



Segmentation d'images de façades de bâtiments acquises d'un point de vue terrestre

Jean-Pascal Burochin

► To cite this version:

Jean-Pascal Burochin. Segmentation d'images de façades de bâtiments acquises d'un point de vue terrestre. Autre. Université Paris-Est, 2012. Français. NNT : 2012PEST1064 . tel-00786332

HAL Id: tel-00786332

<https://theses.hal.science/tel-00786332>

Submitted on 8 Feb 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

pour obtenir le titre de
Docteur d'Université Paris-Est

Spécialité : Traitement du Signal et des Images

Présentée et soutenue publiquement le 2 mai 2012 par

Jean-Pascal BUROCHIN

Segmentation d'images de façades de bâtiments
acquises d'un point de vue terrestre

Composition du jury

Florence TUPIN	TELECOM ParisTech	Président
Matthieu CORD	LIP6	Rapporteur
Peter STURM	INRIA Alpes, Grenoble	Rapporteur
Christian HEIPKE	IPG, Hanovre	Examineur
Caroline BAILLARD	Technicolor, Cesson-Sévigné	Examineur
Olivier TOURNAIRE	CSTB, Marne-la-Vallée	Encadrant
Nicolas PAPARODITIS	IGN, Saint Mandé	Directeur

Cette thèse s'est déroulée à l'Institut géographique national (*IGN*), *laboratoire MATIS, 73 avenue de Paris, 94165 Saint Mandé cedex, France.*

Fiche bibliographique BIB_{TEX}~ :

```
@PHDTHESIS{Burochin2012_thesis,
  author = {Burochin, Jean-Pascal},
  title = {Segmentation g\ '{e}n\ '{e}rique d'images de façades
    de b\ ^{a}timents acquises d'un point de vue terrestre},
  school = {Universit\ '{e} {P}aris-{E}st},
  year = {2012},
  address = {Institut g\ '{e}ographique national ({IGN}),
    laboratoire {MATIS},
    73 avenue de Paris, 94165 {S}aint {M}and\ '{e} cedex, {F}rance},
  month = {feb},
}
```


Remerciements

Les travaux de cette thèse m'ont amené à rencontrer différentes personnes liées de près ou de loin à ma problématique scientifique. Leurs questions, leurs réponses, leurs conseils, leurs encouragements ou leur simple sourire aux moments opportuns, m'ont apporté un soutien non négligeable. La présente page cite un certain nombre de ces personnes. Que les autres soient également assurées de toute ma reconnaissance !

Nicolas Paparoditis, mon directeur de thèse, est l'acteur principal de ma progression scientifique ces dernières années. J'ai particulièrement apprécié son enthousiasme, sa patience et sa parole percutante quotidienne pour me faire avancer : soit pour m'encourager, soit pour me recadrer et toujours avec une capacité d'écoute remarquable pour discuter de concepts que j'étudiais ou de tout autre chose.

Olivier Tournaire s'est rendu disponible durant la majeure partie de ma thèse pour m'encadrer techniquement. Sans sens pratique et ses conseils informatiques m'ont grandement profité. Je le remercie pour la manière pragmatique avec laquelle il m'a expliqué la résolution d'un certain nombre de problèmes que j'ai rencontré.

Je remercie les autres membres du jury : le Pr. Matthieu Cord et le Dr. Peter Sturm pour le temps qu'ils ont consacré à lire et synthétiser ce mémoire et pour leurs remarques pertinentes ainsi que le Dr. Florence Tupin, le Dr. Caroline Baillard et le Pr. Chritian Heipke pour leur implication.

Bruno Vallet m'a également beaucoup apporté scientifiquement. J'ai apprécié sa prise de recul et sa capacité de formalisation des problèmes que je lui soumettais. Qu'il soit assuré de toute ma reconnaissance pour sa disponibilité.

Un merci tout particulier aussi à Mathieu Brédif qui m'a expliqué patiemment et très clairement les rudiments de l'échantillonnage stochastique. J'ai eu beaucoup de plaisir à travailler avec lui pour la rigueur de son enseignement et sa simplicité.

Je remercie l'ensemble du laboratoire MATIS pour le cadre de réflexion productive et très agréable qu'il m'a offert. Chaque chercheur m'a partagé son expertise et les sportifs se sont constamment relayés pour m'entraîner m'aérer la tête.

Merci également à l'ensemble du personnel du service de la recherche de l'IGN ainsi qu'à l'administration si sympathique du PRES Paris-Est.

Il me reste à remercier chaleureusement mon épouse et mon père pour leurs patientes relectures ainsi que Barbara pour mon anglais. Je remercie ces mêmes personnes ainsi que ma mère, ma famille et mes amis pour leurs encouragements.

Résumé

L'analyse de façades (détection, compréhension et reconstruction) à partir d'images acquises depuis la rue est aujourd'hui un thème de recherche très actif en photogrammétrie et en vision par ordinateur de part ses nombreuses applications industrielles. Cette thèse montre des avancées réalisées dans le domaine de la segmentation générique de grands volumes de ce type d'images, contenant une ou plusieurs zones de façades (entières ou tronquées).

Ce type de données se caractérise par une complexité architecturale très riche ainsi que par des problèmes liés à l'éclairage et au point de vue d'acquisition. La généralité des traitements est un enjeu important. La contrainte principale est de n'introduire que le minimum d'*a priori* possible.

Nous basons nos approches sur les propriétés d'alignements et de répétitivité des structures principales de la façade. Nous proposons un partitionnement hiérarchique des contours de l'image ainsi qu'une détection de grilles de structures répétitives par *processus ponctuels marqués*. Sur les résultats, la façade est séparée de ses voisines et de son environnement (rue, ciel). D'autre part, certains éléments comme les fenêtres, les balcons ou le fond de mur, sans être reconnus, sont extraits de manière cohérente. Le paramétrage s'effectue en une seule passe et s'applique à tous les styles d'architecture rencontrés.

La problématique se situe en amont de nombreuses thématiques comme la séparation de façades, l'accroissement du niveau de détail de modèles urbains 3D générés à partir de photos aériennes ou satellitaires, la compression ou encore l'indexation à partir de primitives géométriques (regroupement de structures et espacements entre elles).

Mots-clés images de façades, point de vue terrestre, segmentation en modèles élémentaires 2D, densité de contours, structures répétitives, processus ponctuel marqué, RJMCMC.

Abstract

Ground-Based Building Facade Image Segmentation

Facade analysis (detection, understanding and field of reconstruction) in street level imagery is currently a very active field of research in photogrammetric computer vision due to its numerous applications. This thesis shows some progress made in the field of generic segmentation of a broad range of images that contain one or more facade areas (as a whole or in part).

This kind of data is characterized by different styles of more or less complexity from architectural point of view, non homogeneous lighting conditions and uncontrolled camera pose. Workflow genericity is an important issue. One significant constraint is to be as little biased as possible.

The presented approaches extract the main facade structures based on geometric properties such as alignment and repetitivity. We propose a hierarchic partition of the image contour edges and a detection of repetitive grid patterns based on *marked point processes*. The facade is set apart from its neighboring façades and from its environment (the ground, the sky). Some elements such as windows, balconies or wall backgrounds, are extracted in a relevant way, without being recognized. The parameters regulation is done in one step and refers to all architectural styles encountered.

The problem originates from most themes such as facade separation, the increase of level of details in 3D city models generated from aerial or satellite imagery, compression or indexation based on geometric primitives (structure grouping and space between them).

Keywords facade segmentation, street level imagery, elementary 2D models, contour density, repetitive structures, marked Point Processes, RJMCMC, pattern-based image segmentation

Sommaire

1	Introduction	13
2	État de l’art	27
3	Segmentation hiérarchique	43
4	Grilles de structures répétitives	81
5	Optimisation de l’agencement de grilles	99
6	Conclusion	121
A	Calculs combinatoires	127
	Publications de l’auteur	130
	Bibliographie	130
	Table des figures	137

Introduction

L'analyse de façades (détection, compréhension et reconstruction) à partir d'images acquises depuis la rue est aujourd'hui un thème de recherche très actif en photogrammétrie et en vision par ordinateur de part ses nombreuses applications industrielles.

Le présent chapitre expose le contexte de la thèse en citant différentes applications possibles (section 1.1). Une description synthétique de l'architecture générale des façades étudiées est ensuite présentée (section 1.2), insistant sur la complexité et la richesses des structures. Les données que nous traitons sont alors décrites ainsi que les problèmes d'instabilité liés à leur acquisition (section 1.3). Le chapitre se termine par le détail de la problématique à laquelle répond cette thèse et par le plan du manuscrit (section 1.4). Il s'agit de proposer un modèle générique de segmentation de la façade quel que soit son style d'architecture.

1.1 Applications

Les algorithmes de segmentation de façades pourraient être utilisés dans trois grandes applications. La première et la plus directe serait d'utiliser leurs résultats pour accroître le niveau de détail des modèles 3D de villes générés par imagerie aérienne ou satellitaire (figure 1.1.a).

D'autre part, les images de façades sont souvent acquises de manière automatique en très grande quantité (*cf* système *Stéréopolis* en section 1.3). Une deuxième application pourraient alors se trouver dans la compression de la texture de ces images en vue de leur intégration dans des systèmes embarqués.

Enfin, le développement progressif de plates-formes de simulation des milieux urbains entraîne le besoin d'outils d'indexation d'images. Il devient important de savoir apparier une image acquise par un particulier à sa correspondante la plus ressemblante contenue dans une base géo-localisée. Dans cette idée, une troisième application pourrait être la détection d'indices géométriques comme des régions stables au point de vue d'acquisition (figure 1.1.b).

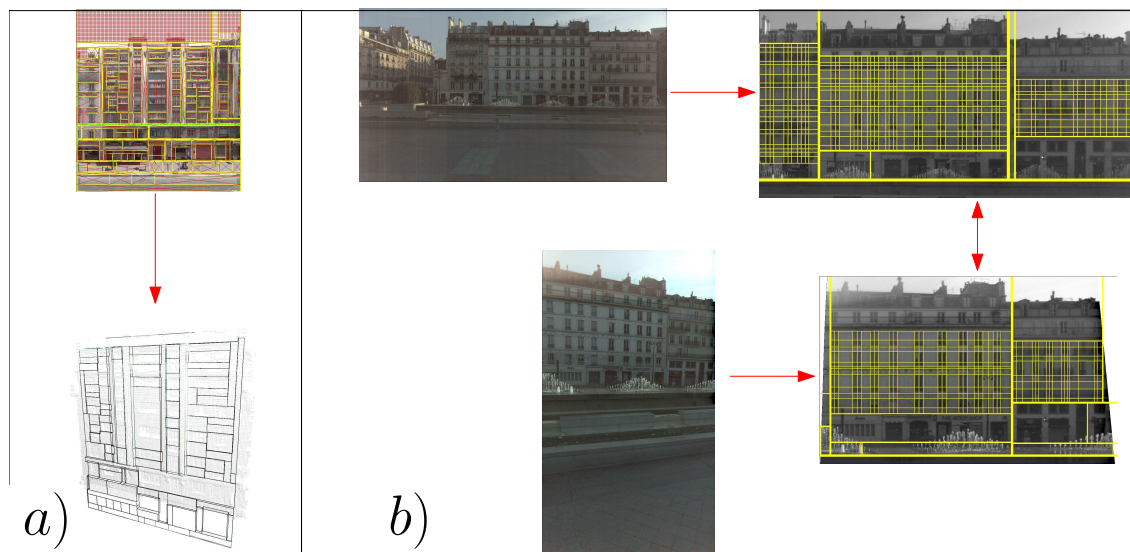


FIGURE 1.1 – Applications de la thèse : a) Découpage réaliste de la façade pour accroître le niveau de détail d'un modèle 3D ; b) Génération d'indices géométriques en vue d'un appariement de façades.

1.2 Architecture de la façade

Depuis le Moyen-Âge, les styles d'architecture des bâtiments se multiplient, principalement au gré des évolutions techniques, des matériaux utilisés, des usages et des lois. Mignot [2004] puis Karantzas et al. [2008] détaillent ce phénomène à travers les monuments de Paris.

La description que nous proposons dans cette section concerne principalement les villes françaises. Ces dernières représentent la cible des traitements menés par le laboratoire *MATIS* de l'*Institut géographique national* dans le cadre duquel s'est effectuée la thèse. Notons que la-dite description peut également s'appliquer aux façades d'autres pays.

Nous montrons que l'architecture se caractérise, d'un côté par une certaine régularité, mais d'un autre, par un nombre de propriétés structurelles qui *camouflent* cette même régularité. Nous verrons que la problématique de cette thèse consiste à segmenter la façade malgré ce *camouflage* qui se retrouve dans la diversité des structures, dans leurs mouvements ainsi que dans l'évolution permanente des styles d'architecture.

1.2.1 Tendances architecturales fréquentes

Il existe certaines tendances architecturales communes à la plupart des façades de style antérieur à 1930 (exemple sur la figure 1.2).

1.2.1.1 Empilement des niveaux traditionnels

Nombre de façades traditionnelles sont constituées de l'empilement de quelques uns des niveaux suivants :

Rez de chaussée comportant les premières habitations ou des commerces ;

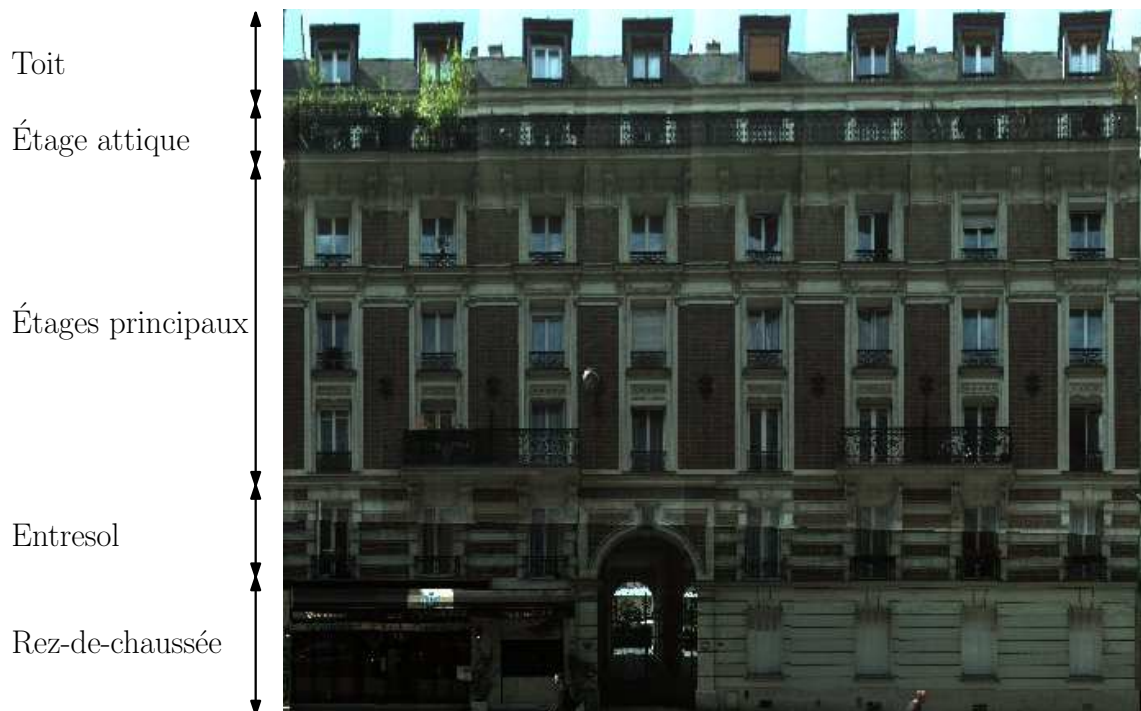


FIGURE 1.2 – Tendances architecturales typiques des façades traditionnelles : empiement en niveaux et rythme des travées.

Entresol si le rez-de-chaussée est suffisamment haut ;

Étages principaux les façades traditionnelles en comptent généralement de 1 à 6, les constructions récentes pouvant en compter beaucoup plus ;

Étage attique plus petit que les autres étages et généralement en arrière plan avec balcon filant éventuel ;

Toit en fonction de l'inclinaison du plan, les baies peuvent être des chiens assis ou des lucarnes.

1.2.1.2 Fréquence des travées

L'étage classique suit une alternance de « pleins », les trumeaux, et de « vides », les baies. Au niveau des étages, la plupart des baies servent de fenêtres, les portes restant souvent au rez de chaussée. Trumeaux et fenêtres se superposent généralement en travées (cf figure 1.2).

1.2.2 Diversité des structures

Même si beaucoup de façades rentrent dans le schéma structurel des étages et travées que nous venons de décrire, la richesse et la variété de leurs structures complexifient considérablement leur architecture. Nous regroupons ces dernières en trois ensembles : les baies, les ornements et les consolidations. La figure 1.3 en présente

un échantillon qui fournit le fil rouge à la brève description que propose cette sous-section. Une présentation beaucoup plus complète, spécifique à la ville de Paris, est proposée par Mignot [2004].

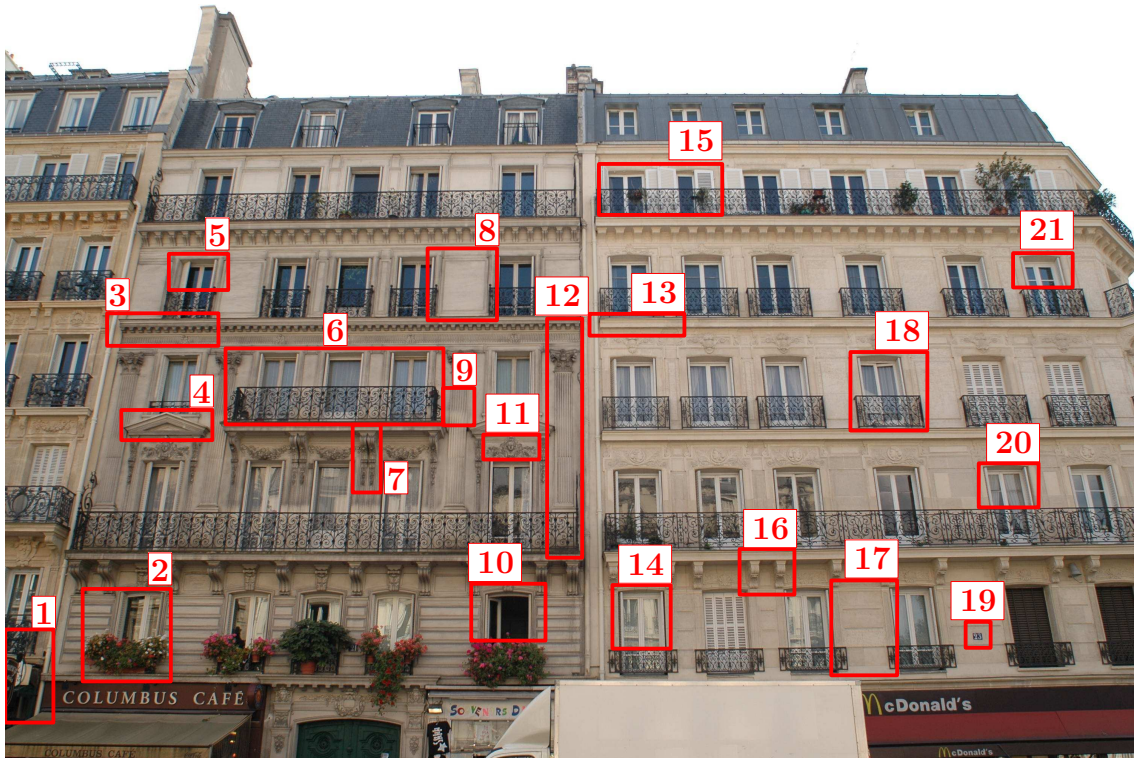


FIGURE 1.3 – Exemples de structures pouvant apparaître sur une façade (cf texte pour la signification de chacune).

1.2.2.1 Baies

La baie est l'élément le plus fréquent sur les façades et le plus complexe. Outre les mouvements qu'elle engendre et que nous étudions plus loin (cf figures 1.8 et 1.9), elle se caractérise par une richesse de matériaux et d'un nombre des plans importants. En ville, elle est close soit par des vitres, soit par un ou deux battants pleins.



FIGURE 1.4 – Diversité des matériaux et des plans des baies : 18) Balustrade en fer forgé ; 6) Balcon filant du troisième étage ; 11) Agrafe au sommet du linteau ; 21) Meneaux séparant les vitres.

Diversité des matériaux La baie est constituée d'un grand nombre de matériaux. Elle est encadrée par un chambranle, traditionnellement en bois ou en pierre. La figure 1.4.11 présente par exemple une agrafe (clé ornée du linteau) sous un fronton. Sur le chambranle sont attachées des vitres qui isolent l'habitat en retrait du plan de la façade. Les fenêtres à *la française* tendent à se répandre à partir du XVIII^e : deux vantaux avec des rangées de petits carreaux séparés par des meneaux (la figure 1.4.21 présente les traverses supérieures des vitres). Depuis la fin du XX^e, les matériaux de type *PVC* se sont standardisés, remplaçant les montants traditionnels en bois.

Assemblage multi-planaire Plusieurs plans peuvent intervenir dans l'architecture des baies avec des balustrades ou des balcons (figures 1.4.18 et 1.4.6). Elles peuvent prendre la forme de « bow-windows » (oriels modernes). Les balcons peuvent être construits en loggias. L'ouverture des vitres et des contrevents engendrent également de nouveaux plans. La complexité sur la figure 1.6.b de la loggia haute entre deux bow-windows, typique de la première décennie du XX^e siècle, ne sont pas des cas isolés à l'échelle d'une ville comme Paris.

1.2.2.2 Ornaments

Le plus simple ornement dont peut se parer le mur de la façade est la nature et la couleur de son matériau. Sur les façades traditionnelles, des cordons, moulures unies et fines, peuvent être placés entre les étages (*cf* figure 1.5.13) ou sous les fenêtres. Un entablement, bandeau très important, peut couronner la façade (soit directement sous les combles, soit avant le retrait de l'étage attique). Des frontons peuvent couronner les baies (figure 1.5.4) ou toute la façade. Ils sont composés d'un cadre mouluré et d'un tympan. Des lignes de refend peuvent être gravées pour décorer ainsi que des tables, surface en relief ou en creux généralement rectangulaire (figure 1.5.8).



FIGURE 1.5 – Ornaments de la façade : 4) Fronton triangulaire au-dessus d'une fenêtre ; 13) Cordon sous le quatrième étage ; 8) Table sur un trumeau ; 19) Plaque du numéro de l'immeuble ; 1) Enseigne en fer forgé (tronquée par le bord de l'image) ; 9) Couche de pollution.

D'autres structures plus spécifiques ou accidentelles apparaissent régulièrement sur les façades comme les plaques (figure 1.5.19), les mascarons, les statues (*cf* figure 1.6.a) et enseignes (panneau, sculpture ou objet caractéristique signalant un

établissement commercial : cf figure 1.5.1). Enfin, la pierre peut être sujette à des dégradations locales (tâches, érosion, etc.) ou plus étendues sous forme de couches sombres de pollution. La figure 1.5.9 illustre, par exemple, le cas d'un contraste probablement dû à un petit ravalement local du bas du pilastre.



FIGURE 1.6 – Diversité des styles de façades : a) Facade néoclassique de 1831 (par Louis Mayet, au 3 rue Soufflot à Paris) b) Style moderne (par Raoul Brandon, au 199-201 rue de Charenton à Paris, primée en 1911) ; c) Immeuble à petit loyer de 1912 (par Louis Bonnier, au 67 rue des Meuniers à Paris).

