximum qui renforcera le rétablissement de la contrainte afin de surmonter les altérations de l'image. La fonction spatio-temporelle du poids de la contrainte de forme est finalement :

$$\lambda_{flex}(\psi(\mathbf{x}), t) = \lambda_a(t) \left(1 - e^{-\left(\frac{\psi(\mathbf{x})}{d(t)}\right)^2} \right) \quad (5.10)$$

où $\lambda_a(t)$ varie entre λ_{min} et λ_{max} avec un profil opposé à celui de l'équation (5.9). La figure 5.2 illustre à trois dimensions la fonction spatio-temporelle $\lambda_{flex}(\psi(\mathbf{x}), t)$.

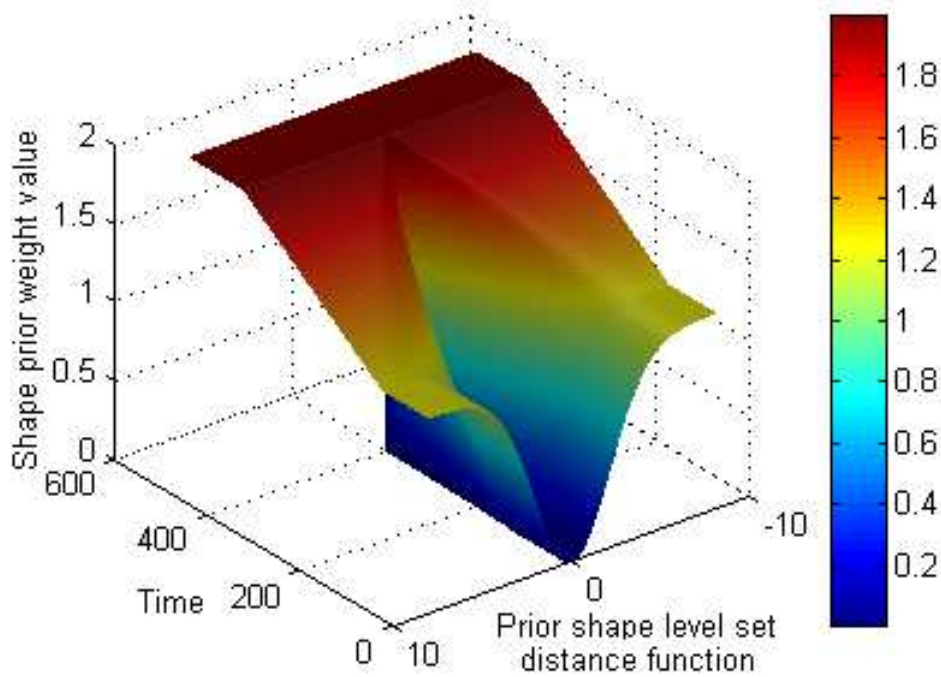


FIG. 5.2 – Variation spatio-temporelle de la fonction pondérant la contrainte de forme. $d_0 = 3$, $t_1 = 100$, $t_2 = 400$, $\lambda_{min} = 1.25$ et $\lambda_{max} = 2$.

5.5 Applications et analyse des résultats expérimentaux

Cette section présente les résultats obtenus avec les trois innovations proposées, à savoir, l'estimation des paramètres de la transformation globale par l'algorithme du simplexe, la fusion de données exogènes dans le terme d'attache aux données des contours actifs et enfin la variation spatio-temporelle de la contrainte de forme. Nous comparerons ces résultats aux modèles plus classiques et examinerons dans quelle mesure les solutions que nous avons proposées permettent de s'affranchir du problème des minima locaux. Enfin, nous ne nous limiterons pas au domaine de la

télédétection et illustrerons des résultats obtenus avec des images utilisées en Vision par Ordinateur et en imagerie médicale. Ces images partagent la même caractéristique que les images satellitaires en milieu urbain : un manque de discrimination globale de l'objet à extraire par rapport au fond de l'image.

5.5.1 Protocole expérimental

Dans toutes les expériences exposées ci-après, le contour actif et la forme *a priori* ont les caractéristiques suivantes :

1. La forme *a priori* représentée par l'ensemble de niveaux $\psi_0(\mathbf{x})$ correspond *exactement* aux frontières de l'objet à segmenter/recaler dans l'image.
2. Le contour actif initial représenté par l'ensemble de niveaux $\phi(\mathbf{x}, t = 0)$ est similaire à la forme *a priori* $\psi_0(\mathbf{x})$. Il est ensuite transformé par une similitude plane directe T_{sim} dont les paramètres sont $\xi_0 = (s_0, \theta_0, \mu_{x,0}, \mu_{y,0})$. Le but d'une telle transformation est de créer un écart géométrique entre le contour actif initial et l'objet cible de l'image afin d'examiner le pouvoir de recalage de la méthode.

Dans certains cas, nous illustrons l'évolution du contour actif de l'état initial à la convergence finale sous la forme d'une séquence allant soit de gauche à droite soit du haut vers le bas.

5.5.2 Comparaison simplex vs. descente de gradient pour l'estimation des paramètres de T_{sim}

Pour comparer les performances des deux algorithmes, considérons l'expérience suivante : une forme *a priori* ψ_0 est choisie parmi les deux exemples de la figure 5.3. Le contour actif représenté par ϕ et transformé par T_{sim} de paramètres ξ_0 n'évoluera pas au cours du temps (aucune force externe due à l'image n'est employée, nous travaillons uniquement dans l'espace des formes).

L'objectif est de constater si les deux méthodes parviennent à estimer ξ_0 par la minimisation de l'énergie donnée en équation (5.3). Le nombre d'itérations pour atteindre la convergence ainsi que

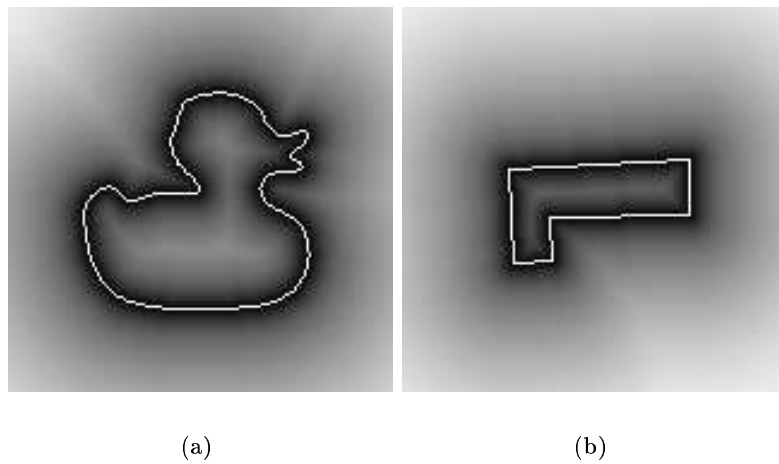


FIG. 5.3 – Formes *a priori* (en blanc, $\{\mathbf{x}/\psi_0(\mathbf{x}) = 0\}$) et leurs ensembles de niveaux associés (niveaux de gris, $|\psi_0(\mathbf{x})|$). (a) jouet (b) bâtiment en L

les temps de calcul seront comparés. Nous menons plusieurs fois la même expérience avec différentes initialisations $(s_0, \theta_0, \mu_{x,0}, \mu_{y,0})$:

1. Translation seule : $\xi_{0,1} = (1, 0, 10, -5)$.
2. Rotation seule : $\xi_{0,2} = (1, \pi/3, 0, 0)$.
3. Échelle seule : $\xi_{0,3} = (1.2, 0, 0, 0)$.
4. Similitude plane directe : $\xi_{0,4} = (1.2, \pi/3, 10, -5)$.

Dans les expériences menées, nous avons privilégié la robustesse des optimisations en travaillant avec une bande étroite de 11 pixels. Les temps de calcul seront élevés, mais le but de ces expériences est la comparaison des deux algorithmes et non leurs performances absolues. Le simplex et la descente de gradient auront convergé lorsqu'ils vérifient le critère : $|J_{shape}^{n+1} - J_{shape}^n| < 10^{-4}$. Les deux algorithmes ont été paramétrés de la façon suivante :

1. Descentes de gradient : $\Delta t_s = 5.10^{-7}$, $\Delta t_\theta = 2, 5.10^{-7}$, $\Delta t_{\mu_x} = 2, 5.10^{-3}$ et $\Delta t_{\mu_y} = 2, 5.10^{-3}$
2. Simplex : $\delta_s = 0.25$, $\delta_\theta = \pi$, $\delta_{\mu_x} = 10$ et $\delta_{\mu_y} = 10$

Les tables 5.1 et 5.2 illustrent les résultats obtenus avec la descente de gradient et le simplex respectivement. Une première constatation est la robustesse du simplex qui estime correctement ξ_0 dans toutes les situations. Le gradient est plus sensible aux minima locaux de la fonctionnelle d'énergie J_{shape} et échoue ainsi dans trois cas à accomplir l'estimation correcte des paramètres. La deuxième remarque concerne les temps de calcul. Le simplex se révèle extrêmement plus efficace que la descente de gradient. Le nombre d'itérations nécessaires pour atteindre la convergence est moindre, avec, dans ces expériences, une rapidité supérieure variant d'un ratio 5 à 68. La principale raison de la lenteur de la descente de gradient est la stabilité numérique. Un pas temporel petit doit être utilisé pour chaque descente afin de satisfaire les conditions de stabilité, au risque de voir le système instable ou diverger. Il est néanmoins important de remarquer qu'un réglage légèrement différent des pas temporels est susceptible de fournir des convergences un peu plus rapides ou plus longues. Les résultats du tableau 5.1 doivent ainsi être nuancés mais donnent un ordre de grandeur des temps de convergence.

Forme	Expérience	Erreur absolue			Itérations	Temps (s)
		$\Delta\theta$ en rad	Δs	$\Delta\mu$ en pixels		
Jouet	$\xi_{0,1}$	0	0	0	440	15
	$\xi_{0,2}$	0.8	1.05	$(-14.3, -3.17)$	1680	58
	$\xi_{0,3}$	0	0	0	550	21
	$\xi_{0,4}$	0.8	1.05	$(-14.3, -3.17)$	2780	98
Bâtiment en L	$\xi_{0,1}$	0	0	0	420	11
	$\xi_{0,2}$	0	0	0	2750	68
	$\xi_{0,3}$	0	0	0	330	8
	$\xi_{0,4}$	0.3	1.51	$(16.47, 12.54)$	1170	30

TAB. 5.1 – Erreur absolue entre ξ_0 et son estimée $\hat{\xi}_0$ par l'algorithme de descente de gradient et temps de calcul.

Forme	Expérience	Erreur absolue			Itérations	Temps (s)
		$\Delta\theta$ en rad	Δs	$\Delta\mu$ en pixels		
Jouet	$\xi_{0,1}$	0	0	0	338	3
	$\xi_{0,2}$	0	0	0	251	3
	$\xi_{0,3}$	0	0	0	204	3
	$\xi_{0,4}$	0	0	0	390	3
Bâtiment en L	$\xi_{0,1}$	0	0	0	201	<1
	$\xi_{0,2}$	0	0	0	312	1
	$\xi_{0,3}$	0	0	0	216	1
	$\xi_{0,4}$	0	0	0	276	1

TAB. 5.2 – Erreur absolue entre ξ_0 et son estimée $\hat{\xi}_0$ par l'algorithme du simplex et temps de calcul.

5.5.3 Contribution de la connaissance de forme *a priori* pour la segmentation

Dans cette section, nous comparons les performances des contours actifs choisis en section 4.3.2 avec et sans incorporation de forme *a priori*. Le but est d'illustrer la nécessité d'utiliser la connaissance de forme afin de segmenter correctement un objet dont la représentation est altérée dans une image.

La figure 5.4 présente les résultats obtenus avec le modèle de Chan et Vese et le modèle Bayésien avec un bâtiment répondant au critère d'homogénéité. Dans cette expérience où la contrainte de forme n'est pas utilisée, le contour actif initial n'a pas subi de transformation ($\xi_0 = (1, 0, 0, 0)$) et se trouve ainsi exactement superposé au bâtiment en forme de U de l'image. En dépit d'une initialisation du contour actif confondue avec la solution désirée, celui-ci échoue à segmenter le bâtiment considéré. On constate que dans ce cas, le contour actif a un comportement attendu sur les zones de faible contraste de l'image : le contour "fuit" et se propage alors dans toute l'image, englobant des zones ayant les mêmes propriétés statistiques. Ces résultats illustrent aussi une segmentation plus sélective du modèle d'attache aux données de Chan et Vese du fait de sa restriction au partitionnement de zones constantes de l'image. Le modèle Bayésien, plus permissif, segmente une zone homogène et plus étendue de l'image. Les expériences présentées en figure 5.5 sont menées avec les mêmes conditions initiales et les mêmes modèles que ceux de la figure 5.4. Cependant, l'incorporation de forme *a priori* est prise en compte. Nous constatons que dans ce cas, les résultats sont similaires quel que soit le modèle utilisé : une segmentation satisfaisante est obtenue et les artefacts dus aux zones de faible contraste sont surmontés. Il s'est avéré nécessaire d'opter pour un poids de contrainte de forme λ différent selon le modèle d'attache aux données utilisé. Le choix empirique de λ est un compromis entre l'information contenue dans l'image et la contrainte de forme *a priori*. Si λ est trop faible, le contour actif est libre d'être influencé par le terme d'attache aux données qui est susceptible d'être corrompu par le bruit, les occlusions ou autres artefacts urbains. A l'inverse, un poids trop élevé reviendrait à minimiser la contribution de l'image qui conduit le contour vers le bâtiment que l'on cherche à segmenter. Le contour actif perdrait ainsi en flexibilité et serait plus enclin à atteindre des minima locaux de la fonctionnelle à minimiser (nous examinerons ce point en section 5.5.4). Si un poids $\lambda = 1$ est convenable pour le modèle de Chan et Vese, il est insuffisant pour le modèle Bayésien. En effet, ce dernier modèle diffère de celui de Chan et Vese par la normalisation des termes relatifs à l'image par la variance à l'intérieur et l'extérieur du contour actif. Alors que l'extérieur du contour actif dans l'image est plus "désordonné" que le bâtiment homogène que l'on cherche à segmenter, on a de façon générale : $\sigma_{in}^2(t) < \sigma_{out}^2(t)$. Lorsque le contour converge vers la cible homogène, on a $\sigma_{in}^2(t) \ll 1$: la division par ce terme accroît alors de façon drastique le terme d'attache aux données à l'intérieur de $\mathbf{C}(t)$. Pour contrecarrer cet effet, nous choisissons

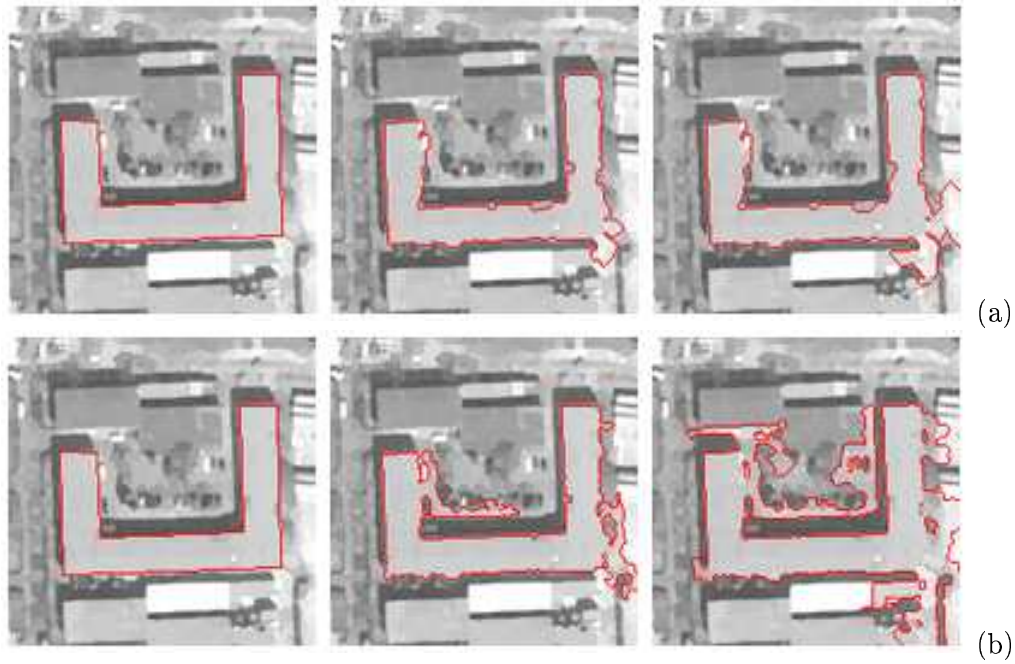


FIG. 5.4 – Segmentation sans insertion de forme *a priori* et $\xi_0 = (1, 0, 0, 0)$. (a) modèle de Chan et Vese (b) modèle Bayésien.

une valeur plus élevée de λ ($\lambda = 10$).

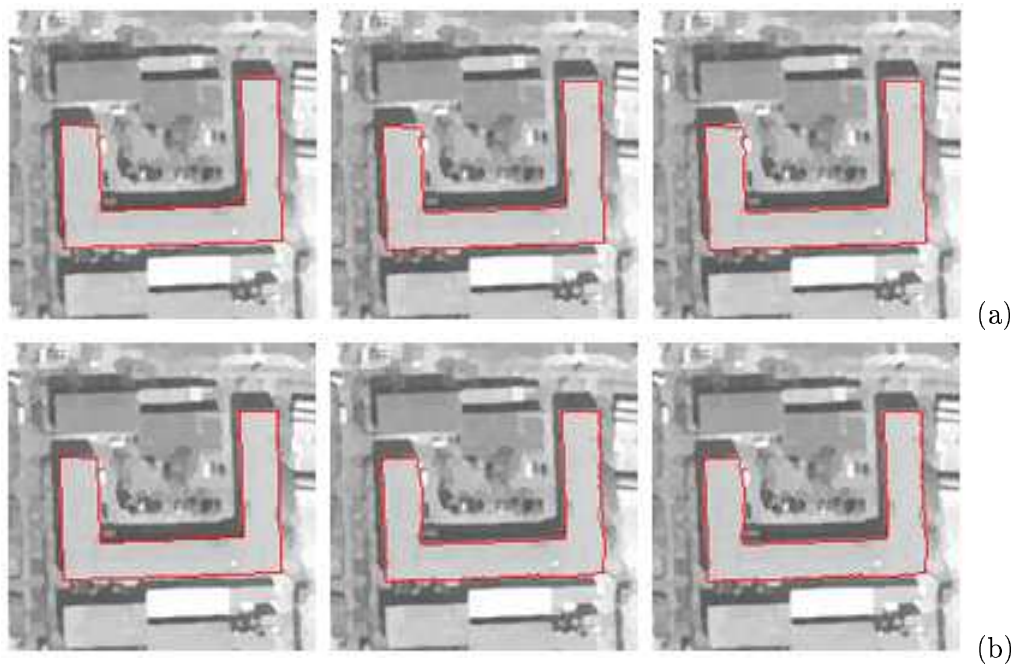


FIG. 5.5 – Segmentation avec insertion de forme *a priori* : (a) modèle de Chan et Vese et $\lambda = 1$ (b) modèle Bayésien et $\lambda = 10$.

La nécessité d'utiliser la connaissance de forme *a priori* est à nouveau illustrée en figure 5.6 avec les contours actifs basés sur l'information diffusée des gradients de l'image (GVF). A la différence des expériences des figures 5.4 et 5.5, le contour actif initial n'est pas exactement superposé à

l'objet à recaler dans l'image : le contour initial a subi une transformation de paramètres $\xi_0 = (s_0 = 1.1, \theta_0 = 0.1, \mu_{x,0} = -4, \mu_{y,0} = 1)$. La figure 5.6.a montre que la segmentation du bâtiment échoue en l'absence de contrainte de forme. Comme nous l'avons déjà remarqué avec les contours actifs basés région, le modèle déformable sans contrainte est trop libre et devient sensible aux artefacts des images urbaines. Dans le cas de la figure 5.6.a, la multitude de détails produit de nombreuses zones de gradient élevé qui attirent le contour actif et l'éloignent des bords du bâtiment. La figure 5.6.b montre l'amélioration apportée par la contrainte de forme : le bâtiment de l'image est correctement segmenté malgré une initialisation éloignée de la cible. L'invariance par similitude plane directe de la contrainte de forme permet de s'affranchir des erreurs initiales de superposition grâce à l'estimation dynamique des paramètres de T_{sim} .

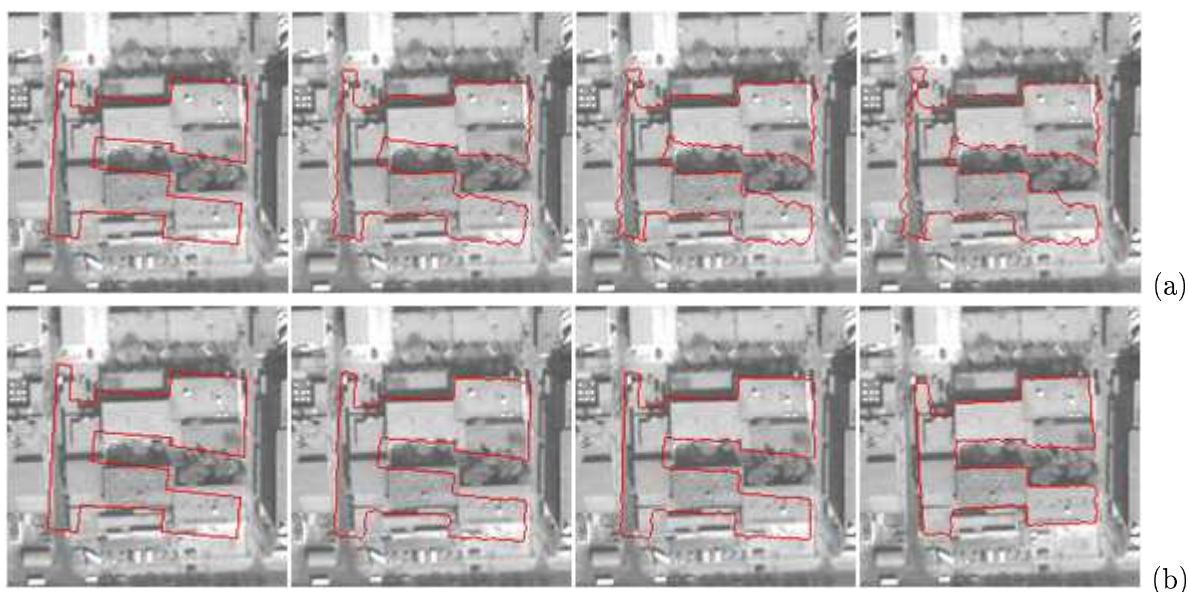


FIG. 5.6 – Segmentation avec un contour actif basé sur l'information de contour (GVF) et $\xi_0 = (s_0 = 1.1, \theta_0 = 0.1, \mu_{x,0} = -4, \mu_{y,0} = 1)$: (a) résultat sans contrainte de forme *a priori*. (b) résultat avec contrainte de forme *a priori* ($\lambda = 5$).

5.5.4 Sensibilité à l'initialisation des contours actifs contraints

L'évolution des contours actifs minimise une fonctionnelle d'énergie par descente de gradient. Cette optimisation n'est pas globale mais locale. Ainsi, toutes les configurations possibles du contour dans l'image ne seront pas essayées pour minimiser la fonctionnelle d'énergie. Au contraire, le contour actif devra être initialement proche de la solution finale pour y converger. Ainsi, un recouvrement non nul entre le contour initial et la cible dans l'image est indispensable. Dans cette section, nous examinons la sensibilité des contours actifs avec contrainte de forme vis-à-vis de l'initialisation afin d'évaluer la robustesse des algorithmes et leur pouvoir à résoudre le problème de la mise en correspondance carte-image. Néanmoins, nous n'examinons qu'une seule classe d'erreur de mise en correspondance initiale : celle modélisable par une similitude plane directe.

5.5.4.1 Comparaison des modèles basés sur l'information de région

Les figures 5.7 à 5.12 illustrent la sensibilité des modèles de Chan et Vese et du modèle Bayésien par rapport à l'initialisation du contour actif. La figure 5.7 montre que le modèle de Chan et Vese n'est pas sensible à une faible translation sur les trois exemples illustrés. Les figures 5.8.a et 5.8.c montrent que si la mise en correspondance des bâtiments rectangulaires ne semble pas poser de

problème avec une rotation initiale de 0.3 rad, on constate en revanche que celle du bâtiment en U échoue (figure 5.8.b). Il existe deux raisons à cela :

- La première est relative aux images que nous utilisons. En milieu urbain dense, un bâtiment considéré ne se démarque pas singulièrement de son environnement. Ainsi, si l’initialisation est trop éloignée, une convergence vers l’objet recherché peut être très délicate. Cette convergence est rendue encore plus hasardeuse pour le modèle de Chan et Vese où la seule quantité statistique examinée est la moyenne. Il est fortement possible qu’au cours du processus itératif, $c_{in}(\phi(\mathbf{x}, t)) \simeq c_{out}(\phi(\mathbf{x}, t))$ ce qui aura pour effet d’immobiliser le contour actif, empêchant la mise en correspondance (minimum local).
- La deuxième raison est inhérente au terme de contrainte de forme. Il est possible que le terme d’attache aux données fasse tendre le contour actif vers un état qui n’appartient pas à l’espace des formes autorisées par le terme énergétique de contrainte. Ainsi, même si l’objet représenté dans l’image est la solution minimisant la fonctionnelle d’énergie, il est probable qu’il en coûte plus pour le terme de contrainte de forme de s’en approcher que le gain enregistré par l’énergie d’attache aux données. On atteint alors un *statu quo* entre ces deux forces antagonistes, faisant converger le contour actif vers un minimum local de la fonctionnelle d’énergie.

Le modèle de Chan et Vese semble très sensible à une variation du facteur d’échelle. Cela est une manifestation probante du manque de discrimination radiométrique du bâtiment à recalcr dans l’image au sens de la moyenne (figures 5.9.a et 5.9.b).

Les résultats obtenus avec le modèle Bayésien en figures 5.10.b et 5.10.c sont similaires à ceux réalisés avec le modèle de Chan et Vese concernant la sensibilité vis-à-vis de la translation. En revanche, on constate que la segmentation du premier bâtiment rectangulaire échoue (figure 5.10.a). Ceci est inhérent au modèle Bayésien, plus permissif et qui autorise la segmentation d’objets d’intensité non uniforme. Cette flexibilité du modèle associée à une contrainte de forme trop élevée ($\lambda = 10$) permet d’expliquer le résultat erroné obtenu. Nous verrons en section 5.4 comment résoudre ce problème avec un poids λ variable. Quant à la rotation (figure 5.11), on constate que la mise en correspondance du bâtiment en U a réussi (figure 5.11.b), tout comme celle des bâtiments rectangulaires, alors que le modèle de Chan et Vese avait échoué (figure 5.8.b). Comme nous l’évoquions auparavant, la normalisation par la variance de l’image à l’intérieur du contour actif permet de rendre prépondérant le terme d’attache aux données pour vaincre le terme de contrainte de forme qui avait tendance à créer des minima locaux. A l’instar du modèle de Chan et Vese, le test de sensibilité vis-à-vis du facteur d’échelle échoue pour les expériences des figures 5.12.a et 5.12.b. Ceci est dû à une mauvaise initialisation et à un manque de discrimination du bâtiment dans l’image. A contrario, le bâtiment clair de la troisième expérience étant singulièrement différent du reste de la scène, la segmentation réussit quel que soit le modèle employé (figure 5.12.c).

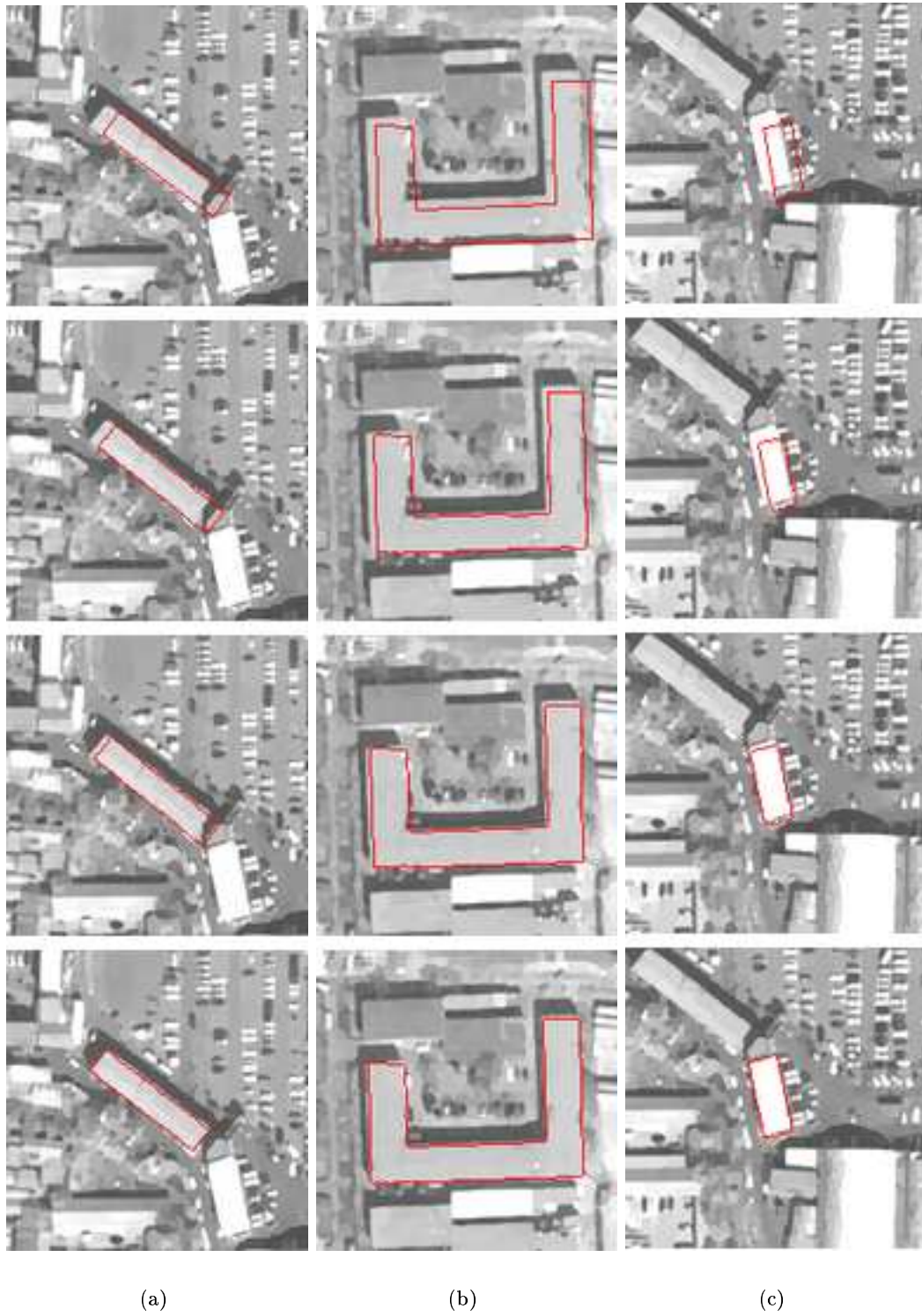


FIG. 5.7 – Sensibilité à la translation : segmentation avec le modèle de Chan et Vese et la contrainte de forme. Le contour actif initial des expériences (a), (b) et (c) a été transformé par une similitude plane directe de paramètres $\xi_0 = (s_0 = 1.0, \theta_0 = 0.0, \mu_{x,0} = 5, \mu_{y,0} = 5)$.

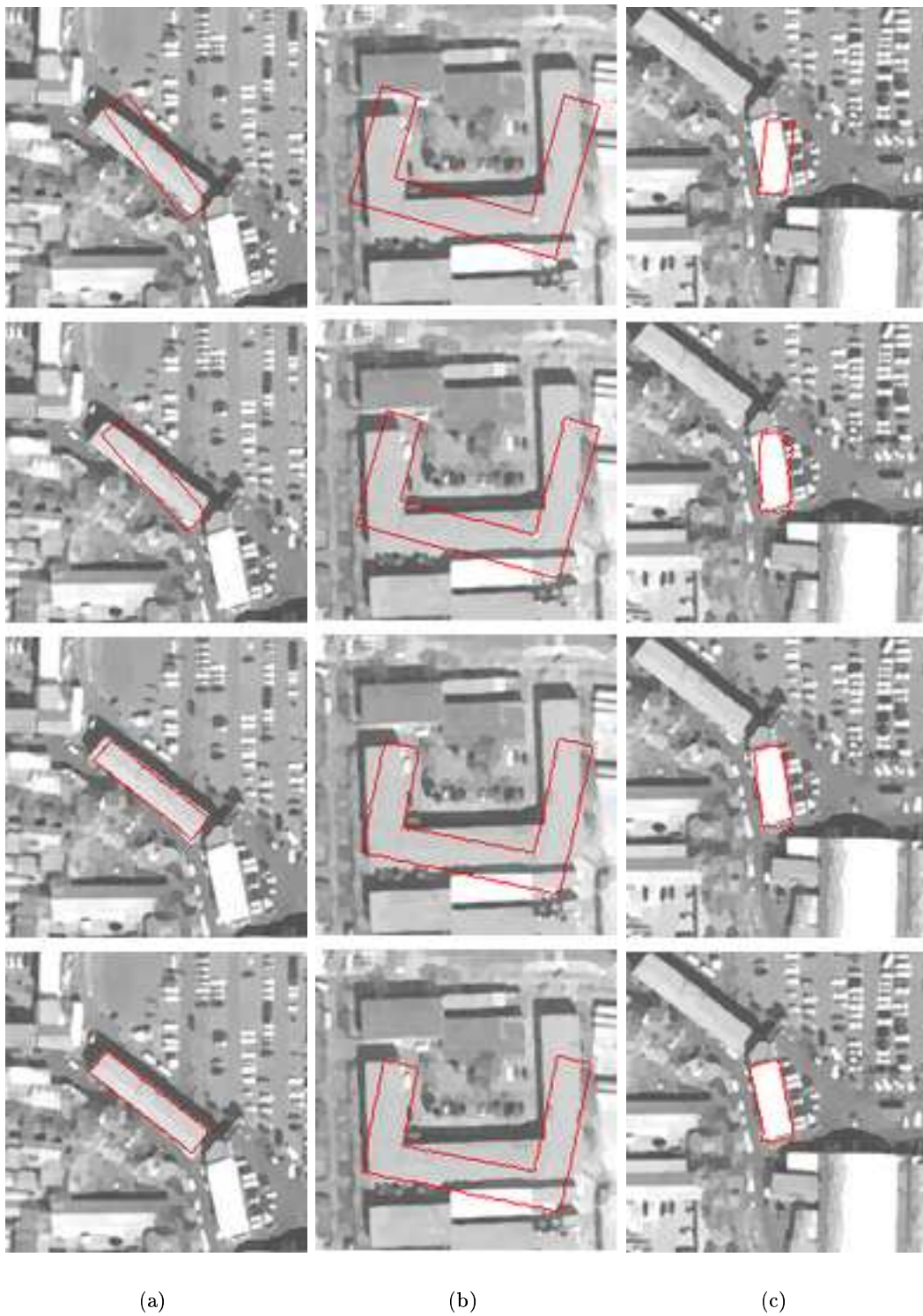


FIG. 5.8 – Sensibilité à la rotation : segmentation avec le modèle de Chan et Vese et la contrainte de forme. Le contour actif initial des expériences (a), (b) et (c) a été transformé par une similitude plane directe de paramètres $\xi_0 = (s_0 = 1.0, \theta_0 = 0.3, \mu_{x,0} = 0, \mu_{y,0} = 0)$.

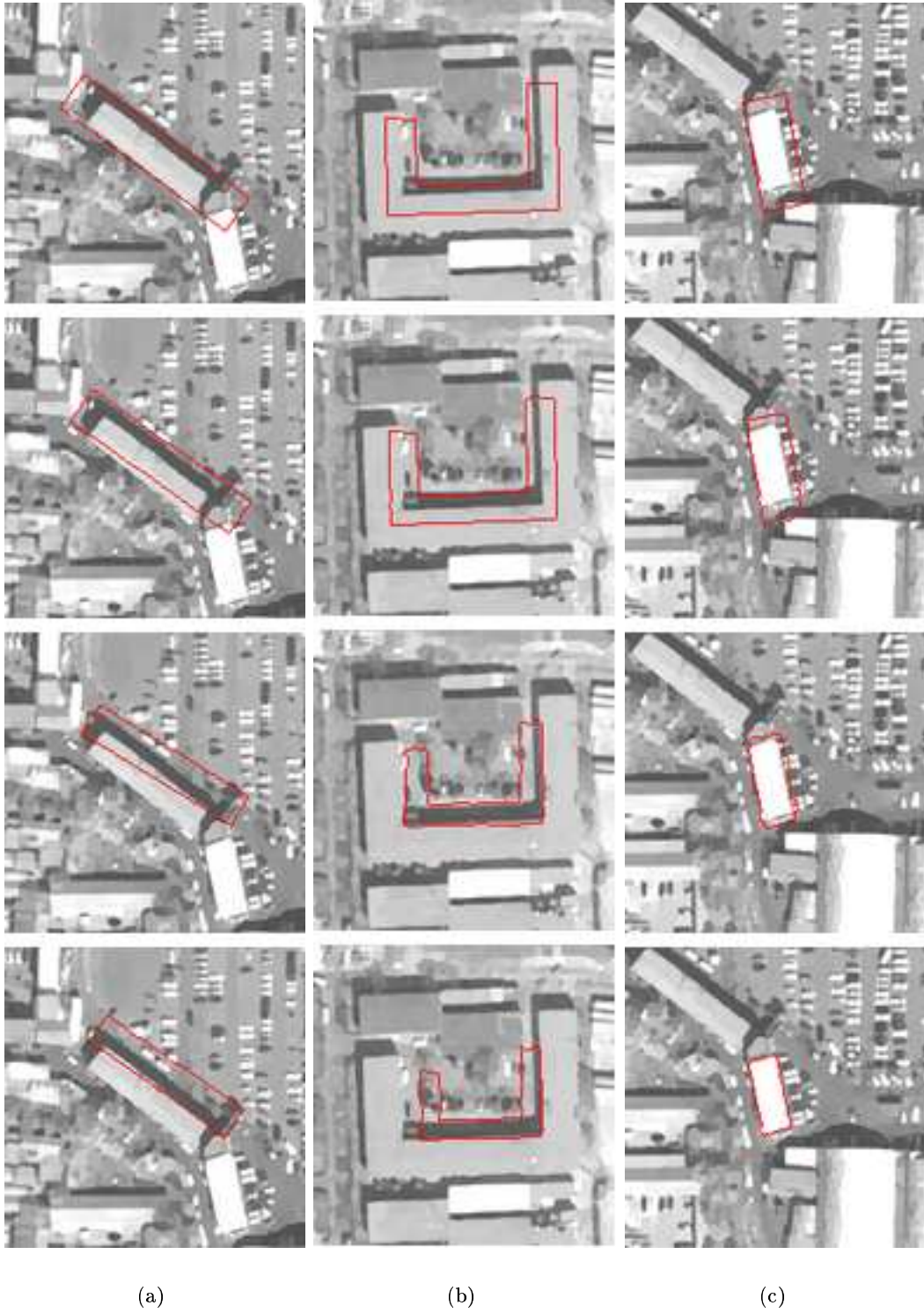


FIG. 5.9 – Sensibilité au facteur d'échelle : segmentation avec le modèle de Chan et Vese et la contrainte de forme. (a) $\xi_0 = (s_0 = 1.5, \theta_0 = 0.0, \mu_{x,0} = 0, \mu_{y,0} = 0)$; (b) $\xi_0 = (s_0 = 0.8, \theta_0 = 0.0, \mu_{x,0} = 0, \mu_{y,0} = 0)$; (c) $\xi_0 = (s_0 = 1.5, \theta_0 = 0.0, \mu_{x,0} = 0, \mu_{y,0} = 0)$.

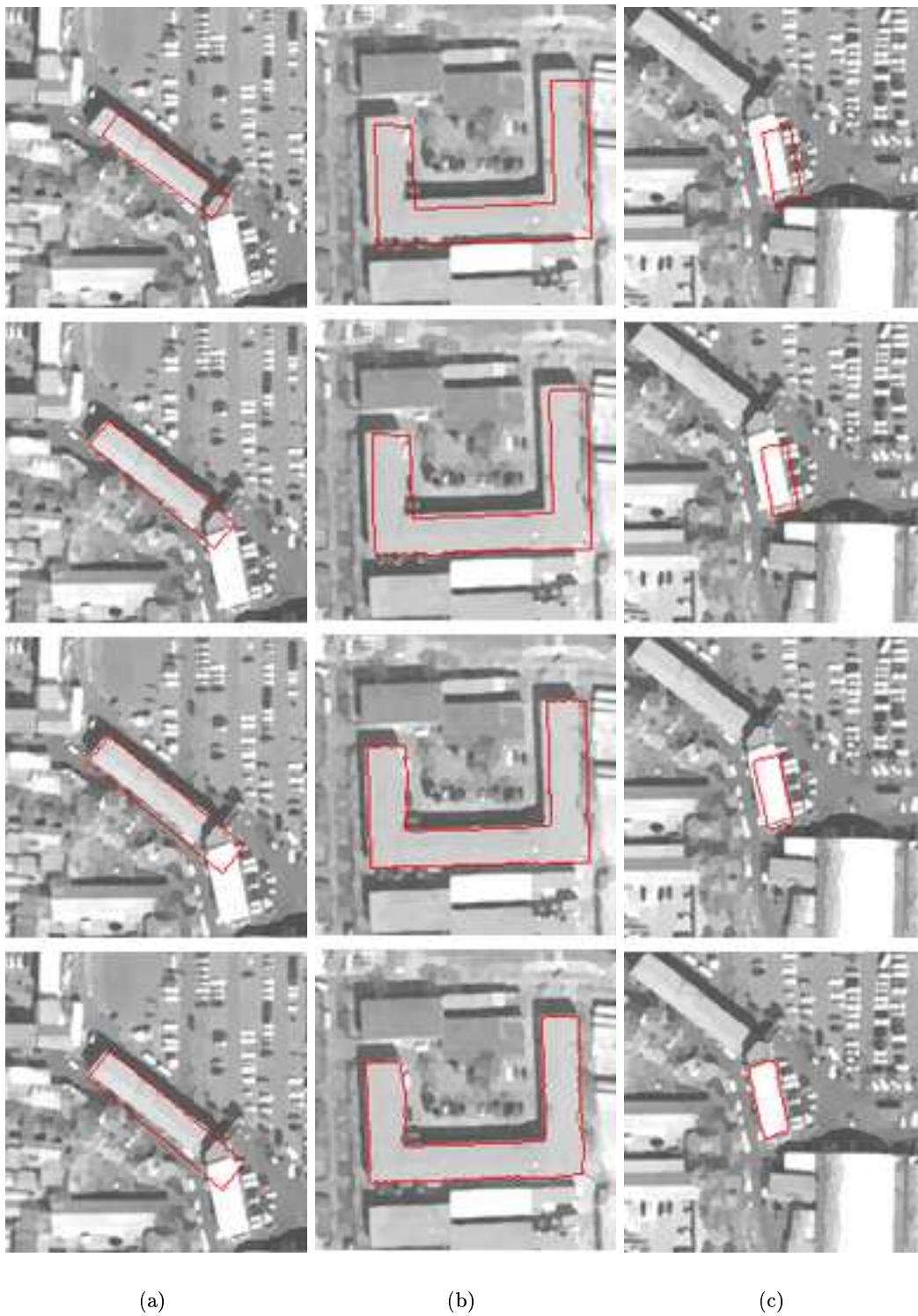


FIG. 5.10 – Sensibilité à la translation : segmentation avec le modèle Bayésien et la contrainte de forme. Le contour actif initial des expériences (a), (b) et (c) a été transformé par une similitude plane directe de paramètres $\xi_0 = (s_0 = 1.0, \theta_0 = 0.0, \mu_{x,0} = 5, \mu_{y,0} = 5)$.

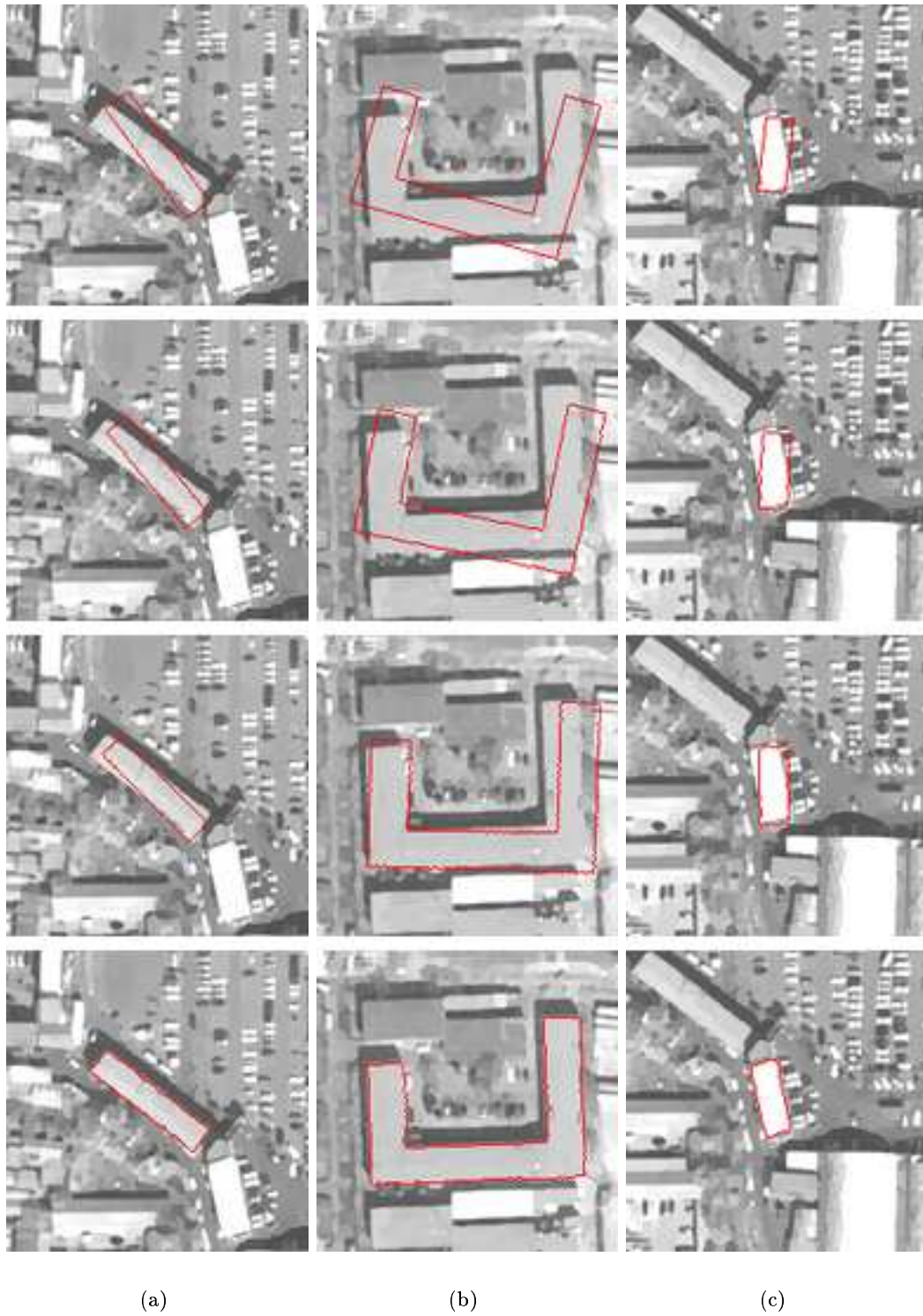


FIG. 5.11 – Sensibilité à la rotation : segmentation avec le modèle Bayésien et la contrainte de forme. Le contour actif initial des expériences (a), (b) et (c) a été transformé par une similitude plane directe de paramètres $\xi_0 = (s_0 = 1.0, \theta_0 = 0.3, \mu_{x,0} = 5, \mu_{y,0} = 5)$.