1.2.2.3 Structures de consolidation

Sur une façade traditionnelle, un chaînage de pierres de taille, de pierres choisies ou de briques peut consolider le mur. La figure 1.7.17 présente un bossage : un alignement de pierres en saillie sur un trumeau. Le toit et certains étages s'appuient



FIGURE 1.7 – Structures de consolidation : 12) Pilastre soutenant la corniche du quatrième étage ; 17) Bossage sur trumeau ; 16) Corbeaux soutenant le balcon du deuxième étage ; 7) Consoles soutenant le balcon du troisième étage ; 3) Corniche à modillons soutenant le quatrième étage.

sur une corniche (*cf* figure 1.7.3). Les saillies peuvent être soutenues par des corbelets, des corbeaux (*cf* figure 1.7.16), des consoles (*cf* figure 1.7.7) ou des pilastres (*cf* figure 1.7.12) solidaires de l'appareil de la façade. Les éléments plus importants peuvent être soutenus par des colonnettes ou même des piliers qui peuvent prendre la forme d'atlantes ou de cariatides. Des ancrs, pièces métalliques reliant deux éléments de construction, peuvent apparaître pour éviter que les éléments ne s'écartent. Des arcs de décharge peuvent être situés au-dessus d'une baie pour réduire la pression du mur sur le linteau.

1.2.3 Mouvements de la façade

L'aspect de la façade évolue au cours de la journée. La principale source de mouvement se retrouve dans les baies de par leur fonction : aération ou isolation, adaptation mécanique et structurelle à la lumière extérieure, protection contre les intempéries ou simplement éléments de décorations. Ces mouvements se manifestent par la position des vitres, leurs reflets, les objets ou les personnes qu'elles laissent apparaître et les systèmes de protection intérieurs et extérieurs.

1.2.3.1 Reflet des vitres

Si elles ne sont cachées par aucun système de protection extérieur, les vitres reflètent la lumière provenant de leur environnement (figure 1.8.14). Ce reflet varie en fonction du degré de leur ouverture, de la localisation des sources d'éclairage et du point de vue de l'acquisition.



FIGURE 1.8 – Variation des baies : 10) Pivotement des vitres d’une pièce sombre du premier étage ; 20) Rideaux à demi tirés à l’intérieur ; 14) Reflet, sur la vitre, d’une fenêtre de l’immeuble en vis à vis ; 2) Fleurs disposées par l’habitant.

1.2.3.2 Position des vitres

Les vitres pivotent sur leur charnière (figure 1.8.10) ou coulissent sur leur châssis pour aérer ou isoler l’habitat.

1.2.3.3 Objets ou personnes de l’habitat

Les baies peuvent montrer les habitants ou des objets laissés plus ou moins provisoirement par ces derniers (pots de fleurs sur un rebord de fenêtre comme par exemple sur la figure 1.8.2).

1.2.3.4 Système intérieur de protection

Des rideaux (figure 1.8.20) ou des stores ornent l’intérieur des fenêtres dans l’habitat et peuvent être tirés pour cacher la lumière extérieure ou simplement abriter des regards extérieurs.

1.2.3.5 Système extérieur de protection



FIGURE 1.9 – Variations des contrevents : 5) Ouverture sur les trumeaux ; 15) Ouverture dans l’embrasure.

Des contrevents sont généralement installés à l’extérieur des baies. Ils ont succédé à partir de 1830 aux volets installés à l’intérieur de l’habitat et se sont généralisés pour protéger les vitres des intempéries. Les battants s’ouvrent en cachant les trumeaux (figure 1.9.5). Un certain nombre sont articulés pour se replier dans l’embrasure (figure 1.9.15). Les contrevents pleins isolent l’habitat de la lumière.

Pour laisser passer cette dernière tout en restant fermés, beaucoup comportent des persiennes. Les architectures contemporaines et la rénovation des immeubles anciens tendent à généraliser l'usage des volets roulants.

1.2.4 Évolution permanente des styles d'architecture

Outre la diversité des structures et leurs mouvements auxquels nous avons fait allusion, de nombreuses façades se distinguent dans leur architecture, à commencer par les monuments publics (hôtel de ville avec parvis, colonnettes, etc.). De plus en plus de styles d'architectures postérieurs à 1930 font apparaître de nouvelles tendances architecturales. Les trumeaux de la barre d'habitations de neuf étages de la figure 1.10.a tendent à disparaître. La façade de style moderne sur la figure 1.10.b est composée de trois grandes tours reliées tous les trois étages par des ponts et avec un étage sur trois traité comme des sculptures abstraites. L'effet de volumes déboîtés est mis en valeur par le contraste entre les surfaces de béton lisses et les fenêtres en bande. L'immeuble de la figure 1.10.c témoigne de l'effort de l'architecte à sortir des schémas classiques : petite villa posée au sommet d'un empilement de longs balcons filants.

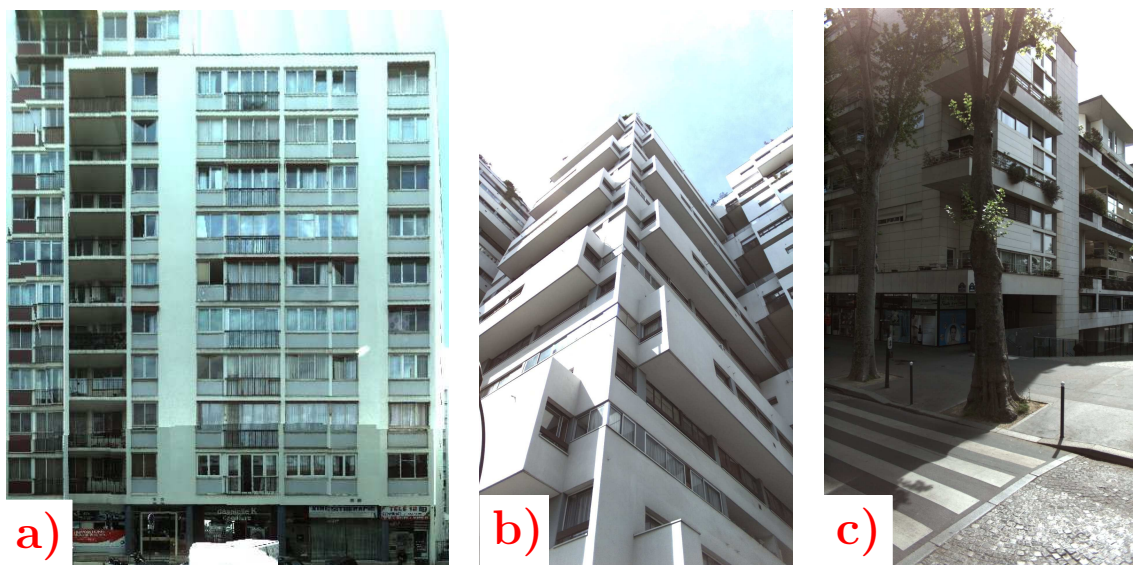


FIGURE 1.10 – Diversité des styles de façades : a) Barre de l'après-guerre (au 238 rue de Charenton à Paris) ; b) Façade style moderne de 1962 (agence Anger-Heymann, au 15-21 rue Erard à Paris) ; c) Immeuble-villa de 1994 (par Christian de Portzamparc au 25 rue de l'Ambroisie à Paris).

Nous pouvons imaginer que les façades de demain n'auront peut-être plus vraiment la même allure que celles d'aujourd'hui. Leur architecture évolue en permanence. Dans les années à venir, une nouvelle tendance urbanistique pourrait voir le jour avec par exemple l'intégration de systèmes techniques permettant la production d'énergie électrique locale (panneaux solaires thermiques, photovoltaïques, éoliennes urbaines). D'autre part, bien que la politique urbanistique française tende à renforcer la cohérence architecturale dans un même voisinage, il n'est pas évident de trouver des quartiers ne contenant que des façades d'une même époque. Les

figures 1.6 et 1.10 présentent des façades de styles totalement différents qui sont pourtant présentes dans le même arrondissement de Paris.

1.3 Données

Nous travaillons exclusivement à partir d’images isolées, acquises d’un point de vue terrestre et d’orientation interne connue. Nous n’utilisons pas d’autre information volontairement comme de l’imagerie 3D (nuage de points ou autre) car, dans le cadre d’applications telles que l’indexation, la reconnaissance d’image ou la géolocalisation, nous pourrions disposer uniquement d’une simple photo acquise par un système nomade (appareil téléphonique mobile, *iPAD*, etc.). Cet aspect du travail de la thèse est important. Contrairement à ce qui est opéré dans la plupart des approches de l’état de l’art, nous n’imposons pas de contraintes sur le contenu de l’image traitée. Cette dernière peut comporter une ou plusieurs façades dans leur intégralité ou bien tronquées. Le cas typique est une image ne possédant qu’une ou deux parties hautes de façades comme sur la figure 1.6.b. Il peut arriver également que plusieurs plans soient visibles comme sur les figures 1.13 et 1.14.

Notre analyse porte majoritairement sur les images du système *Stéréopolis*. Elle doit rester suffisamment générique pour pouvoir s’appliquer à d’autres images, d’une source quelconque.

Cette section décrit d’abord le système *Stéréopolis*, puis présente quelques problèmes d’acquisitions auxquels peuvent être sujettes les données. Le processus de redressement des images dans le plan des façades est ensuite expliqué.



FIGURE 1.11 – Système de numérisation *Stéréopolis*

1.3.1 Données du système Stéréopolis

Les données étudiées dans cette thèse sont principalement les images acquises par le système de numérisation *Stéréopolis* développé par la laboratoire *MATIS* de l'*IGN*. Ce système, installé sur le toit d'un véhicule, permet d'acquérir des données 3D géo-référencées en milieu urbain (images panoramiques, images stéréoscopiques et nuages de points 3D). Il est visible sur la figure 1.11. Dix caméras disposées sur le mat central permettent d'acquérir l'image panoramique (figure 1.12). Un laser rotatif placé au-dessus de ces caméras permet de scanner l'environnement du véhicule et deux scanners à balayage orientés à 45° et placés au pied du mât permettent de scanner les façades lorsque le véhicule est en déplacement. Le système de positionnement se compose d'une centrale inertielle associée à deux antennes *GPS* et à un odomètre placé sur la roue arrière gauche.

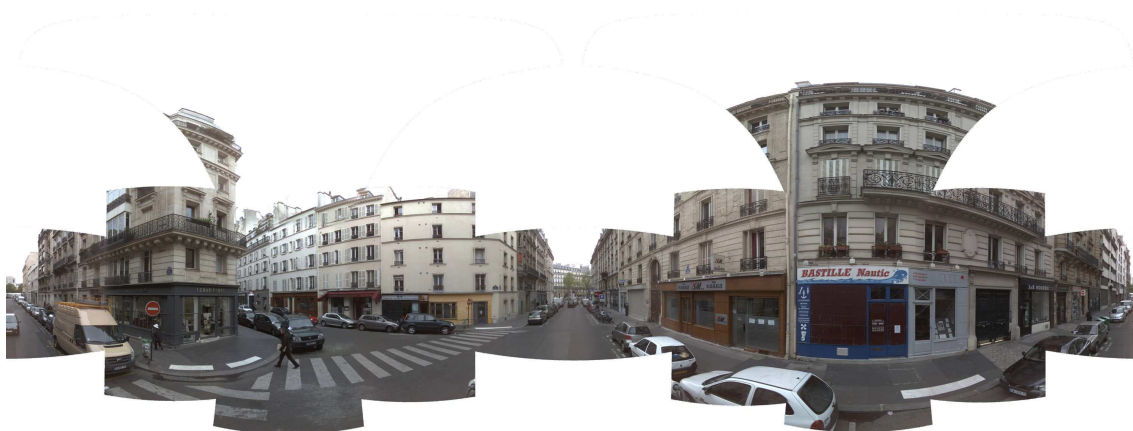


FIGURE 1.12 – Image panoramique de la rue

Les images que nous utilisons sont acquises de jour et à partir des appareils latéraux dépointés vers le haut. Elle représentent pour la plupart une, deux ou plus rarement trois façades ou parties de façades à quelques mètres du camion. Leur résolution est centimétrique pour une dimension de 2 méga-pixels. Il s'agit d'un traitement massif qui concerne 60 000 points de vue, pour ne parler par exemple que du 12^e arrondissement de Paris.

1.3.2 Sources d'instabilité liées à l'acquisition

Certaines propriétés de l'image sont sujettes à des problèmes d'instabilité liés aux conditions d'acquisition. Le *Stéréopolis* n'échappe pas à ces problèmes.

1.3.2.1 Instabilités dues à l'éclairage

La radiométrie des matériaux de la façade dépend principalement de la nature de ces matériaux mais elle dépend également beaucoup de deux propriétés de leur exposition :

L'intensité de la lumière Le climat joue beaucoup sur la couleur des matériaux de la façade.

La localisation des sources de lumière Les saillies et retraits génèrent des ombres du côté non exposé. Pour un éclairage diffus comme sur celui de la figure 1.3, les ombres sont très atténuées. Au contraire, la figure 1.10.b présente une façade éclairée fortement et dont chaque balcon projette des ombres contrastées avec les structures voisines.

De plus, le paramétrage de l'acquisition peut parfois entraîner des phénomènes de saturation de la dynamique du capteur (si le soleil se trouve dans l'objectif).

1.3.2.2 Instabilités dues au point de vue

L'apparence d'un même objet peut dépendre de sa position dans l'image. Nous recensons trois raisons possibles à ce phénomène :

Parties cachées par les saillies Théoriquement, pour que toutes les régions soient visibles, il faudrait que le capteur soit *placé à un point de vue infini* en face de la façade. Ceci n'est pas réalisable donc certaines parties se retrouvent cachées par les saillies.

Reflets sur les matériaux Une réflexion spéculaire se produit sur certains matériaux polis et éclairés, comme les vitres (*cf* figure 1.8.14). Ces reflets dépendent de la position du capteur.

L'avant plan Certains objets situés entre le capteur et la façade cachent cette dernière. Il peut s'agir de végétation, de piétons, de voitures, de lampadaires, de panneaux (signalisation, publicitaire) ou de tout autre mobilier urbain (abris de bus, grue, échafaudage, etc.). Le camion en bas de la figure 1.3 se trouve par exemple devant deux commerces.

1.3.3 Redressement multi-planaire

Nous verrons qu'un des *a priori* qu'utilise cette thèse est l'orthogonalité des alignements de structures de la façade. Pour utiliser cette propriété, nous projetons l'image dans les différents plans qu'elle peut contenir. Le système *Stéréopolis* fournit ces derniers à partir du nuage de points laser et des méta-données de géo-localisation.

Pour une image acquise à partir d'un système ne possédant pas de telles données complémentaires, la géométrie de ces plans peut être estimée à partir d'un couple de points de fuite extraits suivant la méthode de Kalantari et al. [2008]. Cette dernière est indépendante des paramètres intrinsèques de la prise de vue et n'impose pas de contrainte sur le nombre de points de fuite réels dans l'image. Elle est robuste, très précise et estime l'incertitude des contours de l'image sur les résultats de redressement. Son principe consiste à projeter les contours de l'image dans une sphère de Gauss centrée en un point origine O de l'espace. Chaque contour C_i de l'image est associé à un vecteur $\vec{n}_i$ unitaire, normal au plan (O, C_i) et ayant pour origine le centre de la sphère. Une propriété géométrique importante de cet espace assure que les contours associés à un même point de fuite, dans l'image, possèdent des vecteurs normaux coplanaires dans la sphère de Gauss. Ce plan formé par les vecteurs normaux est appelé *plan de fuite* et sa direction normale coupe le plan image au point de fuite. La méthode consiste à détecter les plans de fuite les plus représentés dans

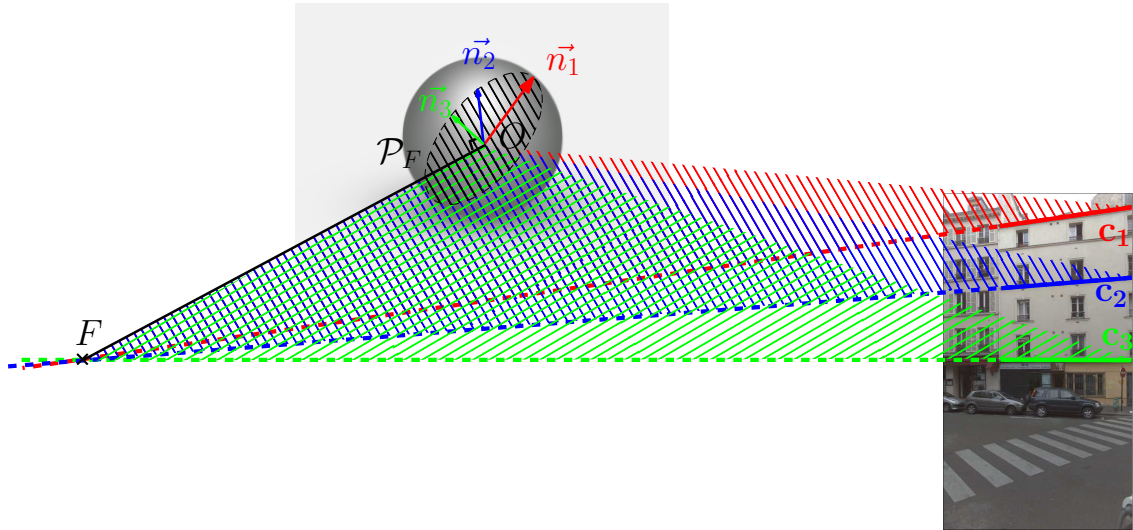


FIGURE 1.13 – Estimation du point de fuite F par projection des contours C_1 , C_2 et C_3 dans la sphère de Gauss.

la sphère de Gauss. Une homographie permet ensuite de projeter l'image dans les plans définis par chacun des couples de points de fuite détectés.

Ce calcul est illustré sur la figure 1.13 pour l'estimation d'un point de fuite F à partir de trois contours de l'image C_1 , C_2 et C_3 . Ces trois contours sont respectivement associés à trois vecteurs normaux dans la sphère de Gauss $\vec{n}_1$, $\vec{n}_2$ et $\vec{n}_3$. Ces trois vecteurs normaux forment le plan de fuite $\mathcal{P}_F$ dont la direction normale coupe le plan image au point F recherché. La figure 1.14 présente un redressement de l'image dans les deux plans principaux qu'elle contient.

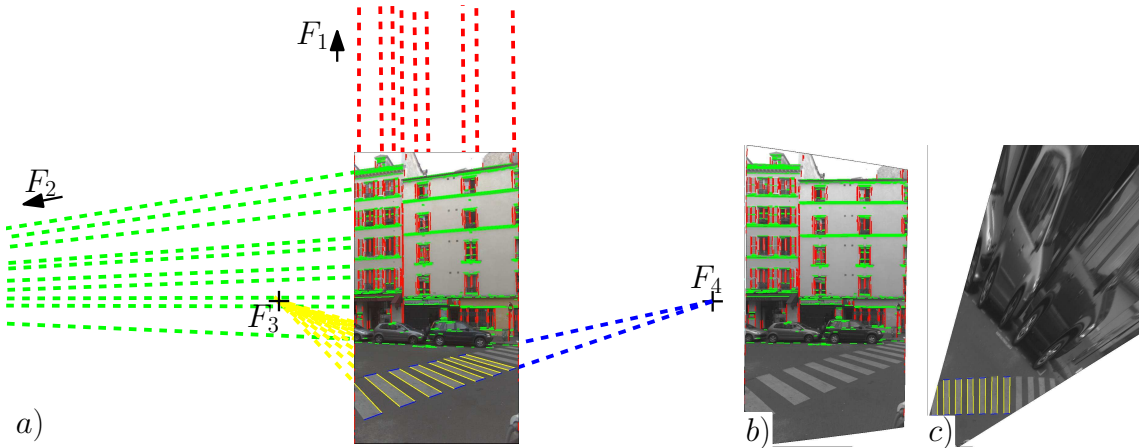


FIGURE 1.14 – Redressement multi-planaire. a) Détection des points de fuite principaux de l'image originale. b) Redressement dans le plan formé par les points de fuite F_1 et F_2 . c) Redressement dans le plan formé par les points de fuite F_3 et F_4 .

Notons que, dans la plupart des cas, l'image redressée obtenue n'est plus rectangulaire. Nous utilisons dans tous nos calculs un masque binaire de pixels significatifs pour son analyse. Les pixels non significatifs ne sont pas pris en compte dans les calculs.

1.4 Problématique et plan du manuscrit

Nous avons vu combien l'architecture de façade est caractérisée, d'un côté, par une certaine tendance à la régularité des étages et travées, et d'un autre, par un ensemble de propriétés structurelles qui camouflent cette même régularité. Nous avons en effet étudié la densité et la diversité des structures de type baies, ornements et éléments de consolidation ainsi que le mouvement que comporte la façade et l'évolution permanente de ses styles d'architecture. Nous avons abordé d'autre part les problèmes d'instabilité de l'acquisition des images liés à l'éclairage et au point de vue.

La problématique de cette thèse répond au besoin d'outils de segmentation générique embrassant cette complexité architecturale et ces problèmes d'instabilité. Les résultats obtenus pourront, comme nous l'avons illustré en début de chapitre, permettre le découpage réaliste de la façade en textures élémentaires pour enrichir un modèle 3D ou bien la compression de la façade de manière compacte ou encore la génération d'indices géométriques en vue d'une indexation. Les images sont traitées isolément. Elles peuvent comporter une ou plusieurs façades, en parties ou bien dans leur intégralité, ce qui est très faiblement étudié dans l'état de l'art. Les traitements sont censés s'appliquer à un grand volume d'images et leur paramétrage doit donc être effectué en une seule fois.

Les éléments de solutions que nous proposons consistent en des approches asémantiques reposant sur les propriétés radiométriques et géométriques des façades comme l'alignement et la répétitivité de leurs structures.

La suite de ce manuscrit présente en premier lieu un état de l'art de la segmentation de façades (chapitre 2). Nous montrons alors que beaucoup d'auteurs ne se concentrent que sur un style donné et que d'autres ont fait le même choix que nous de minimiser l'emploi d'*a priori* sémantiques. Nous exposons ensuite une segmentation hiérarchique que nous proposons pour représenter la façade en modèles élémentaires (chapitre 3). Nous décrivons aussi les avancées que nous avons menées dans la recherche de structures répétitives (chapitres 4 et 5). Nous terminons par une conclusion qui souligne les points forts et les limites des approches proposées et qui énonce quelques perspectives (chapitre 6).

Chapitre 2

État de l'art

Les images de façades acquises d'un point de vue terrestre intéressent différentes communautés de recherche (photogrammétrie, vision par ordinateur, synthèse d'images, reconnaissance d'objets, etc.) pour plusieurs applications :

Segmentation Regroupement de pixels voisins et présentant des propriétés géométriques ou radiométriques analogues en super-pixels pouvant être utiles à beaucoup d'applications ;

Reconstruction Pour apporter du relief et du réalisme aux images de façades ;

Appariement Extraire des descripteurs d'indexation pour reconnaître, dans une base d'images pré-traitées, quelle image ressemble le mieux à une image requête extérieure à la base ;

Compression Pour stocker des textures réalistes dans un espace mémoire restreint ;

Reconnaissance de façades Localiser dans une image les éventuelles zones de façade qu'elle contient.

Le but de cette thèse réside dans la segmentation générique d'un grand volume d'images acquises d'un point de vue terrestre contenant une zone de façade (partielle, entière ou multiple).

Cet état de l'art ne concerne que des images contenant des façades. Si ce n'est pas le cas systématiquement, il existe des techniques pour confirmer ou infirmer leur présence. La voie classique est d'apprendre la texture-type d'une façade à partir d'un ensemble de jeux de façades-type. Drauschke et Mayer [2010] proposent une étude de sept bancs de filtres les plus utilisés pour reconnaître les régions de la façade, de la rue, du ciel et de la végétation dans une image de scène urbaine. Ils comparent leurs résultats avec deux techniques de classification : des forêts aléatoires et une analyse discriminante linéaire. Dans un tout autre ordre d'idée, l'approche de Korč et Förstner [2008a] est non supervisée et est basée sur la détection de régions structurées à partir d'histogrammes de norme et d'orientation de gradients. Ils détectent les régions de façades comme étant celles qui sont structurées en utilisant un champ aléatoire conditionnel (*CRF*).

Ce chapitre présente les principales techniques de la littérature pour l'étiquetage des éléments de la façade (section 2.1), la séparation de façades (section 2.2, l'estimation des profondeurs (section 2.3) et l'appariement de façades (section 2.3). Un

point est enfin donné sur la segmentation générique qui n'introduit que le moins possible d'*a priori* (section 2.5).

2.1 Étiquetage des éléments de la façade

Il est possible de reconnaître les éléments de la façade par leur position et leur apparence.

2.1.1 La position

Les éléments peuvent être extraits suivant un modèle et une position de la façade dans l'image. La partie supérieure est probablement le toit. Si des éléments sont alignés plus ou moins régulièrement dessous, il s'agit certainement des fenêtres.

Beaucoup de techniques se focalisent sur la détection de ces derniers éléments qui sont les plus fréquents. À partir de certains indices locaux (informations pixellaires, points d'intérêt, rectangles, etc.), il s'agit de générer des hypothèses de fenêtres à contraindre dans des alignements et des quadrillages (stratégie de type *bottom-up* : figure 2.1.a). D'autres approches segmentent la façade en étages et travées (figure 2.1.b). La contrainte des alignements permet de compenser les erreurs de localisation des éléments perturbés par le bruit local, par certains problèmes d'éclairage et de perspective ainsi que par les objets mobiles de la façade, la végétation et l'avant plan.

Pour un style de façade donné, il est courant de formaliser toutes les règles architecturales de ce type de manière hiérarchique. Une formalisation procédurale beaucoup utilisée dans la littérature consiste en l'établissement de grammaires de façades (utilisées par exemple par Teboul et al. [2010] ou Ripperda [2008]).

Enfin, indépendamment de leur position absolue dans l'image, la nature des éléments peut se déduire de celle de leurs éléments voisins. Certains auteurs les étiquettent en utilisant des statistiques sur leurs règles topologiques (figure 2.1.c).

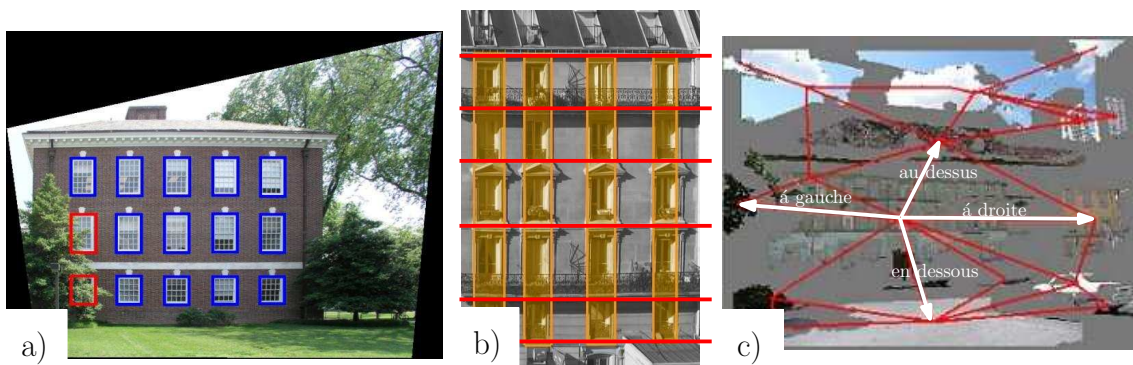


FIGURE 2.1 – Approches de sémantisation de la façade à partir de la position de ses éléments : a) La contrainte du quadrillage (Korah et Rasmussen [2008]) b) La logique des étages et travées (Koutsourakis et al. [2009]) ; c) Les relations topologiques (Recky et Leberl [2009]).

Toutes les approches présentées peuvent éventuellement précéder une phase de validation supervisée.

2.1.1.1 Le quadrillage de fenêtres

Un certain nombre d’approches consistent à générer un ensemble d’hypothèses de fenêtres à contraindre dans des alignements et des quadrillages (stratégie de type *bottom-up*). La figure 2.2 présente trois approches différentes de génération d’hypothèses de fenêtres.

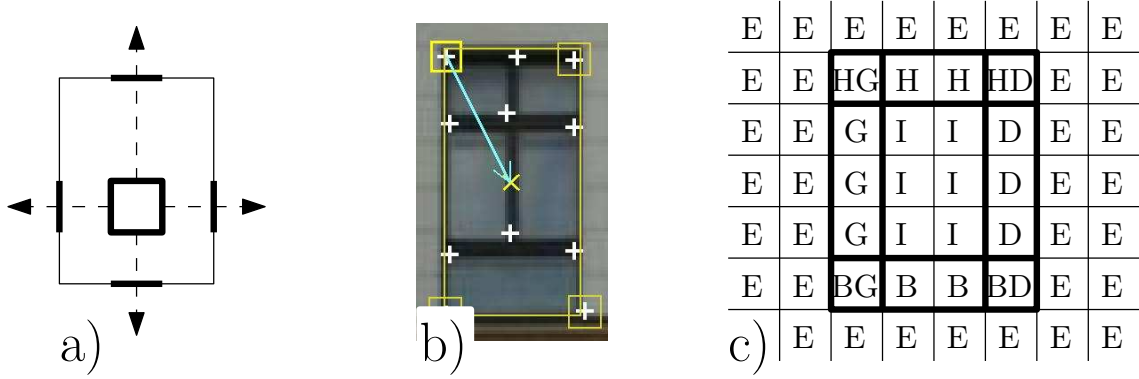


FIGURE 2.2 – Exemple d’approches de génération d’hypothèses de fenêtres : a) Croissance de régions rectangulaires réparties régulièrement sur la surface de la façade (Wang et al. [2002]), b) Estimation supervisée de la position des fenêtres à partir de leurs coins (Reznik et Mayer [2008]), c) Coupe minimale d’un graphe dont chaque nœud est un pixel pouvant être associé à dix classes possibles : extérieur (E), intérieur (I), haut (H), bas (B), gauche (G), droite (D), etc. (Čech et Sára [2007]).

Certains auteurs procèdent en deux phases distinctes (éventuellement alternées). Sur des images ortho-rectifiées, Wang et al. [2002] répartissent régulièrement des *graines de régions rectangulaires*. La croissance de ces graines (figure 2.2.a) aboutit à un ensemble de rectangles qu’ils regroupent par gabarit et périodicité. Toujours avec des images ortho-rectifiées, Reznik et Mayer [2008] disposent d’une base d’images de coins de fenêtres qu’ils corréleront à des points d’intérêts de l’image d’entrée. Chaque coin de la base est associé à une position relative d’un centre de fenêtre (figure 2.2.b). À partir des points d’intérêt similaires aux coins, ils accumulent les hypothèses de centres pour générer des hypothèses de fenêtres. Il s’agit d’un modèle implicite de formes (Leibe et al. [2004]). Ces hypothèses sont alors contraintes par des alignements selon le critère de (Akaike [1973]). Korč et Förstner [2008b] génèrent des hypothèses de rectangles par une segmentation itérative à *dérive moyennée* (de type *mean shift*) en recherchant des rectangles sombres et de radiométrie non uniforme mais symétrique. Ils sélectionnent un ensemble de rectangles non recouvrants de manière supervisée. Sur des images qui ne sont pas nécessairement ortho-rectifiées, Korah et Rasmussen [2008] génèrent plusieurs centaines d’hypothèses de rectangles en perspective (de la même manière que Han et Zhu [2005] avec des paires de segments parallèles). Ils agencent un sous-ensemble optimal de ces rectangles dans un quadrillage modélisé par des textures quasi-régulières (*Near Regular Textures*, Liu et al. [2004]).

D’autres auteurs construisent ce type de quadrillage de fenêtres d’une seule passe en utilisant la théorie des graphes. Sous l’hypothèse d’une apparence de façade simplifiée et ortho-rectifiée, Čech et Sára [2007] associent chaque pixel de l’image à un

nœud du graphe. Ils maximisent une probabilité *a posteriori* associée à un ensemble d'étiquettes pour chaque pixel (figure 2.2.c) avec la position par rapport aux fenêtres recherchées (coin, arête, intérieur ou extérieur). Les fenêtres qu'ils recherchent sont assimilées à des rectangles de radiométrie uniforme, non recouvrants, sombres et sur un fond clair. Sirmacek [2011] construit également un graphe à partir d'une segmentation de type *mean shift* qui regroupe les pixels en super-pixels de propriétés radiométriques semblables sur une image ortho-rectifiée. Les nœuds du graphe sont alors les centres de masse des super-pixels et les arêtes sont déterminées par une triangulation de Delaunay sur la sur-segmentation.

2.1.1.2 Logique des étages et travées

Une autre manière de localiser les éléments s'appuie sur des règles architecturales. Il s'agit d'approches descendantes (*Top-down*) : partir de macro-éléments pour retrouver les éléments plus simples. L'organisation en étages et en travées apporte une information globale sur leur agencement. Les auteurs qui basent leur segmentation sur cet *a priori* ont recours à deux types d'analyse :

Projection d'informations locales sur des profils d'accumulations Ces informations locales peuvent être géométriques ou radiométriques. L'analyse des profils *2D* permet de localiser les étages et travées (avec d'éventuels ajustements locaux *a posteriori*).

Recherche de symétries Il s'agit généralement de mesures de similarité entre deux alignements contigus ou de symétrie axiale.

La tendance des étages et des travées n'est qu'un exemple de règle architecturale pouvant se retrouver dans les façades. Le point 2.4 explique brièvement l'utilisation de *grammaires de façades* pour tenir compte d'une hiérarchie formelle de règles de ce type.

Projection d'informations locales Schindler et Joachim [2003] utilisent des profils de projection des profondeurs d'indentation des plans. Ils estiment ces plans dans un nuage de points de la façade (figure 2.3.a). Ils parviennent à reconstruire la géométrie *3D* à partir d'une séquence d'images. L'hypothèse forte ici est que les seuls éléments en saillie ou en retrait de la façade doivent être les fenêtres.

Plusieurs auteurs utilisent le gradient et les contours d'une image ortho-rectifiée. Il existe deux approches complémentaires pour projeter horizontalement ce type d'information sur un profil vertical d'accumulations. Selon que l'objectif est la détection ou respectivement la localisation plus précise des étages, il faut accumuler les informations verticales (c'est à dire relatives à des contours orthogonaux au sol) ou respectivement horizontales (relatives à des contours parallèles au sol). Nous reprenons la terminologie employée par Lee et Nevatia [2004] et désignons par V_V un profil vertical de projection horizontale d'informations verticales. De même, nous appelons H_H les profils horizontaux de projection verticale d'informations horizontales. Nous définissons réciproquement les profils de type V_H et H_V . Les figures 2.3.b et 2.3.c illustrent cette distinction. Nous définissons de même les projections de type VH et VV pour désigner les projections verticales.

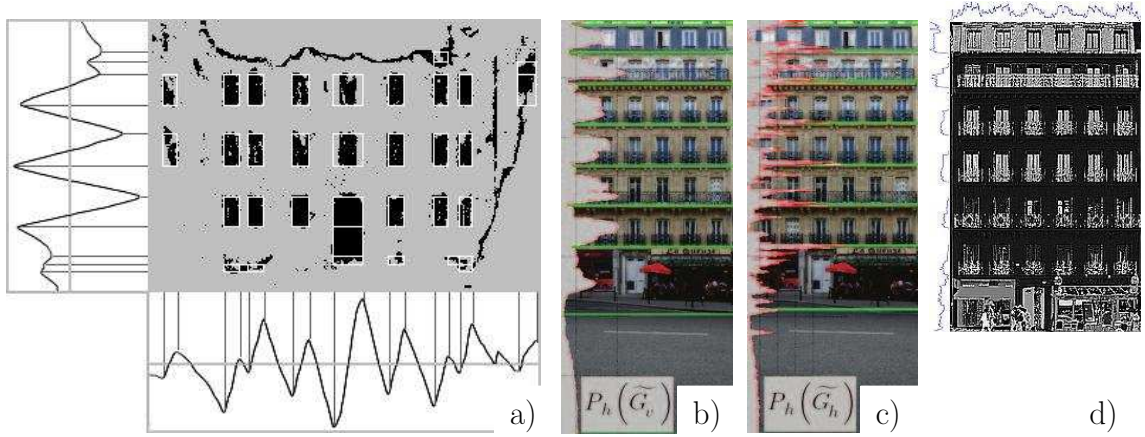


FIGURE 2.3 – Projection de différents types d'informations locales : a) Projection des distances des points par rapport au plan moyen de la façade (Schindler et Joachim [2003]); b) et c) Profils de projection de gradients de type V_V , respectivement de type V_H (Hernández et Marcotegui [2009]); d) Projection des valeurs de la teinte de l'image (Liu et Gagalowicz [2010]).

Lee et Nevatia [2004] utilisent des profils de projection des contours de l'image de type H_H et V_V pour quadriller cette dernière. Haugeard et al. [2009], Hernández et Marcotegui [2009] et Recky et Leberl [2010] accumulent les valeurs de gradient pour séparer les étages puis traiter ces derniers indépendamment de la même manière. Haugeard et al. [2009] s'appuient sur le fait que le nombre de modes détectés sur le profil d'accumulations dépend beaucoup du lissage utilisé. Ils appliquent alors plusieurs filtres de lissage de plus en plus importants et sélectionnent le plus cohérent (le dernier pour lequel le nombre de modes reste stable). Ces calculs ont une très faible complexité car ils sont menés sur des profils en une seule dimension. Hernández et Marcotegui [2009] utilisent un procédé similaire à partir d'opérateurs morphologiques (*filtres séquentiels alternés*) et complètent le traitement par des segmentations de type *ligne de partage des eaux* et *k-means* pour localiser les modes. Recky et Leberl [2010] couplent l'utilisation des profils d'accumulation avec une segmentation de type *k-means* de Anagnostopoulos et al. [2005] et obtiennent des super-pixels dans l'espace de couleurs *CIE-Lab*. Les plus grands blocs sont le fond, les autres sont les fenêtres.

D'autres auteurs accumulent des informations radiométriques. Sur des images de façades de style *haussmannien*, Liu et Gagalowicz [2010] accumulent la teinte (figure 2.3.d). Ils normalisent alors le profil vertical pour atténuer le dégradé de luminosité (lumineux en haut, sombre en bas) qui apparaît dans ce cas. Ils opèrent ensuite une série de filtrages pour localiser les modes (filtrages morphologiques pour supprimer le bruit généré par les petites structures locales et filtrage intégrateur pour obtenir un signal triangulaire). Sur le profil horizontal, ils appliquent un filtre de lissage dont la longueur du voisinage est la largeur moyenne d'une fenêtre de l'image. Cette largeur moyenne est obtenue par le maximum du spectre d'une transformée de Fourier de ce même profil.

Ricard et al. [2008] fusionnent des profils de projection d'informations géométriques et radiométriques. Ils utilisent les contours et les droites de l'image ainsi

que les différences radiométriques pixellaires avec la couleur du fond. Cette dernière est estimée dans *le plus grand rectangle vide* (Chazelle et al. [1986]). Ils effectuent plusieurs traitements sur la combinaison de tous ces profils de projections pour supprimer les hautes fréquences et détecter les modes.

Recherche de symétries En complément aux techniques de segmentation de profils de projections, d'autres auteurs comparent les étages et les travées deux à deux. Pour détecter les étages, Müller et al. [2007] mesurent leur similarité par information mutuelle et sélectionnent le meilleur score parmi différentes positions et différentes hauteurs. Ils agissent de même pour les travées. Ils ajustent leur localisation par projection des gradients sur des profils d'accumulations. Koutsourakis et al. [2009] recherchent la partition optimale d'étages, respectivement de travées, non chevauchants et qui se ressemblent (hauteurs, respectivement largeurs, proches et faible somme des différences radiométriques quadratiques). Les limites doivent se situer sur les forts gradients. Entre les travées, une symétrie axiale doit exister sur le fond de mur et ce dernier doit posséder une radiométrie de faible variance.

2.1.1.3 Étiquetage des autres éléments de la façade

Il existe bien d'autres éléments clés de la façade que les fenêtres comme les portes, les balcons ou les corniches (*cf* section 1.2). Beaucoup d'auteurs s'appuient sur une phase préliminaire de l'agencement des fenêtres pour les localiser.

Le fond de mur, schématiquement, peut être extrait directement comme l'élément complémentaire des fenêtres (Sirmacek [2011] et Ricard et al. [2008]). Korah et Rasmussen [2008] estiment la distribution des pixels du mur en maximisant leur probabilité à l'aide de distributions radiométriques gaussiennes extraites de bandeaux encadrant les fenêtres (censés représenter le fond de mur). Cette détection du mur et des fenêtres leur permet de localiser la façade entière dans une boîte englobante.

Les portes sont potentiellement situées aux mêmes endroits que les fenêtres et il est possible de les reconnaître de manière supervisée (Liu et Gagalowicz [2010]). Les balustrades et balcons en fer forgé peuvent être localisés dans une région d'un certain gabarit en bas des fenêtres. Liu et Gagalowicz [2010] proposent une transformée de Fourier et quelques opérations morphologiques pour détecter les pixels du fer forgé. Hernández et al. [2009] proposent un filtrage morphologique particulier : *les nivellements visqueux*, pour les supprimer directement.

Pour éliminer les fausses détections d'étages, Hernández et Marcotegui [2009] comparent les profils de projections verticales des gradients de chaque niveau de la façade : celui du rez-de-chaussée ou de toute autre région de scène urbaine est nettement moins régulier.

Le ciel possède schématiquement une radiométrie uniforme et caractéristique. Sirmacek [2011] le détecte par les super-pixels qui n'appartiennent pas à la façade et qui ont une différence suffisamment faible avec une couleur moyenne calculée dans un échantillon de test. Hernández et Marcotegui [2009] proposent une *ouverture ultime* avec un descripteur de la couleur bleue en haut de l'image et sur ses côtés.

2.1.1.4 Grammaires

De manière générale, les règles architecturales du type des regroupements de fenêtres dans des quadrillages ou en étages et travées, ainsi que celles relatives à la localisation des autres éléments de la façade, peuvent se formaliser de manière hiérarchique dans une grammaire. Cette représentation rajoute davantage encore de contraintes sémantiques mais offre l'avantage d'une partition exhaustive de la façade.

Un historique détaillé du développement de la modélisation procédurale est proposé par Teboul [2011]. Cette formalisation s'est développée depuis l'introduction des *grammaires formelles*, outils linguistiques, dans le domaine informatique par Chomsky [1956]. Aujourd'hui, elle s'applique très bien à la segmentation de façades. Simon et al. [2010] et Teboul et al. [2010], par exemple, segmentent les façades de style *haussmannien* à l'aide d'une *grammaire de formes*. Au préalable, les éléments sont grossièrement reconnus au niveau pixellaire de manière supervisée. L'algorithme consiste ensuite à générer, selon les règles de la grammaire, une segmentation qui soit cohérente avec cette pré-segmentation : que les régions de balcon encadrent des pixels reconnus comme appartenant au balcon et que les régions de fenêtre, de porte, de mur, suivent également cette même contrainte. La figure 2.4.b illustre le phénomène de convergence de la segmentation.

Ripperda [2008] utilise une grammaire comprenant des règles dont les paramètres suivent des lois statistiques apprises à partir d'un ensemble de façades représentatives du style étudié. Ces *a priori* statistiques guident l'échantillonnage des réalisations de la grammaire en restreignant davantage encore le type de façades concernées mais en accélérant le processus de convergence. La figure 2.4.a montre un exemple de quatre règles de grammaire : génération d'un quadrillage régulier, d'un mur de fenêtres, d'un élément de façade ou d'une partie de façade (étage, rez-de-chaussée, etc.).

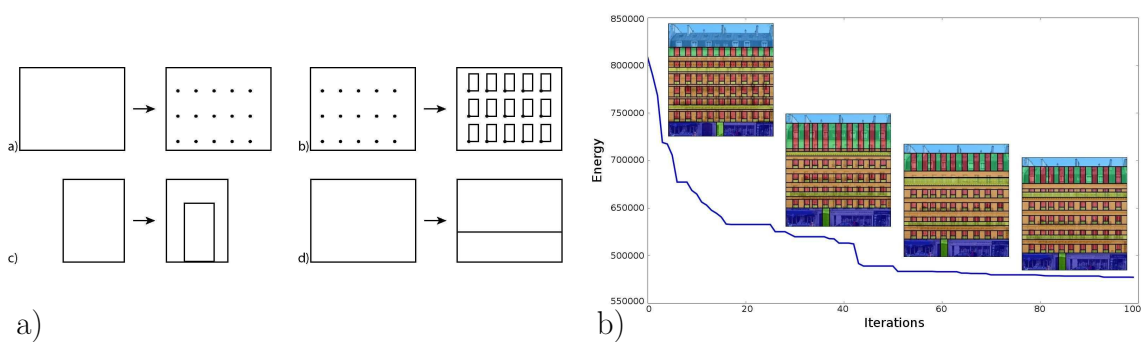


FIGURE 2.4 – A) Exemple de quatre règles de grammaire de façade (Ripperda [2008]) : a) FAÇADE $\rightarrow$ QUADRILLAGERÉGULIER, b) QUADRILLAGERÉGULIER $\rightarrow$ MURFENÊTRES, c) ÉLÉMENTFAÇADE $\rightarrow$ MURPORTE, d) FAÇADE $\rightarrow$ PARTIEFAÇADE PARTIEFAÇADE ; B) Echantillonnage des réalisations de la grammaire pour trouver l'optimale : l'énergie qui caractérise l'adéquation aux données décroît avec les itérations (Simon et al. [2010]).

2.1.1.5 Relations topologiques contextuelles

Parallèlement aux approches précédemment abordées, il est possible d'ajouter des informations statistiques sur la localisation relative à chaque élément les uns par rapport aux autres. Les relations spatiales entre les régions (au-dessus de, à côté, etc.) sont apprises en pré-traitement. Ainsi, Heesch et Petrou [2007] proposent la représentation d'un ensemble quelconque de régions d'intérêt dans des *champs aléatoires de Markov (MRF) non Gibbsien* qui contient ce type d'informations. Recky et Leberl [2009] et Recky et al. [2011] utilisent des champs aléatoires discriminants (*DRF*) pour la segmentation de séquences de façades. Hartz et Neumann [2009] apprennent de la même manière des topologies d'objets de la façade dans un graphe structurel probabiliste à partir d'un petit ensemble d'exemples. Gupta et al. [2010] établissent des relations physiques entre les objets de la façade (3D, poids, etc.).

2.1.2 L'apparence

Il est également possible de reconnaître les éléments par leurs caractéristiques texturales, radiométriques et géométriques. La teinte d'une fenêtre est par exemple généralement plus sombre que le fond de mur.

Beaucoup de ces auteurs utilisent une base de caractéristiques d'éléments-type pour guider (par exemple Korč et Förstner [2008b]) ou valider (par exemple Ricard et al. [2008]) leur détection. D'une manière générale, tous peuvent utiliser ce type d'*a priori* une fois qu'ils détiennent un jeu d'hypothèses d'éléments. Cette sous-section présente différentes caractéristiques utilisées dans l'état de l'art en faisant abstraction de la manière dont sont générées les hypothèses d'éléments. Les trois principaux groupes de caractéristiques concernent des attributs texturaux, radiométriques et géométriques. La figure 2.5 en illustre trois typiques concernant les fenêtres.