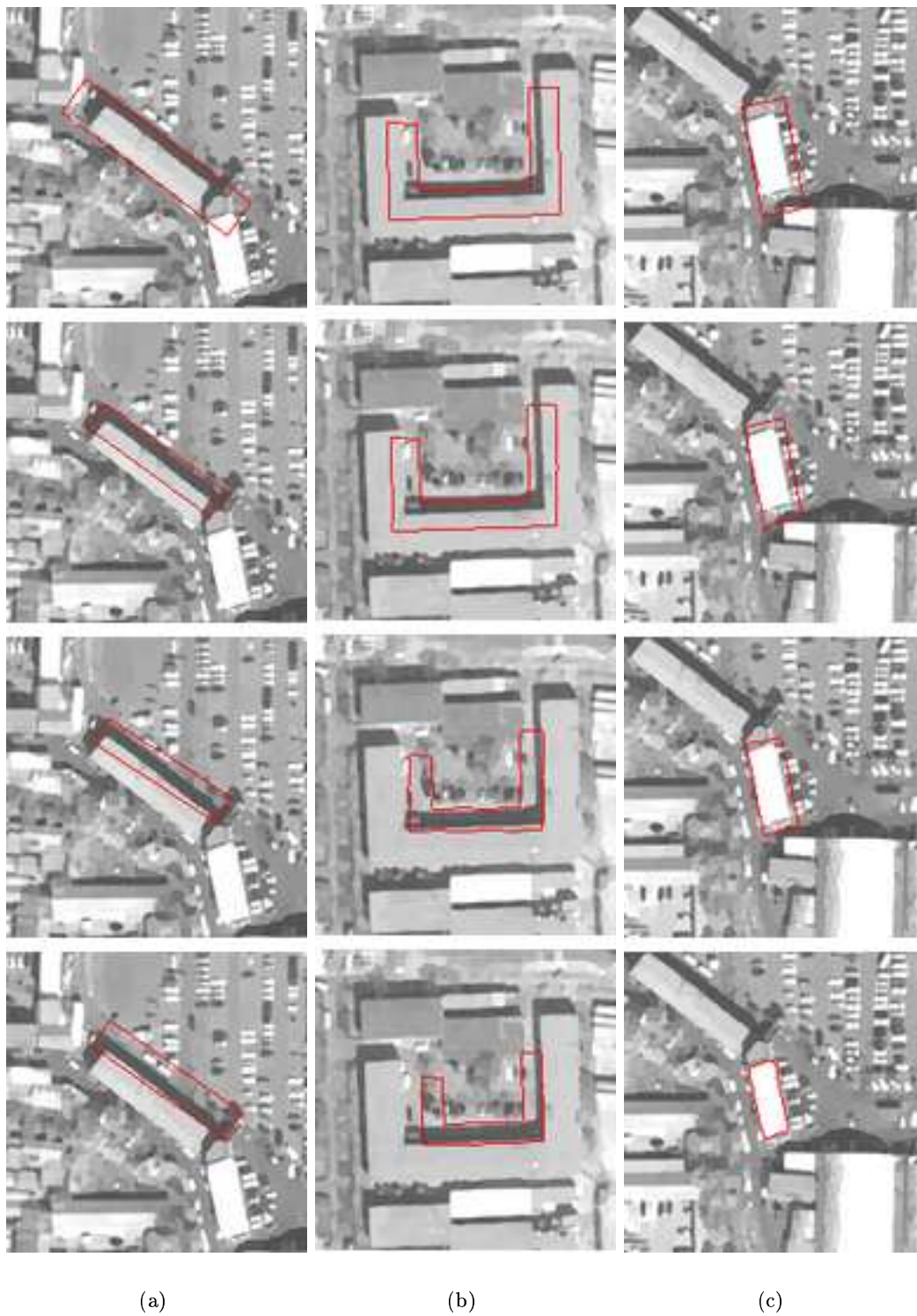


FIG. 5.12 – Sensibilité au facteur d'échelle : segmentation avec le modèle Bayésien et la contrainte de forme.

(a) $\xi_0 = (s_0 = 1.5, \theta_0 = 0.0, \mu_{x,0} = 0, \mu_{y,0} = 0)$

(b) $\xi_0 = (s_0 = 0.8, \theta_0 = 0.0, \mu_{x,0} = 0, \mu_{y,0} = 0)$

(c) $\xi_0 = (s_0 = 1.5, \theta_0 = 0.0, \mu_{x,0} = 0, \mu_{y,0} = 0)$.

5.5.4.2 Modèle basé sur l'information de frontières

La figure 5.13 montre le résultat du recalage avec le contour actif basé sur le GVF de l'image. Les recalages illustrés sur cette figure sont satisfaisants. Il est cependant à noter que l'erreur d'initialisation est faible. Avec une initialisation plus éloignée, la segmentation contrainte échoue (figure 5.14). Ces expériences démontrent que même avec la diffusion spatiale de l'information de gradient, le contour actif reste sensible à l'initialisation. La raison est la profusion de détails entourant le bâtiment et qui piègent facilement le contour actif dans un minimum local de la fonctionnelle d'énergie.

5.5.5 Solutions pour une moindre sensibilité à l'initialisation et aux minima locaux

Les expériences de la section précédente ont montré que les contours actifs utilisés étaient sensibles à une initialisation éloignée par rapport à l'objet à segmenter dans l'image. Nous justifions cette sensibilité par le manque de discrimination de l'objet d'intérêt par rapport au reste de l'image et par la présence de la contrainte de forme, encline à faire converger le contour vers des minima locaux. Dans cette section, nous présentons des résultats relatifs à la fusion du MNS avec l'image satellitaire et la contrainte de forme spatio-temporelle. Nous examinerons expérimentalement dans quelle mesure ces deux solutions résolvent le problème des minima locaux qui empêchent un recalage efficace.



FIG. 5.13 – Sensibilité à l'initialisation : segmentation avec un contour actif basé sur l'information de contour (GVF) et avec contrainte de forme.

(a) $\xi_0 = (s_0 = 1.1, \theta_0 = 0.1, \mu_{x,0} = -4, \mu_{y,0} = 1)$

(b) $\xi_0 = (s_0 = 1.0, \theta_0 = -0.1, \mu_{x,0} = -1, \mu_{y,0} = -2)$

(c) $\xi_0 = (s_0 = 1.0, \theta_0 = 0.1, \mu_{x,0} = 0, \mu_{y,0} = 2)$.

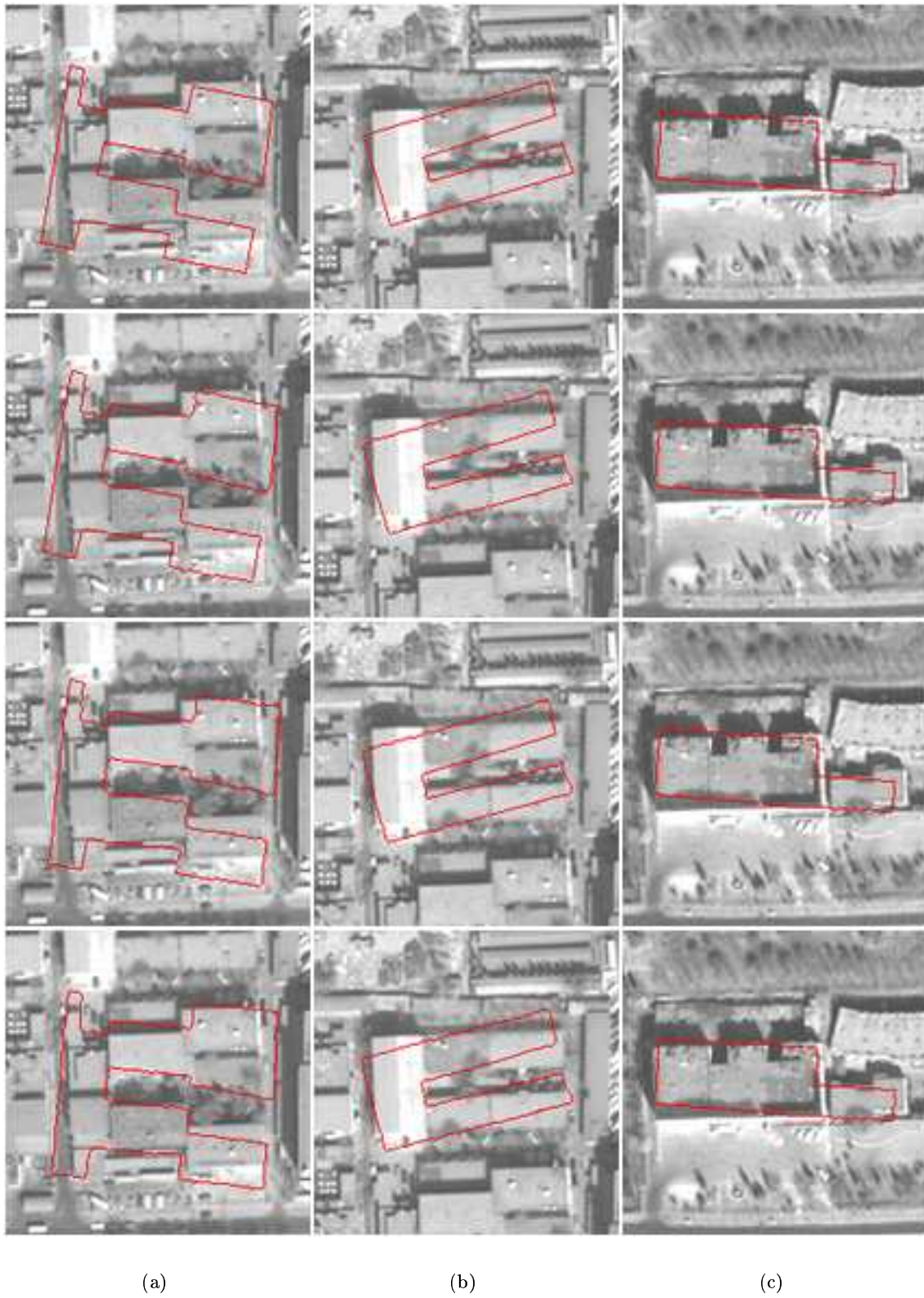


FIG. 5.14 – Sensibilité à une initialisation éloignée : segmentation avec un contour actif basé sur l'information de contour (GVF) et avec contrainte de forme.

(a) $\xi_0 = (s_0 = 1.1, \theta_0 = 0.25, \mu_{x,0} = -4, \mu_{y,0} = 1)$

(b) $\xi_0 = (s_0 = 1.0, \theta_0 = -0.25, \mu_{x,0} = -2, \mu_{y,0} = -3)$

(c) $\xi_0 = (s_0 = 1.0, \theta_0 = 0.1, \mu_{x,0} = -2, \mu_{y,0} = 2)$.

5.5.5.1 Résultats avec fusion de données exogènes dans le terme d'attache aux données

La figure 5.15 illustre le résultat obtenu avec le bâtiment en U et les modèles proposés en équations (5.5) et (5.6). En dépit d'une initialisation volontairement éloignée du bâtiment de l'image, le recalage réussit. Ceci montre que la bonne discrimination du bâtiment dans le MNS parvient à conduire le contour actif vers la cible dans l'image. A la fin de la convergence, la contribution du MNS est relaxée pour donner plus de poids à l'image dans laquelle les bords du bâtiment sont mieux définis. La figure 5.16 illustre la même expérience que la figure 5.15 mais sans intégration du MNS. On constate que dans ce cas, le contour actif ne parvient pas à segmenter le bâtiment.

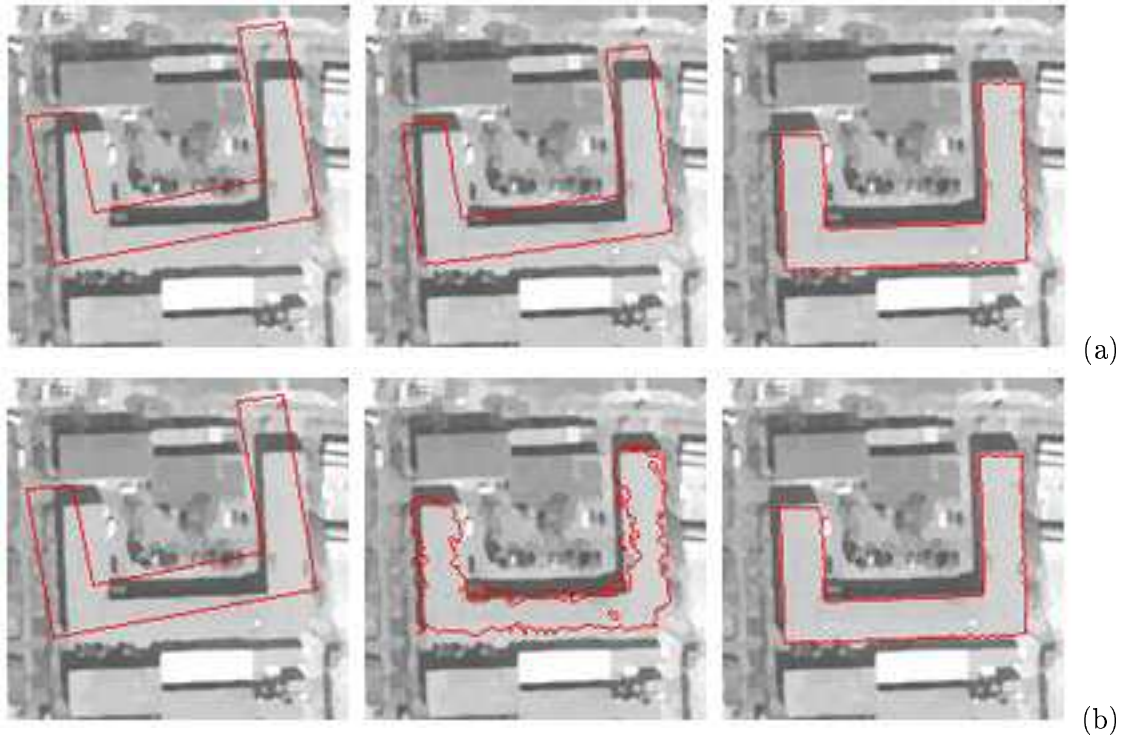


FIG. 5.15 – Recalage réussi avec une initialisation éloignée et l'insertion du MNS exogène dans la fonctionnelle d'énergie. (a) modèle de Chan et Vese (b) modèle Bayésien.



FIG. 5.16 – Échec du recalage avec une initialisation éloignée et le modèle de Chan et Vese (pas d'incorporation du MNS).

La comparaison des figures 5.17 et 5.18 illustre l'efficacité de la méthode proposée. Dans l'expérience de la figure 5.17 où seule l'information de l'image satellitaire dirige le contour actif (modèle

Bayésien), le recalage échoue. L'initialisation éloignée et la présence d'ombres expliquent le résultat obtenu. L'ombre est un artefact absent du MNS qui parvient alors à diriger le contour vers la bonne solution (figure 5.18).

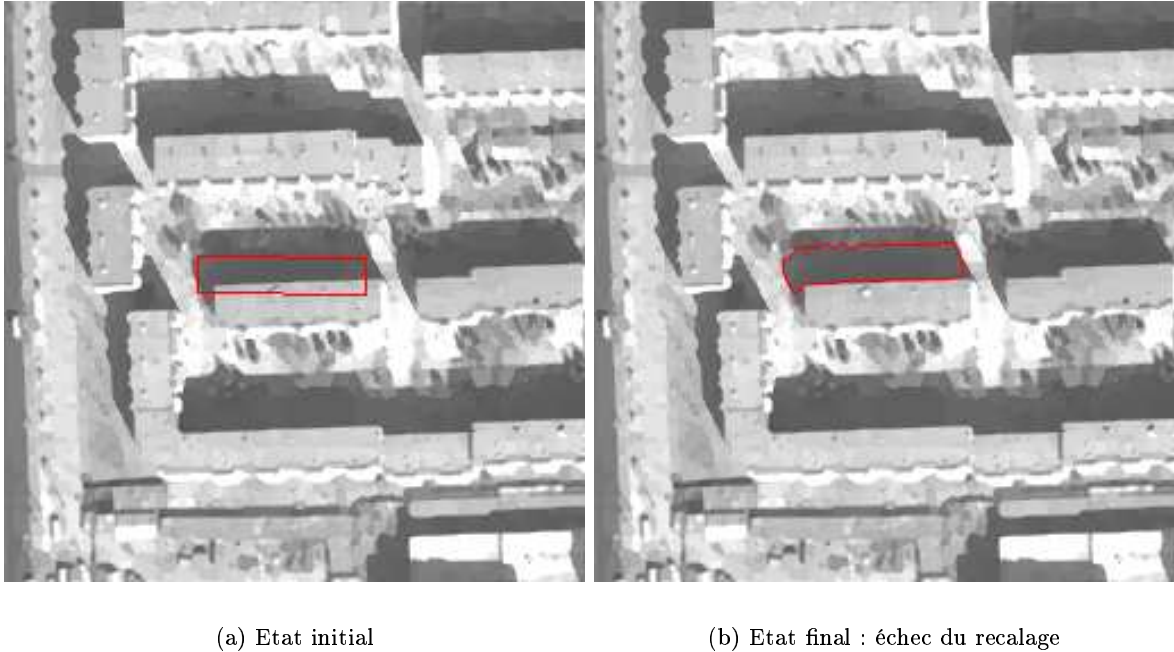


FIG. 5.17 – Échec du recalage sans insertion du MNS, avec une initialisation éloignée et le modèle Bayésien.

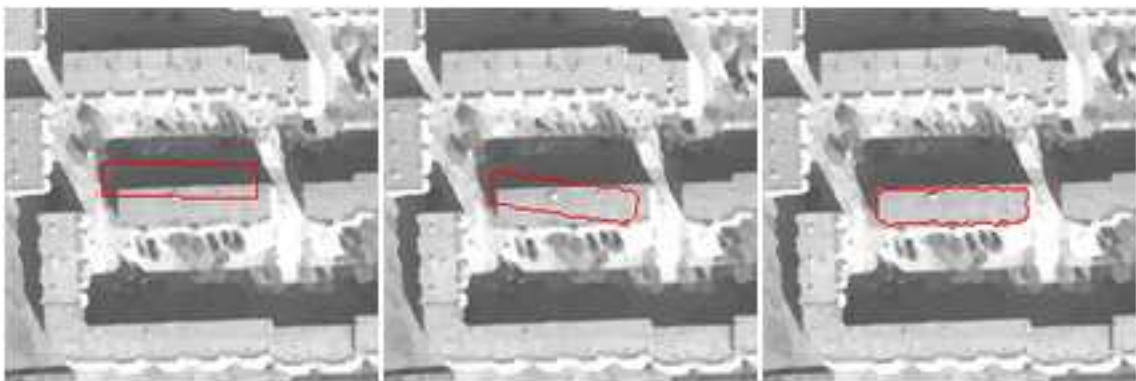


FIG. 5.18 – Recalage réussi avec une initialisation éloignée et l'insertion du MNS exogène dans la fonctionnelle d'énergie (modèle Bayésien pour l'image, modèle de Chan et Vese pour le MNS).

Enfin, l'expérience de la figure 5.19 illustre la complémentarité de l'information de région provenant du MNS (Chan et Vese) et de frontière extraite des segments de l'image (GVF). La force dérivée du MNS parvient à rapprocher le contour actif de la solution finale en surmontant les artefacts susceptibles de piéger un contour actif uniquement mû par les gradients de l'image. Le bâtiment est finalement segmenté avec succès (figure 5.19), contrairement au cas de la figure 5.20 sans incorporation de MNS.

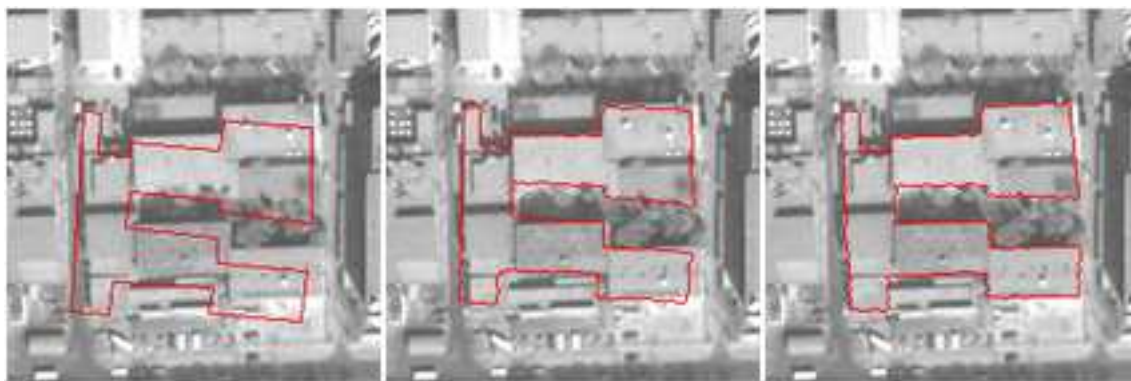
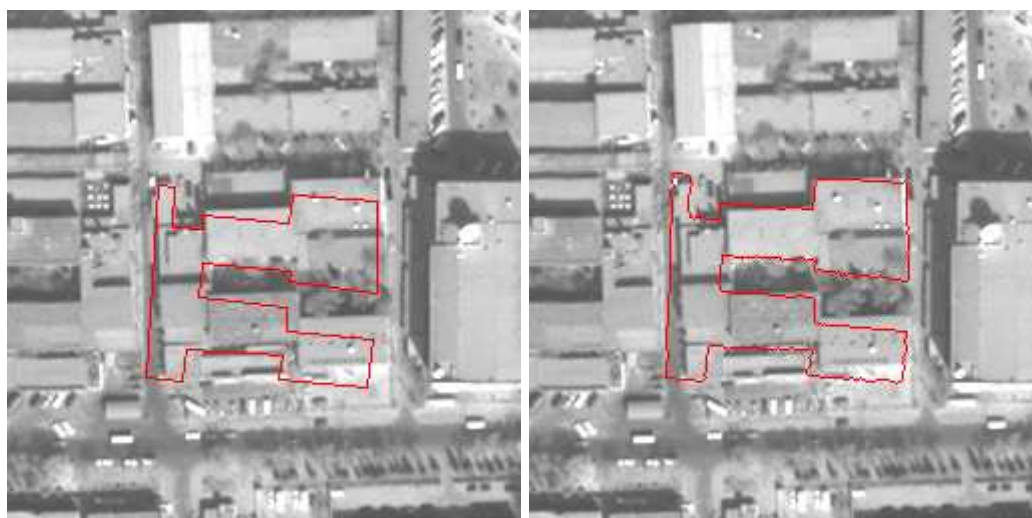


FIG. 5.19 – Recalage réussi avec une initialisation éloignée et l'insertion du MNS exogène dans la fonctionnelle d'énergie (modèle GVF pour l'image, modèle de Chan et Vese pour le MNS).



(a) Etat initial

(b) Etat final : échec du recalage

FIG. 5.20 – Échec du recalage sans insertion du MNS, avec une initialisation éloignée et le modèle GVF.

Le dernier avantage de la fusion de MNS concerne la diminution des temps de convergence. Le tableau 5.3 compare les temps de convergence sans et avec fusion de MNS grâce au calcul de leur ratio. Les expériences réalisées pour établir ces comparaisons ont été menées avec les bâtiments des figures 5.7.a et 5.7.b avec différentes initialisations. Nous constatons que les temps de convergence sont plus courts avec la fusion de MNS. Cette amélioration est due à la forte discrimination du bâtiment dans le MNS comparée à celle de l'image. Ainsi, l'utilisation conjointe du MNS permet de conduire sûrement et plus efficacement le contour actif vers l'objet cible. Les résultats du tableau 5.3 dépendent de l'initialisation du contour actif et de l'image analysée. Néanmoins ils affichent une réelle tendance de gain en efficacité.

Expérience	ratio
Bâtiment en U, expérience 1	1.4
Bâtiment rectangulaire, expérience 1	1.3
Bâtiment en U, expérience 2	2.5
Bâtiment rectangulaire, expérience 2	1.3

TAB. 5.3 – Rapport entre les temps de convergence sans et avec fusion de MNS : $ratio = \frac{t_{sans\ fusion}}{t_{avec\ fusion}}$.

5.5.5.2 Résultats avec une contrainte de forme spatio-temporelle

L'expérience illustrée en figure 5.21 reproduit celle de la figure 5.10.a avec le poids spatio-temporel de contrainte de forme exprimé en équation (5.10). L'amplitude du poids de contrainte varie entre $\lambda_{min} = 3$ et $\lambda_{max} = 30$ alors que la distance de non-interaction initiale est $d_0 = 7$ pixels. Alors que la mise en correspondance du bâtiment rectangulaire échouait avec un poids de contrainte de forme constant et le modèle Bayésien (figure 5.10.a), elle réussit avec sa formulation variable en fonction de l'espace et du temps (figure 5.21). La seconde image de la séquence illustre la première phase de l'algorithme : le poids étant faible, le contour peut changer de topologie afin de segmenter les zones de radiométrie similaire. Cette segmentation est évidemment limitée dans l'espace. La troisième image de la séquence illustre la seconde phase : le poids augmente progressivement et de façon continue ce qui a tendance à pénaliser et donc réduire la surface de la seconde région apparue à l'étape précédente. Finalement, le contour converge vers la bonne solution.



FIG. 5.21 – Segmentation avec une variation spatio-temporelle de la contrainte de forme et le modèle Bayésien ($\lambda_{min} = 3$, $\lambda_{max} = 30$, $d_0 = 7$).

La figure 5.22 présente un recalage réussi avec une initialisation éloignée de la cible dans l'image. La difficulté de la scène est la présence d'un bâtiment voisin (carré) avec des propriétés radiométriques similaires à celles du bâtiment d'intérêt. Dans le cas de l'utilisation d'un poids de contrainte constant (figure 5.23), nous remarquons que ce bâtiment périphérique est aussi segmenté, empêchant un recalage satisfaisant. À l'inverse, la contrainte de forme spatio-temporelle confère suffisamment de flexibilité au contour actif pour autoriser un changement de topologie pour atteindre une segmentation grossière de l'objet (deuxième image de la séquence en figure 5.22). Lorsque la contrainte uniforme est restaurée dans un second temps, les erreurs locales de segmentation situées dans l'espace de relaxation sont éliminées et un recalage réussi est finalement atteint.



FIG. 5.22 – Segmentation avec une variation spatio-temporelle de la contrainte de forme et le modèle Bayésien ($\lambda_{min} = 27$, $\lambda_{max} = 70$, $d_0 = 7$).



FIG. 5.23 – Segmentation avec un poids constant de contrainte de forme et le modèle Bayésien ($\lambda = 50$).

Les expériences présentées en figure 5.24 reproduisent avec une contrainte flexible celles ayant échoué en figure 5.14 avec un poids constant de contrainte de forme et un modèle basé frontière (GVF). L'amélioration due à la variation de la contrainte est plus nuancée avec le modèle GVF. Dans le premier cas (figure 5.24.a), le poids adaptatif ne permet pas un recalage correct, alors qu'il est efficace sur les deux autres exemples. Avec un contour actif basé sur le gradient de l'image, le poids spatio-temporel ne permet pas un recalage réussi avec une initialisation trop éloignée. En effet, ces modèles déformables sont affectés par l'information locale de l'image, ce qui est un problème lorsque la contrainte de forme est relaxée. Ainsi, il est probable que le contour reste piégé durant la période où le poids de contrainte est faible, ce qui l'empêche de se rapprocher du bâtiment à recalier dans l'image.

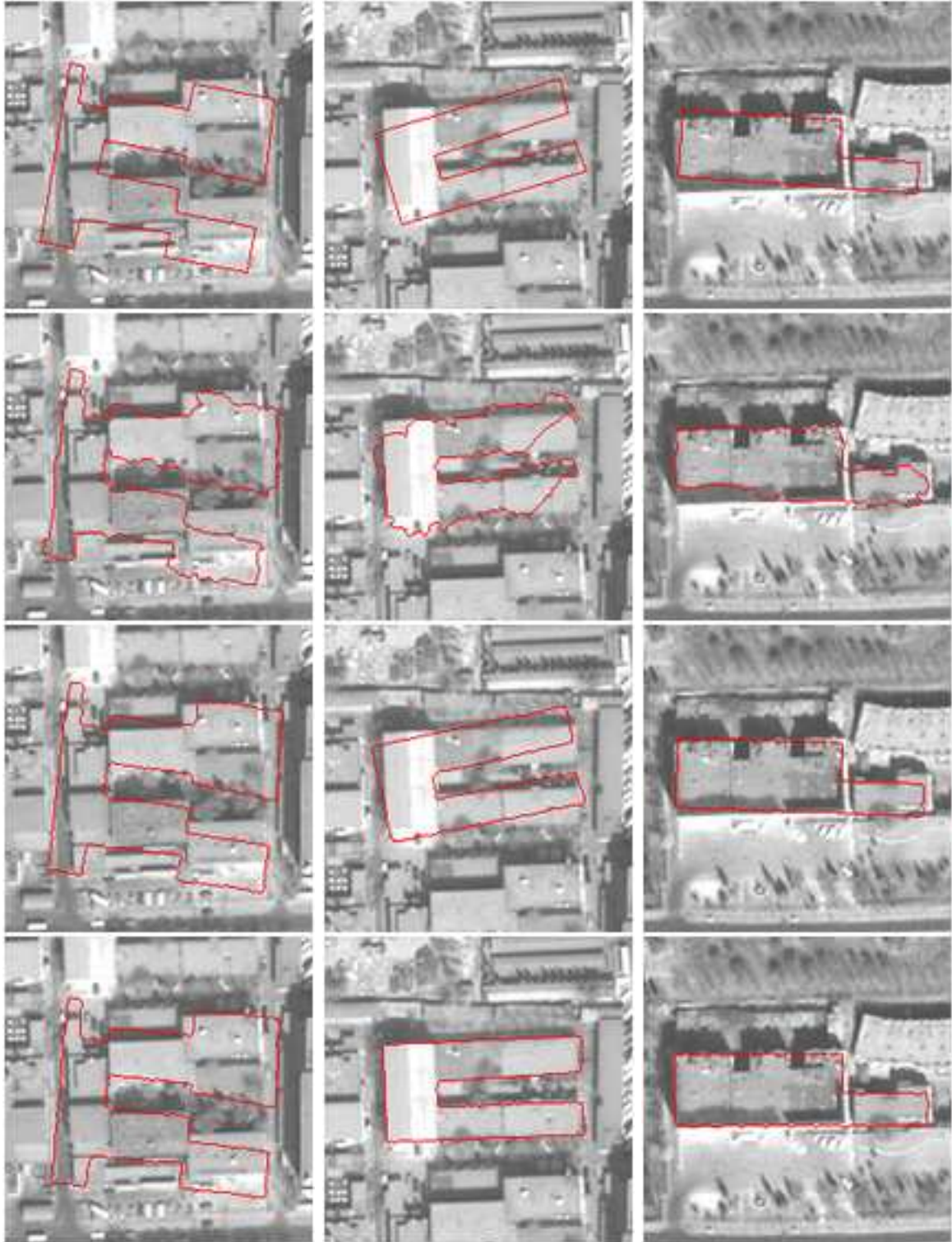


FIG. 5.24 – Variation spatio-temporelle de lambda avec le modèle GVF et $\lambda_{min} = 1$, $\lambda_{max} = 5$, $d_0 = 5$:

- (a) $\xi_0 = (s_0 = 1.1, \theta_0 = 0.25, \mu_{x,0} = -4, \mu_{y,0} = 1)$
- (b) $\xi_0 = (s_0 = 1.0, \theta_0 = -0.25, \mu_{x,0} = -2, \mu_{y,0} = -3)$
- (c) $\xi_0 = (s_0 = 1.0, \theta_0 = 0.1, \mu_{x,0} = -2, \mu_{y,0} = 2)$.

Nous proposons en figure 5.25 de mener des expériences avec des images n'appartenant pas au domaine de la télédétection. Ces images présentent les mêmes difficultés que les données de

télédétection en milieu urbain, à savoir un manque de discrimination de l'objet d'intérêt par rapport à son environnement. Les expériences des images 5.25.a et 5.25.b montrent l'efficacité du poids flexible de contrainte de forme avec le modèle Bayésien d'attache aux données. L'utilisation d'un poids constant empêche un recalage réussi : le manque de flexibilité du contour le fait rapidement sombrer dans un minimum local de la fonctionnelle d'énergie (figure 5.26).

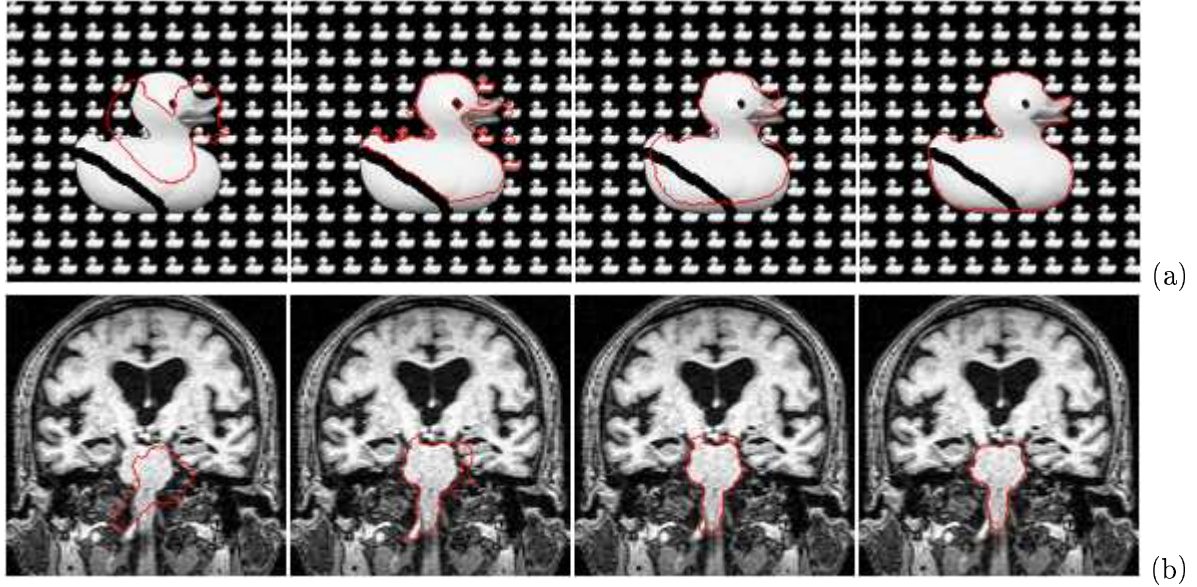


FIG. 5.25 – Segmentation avec variation spatio-temporelle du poids de contrainte de forme (modèle Bayésien) : (a) $\lambda_{min} = 15$, $\lambda_{max} = 100$, $d_0 = 15$ (b) $\lambda_{min} = 50$, $\lambda_{max} = 100$, $d_0 = 10$.

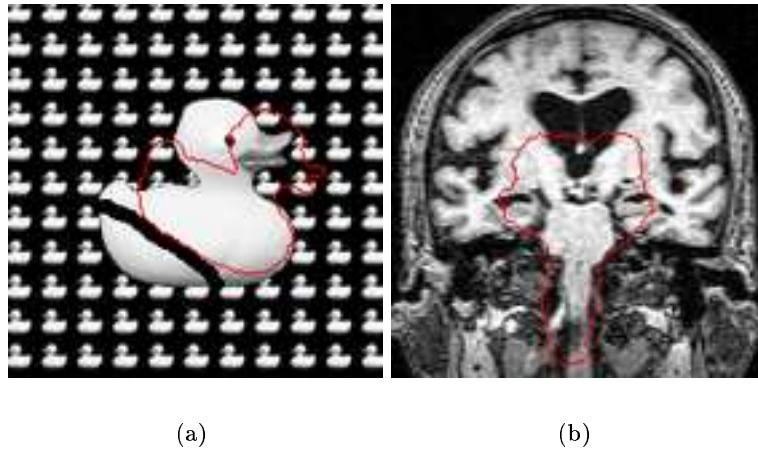


FIG. 5.26 – Échec du recalage avec un poids constant de contrainte de forme (modèle Bayésien) : (a) $\lambda = 20$ (b) $\lambda = 70$.

Enfin, nous proposons des expériences avec différentes façons de faire varier la contrainte de forme en fonction de l'espace et du temps. L'expérience illustrée en figure 5.27.a reproduit celle de la figure 5.22.a mais sans relaxation spatiale de la contrainte, seule la variation temporelle en amplitude de l'équation (5.10) est employée. Nous constatons que contrairement au cas spatio-temporel, la segmentation échoue. Dans ce cas, l'absence de relaxation ne permet pas au contour actif d'être suffisamment flexible pour segmenter grossièrement l'objet d'intérêt avant le rétablissement

de la contrainte uniforme. L'utilisation d'une valeur λ_{min} plus faible pour conférer plus de flexibilité échouerait dans la mesure où la flexibilité accordée ne serait pas spatialement bornée. Le contour segmenterait alors les objets périphériques ayant les mêmes propriétés statistiques que la cible. L'expérience de la figure 5.27.b reproduit celle exposée en figure 5.25.b. Un poids spatio-temporel est aussi utilisé, mais avec une formulation plus simple du profil spatial :

$$\lambda_{flex}(\psi(\mathbf{x}), t) = \lambda_a(t) \begin{cases} \frac{|\psi(\mathbf{x})|}{2d(t)} & \text{si } |\psi(\mathbf{x})| < 2d(t) \\ 1 & \text{sinon} \end{cases} \quad (5.11)$$

où ψ est l'ensemble de niveaux de la contrainte de forme, $d(t)$ et $\lambda_a(t)$ sont définis en équation (5.10). La formulation linéaire par morceaux de la relaxation spatiale de l'équation (5.11) est certainement plus intuitive que celle de l'équation (5.10). Néanmoins, l'expérience de la figure 5.27.b utilisant une telle formulation échoue. Le problème d'un tel profil spatial est un renforcement trop important de la contrainte de forme au voisinage du niveau zéro de ψ . Ainsi, la contrainte n'est pas autant relaxée qu'avec la présence d'un point stationnaire de la formulation exponentielle en équation (5.10). Augmenter la taille de l'espace de relaxation d_0 ne serait pas une solution viable car ce dernier étant trop étendu, le contour actif serait susceptible d'être influencé par les artefacts et objets environnants, empêchant la segmentation grossière de l'objet d'intérêt avant le rétablissement de la contrainte uniforme. L'avantage du profil spatial de l'équation (5.10) est d'autoriser une relaxation forte très proche de la forme *a priori* tout en préservant une contrainte élevée à une distance plus éloignée. Enfin, l'expérience de la figure 5.27.c répète celle de la figure 5.22.a avec une formulation plus simple du profil temporel. Les variations temporelles de $\lambda_a(t)$ et $d(t)$ sont remplacées par des fonctions respectivement croissante et décroissante de type échelon de Heaviside. La restauration de la contrainte de forme est ainsi rétablie brusquement et non progressivement. Dans ce cas, la figure 5.27.c montre que la segmentation est un échec. La raison d'un tel résultat concerne l'estimation dynamique des paramètres de T_{sim} permettant l'invariance de la contrainte de forme par similitude plane directe. Alors que l'espace de relaxation est effectif, le contour actif segmente grossièrement l'objet d'intérêt dans l'image. Les variations spatiales bornées du contour actif par rapport à la forme de référence peuvent biaiser légèrement l'estimation des paramètres de T_{sim} . Ainsi, lorsque la contrainte uniforme est rétablie soudainement, le contour actif ressemble à la forme de référence transformée par T_{sim} dont les paramètres sont mal estimés. En conséquence, le contour actif peut se retrouver loin de l'objet à segmenter dans l'image et avec une contrainte uniforme et forte, ce qui le rend très sensible aux minima locaux que nous cherchions à éviter. A l'inverse, un rétablissement progressif de la contrainte de forme permet à l'estimateur des paramètres de T_{sim} (ici le simplex) d'effectuer une estimation de moins en moins biaisée alors que le contour actif converge vers la cible (figure 5.27.c).

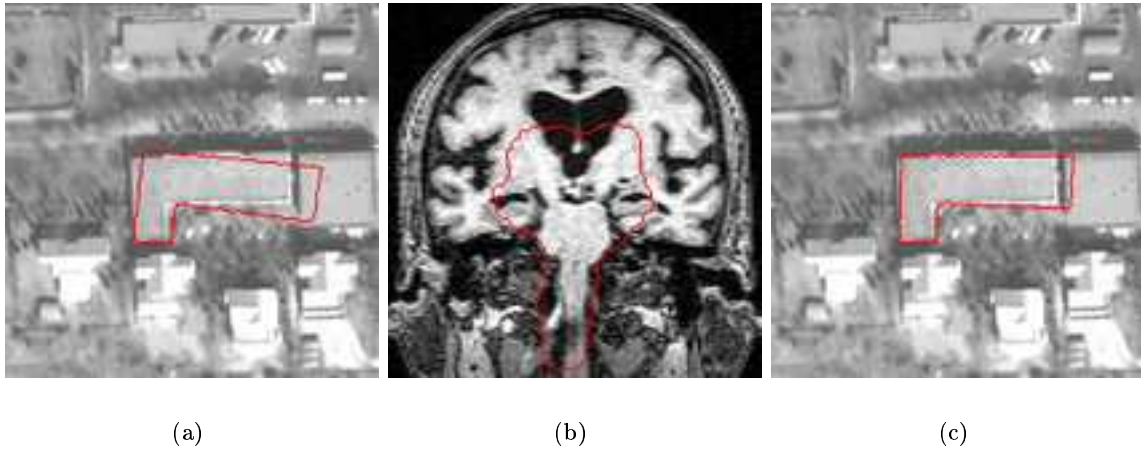


FIG. 5.27 – Échec du recalage (modèle Bayésien) avec : (a) unique variation temporelle de la contrainte de forme avec $\lambda_{min} = 30$, $\lambda_{max} = 70$ (pas de relaxation spatiale) ; (b) variation spatio-temporelle de la contrainte de forme avec un profil spatial défini en équation (5.11) et $\lambda_{min} = 27$, $\lambda_{max} = 70$, $d_0 = 7$; (c) variation spatio-temporelle de la contrainte de forme avec des fonctions échelon de Heaviside pour $\lambda_a(t)$ et $d(t)$.

Outre la possibilité d'assurer une convergence plus robuste, la flexibilité spatio-temporelle du poids de contrainte de forme permet aussi d'accélérer significativement les temps de calcul (tableau 5.4). En effet, lors de la première phase, la faible contrainte de forme confère suffisamment de flexibilité au contour actif pour qu'il s'approche efficacement de la cible dans l'image. La contrainte uniforme et constante qui a tendance à ralentir l'évolution du contour est rétablie uniquement à la fin du processus pour régulariser la segmentation. Le tableau 5.4 compare les temps de convergence sans et avec variation spatio-temporelle de la contrainte grâce au calcul de leur ratio. Comme pour le cas de la fusion de données exogènes, les expériences réalisées pour établir ces comparaisons ont été menées avec les bâtiments des figures 5.7.a et 5.7.b avec différentes initialisations. A l'instar des résultats avec fusion de MNS, la tendance affichée est un gain en efficacité, qui dépend aussi de l'initialisation et de la scène analysée.

Expérience	ratio
Bâtiment en U, expérience 1	1.8
Bâtiment rectangulaire, expérience 1	3.1
Bâtiment en U, expérience 2	4.3
Bâtiment rectangulaire, expérience 2	2.0
Bâtiment en U, expérience 3	2
Bâtiment rectangulaire, expérience 3	1.3

TAB. 5.4 – Rapport entre les temps de convergence sans et avec contrainte de forme spatio-temporelle : $ratio = \frac{t_{contrainte\ uniforme}}{t_{contrainte\ spatio-temporelle}}$.

5.6 Limites du modèle avec contrainte de forme

Les cartes numériques sont souvent entachées d'erreurs car elles résultent d'un traitement manuel, qui est de fait imparfait et subjectif. Si ces erreurs sont modélisables par la composition d'une

translation, rotation et dilatation/rétraction, notre méthode sera capable de s'en affranchir puisque le terme énergétique de contrainte de forme est invariant par similitude plane directe. Le plus souvent, les erreurs observées dans la carte sont locales et ne peuvent être modélisées et corrigées par une transformation globale (cf. section 2.3 détaillant les erreurs de $U_{interprétation}$). Ainsi, il est possible que par endroits, l'objet cartographique et sa représentation homologue dans l'image soient incohérents sans obéir à des règles particulières (figure 5.28).

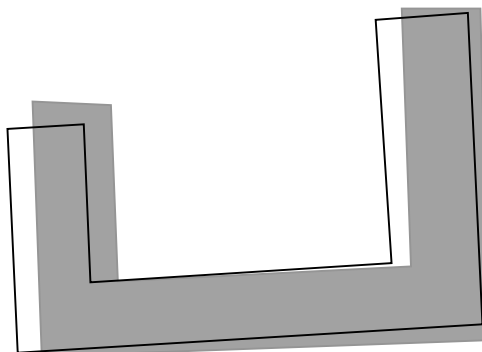


FIG. 5.28 – Superposition de la vérité terrain (grisée) et d'un polygone cartographique localement erroné

Les erreurs cartographiques influenceront et se manifesteront sur les résultats de la segmentation comme l'illustrent les figures 5.29 et 5.31. Sur la figure 5.28, on remarque qu'une correction par une rotation améliorerait la correspondance entre la carte et la vérité terrain réalisée à la main. Cependant, des incohérences locales subsistent : la barre centrale du U n'est pas assez large dans la carte, la barre gauche n'est pas assez longue non plus, alors que la taille de celle de droite semble convenir même si elle paraît trop large. Le résultat du recalage avec une carte localement erronée est alors un compromis entre l'image et la contrainte de forme corrompue. La figure 5.29 illustre l'exemple de segmentation du bâtiment en U avec le modèle Bayésien. Les figures 5.30 et 5.31 illustrent le résultat avec le modèle basé GVF sans erreur de la carte (figure 5.30) et avec erreurs locales de forme (figure 5.31). La présence d'erreurs locales dans la carte est un problème puisque nous tirons parti de celle-ci pour assister et contraindre le processus de mise en correspondance. Nous proposerons dans le chapitre suivant un moyen de surmonter cet obstacle en assouplissant l'incorporation de la contrainte de forme.



FIG. 5.29 – Convergence finale du contour actif avec une forme *a priori* erronée

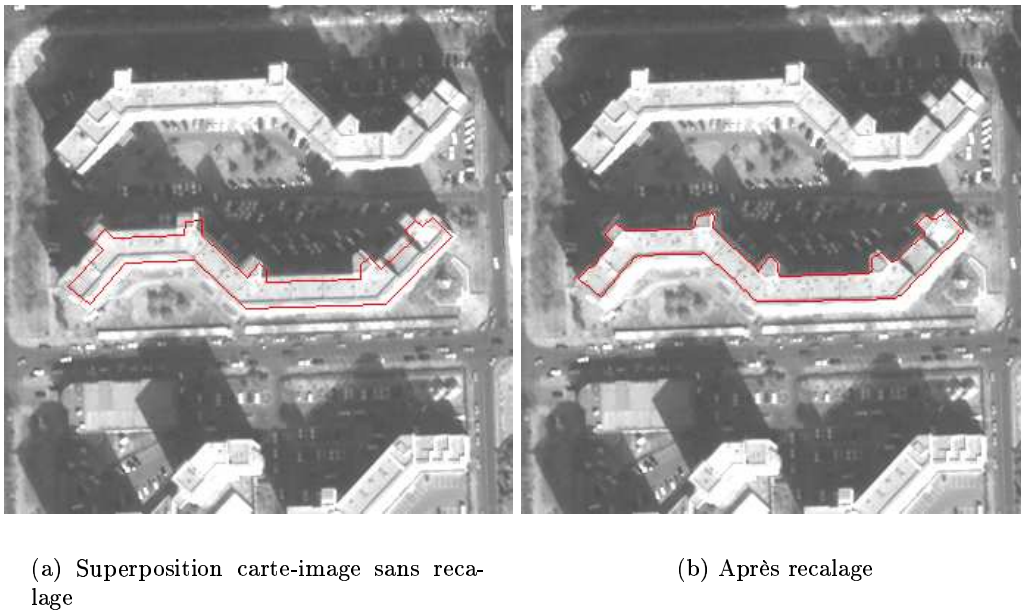


FIG. 5.30 – Recalage par contour actif basé sur l'information de gradient (GVF) et avec une forme *a priori* non erronée.

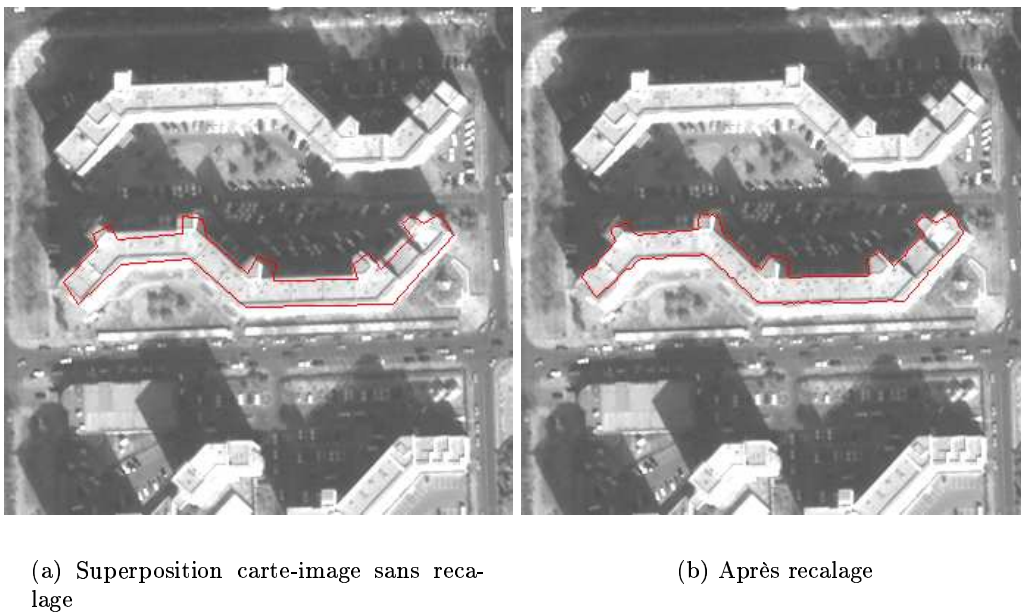


FIG. 5.31 – Recalage par contour actif basé sur l'information de gradient (GVF) et avec une forme *a priori* erronée dérivée de la carte.

5.7 Conclusion

L'objectif de ce chapitre était de montrer la contribution de l'insertion de contrainte de forme pour le recalage fin carte-image par contours actifs. La technique de mise en correspondance proposée permet de recaler individuellement chaque bâtiment symbolisé dans la carte sur sa représentation homologue dans une image satellitaire panchromatique haute résolution. La finalité est de réduire les variabilités exogènes carte-image afin d'améliorer la cohérence entre ces deux représentations et de permettre une détection de non-changement subséquente plus fiable.

Des résultats expérimentaux ont permis de mesurer l'apport de la forme *a priori* de la carte afin de palier les difficultés inhérentes aux images urbaines. Trois types de contours actifs ont été utilisés et ont exhibé une certaine sensibilité à l'initialisation ainsi qu'aux minima locaux de la fonctionnelle d'énergie minimisée. Nous avons proposé deux solutions pour rendre les modèles déformables plus robustes. La première regarde l'injection d'un MNS orthoscopique dans l'énergie d'attache aux données et permet une meilleure discrimination de la couche bâti par rapport au sol de l'image. La force dérivée de la donnée du MNS dirige alors plus sûrement et efficacement le contour actif vers le bâtiment dans l'image. La seconde contribution consiste à relaxer spatialement et temporellement la contrainte de forme au cours du processus de convergence. La flexibilité spatialement bornée conférée au contour actif autorise une segmentation approximative de l'objet dans l'image avant d'être spécifiquement régularisée par le rétablissement d'une contrainte uniforme. Des résultats expérimentaux ont montré l'efficacité de ces deux approches avec une réduction des temps de calcul. Une alternative à la descente de gradient a aussi été proposée afin d'assurer l'invariance de la contrainte de forme par transformation globale. Notre alternative basée sur la technique du simplex s'avère plus robuste et plus rapide que la méthode traditionnelle de descente de gradient.

Enfin, nous avons supposé que la forme *a priori* utilisée comme contrainte représentait exactement l'objet à recaler dans l'image, modulo une similitude plane directe. Nous n'avons jusqu'à présent pas abordé le problème de possibles disparités entre la forme globale du bâtiment dans l'image et celle de l'objet cartographique. La présence de telles variabilités est cependant fréquente, elle est due soit à des erreurs locales de forme de la carte ou à des changements partiels intervenus sur le bâti. Pour traiter le recalage carte-image avec de tels artefacts, de nouvelles énergies de contrainte de forme doivent être formulées.

Chapitre 6

Contrainte de forme autorisant les variations parallèles

6.1 Introduction

Il est possible que les objets cartographiques contiennent des incohérences *locales* de forme par rapport à la réalité de l'image. Nous avons illustré au chapitre précédent la présence de ces erreurs qui empêchent un recalage totalement satisfaisant avec le bâtiment représenté dans l'image satellitaire. Ces incohérences sont aussi dues à des changements locaux et partiels apparus au niveau du bâtiment de l'image. Dans les deux cas, les parties du contour actif, contraintes à tort par la forme *a priori* n'atteignent pas les bords de l'objet vu dans l'image en raison de la façon rigide dont a été insérée la contrainte de forme qui est uniquement invariante par une transformation *globale*. Le problème des écarts locaux par rapport à une forme *a priori* a été essentiellement traité d'un point de vue statistique grâce à la donnée d'échantillons d'apprentissage [114, 28, 66, 105, 93, 103, 30]. Nous proposons une autre solution non heuristique, consistant à autoriser certains écarts du contour actif par rapport à la forme de référence. La classe des écarts autorisés est celle des déplacements contraints par le parallélisme par rapport aux segments de la forme *a priori* et correspond aux incohérences locales carte-image les plus souvent rencontrées dans les données cartographiques. Pour permettre de telles déformations, nous proposons une nouvelle énergie de contrainte de forme. Celle-ci est la somme d'une énergie linéaire comparant les normales du contour actif et de la forme *a priori*, et d'une énergie quadratique imposant au contour actif des contraintes génériques de rectilinéarité et d'orthogonalité. Des résultats expérimentaux préliminaires, mais néanmoins encourageants sont présentés.

6.2 Énergie de contrainte de forme autorisant les variations parallèles

La façon d'introduire la notion de parallélisme dans la fonctionnelle d'énergie de contrainte de forme consiste à imposer l'alignement des normales du contour actif et de la forme de référence. En termes d'ensembles de niveaux, les normales sont $\nabla\phi(\mathbf{x})$ et $\nabla\psi(\mathbf{x})$ ¹². Le nouveau terme énergétique

¹On suppose que la condition $|\nabla\phi| = 1$ est toujours vérifiée, ce qui est en pratique vrai si on réinitialise la fonction ϕ à chaque itération de l'équation d'évolution. La condition $|\nabla\psi_0| = 1$ est toujours vérifiée puisque la forme *a priori* n'est pas mise à jour. Ainsi $|\nabla\psi_0| = |\nabla\psi| = 1$ puisque que les ensembles de niveaux ψ_0 et ψ ne diffèrent que par une transformation isométrique qui par définition n'affecte pas la fonction distance.

²Les déformations locales contraintes par le parallélisme sont susceptibles d'être inhibées par l'invariance par similitude plane directe (échelle) exprimée en équation (4.54). Par conséquent, nous excluons toute invariance par transformation globale dans ce chapitre. En reprenant la notation des chapitres précédents, nous avons alors : $\psi = \psi_0 \circ T_{sim} = \psi_0$ où T_{sim} est l'identité.

formulé sur la base de l'équation (4.54) est alors :

$$J_{shape}(\phi, \psi) = \lambda_{para} \int_{\Omega} |\nabla \phi(\mathbf{x}) - \nabla \psi(\mathbf{x})|^2 (H(\phi(\mathbf{x})) - H(\psi(\mathbf{x})))^2 d\mathbf{x} \quad (6.1)$$

L'équation d'évolution associée à une fonctionnelle $J = J_{image} + J_{Shape}$ nécessite le calcul du gradient de J_{shape} par rapport à ϕ . Le détail des calculs est donné en annexe E, et aboutit à :

$$\begin{aligned} \phi_t(\mathbf{x}, t) = & -\frac{\partial J_{image}}{\partial \phi} \\ & + 4\lambda_{para} \delta_{\alpha}(\psi(\mathbf{x}, t)) (\langle \nabla \phi(\mathbf{x}), \nabla \psi(\mathbf{x}) \rangle - 1) [H_a(\phi(\mathbf{x}, t)) - H_a(\psi(\mathbf{x}, t))] \\ & + 2\lambda_{para} (\Delta \phi(\mathbf{x}) - \Delta \psi(\mathbf{x})) [H_a(\phi(\mathbf{x}, t)) - H_a(\psi(\mathbf{x}, t))]^2 \end{aligned} \quad (6.2)$$

Les deux derniers termes de l'équation (6.2) sont le résultat de l'insertion du nouveau terme d'énergie. Le premier tend vers zéro lorsque la normale au contour actif a le même sens et la même direction que celle de la forme *a priori*. Cependant, ce terme a une influence très locale puisqu'il est pondéré par la distribution de Dirac qui dépend ici de ψ . Ainsi, lorsque le contour actif sera éloigné de la forme de référence (au delà de quelques pixels), ce terme n'aura aucune influence. Le second terme n'est pas local, et impose une contrainte sur la courbure du contour actif par rapport à celle de la forme *a priori*. C'est ce terme qui contribuera principalement à imposer la contrainte de parallélisme. La contrainte de parallélisme proposée s'applique aux lignes de niveau de la forme *a priori* ψ , cependant celles-ci sont courbes loin des singularités géométriques (coins) du niveau 0 de la forme *a priori*. Cela aura pour effet d'arrondir le contour actif lorsque ce dernier s'éloignera de la forme de référence (cf. figure 6.1).

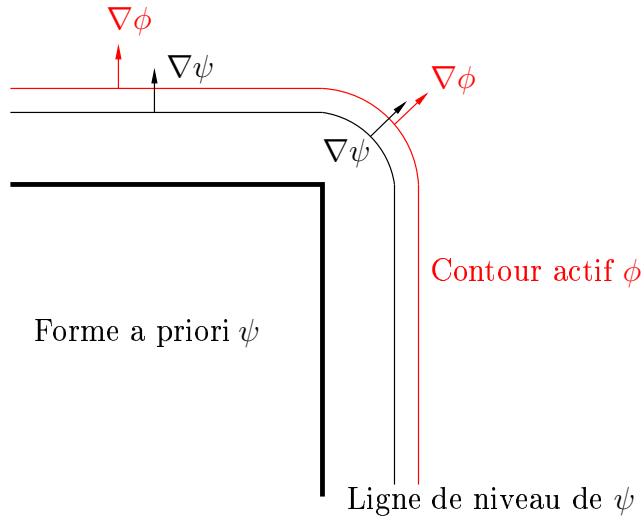


FIG. 6.1 – Effet d'arrondissement au niveau des singularités géométriques dû à la nouvelle énergie de contrainte de forme basée sur l'alignement des normales.

6.3 Connaissance *a priori* générique par modèles quadratiques

Afin de résoudre l'effet indésirable d'arrondissement au niveau des singularités géométriques de type "coins", nous introduisons une énergie correctrice visant à favoriser la formation de coins à 90

degrés et de segments rectilignes. Nous introduisons ainsi une contrainte de forme générique en plus de celle spécifique représentée par la forme *a priori*. Les énergies quadratiques exposées à la section 4.2.3 ont été retenues pour corriger l'arrondissement du contour actif.

6.3.1 Énergie quadratique correctrice

6.3.1.1 Modèle

Nous proposons d'ajouter à l'équation (6.1) l'énergie quadratique de correction suivante :

$$J_{correction}(\mathbf{C}) = \oint \oint ds ds' (\sin 2\theta)^2 \Psi(|\mathbf{C}(s) - \mathbf{C}(s')|) \quad (6.3)$$

où θ est l'angle entre les tangentes au contour³ $\mathbf{C}$ aux points d'abscisse curviligne s et s' (cf. figure 6.2). Ψ est une fonction décroissante tendant asymptotiquement vers 0 et définit le potentiel d'interaction entre $\mathbf{C}(s)$ et $\mathbf{C}(s')$. Notons que le domaine d'intégration de $J_{correction}$ est cette fois le

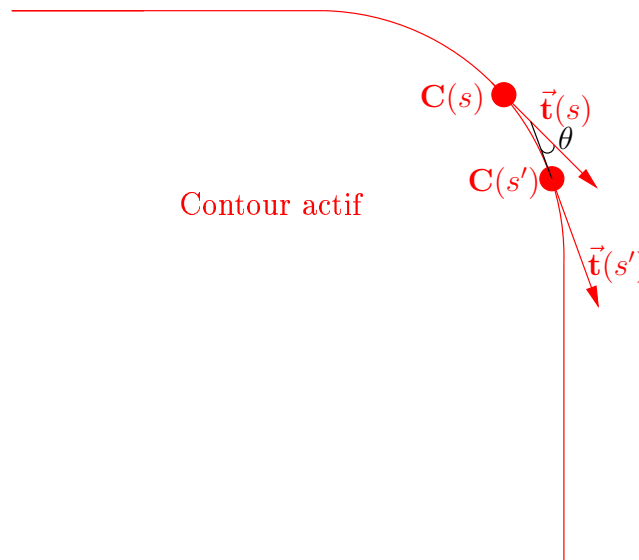


FIG. 6.2 – Interaction à longue distance due à l'énergie correctrice quadratique. Les vecteurs tangents $\vec{t}(s)$ et $\vec{t}(s')$ auront tendance à être orthogonaux ou colinéaires.

contour lui-même et non Ω . Une telle énergie permettra de favoriser la formation de parties rectilignes et de coins au sein du contour actif. L'imposition de contraintes d'orthogonalité a déjà été proposée par H. Oriot dans [84] dans le cadre de la segmentation de bâtiments par *snakes*. Nous nous sommes inspirés de cette idée afin de proposer la formulation donnée en équation (6.3). Cependant, à la différence des travaux de H. Oriot qui n'utilisent pas de forme *a priori*, nous proposons de formuler un critère d'énergie quadratique et non linéaire; de plus, nous représentons le contour actif de façon implicite.

En remarquant que $\sin 2\theta = 2 \sin \theta \cos \theta$, l'énergie de correction peut être réécrite selon :

$$J_{correction}(\mathbf{C}) = 4 \oint \oint ds ds' \{ \cos^2 \theta - \cos^4 \theta \} \Psi(|\mathbf{C}(s) - \mathbf{C}(s')|) \quad (6.4)$$

Cette énergie est minimale pour $\theta = 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$:

³Nous considérerons dans ce document que $\mathbf{C}$ est un contour orienté dans le sens trigonométrique.

- Si $\theta = 0$. Cela implique que les tangentes sont alignées dans le voisinage d'interaction défini par le profil Ψ . Ainsi, la rectilinéarité des objets extraits par le contour actif se trouvera renforcée. Cette propriété est souhaitée puisque les bâtiments exhibent très souvent des bords de toit rectilignes (objets fortement polygonaux).
- Si $\theta = \frac{\pi}{2} \text{ modulo } \pi$. Ce cas représente l'objectif visé par la formulation de l'énergie correctrice : les coins à 90 degrés sont favorisés pour contrecarrer l'effet d'arrondissement.
- Si $\theta = \pi$. Les tangentes sont dans ce cas anti-parallèles. Le voisinage d'interaction défini par Ψ a une étendue spatiale de l'ordre de quelques pixels avant de tendre vers 0. Dans le cas des bâtiments, l'occurrence de tangentes anti-parallèles survient alors que la distance $|\mathbf{C}(s) - \mathbf{C}(s')|$ est grande et où la fonction Ψ est déjà quasi nulle. Ainsi l'effet des tangentes opposées est négligeable dans notre cas. Ce cas surviendrait si l'objet à recaler exhibait des structures fines ou labyrinthiques, ce qui est le cas des routes imagées à petite échelle, mais pas celui des bâtiments à très haute résolution.

Avant de déduire la force quadratique à partir de la dérivation de $J_{\text{correction}}$ par rapport à $\mathbf{C}$, il convient de réécrire l'équation (6.4) dans une forme indépendante du paramétrage du contour. Pour ce faire, considérons un paramètre arbitraire p dont la relation avec l'abscisse curviligne s est :

$$\frac{d\mathbf{C}}{ds} = \frac{d\mathbf{C}}{dp} \frac{dp}{ds} \quad (6.5)$$

ce qui donne

$$\frac{ds}{dp} = \left| \frac{d\mathbf{C}}{dp} \right| \quad (6.6)$$

Dans la suite de ce chapitre, nous adoptons les notations suivantes :

$$\begin{aligned} \mathbf{C} &\equiv \mathbf{C}(p) \\ \mathbf{C}' &\equiv \mathbf{C}(p') \\ R(p, p') &= |\mathbf{C} - \mathbf{C}'| \\ R(p, p') &\equiv R \\ \tilde{\mathbf{R}} &\equiv (\mathbf{C} - \mathbf{C}') R^{-1} \end{aligned}$$

La dérivation par rapport au paramétrage du contour est symbolisée par un point : $\dot{\mathbf{C}} = \frac{d\mathbf{C}}{dp}$. Les vecteurs tangents aux points p et p' deviennent alors : $\dot{\mathbf{C}}$ et $\dot{\mathbf{C}}'$. En utilisant ces notations, l'énergie quadratique à minimiser de l'équation (6.4) devient :

$$J_{\text{correction}}(\mathbf{C}) \sim \oint \oint dp dp' \left\{ \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 |\dot{\mathbf{C}}|^{-1} |\dot{\mathbf{C}}'|^{-1} - \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^4 |\dot{\mathbf{C}}|^{-3} |\dot{\mathbf{C}}'|^{-3} \right\} \Psi(R) \quad (6.7)$$

6.3.1.2 Gradient de l'énergie

En vue de trouver l'équation différentielle régissant l'évolution du contour actif avec correction quadratique, la variation de $J_{\text{correction}}$ par rapport à $\mathbf{C}$ doit être calculée :

$$\Delta J_{\text{correction}} = J_{\text{correction}}(\mathbf{C} + \delta \tilde{\mathbf{C}}) - J_{\text{correction}}(\mathbf{C}) \quad (6.8)$$

où $\Delta J_{correction}$ est la variation d'énergie $J_{correction}$ engendrée par l'écart infinitésimal du contour $\delta\tilde{\mathbf{C}}$. Le calcul de $\Delta J_{correction}$, long et fastidieux est synthétisé en annexe E. En supposant que la courbe est à nouveau paramétrée par l'abscisse curviligne, la variation d'énergie est finalement :

$$\begin{aligned}\Delta J_{correction} &= \oint \oint ds ds' \Psi(R) \langle \tilde{\alpha}(s, s'), \delta\tilde{\mathbf{C}} \rangle \\ &+ \oint \oint ds ds' \Psi'(R) \langle \tilde{\beta}(s, s'), \delta\tilde{\mathbf{C}} \rangle\end{aligned}\quad (6.9)$$

avec

$$\tilde{\alpha}(s, s') = 2 \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 \left(1 - 3 \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 \right) \ddot{\mathbf{C}} + 4 \left(1 - 6 \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 \right) \langle \dot{\mathbf{C}}', \ddot{\mathbf{C}} \rangle \left(\langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle \dot{\mathbf{C}} - \dot{\mathbf{C}}' \right) \quad (6.10)$$

et

$$\begin{aligned}\tilde{\beta}(s, s') &= 2 \left(1 - \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 \right) \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 \tilde{\mathbf{R}} + 4 \langle \dot{\mathbf{C}}, \tilde{\mathbf{R}} \rangle \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle \left(2 \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 - 1 \right) \dot{\mathbf{C}}' \\ &+ 2 \langle \dot{\mathbf{C}}, \tilde{\mathbf{R}} \rangle \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 \left(1 - 3 \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 \right) \dot{\mathbf{C}}\end{aligned}\quad (6.11)$$

En notant que $\ddot{\mathbf{C}} = \kappa \mathbf{n}$ où $\mathbf{n}$ est la normale unitaire à $\mathbf{C}$, et que κ est la courbure signée du contour, la force quadratique $F_{quad}(s)$ calculée au point d'abscisse curviligne s est :

$$F_{quad}(s) = \left\langle \frac{\partial J_{correction}}{\partial \mathbf{C}}, \mathbf{n} \right\rangle = \oint ds' \Psi(R) \langle \tilde{\alpha}(s, s'), \mathbf{n} \rangle + \oint ds' \Psi'(R) \langle \tilde{\beta}(s, s'), \mathbf{n} \rangle \quad (6.12)$$

avec

$$\langle \tilde{\alpha}(s, s'), \mathbf{n} \rangle = 2\kappa \left\{ \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 \left(1 - 3 \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 \right) - 2 \langle \dot{\mathbf{C}}', \mathbf{n} \rangle^2 \left(1 - 6 \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 \right) \right\} \quad (6.13)$$

$$\langle \tilde{\beta}(s, s'), \mathbf{n} \rangle = 2 \left(1 - \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 \right) \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 \langle \tilde{\mathbf{R}}, \mathbf{n} \rangle + 4 \langle \dot{\mathbf{C}}, \tilde{\mathbf{R}} \rangle \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle \left(2 \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 - 1 \right) \langle \dot{\mathbf{C}}', \mathbf{n} \rangle \quad (6.14)$$

Le modèle de contrainte de forme autorisant les déplacements parallèles par rapport à l'*a priori* et avec correction quadratique est en définitive régi par l'équation :

$$\begin{aligned}\phi_t(\mathbf{x}, t) &= -\frac{\partial J_{image}}{\partial \phi} \\ &+ 4\lambda_{para} \delta_\alpha(\psi(\mathbf{x}, t)) (\langle \nabla \phi(\mathbf{x}), \nabla \psi(\mathbf{x}) \rangle - 1) [H_a(\phi(\mathbf{x}, t)) - H_a(\psi(\mathbf{x}, t))] \\ &+ 2\lambda_{para} (\Delta \phi(\mathbf{x}) - \Delta \psi(\mathbf{x})) [H_a(\phi(\mathbf{x}, t)) - H_a(\psi(\mathbf{x}, t))]^2 - \lambda_{quad} F_{quad, ext}(\mathbf{x}) |\nabla \phi(\mathbf{x})|\end{aligned}\quad (6.15)$$

où $F_{quad, ext}(\mathbf{x})$ est la force correctrice étendue en tout point du domaine défini par la bande étroite entourant le niveau zéro de la fonction ϕ selon l'équation (6.18).

6.3.2 Algorithme

L'algorithme d'évolution du contour actif est alors :

1. Construction de $\phi(\mathbf{x}, t = 0)$ dans une bande étroite; construction de $\psi(\mathbf{x}, t = 0)$ sur le domaine Ω de l'image.

2. Extraction du contour $\mathbf{C}$.
3. Calcul de la tangente et de la courbure pour chaque point du contour. Lorsque le point a des coordonnées non entières, les quantités géométriques sont interpolées linéairement entre les deux plus proches voisins entourant le point.
4. Calcul de la force $F_{quad}(s)$ pour chaque point du contour.
5. Extension de la force : $F_{quad}(s) \rightarrow F_{quad,ext}(\mathbf{x})$ selon l'équation (6.18).
6. Évolution du contour actif en utilisant l'équation (6.15).
7. Réinitialisation de la fonction distance de l'ensemble de niveaux ϕ .
8. Boucle les étapes 2 à 7 jusqu'à la convergence.

6.3.2.1 Calcul de la force quadratique

La force $F_{quad}(s)$ est une intégrale sur un contour fermé dont on fait une approximation par la méthode des trapèzes de Newton-Cotes. Si l'on écrit $F_{quad}(s)$ sous la forme :

$$F_{quad}(s) = \oint ds' f(s, s')$$

L'approximation trapézoïdale sera :

$$\hat{F}_{quad}(s) = \sum_{s'=0}^{L-1} \frac{f(s, s') + f(s, s' + 1)}{2} |\mathbf{C}(s') - \mathbf{C}(s' + 1)| \quad (6.16)$$

que l'on peut aussi écrire sous la forme :

$$\hat{F}_{quad}(s) = \sum_{s'=0}^{L-1} \frac{f(s, s')}{2} (|\mathbf{C}(s' - 1) - \mathbf{C}(s')| + |\mathbf{C}(s') - \mathbf{C}(s' + 1)|) \quad (6.17)$$

où : L est la longueur du contour et avec les conditions cycliques : $f(s, 0) = f(s, L)$; $\mathbf{C}(0) = \mathbf{C}(L)$. La formule d'intégration numérique donnée en équation (6.16) impose implicitement que les points $\mathbf{C}(s)$ du contour soient ordonnés afin de pouvoir calculer la quantité $|\mathbf{C}(s') - \mathbf{C}(s' + 1)|$. Après que le contour actif soit extrait par détection des zéros de l'ensemble de niveaux ϕ , l'algorithme de traçage de contour proposé par Pavlidis dans [94] est utilisé pour ordonner les points de $\mathbf{C}$.

6.3.2.2 Extension de la force

L'extension de la force consiste à renseigner $F_{quad,ext}(\mathbf{x})$ à chaque location $\mathbf{x}$ de la bande étroite avec la valeur $\hat{F}_{quad}(s)$ du point du contour le plus proche de $\mathbf{x}$. Ainsi, la valeur étendue de la force est constante le long des normales aux lignes de niveau de ϕ . Les techniques d'extension de force sont référencées dans [111]. La façon la plus populaire consiste à résoudre l'équation aux dérivées partielles proposée par les auteurs de [1] :

$$\frac{\partial F_{quad,ext}(\mathbf{x})}{\partial t} + \text{sign}(\phi) \left\langle \frac{\nabla \phi}{|\nabla \phi|}, \nabla F_{quad,ext}(\mathbf{x}) \right\rangle = 0 \quad (6.18)$$

Les détails relatifs à l'extension de $F_{quad,ext}(\mathbf{x})$ sont disponibles dans [102].

6.3.3 Difficultés d'implantation

La programmation et l'implantation de l'équation d'évolution (6.15) a soulevé de nombreux problèmes de stabilité numérique. En particulier, l'évaluation précise des quantités géométriques

telles que la tangente ou la courbure a été capitale pour atténuer les instabilités. Les mesures listées ci-dessous par ordre décroissant d'importance ont été mises en œuvre pour limiter les problèmes de stabilité :

1. Calcul précis des quantités géométriques.

- (a) Calcul de la courbure selon la formule proposée par les auteurs de [135] dans le cadre de la mécanique des fluides. La formulation proposée est robuste vis-à-vis des singularités géométriques. La courbure se calcule uniquement grâce à l'ensemble de niveaux ϕ .
- (b) Calcul précis des tangentes. Les normales aux points du contour sont estimées, les tangentes associées s'en déduisent par rotation de $\pi/2$ dans le sens trigonométrique. Nous estimons la normale à partir de ϕ comme la moyenne des normales dans quatre directions [111, pp. 70]. Des dérivées aux différences finies sont sollicitées lors de l'estimation. Afin d'accroître la précision des dérivées, nous les calculons par une technique Essentiellement Non-Oscillatoire (ENO) qui consiste à faire une approximation locale de ϕ par un polynôme d'ordre n . Le gradient de ϕ estimé de cette façon est plus robuste au voisinage de singularités [68].

- 2. **Réinitialisation de l'ensemble de niveaux.** La technique de reconstruction de la bande étroite rapide exposée en section 4.3.4.1 s'est révélée trop grossière pour être utilisée avec les modèles quadratiques. En effet, la technique dite de bande étroite rapide déplace le contour (niveau zéro de l'ensemble de niveaux) à une échelle sub-pixellaire, ce qui est source d'instabilités. Nous avons opté pour une autre technique de réinitialisation conservant la surface incluse dans le contour, et qui limite le déplacement du front [68].
- 3. **Estimation de $\nabla\phi$.** Le calcul du gradient de l'ensemble de niveaux représentant le contour actif est utilisé de nombreuses fois lors du calcul de l'énergie quadratique : calcul des tangentes, réinitialisation de ϕ , extension de la force quadratique. Sachant que nous cherchons à créer des singularités géométriques (coins) où le calcul de $\nabla\phi$ est délicat et sensible aux imprécisions de calcul, nous avons employé une méthode ENO avec un polynôme d'ordre trois pour obtenir une mesure robuste de ce gradient.
- 4. **Choix de la fonction d'interaction Ψ .** Nous avons mené des expériences avec le profil d'interaction Ψ de l'équation (4.60). Un tel profil n'est pas forcément adapté pour notre application puisqu'il est dédié à la répulsion de tangentes anti-parallèles dans un voisinage spatial inférieur à d_{min} . De plus, la décroissance de la fonction de l'équation (4.60) est susceptible d'occasionner des perturbations de Gibbs dans le domaine spatial. Le choix d'une fonction de Bessel modifiée de seconde espèce $\Psi(x) = K_0(x/a)$ ($a \in \mathbb{R}^{*+}$) dont l'allure est visible en figure (6.3) a donné de meilleurs résultats, avec des instabilités numériques atténuées. $K_n(x)$ est solution de l'équation différentielle modifiée de Bessel, et sa définition intégrale est :

$$K_n(x) = \frac{\sqrt{\pi}}{(n-1/2)!} \left(\frac{1}{2}x\right)^n \int_1^\infty e^{-zx} (z^2 - 1)^{n-1/2} dz \quad (6.19)$$

Pour $n = 0$, on a :

$$K_0(x) = \int_0^\infty \cos(x \sinh t) dt \quad (6.20)$$

La dérivée de K_0 par rapport à x est :

$$\frac{dK_0(x)}{dx} = -K_1(x)$$

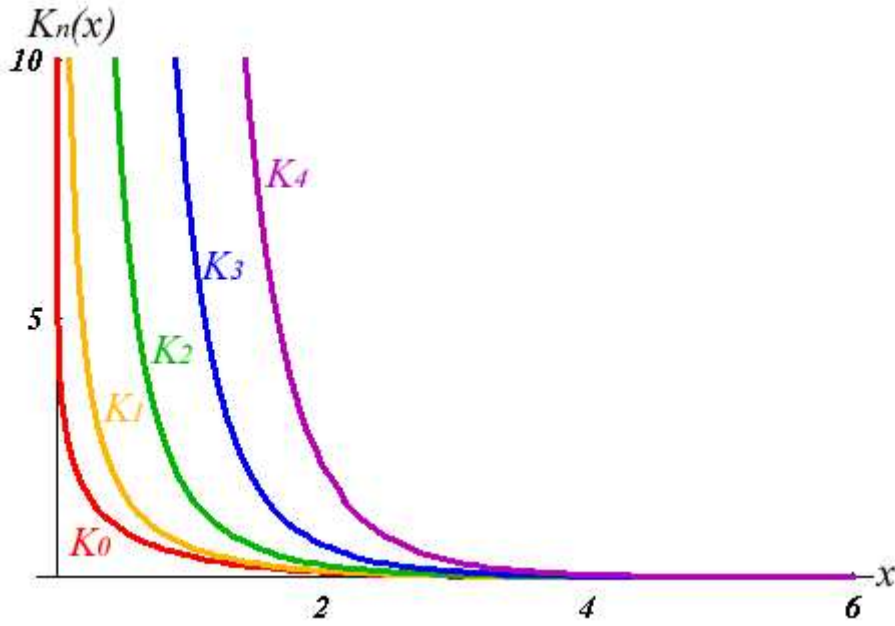


FIG. 6.3 – Fonctions modifiées de Bessel de seconde espèce.

La fonction K_0 décroît asymptotiquement, mais diverge lorsque $x \rightarrow 0$. Cette singularité n'empêche cependant pas l'intégrale de l'équation (6.12) de converger. Toutefois, il faut prendre en compte cette singularité lors du calcul numérique de l'intégration exprimé en équation (6.16). En effet, la divergence de $\Psi(x)$ en 0 fait que le terme $f(s, s' = s)$ de (6.17) n'est plus calculable. Nous faisons une approximation $\hat{D}$ du terme

$$D = \frac{f(s, s)}{2} (|\mathbf{C}(s-1) - \mathbf{C}(s)| + |\mathbf{C}(s) - \mathbf{C}(s+1)|)$$

en sur-échantillonnant les segments $[\mathbf{C}(s-1) \mathbf{C}(s)]$ et $[\mathbf{C}(s) \mathbf{C}(s+1)]$ avec un pas h fixé (figure 6.4). Les valeurs de f aux endroits sur-échantillonnés des segments sont calculées à partir d'interpolations linéaires des quantités géométriques, $\hat{D}$ est calculée par la méthode des trapèzes.

5. **Extraction du contour.** L'extraction du contour par interpolation linéaire peut se révéler inadéquate à cause d'une localisation trop grossière des points du contour, biaisant *in fine* le calcul des quantités géométriques. Nous avons opté pour une méthode ENO faisant une approximation de l'ensemble de niveaux par un polynôme de degré deux [113]. Les racines du polynôme désignent les points du contour $\mathbf{C}$.
6. **Régularisation due à la courbure.** Un terme de régularisation générique de la courbure a été utilisé pour amoindrir les instabilités restantes.

6.4 Expériences et analyse des résultats

6.4.1 Résultats sans correction quadratique

Nous proposons de mener des expériences avec le modèle de l'équation (6.2), sans correction quadratique et avec des images de synthèse. Le but de ces expériences est d'évaluer l'efficacité de l'énergie proposée et de mesurer l'impact de l'effet d'arrondissement des coins. L'image de synthèse illustrée en figure 6.5 présente un objet en forme de U localement altéré par un effacement. Pour les

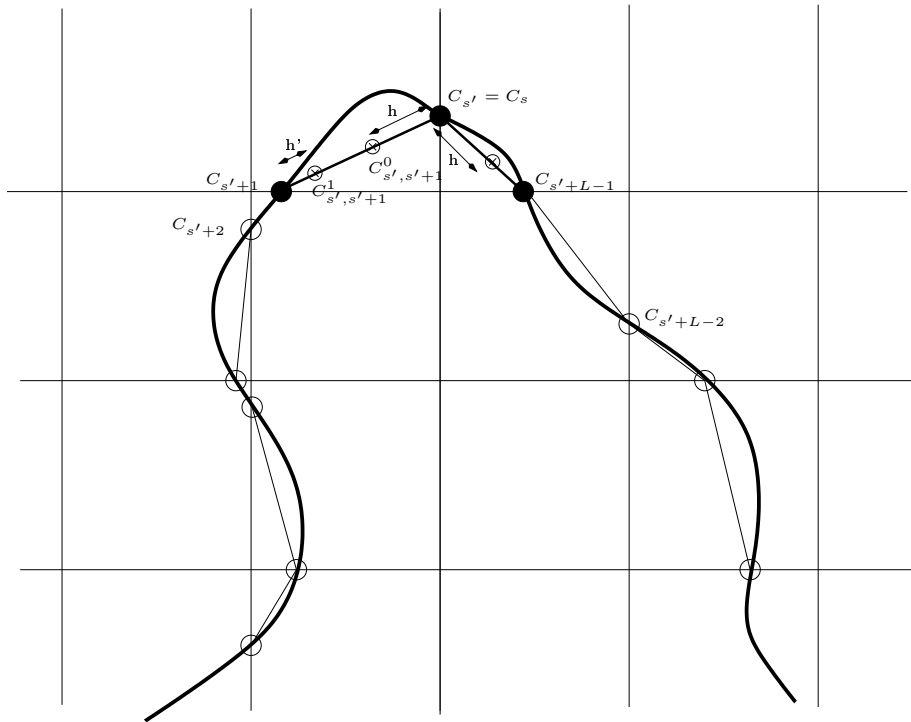


FIG. 6.4 – Schéma décrivant le mode d'échantillonnage avec une fonction de Bessel pour profil Ψ

expériences de cette section, le contour initial est carré en englobe l'objet d'intérêt (figure 6.5). La forme *a priori* utilisée a aussi une forme en U mais avec des différences locales par rapport à l'objet d'intérêt (figure 6.6). Pour faciliter la compréhension des résultats présentés en figures 6.7-6.9, nous affichons en bleu le polygone représentant la forme *a priori* sur l'image.

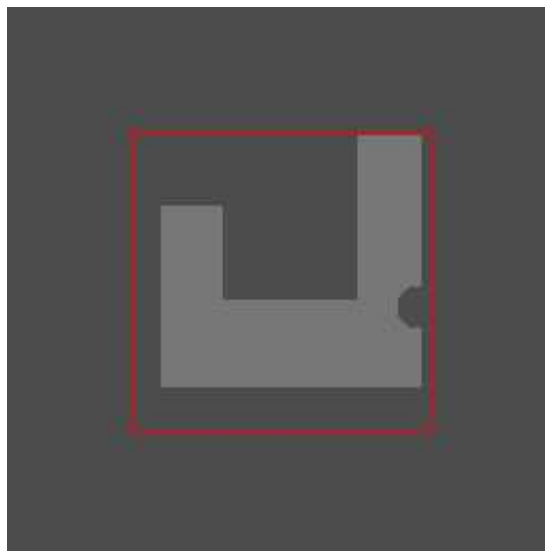


FIG. 6.5 – Image de synthèse avec le contour initial représenté en rouge.

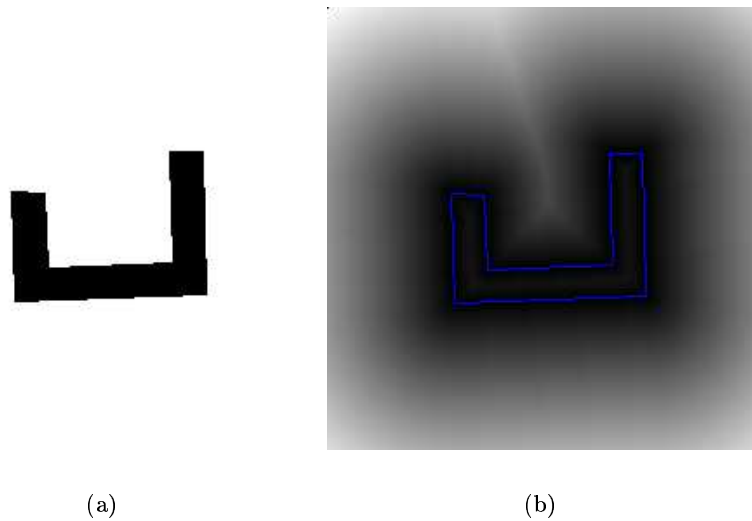


FIG. 6.6 – (a) Forme *a priori* et son ensemble de niveaux associé (b).

L'expérience de la figure 6.7 montre le résultat obtenu avec une contrainte de forme classique exprimée en équation (4.54) et le modèle d'attache aux données de Chan et Vese.

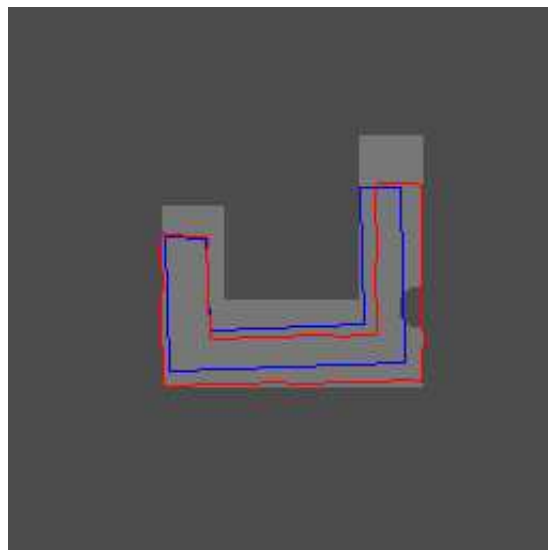


FIG. 6.7 – Résultat avec la contrainte de forme classique

On constate qu'avec une contrainte de forme classique, le contour final est un compromis entre l'attache aux données et la forme de référence à une similitude plane directe près. Aucune déviation locale par rapport à la référence n'est permise. Les expériences des figures 6.8-6.9 illustrent les résultats avec la nouvelle énergie de contrainte et ont été réalisées avec différentes valeurs de λ_{para} afin d'évaluer son impact sur la contrainte de forme. Comme le montrent les figures 6.8-6.9, la forme de contrainte avec le nouveau terme d'énergie permet de s'éloigner de la forme *a priori*, afin de segmenter plus précisément l'objet de l'image.

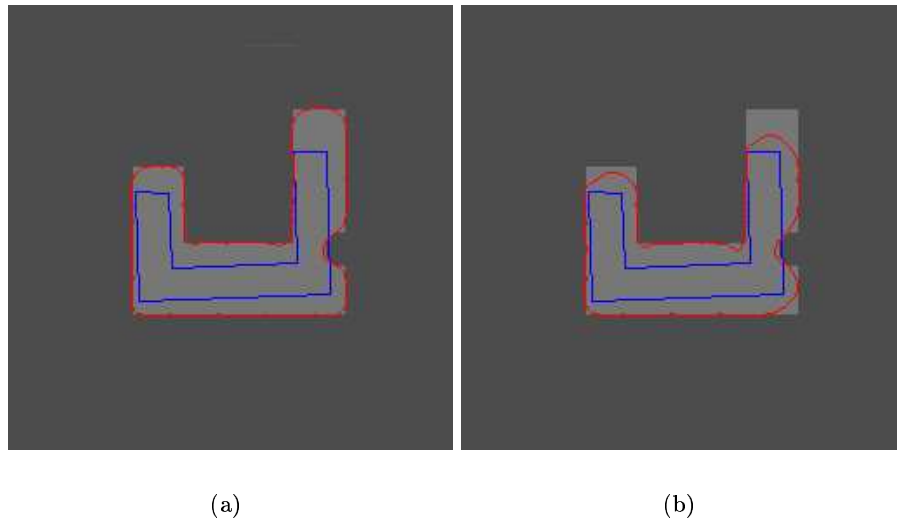


FIG. 6.8 – Segmentation avec contrainte de forme autorisant les variations parallèles : (a) $\lambda_{para} = 5$; (b) $\lambda_{para} = 10$

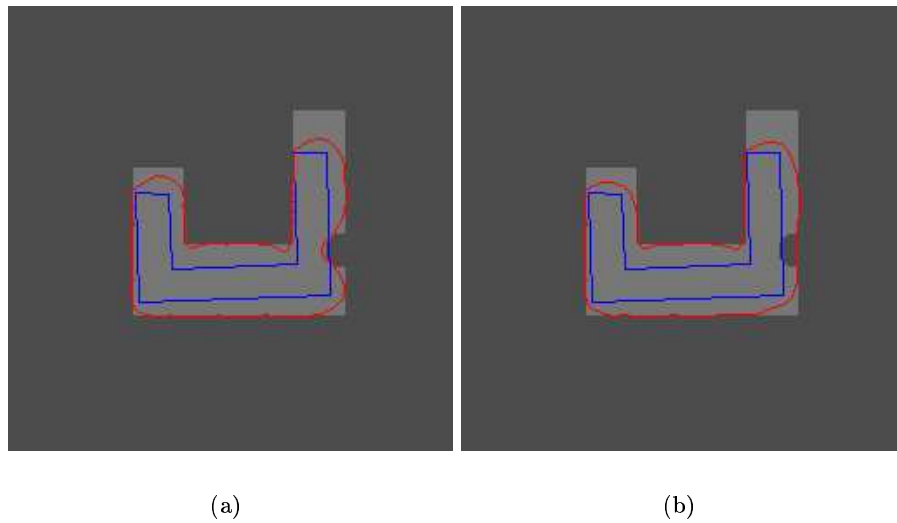


FIG. 6.9 – Segmentation avec contrainte de forme autorisant les variations parallèles : (a) $\lambda_{para} = 15$; (b) $\lambda_{para} = 30$

La barre centrale du U est désormais bien extraite, même si ses bords sont éloignés de la forme *a priori*. On remarque que naturellement, plus le poids λ_{para} est élevé, plus le modèle est à même de s'affranchir des altérations de l'image (effacement de l'objet représenté dans l'image dans le cas présent). Cependant, cette nouvelle énergie introduit des effets indésirables tel que l'arrondissement du contour au niveau des coins extérieurs ; l'effet de distorsion duale est aussi observable sur les coins intérieurs de la référence. Dans le cas des coins externes, les lignes de niveau sont courbes à l'extérieur de la forme implicitement représentée. Ainsi, la contrainte de forme de longue portée impose au contour de respecter cette courbure non nulle et induit l'arrondissement du contour actif.

6.4.2 Résultats avec correction quadratique

Les expériences menées avec la force quadratique sont menées de la façon suivante :

1. Le contour actif initial est semblable à la forme *a priori* (représentée en bleu sur les images) et évolue selon le modèle de l'équation (6.2) sans correction quadratique. Le modèle d'attache aux données est celui de Chan et Vese.
2. Après n_0 itérations, la force de correction quadratique est activée.

Il y a deux raisons pour un tel protocole expérimental : i) cela permet de réduire les temps de calcul. La force quadratique est une intégration à effectuer pour chaque point du contour, ce qui augmente de façon drastique la complexité calculatoire. ii) la force quadratique semble neutraliser la progression du contour. Cet effet n'est actuellement pas compris. L'une des raisons serait le terme proportionnel à la courbure de l'équation (6.13) qui tend à réduire la longueur du contour actif.

La figure 6.10 compare le résultat obtenu sans et avec correction quadratique avec un objet synthétique parfait. On constate en figure 6.10.a l'effet d'arrondissement des bords remarqué en section précédente. Ces arrondissements sont atténués avec l'insertion de la force quadratique de régularisation (figure 6.11.b). On remarque en particulier que le contour actif est rectiligne sur les parties supérieures des deux branches de l'objet en U, les parties intérieures sont aussi correctement segmentées.

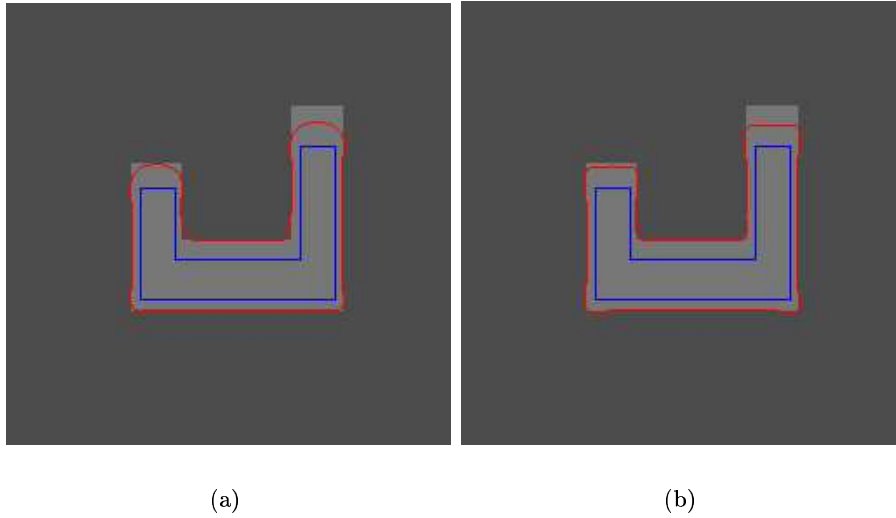


FIG. 6.10 – Segmentation avec le modèle d'attache aux données de Chan et Vese et la contrainte de forme autorisant les variations parallèles : (a) sans correction quadratique et $\lambda_{para} = 20$ à l'instant n_0 ; (b) avec correction quadratique, $\lambda_{para} = 20$, $\lambda_{quad} = 8$ et $a = 0.5$.

La figure 6.11 compare le résultat obtenu sans et avec correction quadratique avec un objet synthétique altéré. Nous constatons la même atténuation des effets d'arrondissement. En revanche, nous remarquons que l'effacement est moins bien corrigé. La figure 6.12 permet de comprendre ce résultat : la partie du contour actif située au voisinage de l'altération reçoit une force négative qui aura tendance à repousser ce dernier vers l'intérieur. Ainsi, le terme d'attache aux données ainsi que la force quadratique conduisent le contour à se contracter et vainquent la force dérivée de la contrainte de forme. Une solution pour résoudre ce problème est d'augmenter le poids de la contrainte de forme λ_{para} , cependant, le contour ne serait alors pas en mesure de s'écarter de la forme *a priori* de façon significative.

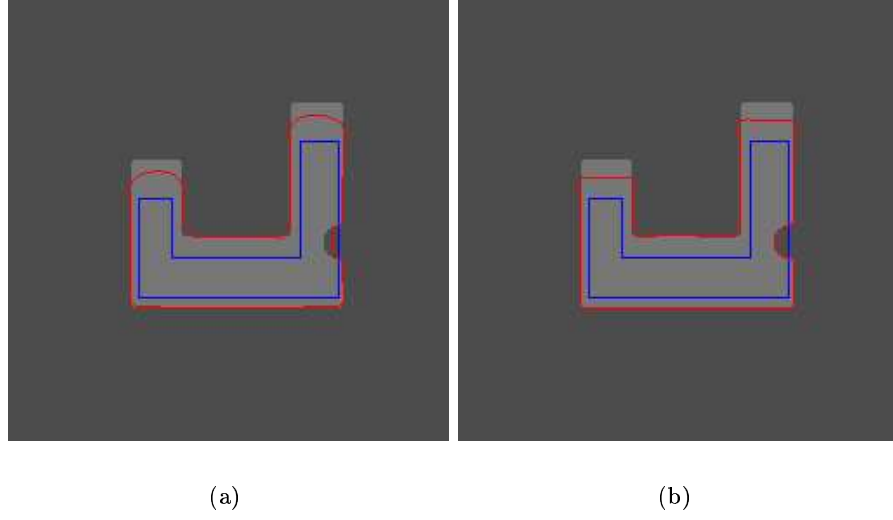


FIG. 6.11 – Segmentation avec le modèle d’attache aux données de Chan et Vese et la contrainte de forme autorisant les variations parallèles : (a) sans correction quadratique et $\lambda_{para} = 50$ à l’instant n_0 ; (b) avec correction quadratique, $\lambda_{para} = 50$, $\lambda_{quad} = 7$ et $a = 0.5$.