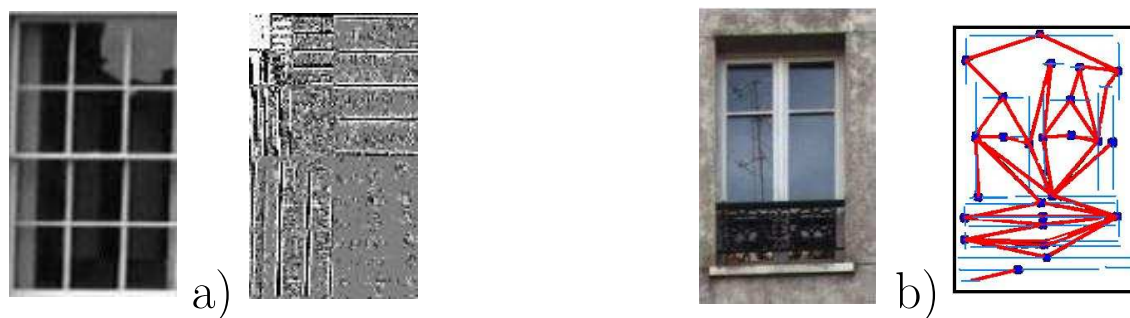


FIGURE 2.5 – Deux caractéristiques de l'apparence d'une fenêtre : a) Représentation de la texture par *ondelettes de Haar* (Dick et al. [2004]) ; b) Représentation de la géométrie par graphe de contours (Haugeard et Philipp-Foliguet [2010]).

Une manière classique de caractériser l'apparence d'une fenêtre est de réduire sa texture à la réponse de bancs de filtres. Ali et al. [2007] utilisent par exemple les *ondelettes de Haar* (exemple de représentation sur la figure 2.5.a) et classifient les fenêtres par un arbre de décision de type *adaboost*.

D'autres auteurs ajoutent des informations radiométriques. Bochko et Petrou [2007] opèrent une classification par séparateurs à vaste marge de la première com-

posante d'une *analyse en composantes principales* des trois canaux rouge, vert, bleu de l'image. Ils comparent leurs résultats à une transformée de *Walsh* (Petrou et Sevilla [2006]) de cette même première composante. Leur objectif est de reconnaître des objets d'intérêt comme des fenêtres, des portes, des tuyaux et des cheminées.

Il est possible de compléter l'apparence des éléments de la façade par des informations géométriques. Bochko et Petrou [2007] comparent leur classification de caractéristiques radiométriques à celle obtenue par la norme du gradient (le maximum des trois canaux). Leurs résultats indiquent que les différentes caractéristiques aboutissent à une qualité globalement similaire. Drauschke [2008] mixe l'utilisation de classifieurs radiométriques et géométriques par un algorithme de type *adaboost* pour reconnaître la façade, les fenêtres, le toit, le ciel et la végétation. Leurs attributs sont basés sur des moyennes et variances radiométriques, des histogrammes d'orientation et de norme du gradient et des relations de proportionnalité des longueurs et des contours. Haugeard et al. [2009] basent également leur reconnaissance de fenêtre sur des contours qu'ils modélisent dans des graphes (figure 2.5.b). Ricard et al. [2008] utilisent des histogrammes radiométriques (*CIELab*), des ondelettes de *Gabor* et des attributs de formes.

2.2 Séparation des façades

Si l'image d'entrée est supposée contenir plus d'une façade, différentes techniques ont été proposées pour les séparer et revenir au problème de la segmentation d'une seule façade.

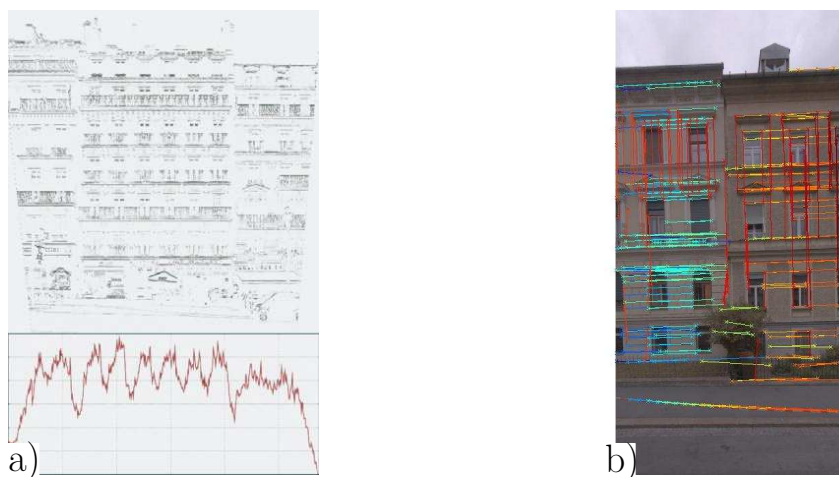


FIGURE 2.6 – Illustration de deux techniques de séparation de façades : a) Gradients de contours horizontaux et leurs projections verticales (Hernández et Marcotegui [2009] ; b) Segments délimités par deux points identiques (Wendel et al. [2010])

Une technique pour détecter les séparations dans une séquence de façades est de se baser sur l'*a priori* que les contours horizontaux sont beaucoup moins fréquents et prononcés entre deux façades qu'ailleurs. Hernández et Marcotegui [2009] proposent de projeter verticalement le gradient des contours horizontaux (figure 2.6.a). Ils localisent alors des limites de façades à partir du profil obtenu par un traitement

morphologique (*filtre séquentiel alterné*) et un algorithme de croissance de région. Ils classent les modes du profil en deux catégories (segmentation *k-mean*) : dans une façade ou entre deux façades.

Une autre technique que proposent Wendel et al. [2010] repose sur l'*a priori* que les points d'intérêt sont davantage similaires entre eux sur la surface d'une même façade qu'ils le sont entre deux façades (figure 2.6.b). Les auteurs projettent verticalement les segments de l'image reliant deux paires de points similaires. Les minima du profil obtenu ont une forte probabilité d'être situés entre deux façades. Les descripteurs des points à comparer sont constitués de profils radiométriques reliant les points entre eux. Les auteurs étudient l'influence de l'espace de couleur et du mode de détection des points dans cette approche. Recky et al. [2011] utilisent cette dernière pour segmenter des séquences de façades.

D'autres techniques sont utilisées dans la littérature. Certaines enrichissent leur traitement par un processus d'apprentissage des propriétés des façades. Par exemple, Zhao et al. [2010] comparent deux blocs de façades, pré-segmentés à partir de contours verticaux, par un algorithme de type *boosting* en prenant en compte la hauteur des bâtiments, des profils d'intensité horizontaux sur toute la largeur des blocs (distance de Kullback Leibler) ainsi que les intersections et l'intensité des contours.

2.3 Estimation des profondeurs

Il arrive que la façade comporte plusieurs éléments qui sortent du plan dominant comme les fenêtres qui sont souvent en retrait ou les balcons, au contraire, en saillie. Beaucoup d'auteurs émettent l'hypothèse que les éléments qui sortent du plan dominant et qui se ressemblent, sont situés dans un autre même plan.

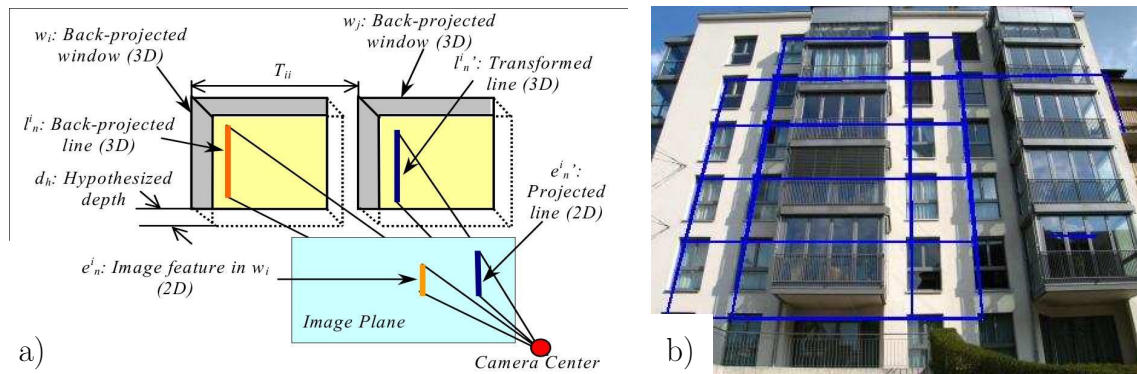


FIGURE 2.7 – Estimation des profondeurs : a) Comparaison de la projection des contours sur des plans de vitres de même profondeur (Lee et Nevatia [2004]) ; b) Groupement de points d'intérêt similaires pour détecter les régions de même profondeur (Van Gool et al. [2007]).

Van Gool et al. [2007] optimisent une carte de profondeur de l'image par la coupe minimale d'un graphe dont les nœuds sont des points d'intérêts, sous l'hypothèse que les points similaires se situent sur un même plan (figure 2.7.b).

Lee et Nevatia [2004] et Reznik et Mayer [2008] possèdent déjà la localisation des fenêtres. Ils estiment leur profondeur par un balayage de plans (figure 2.7.a).

2.4 Appariement de façades

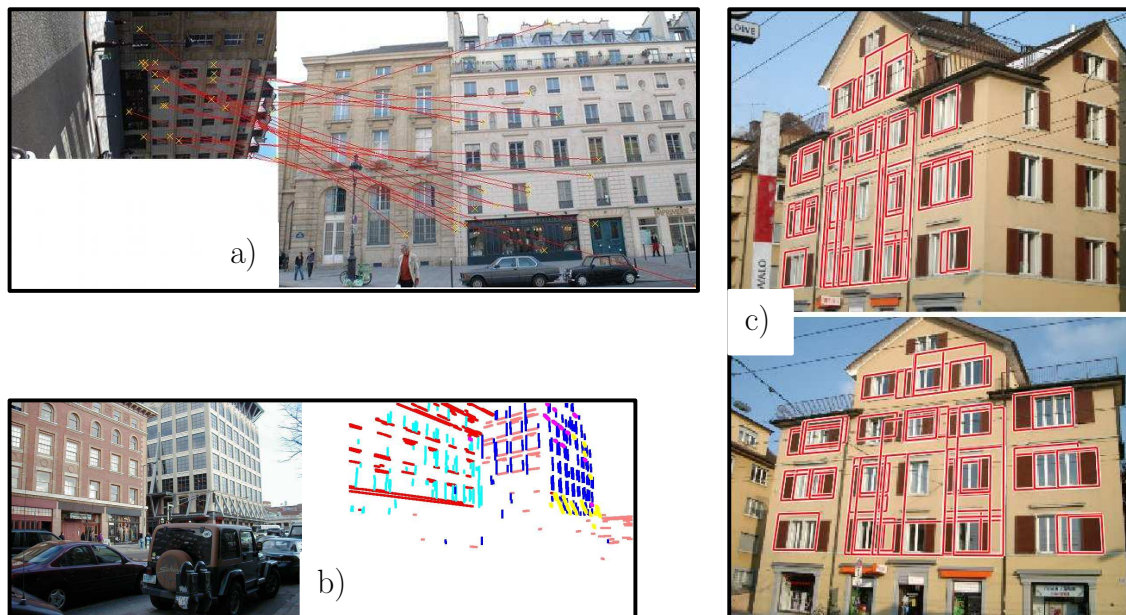


FIGURE 2.8 – Appariement de façades : a) Une image requête (à droite) est appariée par *SIFT* à l'image la plus proche (à gauche) dans une base de données de référence (Bris et Paparoditis [2010]); b) Extraction de contours d'indexation (Li et Shapiro [2002]); c) Extraction de rectangles d'indexation (Mičušík et al. [2008]).

Une méthode classique d'appariement d'images consiste à utiliser des points d'intérêt. Elle s'applique relativement bien à l'appariement de façades. Robertson et Cipolla [2004] utilisent des coins de Harris (Harris et Stephens [1988]) à cinq échelles différentes pour estimer la pose des images. Schindler et al. [2008] calculent des homographies pour rendre carrés des maillages de *SIFT* (Lowe [2004]) et utilisent une représentation canonique de l'apparence du maillage par un motif selon la théorie des *groupes symétriques tordus* de Liu et Collins [2001]. Leurs travaux s'insèrent dans une problématique de *géo-marquage*.

Bris et Paparoditis [2010] proposent d'utiliser des points *SIFT* pour appairer des images ortho-rectifiées (figure 2.8.a). Ils se servent des informations de l'échelle et de l'orientation des points appariés pour palier les problèmes de répétitions.

Li et Shapiro [2002] représentent l'image par des groupements cohérents de contours à partir de leur radiométrie, leur orientation et leur topologie (figure 2.8.b). Ils établissent alors des descripteurs constitués de mesures d'intersections des contours entre les groupes et de chevauchements des contours dans un même groupe. Zhang et Košecká [2005] utilisent les points de fuite et des contours longs. Haugeard et Philipp-Foliguet [2010] utilisent des graphes de contours et des informations de texture.

Mičušík et al. [2008] proposent un étiquetage des contours de l'image : haut, bas, gauche ou droite par un réseau aléatoire markovien (*MRF*). Les nœuds sont les contours alignés sur la direction des points de fuite et les arêtes sont obtenues par une triangulation de Delaunay. Les descripteurs de la façade sont alors calculés par une transformation en cosinus discrète des régions rectangulaires détectées.

Pour accélérer le processus d'appariement, Zhang et Košecká [2006] proposent d'utiliser l'information donnée par les points de fuite pour sélectionner un premier jeu de façades candidates. Ils calculent des descripteurs composés d'histogrammes locaux de la teinte des pixels d'orientation verticale et horizontale. Ils détectent alors la façade la plus semblable grâce à des points *SIFT*.

Zhang et al. [2010] ne se basent pas sur les informations contenues directement dans les images à comparer mais sur le résultat d'un étiquetage de leurs éléments : ils détectent des images comprenant le même agencement d'étiquettes. Il s'agit d'un *transfert supervisé d'étiquettes*. Ils génèrent des hypothèses d'appariement d'étiquettes qu'ils valident par une classification supervisée des correspondances. Ils segmentent l'image avec les contraintes spatiales liées aux étiquettes.

2.5 Segmentation générique

Lorsqu'il s'agit de traiter de grands volumes d'images de scènes de rues contenant une ou plusieurs parties de façades de différents styles d'architecture, certaines approches que nous venons de présenter ne sont plus adaptées pour un traitement automatique. Chaque image nécessiterait en effet un pré-traitement par un opérateur pour la configuration des paramètres ou le choix explicite de la grammaire à utiliser. Liu et Gagalowicz [2010], par exemple, ou Ali et al. [2007], se focalisent sur des façades d'un même style d'architecture ou liées à des styles dont les propriétés ont été *appprises* par le processus.

D'autres approches, au contraire, comportent des modules de traitement générique qui nous intéressent particulièrement. Leurs résultats pourraient guider les applications citées plus haut tout en ne comportant que le moins possible d'*a priori*. Les régions produites pourraient par exemple être exploitées lors d'une phase de sémantisation de la façade ou bien d'extraction de ses surfaces de même profondeur. Nous relevons dans cette section un petit échantillon de ces traitements génériques et nous situons nos travaux par rapport à eux.

2.5.1 Sélection de traitements génériques

De manière globale, une segmentation générique pourrait servir de base à une indexation géométrique. Plusieurs auteurs se sont orientés vers l'extraction de régions stables. Drauschke [2009] détecte les régions de la façade qui changent le moins avec l'échelle. Ils construisent une arborescence de régions en fonction du recouvrement entre deux échelles. La figure 2.9.a schématise ce procédé : l'image de gauche représente le regroupement de ces régions dans la pyramide d'images et un résultat obtenu est illustré à droite. Une segmentation de type *ligne de partage des eaux* sur la norme du gradient est calculée à chaque échelle après suppression des gradients inférieurs à la médiane. Ces travaux sont tout à fait pertinents.

D'autres auteurs retrouvent une certaine stabilité dans les structures répétitives de la façade. Il s'agit d'un *a priori* tout à fait adapté aux traitements génériques dont nous parlons. Zeng et Gool [2008] segmentent des images par l'utilisation d'un graphe dont les nœuds sont des super-pixels et dont les arêtes possèdent des informations de similarité entre eux. Leurs travaux semblent s'appliquer de manière prometteuse aux façades. Wenzel et al. [2008] trouvent des combinaisons linéaires de distances entre points *SIFT*. Ils constituent une enveloppe convexe autour des structures répétitives (figure 2.9.c). Jahangiri et Petrou [2008] comparent des régions d'intérêt susceptibles d'être caractéristiques d'un immeuble (fenêtres, portes, cheminées, balcons, lucarnes, dais). Ils recherchent les similarités par des vecteurs de translation obtenus dans une carte d'accumulations des emplacements de régions. Schaffalitzky et Zisserman [1998] utilisent des transformations affines de Leung et Malik [1996] ainsi que des transformations en perspective.

D'autre part, les lignes de découpe d'une segmentation générique devraient pouvoir faciliter un processus de séparation des façades ou de leurs niveaux et travées. Plusieurs auteurs comme Lee et Nevatia [2004] ou Müller et al. [2007], introduisent un *a priori* sur l'alignement et l'orthogonalité des structures de la façade pour orienter leurs travaux vers cette application. Nous avons vu qu'ils ont recours à des projections orthogonales d'informations locales sur des profils d'accumulation comme décrit plus haut. L'utilisation de l'*a priori* de l'alignement des structures est tout à fait pertinent dans le cadre d'une segmentation générique. La plupart des façades, en effet, possèdent des structures empilées les unes au-dessus des autres et côte à côte parce qu'il s'agit d'un mode de construction très simple et solide des bâtiments.

Dans Müller et al. [2007], les tuiles formées par l'intersection des niveaux et travées sont segmentées de manière hiérarchique par éliminations successives des petits contours périphériques dont le prolongement croisent d'autres contours (figure 2.9.b). Ce procédé se répète de la même manière et indépendamment à chaque itération. Il est utilisé pour extraire finement les fenêtres, ce qui s'applique très bien au type des façades présentées. Il pourrait également être utilisé, d'une manière plus globale et cette fois-ci sur n'importe quel style d'architecture (par exemple avec balcons ou corniches), sur des boîtes englobantes plus grandes à l'échelle de la façade.

Dans le même esprit, d'autres auteurs, comme Alegre et Dellaert [2004] ou Han et Zhu [2005], perfectionnent l'arborescence de la segmentation avec des règles de grammaires génériques. Alegre et Dellaert [2004] utilisent uniquement des règles de compositions de découpes verticales et horizontales pour segmenter la façade en région de radiométrie uniforme. Bien que le critère de modélisation des régions soit très simple, les résultats qu'ils obtiennent sont très cohérents sur les façades qu'ils étudient. Han et Zhu [2005] construisent un graphe mis à jour itérativement à partir d'une grammaire comprenant quatre règles d'agencement de surfaces en 3D pour fournir la description la plus compacte d'une esquisse de la scène (figure 2.9.d). Ce procédé est également générique et leurs résultats sont très satisfaisants.

2.5.2 Notre contribution

Les outils que nous proposons visent à explorer le plus loin possible le domaine de la segmentation générique de façades. Ils se basent sur les *a priori* d'orthogonalité et de répétitivité des structures. Nous proposons deux approches.

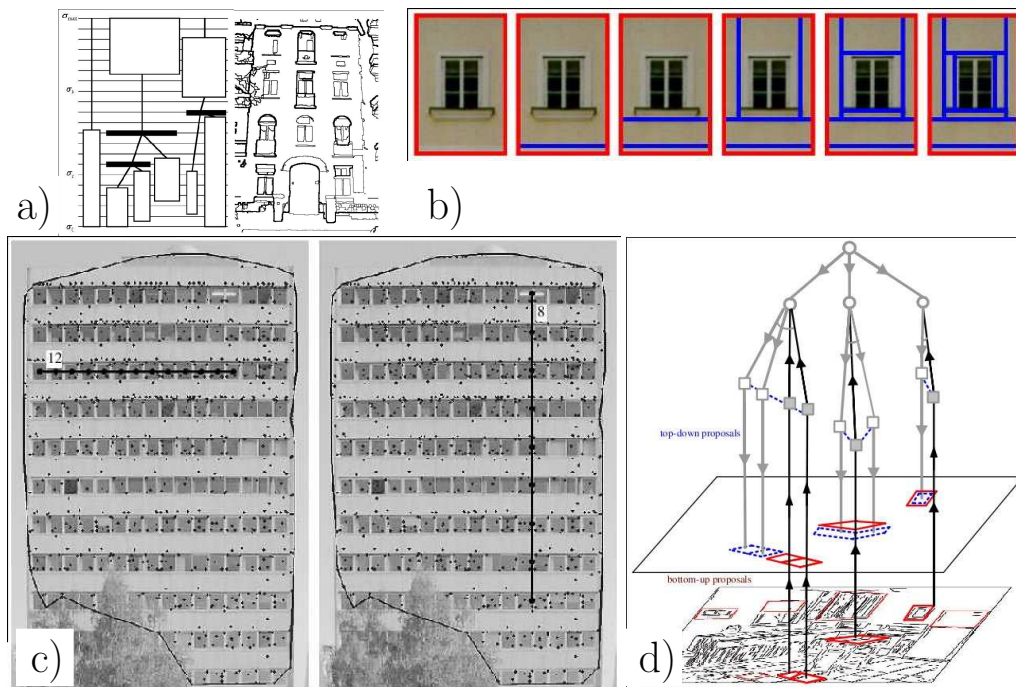


FIGURE 2.9 – Approches génériques : a) Arborescence de régions stables multi-échelle (Drauschke [2009]); b) Segmentation hiérarchique basée sur les contours (Müller et al. [2007]); c) Détection de structures répétitives (Wenzel et al. [2008]); d) Regroupement de surfaces 3D suivant des règles de grammaire (Han et Zhu [2005]).

La première (Burochin et al. [2010]) se base sur une projection de gradients dans le même esprit que Lee et Nevatia [2004] pour découper l'image le long des gradients forts. Il s'agit d'une segmentation hiérarchique de la même manière que le module de segmentation des tuiles de Müller et al. [2007]. Dans notre approche, la région initiale n'est pas une tuile mais l'image entière. Elle est coupée en sous-régions (deux dans notre cas) et le processus se répète de la même manière et indépendamment dans chacune des sous-régions. L'objectif est de localiser des modèles élémentaires dans la façade. En cela, il rejoint la segmentation en régions uniformes de Alegre et Dellaert [2004] : nous n'utilisons pas de grammaire mais des heuristiques de découpe et différents modèles élémentaires. Un de ces modèles élémentaires tient compte des répétitions dans la façade et consiste en un calcul plus simplifié que ce que proposent Wenzel et al. [2008].

La deuxième approche vise à extraire les structures répétitives de la façade. Contrairement aux autres auteurs, nous ne cherchons pas de topologies figées. Nous recherchons de petits ensembles de structures répétitives aléatoirement sur toute la surface de l'image. Ces petits ensembles doivent se superposer et se compléter entre eux pour faire ressortir un plus gros ensemble, la meilleure configuration des répétitions de l'image.

Nous avons encore laissé de côté beaucoup de propriétés intrinsèques des images de façades comme l'utilisation des histogrammes de normes et d'orientation de gradients proposée par exemple dans Korč et Förstner [2008a] ou bien comme la détection de régions stables avec l'échelle du type de Drauschke [2009]. Néanmoins, les

résultats que nous obtenons sont prometteurs.

Segmentation hiérarchique en modèle 2D élémentaires

Ce chapitre présente une segmentation hiérarchique d'une image de façades décrite dans Burochin et al. [2010]. Sa problématique se situe en amont de toutes les autres approches sémantiques de l'état de l'art pour traiter une plus large palette de styles d'architecture. La solution proposée consiste en la construction d'une topologie des structures principales de la façade, notamment celles qui se répètent. Cette topologie doit être robuste par rapport au point de vue et par rapport à l'échelle. Elle doit permettre, au moins, d'isoler la façade de son environnement. Une des applications de cette segmentation peut être par exemple, comme nous le verrons, d'ajouter du réalisme aux modèles 3D des villes.