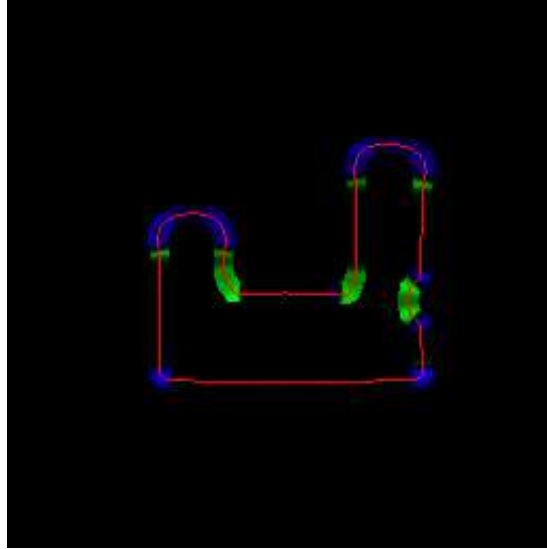


FIG. 6.12 – Représentation de la force quadratique à l’instant n_0 . Les zones bleues correspondent à une force positive (effet de dilatation), celles en vert représentent une force négative (effet de compression).

D’une manière générale, il s’est avéré difficile de régler les poids respectifs à la contrainte et au terme quadratique (λ_{para} , λ_{quad}) par rapport à la force d’attache aux données. L’exemple de la figure 6.13 avec une image réelle illustre cette difficulté. Le bâtiment de l’image étant peu discriminé du reste de la scène, il a fallu diminuer les poids de contrainte (λ_{para} , λ_{quad}). Dans cette expérience, la forme *a priori* est le résultat du recalage du polygone cartographique issu de la carte avec la contrainte de forme classique utilisée au chapitre précédent. Nous remarquons que le résultat avec l’image réel n’est pas satisfaisant. Des oscillations sont visibles le long du contour qui semble sensible aux artefacts de l’image. Nous avons déjà remarqué cette sensibilité avec l’image de synthèse.



FIG. 6.13 – Expérience avec une image réelle, la forme *a priori* est le résultat du recalage avec la contrainte de forme classique. La segmentation est effectuée avec le modèle d'attache aux données de Chan et Vese, la contrainte de forme autorisant les variations parallèles et avec correction quadratique $\lambda_{para} = 4$, $\lambda_{quad} = 0.5$ et $a = 0.5$.

6.5 Conclusion

Nous avons présenté un nouveau modèle énergétique de contrainte de forme. Celui-ci, plus souple, permet des variations locales du contour actif par rapport à la forme *a priori*. Le modèle proposé comprend une énergie linéaire limitant les différences entre le contour et la forme *a priori* à appartenir à la classe des variations parallèles. Une seconde énergie quadratique impose la rectilinéarité du contour et favorise la formation de coins. Ces caractéristiques sont adaptées aux objets fortement polygonaux tels que les bâtiments représentés dans les images de télédétection. Des expériences ont montré des résultats encourageants obtenus avec des images de synthèse. Cependant, l'application du modèle aux images réelles est à ce jour infructueuse. De nombreux aspects relatifs à l'énergie quadratique restent incompris, en particulier le réglage de la pondération de la force quadratique par rapport à la contrainte de forme et à l'image, la forme du profil d'interaction ainsi que sa taille. Toutes ces questions actuellement sans réponses influencent la stabilité numérique du modèle.

Chapitre 7

Analyse de changements carte-image

7.1 Introduction

Le recalage fin par contours actifs nous a permis de réduire les variabilités entre l'objet cartographique et sa représentation dans l'image. Les variabilités que nous avons proposé de corriger au fil des chapitres précédents sont caractérisées soit par une transformation de type similitude, soit un déplacement normal par rapport aux segments composant l'objet cartographique. Ces variabilités sont signes de différences relativement mineures entre la carte et l'image, et peuvent être issues ou non de réels changements entre les deux données. Nous exposons dans ce chapitre une méthodologie pour l'analyse des changements entre une carte de bâtiments et des données de télédétection multi-sources. L'approche que nous proposons se fonde sur la fusion de traitements de la donnée altimétrique, multispectrale et panchromatique et se scinde en deux étapes. La première fusionne des indices de changement destinés à valider uniquement la disparition univoque d'un bâtiment. Ces indices sont calculés à partir de méthodes simples utilisant l'image multispectrale et altimétrique. Dans la majorité des cas où un changement n'a pu être identifié de façon évidente, une seconde phase plus complexe est mise en œuvre afin de caractériser le degré de ressemblance entre l'objet cartographique et sa représentation dans l'image satellitaire panchromatique. Ce traitement tire parti du recalage fin par contours actifs et permet la formulation d'une probabilité de non-changement pour chaque élément de la carte. Des résultats expérimentaux illustreront l'efficacité de la méthode proposée et en particulier la contribution du recalage fin à accroître la confiance accordée à la décision de non-changement. Enfin, nous examinerons la sensibilité de l'approche proposée à des erreurs locales de forme contenues dans la carte grâce à une analyse statistique des résultats.

7.2 Méthodologie

La méthode de détection de changement que nous proposons suppose une superposition globale de la carte avec les données de télédétection et le modèle numérique de surface généré (MNS). Dans notre cas, cette mise en correspondance est accomplie grâce à l'information de géocodage des données (figure 7.2).

Nous proposons ensuite deux niveaux de détection de changement :

1. Le premier opère par fusion d'indices de bas niveau calculés à partir des données auxiliaires (MNS et image multispectrale) susceptibles d'être acquises à des temps différents mais nécessairement plus récentes que la réalisation de la carte (figure 7.1, phase A). Le résultat de cette première étape est de confirmer uniquement des *changements univoques* au niveau du bâti mais ne permet pas de décider des *non-changements*. Les bâtiments non détectés comme

ayant changé sont considérés comme “indéterminés” et nécessitent le traitement plus sophistiqué de la seconde phase.

2. Si le résultat final de la première phase ne peut pas confirmer de changement, nous cherchons en premier lieu à atténuer les variabilités exogènes géométriques de faible amplitude entre l’objet cartographique et sa représentation dans l’image satellitaire panchromatique. Ceci est achevé au moyen des contours actifs présentés aux chapitres précédents (figure 7.1, phase B). Nous proposons alors de nouveaux indices de changement, plus élaborés que ceux de l’étape 1. Le premier indice est une mesure dans un espace de Hough de la ressemblance des segments extraits de l’image et de ceux de l’objet cartographique recalé. Le second évalue la variation géométrique entre l’objet cartographique initial et son repositionnement après le recalage par contours actifs. Ces indicateurs sont fusionnés sous la forme d’une énergie qui permet d’exprimer une probabilité de non-changement dans le cadre probabiliste de Gibbs (figure 7.1, phase C). Ce résultat est un indicateur de changement qui pourra être utilisé pour faciliter et accélérer une procédure de mise à jour de carte par un opérateur.

La méthodologie proposée est synthétisée en figure 7.1.

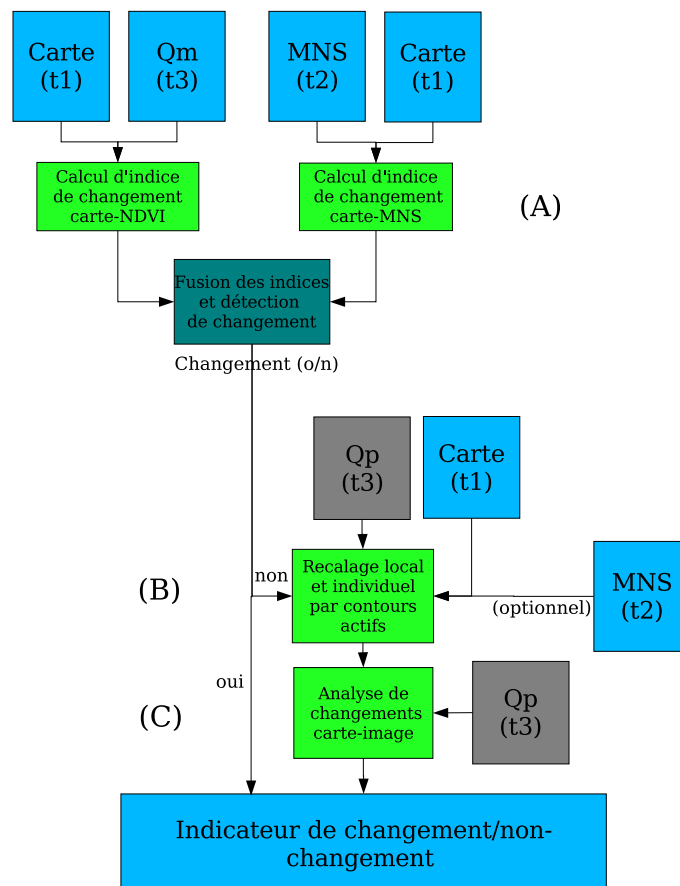


FIG. 7.1 – Méthodologie proposée pour la détection de changement : (A) calcul et fusion d’indices de changement. (B) atténuation des variabilités exogènes carte-image (C) analyse de changements finale. Légende : Qm et Qp sont respectivement l’images satellitaire multispectrale et panchromatique.

7.3 Calcul et fusion d'indices simples de changement

Cette section présente deux méthodes simples permettant de calculer des indices confirmant des changements univoques au niveau du bâti à partir des données auxiliaires de télédétection. La première compare les éléments de la carte avec la donnée altimétrique (MNS) et produit un indice de changement issu de tests statistiques et de seuillages d'altitudes. La seconde utilise l'image multispectrale satellitaire. Une classification par maximum de vraisemblance permet de confirmer la présence de végétation et ainsi l'absence de bâti. Les deux indices produits sont alors fusionnés afin de tirer parti de la complémentarité de la représentation altimétrique et multispectrale.

7.3.1 Indice calculé à partir de la donnée altimétrique (MNS)

7.3.1.1 Limitations de la comparaison carte-MNS

Une méthode de détection de changement entre une carte et un MNS est intrinsèquement restreinte. En effet, un MNS ne fournit qu'une information d'altitude, ce qui est loin d'être exhaustif pour renseigner la réalité d'une scène. Ainsi, il sera difficile de distinguer les bâtiments de la canopée avec la donnée unique d'un MNS issu d'une reconstruction 3D (une analyse d'imagerie LIDAR permet au contraire de faire la différence avec la végétation).

Plus spécifiquement, le MNS que nous avons généré contient des pixels non renseignés inhérents à la méthode de stéréo-corrélation ayant servi à le construire. Ces pixels non renseignés représentent 30% à 40% de l'image et sont localisés au voisinage des bâtiments en raison des occlusions. De plus, ces zones localement indéterminées auront été susceptibles d'être agrandies par les pré-traitements des images de disparité exposés en section 3.4.2. Cette absence d'information d'altitude est une limitation supplémentaire pour la détection de changement qui empêchera la prise de décision. Pour ces raisons, nous avons décidé de concevoir une méthode simple de pré-détection de changement carte-MNS. Une approche plus sophistiquée ne serait pas forcément plus efficace en raison du caractère incomplet de l'information délivrée par le MNS.

Enfin, les erreurs de superposition carte-MNS devront aussi être prises en compte. Il paraît difficile d'appliquer le recalage local par contours actifs avec le MNS tant celui-ci est bruité. Un recalage effectué avec l'image satellitaire panchromatique pour éliminer les variabilités exogènes carte-MNS paraît problématique car ces données sont exogènes : leurs géométries sont différentes ainsi que leurs dates d'acquisition.

7.3.1.2 Approche

L'algorithme que nous proposons s'appuie sur le calcul de critères statistiques globaux de moyenne, médiane et variance à partir des valeurs d'altitude du MNS contenues dans un polygone cartographique ou dans son voisinage. La méthode implantée se limite à deux cas de changements univoques qui correspondent aux situations *a* et *b* du tableau 2.2.

1. Bâtiment totalement détruit et remplacé par un sol nu et plat non incliné.
2. Bâtiment totalement détruit et remplacé par un bâtiment plus grand au toit plat et non incliné.

Nous nous limitons aux toits ou sols plats non inclinés car ces caractéristiques de planitude sont faciles à vérifier avec une mesure de la variance de l'altitude des pixels du MNS à l'intérieur d'un bâtiment de la carte (polygone).

La méthode de pré-détection de changement peut être décrite de la façon suivante : pour chaque polygone cartographique superposé au MNS, le pourcentage de pixels renseignés du MNS et inclus dans le polygone est calculé. Si ce pourcentage n'est pas assez élevé, aucune décision de changement n'est prise et le cas est considéré comme indécidable. Dans le cas contraire, la planitude altimétrique est vérifiée dans le polygone et sur son voisinage. Si la planitude est avérée, un test final mesure la variation d'altitude (au sens de la médiane) entre le polygone et son voisinage. Si cette variation est inférieure à la taille normale d'un bâtiment, un cas de changement sera détecté; sinon le cas est considéré comme ambigu. Le résultat final de cette pré-détection est un score binaire s_{MNS} qui vaut 1 lorsqu'un changement est validé. Les cas indécidables et ambigus sont regroupés pour être représentés par un score $s_{MNS} = 0$.

Si cette méthode est simple, elle se doit d'être fiable et ainsi de commettre le moins d'erreurs possibles. Nous évaluerons le succès de la méthode par rapport à ses différents paramètres. Ceux-ci sont au nombre de trois :

1. $seuil_{rens}$ pourcentage de pixels du MNS renseignés en altitude en dessous duquel la détection de changement n'est pas mise en œuvre.
2. $seuil_{plat}$ seuil de variance altimétrique (en mètres) en dessous duquel une zone du MNS est considérée comme plate et non inclinée.
3. $seuil_{bat}$ seuil altimétrique (en mètres) au dessus duquel un pixel du MNS est considéré comme appartenant au bâti.

L'algorithme de la méthode présentée est alors :

1. Pour chaque polygone du SIG superposé au MNS :
 - (a) Si le pourcentage de pixels non renseignés à l'intérieur du polygone, ou dans son voisinage $< seuil_{rens}$, alors la détection de changement n'est pas mise en œuvre. Le résultat est *indécidable*.
 - (b) Sinon si le polygone et son voisinage ont un écart type $< seuil_{plat}$ ET le voisinage a un taux de pixel non renseignés $< seuil_{rens}$ alors :
 - i. Si la différence d'altitude polygone/voisinage $< seuil_{bat}$, alors le bâtiment est considéré comme ayant *changé*.
 - (c) Sinon le cas est considéré comme *ambigu*
2. Fin.

Nous menons des calculs aussi bien sur le voisinage d'un polygone que sur son intérieur afin d'évaluer des différences relatives d'altitude. Une mesure absolue sous-entend la connaissance de l'altitude du sol de la scène, ce qui nécessite la donnée d'un MNT qui permet alors de normaliser le MNS. Sans la donnée du MNT, un critère de détection absolu aurait échoué si le relief du terrain de la scène n'était pas strictement plat (régions montagneuses ou vallonnées).

Comme nous l'avons précisé, la méthode se doit de tenir compte des imperfections éventuelles de superposition carte-MNS. Le voisinage d'un polygone dans le MNS est par conséquent défini comme étant la couronne résultant de la soustraction d'une dilatation du polygone par un élément structurant 15×15 pixels avec celle d'un élément de 3×3 pixels. Le voisinage n'est donc pas trop près du bord du polygone, évitant ainsi de considérer des pixels appartenant au bâtiment et qui fausseraient les calculs des critères statistiques de la détection de changement.

7.3.1.3 Résultats expérimentaux

La détection de changement illustrée en figure 7.3 a été menée sur 975 bâtiments avec les paramètres suivants :

- $seuil_{rens} = 10 \%$
- $seuil_{plat} = 1 \text{ m}$
- $seuil_{bat} = 3 \text{ m}$

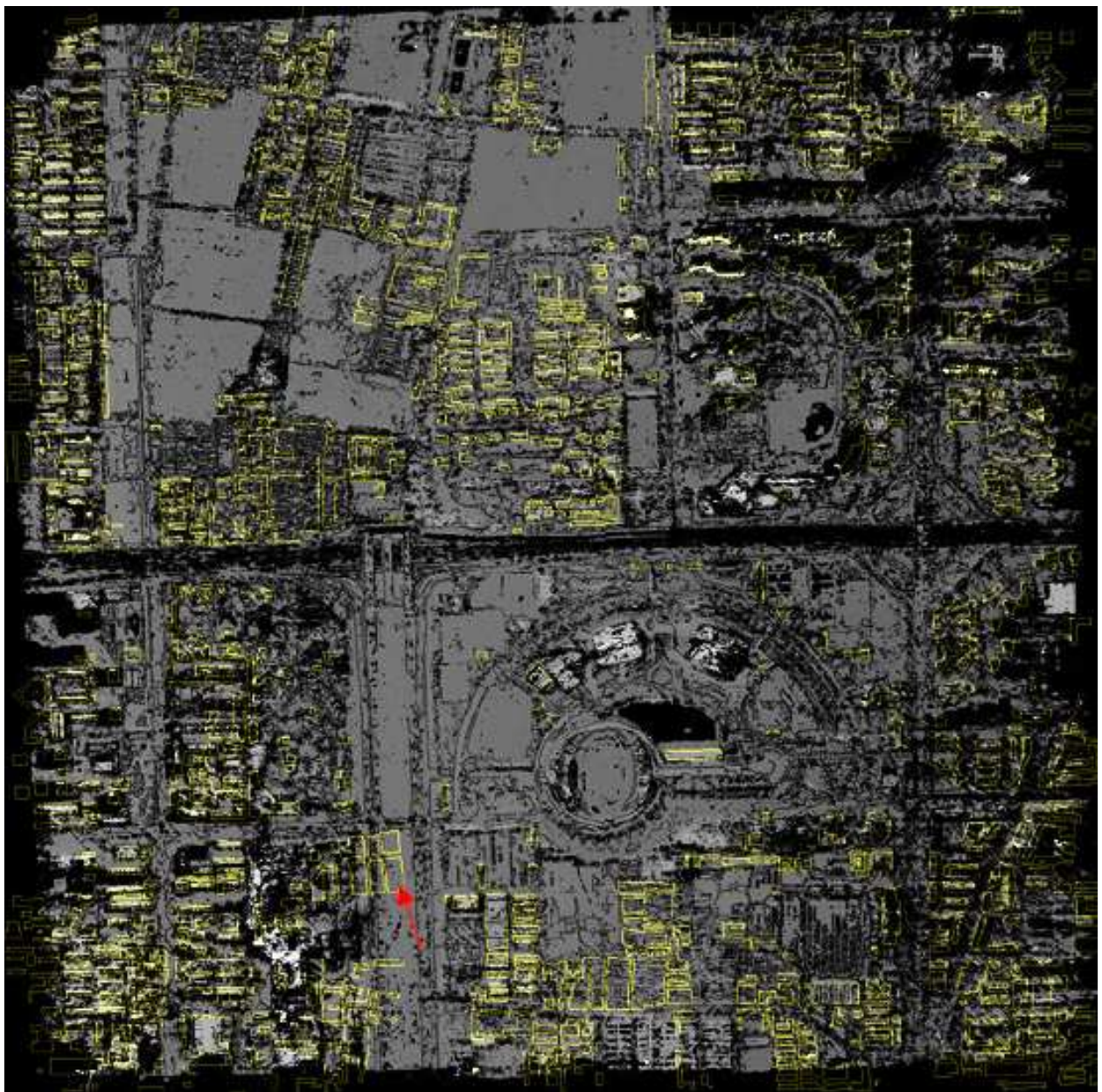


FIG. 7.2 – Superposition de la carte du bâti (en jaune) avec le MNT grâce à l'information de géocodage. La flèche rouge pointe un changement : bâtiment remplacé par un terre-plein.

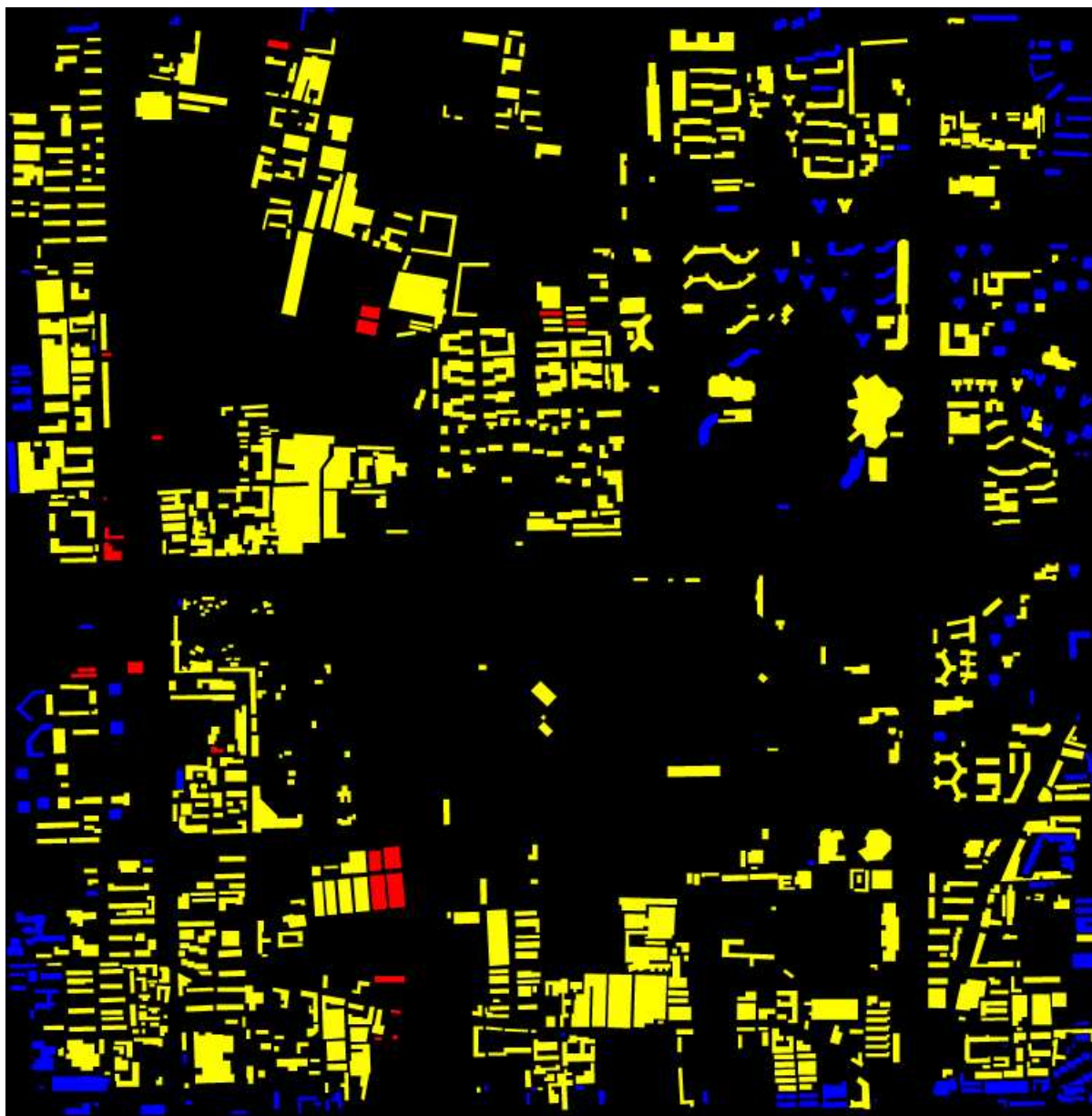


FIG. 7.3 – Résultat de la détection de changement carte-MNS. Légende : rouge (détection de changement), jaune (ambigu), bleu (indécidable). Paramètres : $seuil_{rens} = 10\%$, $seuil_{plat} = 1\text{ m}$, $seuil_{bat} = 3\text{ m}$.

Le choix des valeurs des paramètres est, comme dans toute méthode, une décision délicate. Cependant ceux-ci ont une valeur physique qui ne peut être choisie au hasard. Un seuil de platitude $seuil_{plat}$ supérieur à trois mètres n'a pas de sens. Fixer un seuil de détection de bâtiment trop élevé augmentera le risque de fausses alarmes car tout bâtiment dont l'altitude est inférieure au seuil se verra injustement classé comme ayant changé. Enfin, prendre des décisions avec un pourcentage de pixels déterminés trop faible se verra sanctionné par un risque accru d'erreurs de la méthode.

– Analyse variationnelle des paramètres

Nous avons mené une analyse du succès de la méthode en fonction de la variation de ses trois paramètres (cette analyse a été menée avec les données cartographiques SIG de 1996 illustrées en

figure 7.2). Systématiquement, deux paramètres sont fixés, un seul varie. L'évaluation du succès de la méthode se fait par examen visuel de la carte et d'une image aérienne ayant servi à réaliser le MNS. Les résultats sont résumés en figures 7.4 à 7.6.

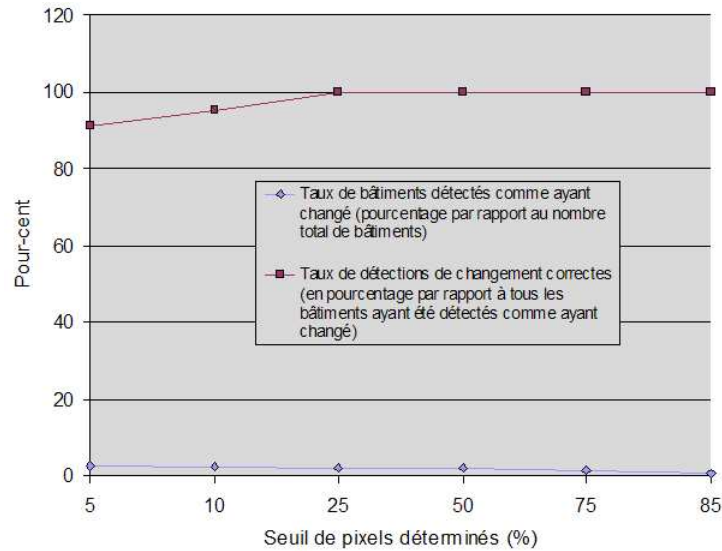


FIG. 7.4 – Succès de la méthode en fonction de $seuil_{rens}$. En rouge figure le taux de bâtiments correctement détectés comme ayant changé. La courbe bleue représente le taux de bâtiments détectés comme ayant changé par rapport au nombre total de bâtiments dans la carte. $seuil_{plat} = 1m$ et $seuil_{bat} = 3m$.

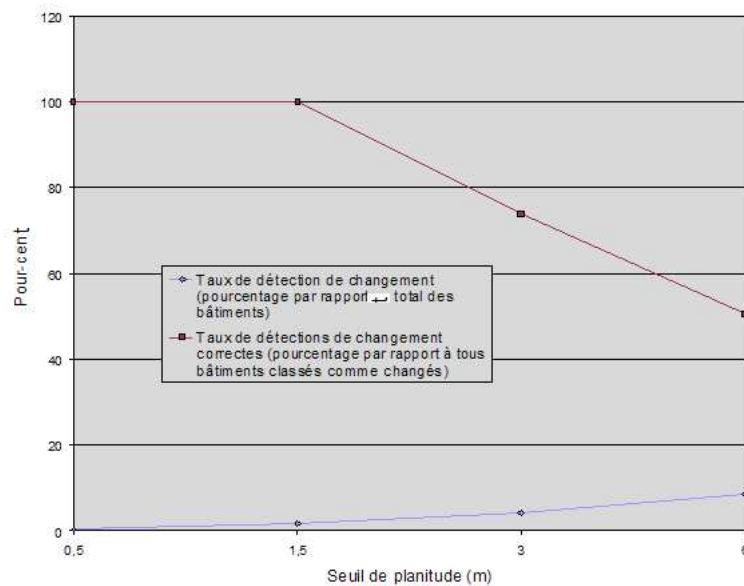


FIG. 7.5 – Succès de la méthode en fonction de $seuil_{plat}$. En rouge figure le taux de bâtiments correctement détectés comme ayant changé. La courbe bleue représente le taux de bâtiments détectés comme ayant changé par rapport au nombre total de bâtiments dans la carte. $seuil_{rens} = 50\%$ et $seuil_{bat} = 3m$.

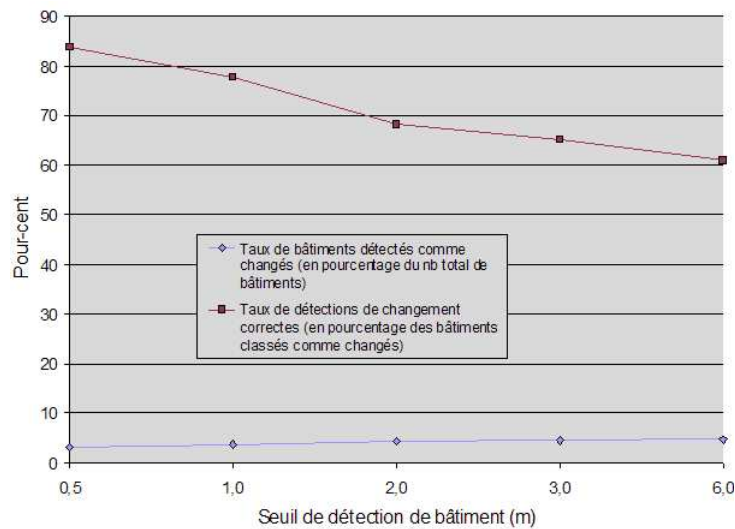


FIG. 7.6 – Succès de la méthode en fonction de $seuil_{bat}$. En rouge figure le taux de bâtiments correctement détectés comme ayant changé. La courbe bleue représente le taux de bâtiments détectés comme ayant changé par rapport au nombre total de bâtiments dans la carte. $seuil_{rens} = 65\%$ et $seuil_{plat} = 1.5m$.

Nous constatons que plus $seuil_{rens}$ augmente, plus le succès de la méthode s'accroît (figure 7.4). Plus $seuil_{plat}$ augmente, plus le succès de la méthode diminue. Il est à noter que la méthode semble être plus sensible vis-à-vis de ce paramètre que par rapport aux autres. Au delà de 1 m - 1.5 m pour $seuil_{plat}$, le succès de la méthode chute de façon critique (figure 7.5). Enfin, plus $seuil_{bat}$ augmente, plus le taux de fausses alarmes augmente, faisant chuter le taux de succès de la méthode (figure 7.6).

L'examen des valeurs numériques des graphiques montre le nombre peu élevé de bâtiments pour lesquels il y a prise de décision de changement. Cela est cohérent avec notre hypothèse initiale de restriction de la méthode à des cas limites de détection de changement mentionnés dans le tableau 2.2. Un problème de la méthode consiste à devoir considérer systématiquement le voisinage d'un polygone pour pouvoir faire le test de présence de bâti en comparant la différence de l'altitude entre le voisinage et le polygone avec $seuil_{bat}$. Si le voisinage n'est pas calculable à cause d'un taux trop élevé de pixels non renseignés (alors que les pixels à l'intérieur du polygone sont assez renseignés), le cas ne sera pas traité. Normaliser le MNS avec un MNT serait une solution pour s'affranchir du calcul du critère relatif d'altitude et de son problème intrinsèque. On pourrait alors utiliser un critère absolu ne testant que l'altitude médiane des pixels inclus dans le polygone par rapport à un seuil.

En conclusion de cette section, nous pouvons nous interroger sur les facteurs influençant la méthode de détection de changement carte-MNS proposée, et examiner si celle-ci est généralisable à d'autres sites que celui de Pékin. Un premier facteur important est le relief de la scène. Si le relief est trop "tourmenté", les conditions de planitude du sol ne seront jamais remplies et par conséquent, les prises de décision seront très peu nombreuses. La méthode n'est donc pas adaptée à des scènes de paysage vallonné ou montagneux, à moins d'avoir auparavant normalisé le MNS par un MNT. Le terrain de Pékin étant très plat, nous avons pu nous affranchir de cette étape de normalisation. Un deuxième facteur relatif à la culture architecturale de la ville concerne la forme des bâtiments. La méthode ne peut que détecter des changements de type *b* avec des bâtiments à toit plat (tableau 2.2). Une ville ne possédant que des bâtiments aux toits bi-pentes et contigus ne se prêterait pas de façon satisfaisante à cette détection de changement.

7.3.2 Indice calculé à partir d'une image multispectrale

Les images multispectrales que nous possédons ont une bande dans le proche infra-rouge qui permet de caractériser la végétation du paysage. En faisant l'hypothèse qu'il n'y a pas de recouvrement entre la végétation et le bâti, il nous sera alors possible de confirmer la disparition d'un bâtiment aux endroits de la carte qui sont recouverts de végétation dans l'image. Comme pour la méthode exposée précédemment, il ne sera pas possible de confirmer un non-changement puisqu'une absence de végétation n'implique pas la présence d'un bâtiment (qui aurait pu être remplacé par une route à titre d'exemple).

Le calcul d'indice de changement entre l'objet cartographique et sa représentation dans l'image multispectrale doit ainsi répondre à deux questions :

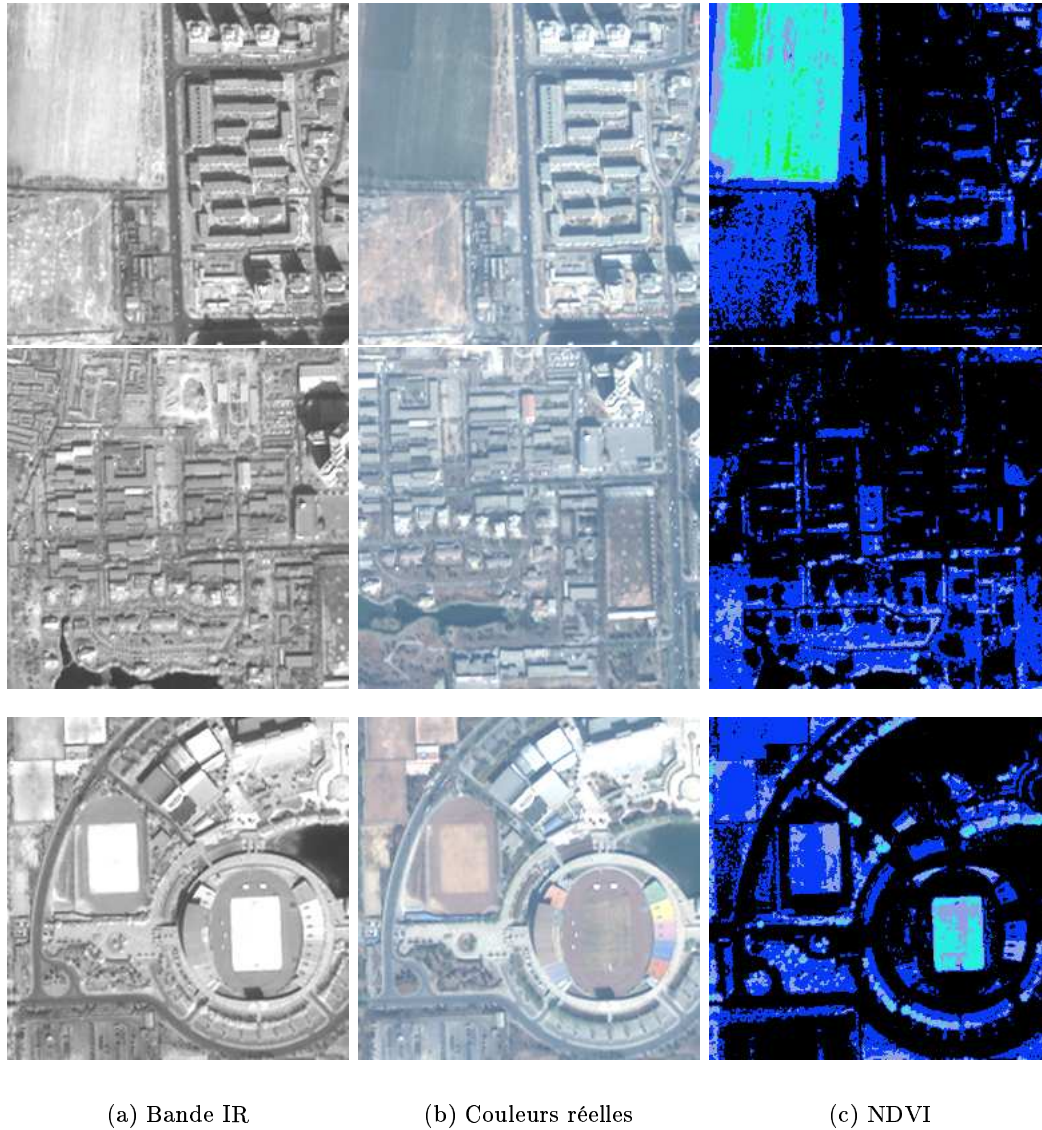
1. Comment caractériser la végétation de la scène ?
2. Comment prendre une décision de changement fiable à partir de la végétation détectée ?

Caractérisation de la végétation. Nous utilisons un indice classiquement utilisé en télédétection depuis 1969 : le *NDVI* (*Normalized Difference Vegetation Index*). Il consiste à faire la différence entre la bande infra-rouge et la rouge, puis de normaliser ce terme par la somme des deux bandes :

$$NDVI = \frac{IR - R}{IR + R}$$

Cet indice varie entre -1 et 1. La végétation a une forte absorption dans le rouge, et une forte réflectance dans l'infra-rouge. Ainsi, une valeur positive du *NDVI* confirme la présence de végétation. Plus la valeur est proche de 1, plus cette présomption est forte. Une valeur négative confirme l'absence de végétation. La figure 7.7 illustre les valeurs de cet indice selon onze classes. On constate que sur la figure 7.7 la végétation est plutôt bien caractérisée, avec néanmoins des valeurs peu élevées du *NDVI* (sur les exemples illustrés, ce dernier ne dépasse pas 0.4). Il existe plusieurs explications à cela : i) l'effet saisonnier a un fort impact sur la valeur du *NDVI* qui mesure la vigueur de la végétation. L'image Quickbird étudiée a été acquise en mars, juste avant le printemps. Ainsi, la biomasse aérienne est peu développée. ii) la densité de poussière ou de pollution (élevée à Pékin) influent aussi sur l'amplitude du *NDVI*.

Prise de décision avec l'indice NDVI. A l'instar de s_{MNS} , nous désirons formuler un critère binaire s_{NDVI} dont la valeur 0 signifie l'indétermination alors que 1 correspond à une forte densité de végétation à l'intérieur du polygone cartographique considéré, i.e. à une indication de changement. Nous utilisons une classification supervisée par maximum de vraisemblance pour déterminer s_{NDVI} . Ce type de méthode a l'avantage d'éviter les écueils des approches empiriques basées sur le seuillage de la valeur du *NDVI* pour caractériser la végétation. En revanche, la méthode est supervisée et nécessite un apprentissage. Dans notre cas, cet apprentissage peut être automatisé grâce à la couche "végétation" d'une carte numérique SIG (en supposant que peu de changements ont affecté la végétation entre la réalisation de la carte et l'acquisition de l'image).

FIG. 7.7 – Bande infra-rouge, couleurs réelles (RVB), et $NDVI$

Considérons n mesures indépendantes de l'indice $NDVI$ à l'intérieur du $j^{ième}$ polygone cartographique. Nous noterons y_i^j ces mesures :

$$y_{i \in \{1, \dots, n\}}^j = NDVI(\mathbf{x}_{i \in \{1, \dots, n\}})$$

où $\mathbf{x}_i$ est la position pixellaire du $i^{ième}$ pixel de l'image $NDVI$ appartenant au $j^{ième}$ polygone de la carte.

Soit $p(y_i^j|\theta)$ la densité de probabilité conditionnelle et normalisée d'observer la mesure y_i^j pour le jeu de paramètres θ donné. L'estimation des paramètres θ se fait par la maximisation de la fonction de vraisemblance :

$$\mathcal{L}(y^j|\theta) = \prod_{i=1}^n p(y_i^j|\theta) \quad (7.1)$$

où $\mathcal{L}$ est la fonction de vraisemblance mesurant la probabilité d'observer le jeu particulier de données y_i^j sachant θ . Dans notre cas, nous considérons les deux classes $\{C_k\}_{k \in \{0,1\}}$ avec C_0 et C_1 dénotant la non-végétation et la végétation respectivement. Selon notre hypothèse de départ, il n'y a pas de recouvrement entre bâti et végétation, ainsi, la classe C_1 est associée au changement des bâtiments, alors que C_0 correspond à l'indétermination.

Nous supposons par ailleurs que $p(y_i^j|\theta)$ suit une loi normale de moyenne m et de variance σ^2 . Dans ce cas particulier, chaque classe est totalement caractérisée par ces deux quantités statistiques qui doivent être mesurées par apprentissage : $\theta_k = (m_k, \sigma_k^2)_{k=0,1}$. Dans le cas des expériences illustrées dans ce chapitre, les zones d'apprentissage ont été déterminées manuellement. Le problème de classification binaire d'un polygone cartographique revient donc à choisir la classe C_k dont la fonction de vraisemblance est maximum à la vue des données mesurées :

$$\mathcal{L}(y^j|\theta_k)_{k=0,1} = \frac{1}{(2\pi\sigma_k^2)^{n/2}} e^{-\frac{\sum_{i=1}^n (y_i^j - m_k)^2}{2\sigma_k^2}} \quad (7.2)$$

L'indicateur de changement relatif au NDVI pour le polygone cartographique j considéré est alors :

$$s_{NDVI}(j) = \arg \max_{k,k=0,1} \mathcal{L}(y^j|\theta_k) \quad (7.3)$$

7.3.3 Fusion de la prise de décision

Les calculs d'indices de changement basées sur le MNS et la végétation présentent de nombreuses similarités : ce sont des méthodes simples ne pouvant qu'infirmar la présence d'un bâtiment de la carte de façon binaire. Ce sont aussi des méthodes utilisant des données non redondantes et avec des résultats dont nous pouvons utiliser la complémentarité pour la prise de décision. Nous fusionnons les indices de changement carte-MNS et carte-NDVI selon la stratégie suivante :

1. Le MNS et l'image NDVI sont réalisés grâce à des données acquises à des dates différentes. Il ne semble néanmoins pas judicieux de faire intervenir la donnée temporelle pour mettre en balance les résultats des deux méthodes. En effet, il serait incorrect d'affaiblir ou renforcer une décision confirmée de changement en fonction de la date d'acquisition car un changement avéré perdure dans le temps. Il est préférable de déterminer la contribution d'une méthode en fonction de la confiance qui lui est accordée. Pour notre application, nous n'avons pas de connaissance *a priori* permettant d'affirmer que l'une ou l'autre est plus fiable, ainsi, nous leur accordons une confiance égale lorsque les deux méthodes confirment un changement.
2. Lorsque l'une de deux méthodes renvoie un résultat indéterminé alors que l'autre confirme un changement, nous choisissons de ne pas pénaliser celle ayant détecté le changement dans la fusion des résultats. Les deux méthodes basées sur la comparaison MNS-carte et NDVI-carte sont simples et traitent des cas limités de changement. Ainsi, une indétermination n'implique pas une forte présomption de non-changement. Elle est souvent due à une incapacité de la méthode à prendre une décision.
3. Lorsque les deux méthodes renvoient un résultat indéterminé, la fusion des résultats est aussi indéterminée.

En résumé, la fusion des indices de changement se formalise simplement sous la forme suivante pour l'élément cartographique j considéré et peut être considérée comme un "OU" logique :

$$s_{fusion,MNS-NDVI}(j) = \max(s_{MNS}(j), s_{NDVI}(j))$$

7.4 Analyse de changements carte-image

Alors que la section précédente ne traitait que des disparitions univoques des bâtiments symbolisés dans la carte, la méthode d'analyse de changement entre la carte et l'image satellitaire panchromatique est destinée à prendre également en compte les cas de non-changement. Cette méthode suppose qu'une mise en correspondance locale et individuelle de chaque bâtiment avec l'image soit effectuée au préalable par la technique des contours actifs (chapitre 5). Ce pré-traitement visant à réduire les variabilités exogènes carte-image accroît la cohérence de l'objet cartographique avec sa représentation homologue dans l'image si aucun changement n'est survenu. La méthode d'analyse de changements carte-image est ensuite appliquée sur chaque bâtiment cartographique recalé. Pour chaque objet cartographique considéré, la méthode proposée est basée sur deux types d'information :

1. **La mesure de la cohérence des primitives segment extraites de l'image et de celles issues de la carte.** La cohérence est mesurée par un vote dans l'espace de Hough qui estime la translation mettant en correspondance les deux représentations. Nous verrons que la valeur maximale de l'accumulateur de Hough est un indicateur de ressemblance entre l'objet cartographique et celui de l'image. Ceci nous permettra d'en déduire si un changement ou un non-changement a eu lieu. Les segments extraits de l'image sont le résultat d'un chaînage rectiligne des pixels de fort gradient de la sous-image encadrant le bâtiment cartographique considéré. Ces segments de l'image sont ceux utilisés par les contours actifs basés sur l'information de frontière (cf. section 4.2.1.3). Les segments de l'objet cartographique affiné par la technique des contour actifs est le résultat d'une vectorisation du contour final.
2. **Variation géométrique lors du recalage fin carte-image.** Une variation géométrique trop grande entre l'objet cartographique original et sa version affinée par contours actifs est aussi un signe de changement. Cette variation est mesurable grâce à la donnée des ensembles de niveaux ψ_{final} et ψ_0 qui représentent respectivement l'objet cartographique recalé (résultat du recalage fin) et l'objet cartographique initial.

Nous formulerons ces deux types d'information comme des termes d'attache aux données pour définir une énergie de Gibbs. Ceci nous permettra *in fine* de calculer une probabilité $p_{NC}(j)$ de non-changement pour l'objet cartographique j .

7.4.1 Mise en correspondance de segments carte-image par vote de Hough

7.4.1.1 Principe

Soit S et S' deux ensembles de segments. S est l'ensemble des segments d'un bâtiment cartographique considéré. Cet objet a été affiné par le recalage des contour actifs. S' est l'ensemble des segments extraits d'une sous-image entourant l'objet cartographique.

Nous proposons d'utiliser un espace de Hough pour accumuler l'évidence de la ressemblance entre S et S' [52]. Ce type de méthode ascendante est robuste et permet de s'affranchir de la contribution des segments extraits de l'image n'appartenant pas au bâtiment d'intérêt. Le principe de la méthode est de considérer individuellement chaque paire de segments carte-image et d'estimer la translation permettant de les superposer. Cette translation est projetée dans un accumulateur de Hough qui permet ainsi de déterminer la translation μ_{Hough} la plus probable permettant de superposer S et S' . La normalisation du maximum de l'accumulateur de Hough par le périmètre de l'objet cartographique fournit un score de ressemblance s_{Hough} entre S et S' variant entre 0 et 1. Si $s_{Hough} \approx 1$ et $\mu_{Hough} \approx \mathbf{0}$ alors la cohérence de l'objet cartographique avec l'image est avérée. En effet, le vote majoritaire est important et l'estimation de la translation est cohérente avec le fait

que l'objet cartographique affiné est recalé sur sa représentation dans l'image. Si $s_{Hough} \ll 1$ ou $\mu_{Hough} \neq 0$ alors une absence de ressemblance est détectée, ce qui correspond à un changement. Cette méthode basée sur le gradient de l'image sera sensible aux ombres des bâtiments. Il est probable que l'appariement se fasse sur l'ombre portée due au bâtiment et non sur ses bords propres. Nous filtrons donc au préalable les segments d'ombre de l'image avant l'estimation de la translation entre chaque paire de segments carte-image.

7.4.1.2 Algorithme

L'algorithme d'évaluation de s_{Hough} et μ_{Hough} est le suivant :

1. Soit $\mathbf{u}_{ombre}$ le vecteur normé connu *a priori* donnant le sens et la direction de l'ombre dans l'image. Ce vecteur a le même sens et la même direction que les rayons solaires. Il peut être calculé par la connaissance des paramètres de prise de vue et de la date d'acquisition de l'image de télédétection. Dans notre cas, nous le mesurons manuellement. Le vecteur mesuré est quasi-constant sur toute l'image.
2. Soit S et S' deux ensembles de segments de taille n et m respectivement.

$$S = \{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n\} ; S' = \{\mathbf{s}'_1, \mathbf{s}'_2, \dots, \mathbf{s}'_m\} \quad (7.4)$$

3. Pour chaque segment $\mathbf{s}_i \in S$:

(a) Pour chaque segment $\mathbf{s}'_j \in S'$:

- i. Si la moyenne de la radiométrie le long du profil $\mathbf{s}'_j + \mathbf{u}_{ombre}$ est inférieure à celle de $\mathbf{s}'_j - \mathbf{u}_{ombre}$ plus un seuil, alors :

A. Si les segments $\mathbf{s}_i$ et $\mathbf{s}'_j$ sont quasi colinéaires ($\frac{\langle \mathbf{s}_i, \mathbf{s}'_j \rangle}{|\mathbf{s}_i| |\mathbf{s}'_j|} \leq \cos \alpha_s$, avec α_s un seuil donné), on estime la translation minimale $\mu_{ij, \min}$ et maximale $\mu_{ij, \max}$ permettant de superposer chacune des extrémités des deux segments (figure 7.8).

B. Dans un accumulateur à deux dimensions (μ_x, μ_y) dont les axes représentent l'espace des translations possibles dans les directions Ox et Oy , on ajoute le segment $[\mu_{ij, \min}, \mu_{ij, \max}]$ avec un poids égal à $\min(|\mathbf{s}_i|, |\mathbf{s}'_j|)$, où $|\mathbf{s}_i|$ représente la longueur du segment $\mathbf{s}_i$. Pour rendre la méthode plus robuste aux variations locales des segments des deux représentations (incertitudes dues à la méthode d'extraction de segments, erreurs présentes dans l'une ou l'autre des représentations), le segment est élargi de l pixels pour former une bande dans l'accumulateur (figure 7.9).

(b) Fin

4. La translation μ_{Hough} est estimée à l'endroit du maximum de l'accumulateur. Ce maximum est normalisé par le périmètre P du polygone cartographique ($P = \sum_{i=1}^n |\mathbf{s}_i|$) pour obtenir s_{Hough} .

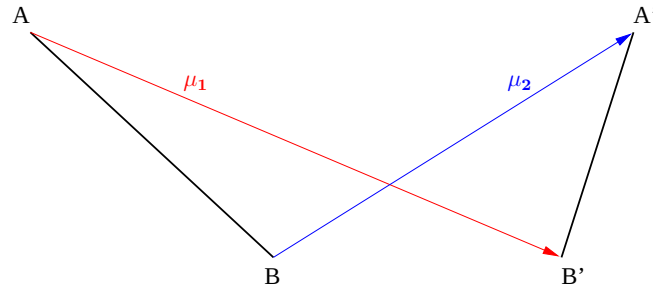


FIG. 7.8 – Estimation de la translation maximale μ_1 et minimale μ_2 entre le segment cartographique $[AB]$ et le segment-image $[A'B']$.

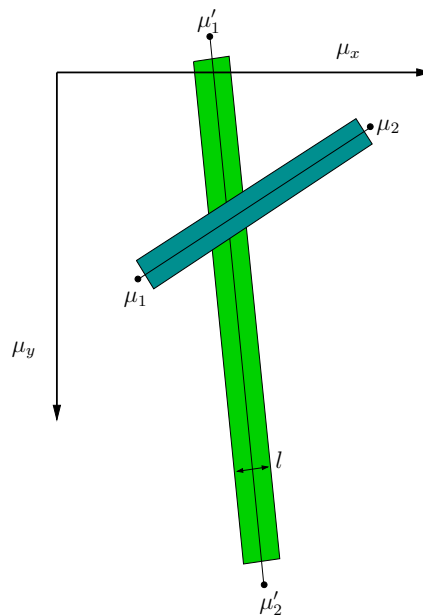


FIG. 7.9 – Accumulateur de Hough à 2 dimensions (μ_x, μ_y) . L'accumulation entre la translation minimale et maximale est élargie d'une bande de l pixels.

7.4.1.3 Énergie associée au vote de Hough

En vue d'insérer la décision du vote de Hough dans le cadre probabiliste de Gibbs, nous formulons une énergie E_{Hough} mesurant la cohérence entre un objet cartographique et l'image. L'énergie proposée est la suivante :

$$E_{Hough} = \frac{1}{P} \oint (1 - \alpha(\mathbf{C}_{\mu_{Hough}}(p))) dp \quad (7.5)$$

où P est le périmètre du polygone cartographique $\mathbf{C}$ paramétré par p . $\mathbf{C}_{\mu_{Hough}}$ est le contour $\mathbf{C}$ translaté par l'estimée μ_{Hough} . La quantité $\alpha(\mathbf{x})$ est telle que :

$$\alpha(\mathbf{x}) = \begin{cases} 1 & \text{si } I(\mathbf{x}) \in S'_{Hough} \\ 0 & \text{sinon} \end{cases} \quad (7.6)$$

où $I(\mathbf{x})$ est l'image satellitaire, et S'_{Hough} est l'ensemble des segments extraits de l'image ayant contribué au vote majoritaire dans l'espace de Hough, i.e. ayant contribué à l'estimation de μ_{Hough} . L'énergie de l'équation (7.5) sera minimale lorsque tous les pixels de l'image se trouvant sous la trace du polygone cartographique translaté auront contribué au vote de Hough. Dans le cas contraire, peu de pixels de l'image auront contribué au vote, ce qui est une indication de changement. L'énergie est alors maximale. La normalisation par le périmètre P permet de rendre l'énergie de l'équation (7.5) invariante par rapport à la taille du polygone cartographique considéré. L'énergie de l'équation (7.5) peut être écrite plus simplement :

$$E_{Hough} = 1 - s_{Hough} \quad (7.7)$$

où s_{Hough} est le score normalisé de Hough défini en section précédente.

Enfin, rappelons que la condition $E_{Hough} = 0$ n'est pas suffisante pas caractériser un non-changement. En effet, celle-ci doit être associée à une estimée μ_{Hough} de norme faible. Pour rendre compte de cette propriété, nous utilisons l'énergie associée à un ressort de raideur k qui pénalisera l'estimation d'une translation trop importante :

$$E_{\mu_{Hough}} = k |\mu_{Hough}|^2 \quad (7.8)$$

7.4.2 Mesure de la variation géométrique due au recalage par contours actifs

Lorsque la superposition initiale carte-image est de mauvaise qualité, l'invariance par transformation affine des contours actifs permet de corriger cet effet par rotation, translation et dilatation/réduction par facteur d'échelle. Cette correction par contours actifs fait partie de la réduction des variabilités exogènes indépendantes d'un changement effectif et ne devraient pas intervenir dans la détection de changement. Néanmoins, si la variation géométrique est trop importante, il se peut qu'elle ne soit pas imputable à la mauvaise qualité de la superposition initiale, mais plutôt à un changement (figure 7.11). Nous proposons de formuler un critère énergétique E_{geom} mesurant la variation géométrique entre l'objet cartographique initial et son transformé par $T_{sim,final}$ qui est la similitude plane estimée par le recalage des contours actifs :

$$E_{geom}(\psi_{final}, \psi_0) = \frac{1}{A_{\psi_{final}} + A_{\psi_0}} \int_{\Omega} (H(\psi_{final}(\mathbf{x})) - H(\psi_0(\mu_{persp} + \mathbf{x})))^2 d\mathbf{x} \quad (7.9)$$

où ψ_0 est l'ensemble de niveaux représentant le polygone cartographique initial ; ψ_{final} est l'ensemble de niveaux représentant l'objet cartographique transformé par $T_{sim,final}$: $\psi_{final} = \psi_0(T_{sim,final}\mathbf{x})$;

$A_{\psi_{final}} = \int_{\Omega} H(\psi_{final}(\mathbf{x})) d\mathbf{x}$ et $A_{\psi_0} = \int_{\Omega} H(\psi_0(\mathbf{x})) d\mathbf{x}$. Le terme énergétique de l'équation (7.9) est inspiré de celui utilisé pour la contrainte de forme *a priori* de l'équation (4.54). Deux différences sont cependant à noter. Premièrement, l'énergie E_{geom} est rendue indépendante de la taille des objets comparés grâce à la normalisation par la somme des aires incluses dans le polygone cartographique (A_{ψ_0}) et dans son transformé par $T_{sim,final}$ ($A_{\psi_{final}}$). Enfin, cette énergie est rendue invariante par translation μ_{persp} uniquement. Cette invariance est destinée à ne pas pénaliser les translations entre la carte et le contour actif qui sont colinéaires à la direction $\mathbf{V}$ de distorsion due à la perspective. Rappelons que l'image satellitaire n'est pas rectifiée, il existe ainsi une variation géométrique entre l'empreinte au sol du bâtiment (sur laquelle l'objet cartographique est susceptible d'être superposé) et le toit segmenté par le contour actif. Cette distorsion géométrique est intrinsèque à l'image et ne doit pas être interprétée comme un changement. Sur l'image satellitaire Quickbird, la distorsion peut être modélisée par une translation dont la direction est constante sur toute l'image. Le sens et la direction de la distorsion est donnée par le vecteur unitaire $\mathbf{V}$ que nous avons mesuré manuellement. Le vecteur μ_{persp} est alors défini de la façon suivante :

$$\mu_{persp} = w_{per} \mu_{final} \quad (7.10)$$

avec μ_{final} l'estimation de la translation entre le contour actif final et la forme de contrainte *a priori*. La forme *a priori* est similaire à l'objet cartographique initial et μ_{final} est estimée par l'algorithme du simplex (cf. section 5.2). Le poids w_{per} favorise l'invariance par μ_{final} uniquement si ce dernier est colinéaire à $\mathbf{V}$:

$$w_{per} = \begin{cases} \left\langle \frac{\mu_{final}}{|\mu_{final}|}, \mathbf{V} \right\rangle^2 & \text{si } \left\langle \frac{\mu_{final}}{|\mu_{final}|}, \mathbf{V} \right\rangle > 0 \\ 0 & \text{sinon} \end{cases} \quad (7.11)$$

Les variations locales et parallèles du contour actif (cf. chapitre 6) par rapport à la forme cartographique *a priori* témoignent aussi des incohérences carte-image. Dans ce cas, il est cependant difficile de discerner si la variation locale (au sens de l'objet cartographique) carte-image est due à un changement effectif représenté dans l'image, ou à une erreur présente dans la carte. C'est en raison du caractère mal posé de ce problème que nous avons décidé de ne pas comparer le contour actif final ayant permis le recalage avec l'objet cartographique (transformé ou non).

7.4.3 Probabilité de détection de non-changement

Nous proposons de fusionner l'information du vote de Hough et de la variation géométrique des contours actifs pour chaque bâtiment j symbolisé dans la carte. L'énergie totale mesurant le non-changement $E_{NC}(j)$ est alors :

$$E_{NC}(j) = \lambda_{Hough} E_{Hough}(j) + \lambda_{\mu_{Hough}} E_{\mu_{Hough}}(j) + \lambda_{geom} E_{geom}(j) \quad (7.12)$$

où λ_{Hough} , λ_{geom} et $\lambda_{\mu_{Hough}}$ sont des constantes réelles positives. Utilisant le formalisme des distributions de Gibbs pour définir une densité de probabilité de non-changement $p_{NC}(j)$ associée à $E_{NC}(j)$, nous pouvons écrire :

$$p_{NC}(j) = \frac{1}{Z} e^{-E_{NC}(j)} \quad (7.13)$$

où Z est la constante de partition. En définissant $p_{NC}(j) + p_C(j) = 1$ (où $p_C(j)$ est la probabilité de changement), la constante Z prend alors pour valeur 1. La probabilité $p_{NC}(j)$ est un indicateur de non-changement lorsqu'elle est supérieure à 0.5, sinon un changement est détecté.

7.5 Résultats expérimentaux

Nous montrons en section 7.5.2 quelques résultats expérimentaux pour différents cas de changements et de non-changements afin de mesurer l'efficacité de la méthode ainsi que d'illustrer les difficultés ne pouvant être surmontées. En section 7.5.3 nous proposons d'analyser statistiquement des résultats réalisés sur un plus grand nombre de cas et avec des cartes de qualités différentes.

7.5.1 Protocole expérimental

Toutes les expériences menées dans cette section ont été réalisées dans les conditions suivantes :

- **Données** : des cartes de 1996 et 2002 ont été utilisées. Lorsqu'une carte porte la mention "parfaite", cela signifie que les erreurs de la carte ont été corrigées manuellement. Dans ce cas spécifique, l'objet cartographique est cohérent avec sa représentation dans l'image ; nous simulons alors une erreur de recalage modélisable sous la forme d'une similitude plane directe de paramètres $\xi_0 = (s_0, \theta_0, \mu_{x,0}, \mu_{y,0})$. L'image satellitaire Quickbird (panchromatique et multispectrale) datant de 2002 a été utilisée dans les expériences.
- **Procédure** : Lorsque le test préliminaire basé sur l'information MNS-NDVI confirme un changement, le recalage fin par contours actifs et le vote de Hough ne sont pas mis en œuvre. Lorsqu'un changement n'a pas pu être validé par la méthode basée sur le MNS-NDVI, nous illustrons l'état initial de l'objet cartographique superposé à l'image ainsi que l'état final du recalage (les illustrations ne concernent que la section 7.5.2).
- **Initialisation des contours actifs** : la forme *a priori* ainsi que le contour actif initial sont similaires à l'objet cartographique. L'objet cartographique est superposé à l'image satellitaire plus récente grâce à l'information de géocodage des deux données.
- **Vote de Hough** : la taille de l'accumulateur utilisé dans les expériences est de : ($\mu_x = -30...30 \text{ pixels}$, $\mu_y = -30...30 \text{ pixels}$). La largeur l des segments ajoutés dans l'accumulateur est $l = 2 \text{ pixels}$.
- **Vectorisation** : pour pouvoir participer au vote de Hough, le contour actif final ayant permis le recalage fin doit être discrétisé sous forme de segments. Ce contour actif est originellement une liste de points issus de la détection des zéros de l'ensemble de niveaux le représentant. Ces points sont premièrement ordonnés en une liste $\{\mathbf{C}(i)\}$, puis chaînés pour former une séquence de segments. L'algorithme de chaînage consiste à former des segments $[\mathbf{C}(j), \mathbf{C}(j+k)]$ tant que la distance des points $\{\mathbf{C}(i)\}$ (situés entre $\mathbf{C}(j)$ et $\mathbf{C}(j+k)$) au segment $[\mathbf{C}(j), \mathbf{C}(j+k)]$ est inférieure à un seuil ε_{vect} . Dans notre application, la finesse de la vectorisation est déterminante pour une mise en correspondance efficace entre l'objet cartographique affiné et les segments extraits de l'image. Nous avons ainsi choisi un seuil restrictif $\varepsilon_{vect} = 0.3$.
- **Paramètres des énergies de non-changement** : Les paramètres de la détection de changement utilisés sont : $\lambda_{Hough} = \lambda_{\mu_{Hough}} = 1$, $\lambda_{geom} = 2$ et $k = 0.01$. La valeur de λ_{geom} reflète la connaissance *a priori* disponible sur la qualité de superposition initiale carte-image. Un poids élevé signifie une confiance élevée en la qualité, ainsi, une faible variation de E_{geom} se verra sanctionnée par une forte hausse de l'énergie totale de non-changement pour finalement diminuer la probabilité p_{NC} . La raideur du ressort k peut être fixée en fonction de l'incertitude associée à l'estimation de μ_{Hough} . Rappelons que les segments de l'accumulateur de Hough sont élargis de l pixels, ce qui détériore la qualité de l'estimation de la translation.

Nous choisissons une faible valeur de k pour ne pas pénaliser la décision finale de changement vis-à-vis de ces imprécisions.

– **Décision de (non)-changement :**

- Nous considérons un changement avéré si $s_{fusion,MNS-NDVI} = 1$ ou $p_{NC} < 0.4$.
- Nous considérons la détection d'un non-changement si $s_{fusion,MNS-NDVI} = 0$ et $p_{NC} > 0.6$.
- Nous considérons la prise de décision impossible (indétermination) si $s_{fusion,MNS-NDVI} = 0$ et $p_{NC} > 0.4$ et $p_{NC} < 0.6$.

7.5.2 Étude de quelques cas de changements/non-changements

7.5.2.1 Détection de changement

Les figures ci-après illustrent de correctes détections de changement menées avec la carte numérique de 1996. Dans les figures 7.10.a et 7.10.b, l'absence de bâtiment est validée par la pré-détection de changement mise en œuvre grâce à la donnée de l'image multispectrale. Ce sont des cas univoques de changement où les bâtiments ont été totalement détruits et remplacés par un sol couvert de végétation (cas *a* du tableau 2.2). Dans les deux cas, la classification par maximum de vraisemblance identifie les zones de l'image incluses dans le polygone cartographique comme de la végétation, ce qui infirme la présence de bâtiments : $s_{fusion,MNS-NDVI} = 1$.

Les deux cas suivants de changement ont été simulés (objets cartographiques créés manuellement). La prise de décision est impossible par les pré-traitements ($s_{fusion,MNS-NDVI} = 0$), elle résulte par conséquent de la méthode comparant uniquement la carte et l'image satellitaire pan-chromatique. La figure 7.11 illustre la contribution de la prise en compte de E_{geom} dans le modèle de prise de décision. Dans le cas illustré, le score final de Hough est très élevé car le nouveau bâtiment de l'image a la même forme que l'objet cartographique, ce qui tend à confirmer (à tort) un non-changement. Cependant, la variation spatiale est trop importante pour être imputable aux variabilités exogènes, E_{geom} est ainsi élevée et le bâtiment est détecté comme ayant changé ($p_{NC} = 0.3$).

Enfin, la figure 7.12 illustre un autre type de changement simulé et correctement détecté où un bâtiment est remplacé par un autre. Dans ce cas, le recalage par contours actifs échoue en raison de l'absence du bâtiment dans l'image. Le bâtiment initial ayant disparu, peu de primitives segments de l'image confortent l'hypothèse de la présence du bâti sous la trace de l'objet cartographique modifié par le contour actif. Ainsi, le vote de Hough désigne un autre emplacement où les segments de l'image sont plus cohérents avec la représentation cartographique. Alors que cet emplacement éloigné ($\mu_{Hough} = (-6, -26)$) est un meilleur candidat, le score du vote reste faible $s_{Hough} = 0.27$ puisqu'aucun bâtiment similaire n'est retrouvé dans l'image. La conjonction d'un score de Hough faible avec une translation estimée importante fait tendre p_{NC} vers 0 de façon drastique.

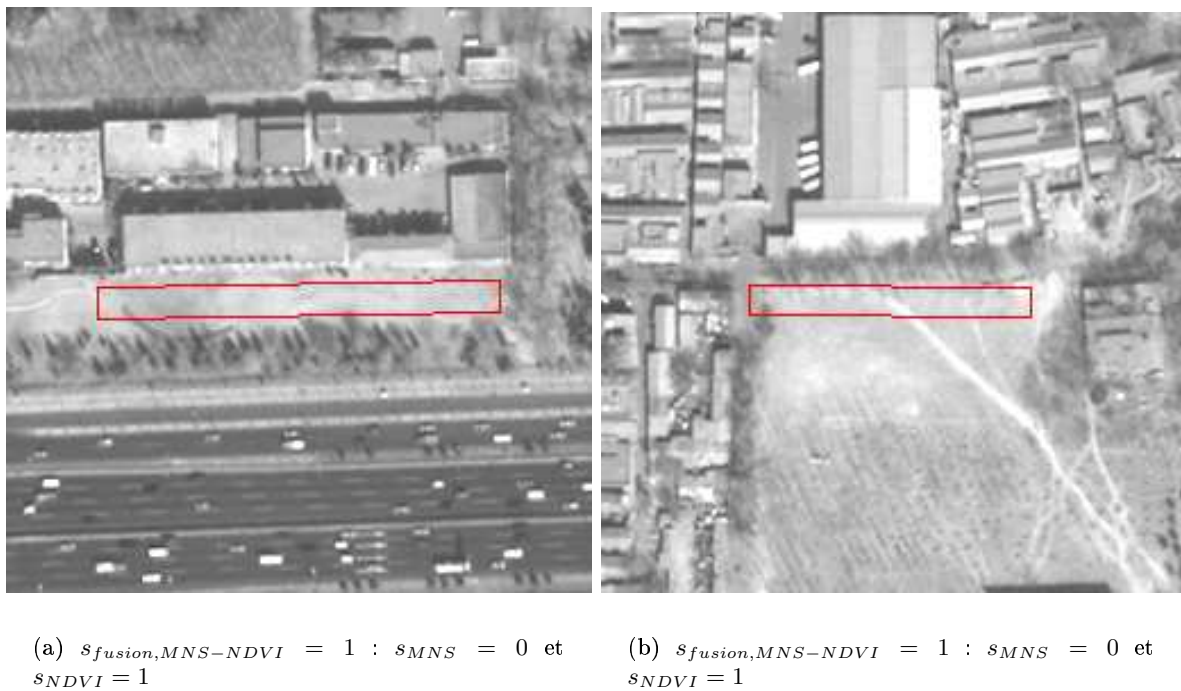


FIG. 7.10 – Détection de changements univoques. Les changements ont été détectés par la comparaison avec l'image NDVI. Les figures illustrent les objets cartographiques superposés à une image satellitaire plus récente.

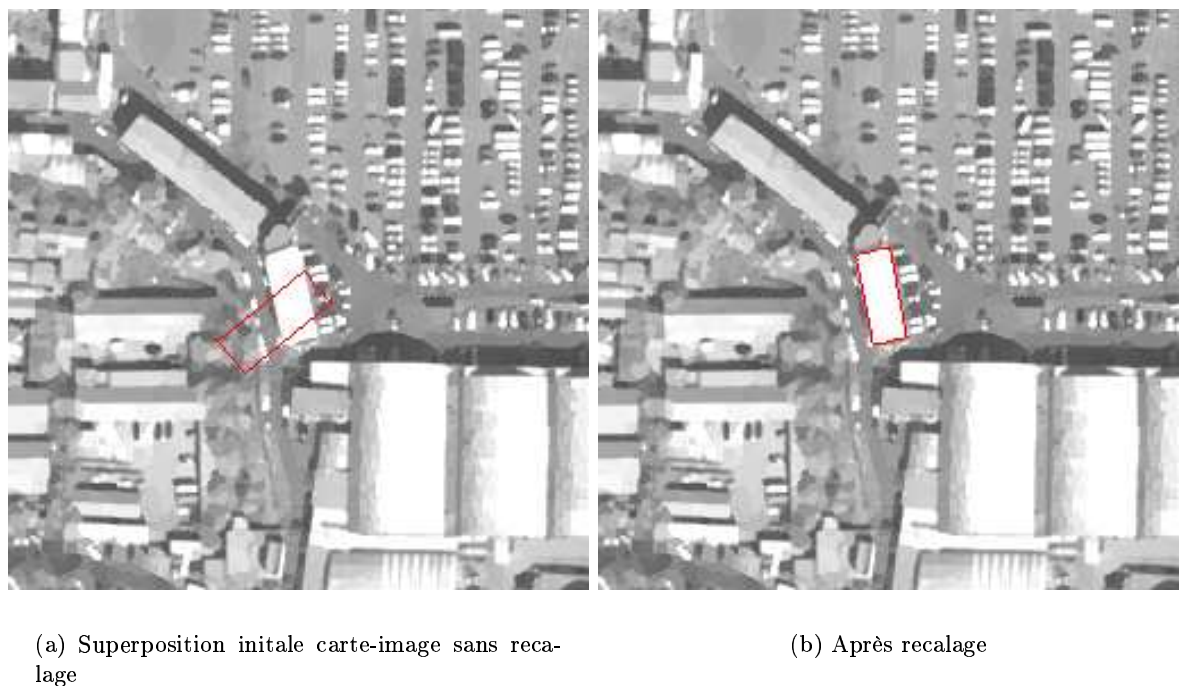
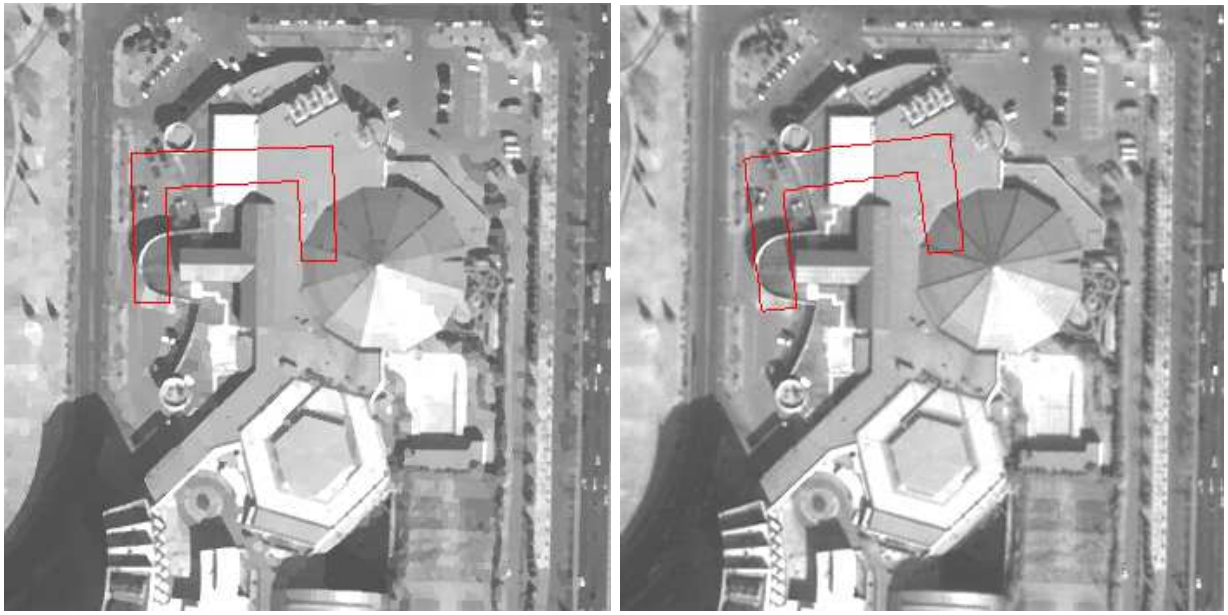


FIG. 7.11 – Détection de changement en raison d'une forte variation géométrique : $E_{geom} = 0.54$ et $p_{NC} = 0.30$. Le changement est simulé.



(a) Superposition initiale carte-image sans recalage

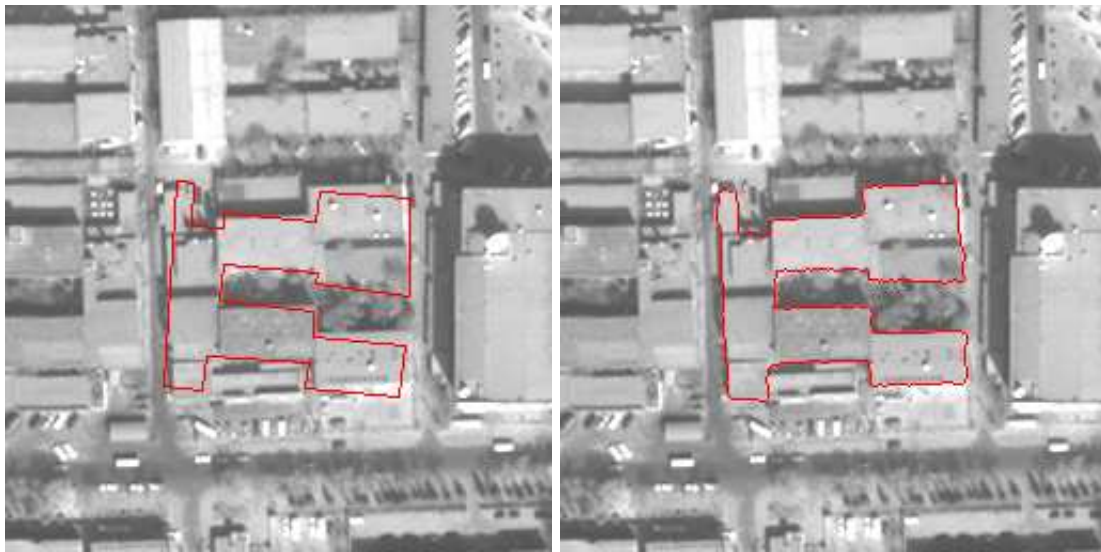
(b) Echec du recalage

FIG. 7.12 – Détection de changement en raison d'une translation μ_{Hough} élevée. $s_{\text{Hough}} = 0.27$, $\mu_{\text{Hough}} = (-6, -26)$ et $p_{NC} = 0.00$. Le changement est simulé.

7.5.2.2 Détection de non-changement

Carte parfaite

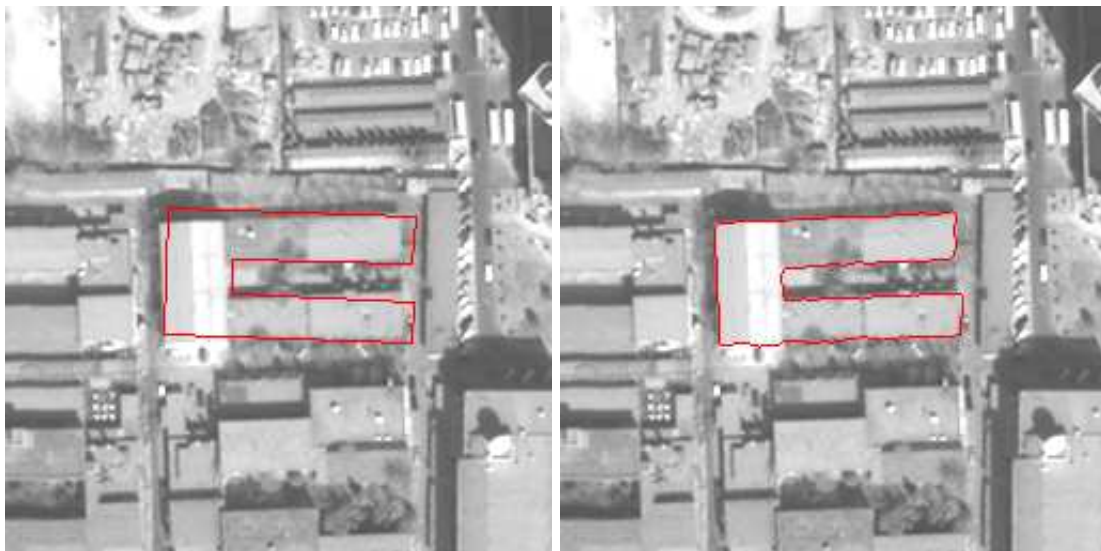
Nous proposons de mener des tests de non-changement avec une carte parfaite datant de 1996. Dans cette carte, les objets cartographiques ont été corrigés afin d'être similaires avec leur représentation dans l'image. Une probabilité de non-changement élevée est alors garantie si le recalage carte-image est effectif. Les figures 7.13 à 7.16 illustrent une réduction efficace des variabilités carte-image qui permet d'accroître la cohérence entre ces deux représentations. Le score final de Hough est ainsi largement accru par rapport à celui de l'état initial (table 2.2). Il en résulte une détection correcte des non-changements du bâti. En figure 7.13, une rotation de 0.1 rad est appliquée à l'objet cartographique, simulant une erreur de superposition carte-image. Le recalage par contours actifs est efficace et permet d'accroître le score du vote de Hough final ($s_{Hough} = 0.76$). La probabilité $p_{NC} = 0.63$ confirme un non-changement avec néanmoins une valeur moyennement élevée. Cela est dû à une variation géométrique $E_{geom} = 0.1$ non négligeable qui amoindrit la valeur de la probabilité. L'expérience de la figure 7.14 illustre le même phénomène : un score final de Hough relativement élevé ($s_{Hough} = 0.76$) avec une faible translation estimée ($\mu_{Hough} = (1, 0)$) confirment un non-changement. Néanmoins, la variation géométrique revoit à la baisse la valeur de la probabilité de non-changement : $p_{NC} = 0.63$. Avec l'expérience de la figure 7.15, la présomption de non-changement est plus forte ($p_{NC} = 0.76$) en raison d'un score de Hough plus élevé et d'une faible variation géométrique. Enfin, le résultat de la figure 7.16 illustre la nécessité de ne pas pénaliser les translations colinéaires avec la direction $\mathbf{V}$ de distorsion de perspective. La figure 7.16.a montre un polygone cartographique proche de l'empreinte au sol d'un bâtiment moyennement élevé de l'image. La distorsion est suffisamment faible pour que le contour actif puisse segmenter le toit. Le vecteur de translation entre l'état initial et le recalage final est quasiment colinéaire à $\mathbf{V}$. Ainsi, la variation géométrique n'est pas pénalisante ($E_{geom} = 0.01$). Le score de Hough est élevé ($s_{Hough} = 0.84$) avec une faible translation $\mu_{Hough} = (-1, 0)$, ce qui explique la valeur élevée de la probabilité de non-changement calculée $p_{NC} = 0.83$ (qui est aussi due au caractère parfait de la carte).



(a) Superposition carte-image sans recalage

(b) Après recalage

FIG. 7.13 – Détection de non-changement avec un objet cartographique parfait transformé par similitude de paramètres $\xi_0 = (s_0 = 1, \theta_0 = 0.1, \mu_{x,0} = 1, \mu_{y,0} = 1)$. Le non-changement est confirmé avec $s_{Hough} = 0.76$, $\mu_{Hough} = (1, 0)$, $E_{geom} = 0.1$ et $p_{NC} = 0.63$.



(a) Superposition carte-image sans recalage

(b) Après recalage

FIG. 7.14 – Détection de non-changement avec un objet cartographique parfait transformé par similitude de paramètres $\xi_0 = (s_0 = 1.05, \theta_0 = 0.075, \mu_{x,0} = 4, \mu_{y,0} = -1)$. Le non-changement est confirmé avec $s_{Hough} = 0.72$, $\mu_{Hough} = (1, 0)$, $E_{geom} = 0.08$ et $p_{NC} = 0.63$.



(a) Superposition carte-image sans recalage

(b) Après recalage

FIG. 7.15 – Détection de non-changement avec un objet cartographique parfait transformé par similitude de paramètres $\xi_0 = (s_0 = 1.0, \theta_0 = -0.02, \mu_{x,0} = -2, \mu_{y,0} = 1)$. Le non-changement est confirmé avec $s_{Hough} = 0.81$, $\mu_{Hough} = (0, 0)$, $E_{geom} = 0.04$ et $p_{NC} = 0.76$.



(a) Superposition carte-image sans recalage

(b) Après recalage

FIG. 7.16 – Détection de non-changement avec un objet cartographique parfait transformé par similitude de paramètres $\xi_0 = (s_0 = 1.0, \theta_0 = 0.0, \mu_{x,0} = 5, \mu_{y,0} = 5)$. Le non-changement est confirmé avec $s_{Hough} = 0.84$, $\mu_{Hough} = (-1, 0)$, $E_{geom} = 0.01$ et $p_{NC} = 0.83$.

Carte réelle

Nous traitons dorénavant des objets cartographiques issus d'une carte réelle de 1996, sans correction manuelle des erreurs potentielles. Les erreurs par rapport à l'image sont soit modélisables par une transformation globale (au sens de l'objet cartographique), soit locales. Dans le cas d'erreurs locales de forme des objets cartographiques, le recalage par contours actifs est alors un compromis entre l'information *a priori* erronée de la carte et la réalité de l'image. Les figures 7.17 et 7.18 montrent des résultats corrects de détection de non-changement dans le cas des erreurs de mise en correspondance modélisables par des transformations globales. Le recalage fin par contours actifs permet alors une superposition quasiment parfaite entre l'objet cartographique et le bâtiment de l'image. Cela permet d'accroître la cohérence entre les deux représentations et par conséquent le score de Hough avant et après le recalage. La probabilité p_{NC} est moins élevée dans le cas de la figure 7.17 en raison d'une variation géométrique plus importante ($p_{NC} = 0.68$ en figure 7.17 et $p_{NC} = 0.83$ en figure 7.18). Les figures 7.19 à 7.21 illustrent des résultats d'expériences menées avec des objets cartographiques contenant des erreurs de forme locales et de faible amplitude. Nous constatons que pour chacune de ces expériences, le résultat de recalage fin permet une meilleure correspondance, bien que les erreurs locales persistent et ne peuvent être amoindries. La présence de ces erreurs diminue le score du vote de Hough. En effet, certaines parties de l'objet cartographique affiné ne coïncident pas avec les gradients de l'image et ne peuvent ainsi recevoir la contribution des segments pour la translation μ_{Hough} recherchée. Dans le cas présent, les erreurs locales sont minoritaires et de faible amplitude. Elles ne détériorent pas le vote de Hough au point de rendre la décision de changement indéterminée ou incorrecte ($p_{NC} = 0.66$ en figure 7.19, $p_{NC} = 0.71$ en figure 7.20 et 7.21). Ces exemples illustrent la robustesse de l'approche proposée vis-à-vis des erreurs locales de forme mineures présentes dans la carte. Cette robustesse se manifeste à la fois dans le processus de recalage ainsi que durant le vote de Hough.

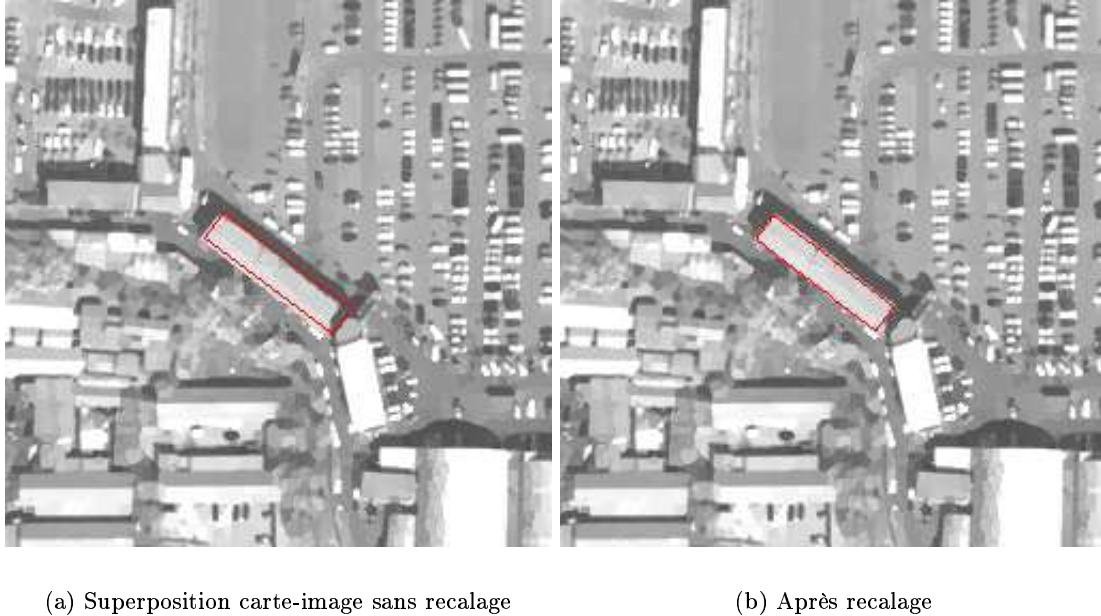
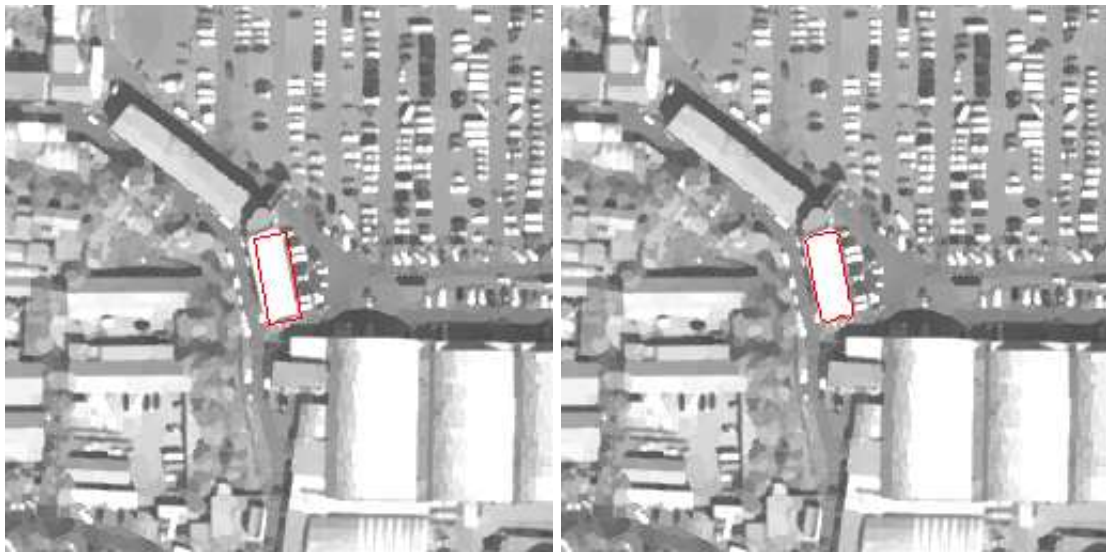


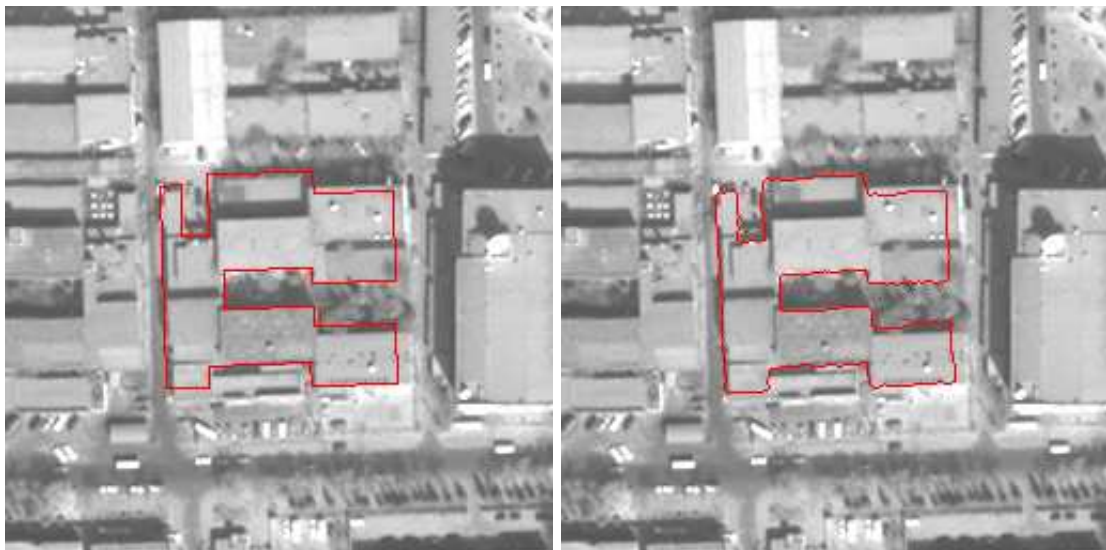
FIG. 7.17 – Détection de non-changement avec une erreur de mise en correspondance modélisable par une transformation globale. Le non-changement est confirmé avec $s_{\text{Hough}} = 0.86$, $\mu_{\text{Hough}} = (0, -1)$, $E_{\text{geom}} = 0.12$ et $p_{NC} = 0.68$.



(a) Superposition carte-image sans recalage

(b) Après recalage

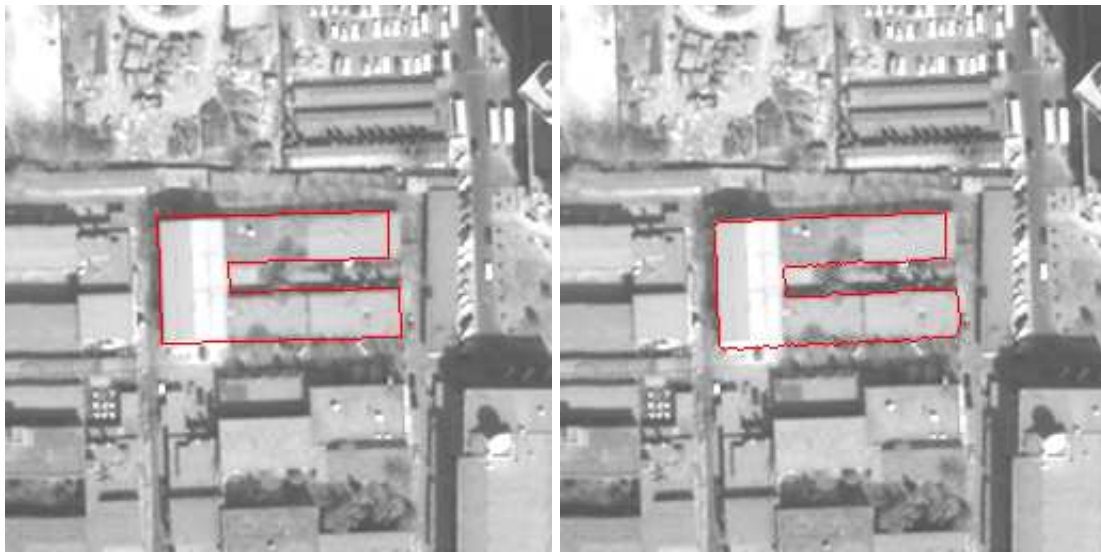
FIG. 7.18 – Détection de non-changement avec une erreur de mise en correspondance modélisable par une transformation globale. Le non-changement est confirmé avec $s_{Hough} = 0.88$, $\mu_{Hough} = (0, 0)$, $E_{geom} = 0.03$ et $p_{NC} = 0.83$.



(a) Superposition carte-image sans recalage

(b) Après recalage

FIG. 7.19 – Détection de non-changement avec des erreurs locales de forme. Le non-changement est confirmé avec $s_{Hough} = 0.62$, $\mu_{Hough} = (1, 0)$, $E_{geom} = 0.02$ et $p_{NC} = 0.66$.



(a) Superposition carte-image sans recalage

(b) Après recalage

FIG. 7.20 – Détection de non-changement avec des erreurs locales de forme. Le non-changement est confirmé avec $s_{Hough} = 0.73$, $\mu_{\mathbf{Hough}} = (1, 0)$, $E_{geom} = 0.03$ et $p_{NC} = 0.71$.



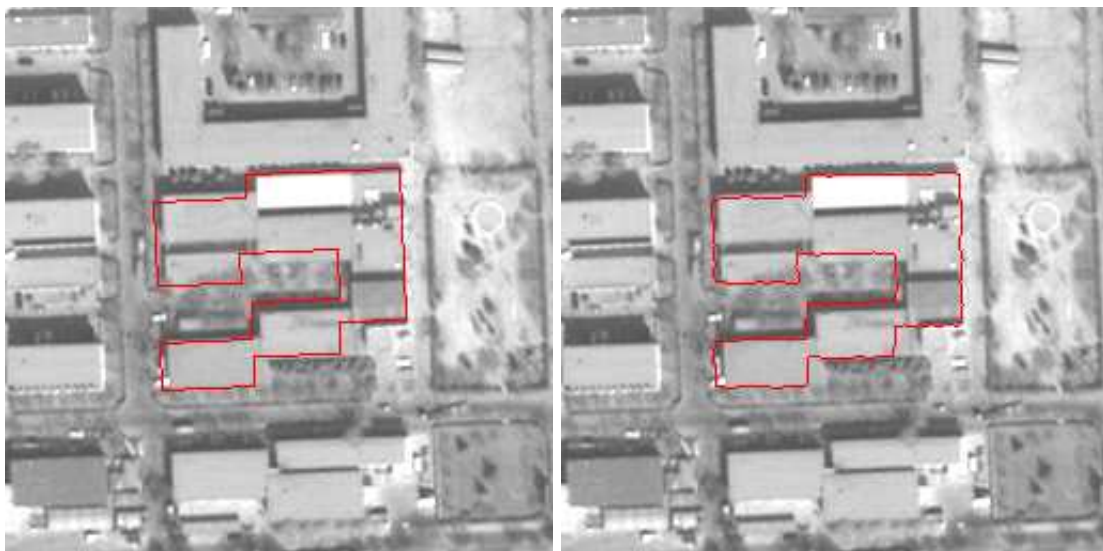
(a) Superposition carte-image sans recalage

(b) Après recalage

FIG. 7.21 – Détection de non-changement avec des erreurs locales de forme. Le non-changement est confirmé avec $s_{Hough} = 0.75$, $\mu_{\mathbf{Hough}} = (1, 0)$, $E_{geom} = 0.05$ et $p_{NC} = 0.71$.

7.5.2.3 Ambiguïté et fausse alarme avec une carte entachée d'erreurs locales

Lorsque les erreurs locales de forme de la carte sont plus sévères, la prise de décision finale devient soit ambiguë (cf. figure 7.22 avec une probabilité proche de 0.5, $p_{NC} = 0.55$), soit fausse (cf. figure 7.23 où $p_{NC} = 0.39$ alors qu'aucun changement n'est survenu). Dans le cas où des erreurs de forme plus significatives existent dans la carte, les incohérences carte-image subsistent après le recalage par contours actifs et sont trop nombreuses pour permettre un score de Hough important. Les scores de Hough des expériences des figures 7.22 et 7.23 sont équivalents ($s_{Hough} = 0.45$ et $s_{Hough} = 0.47$) et moins bons que ceux obtenus avec des erreurs de forme mineures (cf. figures 7.19-7.21). Alors que dans l'expérience de la figure 7.22, la décision finale est indéterminée, dans le second cas, une détection erronée de changement est produite à cause d'une variation géométrique plus importante ($E_{geom} = 0.19$). Ces deux cas illustrent les limites du modèle proposé et son incapacité à surmonter des erreurs de forme locales trop importantes. Ceci est inhérent au processus de recalage fin qui ne permet pas de variations locales significatives du contour actif par rapport à la forme *a priori* dérivée de la carte. Les travaux exposés au chapitre 6 sont une tentative de palier ce problème et permettraient ainsi d'obtenir un score de Hough élevé en dépit des erreurs locales, pour le peu qu'elles soient modélisables par des variations parallèles vis-à-vis de la forme de référence.