Le processus repose sur les deux hypothèses suivantes :

- Les principales structures de la façade sont rectangulaires.
- Leurs contours sont alignés horizontalement et verticalement.

L'algorithme de segmentation est schématisé par l'arbre de décision de la figure 3.1. Il s'agit de rechercher des modèles radiométriques et géométriques en deux dimensions de manière dichotomique. L'algorithme prend l'image entière en entrée et tente l'appariement à un des trois modèles utilisés. Si elle ne correspond à aucun d'eux, elle est coupée en deux régions. La même démarche est effectuée récursivement et indépendamment dans les deux régions. Le processus s'arrête lorsqu'il détecte un modèle ou que la région considérée est trop petite.

La découpe s'effectue le long des alignements principaux des contrastes des structures et se base sur une analyse de la distribution des contours. Les modèles sont liés à la radiométrie et à la géométrie des éléments de la façade. Un modèle planaire suit une radiométrie uniforme. Un modèle cylindrique suit une radiométrie en bandes verticales ou horizontales. Un modèle périodique rend compte de l'existence d'une périodicité des structures.

Une propriété géométrique importante des façades intervient dans la détection du modèle périodique et dans le processus de découpe de l'image. Il s'agit des *alignements dominants*, détaillés en section 3.1. La modélisation des structures de la façade est ensuite présentée à travers les modèles plans, cylindriques et périodiques (section 3.2). Le processus de découpe est expliqué en section 3.3 : il tend à séparer les régions de l'image en fonction de la distribution de leurs contours. Enfin

une conclusion synthétise les atouts et les limites de la segmentation, les perspectives d'approfondissement ainsi que quelques applications possibles (section 3.4).

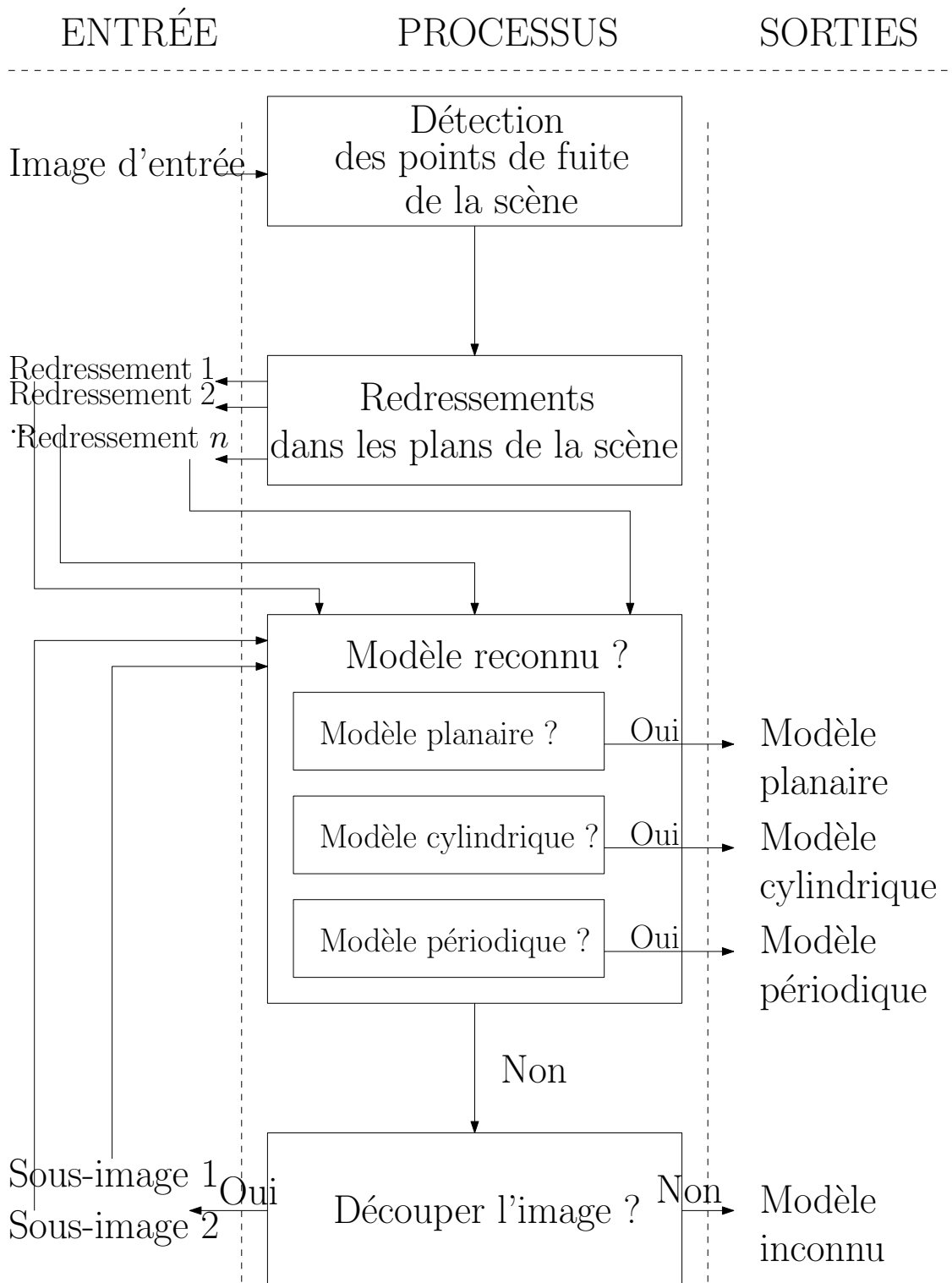


FIGURE 3.1 – Segmentation hiérarchique : l'algorithme confronte récursivement les données aux modèles. Les régions qui ne sont associées à aucun modèle proposé sont découpées en deux sous régions.

3.1 Alignements dominants

Les *alignements dominants* (définition 3.1) sont une propriété très importante des façades. Ils rendent compte de la répétitivité des structures alignées horizontalement et verticalement. Nous verrons qu'ils sont utilisés dans le processus de découpe et dans la détection des régions de l'image qui suivent un modèle périodique.

3.1.1 Mode de calcul

Nous avons vu dans la sous-section 2.1.1.2 qu'il existait dans la littérature des techniques permettant d'extraire les informations géométriques globales d'une image en s'affranchissant de l'influence du bruit local d'acquisition et des petites structures parasites isolées. Sur des images ortho-rectifiées, cette extraction est réalisable par des projections verticales et horizontales d'informations locales comme les gradients ou les contours. Le fait d'accumuler ces informations le long de toute la largeur ou la hauteur de l'image renforce les contrastes qui ne sont pas nécessairement les plus forts mais qui se répètent.

Dans nos approches de segmentation, nous avons recours à ce type d'information en utilisant les *alignements dominants* décrits par la définition 3.1. Ils fournissent la localisation fine d'alignements de contrastes plus ou moins marqués des structures de la façade.

Définition 3.1 Soit $V_H(G)$ (respectivement $H_V(G)$) le profil vertical de projection horizontale (respectivement le profil horizontal de projection verticale) d'un gradient G des contours horizontaux (respectivement des contours verticaux), selon la terminologie de Lee et Nevatia [2004] utilisée pour les contours de l'image.

Les alignements dominants horizontaux (respectivement verticaux) de l'image sont les droites horizontales qui coupent le profil $V_H(G)$ (respectivement $H_V(G)$) au niveau de ses maxima locaux.

Les valeurs que nous accumulons sont les valeurs absolues d'un gradient G obtenu par un opérateur discret du premier ordre. Une fenêtre glissante de taille fixe le long du profil permet de localiser les maxima locaux. La taille de cette fenêtre reste constante tout au long du processus et est associée à la longueur des structures recherchées dans la façade. Si l'objectif est de limiter le processus à l'encadrement des fenêtres, cette longueur doit être de l'ordre du mètre. Il faut la réduire si l'objectif est de poursuivre jusqu'à de plus petits éléments. Tous les maxima n'ont pas besoin d'être retenus.

3.1.2 Propriétés intéressantes

Le quadrillage formé par l'intersection des alignements dominants horizontaux et verticaux est intéressant. Chaque cellule a une probabilité d'encadrer une structure qui suit un modèle que nous recherchons. La figure (3.2) schématise ce quadrillage.

L'intersection de ces alignements forme un ensemble de points d'intérêt utilisé dans la détection des régions suivant un modèle périodique. Ces points d'intérêt présentent l'avantage d'un calcul très rapide et d'une répartition régulière.

D'autre part, les segments du quadrillage ont une probabilité d'être des contours principaux. Le processus de découpe s'appuie sur cette considération.

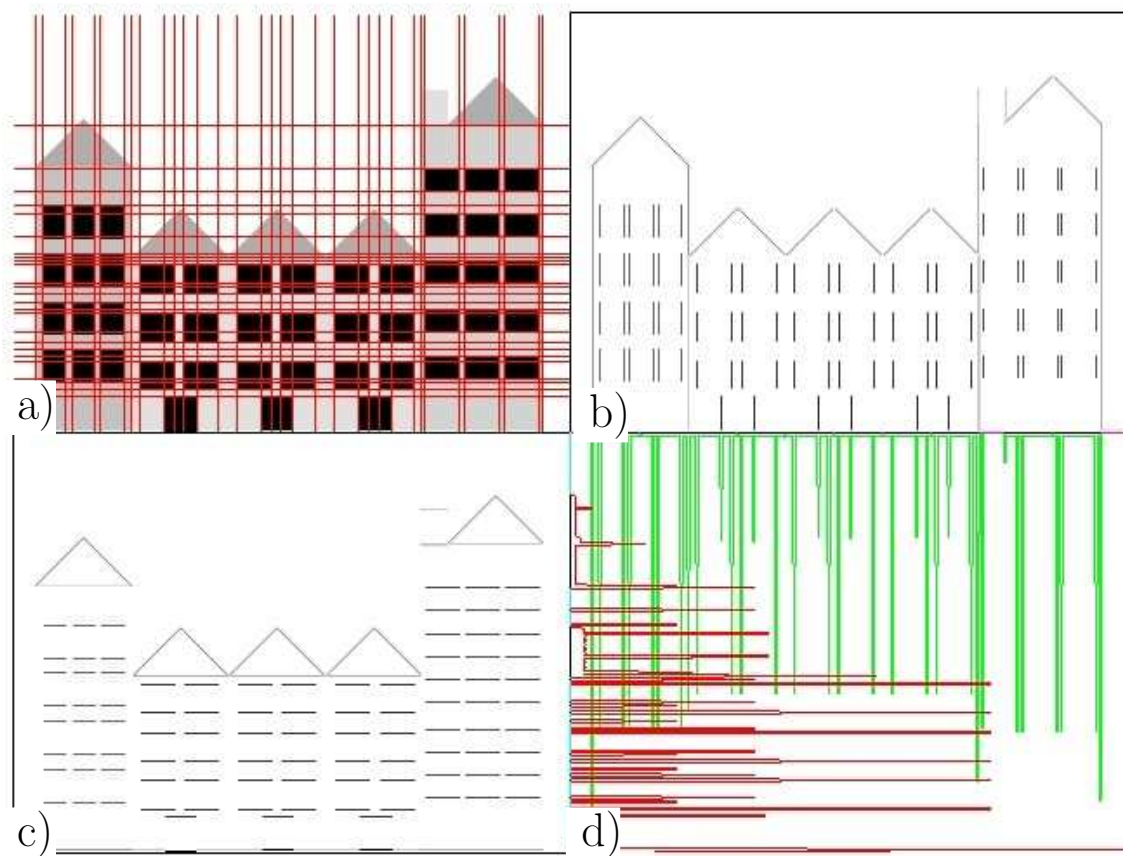


FIGURE 3.2 – Calcul des alignements dominants. a) Quadrillage de la façade formé par les alignements dominants. b) Valeur absolue des gradients associés aux contours verticaux ; c) Valeur absolue des gradients associés aux contours horizontaux ; d) Profils d'accumulation des gradients : H_V en vert et V_H en rouge.

3.1.3 Discussion sur la détection et la localisation

Lors de nos tests nous n'avons pas cherché à purger les alignements dominants qui ne contiendraient que des contours de très fort contraste mais en nombre non significatifs. Mise à part un certain ralentissement, cette sur-détection ne gêne pas la suite du traitement.

D'autre part, il peut arriver que les contours des grandes structures ne soient pas strictement horizontaux ou verticaux dans l'image redressée. Cela peut s'expliquer par le processus d'extraction des points de fuite perturbé par le bruit. Pour compenser ce phénomène, le profil d'accumulation est lissé par un filtre gaussien. Ce procédé améliore la localisation des extrema locaux mais introduit une légère délocalisation. Dans la suite du processus, chaque alignement est alors associé à un voisinage d'incertitude.

3.2 Reconnaissance de modèles

L'objectif est de modéliser les structures géométriques par des rectangles dont la radiométrie suit un modèle particulier.

L'image est comparée à trois types de modèles. Mais certaines régions pourraient accepter plusieurs modèles. Notre choix est de sélectionner le plus simple (le moins de degrés de liberté). Il s'agit d'un compromis entre la qualité de la représentation et le taux de compression. L'image est d'abord comparée au modèle planaire. Si elle en est suffisamment proche, le processus s'arrête. Sinon, seulement après, elle est comparée au modèle cylindrique, puis enfin éventuellement au modèle périodique. La distance entre les données et le modèle consiste en un dénombrement des différences radiométriques locales.

Le processus de segmentation s'arrête lorsque la région correspond à un des modèles. Si ce n'est pas le cas, il partage la région en deux sous-régions et leur applique ce même traitement (*cf* section 3.3).

Soit I_k la sous-image correspondant à la région R_k d'une image de façade I . La sous-image I_k est décrite par le modèle $\mathcal{M}$ lorsque son écart $N_{\mathcal{M}}(I_k)$ est suffisamment petit et que le modèle $\mathcal{M}$ est le plus simple possible. L'écart $N_{\mathcal{M}}(I_k)$ se définit par le nombre de pixels dont la radiométrie diffère trop du modèle.

3.2.1 Modèles planaires et cylindriques

Nous définissons d'abord deux types de modèles pour décrire les régions de distributions radiométriques uniformes (uniforme dans toute la région pour le modèle planaire ou par bandes pour le modèle cylindrique) : définitions 3.2, 3.3 et 3.4.

Définition 3.2 *Le modèle planaire suppose que la radiométrie de la sous-image I_k au pixel p suive une radiométrie uniforme λ .*

$$I_k(p) = \lambda + \epsilon(p)$$

où $\epsilon(p)$ est le bruit (petits détails locaux, occlusions éparses ou bruit gaussien).

Définition 3.3 *Le modèle cylindrique en colonnes suppose que l'intensité de chacune des colonnes de pixels de la région suive une distribution uniforme $\lambda(x)$ en colonnes.*

$$I_k(x, y) = \lambda(x) + \epsilon(x, y)$$

où $\epsilon(x, y)$ est le bruit (détails locaux, occlusions éparses).

Définition 3.4 *Le modèle cylindrique en lignes suppose que l'intensité de chaque ligne de pixels suit une distribution uniforme $\lambda(y)$ en lignes.*

$$I_k(x, y) = \lambda(y) + \epsilon(x, y)$$

où $\epsilon(x, y)$ est le bruit (détails locaux, occlusions éparses).

La figure 3.3 présente des exemples de ces modèles.

Nous apparions ces modèles en fixant leurs radiométries aux valeurs médianes des régions concernées dans l'image, sous l'hypothèse que ce sont elles les plus représentées. Cette modélisation dépend de deux seuils : la distance radiométrique maximale au modèle et le nombre maximal de pixels suffisamment proches de ce modèle. Ces deux seuils sont fixés empiriquement en pré-traitement.

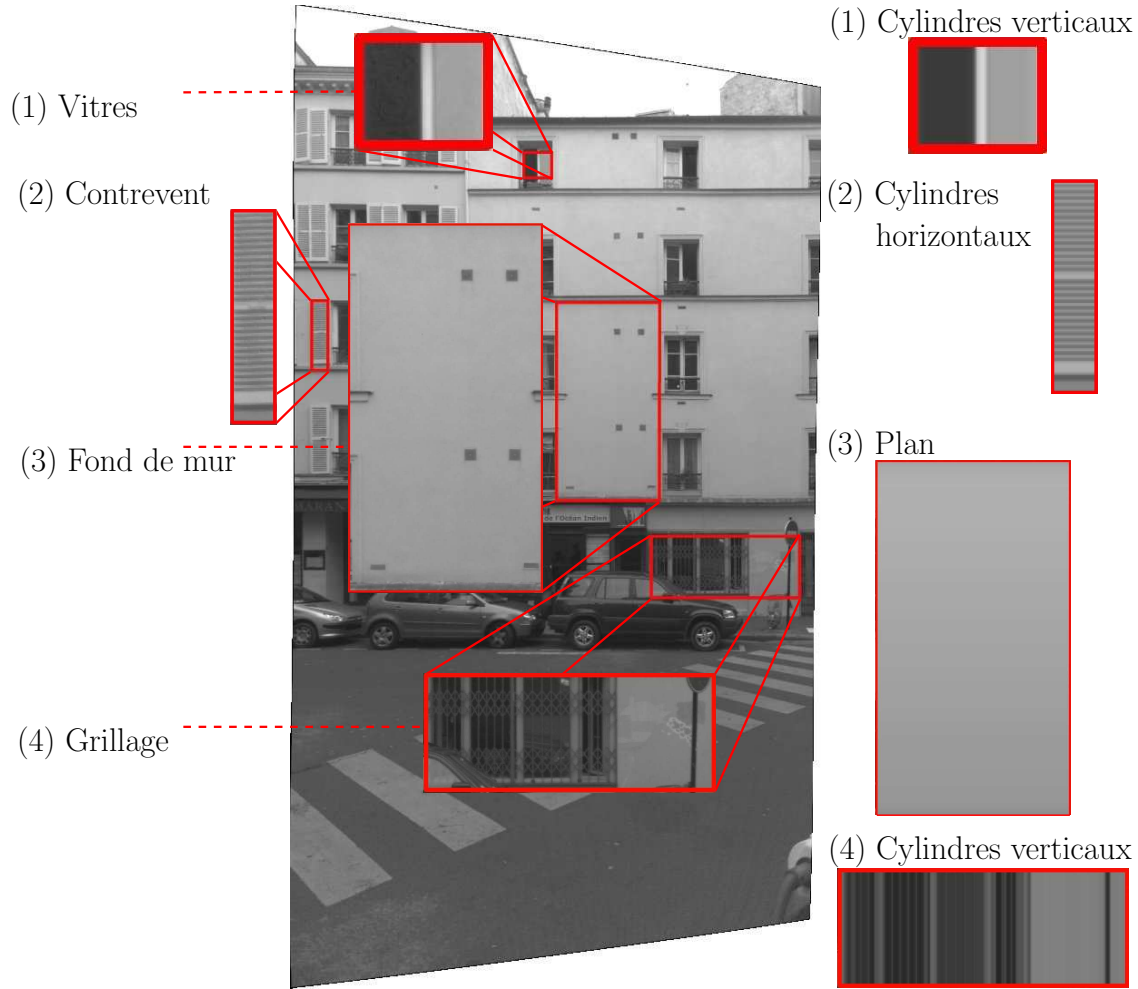


FIGURE 3.3 – Exemple de modèles radiométriques se retrouvant sur une façade (modèles synthétiques présentés à droite) : modèle plan au niveau du fond de mur des deux premiers étages ; cylindres verticaux sur une fenêtre et un grillage ; cylindres horizontaux sur les persiennes.

3.2.2 Modèle périodique

Nous proposons par la définition 3.5 un *modèle périodique* pour décrire les régions à forte répétitivité. Contrairement aux modèles planaire et cylindrique, il ne prend pas en compte l'intégralité de la région, qui peut être perturbée par des effets de perspective et certaines instabilités structurelles de la façade (ouverture des vitres et contrevents, objets sur le rebord, etc.). Il se restreint à la radiométrie du voisinage de certains pixels d'intérêt. Nous choisissons ces pixels d'intérêt parmi les intersections, Π , des alignements dominants (définition 3.1), qui offrent l'avantage d'une distribution homogène et régulière dans la région I_k .

Définition 3.5 *Considérant un ensemble, Π , de points d'intérêt régulièrement répartis, le modèle périodique suppose que les points de la région qui sont séparés d'une même distance $||\vec{t}||$ se trouvent sur des zones radiométriquement semblables.*

$$\exists \vec{t} : \forall i, \left[\exists j : \Pi[i] \simeq \Pi[j] + \vec{t} \right] \Rightarrow \left[I_k(\Pi[i]) \simeq I_k(\Pi[j]) \right]$$

Les paragraphes qui suivent décrivent la génération d'hypothèses de périodes $\vec{t}$ ainsi que la sélection de celle qui maximise la proportion de bonnes corrélations.

3.2.2.1 Génération des hypothèses de période

Les hypothèses de période sont obtenues par les coordonnées des alignements dominants dont la distribution est régie par la répétition des structures principales de la façade. Les hypothèses horizontales et verticales sont traitées indépendamment. Pour les hypothèses verticales, nous accumulons toutes les distances entre les alignements dominants horizontaux pour établir un histogramme de distances. La figure 3.4 montre un exemple d'histogramme de la région analysée sur la figure 3.5. L'histogramme des distances horizontales est calculé de la même manière.

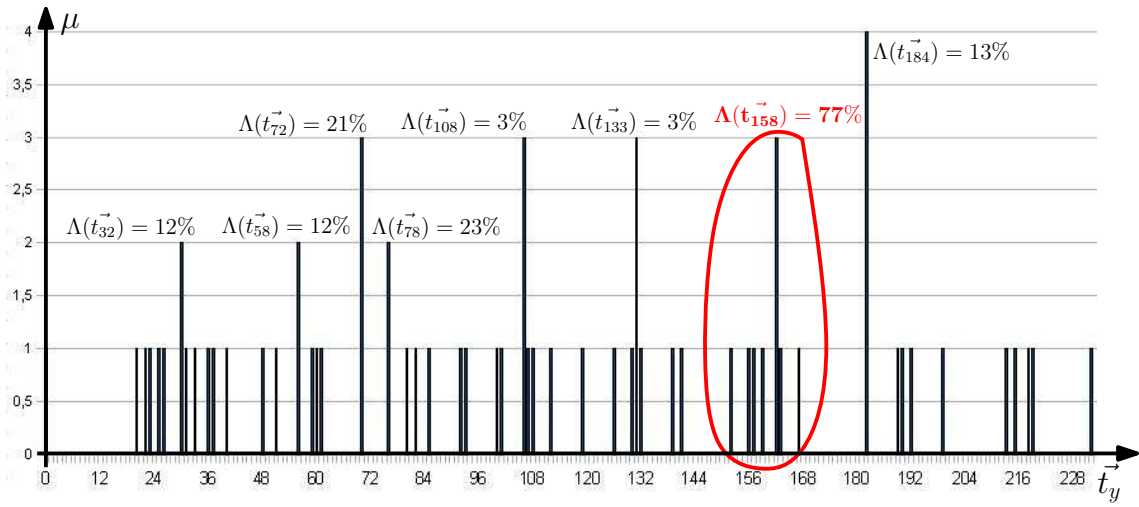


FIGURE 3.4 – Histogramme de distances entre les alignements dominants horizontaux de la région analysée sur la figure 3.5. Chaque mode est relatif à la taille d'une structure répétitive : la proportion des appariements est indiquée pour chacun. Le meilleur est entouré en rouge (hauteur des étages).

Chaque mode de ces histogrammes est relatif à la taille fixe d'une structure qui se répète dans la façade. Les modes sont détectés comme maxima dans un voisinage fixé après un filtrage gaussien.