(a) Superposition carte-image sans recalage

(b) Après recalage

FIG. 7.22 – Détection ambiguë (cas indéterminé). $s_{Hough} = 0.45$, $\mu_{Hough} = (0, 0)$, $E_{geom} = 0.02$ et $p_{NC} = 0.55$.



(a) Superposition carte-image sans recalage

(b) Après recalage

FIG. 7.23 – Détection de changement erronée due aux erreurs locales de forme de la carte. $s_{Hough} = 0.47$, $\mu_{Hough} = (-1, -1)$, $E_{geom} = 0.19$ et $p_{NC} = 0.39$.

7.5.2.4 Synthèse des résultats

La table 7.1 résume les quantités calculées pour la détection de changement/non-changement des expériences présentées.

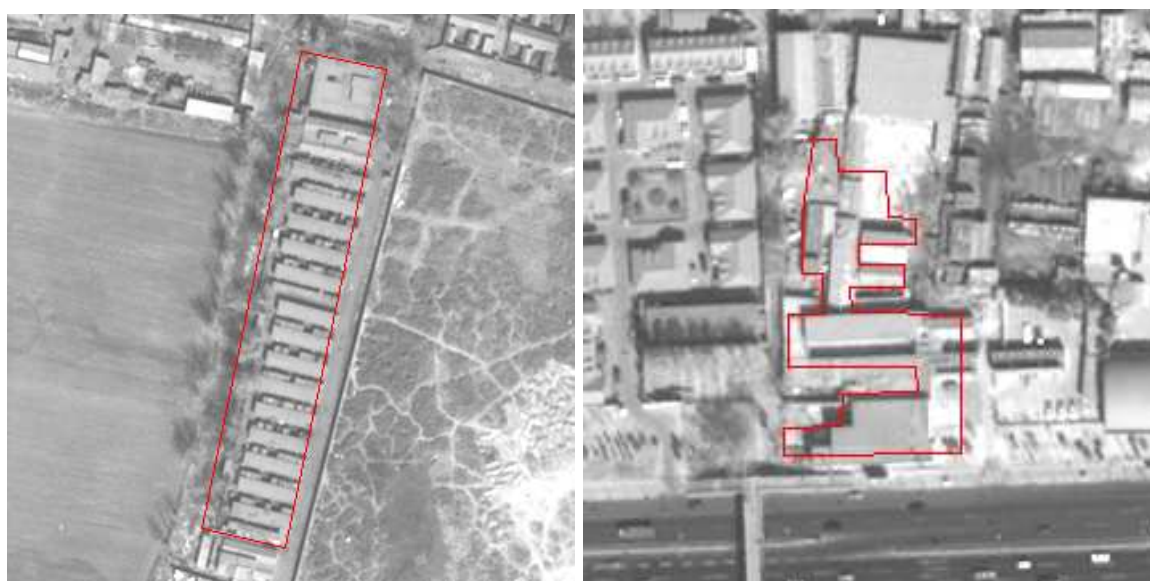
	<i>Expérience</i>	Avant recalage		Après recalage			
		s_{fusion}	s_{Hough}	s_{Hough}	μ_{Hough}	E_{geom}	p_{NC}
Changement	figures 7.10.a-b	1	/	/	/	/	/
	figure 7.11	0	0.22	0.91	(0, -1)	0.54	0.30
	figure 7.12	0	0.12	0.27	(-6, -26)	0.13	0.00
Non-changement	figure 7.13	0	0.34	0.76	(1, 0)	0.1	0.63
	figure 7.14	0	0.30	0.72	(1, 0)	0.08	0.63
	figure 7.15	0	0.54	0.81	(0, 0)	0.04	0.76
	figure 7.16	0	0.69	0.84	(-1, 0)	0.01	0.83
	figure 7.17	0	0.74	0.86	(0, -1)	0.12	0.68
	figure 7.18	0	0.78	0.88	(0, 0)	0.03	0.83
	figure 7.19	0	0.33	0.62	(1, 0)	0.02	0.66
	figure 7.20	0	0.35	0.73	(1, 0)	0.03	0.71
	figure 7.21	0	0.52	0.75	(1, 0)	0.05	0.71
Indéterminé	figure 7.22	0	0.35	0.45	(0, 0)	0.02	0.55
Fausse alarme	figure 7.23	0	0.41	0.47	(-1, -1)	0.19	0.39

TAB. 7.1 – Quantification des scores de changement/non-changement. p_{NC} est la probabilité de non-changement. s_{Hough} est le score de Hough normalisé.

Nous constatons que l'atténuation des variabilités exogènes grâce au recalage par contour actifs est efficace. En effet le score s_{Hough} est toujours amélioré après le recalage. Dans le cas de non-changement, l'amélioration varie entre 10% et 40%.

7.5.3 Analyse statistique

Dans cette section nous analysons le succès de la méthode proposée sur un plus grand nombre de cas de changements/non-changements à partir des données cartographiques disponibles. Une première série de tests est effectuée avec une carte de bonne qualité. Ce que nous entendons par “bonne qualité” concerne un faible taux d’erreurs locales de forme des objets cartographiques, ainsi que peu d’effets de généralisation et simplification. Une seconde série de tests est réalisée avec une carte de qualité moyenne. Dans cette carte, des erreurs locales de forme sévères existent. Elles sont soit dues à une mauvaise délimitation des bâtiments vus dans les images de télédétection ayant servies à la réalisation de la carte, soit à des effets de simplification/généralisation significatifs (cf. figure 7.24).



(a) Effet de généralisation

(b) Généralisation, erreurs de délimitation et présence de changements partiels

FIG. 7.24 – Illustration des erreurs de délimitation et de l’effet de généralisation dans la carte de 1996. L’objet cartographique est superposé sur l’image satellitaire de 2002.

Pour le recalage par contours actifs, nous avons choisi le modèle d’attache aux données basé sur l’information de frontière (GVF, cf. section 4.3.2.2). C’est en effet le modèle qui est capable de traiter le plus grand nombre de cas en raison de son insensibilité à l’hétérogénéité des toits et à l’effet de généralisation de la carte. Nous utilisons la contrainte de forme spatio-temporelle avec $\lambda_{min} = 1$, $\lambda_{max} = 3$ et $d_0 = 2$ (cf. section 5.4). La contribution du MNS n’est pas utilisée dans ces expériences.

7.5.3.1 Résultats avec une carte de bonne qualité

La carte utilisée a été réalisée par le Beijing Institute of Surveying and Mapping (BISM) grâce à l’image satellitaire Quickbird de 2002 utilisée dans les expériences. Ceci a pour effet d’augmenter la cohérence carte-image, bien que des effets de simplification, généralisation et d’erreur soient présents. D’autre part, aucun cas de changement ne doit être détecté avec une telle carte. Nous menons une première série de tests avec 40 bâtiments visant à quantifier les performances de la méthode vis-à-vis des non-changements. Dans un second temps, nous translatons la carte de 30

mètres dans la direction Est. La carte et l'image ne sont alors plus en correspondance, ce qui nous permet de détecter uniquement des changements (sur 37 cas). Les résultats sont ensuite rassemblés et classés selon la nomenclature suivante :

1. Détection de non-changement correcte (*true positive*)
2. Détection de changement correcte (*true negative*)
3. Détection de changement erronée (la méthode considère le bâtiment changé alors qu'il est inchangé dans la réalité)
4. Détection de non-changement erronée (fausse alarme : la méthode considère le bâtiment inchangé alors qu'il a disparu dans l'image)
5. Cas indéterminé

Les résultats menés avec la carte de 2002 sont présentés dans le tableau 7.2. Nous constatons que le taux d'indétermination relatif à la méthode est raisonnablement élevé. Ainsi, 87% des cas se sont vus assigner une décision de changement/non-changement. Parmi la totalité des cas considérés, 5.2% ont été interprétés à tort comme un changement. Ces cas de détections de non-changement manquées sont peu nombreux et s'expliquent par la présence d'erreurs locales de forme dans la carte. Comme nous l'avons montré en section 7.5.2.3, la présence d'erreurs sévères affaiblit le score du vote de Hough au point de rendre la décision indéterminée ou erronée. Ces erreurs sont responsables des détections manquées et des cas indéterminés. Les reste des cas représente 81.8% de détections correctes de changements et non-changements.

Détections de non-changement correctes :	Détections de non-changement erronées :
42.8%	0%
Détections de non-changement manquées :	Détections de changement correctes :
5.2%	39%
Cas indéterminés : 13%	

TAB. 7.2 – Quantification des résultats de la détection de changement avec la carte de 2002. Les résultats sont exprimés en pour-cent sur un total de 77 cas.

7.5.3.2 Résultats avec une carte de qualité moyenne

La carte de qualité moyenne que nous considérons date de 1996. L'écart temporel entre la carte et l'image satellitaires (2002) étant assez important, nous n'avons pas recours à la simulation de changements comme pour les tests précédents. Les résultats menés sur 105 cas sont présentés dans le tableau 7.3. Par rapport aux résultats réalisés avec la carte de 2002, nous constatons une légère hausse des cas indéterminés (14.3%). En revanche, la plus forte progression est enregistrée par le taux de détections de non-changement manquées qui atteint 17%. Ce taux élevé est cohérent avec les conditions expérimentales : la carte utilisée comporte des erreurs plus sévères que celle de 2002. Les cas difficiles de non-changement n'ont pas été détectés et au lieu de basculer dans la catégorie des cas indéterminés, ils ont été à tort interprétés comme des changements. Outre les erreurs de la carte, les cas de cette seconde série de tests étaient plus délicats à analyser : l'ombre portée de bâtiments élevés masquait complètement le bâtiment d'intérêt, le rendant ainsi non détectable ; la canopée recouvrait totalement certains bâtiments de petite taille. Comme pour le cas de la carte de 2002, les détections de non-changement erronées restent négligeables. Elles surviennent lorsque des segments de l'image sont cohérents avec la géométrie du bâtiment de la carte qui n'est plus présent

dans l'image. Ces cas surviennent mais sont très rares. Parmi tous les cas considérés, le taux de détections de changement/non-changement correctes s'élève à 66.7%, ce qui revient à un taux de 77.8% parmi les cas où il y a eu prise de décision.

Détections de non-changement correctes : 50.5%	Détections de non-changement erronées : 2%
Détections de non-changement manquées : 17%	Détections de changement correctes : 16.2%
Cas indéterminés : 14.3%	

TAB. 7.3 – Quantification des résultats de la détection de changement avec la carte de 1996. Les résultats sont exprimés en pour-cent sur un total de 105 cas.

7.6 Conclusion

Nous avons présenté une méthodologie de détection de changement entre une carte et des données de télédétection. Cette approche se décompose en deux temps et tire parti à la fois de la fusion de données et de la réduction des variabilités exogènes carte-image par la technique des contours actifs. La première phase consiste à calculer et fusionner des indices de changement respectivement basés sur la comparaison de la carte avec un modèle numérique de surface (MNS) et une image multispectrale. Les méthodes mises en œuvre sont simples, rapides mais permettent uniquement la détection d'un changement univoque du bâti urbain.

Lorsque cette phase n'est pas en mesure de décider de l'absence d'un bâtiment cartographique considéré, une seconde phase le compare à sa représentation dans une image satellitaire panchromatique. Les erreurs de superposition carte-image sont premièrement diminuées grâce à la mise en correspondance achevée par les contours actifs. Une comparaison subséquente entre l'image panchromatique et l'objet cartographique recalé est effectuée au moyen de deux mesures : une mesure par vote de Hough (qui quantifie la cohérence entre les segments du polygone et ceux extraits de l'image) et une mesure de déformation géométrique entre l'objet cartographique avant et après le recalage. Associant ces deux mesures à une énergie de Gibbs, une densité de probabilité de non-changement est formulée et évaluée en chaque bâtiment de la carte.

Des résultats expérimentaux ont illustré l'efficacité de la méthode sur des cas où les erreurs locales de forme des objets cartographiques sont faibles. Une analyse statistique a montré que dans ces circonstances, 82% des cas de changement/non-changement étaient correctement traités. Nous avons aussi pu constater que le recalage fin par contours actifs pouvait augmenter significativement la cohérence carte-image afin d'améliorer la confiance accordée à la décision de non-changement. Les résultats exposés ont par ailleurs mis en lumière les limitations de l'approche proposée : la présence d'erreurs sévères dans la carte met en échec le recalage fin par contours actifs et ne permet pas un vote de Hough significatif. Nous avons montré par une analyse statistique que dans ce cas, le taux de détections manquées de non-changement augmente sensiblement (+12%).

Chapitre 8

Conclusion

8.1 Synthèse

L'objectif de ce travail de thèse était le développement d'une méthode d'analyse de changements entre une carte numérique du bâti urbain et des données multi-sources de télédétection de haute résolution. Nous nous sommes en particulier attachés à résoudre le problème des variabilités exogènes entre les objets cartographiques et leurs représentations dans l'image. Ces variabilités sont dues soit à une imprécision de la superposition carte-image, soit à des erreurs de saisie des objets cartographiques, soit aux effets de simplification et de généralisation de la carte. Ces différences exogènes sont indépendantes d'un changement effectif et se doivent d'être atténuées pour garantir une détection de changement fiable. La correction de ces incohérences carte-image est rendue difficile en raison de la multitude d'artefacts présents dans les images et qui est propre à notre application focalisée sur les milieux urbains.

L'insertion de contrainte de forme dérivée de la carte dans les modèles de contours actifs a montré son efficacité à surmonter certains de ces artefacts. Cette connaissance *a priori* extraite de la carte est par ailleurs spécifique à la scène étudiée et permet de s'affranchir de l'endroit géographique analysé. Nous avons expérimentalement montré qu'en présence de contrainte de forme, les contours actifs étaient sensibles à une initialisation éloignée de l'objet cible. Ceci nous a mené à proposer deux solutions visant à conduire plus sûrement le contour actif vers la solution et à éviter le problème des minima locaux. La première consiste à fusionner l'information d'altitude d'un modèle numérique de surface (MNS) avec celle d'une image satellitaire panchromatique. La complémentarité de ces deux représentations a montré son efficacité. La seconde solution consiste à relaxer spatialement et dynamiquement la contrainte de forme durant le processus de convergence du contour actif. Une solution approximative et non régulière est d'abord atteinte alors que la contrainte est spatialement relaxée. Le rétablissement ultérieur et spatialement uniforme de la contrainte permet de régulariser le contour, évitant ainsi les artefacts urbains. Nous avons également proposé une nouvelle énergie de contrainte de forme, linéaire et quadratique, autorisant une classe de déformations du contour actif par rapport à la forme *a priori*. Ce modèle permet au contour des variations parallèles par rapport aux segments de la forme polygonale de l'objet cartographique et vise à corriger les erreurs locales de forme de ce dernier. Alors que des résultats encourageants sont obtenus avec des images de synthèse, cette dernière technique n'est pas encore mature et demande une compréhension plus approfondie des modèles et de leurs implantations afin d'être efficace sur des images réelles.

Enfin, nous avons proposé une méthode d'analyse de changement entre la carte et différentes données de télédétection. Des indices de changement simplement calculés à partir du MNS et d'une image satellitaire multispectrale sont fusionnés pour effectuer la détection des changements uni-

voques du bâti urbain. Dans la majorité des cas où un changement n'a pu être identifié de façon évidente, une seconde phase plus complexe est mise en œuvre afin de caractériser le degré de ressemblance entre l'objet cartographique et sa représentation dans l'image satellitaire panchromatique. Ce traitement tire parti du recalage fin par contours actifs et permet la formulation d'une probabilité de non-changement pour chaque bâtiment de la carte. Cette probabilité est dérivée d'une énergie de Gibbs. Celle-ci est calculée à partir de la déformation géométrique enregistrée lors du recalage par contours actifs et d'un vote de Hough destiné à mesurer la ressemblance des segments de l'image et de l'objet cartographique affiné. Une analyse quantitative a montré que le recalage fin améliore la robustesse de la décision de non-changement. Des tests menés sur un plus grand nombre de bâtiments ont statistiquement démontré les performances satisfaisantes de la méthode lorsque la carte contient peu d'erreurs locales de forme.

De ce travail de thèse, nous pouvons retenir les points suivants :

- La variation spatio-temporelle du poids de contrainte de forme a montré son efficacité à s'affranchir des minima locaux dus soit à une contrainte de forme uniforme et élevée, soit aux altérations de l'image. Elle permet de déplacer le choix critique d'une valeur constante du poids de contrainte vers celui d'une famille de fonctions paramétrées. Le résultat de la segmentation est alors moins sensible vis-à-vis de la valeur de ces paramètres que celle d'un poids constant. Cependant, notre formulation de la contrainte spatio-temporelle est issue d'observations empiriques et n'a pas de fondement théorique. Néanmoins, elle peut montrer la voie d'une formulation plus rigoureuse menant à la détermination de la fonction spatio-temporelle selon les informations dérivées de l'image et de la forme *a priori* et non par la définition de profils *ad hoc* dont le réglage des paramètres est une limitation.
- La contribution la plus innovante de ce travail de thèse est certainement celle concernant la formulation de nouvelles énergies visant à réduire les erreurs locales de forme de la carte. Néanmoins, de nombreux aspects relatifs à la stabilité numérique et l'influence de l'énergie quadratique restent mal compris. Les résultats de cette approche, non mature à ce jour, sont cependant encourageants.
- L'algorithme d'optimisation du simplex a montré son utilité dans deux situations : le calibrage de caméra pour la génération du MNS et l'estimation des paramètres de la similitude rendant la contrainte de forme des contours actifs invariante par rotation, translation et facteur d'échelle. Dans les deux cas, cette méthode s'est révélée efficace et robuste. Le simplex permet par ailleurs d'éviter la linéarisation (calibrage) ou la dérivation (descente de gradient) de la fonction de coût à minimiser. Ainsi, son utilisation est transposable à tout type de fonctionnelle, supposé que le nombre de paramètres à estimer n'est pas trop grand. Enfin, nous avons démontré que le simplex, contrairement à l'intuition, convergeait plus rapidement que la descente de gradient. Ce fait s'explique par l'estimation simultanée des paramètres et grâce à la réduction de la complexité calculatoire par la technique de la bande étroite.
- L'usage que nous avons fait des contours actifs est plutôt paradoxal : alors que ces outils sont connus pour leur aptitude à se déformer, nous leur avons imposé une contrainte de forme inhibant cette propriété de flexibilité. Ainsi, la mise en correspondance par contours actifs s'apparenterait à un recalage rigide qui aurait pu être achevé par exemple, par la technique de vote de Hough présentée. Cependant, en dépit d'une contrainte de forme suffisamment forte pour palier les artefacts urbains, les contours actifs gardent une flexibilité très locale qui les rendent *in fine* plus performants que les techniques de recalage rigide.

- Les indices de changement proposés dans cette thèse, issus d’approches ascendantes, sont relativement simples à calculer. Le score de Hough utilisé pour mesurer la similarité entre l’objet cartographique et l’image est robuste et complémentaire de l’approche descendante des contours actifs.
- Enfin, nous avons pu vérifier que les contours actifs étaient des outils flexibles et puissants de segmentation, capables à la fois d’intégrer des contraintes géométriques génériques et spécifiques et de mener à bien la fusion de données.

8.2 Discussion des limitations

8.2.1 Limites du modèle de contrainte de forme

La limitation la plus contraignante de l’approche proposée est certainement l’incapacité des contours actifs à s’affranchir de la forme *a priori* dérivée de la carte afin de corriger les erreurs locales de forme qu’elle comprend. Les modèles de contrainte de forme utilisés sont invariants par similitude plane directe et permettent ainsi de corriger des erreurs globales de rotation, translation et d’échelle modulo des variations locales de très faible amplitude dépendant du poids alloué à la contrainte par rapport à l’information de l’image. Nous avons montré dans cette étude que l’invariance par similitude n’était pas à même de résoudre le problème des erreurs locales, ce qui nous a conduit à proposer une façon plus flexible d’imposer la contrainte de forme avec la possibilité de déformer localement et parallèlement le contour par rapport à la forme *a priori*. Cette flexibilité est néanmoins restreinte à une *classe* de déformations : toute tentative d’incorporation de plus de souplesse dans le modèle se doit d’être contrainte et doit obéir à certaines règles à défaut de quoi le contour actif sera sensible aux artefacts de l’image.

Notre choix de conférer au contour actif une liberté restreinte et contrôlée par le parallélisme reflète notre connaissance *a priori* des erreurs locales de forme les plus souvent rencontrées dans la carte. Cependant, cette connaissance est clairement insuffisante pour modéliser la totalité des variabilités exogènes carte-image. En effet, celles-ci sont dues soit aux imprécisions de tracé, soit aux effets de simplification et de généralisation de la carte. Dans la généralité des cas, ces variabilités obéissent plus à la subjectivité du photo-interpréteur qu’à des règles bien définies, elles semblent ainsi impossibles à modéliser avec des méthodes non heuristiques. Par conséquent, les modèles de contours actifs proposés dans cette thèse ne seront jamais en mesure d’éradiquer totalement les erreurs locales des objets cartographiques.

8.2.2 Limites de l’analyse de changements

Les erreurs locales de forme de la carte, lorsqu’elles sont de faible amplitude, ont peu d’impact sur la méthode d’analyse de changements proposée. Nous avons néanmoins montré quantitativement que la présence d’erreurs sévères accroît sensiblement le taux de détections de changement erronées. Ceci est la conséquence directe de l’inaptitude des contours actifs à s’affranchir des erreurs locales de forte amplitude. Ainsi, dans le cas d’erreurs cartographiques notables, la comparaison entre l’objet cartographique “affiné” et les segments de l’image par vote de Hough produit un score médiocre de similarité, sans qu’il nous soit possible d’affirmer si ce manque de ressemblance est dû à un changement ou à des erreurs de la carte.

8.2.3 Limites des contours actifs

L'utilisation des contours actifs contraints nécessite le réglage de poids déterminant l'influence de l'attache aux données et de la contrainte de forme. Ce réglage est empirique ce qui représente une limite au caractère totalement automatique de la méthode. Nous avons néanmoins remarqué qu'un seul réglage permettait d'obtenir des résultats satisfaisants quel que soit l'endroit de l'image analysé.

Le choix du modèle d'attache aux données paraît en revanche plus délicat. Nous avons illustré dans cette étude que les modèles basés sur l'information de région étaient plus performants mais uniquement dédiés aux bâtiments avec un toit homogène. Les contours actifs basés sur l'information de frontière sont plus sensibles à l'initialisation mais sont applicables à tout type de bâtiment. Pouvoir choisir automatiquement entre l'un et l'autre des modèles serait un gain en robustesse et efficacité. Cette question reste néanmoins ouverte car déterminer à l'avance si un bâtiment est homogène ou non est une tâche difficile dans la mesure où le recalage initial carte-image est imprécis.

Les contours actifs sont connus pour leur sensibilité à l'initialisation et leurs temps de convergence longs. La première limitation nous a contraint à traiter uniquement les bâtiments peu à moyennement élevés de l'image. La seconde représente un obstacle au caractère opérationnel de l'approche proposée dans la thèse. Bien que les procédures aient été optimisées, le traitement d'un bâtiment varie de quelques secondes à quelques minutes ; les temps de calcul menés avec les modèles quadratiques sont prohibitifs.

8.3 Perspectives

Avant de clore ce document, nous souhaitons énumérer quelques voies possibles pour de futurs travaux :

- Nous venons de voir que la robustesse de l'analyse de changements carte-image dépend du succès des contours actifs à surmonter les erreurs locales de forme des données cartographiques. Nous avons par ailleurs constaté qu'une correction exhaustive des erreurs était illusoire. Cependant, une suppression totale des erreurs n'est pas indispensable pour accroître la robustesse de l'analyse de changements. En effet, la réduction de l'amplitude de l'erreur serait déjà suffisante pour réduire le taux de détections de changement erronées. Ainsi, le travail entrepris au sujet des énergies quadratiques mérite d'être poursuivi. Alternativement, il serait intéressant de voir dans quelle mesure une incorporation d'un modèle statistique de contrainte de forme (à partir d'un apprentissage des erreurs) pourrait améliorer la technique.
- Les critères énergétiques de non-changement formulés dans notre approche sont globaux à l'échelle de l'objet cartographique. En effet, le score de Hough est le résultat de la contribution totale des segments ayant participé au vote ; l'énergie de variation géométrique est un calcul intégral sur le domaine de l'image, de fait global. Il serait judicieux de mener une analyse plus locale du changement/non-changement afin de valider les parties de l'objet cartographique affiné qui sont cohérentes avec l'image. Les parties incohérentes qui sont soit dues à des erreurs de la carte, soit à des changements partiels effectifs seraient alors proposées à l'examen d'un photo-interpréteur.
- Notons que l'information d'ombre, caractéristique des bâtiments, n'a pas été utilisée pour la détection de changement. Celle-ci mérite d'être considérée afin d'accroître la robustesse de la décision finale de changement/non-changement.

- En vue d'améliorer la robustesse des contours actifs, il est concevable d'utiliser l'image satellitaire multispectrale préalablement à l'image panchromatique (de résolution quatre fois inférieure à celle de l'image multispectrale). Ceci représenterait un traitement multi-échelles susceptible d'améliorer les temps de calcul. L'information de couleur est directement utilisable dans le modèle Bayésien selon les travaux de M. Rousson [103]. Bien que la couleur paraît une caractéristique ne discriminant que moyennement le bâti du reste de la scène, cette piste mérite d'être explorée.
- Enfin, la technique du vote de Hough peut être utilisée pour placer l'objet cartographique sur le toit des bâtiments élevés (effet de perspective important) avant d'appliquer le recalage plus fin et plus précis des contours actifs. L'approche basée sur le vote de Hough est non locale, rapide et permettrait de généraliser la méthode proposée à tout type de bâtiment. Elle nécessite néanmoins d'être étendue à l'invariance par rotation et facteur d'échelle afin de tenir compte des erreurs de superposition initiale carte-image.

Nous espérons que ce travail de thèse aura pu apporter une modeste contribution à l'analyse (semi-) automatique de scènes urbaines pour assister et aider tout photo-interpréteur dans sa tâche de mise à jour de cartes. Au delà de l'application immédiate, nous espérons que ce travail profitera à la société et au développement durable des zones urbaines qui sont en passe de regrouper près de la moitié de la population mondiale.

Annexe A

Présentation et pré-traitements des données

Cette section propose de décrire les données cartographiques et de télédétection utilisées lors de ce travail de thèse. Ces données représentent toutes la ville de Pékin à différentes dates. Les images de télédétection sont optiques et de deux sortes. La première concerne des couples stéréoscopiques d'images aériennes à partir desquels nous générerons un modèle numérique de surface de la scène (MNS). L'autre type d'images de télédétection regarde l'imagerie satellitaire très haute résolution. Nous présenterons les caractéristiques des images réalisées par le satellite Quickbird qui ont la meilleure résolution à l'heure actuelle dans le domaine civil. Enfin, les données cartographiques utilisées sont disponibles sous forme numérique et sont issues d'un système d'information géographique (SIG). Nous nous attacherons à décrire uniquement celle du bâti.

A.1 Présentation des données

A.1.1 Données de télédétection

A.1.1.1 Images aériennes

Nous disposons de trois images de Pékin couvrant la zone du future village olympique de 2008 (figure A.1). Les images sont analogiques et représentent la scène à une échelle de 1 :10,000. Elles ont ensuite été scannées par les soins du BISM à un pas de 21 microns. Les images numériques résultantes ont ainsi une résolution de 21 cm par pixel. Les autres caractéristiques de ces images sont les suivantes :

- date d'acquisition : automne 1999
- acquises par un appareil analogique « RC 30 Leica Systems Camera »
- capteur optique à géométrie conique
- images multispectrales (3 bandes : RVB)
- dimensions d'une image : 11,000 par 11,000 pixels (soit une zone couverte de $2.3 \times 2.3 \text{ km}^2$)
- le système optique assurant l'acquisition de la photographie a une distorsion négligeable.
- pas de système de géo-référencement par GPS ou gyroscope pour la détermination des paramètres de prise de vue.



FIG. A.1 – Image aérienne représentée à 5% de sa taille originale. Le futur village olympique se situe au nord-est de l'image.

Suite à l'énoncé de ces caractéristiques, plusieurs remarques s'imposent. La première concerne la taille élevée des images. Dans l'optique d'utiliser celles-ci avec le logiciel de reconstruction 3D pour générer un MNS, il sera nécessaire de les sous-échantillonner au risque d'être confronté à des temps de calcul prohibitifs (un MNS généré par le logiciel résulte d'une corrélation de deux images stéréoscopiques). Le temps n'est pas le seul argument quant au choix du sous-échantillonnage. En effet, ce dernier permettra de réduire le bruit inhérent au scannage de la photographie ainsi que le niveau de détail de l'image qui est source d'erreurs lors de la génération d'un MNS lorsqu'il est trop élevé.

Le logiciel de reconstruction impose une autre contrainte quant aux images aériennes. Celles-ci doivent être mono-bandes. Nous avons choisi de convertir les images originales tri-bandes (RVB) par le calcul de la moyenne des trois bandes pour chaque pixel considéré. L'image mono-bande ré-

sultante se rapproche ainsi le plus du spectre panchromatique. Cette conversion est effectuée avant le sous-échantillonnage des images.

A.1.1.2 Images satellitaires

Nous possédons deux images Quickbird brutes (produit “basic bundle” de DigitalGlobe) acquises en 2002 et 2003 et dont les caractéristiques sont les suivantes :

- Date d’acquisition : mars 2002 / septembre 2003.
- Image corrigée radiométriquement uniquement.
- Image panchromatique à 0.6 m/pixel.
- Image multispectrale à quatre bandes (trois canaux rouge, vert, bleu traditionnels avec une quatrième bande en infra-rouge proche) avec 2.8 mètres de pas au sol.
- Latitude/Longitude de la zone rectangulaire couverte :
 - Coin nord-ouest : Latitude : 40.0667° (2002) / 40.0808° (2003); Longitude : 116.2608° (2002) / 116.259 (2003)
 - Coin sud-est : Latitude : 39.9475 (2002) / 39.919 (2003); Longitude : 116.4247° (2002) / 116.455 (2003)
- Chaque canal spectral est codé sur 16 bits.
- L’image est brute et ne contient aucune information de géocodage.
- Angle par rapport au nadir : 5° (2002) / 9° (2003).

A.1.2 Données cartographiques

Nous possédons deux jeux de données cartographiques réalisés en 1996, 2001 et 2002 par le BISM. Les données de 2001 et 2002 sont réalisées par photo-interprétation d’images Ikonos (1m/pixel) et Quickbird respectivement, celles de 1996 le sont par photo-interprétation de couples d’images aériennes. Ces données de type SIG ont les caractéristiques suivantes :

- Échelle 1 :10,000
- Une carte est partagée sous la forme de tuiles de 4 km (nord-sud) × 5 km (est-ouest)
- Organisées en couches : route, bâti, rivières, lacs, espaces verts (figure A.2).
- Données uniquement à *deux dimensions* et géocodées dans le système cartographique propre à Pékin avec une précision inférieure ou égale au mètre.
- Format vectoriel (au niveau du bâti) :
 - Chaque bâtiment est représenté par un polygone simple quelconque (pas d’intersection des côtés) résultant du détournage de l’objet par un opérateur. Il se peut qu’un ou plusieurs polygones soient inclus dans un autre polygone pour représenter des cours intérieures, ou des super-structures.
 - Les polygones ne s’intersectent pas de façon partielle.
 - Le détournage du bâtiment par l’opérateur varie en fonction de l’effet de perspective (façades visibles). Dans le cas où le bâtiment est peu élevé (distorsion due à la perspective négligeable pour l’échelle considérée), le contour du toit est saisi et inséré dans la carte. Lorsque l’effet de perspective est important (le toit et l’empreinte au sol du bâtiment ne coïncident pas), le toit est détourné puis translaté sur l’empreinte au sol.
- Les données SIG sont manipulables sous la forme de fichiers ASCII de type DXF (format créé par les développeurs d’AUTOCAD).
- Le fichier relatif à une tuile contient une liste de polygones. Chaque polygone contient une sous-liste de points désignant les sommets de ce dernier. Cette liste est cyclique (le dernier

point de la liste est identique au premier), chaque point est exprimé dans le système de coordonnées de Pékin.

- Simplification lors de la création des données :
 - **Généralisation.** Il est difficile de détourner chaque bâtiment lorsqu'ils sont agglomérés entre eux, nombreux et de petite taille. La solution retenue par le BISM consiste à détourner le bloc composé des bâtiments. Les blocs sont séparés par les routes principales. La figure 2.4 illustre cet effet de généralisation.
 - **Simplification individuelle.** Certains bâtiments ont une forme complexe qu'il est inutile de transcrire avec précision dans la carte pour l'échelle considérée. La forme est ainsi simplifiée de façon plus ou moins arbitraire par l'opérateur.

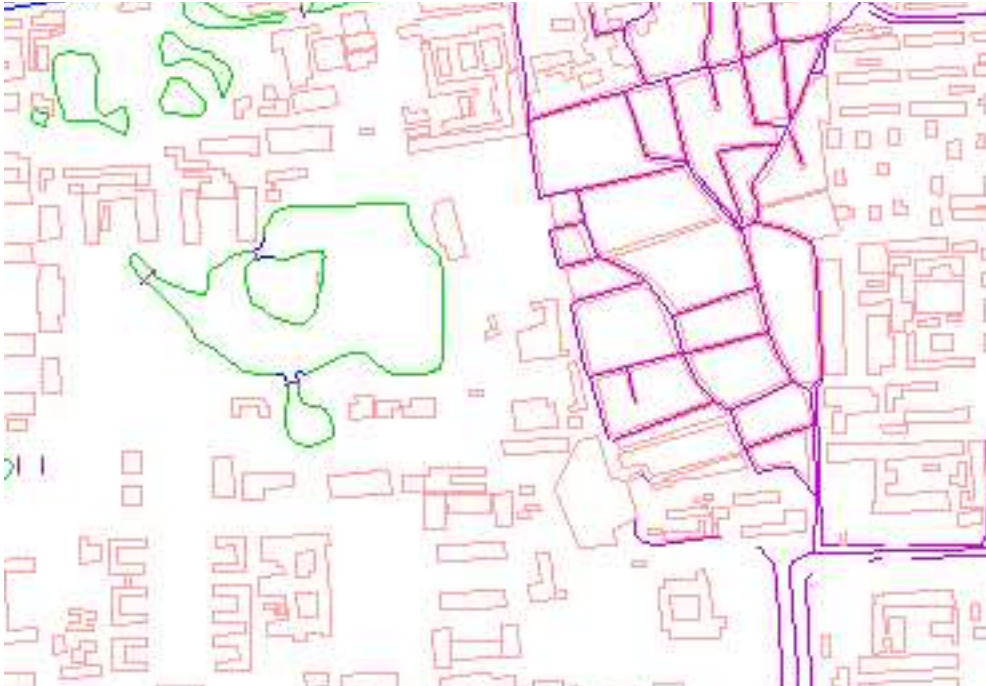


FIG. A.2 – Extrait des données SIG d'échelle 1 :10,000 de la ville de Pékin. Le bâti est en rose, le réseau routier en violet, les rivières et canaux en bleu, les lacs en vert.

A.2 Pré-traitements des données de télédétection

A.2.1 Sous-échantillonnage des images aériennes

Le sous-échantillonnage d'une image est une application de base du traitement d'images. De nombreuses méthodes permettent de le réaliser (plus proche voisin, bilinéaire, bicubique,...). Une opération de sous-échantillonnage peut être formalisée par un filtrage idéal passe-bas [14] :

$$\hat{I}(i, j) = \sum_{k, l \in \mathbb{Z}} \text{sinc}(\nu_e^c i - k) \text{sinc}(\nu_e^l j - l) I(k, l) \quad (\text{A.1})$$

où : ν_e^c et ν_e^l sont les fréquences d'échantillonnage de l'image I selon les colonnes et les lignes respectivement.

Un tel filtrage est réalisé dans le domaine spatial par une convolution avec un sinus cardinal sur une échelle infinie. Dans la pratique, il existe plusieurs méthodes pour rendre le fenêtrage spatial fini. La manière la plus brutale consiste à multiplier le signal par une fenêtre rectangle, engendrant ainsi des effets de Gibbs (cela étant dû à la pente infinie de cette fenêtre à ses bords). De multiples fenêtres aux dérivées finies ont ensuite été proposées (Blackman, Hanning, Hamming, Kaiser,...). [50] est l'article de référence concernant ces fenêtres d'apodisation.

Dans notre cas, une attention particulière sera portée à la préservation des contours de bâtiments puisque ce sont ces derniers que préserve la technique de corrélation développée dans le logiciel de reconstruction 3D [86, 87]. Dans la majorité des cas, un contour de bâtiment se traduit par une différence de luminance sur l'image. C'est une transition de type échelon ou rampe, caractérisée par de hautes fréquences dans le domaine spectral. Dans [14], Philippe Blanc a mené une étude comparative des filtres passe-bas associés à différentes techniques d'échantillonnage. Il en ressort que la technique de sous-échantillonnage dite d'interpolation de Shannon apodisée est celle permettant de mieux conserver les hautes fréquences d'une image. C'est cette méthode que nous avons par conséquent choisie pour le sous-dimensionnement des images aériennes de Pékin. L'interpolation par noyau de Shannon apodisée est une convolution de l'image par un sinus cardinal fenêtré par les fonctions d'apodisation discutées dans [50]. Nous avons choisi de sous-échantillonner les images par un facteur trois afin d'obtenir une résolution de 0,63 mètres par pixel ce qui est la meilleure résolution atteinte par les satellites d'observation dans le domaine civil actuellement (Quickbird 2). Ce facteur de sous-échantillonnage permet aussi de diminuer la complexité des images et par conséquent le taux d'erreurs lors de leur corrélation. Le bruit introduit par le scanner caractérisé par de hautes fréquences a pu être réduit par le filtrage passe-bas du sous-échantillonnage. La convolution de Shannon apodisée a été menée sur un domaine de 97×97 pixels avec une fenêtre de Blackman.

A.2.2 Rectification des images satellitaires

Les images satellitaires ont été ultérieurement rectifiées des variations altimétriques du terrain par un MNT. Ce traitement, effectué par le BISM, permet le géocodage des données dans le système cartographique de Pékin. La rectification est une phase critique car sa précision influencera la qualité de la carte mise à jour. En effet, nous considérons que les images satellitaires sont les données récentes de référence à partir desquelles nous mettrons à jour la carte. Cinquante points de contrôle ont été utilisés par le BISM pour évaluer la qualité de géocodage les images Quickbird. La précision globale atteinte est de 0,65 pixel, i.e. 0,4 m.

Annexe B

Algorithme du Simplex

La méthode du simplex consiste à trouver un jeu de paramètres optimaux minimisant une fonction de coût. C'est une technique d'ordre 0 ne nécessitant pas le calcul du gradient de la fonction de coût. Dans cette méthode due à Nelder et Mead [77], on recherche des valeurs de la fonction de coût sur un polyèdre à $N + 1$ sommets dans un espace de dimension N (un triangle dans le plan, un tétraèdre dans l'espace, ...). Le premier sommet est l'évaluation de la fonction de coût avec les paramètres initiaux, les autres sommets sont construits en faisant varier ces paramètres.

A chaque itération, on calcule un nouveau simplex qui permettra de se rapprocher efficacement du minimum. Pour ce faire, on commence par classer les sommets du simplex courant selon les valeurs de la fonction sur ces sommets. Appelons $\mathbf{m}$ le meilleur point, $\mathbf{p}$ le pire, $\mathbf{a}$ l'avant-dernier immédiatement meilleur que $\mathbf{p}$. Si l'on cherche à avancer, la direction générale doit correspondre à un éloignement de $\mathbf{p}$. On choisira comme direction celle de la droite $\{\mathbf{p}, \mathbf{c}\}$, $\mathbf{c}$ étant le centre de la face opposée à $\mathbf{p}$ (ses coordonnées sont les moyennes des coordonnées des autres points). Sur cette droite, on essaie le point $\mathbf{e}$ symétrique de $\mathbf{p}$ par rapport à $\mathbf{c}$:

$$\mathbf{e} = \mathbf{c} + (\mathbf{c} - \mathbf{p}) = 2\mathbf{c} - \mathbf{p}$$

Si $\mathbf{e}$ est meilleur que $\mathbf{m}$, on peut essayer d'aller encore plus loin dans cette direction :

$$\mathbf{e}' = \mathbf{c} + \gamma(\mathbf{c} - \mathbf{p}), \gamma > 1$$

Si $\mathbf{e}'$ est meilleur que $\mathbf{e}$, le nouveau simplex sera obtenu en remplaçant $\mathbf{p}$ par $\mathbf{e}'$ (*expansion* du simplex, figure B.2). Sinon, on remplacera $\mathbf{p}$ par $\mathbf{e}$ (*réflexion* du simplex cf. figure B.1)

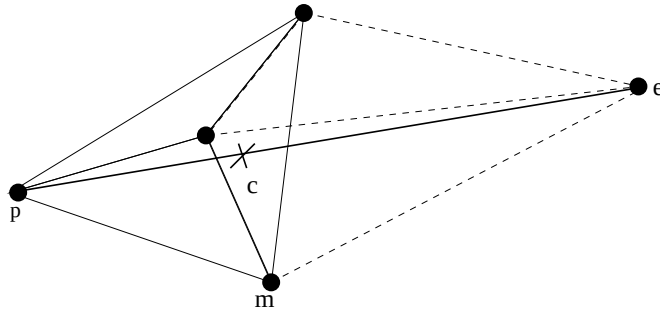


FIG. B.1 – simplex initial et réflexion

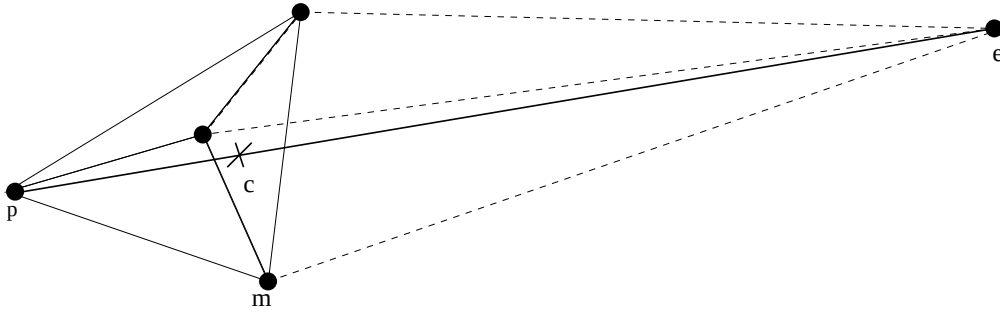


FIG. B.2 – Expansion

Si e n'est pas meilleur que m on essaie e'' et e''' situés sur la droite $\{c, p\}$ de part et d'autre de c selon :

$$e'' = c + \frac{1}{2}(c - p)$$

$$e''' = c - \frac{1}{2}(c - p)$$

Sinon, si le meilleur de ces points est meilleur que a , on va le choisir en remplacement de p (*contraction* du simplex, figure B.3)

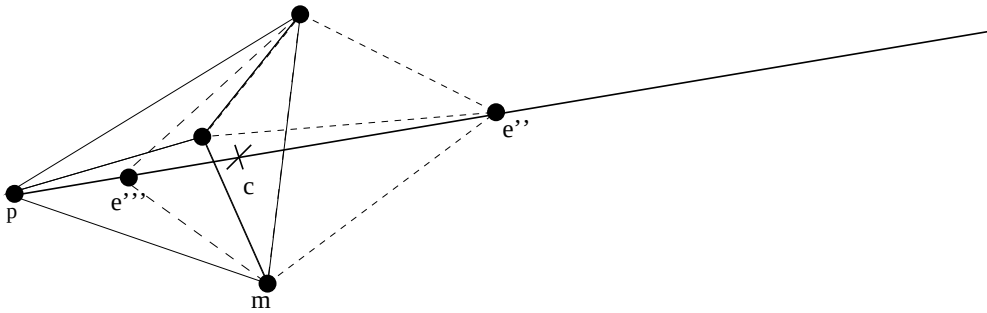


FIG. B.3 – Contractions possibles

Sinon, c'est que le point m est vraiment proche du minimum : on rétrécit alors le simplex en gardant m et en remplaçant tous les points s_i par les milieux des segments $[s_i, m]$.

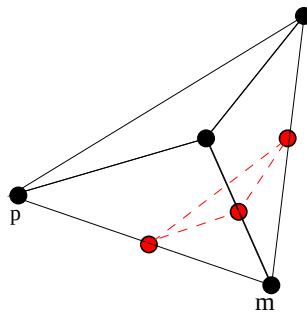


FIG. B.4 – Rétrécissement

La méthode est très simple et converge avec une grande robustesse. Les itérations s'arrêtent lorsque la dimension du simplex est suffisamment petite.

Annexe C

Modélisation de localisation directe et inverse

C.1 Modélisation inverse

Cette modélisation consiste à déterminer la position dans l'image d'un point initialement exprimé en coordonnées tridimensionnelles dans un référentiel $\mathfrak{R}$. Que $\mathfrak{R}$ soit un référentiel géocentré, ou associé à une projection cartographique, cela n'affecte par l'écriture des équations de la modélisation inverse. Nous supposons uniquement que $\mathfrak{R}$ soit un repère direct orthonormé.

1. Soient $(X_{sol}, Y_{sol}, Z_{sol})$ les coordonnées d'un point de la scène exprimées dans le référentiel $\mathfrak{R}$. Les coordonnées du vecteur de visée dans $\mathfrak{R}$ sont : $(X_{sol} - X_{cam}, Y_{sol} - Y_{cam}, Z_{sol} - Z_{cam})$. Les coordonnées indexées par “cam” sont celles de la position du centre optique de la caméra dans $\mathfrak{R}$.
2. Ce même vecteur, exprimé dans le repère caméra, vaut :

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = M \begin{pmatrix} X_{sol} - X_{cam} \\ Y_{sol} - Y_{cam} \\ Z_{sol} - Z_{cam} \end{pmatrix}$$

avec M la matrice de rotation faisant passer du repère $\mathfrak{R}$ au repère caméra. L'expression de M est donnée par :

$$M = \begin{pmatrix} \cos \alpha_z & -\sin \alpha_z & 0 \\ \sin \alpha_z & \cos \alpha_z & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \alpha_y & 0 & -\sin \alpha_y \\ 0 & 1 & 0 \\ \sin \alpha_y & 0 & \cos \alpha_y \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha_x & -\sin \alpha_x \\ 0 & \sin \alpha_x & \cos \alpha_x \end{pmatrix} \quad (C.1)$$

où $(\alpha_x, \alpha_y, \alpha_z)$ sont les angles de rotation du repère caméra par rapport aux axes $(0X)$, $(0Y)$ et $(0Z)$ du référentiel $\mathfrak{R}$.

3. Ces coordonnées dans le repère caméra, centrées sur le Point Principal de Symétrie (PPS), sans distorsion, du point correspondant sont :

$$\begin{pmatrix} -\frac{U}{W}f \\ -\frac{V}{W}f \end{pmatrix}$$

où f est la distance focale du système optique.

4. En négligeant la distorsion, et en recentrant dans le repère caméra (direct et orthonormé), ces

coordonnées deviennent :

$$\begin{cases} x = X_{PPS} - \frac{U}{W}f \\ y = Y_{PPS} - \frac{V}{W}f \end{cases}$$

Le passage des coordonnées exprimées dans le repère caméra vers les coordonnées pixels de l'image sollicitent les paramètres d'affinité :

$$\begin{cases} ligne = T_{lig} + a_{00}x + a_{01}y \\ colonne = T_{col} + a_{10}x + a_{11}y \end{cases}$$

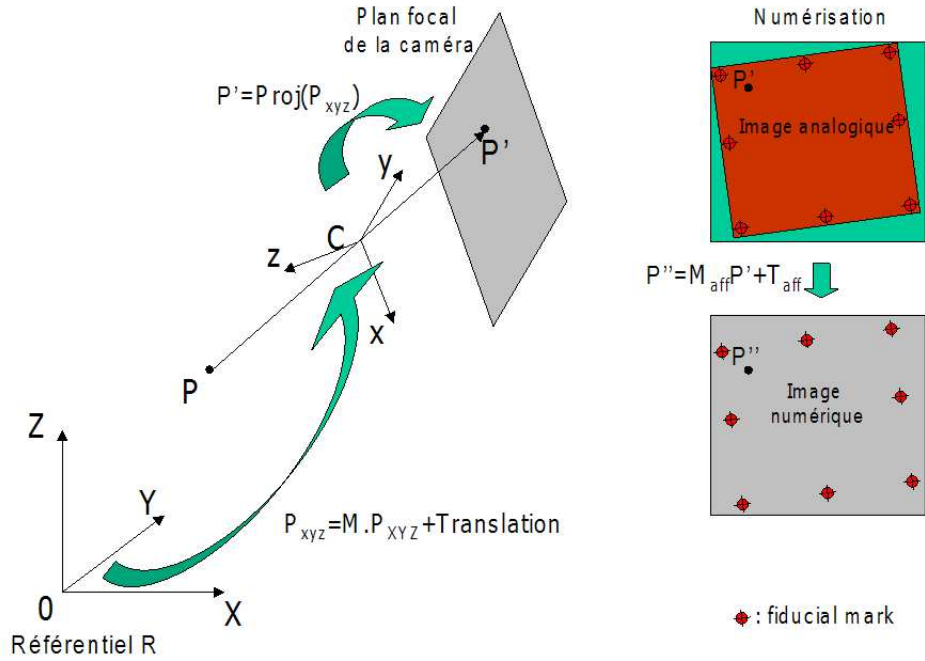


FIG. C.1 – Schéma illustrant la modélisation inverse. Les repères terrain et caméra sont directs.