3.2.2.2 Sélection de la meilleure période

Nous devons à présent sélectionner la meilleure hypothèse de période : celle associée à un nombre significatif de paires de points d'intérêt bien corrélés. Nous utilisons alors la mesure de corrélation de *Pearson*. Les descripteurs radiométriques sont suffisants pour appairer les points similaires parce qu'ils sont comparés dans une même image qui est rectifiée. Pour une hypothèse $\vec{t}_y$ (respectivement $\vec{t}_x$), nous calculons des corrélations entre toutes les paires de points $(\Pi[i], \Pi[j])$ telles que $\Pi[j] \simeq \Pi[i] + \vec{t}_y$ (respectivement $\Pi[j] \simeq \Pi[i] + \vec{t}_x$). Nous sélectionnons l'hypothèse $\vec{t}_y^*$ (respectivement $\vec{t}_x^*$) dont la proportion d'appariements Λ^* est la meilleure dans la région. Si $\Lambda^*(\vec{t})$ est suffisamment élevé pour la période $\vec{t}_y^*$ ou la période $\vec{t}_x^*$, alors l'image I_k est considérée comme suivant un modèle périodique dont les dimensions du noyau sont données

par $(\|\vec{t}_x\|, \|\vec{t}_y\|)$. Si une seule dimension X ou Y est détectée, alors nous utilisons respectivement la hauteur ou la largeur de la région pour construire le noyau. Si aucune bonne proportion d'appariements n'est détectée, la région n'est pas considérée comme suivant un modèle périodique.

La figure 3.5 présente un ensemble d'étages similaires dont les fenêtres ne sont pas disposées régulièrement dans la région. La proportion $\Lambda(\vec{t}_y^*)$ est suffisante pour considérer la distance $\|\vec{t}_y^*\|$ - la hauteur des étages - comme une période tandis que la proportion $\|\vec{t}_x^*\|$ est trop faible. La région est alors supposée suivre un modèle périodique dont le noyau est l'étage.

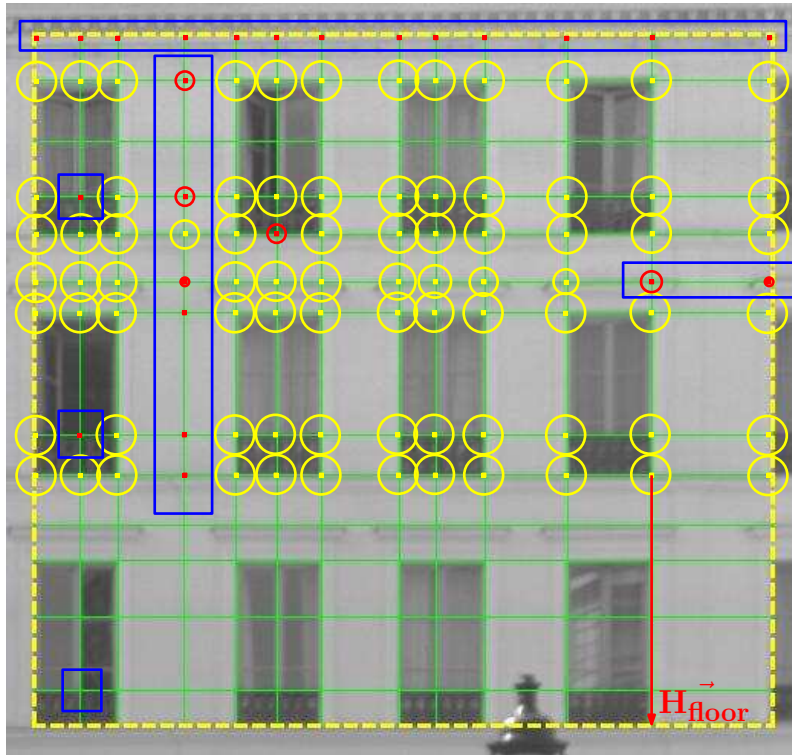


FIGURE 3.5 – Appariements associés à la période $\vec{t}_y^* = H_{\text{etage}}^{\rightarrow}$. Les alignements dominants sont représentés en vert. La longueur du rayon du cercle autour du point $\Pi[i]$ est relatif au score de corrélation $(\Pi[i] * \Pi[j])$, où $\Pi[j] = \Pi[i] + \vec{t}_y^*$. Un score de corrélation négatif est représenté par un point. Les bons scores de corrélation (plus grands que 80%) sont encadrés en jaune ; les mauvais scores sont encadrés en rouge. Sur 104 paires significatives de points, 80 appariements sont corrects. Les rectangles bleus sont des cas particuliers commentés dans le texte. b) Tentative de pénalisation de la découpe par des informations de périodicité (profil en jaune).

3.2.2.3 Discussion sur le modèle périodique

Le modèle périodique que nous utilisons permet d'obtenir des résultats satisfaisants. Pour atténuer l'impact des occlusions, nous ne prenons en compte que de petites parties de la façade que nous considérons assez fiables et nous supposons leur nombre suffisant pour compenser les perturbations. Ces petites parties suffisamment fiables sont les voisinages des points d'intérêt décrits précédemment sous l'hypothèse

d'une répartition uniforme des alignements dominants. Le reste de la façade n'est pas pris en compte. Ainsi les objets mobiles, comme les vitres ou les pots de fleurs peuvent différer d'une fenêtre à l'autre : s'ils ne sont pas situés sur un point d'intérêt, ils ne perturbent pas le processus. Nous supposons leur distribution chaotique sur l'ensemble de la région.

De plus, le rejet des alignements dominants non périodiques permet de réduire l'influence des perturbations liées aux objets situés en arrière plan et qui sont délocalisés par la perspective. La figure 3.6.a présente une façade composée d'un quadrillage de fenêtres. Le processus détecte un modèle périodique directement sur l'image entière malgré les effets de perspective, les reflets et néons visibles derrière les vitres.

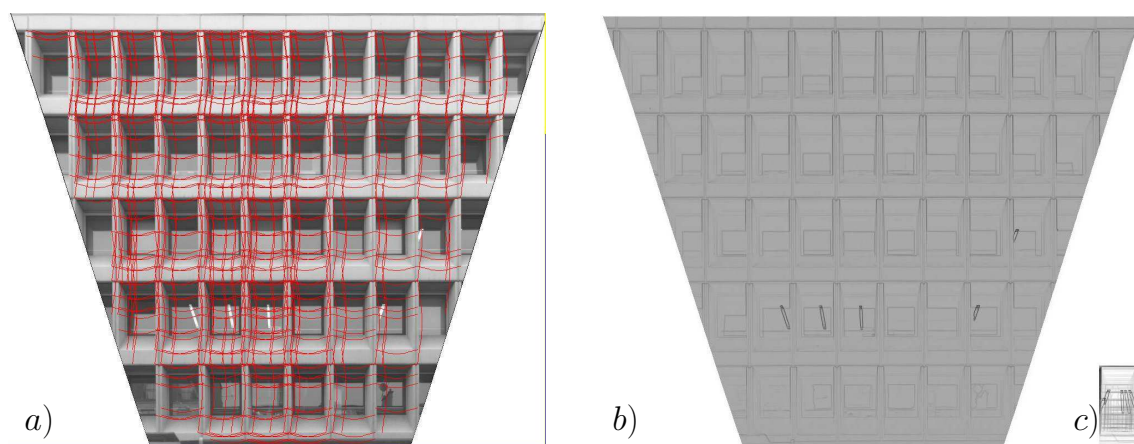


FIGURE 3.6 – a) Modèle périodique malgré les effets de perspective ; b) Norme du gradient ; c) Carte moyenne des différences entre le noyau de périodicité moyen et la norme du gradient.

Initialement, dans le même esprit que les modèles planaires et cylindriques, nous avons commencé par travailler sur la totalité de la région concernée et non sur des points d'intérêt. Mais nous nous sommes heurtés aux effets de perspectives difficilement résolubles en deux dimensions. Nous voulions construire un modèle formant une partition de l'image avec un même noyau constant se répétant périodiquement par translation. Ce noyau constant était estimé par la moyenne de toutes les sous-régions du quadrillage ainsi formé. Il nous suffisait alors de dénombrer les différences entre l'image et le partitionnement de ce noyau constant pour obtenir la distance avec le modèle périodique. La figure 3.6.c montre un résultat d'une de ces tentatives. Le gabarit du noyau est ici celui d'une fenêtre et l'information utilisée dans l'image est la norme du gradient (plus robuste aux effets d'éclairage). Il s'agit de la carte moyenne des différences entre ce noyau moyen et chacune des sous-régions. Dans l'idéal, les valeurs de cette carte de différences devrait être nulles et cela aurait validé la présence du modèle périodique. Mais les effets de perspectives et les structures locales isolées (les néons, par exemple) sont tellement marqués, que ce procédé ne fonctionne pas. Nous n'avons pas trouvé de représentation dense de l'image avec la radiométrie ou le gradient sur laquelle dénombrer de manière pertinente des différences avec un noyau périodique. C'est ainsi que nous avons eu recours à des points d'intérêt.

Cependant, les points d'intérêt présentent également des faiblesses sur les façades

qui possèdent des effets de perspective marqués. Nous avons recensé trois situations critiques.

Coins en perspective Les points situés sur des coins d’objets en perspective (figure 3.7) sont dissemblables entre eux. Pour résoudre ce problème, nous pourrions partager les voisinages de points en deux régions : l’intérieur et l’extérieur de l’objet périodique. L’appariement des points pourrait aussi se calculer à partir d’autres descripteurs : histogramme d’orientation du gradient par exemple. La modélisation pourrait alors comprendre, en plus, la localisation de l’objet périodique supposé en perspective.

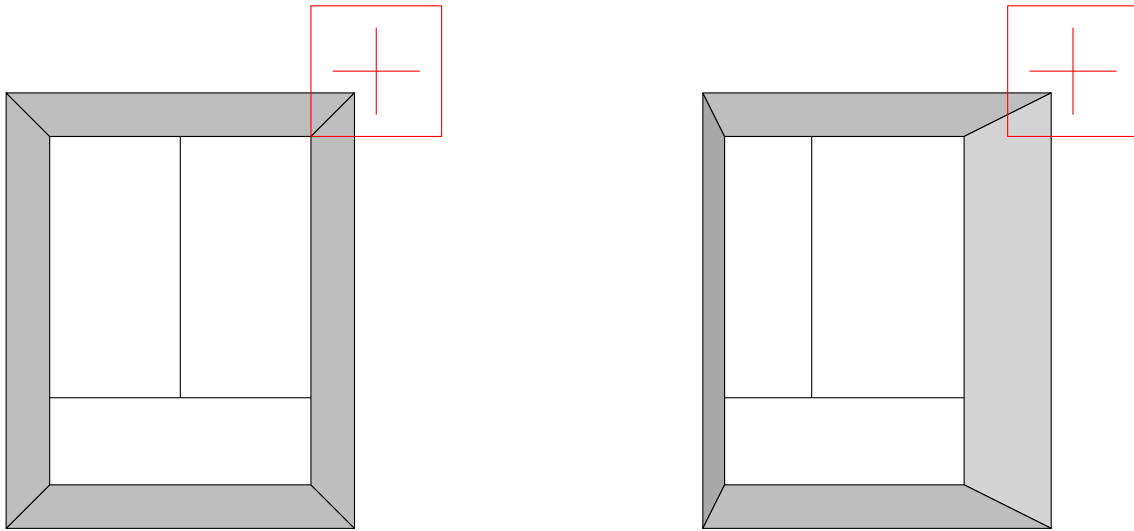


FIGURE 3.7 – Effets de perspective sur le voisinage de points d’intérêts situés au coin supérieur droit de deux fenêtres identiques.

Zones de radiométrie uniforme Les corrélations des points dont le voisinage possède une radiométrie uniforme sont très mauvaises. Le rectangle bleu vertical sur la figure 3.5 illustre cette situation. Il faudrait écarter ces points.

Points délocalisés par le redressement Le processus de redressement occasionne parfois des problèmes d’alignements dans l’image : certains points sont alors délocalisés. La figure 3.5 présente par exemple deux points $\Pi[i_1]$ et $\Pi[i_2]$ dans le rectangle bleu à droite qui ne sont pas appariés avec les points correspondants $(\Pi[i_1] + \vec{t}_y^*)$ et $(\Pi[i_2] + \vec{t}_y^*)$. Pour résoudre ce problème, nous pourrions procéder à un ajustement à d’autres types de points d’intérêt de détection plus locale. Des coins de Harris (Harris et Stephens [1988]), par exemple, pourraient jouer ce rôle. Leur sensibilité aux ombres poseraient néanmoins d’autres problèmes. Mais ils pourraient au moins permettre d’écarter les zones de radiométrie uniforme du problème précédent. L’utilisation de *SIFT* (Lowe [2004]), également, pourrait être intéressante même si l’irrégularité de leur répartition spatiale nécessiterait une adaptation conséquente de notre algorithme.

Nous avons vu que d'autres approches pour modéliser les répétitions dans les façades ont déjà été proposées dans la littérature. Wenzel et al. [2008] localisent justement des régions périodiques directement dans l'espace des *SIFT* de l'image. Ils modélisent la distribution des *SIFT* par des combinaisons linéaires de périodes dans les directions les plus fréquentes à partir d'un critère de moindre longueur de description (*MDL*). Ils détectent ainsi des blocs d'objets répétitifs comme des empilements d'étages ou des travées. La forme de ces blocs est une enveloppe convexe autour des *SIFT*. L'approche est plus complète parce qu'elle hiérarchise les périodes : par exemple une répétition de paires de travées où les paires de travées sont séparées d'une distance d_1 et où les travées d'une même paire sont séparées d'une distance $d_2 \neq d_1$. Wendel et al. [2010] étudient l'appariement de profils radiométriques entre points d'intérêt.

De notre côté, nous ne cherchons aucune hiérarchie dans les périodes. Nous sélectionnons uniquement les plus fréquentes dans les directions horizontales et verticales. Nous considérons cette restriction suffisante puisque le processus est récursif. Nous n'utilisons par les mêmes points d'intérêts non plus : les nôtres sont formés par les intersections des alignements dominants.

Les périodes que nous détectons diffèrent de la *région irréductible* proposée par Müller et al. [2007] car elles ne prétendent pas représenter la longueur du noyau élémentaire mais elles peuvent elles-même se composer de sous-périodes.

3.3 Processus de découpe dichotomique

Les régions qui ne correspondent à aucun modèle sont découpées horizontalement ou verticalement le long d'un des alignements dominants horizontaux ou verticaux (définition 3.1). Le processus sélectionne l'hypothèse de découpe qui maximise la distance entre les deux sous-régions générées de part et d'autre de cette découpe. Nous utilisons une distance locale ainsi qu'une distance globale, toutes les deux relatives à la répartition des arêtes de contours sur l'ensemble de la région concernée. La présente section définit l'espace de ces arêtes et donne la formalisation de chacune des distances utilisées.

3.3.1 Espace des arêtes

Le quadrillage des alignements dominants forme un espace d'arêtes dont la distribution est très intéressante pour localiser la découpe optimale. Cet espace est construit à partir de la définition 3.6.

Définition 3.6 Soit $(x_j)_{j \in [1, n]}$, l'ensemble des alignements dominants verticaux (définition 3.1) tel que $x_{n+1} - x_0$ soit la largeur de la région considérée.

Soit $(y_i)_{i \in [1, m]}$, l'ensemble des alignements dominants horizontaux (définition 3.1) tel que $y_{m+1} - y_0$ soit la hauteur de cette même région.

Nous appelons $h_{i,j} = \left[\begin{pmatrix} x_j \\ y_i \end{pmatrix}; \begin{pmatrix} x_{j+1} \\ y_i \end{pmatrix} \right]$ l'arête située le long de l'alignement dominant y_i entre les alignements x_j et x_{j+1} . Notons $\overline{h_j}$ sa longueur.

Soit c_h la distribution binaire dans $\{0, 1\}$ des arêtes de contour le long des alignements dominants horizontaux : $c_h(i, j)$ vaut 1 lorsque l'arête $h_{i,j}$ est situé sur un

contour de l'image, 0 autrement. Une arête qui n'est pas située sur un contour est appelée arête de fond.

Les arêtes verticales $h_{i,j}$ et leur distribution c_v se définissent de manière analogue. Nous notons également c , la distribution des arêtes de contour le long de l'ensemble des alignements horizontaux et verticaux.

Ainsi, les données sont-elles réduites à une matrice de valeurs binaires (arête de contour ou de fond). La distance entre les deux sous-régions engendrées par la découpe se calcule dans cet espace : la complexité de l'algorithme reste faible et ne dépend que du nombre d'alignements dominants. La classification des arêtes de contour et de fond est détaillée dans la suite de la section.

3.3.2 Classification des arêtes

Dire si une arête du quadrillage formé par les alignements dominants est située ou non sur un contour de l'image est un problème délicat qui détermine la cohérence de toute la suite du processus. Il s'agit d'estimer la distribution binaire c qui est à la base du calcul de la distance entre les deux sous-régions à séparer dans l'image.

Il est possible de réaliser cette classification binaire (arête de contour ou arête de fond), de manière satisfaisante, par un simple algorithme de seuillage. Chaque arête est associée à un poids qui dépend des informations géométriques contenues dans son voisinage. Les arêtes de contour peuvent alors se définir par une valeur de ce poids supérieure à un certain seuil. Les autres sont les arêtes de fond. Le seuil concerné peut être recalculé à chaque itération pour dépendre de la localisation et de l'échelle de la région analysée.

Actuellement, c'est cette stratégie que nous utilisons. Elle permet d'obtenir une classification satisfaisante pour notre application, aux conditions que le calcul du poids des arêtes soit suffisamment représentatif des contours réels de l'image et que le seuil soit choisi de manière adéquate. Nous simplifions le problème en posant l'hypothèse de l'équi-probabilité qu'une arête se trouve sur un contour. Nous fixons alors le seuil de la classification à la médiane des poids. Pour la définition de ces mêmes poids, nous utilisons la médiane du gradient se trouvant dans le voisinage de l'arête.

Si les arêtes sont suffisamment nombreuses, la redondance d'informations permet de compenser les fausses détections. Nous verrons, à la fin du chapitre, qu'une des perspectives de gain de robustesse de la segmentation, consisterait en une définition plus fine du poids des arêtes (à partir d'autres informations que le seul gradient) ainsi qu'en une analyse plus sophistiquée de leur distribution pour déterminer le seuil (avec les modes de l'histogramme des poids, par exemple).

3.3.3 Définition de la distance locale

La distance locale δ entre les deux sous-régions de la découpe est l'un des deux termes à maximiser par le processus. Elle évalue les contrastes locaux. La définition 3.7 formalise la distance locale le long d'une hypothèse de découpe verticale.

Définition 3.7 *Le long d'une hypothèse de découpe verticale d'abscisse x_k , la distance locale $\delta(k, c_v)$ est la somme des longueurs des arêtes de contour.*

$$\delta(k, c_v) = \sum_{j=0}^m c_v(k, j) \overline{v_{k,j}}$$

où $\overline{v_{i,j}}$ est la longueur de l'arête $v_{i,j}$.

Nous définissons la distance locale des hypothèses horizontales d'une manière analogue.

3.3.4 Définition de la distance globale

Le choix du calcul d'une distance globale entre les deux sous-régions est plus délicat. Toutes les informations, selon leur localisation dans la région, n'ont pas la même pertinence. Dans cette sous-section, nous présentons d'abord la manière dont nous pouvons définir cette distance, puis celle dont nous pouvons compenser les problèmes de bord.

La distance globale $\Delta(k, c_h)$ entre les deux régions séparées par l'hypothèse de découpe d'abscisse x_k mesure la différence des densités d'arêtes de contour entre les deux sous-régions. Les définitions 3.8 et 3.9 proposent deux modes de calcul de cette distance. Le présent paragraphe traite des hypothèses verticales, la distance globale le long des hypothèses horizontales se calculant d'une manière analogue.

Définition 3.8 *Le long d'une hypothèse de découpe verticale d'abscisse x_k , la distance globale $\Delta^1(k, c_h)$ est la somme des valeurs absolues des différences de densités d'arêtes de contour entre les deux sous-régions.*

$$\Delta^1(k, c_h) = \sum_{i=1}^m \left| \frac{\sum_{j=0}^k c_h(i, j) \overline{h_j}}{x_k - x_0} - \frac{\sum_{j=k+1}^n c_h(i, j) \overline{h_j}}{x_n - x_k} \right|$$

Définition 3.9 *Le long d'une hypothèse de découpe verticale d'abscisse x_k , la distance globale $\Delta^2(k, c_h)$ est la somme quadratique des différences de densités d'arêtes de contour entre les deux sous-régions.*

$$\Delta^2(k, c_h) = \sum_{i=1}^m \left(\frac{\sum_{j=0}^k c_h(i, j) \overline{h_j}}{x_k - x_0} - \frac{\sum_{j=k+1}^n c_h(i, j) \overline{h_j}}{x_n - x_k} \right)^2$$

La figure 3.8 illustre le calcul de la distance globale Δ^2 pour l'hypothèse x_{12} avec l'analyse (à droite de l'image) d'un alignement dominant horizontal dont la densité de contours droite est négligeable devant celle de gauche. Le profil vert en bas de l'image représente toutes les différences de densité d'arêtes de contour horizontales. Le processus doit sélectionner la plus élevée.

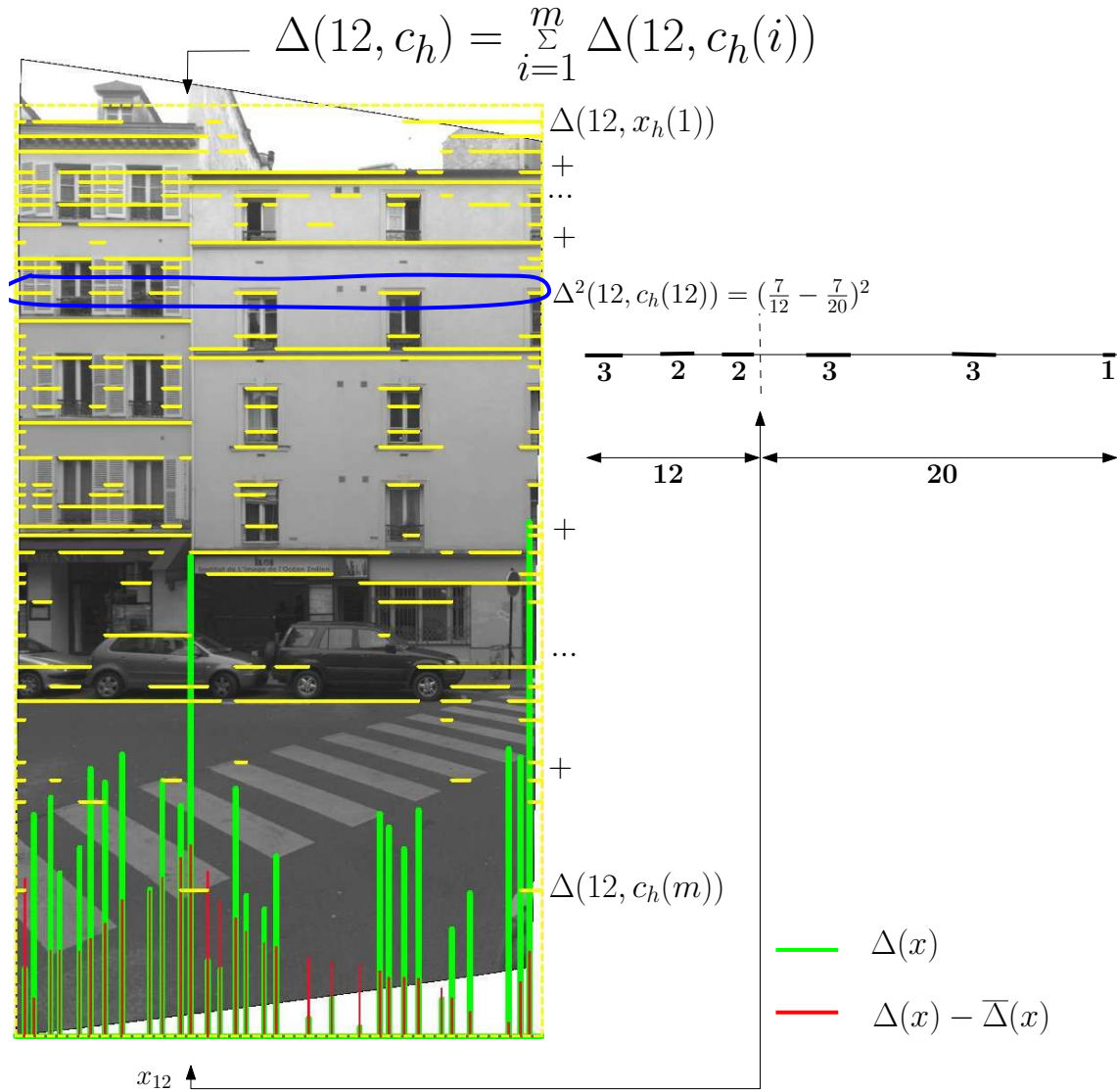


FIGURE 3.8 – Distance globale. Les lignes jaunes sont les arêtes de contour horizontales. Les profils vert et rouge en bas de l'image représentent respectivement les valeurs brutes et pondérées de la distance globale. Un test est effectué sur l'hypothèse de découpe x_{12} au niveau de l'alignement horizontal (cyan) sur la droite de l'image.