C.2 Modélisation directe

La modélisation directe permet de calculer à partir d'un point de l'image sa position tridimensionnelle dans $\mathfrak{R}$ à une altitude h donnée au-dessus d'un ellipsoïde.

1. Soient (l, c) les coordonnées ligne/colonne du pixel à localiser. Les coordonnées (x, y) de ce même point dans le repère caméra sont données par :

$$\begin{cases} x = \frac{(l - T_{lig})a_{11} - (c - T_{col})a_{01}}{a_{00}a_{11} - a_{01}a_{10}} \\ y = \frac{-(l - T_{lig})a_{10} + (c - T_{col})a_{00}}{a_{00}a_{11} - a_{01}a_{10}} \end{cases}$$

où : (T_{lig}, T_{col}) sont les paramètres d'affinité en translation ; (a_{ij}) sont les paramètres d'affinité de rotation/facteur d'échelle.

2. Les coordonnées (x', y') dans ce même repère mais centrées sur le PPS sont :

$$\begin{cases} x' = x - X_{PPS} \\ y' = y - Y_{PPS} \end{cases}$$

3. En négligeant la distorsion, nous pouvons ensuite exprimer la direction de visée V dans le repère caméra :

$$\begin{pmatrix} \frac{x'}{f} \\ \frac{y'}{f} \\ -1 \end{pmatrix}$$

4. Cette même direction exprimée dans $\mathfrak{R}$ est :

$$V' = M^{-1}V$$

où M est la matrice de rotation de l'équation (C.1).

La détermination du point 3D de hauteur h dans $\mathfrak{R}$ se fait par intersection de la direction de visée avec l'ellipsoïde de révolution de rayons $(a + h, b + h)$ (a : grand rayon, b : petit rayon). L'équation de l'ellipsoïde de révolution modélisant la surface de la terre et exprimée dans $\mathfrak{R}$ revient à :

$$\frac{x^2}{(a + h)^2} + \frac{y^2}{(a + h)^2} + \frac{z^2}{(b + h)^2} = 1 \quad (\text{C.2})$$

Le point d'intersection I entre la direction de visée V' et l'ellipsoïde vérifie :

$$I = C + uV'$$

où : C est la position du centre optique de la caméra exprimée dans $\mathfrak{R}$; u est un scalaire. Trouver l'intersection I revient ainsi à déterminer le scalaire u . L'équation précédente peut encore s'écrire sous la forme :

$$\begin{cases} x_I = x_C + uV'_x \\ y_I = y_C + uV'_y \\ z_I = z_C + uV'_z \end{cases} \quad (\text{C.3})$$

Il suffit d'injecter les équations du système (C.3) dans l'équation de l'ellipsoïde (C.2) et de trouver les racines d'un polynôme du second degré :

$$\alpha u^2 + \beta u + \gamma = 0$$

avec :

$$\begin{cases} \alpha = (a + h)^2 V_z'^2 + (b + h)^2 (V_x'^2 + V_y'^2) \\ \beta = 2 \left\{ (a + h)^2 z_C V_z'^2 + (b + h)^2 (x_C V_x' + y_C V_y') \right\} \\ \gamma = (a + h)^2 z_C^2 + (b + h)^2 (x_C^2 + y_C^2) - (a + h)^2 (b + h)^2 \end{cases} \quad (\text{C.4})$$

Les deux racines du polynôme de degré deux correspondent aux deux intersections de la visée avec l'ellipsoïde. Nous sommes à la recherche de l'intersection la plus proche du centre optique de la caméra, i.e. celle correspondant à la plus petite racine. La quantité α étant strictement positive, la plus petite racine est :

$$u_0 = \frac{-\beta - \sqrt{\beta^2 - 4\alpha\gamma}}{2\alpha}$$

L'intersection I d'altitude h est donc :

$$\begin{cases} x_I = x_C + u_0 V'_x \\ y_I = y_C + u_0 V'_y \\ z_I = z_C + u_0 V'_z \end{cases} \quad (\text{C.5})$$

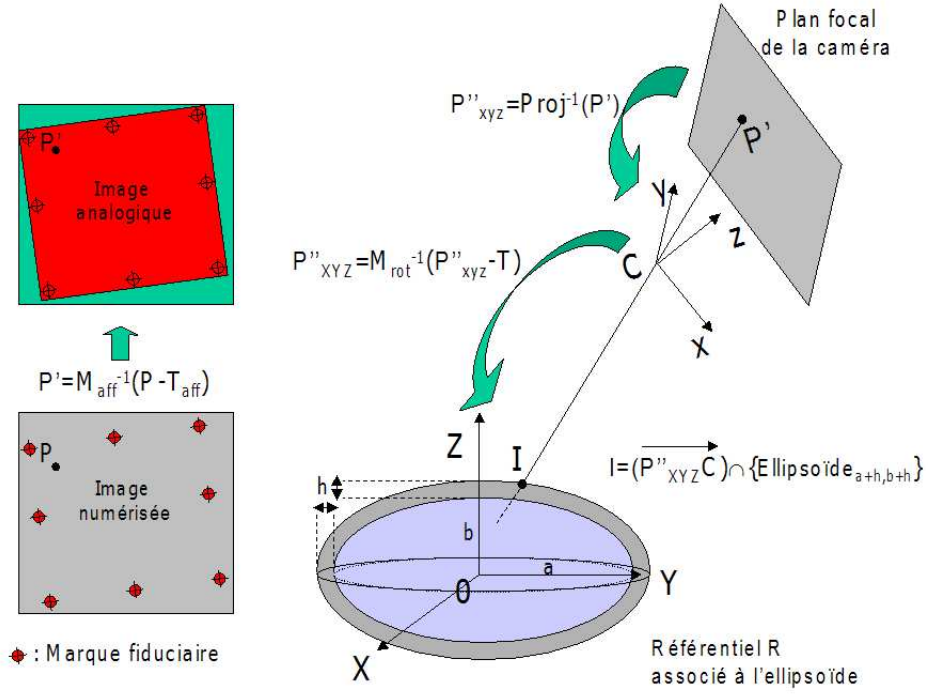


FIG. C.2 – Schéma illustrant la modélisation directe. Les repères terrain et caméra sont directs.

Annexe D

Calcul rapide de la moyenne et de la variance à l'intérieur et l'extérieur d'un contour actif représenté par un ensemble de niveaux

Comme le propose M. Rousson dans [103], il est inutile de calculer intégralement la moyenne et la variance des pixels de l'image à l'intérieur et l'extérieur du contour actif à chaque itération. En effet, cette opération nécessite d'être effectuée une seule fois au temps $t = 0$. Pour l'estimation ultérieure des quantités statistiques, il suffit de repérer les pixels de l'image qui entrent et sortent de la région délimitée par $\mathbf{C}(t)$ afin de corriger les valeurs de la moyenne et de la variance obtenues au temps antérieur. Soit $\Omega_{out,in}$ l'ensemble des pixels exclus du contour à l'instant t et inclus dans le contour à $t + 1$. Cet ensemble est de cardinal $n_{out,in}$. Soit $\Omega_{in,out}$ le dual de $\Omega_{out,in}$ et de cardinal $n_{in,out}$. La moyenne à l'intérieur du contour au temps $t + 1$ est par définition :

$$c_{in}(t+1) = \frac{1}{n_{in}(t+1)} \int_{\Omega_{in}(t+1)} I(\mathbf{x}) d\mathbf{x}$$

On exprime cette quantité en fonction de $c_{in}(t)$ de la façon suivante :

$$c_{in}(t+1) = \frac{1}{n_{in}(t) + n_{out,in} - n_{in,out}} \left\{ n_{in}(t) c_{in}(t) + \int_{\Omega_{out,in}} I(\mathbf{x}) d\mathbf{x} - \int_{\Omega_{in,out}} I(\mathbf{x}) d\mathbf{x} \right\} \quad (\text{D.1})$$

D'une façon similaire, on trouve l'expression de la moyenne à l'extérieur :

$$c_{out}(t+1) = \frac{1}{n_{out}(t) - n_{out,in} + n_{in,out}} \left\{ n_{out}(t) c_{out}(t) - \int_{\Omega_{out,in}} I(\mathbf{x}) d\mathbf{x} + \int_{\Omega_{in,out}} I(\mathbf{x}) d\mathbf{x} \right\} \quad (\text{D.2})$$

En remarquant que la variance d'une variable aléatoire X est :

$$\sigma^2(X) = E(X^2) - E(X)^2$$

On a alors

$$\sigma_{in}^2(t+1) = \frac{1}{n_{in}(t+1)} \int_{\Omega_{in}(t+1)} I^2(\mathbf{x}) d\mathbf{x} - c_{in}^2(t+1)$$

que l'on exprime en fonction de $\sigma_{in}^2(t)$ et de $c_{in}(t)$ grâce à (D.1) :

$$\begin{aligned} \sigma_{in}^2(t+1) = \frac{1}{n_{in}(t) + n_{out,in} - n_{in,out}} & \left\{ n_{in}(t) [c_{in}^2(t) + \sigma_{in}^2(t)] + \int_{\Omega_{out,in}} I^2(\mathbf{x}) d\mathbf{x} \right. \\ & \left. - \int_{\Omega_{in,out}} I^2(\mathbf{x}) d\mathbf{x} \right\} - c_{in}^2(t+1) \end{aligned} \quad (D.3)$$

L'expression duale $\sigma_{out}^2(t+1)$ est analogue :

$$\begin{aligned} \sigma_{out}^2(t+1) = \frac{1}{n_{out}(t) - n_{out,in} + n_{in,out}} & \left\{ n_{out}(t) [c_{out}^2(t) + \sigma_{out}^2(t)] - \int_{\Omega_{out,in}} I^2(\mathbf{x}) d\mathbf{x} \right. \\ & \left. + \int_{\Omega_{in,out}} I^2(\mathbf{x}) d\mathbf{x} \right\} - c_{out}^2(t+1) \end{aligned} \quad (D.4)$$

L'avantage d'une telle technique de calcul est de travailler uniquement sur les espaces $\Omega_{in,out}$ et $\Omega_{out,in}$ dont les cardinaux sont très faibles devant ceux de Ω_{in} et Ω_{out} . La réduction en temps de calcul est notable puisque la complexité du calcul devient proportionnelle à la longueur du contour actif. Dans [103], Rousson illustre expérimentalement le gain obtenu en temps de calcul.

Annexe E

Calcul de variations d'énergies de contrainte de forme et quadratique

E.1 Équation d'évolution dérivée de la fonctionnelle de contrainte de forme par différence de normales

Considérons le terme énergétique de contrainte de forme basé sur la différence entre la normale unitaire du contour actif implicitement représenté par l'ensemble de niveaux ϕ et celle d'une forme de référence représentée par ψ :

$$J_{shape}(\phi, \psi) = \int_{\Omega} |\nabla \phi(\mathbf{x}) - \nabla \psi(\mathbf{x})|^2 (H(\phi(\mathbf{x})) - H(\psi(\mathbf{x})))^2 d\mathbf{x} \quad (\text{E.1})$$

L'équation d'évolution consiste à trouver des valeurs de ϕ qui minimiseront la fonctionnelle $J_{shape}(\phi, \psi, t)$. Pour cela, on adopte le schéma classique de minimisation locale par descente de gradient :

$$\frac{\partial J_{shape}(\phi, \psi, t)}{\partial \phi} = -\phi_t(\mathbf{x}, t) \quad (\text{E.2})$$

L'expression de $\frac{\partial J_{shape}(\phi, \psi, t)}{\partial \phi}$ se déduit de la dérivée de Gâteaux de $J_{shape}(\phi, \psi, t)$ par rapport à ϕ dans la direction $\tilde{\phi}$:

$$\left(\frac{\partial J_{shape}(\phi, \psi, t)}{\partial \phi} \right)_{\phi=\tilde{\phi}} = \left\langle J_{shape}(\phi, \psi, t), \tilde{\phi} \right\rangle = \int_{\Omega} \frac{\partial J_{shape}(\phi, \psi, t)}{\partial \phi} \tilde{\phi}(\mathbf{x}, t) d\mathbf{x} \quad (\text{E.3})$$

Or la dérivée de Gâteaux de la fonctionnelle s'exprime sous la forme :

$$\left(\frac{\partial J_{shape}(\phi, \psi, t)}{\partial \phi} \right)_{\phi=\tilde{\phi}} = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \left(J_{shape}(\phi(\mathbf{x}, t) + \varepsilon \tilde{\phi}(\mathbf{x}, t), \psi, t) - J_{shape}(\phi(\mathbf{x}, t), \psi, t) \right) \quad (\text{E.4})$$

Détaillons le calcul de l'équation (E.4) :

$$\begin{aligned} J_{shape}(\phi(\mathbf{x}, t) + \varepsilon \tilde{\phi}(\mathbf{x}, t), \psi, t) = \\ \int_{\Omega} \left\langle \nabla \phi(\mathbf{x}) + \varepsilon \nabla \tilde{\phi}(\mathbf{x}, t) - \nabla \psi(\mathbf{x}), \nabla \phi(\mathbf{x}) + \varepsilon \nabla \tilde{\phi}(\mathbf{x}, t) - \nabla \psi(\mathbf{x}) \right\rangle \left(H(\phi(\mathbf{x}) + \varepsilon \tilde{\phi}(\mathbf{x}, t)) - H(\psi(\mathbf{x})) \right)^2 d\mathbf{x} \end{aligned} \quad (\text{E.5})$$

En remarquant que :

$$\delta_a(\phi(\mathbf{x}, t)) = \lim_{\epsilon \rightarrow 0} \frac{H_a(\phi(\mathbf{x}, t) + \epsilon \tilde{\phi}(\mathbf{x}, t)) - H_a(\phi(\mathbf{x}, t))}{\epsilon \tilde{\phi}(\mathbf{x}, t)} \quad (\text{E.6})$$

et en développant l'équation (E.5), on a :

$$\begin{aligned} J_{shape}(\phi(\mathbf{x}, t) + \epsilon \tilde{\phi}(\mathbf{x}, t), \psi, t) = \\ J_{shape}(\phi(\mathbf{x}, t)) \\ + 2\epsilon \int_{\Omega} \tilde{\phi}(\mathbf{x}, t) \delta_a(\phi(\mathbf{x}, t)) (H(\phi(\mathbf{x})) - H(\psi(\mathbf{x}))) \langle \nabla \phi(\mathbf{x}) - \nabla \psi(\mathbf{x}), \nabla \phi(\mathbf{x}) - \nabla \psi(\mathbf{x}) \rangle d\mathbf{x} \\ + 2\epsilon \int_{\Omega} (H(\phi(\mathbf{x})) - H(\psi(\mathbf{x})))^2 \langle \nabla \phi(\mathbf{x}) - \nabla \psi(\mathbf{x}), \nabla \tilde{\phi}(\mathbf{x}, t) \rangle \end{aligned} \quad (\text{E.7})$$

En intégrant par parties le dernier terme de l'équation (E.7), on trouve :

$$\begin{aligned} J_{shape}(\phi(\mathbf{x}, t) + \epsilon \tilde{\phi}(\mathbf{x}, t), \psi, t) - J_{shape}(\phi(\mathbf{x}, t)) = \\ + 2\epsilon \int_{\Omega} \tilde{\phi}(\mathbf{x}, t) \delta_a(\phi(\mathbf{x}, t)) (H(\phi(\mathbf{x})) - H(\psi(\mathbf{x}))) \left(|\nabla \phi(\mathbf{x})|^2 + |\nabla \psi(\mathbf{x})|^2 - 2 \langle \nabla \phi(\mathbf{x}), \nabla \psi(\mathbf{x}) \rangle \right) d\mathbf{x} \\ + 2\epsilon \left\{ \left[(H(\phi(\mathbf{x})) - H(\psi(\mathbf{x})))^2 (\phi_x(\mathbf{x}, t) + \phi_y(\mathbf{x}, t) - \psi_x(\mathbf{x}, t) - \psi_y(\mathbf{x}, t)) \tilde{\phi}(\mathbf{x}, t) \right]_{\Omega} \right. \\ + \int_{\Omega} \tilde{\phi}(\mathbf{x}, t) [2(H(\phi(\mathbf{x})) - H(\psi(\mathbf{x}))) ((\delta_a(\phi(\mathbf{x}, t)) + \delta_a(\psi(\mathbf{x}, t))) \langle \nabla \phi(\mathbf{x}), \nabla \psi(\mathbf{x}) \rangle \\ \left. \delta_a(\phi(\mathbf{x}, t)) |\nabla \phi(\mathbf{x})|^2 - \delta_a(\psi(\mathbf{x}, t)) |\nabla \psi(\mathbf{x})|^2 + (H(\phi(\mathbf{x})) - H(\psi(\mathbf{x})))^2 (\Delta \psi(\mathbf{x}) - \Delta \phi(\mathbf{x})) \right] \} \end{aligned} \quad (\text{E.8})$$

Le terme $\left[(H(\phi(\mathbf{x})) - H(\psi(\mathbf{x})))^2 (\phi_x(\mathbf{x}, t) + \phi_y(\mathbf{x}, t) - \psi_x(\mathbf{x}, t) - \psi_y(\mathbf{x}, t)) \tilde{\phi}(\mathbf{x}, t) \right]_{\Omega}$ est nul puisqu'il est toujours possible de trouver un espace rectangulaire Ω qui puisse inclure les formes représentées par ϕ et ψ . Ainsi, loin du niveau 0 des ensembles de niveaux, on a $H(\phi(\mathbf{x})) = 0 = H(\psi(\mathbf{x}))$. La différence des énergies devient :

$$\begin{aligned} J_{shape}(\phi(\mathbf{x}, t) + \epsilon \tilde{\phi}(\mathbf{x}, t), \psi, t) - J_{shape}(\phi(\mathbf{x}, t)) = \\ 2\epsilon \int_{\Omega} \tilde{\phi}(\mathbf{x}, t) \left\{ (H(\phi(\mathbf{x})) - H(\psi(\mathbf{x}))) \left(\delta_a(\phi(\mathbf{x}, t)) (|\nabla \psi(\mathbf{x})|^2 - |\nabla \phi(\mathbf{x})|^2) - 2\delta_a(\psi(\mathbf{x}, t)) |\nabla \psi(\mathbf{x})|^2 \right. \right. \\ \left. \left. + 2\delta_a(\psi(\mathbf{x}, t)) \langle \nabla \phi(\mathbf{x}), \nabla \psi(\mathbf{x}) \rangle + (H(\phi(\mathbf{x})) - H(\psi(\mathbf{x})))^2 (\Delta \psi(\mathbf{x}) - \Delta \phi(\mathbf{x})) \right) \right\} \end{aligned} \quad (\text{E.9})$$

Le gradient d'énergie vient alors directement, et en remarquant que les normales sont unitaires ($|\nabla \psi(\mathbf{x})|^2 = 1 = |\nabla \phi(\mathbf{x})|^2$) :

$$\begin{aligned} \frac{\partial J_{shape}(\phi, \psi, t)}{\partial \phi} = 4(H(\phi(\mathbf{x})) - H(\psi(\mathbf{x}))) \{ \delta_a(\psi(\mathbf{x}, t)) (\langle \nabla \phi(\mathbf{x}), \nabla \psi(\mathbf{x}) \rangle - 1) \} \\ + 2(H(\phi(\mathbf{x})) - H(\psi(\mathbf{x})))^2 (\Delta \psi(\mathbf{x}) - \Delta \phi(\mathbf{x})) \end{aligned} \quad (\text{E.10})$$

L'équation d'évolution est déduite par changement de signe selon l'équation (E.2).

E.2 Dérivée de l'énergie de correction quadratique

L'énergie introduite en équation (6.7) peut être scindée selon :

$$J_{\text{correction}}(\mathbf{C}) = J_1(C) - J_2(C) \quad (\text{E.11})$$

avec

$$J_1(C) = \oint \oint dp dp' \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 |\dot{\mathbf{C}}|^{-1} |\dot{\mathbf{C}}'|^{-1} \Psi(R) \quad (\text{E.12})$$

et

$$J_2(C) = \oint \oint dp dp' \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^4 |\dot{\mathbf{C}}|^{-3} |\dot{\mathbf{C}}'|^{-3} \Psi(R) \quad (\text{E.13})$$

E.2.1 Énergie $J_1(C)$

$$J_1(C + \delta\tilde{\mathbf{C}}) = \oint \oint dp dp' \langle \dot{\mathbf{C}} + \delta\dot{\tilde{\mathbf{C}}}, \dot{\mathbf{C}}' + \delta\dot{\tilde{\mathbf{C}}}' \rangle^2 |\dot{\mathbf{C}} + \delta\dot{\tilde{\mathbf{C}}}|^{-1} |\dot{\mathbf{C}}' + \delta\dot{\tilde{\mathbf{C}}}'|^{-1} \Psi(\tilde{R}) \quad (\text{E.14})$$

La quantité $\Psi(\tilde{R})$ se réécrit selon :

$$\Psi(\tilde{R}) \approx \Psi(R) + \langle \tilde{\mathbf{R}}, \delta\tilde{\mathbf{C}} - \delta\tilde{\mathbf{C}}' \rangle \Psi'(R) \quad (\text{E.15})$$

et $|\dot{\mathbf{C}} + \delta\dot{\tilde{\mathbf{C}}}|^{-1}$ est :

$$|\dot{\mathbf{C}} + \delta\dot{\tilde{\mathbf{C}}}|^{-1} \approx |\dot{\mathbf{C}}|^{-1} \left(1 - \langle \dot{\mathbf{C}}, \delta\dot{\tilde{\mathbf{C}}} \rangle |\dot{\mathbf{C}}|^{-2} \right) \quad (\text{E.16})$$

En développant le produit scalaire au carré, en remplaçant $\Psi(\tilde{R})$ par (E.15), en remplaçant $|\dot{\mathbf{C}} + \delta\dot{\tilde{\mathbf{C}}}|^{-1}$ par (E.16) et en négligeant les variations infinitésimales d'ordre deux, on a :

$$\begin{aligned} J_1(C + \delta\tilde{\mathbf{C}}) &= J_1(C) + 2 \oint \oint dp dp' \Psi'(R) \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 \langle \tilde{\mathbf{R}}, \delta\tilde{\mathbf{C}} \rangle |\dot{\mathbf{C}}|^{-1} |\dot{\mathbf{C}}'|^{-1} \\ &\quad - 2 \oint \oint dp dp' \Psi(R) \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 |\dot{\mathbf{C}}|^{-3} |\dot{\mathbf{C}}'|^{-1} \langle \dot{\mathbf{C}}, \delta\dot{\tilde{\mathbf{C}}} \rangle \\ &\quad + 4 \oint \oint dp dp' \Psi(R) \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle |\dot{\mathbf{C}}|^{-1} |\dot{\mathbf{C}}'|^{-1} \langle \dot{\mathbf{C}}', \delta\dot{\tilde{\mathbf{C}}} \rangle \end{aligned} \quad (\text{E.17})$$

En intégrant le dernier terme par parties, on trouve¹ :

$$\begin{aligned} J_1(C + \delta\tilde{\mathbf{C}}) - J_1(C) &= 2 \oint \oint dp dp' \Psi'(R) \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 |\dot{\mathbf{C}}|^{-1} |\dot{\mathbf{C}}'|^{-1} \langle \tilde{\mathbf{R}}, \delta\tilde{\mathbf{C}} \rangle \\ &\quad + 2 \oint \oint dp dp' \left[\Psi(R) \left(2 \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle \langle \ddot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle |\dot{\mathbf{C}}|^{-3} |\dot{\mathbf{C}}'|^{-1} - 3 |\dot{\mathbf{C}}|^{-5} |\dot{\mathbf{C}}'|^{-1} \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 \langle \ddot{\mathbf{C}}, \dot{\mathbf{C}} \rangle \right) \right. \\ &\quad \left. + \Psi'(R) \langle \dot{\mathbf{C}}, \tilde{\mathbf{R}} \rangle \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 |\dot{\mathbf{C}}|^{-3} |\dot{\mathbf{C}}'|^{-1} \right] \langle \dot{\mathbf{C}}, \delta\tilde{\mathbf{C}} \rangle \end{aligned}$$

¹Ce résultat est obtenu en injectant : $\frac{\partial \Psi}{\partial p} = \Psi'(R) \langle \tilde{\mathbf{R}}, \dot{\mathbf{C}} \rangle$

$$\begin{aligned}
& +2 \oint \oint dp dp' \Psi(R) \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 |\dot{\mathbf{C}}|^{-3} |\dot{\mathbf{C}}'|^{-1} \langle \ddot{\mathbf{C}}, \delta \tilde{\mathbf{C}} \rangle \\
& -4 \oint \oint dp dp' \left[\Psi(R) \left(\langle \dot{\mathbf{C}}', \ddot{\mathbf{C}} \rangle |\dot{\mathbf{C}}|^{-1} |\dot{\mathbf{C}}'|^{-1} - \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle \langle \ddot{\mathbf{C}}, \dot{\mathbf{C}} \rangle |\dot{\mathbf{C}}|^{-3} |\dot{\mathbf{C}}'|^{-1} \right) \right. \\
& \quad \left. + \Psi'(R) \langle \dot{\mathbf{C}}, \tilde{\mathbf{R}} \rangle \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle |\dot{\mathbf{C}}|^{-1} |\dot{\mathbf{C}}'|^{-1} \right] \langle \dot{\mathbf{C}}', \delta \tilde{\mathbf{C}} \rangle
\end{aligned} \tag{E.18}$$

E.2.2 Énergie $J_2(C)$

$$J_2(C + \delta \tilde{\mathbf{C}}) = \oint \oint dp dp' \langle \dot{\mathbf{C}} + \delta \dot{\mathbf{C}}, \dot{\mathbf{C}}' + \delta \dot{\mathbf{C}}' \rangle^4 |\dot{\mathbf{C}} + \delta \dot{\mathbf{C}}|^{-3} |\dot{\mathbf{C}}' + \delta \dot{\mathbf{C}}'|^{-3} \Psi(\tilde{R}) \tag{E.19}$$

avec

$$|\dot{\mathbf{C}} + \delta \dot{\mathbf{C}}|^{-3} \approx |\dot{\mathbf{C}}|^{-3} \left(1 - 3 \langle \dot{\mathbf{C}}, \delta \dot{\mathbf{C}} \rangle |\dot{\mathbf{C}}|^{-2} \right)$$

En développant le produit scalaire, en injectant le résultat de (E.15) et en négligeant les termes d'ordre deux, on a :

$$\begin{aligned}
J_2(C + \delta \tilde{\mathbf{C}}) - J_2(C) &= 2 \oint \oint dp dp' \Psi'(R) \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^4 |\dot{\mathbf{C}}|^{-3} |\dot{\mathbf{C}}'|^{-3} \langle \tilde{\mathbf{R}}, \delta \tilde{\mathbf{C}} \rangle \\
&\quad - 6 \oint \oint dp dp' \Psi(R) \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^4 |\dot{\mathbf{C}}|^{-5} |\dot{\mathbf{C}}'|^{-3} \langle \dot{\mathbf{C}}, \delta \dot{\mathbf{C}} \rangle \\
&\quad + 8 \oint \oint dp dp' \Psi(R) \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^3 |\dot{\mathbf{C}}|^{-3} |\dot{\mathbf{C}}'|^{-3} \langle \dot{\mathbf{C}}', \delta \dot{\mathbf{C}} \rangle
\end{aligned} \tag{E.20}$$

En intégrant par parties, il vient :

$$\begin{aligned}
J_2(C + \delta \tilde{\mathbf{C}}) - J_2(C) &= 2 \oint \oint dp dp' \Psi'(R) \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^4 |\dot{\mathbf{C}}|^{-3} |\dot{\mathbf{C}}'|^{-3} \langle \tilde{\mathbf{R}}, \delta \tilde{\mathbf{C}} \rangle \\
&+ 2 \oint \oint dp dp' \left[\Psi(R) \left(12 \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^3 \langle \ddot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle |\dot{\mathbf{C}}|^{-5} |\dot{\mathbf{C}}'|^{-3} - 15 |\dot{\mathbf{C}}|^{-7} |\dot{\mathbf{C}}'|^{-3} \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^4 \langle \ddot{\mathbf{C}}, \dot{\mathbf{C}} \rangle \right) \right. \\
&\quad \left. + 3 \Psi'(R) \langle \dot{\mathbf{C}}, \tilde{\mathbf{R}} \rangle \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^4 |\dot{\mathbf{C}}|^{-5} |\dot{\mathbf{C}}'|^{-3} \right] \langle \dot{\mathbf{C}}, \delta \tilde{\mathbf{C}} \rangle \\
&\quad + 6 \oint \oint dp dp' \Psi(R) \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^4 |\dot{\mathbf{C}}|^{-5} |\dot{\mathbf{C}}'|^{-3} \langle \ddot{\mathbf{C}}, \delta \tilde{\mathbf{C}} \rangle \\
&- 4 \oint \oint dp dp' \left[\Psi(R) \left(6 \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 \langle \dot{\mathbf{C}}', \ddot{\mathbf{C}} \rangle |\dot{\mathbf{C}}|^{-3} |\dot{\mathbf{C}}'|^{-3} - 6 \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^3 \langle \ddot{\mathbf{C}}, \dot{\mathbf{C}} \rangle |\dot{\mathbf{C}}|^{-5} |\dot{\mathbf{C}}'|^{-3} \right) \right. \\
&\quad \left. + 2 \Psi'(R) \langle \dot{\mathbf{C}}, \tilde{\mathbf{R}} \rangle \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^3 |\dot{\mathbf{C}}|^{-3} |\dot{\mathbf{C}}'|^{-3} \right] \langle \dot{\mathbf{C}}', \delta \tilde{\mathbf{C}} \rangle
\end{aligned} \tag{E.21}$$

E.2.3 Variation totale de l'énergie $J_{\text{correction}}(C)$

En groupant les résultats de (E.18) et (E.21), on a :

$$J_{\text{correction}}(\mathbf{C} + \delta \tilde{\mathbf{C}}) - J_{\text{correction}}(\mathbf{C}) = \oint \oint dp dp' \Psi(R) \{$$

$$\begin{aligned}
& 2 \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 |\dot{\mathbf{C}}|^{-3} |\dot{\mathbf{C}}'|^{-1} \left(1 - 3 \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 |\dot{\mathbf{C}}|^{-2} |\dot{\mathbf{C}}'|^{-2} \right) \langle \ddot{\mathbf{C}}, \delta \tilde{\mathbf{C}} \rangle \\
& + 4 \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle \langle \ddot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle |\dot{\mathbf{C}}|^{-3} |\dot{\mathbf{C}}'|^{-1} \left(1 - 6 \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 |\dot{\mathbf{C}}|^{-2} |\dot{\mathbf{C}}'|^{-2} \right) \langle \dot{\mathbf{C}}, \delta \tilde{\mathbf{C}} \rangle \\
& - 6 \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 \langle \ddot{\mathbf{C}}, \dot{\mathbf{C}} \rangle |\dot{\mathbf{C}}|^{-5} |\dot{\mathbf{C}}'|^{-1} \left(1 - 5 \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 |\dot{\mathbf{C}}|^{-2} |\dot{\mathbf{C}}'|^{-2} \right) \langle \dot{\mathbf{C}}, \delta \tilde{\mathbf{C}} \rangle \\
& - 4 \left(1 - 6 \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 |\dot{\mathbf{C}}|^{-2} |\dot{\mathbf{C}}'|^{-2} \right) |\dot{\mathbf{C}}|^{-1} |\dot{\mathbf{C}}'|^{-1} \left[\langle \ddot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle - \langle \ddot{\mathbf{C}}, \dot{\mathbf{C}} \rangle \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle |\dot{\mathbf{C}}|^{-2} \right] \langle \dot{\mathbf{C}}', \delta \tilde{\mathbf{C}} \rangle \\
& + \oint \oint dp dp' \Psi'(R) \left\{ 2 \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 |\dot{\mathbf{C}}|^{-1} |\dot{\mathbf{C}}'|^{-1} \left(1 - \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 |\dot{\mathbf{C}}|^{-2} |\dot{\mathbf{C}}'|^{-2} \right) \langle \tilde{\mathbf{R}}, \delta \tilde{\mathbf{C}} \rangle \right. \\
& \quad + 2 \langle \dot{\mathbf{C}}, \tilde{\mathbf{R}} \rangle \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 |\dot{\mathbf{C}}|^{-3} |\dot{\mathbf{C}}'|^{-1} \left(1 - 3 \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 |\dot{\mathbf{C}}|^{-2} |\dot{\mathbf{C}}'|^{-2} \right) \langle \dot{\mathbf{C}}, \delta \tilde{\mathbf{C}} \rangle \\
& \quad \left. - 4 \langle \dot{\mathbf{C}}, \tilde{\mathbf{R}} \rangle \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle |\dot{\mathbf{C}}|^{-1} |\dot{\mathbf{C}}'|^{-1} \left(1 - 2 \langle \dot{\mathbf{C}}, \dot{\mathbf{C}}' \rangle^2 |\dot{\mathbf{C}}|^{-2} |\dot{\mathbf{C}}'|^{-2} \right) \langle \dot{\mathbf{C}}', \delta \tilde{\mathbf{C}} \rangle \right\} \quad (\text{E.22})
\end{aligned}$$

En remarquant que $\langle \ddot{\mathbf{C}}, \dot{\mathbf{C}} \rangle = 0$, certains termes s'annulent. En supposant que p est l'abscisse curviligne, cette dernière équation se simplifie et revient à (6.9).

Annexe F

Représentation 3D d'une scène urbaine

F.1 Objectif et rappel de la méthodologie

Les résultats du chapitre 3 ont montré que nous étions en mesure de générer un Modèle Numérique de Surface (MNS) orthorectifié d'un paysage urbain. Cette information d'altitude peut être transférée dans la carte numérique du bâti afin de la rendre à trois dimensions. Ce chapitre montre l'élaboration d'une représentation 3D des bâtiments de moyen niveau et programmée en VRML (*Virtual Reality Modeling Language*) par Andéol Ayzac, stagiaire au LIAMA d'avril à août 2004.

F.2 Principe de la modélisation

Les bâtiments sont modélisés par des parallélépipèdes dont la base polygonale est dérivée de la carte. La hauteur assignée à chaque bâtiment est la moyenne des altitudes du MNS à l'intérieur du polygone cartographique. Une telle modélisation est clairement insuffisante pour la seule finalité d'une représentation réaliste (visite virtuelle d'une ville). Néanmoins, elle est entièrement automatique et suffit à la modélisation de pollution atmosphérique ou sonore.

Pour rendre la scène plus réaliste qu'un amas de "boîtes extrudées", cette dernière peut être habillée d'une image orthoscopique réalisée conjointement au MNS, et de textures au niveau des façades. Une texture de façade est une collection de motifs identiques dont la densité par pixel est un paramètre du programme. Les motifs sont soit de vraies photographies de façades de Pékin, soit des images de synthèse. Il existe approximativement une dizaine de classes de motifs placés aléatoirement sur les bâtiments. Une dernière option du programme est la possibilité de modéliser les variations du terrain si un Modèle Numérique de Terrain (MNT) est disponible. Ce dernier est alors modélisé par des B-splines dont l'effet lissant n'est pas gênant en milieu urbain car le relief y est rarement accidenté.

F.3 Résultats sur la ville de Pékin

La figure F.1 illustre la représentation 3D du complexe sportif construit pour les jeux asiatiques organisés à Pékin en 1990.

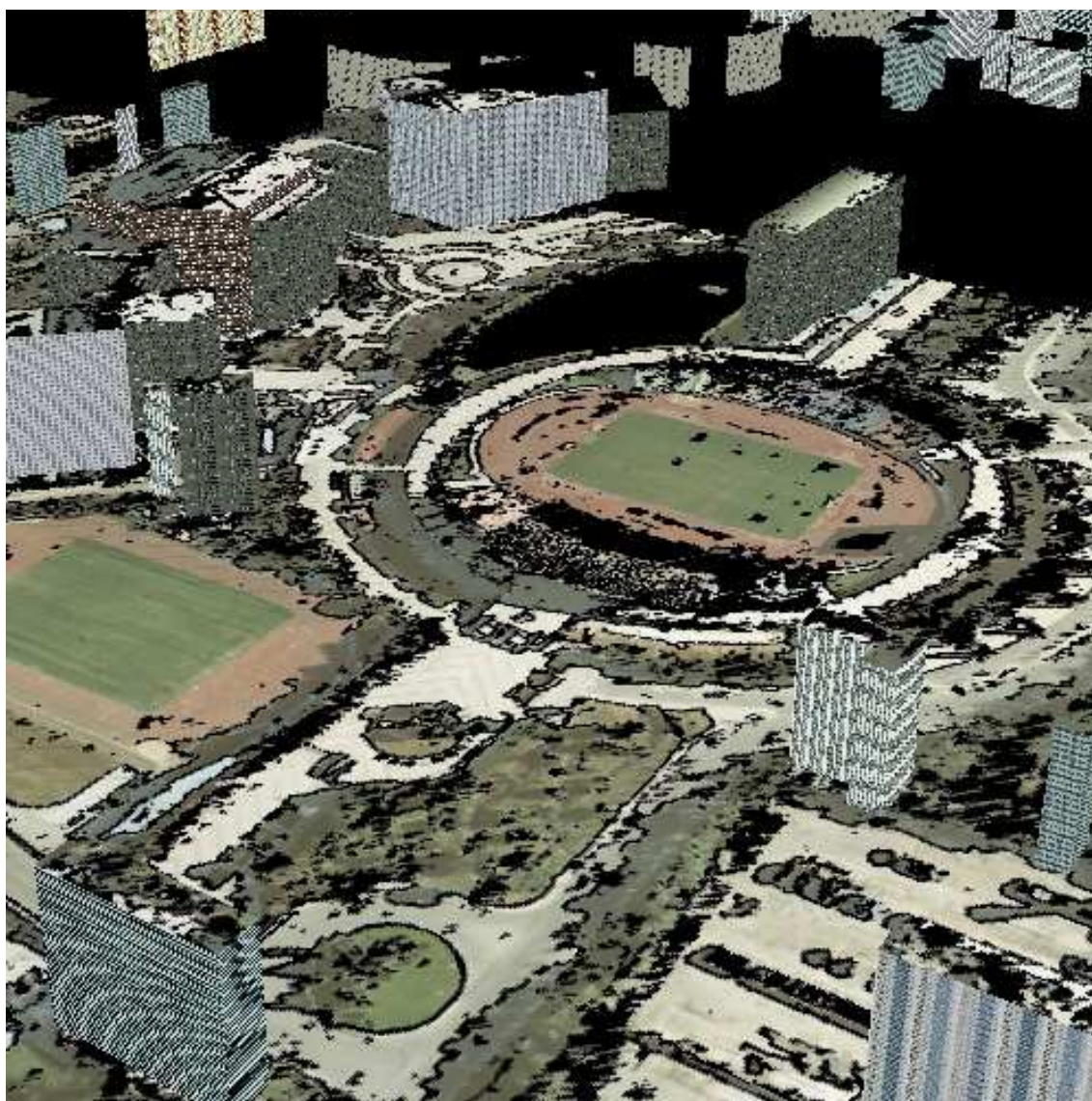


FIG. F.1 – Visualisation 3D de la ville de Pékin - avec textures et sans MNT.

Bibliographie

- [1] D. Adalsteinsson and J. Sethian, *The fast construction of extension velocities in level set methods*, Journal of Computational Physics **148** (1999), 2–22.
- [2] P. Agouris, K. Beard, G. Mountrakis, and A. Stefanidis, *Capturing and modeling geographic object change : a spatiotemporal gazetteer framework*, Photogrammetric Engineering & Remote Sensing **66** (2000), no. 10, 1224–1250.
- [3] P. Agouris, G. Mountrakis, and A. Stefanidis, *Automated spatiotemporal change detection in digital aerial imagery*, Aerosense 2000, SPIE Proceedings, vol. 4054, 2000, pp. 2–12.
- [4] P. Agouris, A. Stefanidis, and S. Gyftakis, *Differential snakes for change detection in road segments*, Photogrammetric Engineering & Remote Sensing **67** (2001), no. 12, 1391–1399.
- [5] G. Ball and D.J. Hall, *A novel method of data analysis and pattern classification*, Tech. report, Stanford Research Institute, Menlo Park, CA, USA, 1965.
- [6] E. Baltsavias, S. Mason, and D. Stallmann, *Use of DTMs/DSMs and orthoimages to support building extraction*, Proc. Workshop on "Automatic extraction of man-made objects from aerial and space images" (A. Gruen O. Kuebler, P. Agouris, ed.), Birkhauser Verlag, 1995.
- [7] E. Baltsavias, L. Zhang, and H. Eisenbeiss, *DSM generation and interior orientation determination of Ikonos images using a testfield in Switzerland*, Proc. ISPRS Hannover Workshop on "High-Resolution Earth Imaging for Geospatial Information", 2005.
- [8] E. P. Baltsavias, A. Grün, and L. Van Gool (eds.), *Automatic extraction of man-made objects from aerial and space images*, vol. 3, A.A. Balkema, 2001.
- [9] E.P. Baltsavias, *Object extraction and revision by image analysis using existing geospatial data and knowledge : State-of-the-art and steps towards operational systems*, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **34** (2002), no. 2, 13–22.
- [10] J. Barron, D. Fleet, and S. Beauchemin, *Performance of optical flow techniques*, Int. J. Comput. Vis **12** (1994), no. 1, 43–77.
- [11] M. Bäümker and F.J. Heimes, *New calibration and computing method for direct georeferencing of image and scanner data using the position and angular data of an hybrid inertial navigation system*, Proc. OEEPE-Workshop on Integrated Sensor Orientation, 2001.
- [12] L. Bentabet, S. Jodouin, D. Ziou, and J. Vaillancourt, *Road vectors update using SAR imagery : a snake based method*, IEEE Transactions in Geoscience and Remote Sensing **41** (2003), no. 8, 1785–1803.
- [13] A. Blake and M. Isard, *Active contours*, Springer-Verlag, 1998.
- [14] P. Blanc, *Développement de méthodes pour la détection de changement*, Ph.D. thesis, Ecole des Mines de Paris, 1999.
- [15] D. Boldo and H. Le Men, *Remote sensing model adaptation to very high resolution digital images of urban areas*, IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, 2001, pp. 20–24.

- [16] A. Busch, *Revision of built-up areas in a GIS using satellite imagery and GIS data*, Proc. ISPRS Symposium on GIS - Between visions and applications, vol. 32/4, 1998.
- [17] V. Caselles, F. Catte, T. Coll, and F. Dibos, *A geometric model for active contours*, Numerische Mathematik **66** (1993), no. 1, 1–31.
- [18] V. Caselles, R. Kimmel, and G. Sapiro, *On geodesic active contours*, International Journal of Computer Vision **22** (1997), no. 1, 61–79.
- [19] T. Chan and L. Vese, *Active contours without edges*, IEEE Trans. on Image Processing **10** (2001), no. 2, 266–277.
- [20] ———, *A level set algorithm for minimizing the Mumford-Shah functional in image processing*, 1st IEEE Workshop on Variational and Level Set Methods in Computer Vision, 2001, pp. 161–168.
- [21] T. Chan and W. Zhu, *Level set based shape prior segmentation*, Tech. report, UCLA, 2003.
- [22] Y. Chen, F. Huang, H. Tagare, M. Rao, D. Wilson, and E. A. Geiser, *Using prior shape and intensity profile in medical image segmentation*, Proc. ICCV'2003, vol. 2, 2003, p. 1117.
- [23] Y. Chen, H. Tagare, S. Thiruvankadam, F. Huang, D. Wilson, K. Gopinath, R. Briggs, and E. Geiser, *Using prior shapes in geometric active contours in a variational framework*, International Journal of Computer Vision **50** (2002), no. 3, 315–328.
- [24] Y. Chen, S. Thiruvankadam, H. Tagare, F. Huang, D. Wilson, and E. A. Geiser, *On the incorporation of shape priors into geometric active contours*, IEEE 1st Workshop on Variational Framework and Level Sets methods, 2001, pp. 145–152.
- [25] Y. Chen, S.R. Thiruvankadam, K.S. Gopinath, and R.W. Briggs, *FMR image registration using the Mumford-Shah functional and shape information*, Intl Conference on Information Systems, Analysis and Synthesis, 2002.
- [26] D. L. Chopp, *Computing minimal surfaces via level set curvature flow*, Journal of Computational Physics **106** (1993), no. 1, 77–91.
- [27] L. D. Cohen, *On active contour models and balloons*, Computer Vision, Graphics, and Image Processing. Image Understanding **53** (1991), no. 2, 211–218.
- [28] T.F. Cootes, D. Cooper, C.J. Taylor, and J. Graham, *Active shape models - their training and application*, Computer Vision and Image Understanding **61** (1995), no. 1, 38–59.
- [29] D. Cremers, *Statistical shape knowledge in variational image segmentation*, Ph.D. thesis, Department of Mathematics and Computer Science, University of Mannheim, Germany, 2002.
- [30] D. Cremers, S. Osher, and S. Soatto, *Kernel density estimation and intrinsic alignment for knowledge-driven segmentation : teaching level sets to walk*, Proc. DAGM'04, 26th Pattern Recognition Symposium, LNCS 3175, Springer, 2004, pp. 36–44.
- [31] D. Cremers and S. Soatto, *A pseudo-distance for shape priors in level set segmentation*, 2nd IEEE Intl. Workshop on Variational, Geometric and Level Set (2nd IEEE Intl. Workshop on Variational Geometric O. Faugeras, N. Paragios (Eds.) and O. Faugeras N. Paragios (Eds.) Level Set Methods (VLSM), Nice, eds.), 2003, pp. 169–176.
- [32] D. Cremers, N. Sochen, and C. Schnörr, *Towards recognition-based variational segmentation using shape priors and dynamic labeling*, Intl. Conf. on Scale-Space Theories in Computer Vision, LNCS 2695, Springer, 2003, pp. 388–400.
- [33] D. Cremers, F. Tischhauser, J. Weickert, and C. Schnörr, *Diffusion snakes : introducing statistical shape knowledge into the mumford-shah functional*, International Journal of Computer Vision **50** (2002), no. 3, 295–313.

- [34] E. P. Crist and R. C. Cicone, *A physically-based transformation of Thematic Mapper data - The TM tasseled cap*, IEEE Trans. Geosci. Remote Sens. **22** (1984), no. 2, 256–263.
- [35] M.E de Gunst, *Knowledge based interpretation of aerial images for updating of road maps*, Ph.D. thesis, Technical University of Delft, 1996.
- [36] M.E de Gunst and J.E. den Hartog, *Knowledge-based updating of maps by interpretation of aerial images*, Proc. ICPR, 1994, pp. 811–814.
- [37] Y. Dufournaud, C. Schmid, and R. Horaud, *Image matching with scale adjustment*, Computer Vision and Image Understanding **93** (2004), 175–194.
- [38] Ch. Eidenbenz, Ch. Kaeser, and E. Baltsavias, *ATOMI - Automated reconstruction of topographic objects from aerial images using vectorized map information*, International Archives of Photogrammetry and Remote Sensing, vol. 23, 2000.
- [39] EUROCONTROL and IfEN, *WGS84 implementation manual, version 2.4*, 1998.
- [40] A. Foulonneau, *Une contribution à l'introduction de contraintes géométriques dans les contours actifs orientés région*, Ph.D. thesis, Université Louis Pasteur - Strasbourg I, December 2004.
- [41] A. Foulonneau, P. Charbonnier, and F. Heitz, *Geometric shape priors for region-based active contours*, IEEE Int. Conf. Image Processing, ICIP 2003, September 2003.
- [42] M. Gerke and A. Busch, *Verification of a digital road database using Ikonos imagery*, Proc. ISPRS Hannover Workshop on "High-Resolution Earth Imaging for Geospatial Information", May 2005.
- [43] J. Gomes and O. Faugeras, *Reconciling distance functions and level sets*, Tech. Report 3666, INRIA, 1999.
- [44] A. Grün, *Adaptative least square correlation : a powerful image matching technique*, South African Journal of Photogrammetry, Remote Sensing and Cartography **14** (1985), no. 3, 175–187.
- [45] A. Grün, E.P. Baltsavias, and O. Henricsson (eds.), *Automatic extraction of man-made objects from aerial and space images*, vol. 2, Birkhauser Verlag Basel, 1997.
- [46] T. Guo and Y. Yasuoka, *Combining high resolution satellite imagery and airborne laser scanning data for generating bare land DEM in urban areas*, Proc. International Workshop on Visualization and Animation Of Landscape, 2002.
- [47] ———, *Snake-based approach for building extraction from high-resolution satellite images and height data in urban areas*, Proc. of the 23rd Asian Conference on Remote Sensing, 2002.
- [48] N. Haala and V. Walter, *Automatic classification of urban environments for database revision using LIDAR and color aerial imagery*, IAPRS, vol. 32, 1999.
- [49] S. Haker, S. Angenent, A. Tannenbaum, and R. Kikinis, *Nondistorting flattening maps and the 3D visualisation of colon CT images*, Proc. MICCAI (2000), 358–366.
- [50] F.J. Harris, *On the use of windows for harmonic analysis with the discrete fourier transform*, Proc. IEEE, vol. 66, pp. 51–88 (1978).
- [51] J.J Helmsen, *A comparison of the three-dimensional photolithography development methods*, Ph.D. thesis, EECS, University of California, Berkeley, 1994.
- [52] P.V.C. Hough, *Machine analysis of bubble chamber pictures*, International Conference on High Energy Accelerators and Instrumentation, CERN, 1959.
- [53] A. Huertas and R. Nevatia, *Detecting changes in aerial views of manmade structures*, Image Visi. Comput. **18** (2000), no. 8, 583–596.
- [54] O. Jamet, *Comparaison vecteurs-images pour la détection de changements d'une base de données topographiques*, Ph.D. thesis, ENST, 1998.

- [55] R. Jeansoulin, *Les images multi-sources en télédétection*, Ph.D. thesis, Université Paul Sabatier, Toulouse, France, 1982.
- [56] S. Jehan-Besson, *Modèles de contours actifs basés régions pour la segmentation d'images et de vidéos*, Ph.D. thesis, Université de Nice-Sophia Antipolis, 2003.
- [57] J. Jensen, *Introductory digital image processing, a remote sensing perspective*, Upper Saddle River, NJ : Prentice-Hall, 1996.
- [58] M. Kass, A. Witkin, and D. Terzopoulos, *Snakes : active contour models*, 1st International Conference on Computer Vision, pp. 259-268 (1987).
- [59] R. Keck, *Reinitialization for level set methods*, Ph.D. thesis, University of Kaiserslautern, 1998.
- [60] R. Keriven, *Equations aux dérivées partielles, evolutions de courbes et de surfaces et espaces d'échelle : Applications à la vision par ordinateur*, Ph.D. thesis, Ecole Nationale des Ponts et Chaussées, December 1997.
- [61] T. Knudsen and A. Nielsen, *Detection of buildings through multivariate analysis of spectral, textural, and shape based features*, Proc. IGARSS, 2004.
- [62] T. Knudsen and B. Olsen, *Automated change detection for updates of digital map databases*, Photogrammetric Engineering & Remote Sensing **69** (2003), no. 11, 1289–1296.
- [63] O. Kuebler, P. Agouris, and A. Grün (eds.), *Automatic extraction of man-made objects from aerial and space images*, Birkhauser Verlag, 1995.
- [64] F. Laporterie, *Evaluation of the quality of panchromatic / multispectral fusion algorithms performed on images simulating the future pleiades satellites*, Proc. 2nd GRSS/ISPRS joint workshop on remote sensing data fusion over urban areas (URBAN 2003), May 2003.
- [65] J. Leitloff, S. Hinz, and U. Stilla, *Vehicle queue detection in complex urban areas by extraction and analysis of linear features*, Proc. ISPRS Hanover Workshop on "High-Resolution Earth Imaging for Geospatial Information", 2005.
- [66] M. Leventon, E. Grimson, and O. Faugeras., *Statistical shape influence in geodesic active contours*, Comp. Vision and Patt. Recon. (CVPR), 2000, pp. 316–323.
- [67] R. Lillestrand, *Techniques for change detection*, IEEE Trans. Comput. **21** (1972), no. 7, 654–659.
- [68] M. and E. Fatemi, *An efficient, interface-preserving level set redistancing algorithm and its application to interfacial incompressible fluid flow*, SIAM J. Sci. Comput. **20** (1999), no. 4, 1165–1191.
- [69] R. Malladi and J.A Sethian, *Level set methods for curvature flow, image enhancement, and shape recovery in medical images*, Proc. Conf. on Visualization and Mathematics, 1995, pp. 329–345.
- [70] R. Malladi, J.A. Sethian, and B.C. Vermuri, *Shape modeling with front propagation : a level set approach*, IEEE Trans. on PAMI **17** (1995), no. 2, 158–175.
- [71] C.R. Maurer, R. Qi, and V. Raghavan, *A linear time algorithm for computing exact euclidean distance transforms of binary images in arbitrary dimensions*, IEEE Transactions on Pattern Analysis and Machine Intelligence **25** (2003), no. 2, 265–270.
- [72] H. Mayer, *Automatic object extraction from aerial imagery - A survey focusing on buildings*, Comput. Vis. Image Underst. **74** (1999), no. 2, 138–149.
- [73] T. McInerney and D. Terzopoulos, *T-snakes : Topology adaptive snakes*, Medical Image Analysis **4** (2000), no. 2, 73–91.

- [74] J. B. Mena, *State of the art on automatic road extraction for GIS update : a novel classification*, Pattern Recogn. Lett. **24** (2003), no. 16, 3037–3058.
- [75] D. Mumford and J. Shah, *Boundary detection by minimizing functionals*, Proc. CVPR'85, 1985, pp. 22–26.
- [76] S. Negahdaripour, *Revised definition of optical flow : Integration of radiometric and geometric cues for dynamic scene analysis*, IEEE Trans. Pattern Anal. Mach. Intell. **20** (1998), no. 9, 961–979.
- [77] J.A. Nelder and R. Mead, *A simplex method for function minimization*, Computer Journal **7** (1965), no. 4, 308–313.
- [78] R. Nevatia, C. Lin, and A. Huertas, *A system for building detection from aerial images*, Workshop on "Automatic extraction of man-made objects from aerial and space images", Birkhauser Verlag Basel, 1997, pp. 77–86.
- [79] M. Niederöst, *Reliable reconstruction of buildings for digital map revision*, International Archives of Photogrammetry and Remote Sensing (Amsterdam), vol. 23, 2000.
- [80] ———, *Automated update of building information in maps using medium-scale imagery (1 : 15,000)*, Proc. Workshop on "Automatic Extraction of Man-Made Objects from Aerial and Space Images " (E. P. Baltsavias, A. Grün, and L. Van Gool, eds.), vol. 3, A.A. Balkema, 2001.
- [81] S. Noronha and R. Nevatia, *Detection and description of buildings from multiple aerial images*, DARPA Image Understanding Workshop, 1996, pp. 469–478.
- [82] K. Novak, *Rectification of digital imagery*, Photogrammetric Engineering & Remote Sensing, vol. 58, no 3, pp. 339–344 (1992).
- [83] B.P. Olsen, *Automatic change detection for validation of digital map databases*, ISPRS XXth congress, IAPRS, vol. 24, 2004, pp. 120–125.
- [84] H. Oriot, *Statistical snakes for building extraction from stereoscopic aerial images*, Proc. of the ISPRS Workshop on "Photogrammetric image analysis" (2003).
- [85] S. Osher and J.A. Sethian, *Fronts propagating with curvature-dependent speed : algorithms based on Hamilton-Jacobi formulations*, Journal of Computational Physics, 79, pp. 12–49 (1988).
- [86] N. Paparoditis, *Reconstruction tridimensionnelle de paysages urbains en imagerie stéréoscopique satellitale haute résolution*, Ph.D. thesis, Université de Nice-Sophia Antipolis, 1998.
- [87] N. Paparoditis, M. Cord, M. Jordan, and J-P. Cocquerez, *Building detection and reconstruction from mid- and high-resolution aerial imagery*, Computer Vision and Image Understanding, Vol. 72, Issue 2, pp. 122–142 (1998).
- [88] N. Paragios, *Geodesic active regions and level set methods : contributions and applications in artificial vision*, Ph.D. thesis, Université de Nice-Sophia Antipolis, January 2000.
- [89] N. Paragios and R. Deriche, *A PDE-based level-set approach for detection and tracking of moving objects*, Tech. Report 3172, INRIA, 1997.
- [90] ———, *Geodesic active contours and level sets for the detection and tracking of moving objects*, IEEE Transactions on Pattern Analysis and Machine Intelligence **22** (2000), no. 3, 266–280.
- [91] ———, *Geodesic active regions : A new paradigm to deal with frame partition problems in Computer Vision*, Journal of Visual Communication and Image Representation **13** (2002), no. 1/2, 249–268.
- [92] N. Paragios, O. Mellina-Gottardo, and V. Ramesh., *Gradient vector flow fast geometric active contours*, IEEE Transactions on Pattern Analysis and Machine Intelligence **26** (2004), no. 3, 402–407.

- [93] N. Paragios, M. Rousson, and V. Ramesh, *Matching distance functions : A shape-to-area variational approach for global-to-local registration*, European Conference in Computer Vision, LNCS 2351, Springer-Verlag, 2002, pp. 775–789.
- [94] T. Pavlidis, *Algorithms for graphics and image processing*, Computer Science Press, Inc., 1982.
- [95] R. Péteri and T. Ranchin, *Extraction and update of street network in urban areas from high resolution satellite images*, Proc. ISPRS Commission IV symposium on "Geospatial Theory, Processing and Applications" (C. Armenakis and Y.C. Lee, eds.), vol. 14, 2002, pp. 768–773.
- [96] ———, *Multiresolution Snakes for urban road extraction from IKONOS and Quickbird images*, 23rd EARSeL Annual Symposium "Remote Sensing in Transition" (Goossens, ed.), Millpress, Rotterdam, Netherlands, 2003, pp. 69–76.
- [97] ———, *Urban street mapping using quickbird and ikonos images*, Proc. IGARSS, vol. 3, 2003, pp. 1721–1723.
- [98] R. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, *Image change detection algorithms : A systematic survey*, IEEE Transactions on Image Processing **14** (2005), no. 3, 294–307.
- [99] T. Riklin-Raviv, N. Kiryati, and N. Sochen, *Unlevel sets : Geometry and prior-based segmentation*, Proc. 8th European Conference on Computer Vision (ECCV'2004) (T. Pajdla and J. Matas, eds.), LNCS 3024, Springer, 2004, pp. 50–61.
- [100] M. Rochery, I.H. Jermyn, and J. Zerubia, *Higher order active contours and their application to the detection of line networks in satellite imagery*, Proc. 2nd IEEE Workshop on Variational, Geometric and Level Set Methods in Computer Vision, 2003.
- [101] M. Rochery, I.H. Jermyn, and J. Zerubia, *Étude d'une nouvelle classe de contours actifs pour la détection de routes dans des images de télédétection*, Proc. GRETSI Symposium on Signal and Image Processing, 2003.
- [102] ———, *Higher-order active contours*, Tech. Report 5656, INRIA, 2005, to be published in the International Journal of Computer Vision.
- [103] M. Rousson, *Cue integration and front evolution in image segmentation*, Ph.D. thesis, Université de Nice-Sophia Antipolis, December 2004.
- [104] M. Rousson and R. Deriche, *A variational framework for active and adaptative segmentation of vector valued images*, Tech. report, INRIA, 2002.
- [105] M. Rousson and N. Paragios, *Shape priors for level set representations*, ECCV '02 : Proceedings of the 7th European Conference on Computer Vision-Part II, LNCS 2351, Springer-Verlag, 2002, pp. 78–92.
- [106] M. Roux, *Recalage d'images multi-sources. Application au recalage d'une image SPOT et d'une carte*, Ph.D. thesis, ENST, 1992.
- [107] M. Roux and H. Maître, *Three-dimensionnal description of dense urban areas using maps and aerial images*, Proc. Workshop on "Automatic extraction of man-made objects from aerial and space images", Birkhauser Verlag Basel, 1997.
- [108] M. Roux and D. McKeown, *Feature matching for building extraction from multiple views*, DARPA Image Understanding Workshop, 1994, pp. 331–349.
- [109] G. Sapiro, *Geometric partial differential equations and image analysis*, Cambridge University Press, 2001.
- [110] S. Servigne, *Base de données géographiques et images aériennes : de l'appariement à la mise à jour*, Ph.D. thesis, INSA Lyon, 1993.
- [111] J.A. Sethian, *Level set methods and fast marching methods : evolving interfaces in computational geometry, fluid mechanics, computer vision and materials science*, Cambridge University Press, 1999.

- [112] G. Shafer, *A mathematical theory of evidence*, Princeton University Press, New Jersey, USA, 1976.
- [113] K. Siddiqi, B. B. Kimia, and C-W. Shu, *Geometric shock-capturing ENO schemes for subpixel interpolation, computation and curve evolution*, Graph. Models Image Process. **59** (1997), no. 5, 278–301.
- [114] L. H. Staib and J. S. Duncan, *Boundary finding with parametrically deformable models*, IEEE Transactions on Pattern Analysis and Machine Intelligence **14** (1992), no. 11, 1061–1075.
- [115] C. Steger, W. Eckstein, and C. Wiedemann, *Update of roads in GIS by automatic extraction from aerial imagery*, Proceedings of the Second International Airborne Remote Sensing Conference and Exhibition, vol. 3, Environmental Research Institute of Michigan, 1996, pp. 308–317.
- [116] M. Sussman, E. Fatemi, P. Smereka, and S. Osher, *An improved level set method for incompressible two-phase flows*, Journal of Computers and Fluids **27** (1998), no. 5-6, 663–680.
- [117] A. Tsai, A. Yezzi, and A.S. Willsky, *Curve evolution implementation of the Mumford-Shah functional for image segmentation, denoising, interpolation, and magnification*, IEEE Transactions on Image Processing **10** (2001), no. 8, 1169–1186.
- [118] M. S. Ulstad, *An algorithm for estimating small scale differences between two digital images*, Pattern Recognit. **5** (1973), 323–333.
- [119] S. Vinson, L.D. Cohen, and F. Perlant, *Extraction of rectangular buildings using DEM and orthoimage*, Proc. of Scandinavian Conference on Image Analysis (SCIA'01), 2001.
- [120] G. Vosselman and M.E. de Gunst, *Updating road maps by contextual reasoning.*, Proc. Workshop on "Automatic Extraction of Man-Made Objects from Aerial and Space Images" (A. Gruen, E.P. Baltsavias, and O. Henricsson, eds.), Birkhäuser Verlag, 1997, pp. 267–276.
- [121] L. Wald, *Definitions and terms of reference in data fusion*, IAPRS (Valladolid, Spain), vol. 32, June 1999.
- [122] V. Walter, *Automatic classification of remote sensing data for GIS database revision*, IAPRS, vol. 32, 1998, pp. 641–648.
- [123] ———, *Automatic change detection in GIS databases based on classification of multispectral data*, Proc. IAPRS, vol. 23, 2000, pp. 1138–1145.
- [124] V. Walter and D. Fritsch, *Automatic verification of GIS data using high resolution multispectral data*, IAPRS, vol. 32, 1998, pp. 485–490.
- [125] Z. Wang, D. Ziou, C. Armenakis, D. Li, and Q. Li, *A comparative analysis of image fusion methods*, IEEE Transactions on Geoscience and Remote Sensing **43** (2005), no. 6, 1391–1402.
- [126] D.J. Williams and M. Shah, *A fast algorithm for active contours and curvature estimation*, Image Understanding **55** (1992), no. 1, 14–26.
- [127] B. Xie, V. Ramesh, and T. Boult, *Sudden illumination change detection using order consistency*, Image Vis. Comput. **22** (2004), no. 2, 117–125.
- [128] C. Xu and J. Prince, *Gradient vector flow : A new external force for snakes*, Proc.Conf. on Computer Vision and Pattern Recognition (CVPR '97), 1997, pp. 66–71.
- [129] G. Xu and Z. Zhang, *Epipolar geometry in stereo, motion and object recognition - a unified approach*, Computational Imaging and Vision, vol. 6, Kluwer Academic Publishers, 1996.
- [130] Z. Xue, S. Li, and E.K. Teoh, *Bayesian shape model for facial feature extraction and recognition*, Pattern Recognition **36** (2003), no. 12, 2819–2833.
- [131] Zhong Xue, Stan Z. Li, and E.K. Teoh, *AI-EigenSnake : an affine-invariant deformable contour model for object matching*, Image and Vision Computing **20** (2002), no. 2, 77–84.

- [132] S. Yui, K. Hara, H. Zha, and T. Hasegawa, *A fast narrow band method and its application in topology-adaptative 3D modeling*, 16th Intl Conf. on Pattern Recognition (ICPR'02), vol. 4, 2002.
- [133] C. Zhang, E. Baltsavias, and A. Gruen, *Knowledge-based image analysis for 3D road reconstruction*, Asian Journal of Geoinformatics **1** (2001), no. 4, 3–14.
- [134] ———, *Updating of cartographic road databases by image analysis*, Proc. workshop on "Automatic Extraction of Man-Made Objects from Aerial and Space Images " (E. P. Baltsavias, A. Gruen, and L. Van Gool, eds.), vol. 3, A.A. Balkema, 2001.
- [135] H.K. Zhao, T. Chan, B. Merriman, and S. Osher, *A variational level set approach to multiphase motion*, Journal of Computational Physics **127** (1996), no. 167, 179–195.
- [136] L. Zhu, T. Sasagawa, K. Tachibana, and Y. Fukuzawa, *Triplet image matching for airborne digital sensor ADS40*, Proc. ISPRS XXth ISPRS Congress, 2004.
- [137] S.C. Zhu and A.L. Yuille, *Region competition : Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation*, IEEE Transactions on Pattern Analysis and Machine Intelligence **18** (1996), no. 9, 884–900.

Liste des publications

1. T. Bailloeul, J. Duan, V. Prinet and B. Serra, *Urban digital map updating from satellite high resolution images using GIS data as a priori knowledge*, in Proc. 2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion on Urban Areas, URBAN 2003, Berlin, Germany, May 2003.
2. T. Bailloeul, V. Prinet, B. Serra, P. Marthon, P. Chen and H. Zhang, *Digital building map refinement from knowledge-driven active contours and very high resolution optical imagery*, in Proc. ISPRS Hannover Workshop on "High-Resolution Earth Imaging for Geospatial Information", Hannover, Germany, May 2005.
3. T. Bailloeul, V. Prinet, B. Serra, P. Marthon, P. Chen and H. Zhang, *Urban building land use change mapping from high resolution satellite imagery, active contours and Hough voting*, in Proc. ISPMSRS, Beijing, China, October 2005.
4. T. Bailloeul, V. Prinet, B. Serra and P. Marthon, *Spatio-temporal prior shape constraint for level set segmentation*, in Proc. EMMCVPR, St Augustine, USA, LNCS 3757 Springer-Verlag Berlin Heidelberg, pp. 503-519, November 2005.
5. T. Bailloeul, V. Prinet, B. Serra, P. Marthon, P. Chen and H. Zhang, *Digital building map refinement from knowledge-driven active contours and very high resolution optical imagery*, in German Journal of Photogrammetry, Remote Sensing and GeoInformation - PFG Photogrammetrie-Fernerkundung-Geoinformation , pp. 511-522, November 2005.

Table des figures

2.1	Erreurs de délinéation des contours : différentes segmentations (en noir) du même objet (vérité terrain en blanc) par plusieurs photo-interpréteurs. Le schéma est extrait de [80].	35
2.2	Effet de généralisation dû à l'échelle. La même scène est représentée à différentes échelles : 1 :25,000 / 1 :50,000 / 1 :100,000 (de gauche à droite). Le schéma est extrait de [80].	35
2.3	Superposition d'une carte du bâti (polygones jaunes) avec une image satellitaire. Les flèches blanches pointent des inexactitudes locales de forme de la carte.	37
2.4	Superposition d'une carte du bâti (polygones jaunes) avec une image satellitaire illustrant l'effet de généralisation : un seul polygone cartographique englobe de nombreux bâtiments contigus.	37
3.1	Exemple d'un couple d'images stéréoscopiques. Les images sont des données satellitaires simulées à partir d'images aériennes.	42
3.2	Pour la restitution stéréoscopique, deux images sont acquises de deux points de vue différents. Un point P de la scène se projette sur chaque image à l'intersection du plan image et de la droite joignant le point P et le centre optique C_A ou C_B	43
3.3	Pour deux points conjugués P_A et P_B , l'orientation relative permet de calculer la position du point P dans le système de coordonnées lié au couple d'images. P est défini comme le centre du plus petit segment reliant les droites (P_A, C_A) et (P_B, C_B)	44
3.4	Exemple de surface de corrélation.	46
3.5	Paire d'images en géométrie quelconque	47
3.6	Paire d'images en géométrie épipolaire	47
3.7	Masque adaptatif de corrélation.	48
3.8	Erreur dans le cas de zones occultées avec une fenêtre de corrélation fixe. (a) un pixel de disparité nulle est considéré. (b) son homologue dans l'image conjuguée n'est pas retrouvé à cause de la prise en compte d'une partie du bâtiment dans la fenêtre de corrélation.	49
3.9	Application du masque sur la vignette de contexte. La partie du bâtiment est retirée de la fenêtre de corrélation grâce à la géométrie adaptative du masque. Le pixel homologue est retrouvé.	49
3.10	Image de disparité générée à partir du couple de la figure 3.1 : (a) corrélation classique (b) corrélation par masques adaptatifs.	50
3.11	Synthèse du processus de reconstruction 3D.	51
3.12	Définition des angles d'attitude (roulis, tangage, lacet) d'une caméra embarquée sur une plate-forme aérienne (en supposant que les axes de rotation de l'avion et de la caméra sont identiques). Le schéma est extrait de [11].	52

3.13	Illustration des angles $(\varphi, \omega, \kappa)$ utilisés en photogrammétrie. Ces angles expriment la rotation entre le référentiel caméra (x^B, y^B, z^B) et le référentiel terrain (x^E, y^E, z^E) associé à une projection cartographique. Le schéma est extrait de [11].	52
3.14	Huit marques fiduciaires représentées dans le plan focal de la caméra. Le repère du plan focal est centré sur le point principal de symétrie (PPS) du système optique. . .	53
3.15	Synthèse algorithmique de l'optimisation globale des paramètres de prise de vue de n images aériennes.	58
3.16	Couple d'images stéréoscopiques rectifiées en géométrie épipolaire.	60
3.17	MNS en géométrie épipolaire. Les zones non renseignées sont en noir ; la luminance est proportionnelle à l'altitude.	60
3.18	Grille orthoscopique associée à la projection cartographique.	63
3.19	Projection d'un triangle du MNS épipolaire vers la grille orthoscopique.	64
3.20	Multiples triangles projetés dans la grille orthoscopique et intersectant le même pixel. . .	64
3.21	MNS orthorectifié avec un pas $R = 0.65$ m et $\varepsilon_{corr} = 0.4$	65
3.22	MNS orthorectifiés avec différents seuils de corrélation.	66
3.23	MNS ortho réalisé avec une image de disparité filtrée avec un filtre médian.	67
3.24	MNS ortho réalisé avec une image de disparité filtrée avec un filtre de réjection de taille 7×7	68
3.25	MNS fusionné $MNS_1 + MNS_2$	70
3.26	Agrandissement de la zone rectangulaire rouge de la figure 3.25.	71
4.1	Sous-échantillonnage d'un contour représenté de façon explicite et paramétré par son abscisse curviligne.	75
4.2	Représentation à trois dimensions de l'intersection d'un ensemble de niveaux avec le plan de l'image (niveau zéro). Le contour représenté implicitement est celui de la figure 4.1. La légende représente la distance Euclidienne signée au contour.	77
4.3	Exemple d'extraction de segments (a) et du gradient vector flow associé (b).	95
4.4	Voisinages circulaires utilisés pour la construction de la bande étroite rapide. La légende fournit une approximation de la distance pixellaire au centre.	98
4.5	Approximation linéaire d'un contour.	99
4.6	Diagramme de l'algorithme d'évolution du contour actif	100
5.1	Représentations différentes de la même scène : image satellitaire et reconstruction 3D. . .	106
5.2	Variation spatio-temporelle de la fonction pondérant la contrainte de forme. $d_0 = 3$, $t_1 = 100$, $t_2 = 400$, $\lambda_{min} = 1.25$ et $\lambda_{max} = 2$	110
5.3	Formes <i>a priori</i> (en blanc, $\{\mathbf{x}/\psi_0(\mathbf{x}) = 0\}$) et leurs ensembles de niveaux associés (niveaux de gris, $ \psi_0(\mathbf{x}) $). (a) jouet (b) bâtiment en L	111
5.4	Segmentation sans insertion de forme <i>a priori</i> et $\xi_0 = (1, 0, 0, 0)$. (a) modèle de Chan et Vese (b) modèle Bayésien.	114
5.5	Segmentation avec insertion de forme <i>a priori</i> : (a) modèle de Chan et Vese et $\lambda = 1$ (b) modèle Bayésien et $\lambda = 10$	114
5.6	Segmentation avec un contour actif basé sur l'information de contour (GVF) et $\xi_0 = (s_0 = 1.1, \theta_0 = 0.1, \mu_{x,0} = -4, \mu_{y,0} = 1)$: (a) résultat sans contrainte de forme <i>a priori</i> . (b) résultat avec contrainte de forme <i>a priori</i> ($\lambda = 5$).	115
5.7	Sensibilité à la translation : segmentation avec le modèle de Chan et Vese et la contrainte de forme. Le contour actif initial des expériences (a), (b) et (c) a été transformé par une similitude plane directe de paramètres $\xi_0 = (s_0 = 1.0, \theta_0 = 0.0, \mu_{x,0} = 5, \mu_{y,0} = 5)$	117
5.8	Sensibilité à la rotation : segmentation avec le modèle de Chan et Vese et la contrainte de forme. Le contour actif initial des expériences (a), (b) et (c) a été transformé par une similitude plane directe de paramètres $\xi_0 = (s_0 = 1.0, \theta_0 = 0.3, \mu_{x,0} = 0, \mu_{y,0} = 0)$	118

5.9	Sensibilité au facteur d'échelle : segmentation avec le modèle de Chan et Vese et la contrainte de forme. (a) $\xi_0 = (s_0 = 1.5, \theta_0 = 0.0, \mu_{x,0} = 0, \mu_{y,0} = 0)$; (b) $\xi_0 = (s_0 = 0.8, \theta_0 = 0.0, \mu_{x,0} = 0, \mu_{y,0} = 0)$; (c) $\xi_0 = (s_0 = 1.5, \theta_0 = 0.0, \mu_{x,0} = 0, \mu_{y,0} = 0)$.	119
5.10	Sensibilité à la translation : segmentation avec le modèle Bayésien et la contrainte de forme. Le contour actif initial des expériences (a), (b) et (c) a été transformé par une similitude plane directe de paramètres $\xi_0 = (s_0 = 1.0, \theta_0 = 0.0, \mu_{x,0} = 5, \mu_{y,0} = 5)$.	120
5.11	Sensibilité à la rotation : segmentation avec le modèle Bayésien et la contrainte de forme. Le contour actif initial des expériences (a), (b) et (c) a été transformé par une similitude plane directe de paramètres $\xi_0 = (s_0 = 1.0, \theta_0 = 0.3, \mu_{x,0} = 5, \mu_{y,0} = 5)$.	121
5.12	Sensibilité au facteur d'échelle : segmentation avec le modèle Bayésien et la contrainte de forme. (a) $\xi_0 = (s_0 = 1.5, \theta_0 = 0.0, \mu_{x,0} = 0, \mu_{y,0} = 0)$ (b) $\xi_0 = (s_0 = 0.8, \theta_0 = 0.0, \mu_{x,0} = 0, \mu_{y,0} = 0)$ (c) $\xi_0 = (s_0 = 1.5, \theta_0 = 0.0, \mu_{x,0} = 0, \mu_{y,0} = 0)$.	122
5.13	Sensibilité à l'initialisation : segmentation avec un contour actif basé sur l'information de contour (GVF) et avec contrainte de forme. (a) $\xi_0 = (s_0 = 1.1, \theta_0 = 0.1, \mu_{x,0} = -4, \mu_{y,0} = 1)$ (b) $\xi_0 = (s_0 = 1.0, \theta_0 = -0.1, \mu_{x,0} = -1, \mu_{y,0} = -2)$ (c) $\xi_0 = (s_0 = 1.0, \theta_0 = 0.1, \mu_{x,0} = 0, \mu_{y,0} = 2)$.	124
5.14	Sensibilité à une initialisation éloignée : segmentation avec un contour actif basé sur l'information de contour (GVF) et avec contrainte de forme. (a) $\xi_0 = (s_0 = 1.1, \theta_0 = 0.25, \mu_{x,0} = -4, \mu_{y,0} = 1)$ (b) $\xi_0 = (s_0 = 1.0, \theta_0 = -0.25, \mu_{x,0} = -2, \mu_{y,0} = -3)$ (c) $\xi_0 = (s_0 = 1.0, \theta_0 = 0.1, \mu_{x,0} = -2, \mu_{y,0} = 2)$.	125
5.15	Recalage réussi avec une initialisation éloignée et l'insertion du MNS exogène dans la fonctionnelle d'énergie. (a) modèle de Chan et Vese (b) modèle Bayésien.	126
5.16	Échec du recalage avec une initialisation éloignée et le modèle de Chan et Vese (pas d'incorporation du MNS).	126
5.17	Échec du recalage sans insertion du MNS, avec une initialisation éloignée et le modèle Bayésien.	127
5.18	Recalage réussi avec une initialisation éloignée et l'insertion du MNS exogène dans la fonctionnelle d'énergie (modèle Bayésien pour l'image, modèle de Chan et Vese pour le MNS).	127
5.19	Recalage réussi avec une initialisation éloignée et l'insertion du MNS exogène dans la fonctionnelle d'énergie (modèle GVF pour l'image, modèle de Chan et Vese pour le MNS).	128
5.20	Échec du recalage sans insertion du MNS, avec une initialisation éloignée et le modèle GVF.	128
5.21	Segmentation avec une variation spatio-temporelle de la contrainte de forme et le modèle Bayésien ($\lambda_{min} = 3, \lambda_{max} = 30, d_0 = 7$).	129
5.22	Segmentation avec une variation spatio-temporelle de la contrainte de forme et le modèle Bayésien ($\lambda_{min} = 27, \lambda_{max} = 70, d_0 = 7$).	130
5.23	Segmentation avec un poids constant de contrainte de forme et le modèle Bayésien ($\lambda = 50$).	130
5.24	Variation spatio-temporelle de lambda avec le modèle GVF et $\lambda_{min} = 1, \lambda_{max} = 5, d_0 = 5$: (a) $\xi_0 = (s_0 = 1.1, \theta_0 = 0.25, \mu_{x,0} = -4, \mu_{y,0} = 1)$ (b) $\xi_0 = (s_0 = 1.0, \theta_0 = -0.25, \mu_{x,0} = -2, \mu_{y,0} = -3)$ (c) $\xi_0 = (s_0 = 1.0, \theta_0 = 0.1, \mu_{x,0} = -2, \mu_{y,0} = 2)$.	131

5.25	Segmentation avec variation spatio-temporelle du poids de contrainte de forme (modèle Bayésien) : (a) $\lambda_{min} = 15$, $\lambda_{max} = 100$, $d_0 = 15$ (b) $\lambda_{min} = 50$, $\lambda_{max} = 100$, $d_0 = 10$	132
5.26	Échec du recalage avec un poids constant de contrainte de forme (modèle Bayésien) : (a) $\lambda = 20$ (b) $\lambda = 70$	132
5.27	Échec du recalage (modèle Bayésien) avec : (a) unique variation temporelle de la contrainte de forme avec $\lambda_{min} = 30$, $\lambda_{max} = 70$ (pas de relaxation spatiale); (b) variation spatio-temporelle de la contrainte de forme avec un profil spatial défini en équation (5.11) et $\lambda_{min} = 27$, $\lambda_{max} = 70$, $d_0 = 7$; (c) variation spatio-temporelle de la contrainte de forme avec des fonctions échelon de Heaviside pour $\lambda_a(t)$ et $d(t)$. . .	134
5.28	Superposition de la vérité terrain (grisée) et d'un polygone cartographique localement erroné	135
5.29	Convergence finale du contour actif avec une forme <i>a priori</i> erronée	135
5.30	Recalage par contour actif basé sur l'information de gradient (GVF) et avec une forme <i>a priori</i> non erronée.	136
5.31	Recalage par contour actif basé sur l'information de gradient (GVF) et avec une forme <i>a priori</i> erronée dérivée de la carte.	136
6.1	Effet d'arrondissement au niveau des singularités géométriques dû à la nouvelle énergie de contrainte de forme basée sur l'alignement des normales.	140
6.2	Interaction à longue distance due à l'énergie correctrice quadratique. Les vecteurs tangents $\vec{t}(s)$ et $\vec{t}(s')$ auront tendance à être orthogonaux ou colinéaires.	141
6.3	Fonctions modifiées de Bessel de seconde espèce.	146
6.4	Schéma décrivant le mode d'échantillonnage avec une fonction de Bessel pour profil Ψ	147
6.5	Image de synthèse avec le contour initial représenté en rouge.	147
6.6	(a) Forme <i>a priori</i> et son ensemble de niveaux associé (b).	148
6.7	Résultat avec la contrainte de forme classique	148
6.8	Segmentation avec contrainte de forme autorisant les variations parallèles : (a) $\lambda_{para} = 5$; (b) $\lambda_{para} = 10$	149
6.9	Segmentation avec contrainte de forme autorisant les variations parallèles : (a) $\lambda_{para} = 15$; (b) $\lambda_{para} = 30$	149
6.10	Segmentation avec le modèle d'attache aux données de Chan et Vese et la contrainte de forme autorisant les variations parallèles : (a) sans correction quadratique et $\lambda_{para} = 20$ à l'instant n_0 ; (b) avec correction quadratique, $\lambda_{para} = 20$, $\lambda_{quad} = 8$ et $a = 0.5$	150
6.11	Segmentation avec le modèle d'attache aux données de Chan et Vese et la contrainte de forme autorisant les variations parallèles : (a) sans correction quadratique et $\lambda_{para} = 50$ à l'instant n_0 ; (b) avec correction quadratique, $\lambda_{para} = 50$, $\lambda_{quad} = 7$ et $a = 0.5$	151
6.12	Représentation de la force quadratique à l'instant n_0 . Les zones bleues correspondent à une force positive (effet de dilatation), celles en vert représentent une force négative (effet de compression).	151
6.13	Expérience avec une image réelle, la forme <i>a priori</i> est le résultat du recalage avec la contrainte de forme classique. La segmentation est effectuée avec le modèle d'attache aux données de Chan et Vese, la contrainte de forme autorisant les variations parallèles et avec correction quadratique $\lambda_{para} = 4$, $\lambda_{quad} = 0.5$ et $a = 0.5$	152

7.1	Méthodologie proposée pour la détection de changement : (A) calcul et fusion d'indices de changement. (B) atténuation des variabilités exogènes carte-image (C) analyse de changements finale. Légende : Qm et Qp sont respectivement l'images satellitaire multispectrale et panchromatique.	154
7.2	Superposition de la carte du bâti (en jaune) avec le MNS grâce à l'information de géocodage. La flèche rouge pointe un changement : bâtiment remplacé par un terre-plein.	157
7.3	Résultat de la détection de changement carte-MNS. Légende : rouge (détection de changement), jaune (ambigu), bleu (indécidable). Paramètres : $seuil_{rens} = 10 \%$, $seuil_{plat} = 1 \text{ m}$, $seuil_{bat} = 3 \text{ m}$	158
7.4	Succès de la méthode en fonction de $seuil_{rens}$. En rouge figure le taux de bâtiments correctement détectés comme ayant changé. La courbe bleue représente le taux de bâtiments détectés comme ayant changé par rapport au nombre total de bâtiments dans la carte. $seuil_{plat} = 1 \text{ m}$ et $seuil_{bat} = 3 \text{ m}$	159
7.5	Succès de la méthode en fonction de $seuil_{plat}$. En rouge figure le taux de bâtiments correctement détectés comme ayant changé. La courbe bleue représente le taux de bâtiments détectés comme ayant changé par rapport au nombre total de bâtiments dans la carte. $seuil_{rens} = 50\%$ et $seuil_{bat} = 3 \text{ m}$	159
7.6	Succès de la méthode en fonction de $seuil_{bat}$. En rouge figure le taux de bâtiments correctement détectés comme ayant changé. La courbe bleue représente le taux de bâtiments détectés comme ayant changé par rapport au nombre total de bâtiments dans la carte. $seuil_{rens} = 65\%$ et $seuil_{plat} = 1.5 \text{ m}$	160
7.7	Bande infra-rouge, couleurs réelles (RVB), et $NDVI$	162
7.8	Estimation de la translation maximale μ_1 et minimale μ_2 entre le segment cartographique [AB] et le segment-image [A'B'].	166
7.9	Accumulateur de Hough à 2 dimensions (μ_x, μ_y). L'accumulation entre la translation minimale et maximale est élargie d'une bande de l pixels.	166
7.10	Détection de changements univoques. Les changements ont été détectés par la comparaison avec l'image NDVI. Les figures illustrent les objets cartographiques superposés à une image satellitaire plus récente.	171
7.11	Détection de changement en raison d'une forte variation géométrique : $E_{geom} = 0.54$ et $p_{NC} = 0.30$. Le changement est simulé.	171
7.12	Détection de changement en raison d'une translation μ_{Hough} élevée. $s_{\text{Hough}} = 0.27$, $\mu_{\text{Hough}} = (-6, -26)$ et $p_{NC} = 0.00$. Le changement est simulé.	172
7.13	Détection de non-changement avec un objet cartographique parfait transformé par similitude de paramètres $\xi_0 = (s_0 = 1, \theta_0 = 0.1, \mu_{x,0} = 1, \mu_{y,0} = 1)$. Le non-changement est confirmé avec $s_{\text{Hough}} = 0.76$, $\mu_{\text{Hough}} = (1, 0)$, $E_{geom} = 0.1$ et $p_{NC} = 0.63$	174
7.14	Détection de non-changement avec un objet cartographique parfait transformé par similitude de paramètres $\xi_0 = (s_0 = 1.05, \theta_0 = 0.075, \mu_{x,0} = 4, \mu_{y,0} = -1)$. Le non-changement est confirmé avec $s_{\text{Hough}} = 0.72$, $\mu_{\text{Hough}} = (1, 0)$, $E_{geom} = 0.08$ et $p_{NC} = 0.63$	174
7.15	Détection de non-changement avec un objet cartographique parfait transformé par similitude de paramètres $\xi_0 = (s_0 = 1.0, \theta_0 = -0.02, \mu_{x,0} = -2, \mu_{y,0} = 1)$. Le non-changement est confirmé avec $s_{\text{Hough}} = 0.81$, $\mu_{\text{Hough}} = (0, 0)$, $E_{geom} = 0.04$ et $p_{NC} = 0.76$	175
7.16	Détection de non-changement avec un objet cartographique parfait transformé par similitude de paramètres $\xi_0 = (s_0 = 1.0, \theta_0 = 0.0, \mu_{x,0} = 5, \mu_{y,0} = 5)$. Le non-changement est confirmé avec $s_{\text{Hough}} = 0.84$, $\mu_{\text{Hough}} = (-1, 0)$, $E_{geom} = 0.01$ et $p_{NC} = 0.83$	175

7.17	Détection de non-changement avec une erreur de mise en correspondance modélisable par une transformation globale. Le non-changement est confirmé avec $s_{Hough} = 0.86$, $\mu_{Hough} = (0, -1)$, $E_{geom} = 0.12$ et $p_{NC} = 0.68$	176
7.18	Détection de non-changement avec une erreur de mise en correspondance modélisable par une transformation globale. Le non-changement est confirmé avec $s_{Hough} = 0.88$, $\mu_{Hough} = (0, 0)$, $E_{geom} = 0.03$ et $p_{NC} = 0.83$	177
7.19	Détection de non-changement avec des erreurs locales de forme. Le non-changement est confirmé avec $s_{Hough} = 0.62$, $\mu_{Hough} = (1, 0)$, $E_{geom} = 0.02$ et $p_{NC} = 0.66$	177
7.20	Détection de non-changement avec des erreurs locales de forme. Le non-changement est confirmé avec $s_{Hough} = 0.73$, $\mu_{Hough} = (1, 0)$, $E_{geom} = 0.03$ et $p_{NC} = 0.71$	178
7.21	Détection de non-changement avec des erreurs locales de forme. Le non-changement est confirmé avec $s_{Hough} = 0.75$, $\mu_{Hough} = (1, 0)$, $E_{geom} = 0.05$ et $p_{NC} = 0.71$	178
7.22	Détection ambiguë (cas indéterminé). $s_{Hough} = 0.45$, $\mu_{Hough} = (0, 0)$, $E_{geom} = 0.02$ et $p_{NC} = 0.55$	179
7.23	Détection de changement erronée due aux erreurs locales de forme de la carte. $s_{Hough} = 0.47$, $\mu_{Hough} = (-1, -1)$, $E_{geom} = 0.19$ et $p_{NC} = 0.39$	180
7.24	Illustration des erreurs de délinéation et de l'effet de généralisation dans la carte de 1996. L'objet cartographique est superposé sur l'image satellitaire de 2002.	181
A.1	Image aérienne représentée à 5% de sa taille originale. Le futur village olympique se situe au nord-est de l'image.	192
A.2	Extrait des données SIG d'échelle 1 :10,000 de la ville de Pékin. Le bâti est en rose, le réseau routier en violet, les rivières et canaux en bleu, les lacs en vert.	194
B.1	simplex initial et réflexion	197
B.2	Expansion	198
B.3	Contractions possibles	198
B.4	Rétrécissement	198
C.1	Schéma illustrant la modélisation inverse. Les repères terrain et caméra sont directs.	200
C.2	Schéma illustrant la modélisation directe. Les repères terrain et caméra sont directs.	202
F.1	Visualisation 3D de la ville de Pékin - avec textures et sans MNT.	212

Liste des tableaux

2.1	Classification des artefacts urbains en fonction des différents univers.	36
2.2	Changements possibles d'un bâtiment dans $U_{objectif}$	36
3.1	Résultats en pixels de la quantification de la précision absolue des paramètres de prise de vue pour les trois images aériennes.	54
3.2	Résultats en pixels de la quantification de la précision relative des paramètres de prise de vue pour chacune des paires d'images aériennes.	55
3.3	Résultats de l'optimisation des paramètres de prise de vue avec les images à 0.21 m/pixel de résolution.	59
3.4	Résultats de l'optimisation des paramètres de prise de vue avec les images sous-échantillonnées par un facteur trois.	59
3.5	Erreur absolue entre les GCPs et la reconstruction 3D du MNS_1 . ΔN : erreur dans la direction nord, ΔE : erreur dans la direction est, ΔZ erreur altimétrique, erreur dans le plan : $\Delta Plan = \sqrt{\Delta N^2 + \Delta E^2}$	62
3.6	Erreur absolue entre les GCPs et la reconstruction 3D du MNS_2 . ΔN : erreur dans la direction nord, ΔE : erreur dans la direction est, ΔZ erreur altimétrique, erreur dans le plan : $\Delta Plan = \sqrt{\Delta N^2 + \Delta E^2}$	62
5.1	Erreur absolue entre ξ_0 et son estimée $\hat{\xi}_0$ par l'algorithme de descente de gradient et temps de calcul.	112
5.2	Erreur absolue entre ξ_0 et son estimée $\hat{\xi}_0$ par l'algorithme du simplex et temps de calcul.	113
5.3	Rapport entre les temps de convergence sans et avec fusion de MNS : $ratio = \frac{t_{sans\ fusion}}{t_{avec\ fusion}}$	129
5.4	Rapport entre les temps de convergence sans et avec contrainte de forme spatio-temporelle : $ratio = \frac{t_{contrainte\ uniforme}}{t_{contrainte\ spatio-temporelle}}$	134
7.1	Quantification des scores de changement/non-changement. p_{NC} est la probabilité de non-changement. s_{Hough} est le score de Hough normalisé.	180
7.2	Quantification des résultats de la détection de changement avec la carte de 2002. Les résultats sont exprimés en pour-cent sur un total de 77 cas.	182
7.3	Quantification des résultats de la détection de changement avec la carte de 1996. Les résultats sont exprimés en pour-cent sur un total de 105 cas.	183

CONTOURS ACTIFS ET INFORMATION A PRIORI POUR L'ANALYSE DE CHANGEMENTS : APPLICATION À LA MISE À JOUR DE CARTES NUMÉRIQUES DU BÂTI URBAIN À PARTIR D'IMAGES OPTIQUES DE TÉLÉ-DÉTECTION HAUTE RÉOLUTION

Résumé :

Cette thèse a pour objectif de proposer une méthodologie visant à analyser les changements entre une carte numérique de bâtiments et des données de télédétection optiques haute résolution plus récentes. L'idée centrale de notre approche est d'utiliser l'information géométrique et spécifique des bâtiments symbolisés dans la carte afin de faciliter leur reconnaissance dans une image satellitaire panchromatique et de s'affranchir des difficultés du milieu urbain. Cette connaissance *a priori* est incorporée comme une contrainte de forme dans un modèle variationnel de recalage fin carte-image par contours actifs. Nous proposons des solutions nouvelles pour l'amélioration de la robustesse des contours actifs contraints : la fusion de l'image avec un Modèle Numérique de Surface (MNS) et la variation spatio-temporelle de la contrainte de forme permettent de surmonter le problème des minima locaux. Une nouvelle énergie de contrainte de forme, linéaire et quadratique, permet aux contours actifs de s'affranchir localement et parallèlement de la forme *a priori* de la carte pour mieux segmenter l'objet dans l'image. L'alignement des données par contours actifs résout le problème des variabilités exogènes carte-image indépendantes d'un changement effectif et rend ainsi l'analyse de changements subséquente plus fiable. Des indices de changement simplement calculés à partir du MNS et d'une image satellitaire multispectrale sont fusionnés pour la détection de changements univoques du bâti urbain. En l'absence de changement univoque, une méthode basée sur le vote de Hough quantifie la ressemblance entre le polygone cartographique recalé et des primitives segments extraites de l'image panchromatique. La fusion du score de Hough avec la variation géométrique enregistrée par le recalage des contours actifs permet la formulation d'une probabilité de non-changement visant à assister un opérateur dans sa tâche de révision de carte.

Mots-clés : *changement, contour actif, connaissance a priori, optimisation, fusion, image de télédétection, haute résolution, Système d'Information Géographique, cartographie.*

ACTIVE CONTOURS AND PRIOR KNOWLEDGE FOR CHANGE ANALYSIS : APPLICATION TO DIGITAL URBAN BUILDING MAP UPDATING FROM OPTICAL HIGH RESOLUTION REMOTE SENSING IMAGES

Abstract :

This thesis aims at proposing a methodology to analyze changes between a digital map of buildings and more recent high resolution optical remote sensing data. The key idea of the proposed approach is to use the specific geometrical information derived from the buildings symbolized in the map in order to ease their recognition in a panchromatic satellite image while overcoming urban difficulties. This prior knowledge is embedded in active contours models as a shape constraint formulated in a variational approach. We propose new solutions to improve the robustness of shape constrained active contours : the fusion of the image with a digital surface model (DSM) and the spatio-temporal variation of the shape constraint allow to overcome the critical issue of local minima. A new shape constraint energy - linear and quadratic - allows local and parallel discrepancies of the active contours from the cartographic shape prior to better segment the object in the image. The fine map-to-image matching achieved by active contours solve the issue of exogenous variabilities between the map and the image and therefore makes the subsequent change detection decision more reliable. Change indicators simply calculated from the DSM and a multispectral satellite image are merged to detect obvious changes of buildings. When no obvious change is detected, a method based on Hough voting quantifies the similarity between the refined cartographic object and segment primitives extracted from the panchromatic image. The fusion of the Hough voting score with the geometric variation due to the active contours matching enables to formulate a "non-change" probability aiming at assisting a photo-interpreter performing map updating.

Key words : *change, active contour, prior knowledge, optimization, fusion, remote sensing image, high resolution, Geographic Information System, cartography.*

Discipline : Traitement des Images - Reconnaissance des Formes - Vision par Ordinateur.

Laboratoires d'accueil :

LIMA, site ENSEEIHT de l'IRIT-UMR CNRS 5505. 2, rue Charles Camichel BP 7122 - 31071 TOULOUSE CEDEX 7, France.

LIAMA, Institut d'Automatique, Académie des Sciences de Chine. ZhongGuanCun DongLu, PoBox 2728, Beijing, Chine.