3.3.5 Correction de la distance globale

Malheureusement la distance globale est biaisée au niveau des bords de la région mère : son espérance sur toute les distributions $c_h(i, j)$ n'est pas constante avec j . Ses valeurs périphériques sont sur-élevée par rapport aux autres. Ce phénomène se comprend facilement en considérant par exemple le cas de la découpe le long de l'alignement vertical le plus à droite : chaque alignement horizontal de la sous-région droite ne contient qu'une arête. Si, sur un alignement donné, il s'agit d'une arête de fond, la densité sera nulle. S'il s'agit d'une arête de contour, elle sera maximale. Ainsi, d'emblée, à cet endroit la distance compare une densité extrême (nulle ou maximale) à une autre densité. D'où le biais. Le pic visible sur le bord droit de l'image de la

figure 3.8 témoigne de ce phénomène. Nous proposons alors de corriger la distance globale en lui retranchant son espérance sur toutes les distributions $c_h(i, j)$ (profil rouge sur la figure).

Pour calculer l'espérance $\bar{\Delta}(k, c_h) = \sum_{c_h} p(c_h) \Delta(k, c_h)$ de la distance globale, nous avons besoin de connaître la loi de probabilité $p(c_h)$ des distributions d'arêtes le long d'un alignement horizontal. Il est possible de considérer les différents alignements horizontaux comme autant de réalisations indépendantes d'une distribution d'arêtes. Nous pouvons alors utiliser un estimateur basé sur la fréquence d'occurrence des contours sur les arêtes d'une même colonne. L'équation 3.1 propose ainsi, le long d'un alignement horizontal, la fonction de masse d'occurrence des distributions d'arêtes.

$$\begin{cases} p(c_h) = \prod_{j=0}^n [p(c_h(j))]^{c_h(j)} [1 - p(c_h(j))]^{(1-c_h(j))} \\ p(c_h(j)) = \sum_{i=0}^m \frac{c_h(i, j)}{m+1} \end{cases} \quad (3.1)$$

Ainsi la distance globale est-elle compensée statistiquement de manière équilibrée. Cette compensation assure d'une part que les valeurs sont rabaisées artificiellement aux bords. Elle garantit d'autre part que si une rupture de densité de contours se trouve effectivement au bord, alors, malgré la pénalisation, les valeurs présenteront tout de même un pic à ce bord (*cf* figure 3.9.a). D'autre part, si la région considérée présente deux ruptures de densité de contours, alors la distance présentera deux pics à leur niveau (*cf* figure 3.9.b).

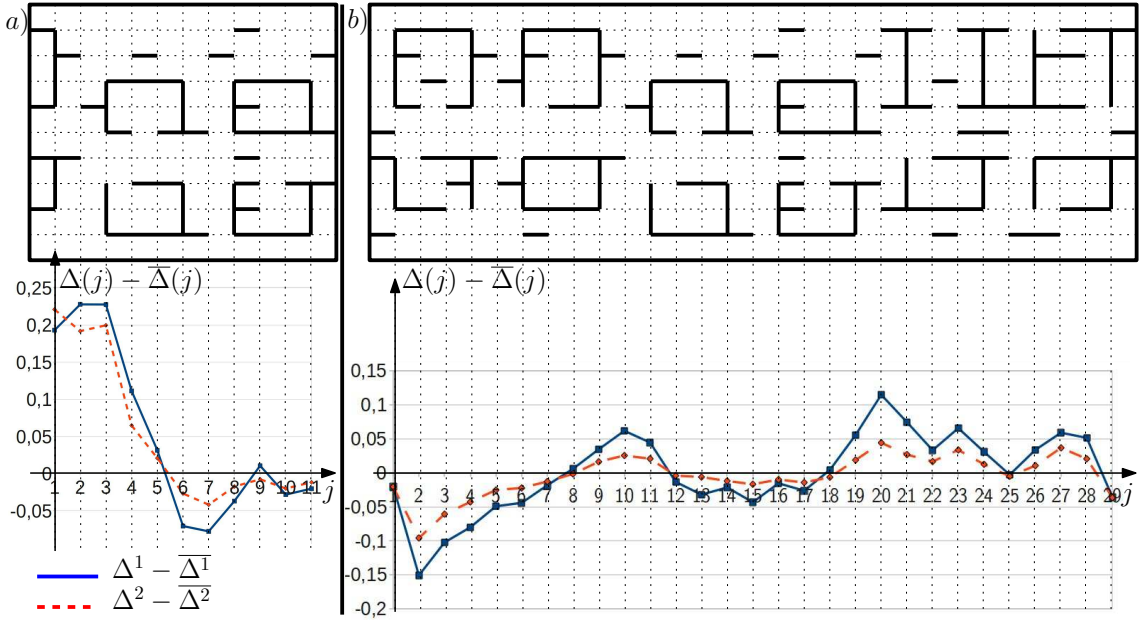


FIGURE 3.9 – Allure de la distance globale dans les cas limites : a) Pic de la distance globale au bord en présence d'une rupture de densité de contours malgré la pénalisation ; b) Deux pics de la distance en cas de double rupture de densité de contours.

En pratique, c'est la distance Δ^2 (définition 3.9) que nous utilisons sous l'hy-

pothèse de l'uniformité des longueurs d'arêtes et de celle de la probabilité d'occurrence d'un contour ($\overline{h_j}$ et $p(c_h(j))$ constantes). Ce choix est justifié dans la sous-section 3.3.7 : en réalité, le calcul n'est pas réalisable pour les images que nous utilisons de la manière précédemment exposée. Sa complexité est exponentielle et les hypothèses d'uniformités permettent de la linéariser.

3.3.6 Distance somme

La meilleure hypothèse de découpe verticale d'indice k^* est donnée par la définition 3.10. Un calcul similaire permet de sélectionner la meilleure hypothèse de découpe horizontale.

Définition 3.10 *La distance somme est une combinaison linéaire des distances locales et globales. L'hypothèse de découpe verticale optimale k^* est celle qui maximise cette distance :*

$$k^* = \arg \max_k (\alpha f(\Delta(k, c_h)) + (1 - \alpha)g(\delta(k, c_v)))$$

où les fonctions f et g définissent respectivement l'influence des distance Δ et δ . Le paramètre $\alpha \in [0; 1]$ permet d'équilibrer ces deux termes.

En pratique, afin d'éviter une compensation excessive du problème de bords, la fonction f doit être nulle si le nombre d'alignements dominants orthogonaux aux hypothèses de découpe n'est pas significatif (dans notre implémentation : inférieur à 6).

Tout notre propos s'applique de la même manière aux découpes horizontales. La découpe sélectionnée par le processus est celle dont la distance somme est la plus élevée indépendamment de l'orientation de cette découpe.

Discussion sur le paramètre α La figure 3.10 montre l'influence des deux distances. Si la distance globale Δ est pénalisée au profit de la distance locale δ , les découpes ont une plus grande tendance à s'effectuer sur les forts gradients. Réciproquement, la pénalisation de la distance Δ entraîne une répartition plus uniforme des lignes de découpe. C'est pourquoi la figure 3.10.c ($\alpha = 20\%$) montre une segmentation qui commence par séparer les étages, alors que les lignes de découpe de la segmentation de la figure 3.10.a ($\alpha = 80\%$) sont distribuées d'une manière plus indépendante des gradients.

Actuellement, le paramètre α est fixé empiriquement et de manière fixe à 50% pour nos données. Il pourrait être envisageable de le faire dépendre de la densité de contours de la région mère ou de la détection des alignements dominants. Pénaliser par exemple davantage la distance globale lorsqu'aucun maximum net ne sort du profil d'accumulation de gradient.

3.3.7 Discussion sur la distance globale

La problématique de correction périphérique de la distance globale est un point clé de la segmentation car il est important de travailler sur toutes les régions de l'image. En effet, les données que nous traitons ne sont pas centrées sur l'intégralité

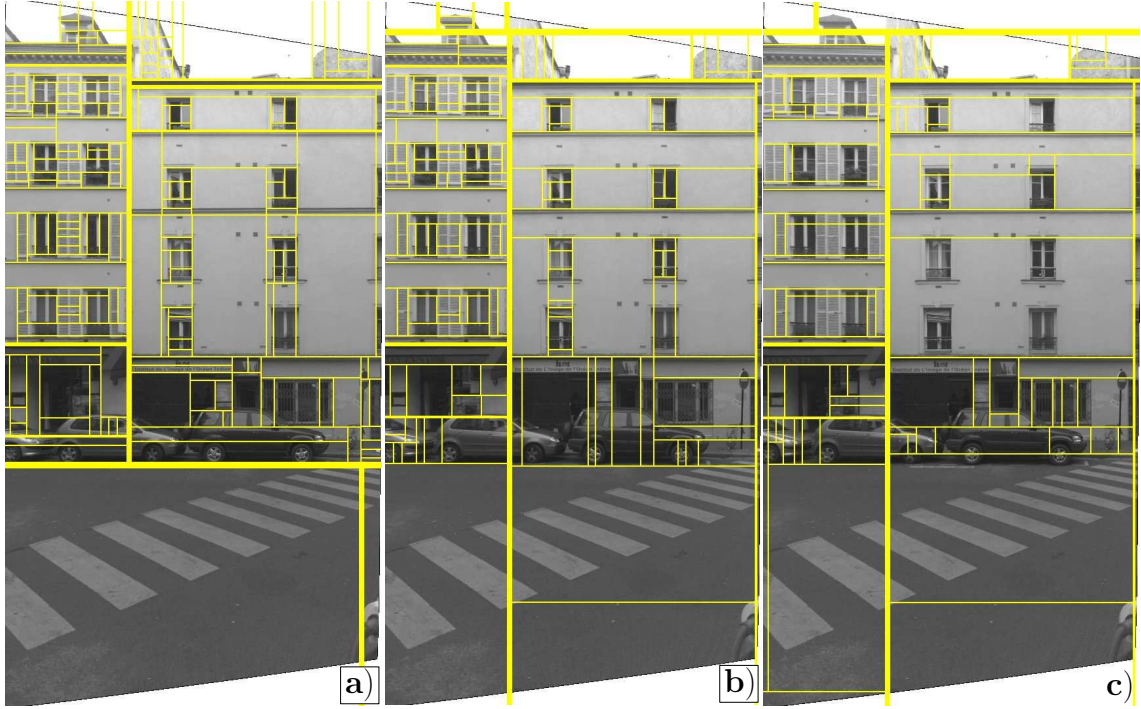


FIGURE 3.10 – Influence du paramètre α sur la segmentation. a) $\alpha = 20\%$ b) $\alpha = 50\%$ c) $\alpha = 80\%$

d'une unique façade mais concernant une, deux, voir trois façades généralement tronquées par les bords de l'image.

Nous avons défini une distance globale Δ^1 (définition 3.8) et une distance globale Δ^2 (définition 3.9). Il s'agit de déterminer laquelle est la mieux adaptée à notre problématique. La première est linéaire avec la différence de densité de contours entre les deux sous-régions (en bleu sur la figure 3.11). La deuxième est quadratique (en rouge sur la figure 3.11). La distance Δ^1 est globalement beaucoup plus élevée que la distance Δ^2 (qui a tendance à *écraser* les faibles valeurs) mais leur allure générale est similaire.

Elles présentent toutes les deux des problèmes de bords, leur espérance sur toute les distributions $c_h(i, j)$ n'étant pas constante avec j . Nous avons alors proposé de les corriger en leur retranchant leur espérance sur toutes les distributions $c_h(i, j)$ (en vert pour $(\Delta^1 - \bar{\Delta}^1)$ et en jaune pour $(\Delta^2 - \bar{\Delta}^2)$ sur la figure 3.11).

Pour calculer cette espérance, nous avons proposé d'utiliser la fréquence verticale des contours comme fonction de masse de l'occurrence des arêtes le long d'un alignement horizontal (en bleu sur la figure 3.12).

Mais, tel quel, le calcul de $\bar{\Delta}$ est irréalisable. Il nécessite en effet le parcours de toutes les configurations possibles d'arêtes le long d'un alignement horizontal, ce qui représente une complexité de l'ordre de 2^n si n est le nombre d'alignements dominants verticaux. Ce problème est d'autant plus bloquant que le calcul est à entreprendre à chaque itération de la segmentation.

En pratique, nous utilisons une approximation en imposant certaines conditions. Dans la suite nous proposons une approximation de $\bar{\Delta}^1$ et de $\bar{\Delta}^2$ ainsi qu'une discussion sur le choix de l'une de ces deux distances.

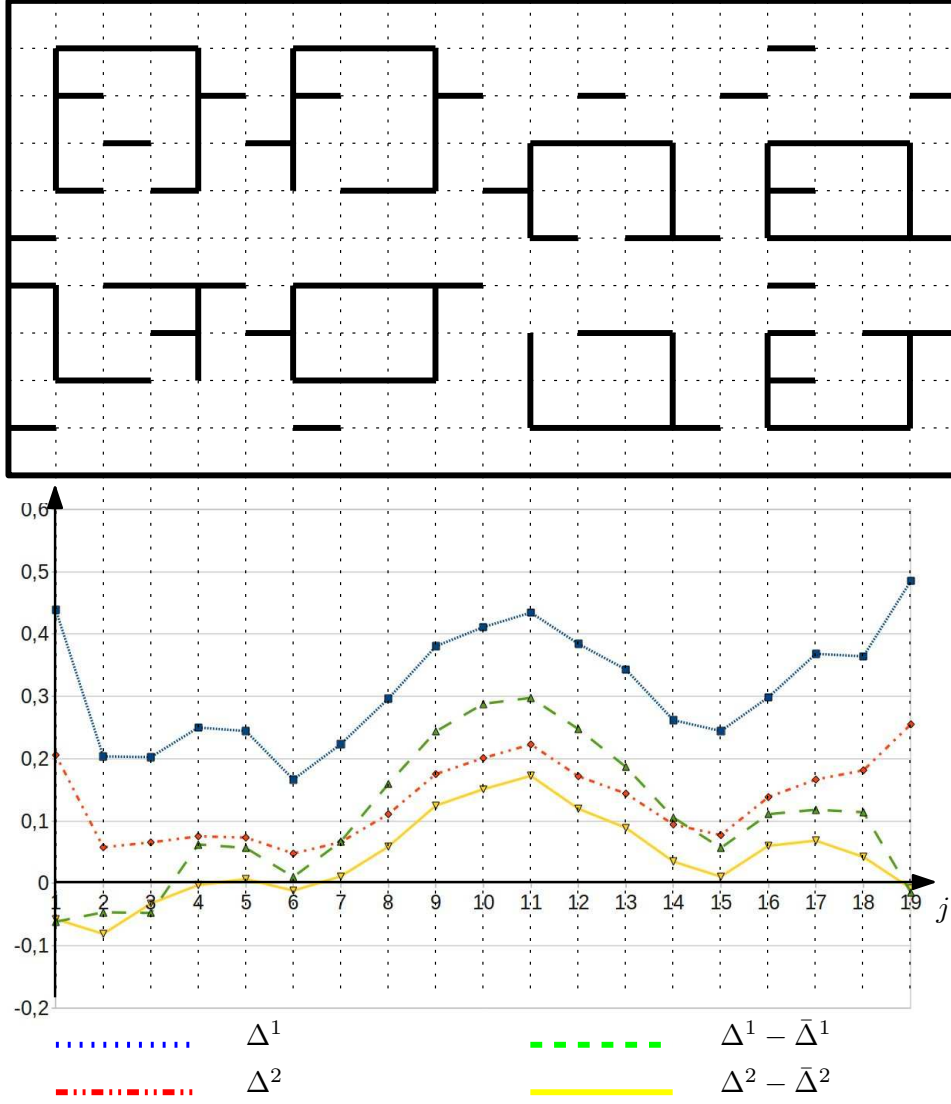


FIGURE 3.11 – Allure de la distance globale. Courbes bleue et rouge : respectivement distances Δ^1 (définition 3.8) et Δ^2 (définition 3.9). Courbes verte et jaune : corrections.

3.3.7.1 Espérance $\bar{\Delta}^1$, cas homogène, équiprobable, aux bords

Pour approcher l'espérance $\bar{\Delta}^1$, nous posons les hypothèses suivantes :

1. Les alignements dominants sont uniformément espacés (cas homogène : $\bar{h}_j$ constante) ;
2. La probabilité qu'une arête soit située sur un contour est exactement $\frac{1}{2}$ (cas équiprobable) ;
3. La densité de contours de la sous-région la plus grande vaut $\frac{1}{2}$.

Notons $\bar{\Delta}^{1'}(k)$ cette approximation. Nous exposons le calcul pour le bord gauche. La même démarche s'applique symétriquement au bord droit.

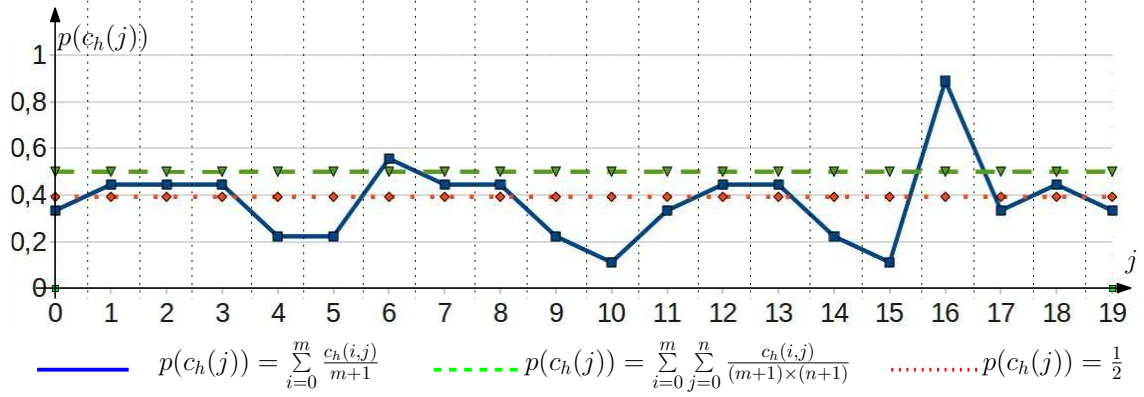


FIGURE 3.12 – Estimation empirique de la fonction de masse de l'occurrence des arêtes de contour le long d'un alignement horizontal de l'image synthétique de la figure 3.11.

$$\overline{\Delta^{1'}}(k, c_h) = \sum_c p(c, c_h) \Delta^{1'}(k, c_h) = \sum_{g=0}^k p(c_g) \Delta^{1'}(k, c_g)$$

où c_g désigne l'ensemble des distributions d'arêtes horizontales d'exactly g arêtes de contours dans la sous-région gauche.

$$\overline{\Delta^{1'}}(k, c_h) = \sum_{g=0}^k \frac{1}{2^k} C_k^g \left| \frac{g}{k} - \frac{1}{2} \right| = \frac{1}{2^k} 2 \sum_{g=0}^{\lfloor k/2 \rfloor} C_k^g \left(\frac{1}{2} - \frac{g}{k} \right)$$

$$\overline{\Delta^{1'}}(k, c_h) = \frac{1}{2^k} \left[\left(\sum_{g=0}^{\lfloor k/2 \rfloor} C_k^g \right) - \frac{2}{k} \left(\sum_{g=0}^{\lfloor k/2 \rfloor} g C_k^g \right) \right]$$

$$\overline{\Delta^{1'}}(k, c_h) = \frac{1}{2^k} \left[H^0(k) - \frac{2}{k} H^1(k) \right]$$

où H^0 et H^1 sont les demi-sommes définies en annexe A.

$$\begin{cases} \overline{\Delta^{1'}}(2k', c_h) = \frac{1}{2^{2k'}} \left[\frac{1}{2} (2^{2k'} + C_{2k'}^{k'}) - \frac{2}{2k'} (k' 2^{2k'-1}) \right] \\ \overline{\Delta^{1'}}(2k' + 1) = \frac{1}{2^{2k'+1}} \left[2^{2k'} - \frac{2}{2k'+1} \left(\frac{2k'+1}{2} (2^{2k'} - C_{2k'}^{k'}) \right) \right] \end{cases}$$

d'après les résultats de H^0 et H^1 démontrés en annexe A.

$$\boxed{\overline{\Delta^{1'}}(k, c_h) = \frac{C_{2k'}^{k'}}{2^{2k'+1}} \quad \text{avec} \quad k' = \lfloor k/2 \rfloor} \quad (3.2)$$

3.3.7.2 Espérance $\overline{\Delta^2}$, cas homogène et uniforme

Nous proposons une approximation de l'espérance $\overline{\Delta^2}$ sous des hypothèses moins contraignantes :

1. Les alignements dominants sont uniformément espacés ($\overline{h_j}$ constante) ;

2. La fonction de masse de l'occurrence des arêtes de contours le long d'un alignement horizontal est uniforme.

Notons $\overline{\Delta_p^{2'}}$ cette approximation et p la probabilité d'une arête de se trouver sur un contour.

$$\overline{\Delta_p^{2'}}(k, c_h) = \sum_c p(c) \Delta_p^{2'}(k, c_h) = \sum_{g=0}^k \sum_{d=0}^{n-k} p(c_{g,d}) \Delta_p^{2'}(k, c_{g,d})$$

où $c_{g,d}$ désigne l'ensemble des distributions d'arêtes horizontales d'exactly g arêtes de contour dans la sous-région gauche et d arêtes de contour dans celle de droite.

$$\overline{\Delta_p^{2'}}(k, c_h) = \sum_{g=0}^k C_k^g p^g (1-p)^{k-g} \sum_{d=0}^{n-k} C_{n-k}^d p^d (1-p)^{n-k-d} \left(\frac{g}{k} - \frac{d}{n-k} \right)^2$$

$$\overline{\Delta_p^{2'}}(k, c_h) = \frac{\left(\sum_{g=0}^k g^2 C_k^g p^g (1-p)^{k-g} \right) \left(\sum_{d=0}^{n-k} C_{n-k}^d p^d (1-p)^{n-k-d} \right)}{k^2}$$

$$- 2 \frac{\left(\sum_{g=0}^k g C_k^g p^g (1-p)^{k-g} \right) \left(\sum_{d=0}^{n-k} d C_{n-k}^d p^d (1-p)^{n-k-d} \right)}{k(n-k)}$$

$$+ \frac{\left(\sum_{g=0}^k C_k^g p^g (1-p)^{k-g} \right) \left(\sum_{d=0}^{n-k} d^2 C_{n-k}^d p^d (1-p)^{n-k-d} \right)}{(k-n)^2}$$

$$\overline{\Delta_p^{2'}}(k, c_h) = \frac{S_p^2(k) S_p^0(n-k)}{k^2} - 2 \frac{S_p^1(k) S_p^1(n-k)}{k(n-k)} + \frac{S_p^0(k) S_p^2(n-k)}{(k-n)^2}$$

où S_p^0 , S_p^1 et S_p^2 sont les sommes définies en annexe A.

$$\overline{\Delta_p^{2'}}(k, c_h) = \frac{kp(1-p+kp)}{k^2} - 2 \frac{kp(n-k)p}{k(n-k)} + \frac{(n-k)p(1-p+(n-k)p)}{(k-n)^2}$$

d'après les résultats de S_p^0 , S_p^1 et S_p^2 démontrés en annexe A.

$$\boxed{\overline{\Delta_p^{2'}}(k, c_h) = p(1-p) \left(\frac{1}{k} + \frac{1}{n-k} \right)} \quad (3.3)$$

En pratique, nous utilisons l'estimateur $p = \sum_{i=0}^m \sum_{j=0}^n \frac{c_h(i,j)}{(m+1)(n+1)}$ (figure 3.12).

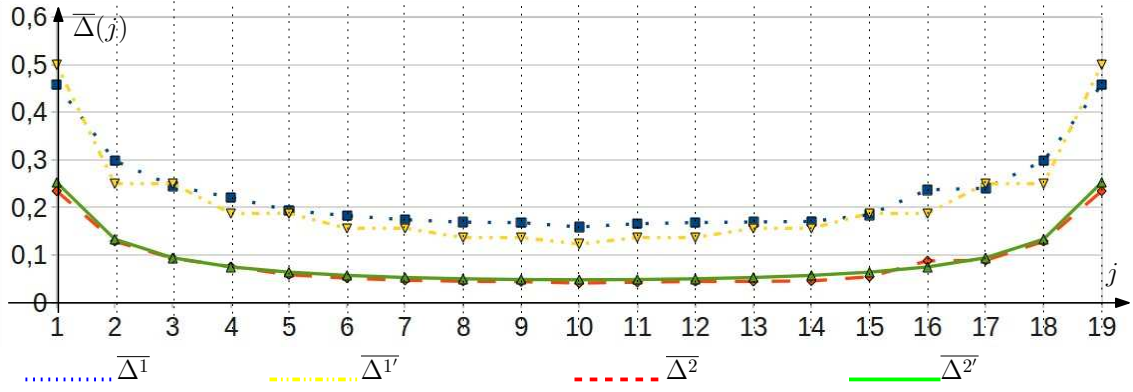


FIGURE 3.13 – Comparaison des espérances approchées $\overline{\Delta}^{1'}$ et $\overline{\Delta}^{2'}$ avec les espérances empiriques $\overline{\Delta}^1$ et $\overline{\Delta}^2$ de la figure 3.11.

3.3.7.3 Choix de la distance globale

C'est la distance quadratique Δ^2 que nous avons choisi d'utiliser car ses hypothèses sont moins contraintes. La figure 3.13 montre ainsi qu'elle se rapproche davantage de l'espérance empirique de l'équation 3.1.

Cette formalisation a été publiée dans Burochin et al. [2010]. Elle fait suite à une définition primitive de la distance globale utilisée dans Burochin et al. [2009]. Initialement, en effet, nous basions cette distance non sur une différence de densités de contours mais sur la non-continuité des alignements dominants dans les deux sous-régions. Pour une hypothèse de découpe verticale, des sous-alignements dominants horizontaux étaient en effet détectés séparément de part et d'autre. Deux sous-quadrillages étaient ainsi constitués à partir des alignements dominants verticaux et des sous-alignements dominants horizontaux. Nous procédions de même pour les hypothèses de découpe horizontales. La meilleure hypothèse de découpe était celle qui minimisait la longueur des arêtes de fond dans chacune des deux sous-régions selon la définition 3.11. Nous appelons Δ^0 cette définition primitive. La figure 3.14a montre ce processus d'optimisation.

Définition 3.11 *Le long d'une hypothèse de découpe verticale d'indice j , la distance globale $\Delta^0(j)$ est la différence entre la longueur des arêtes de fond de la région mère et celle des arêtes de fond de chacune des deux sous-régions.*

$$\Delta^0(j) = \frac{\left[\sum_{h \in H_0} (1 - c_{H_0}(h)) \bar{h} \right] - \left[\sum_{h \in H_1} (1 - c_{H_1}(h)) \bar{h} \right] - \left[\sum_{h \in H_2} (1 - c_{H_2}(h)) \bar{h} \right]}{\max \left(\sum_{h \in H_0} (1 - c_{H_0}(h)) \bar{h}, \sum_{v \in V_0} (1 - c_{V_0}(v)) \bar{v} \right)}$$

où H_0 , H_1 et H_2 sont respectivement les ensembles d'arêtes horizontales de la région mère, de la sous-région de gauche et de la sous-région de droite. De même, V_0 , V_1 et V_2 sont respectivement les ensembles d'arêtes verticales de la région mère, de la sous-région du haut et de la sous-région du bas.

Cependant, la génération des sous-alignements dominants ne s'effectuait pas à la même échelle dans les deux sous-régions. La comparaison des arêtes de contour ne

concernait alors pas le même type de structures. La figure 3.14.a montre par exemple des arêtes de fond sur la région du haut qui correspondent à peu près avec celles de la région mère alors que la région du bas contient de longues arêtes de fond parasites générées par de forts contrastes locaux, insignifiants dans la région mère. Les hypothèses de découpe le long de ces arêtes parasites sont associées à des distances négatives. C'est pourtant à ces endroits que se situe la découpe optimale. Dans ce cas typique, une information relative à un segment de panneau de signalisation isolé est comparée à une séquence de bords de fenêtre.

Sur la même image, la figure 3.15 présente une comparaison de la segmentation hiérarchique de Burochin et al. [2009] avec celle de Burochin et al. [2010]. La première sépare d'abord les deux façades alors que la deuxième sépare les façades de la rue. Les deux choix sont pertinents mais notons la nette amélioration de la deuxième segmentation qui réduit de beaucoup la tendance à la sur-segmentation.

3.4 Conclusion

Nous venons de présenter une segmentation hiérarchique de façade en modèles élémentaires 2D. Nous avons vu que selon l'échelle et la localisation de la façade dans l'image (en bas, en haut, tronquée, etc.), l'ordre des découpes peut varier mais que chacun de ces modèles encadre une structure cohérente. La façade est séparée de ses éventuelles voisines et de son environnement. Le processus est simple et rapide (de l'ordre de la seconde pour une image de 2 méga-pixels sur un processeur de 2,83 GHz) mais pourrait encore être optimisé.

En outre, les trois modules de segmentation constituent chacun un atout très important du processus :

Alignements dominants Ces éléments géométriques représentent une propriété fondamentale de la segmentation (définition 3.1). Ils sont très rapides à calculer car il s'agit de la manipulation de profils de projection en une seule dimension. Par nature, la régularité de leur distribution permet de travailler dans un espace restreint très intéressant de l'image. Cet espace fournit un ensemble d'hypothèses pertinentes de découpe et de périodicités ainsi qu'une distribution uniforme de points d'intérêt et de contours.

Modèles élémentaires Les modèles que nous cherchons sont très simples. Ils offrent une découpe de façade tout à fait cohérente. Le modèle périodique, en particulier, permet de contourner le problème des structures parasites isolées et les effets de perspectives en ne se focalisant que sur des points d'intérêt localisés dans des zones fiables.

Distance géométrique La distance géométrique que maximise le processus à chaque itération entre les deux sous-régions optimales repose sur une formalisation inédite dans la littérature. Elle rend compte de manière équilibrée de la différence de densités de contours et propose une compensation des problèmes de bords.

La figure 3.16 présente un résultat de segmentation où la façade est séparée de la rue puis de sa voisine. La figure 3.17 montre la segmentation de plusieurs façades sur une image à plus petite échelle.

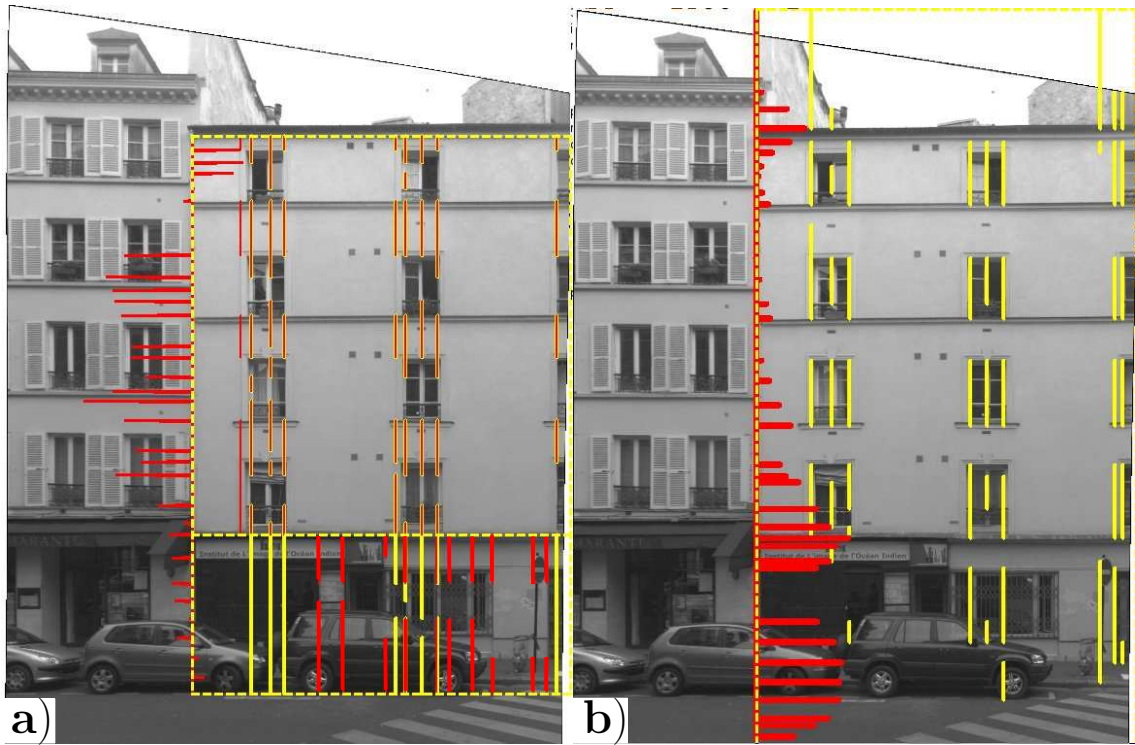


FIGURE 3.14 – a) Problème d'échelle rencontré par la distance globale primitive pour le cas de l'hypothèse de découpe juste au-dessus du rez-de-chaussée : en rouge, à gauche, le profil de valeurs de $\Delta^0(i)$; en rouge sur l'image, les arêtes de fond le long des sous-alignements dominants ; en jaune, les arêtes de fond de la région mère. b) Profil de valeurs de $\Delta^{2'}(i)$ en rouge, à gauche ; en jaune, les arêtes de contour.

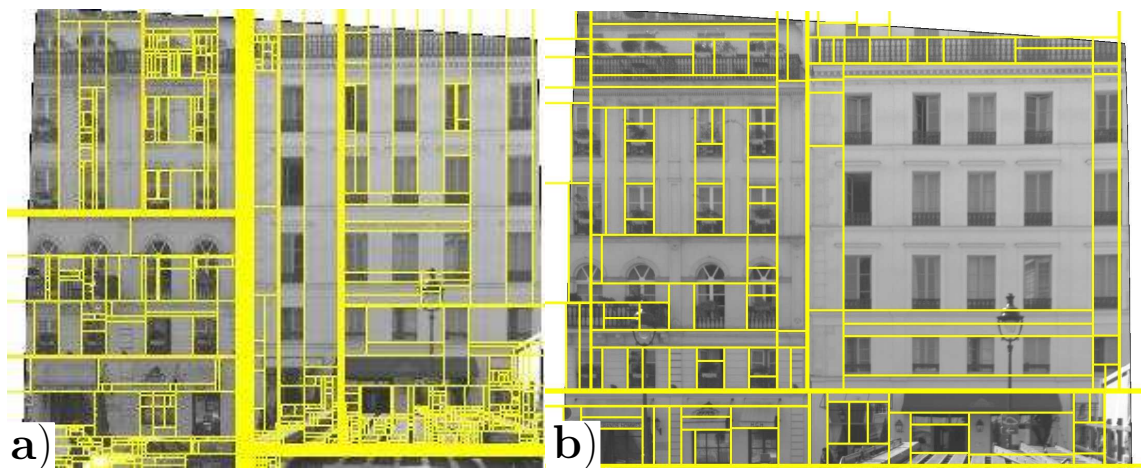


FIGURE 3.15 – a) Segmentation hiérarchique basée sur la distance Δ^0 et sans le modèle périodique (Burochin et al. [2009]). b) Segmentation hiérarchique basée sur la distance $\Delta^{2'}$ et utilisant le modèle périodique (Burochin et al. [2010]).

En revanche, il existe une tendance à la sur-segmentation dans les découpes profondes (au delà du troisième niveau de profondeur). Il peut arriver parallèlement que des structures de même type soient séparées différemment sur une même image. La figure 3.19 présente une façade acquise de deux points de vue différents. Dans le premier cas, l'image provient d'une région de la grande image de la figure 3.18. L'autre est acquise d'un point de vue plus proche de telle sorte que seule la partie gauche de la façade est visible. Remarquons la tendance à séparer les étages dans les deux cas. Mais la segmentation à l'intérieur des étages diffère beaucoup. Les troisième et quatrième étages sont par exemple découpés en petites parties sur l'image 3.19.a tandis qu'ils sont groupés dans un modèle périodique sur l'image 3.19.b. Ce phénomène peut s'expliquer par la différence de situation initiale pour les deux segmentations : la largeur des étages considérés dans l'image 3.19.b est deux fois plus petite que celles des étages de l'image 3.18.

De manière générale, le processus découpe parfois *aveuglément* des régions présentant des périodicités avant même d'encadrer ces dernières et de les apparier au modèle périodique. La distance utilisée alors ne prend absolument pas en compte d'informations de périodicité.

3.4.1 Perspectives

Beaucoup de points du processus de segmentation peuvent encore être améliorés.

3.4.1.1 Perfectionnement de la modélisation

De manière générale, les modèles que nous avons présentés sont très simples. Ils ne modélisent que très grossièrement la façade. La figure 3.20 présente le résidu entre l'image de la figure 3.16 et les modèles détectés. Ces résidus pourraient être exploités pour accroître encore davantage le niveau de détail de la modélisation.

Concernant les modèles planaires et cylindriques, une modélisation plus fine pourrait consister en une analyse de l'histogramme des radiométries de la région concernée. Ainsi, par exemple, le modèle planaire pourrait être détecté par la présence d'un mode compact dans cet histogramme. En effet, parfois, le mode principal ne se situe pas nécessairement autour de la médiane car la région peut être bruitée ou comporter quelques occlusions locales. Un calcul analogue permettrait la détection des modèles cylindriques. Cette analyse reposerait sur celle d'un échantillon représentatif des régions de l'image de la façade. Elle permettrait d'enlever un seuil dans la modélisation.

D'autre part, il n'est pas intéressant de travailler à pleine résolution car cet appariement compare le modèle à l'image *pixel à pixel*. Le procédé est non seulement très lent mais demeure surtout sensible au bruit de l'image. Il faut donc travailler sur une image lissée, puis sous-échantillonnée (d'un facteur 4 dans les deux dimensions, par exemple, pour une image de façade d'une résolution de 2 méga-pixels). Ainsi, l'appariement est beaucoup plus rapide et est moins sensible au bruit.

D'autre part, nous avons vu pour le cas du modèle périodique, que nous pourrions prendre la perspective en compte en manipulant du $2, 5D$. Un ajustement des points d'intérêt sur des informations locales peut aussi être pertinent, notamment pour écarter ceux situés dans des zones de radiométrie uniforme.



FIGURE 3.16 – Résultat de segmentation hiérarchique : la façade est d’abord séparée de la rue, puis de sa voisine. Découpes en jaune (épaisseur du trait fonction de la profondeur de la segmentation). Modèles planaires : quadrillages rouges. Modèles cylindriques : bandes rouges. Périodes : arcs rouges. Régions trop petites : croix rouges.

Le noyau de la périodicité pourrait aussi être segmenté à nouveau, afin de parvenir à une région irréductible semblable à celle proposée par Müller et al. [2007].

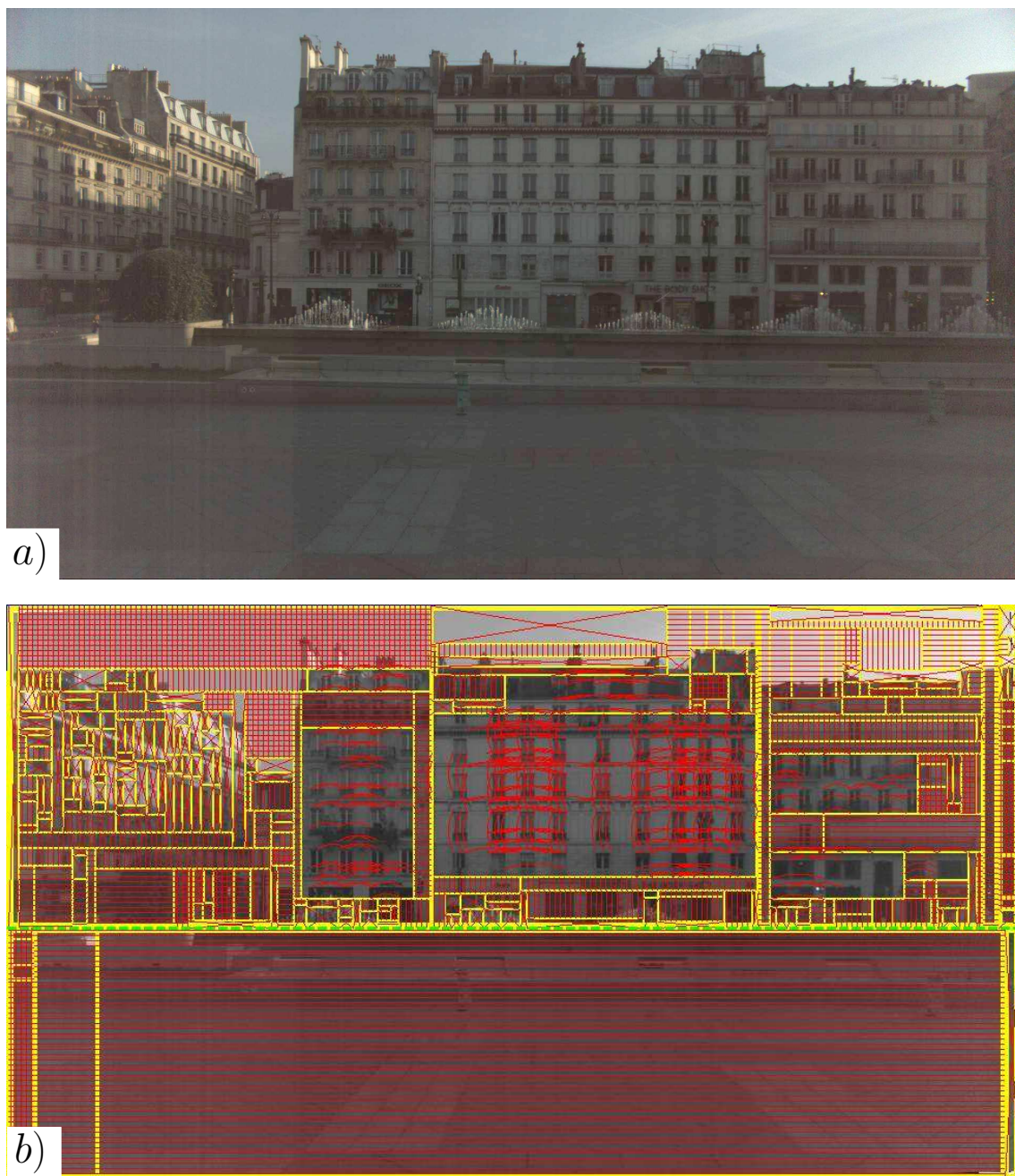


FIGURE 3.17 – Segmentation hiérarchique d’une image de plusieurs façades. a) Image d’entrée ; b) Segmentation (découpes en jaune et modèles en rouge).



FIGURE 3.18 – Segmentation hiérarchique de deux façades mitoyennes. Découpes en rouge (épaisseur du trait fonction de la profondeur de la segmentation).

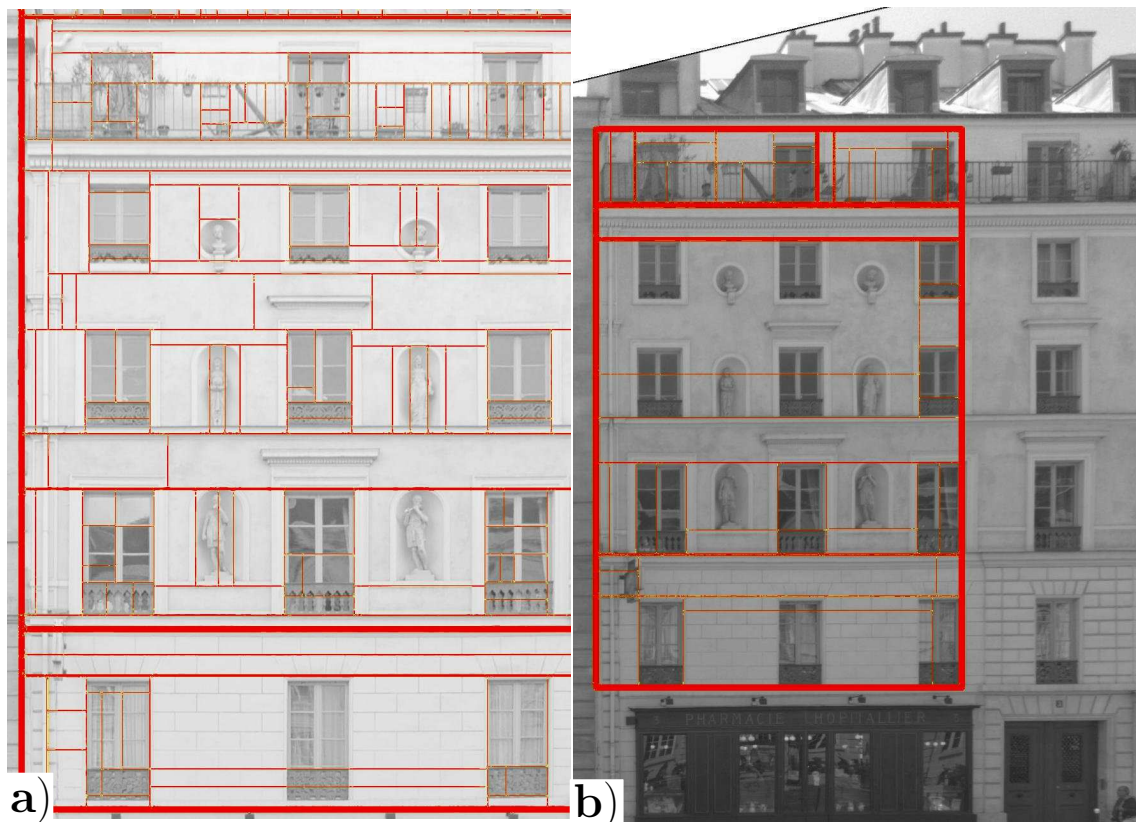


FIGURE 3.19 – Segmentation hiérarchique d'une façade acquise de deux différents points de vue et de différents appareils. a) Segmentation d'un partie de la façade présentée sur la figure 3.18. b) Segmentation de la même façade acquise d'un point de vue plus proche de telle sorte que seule la partie gauche est visible.

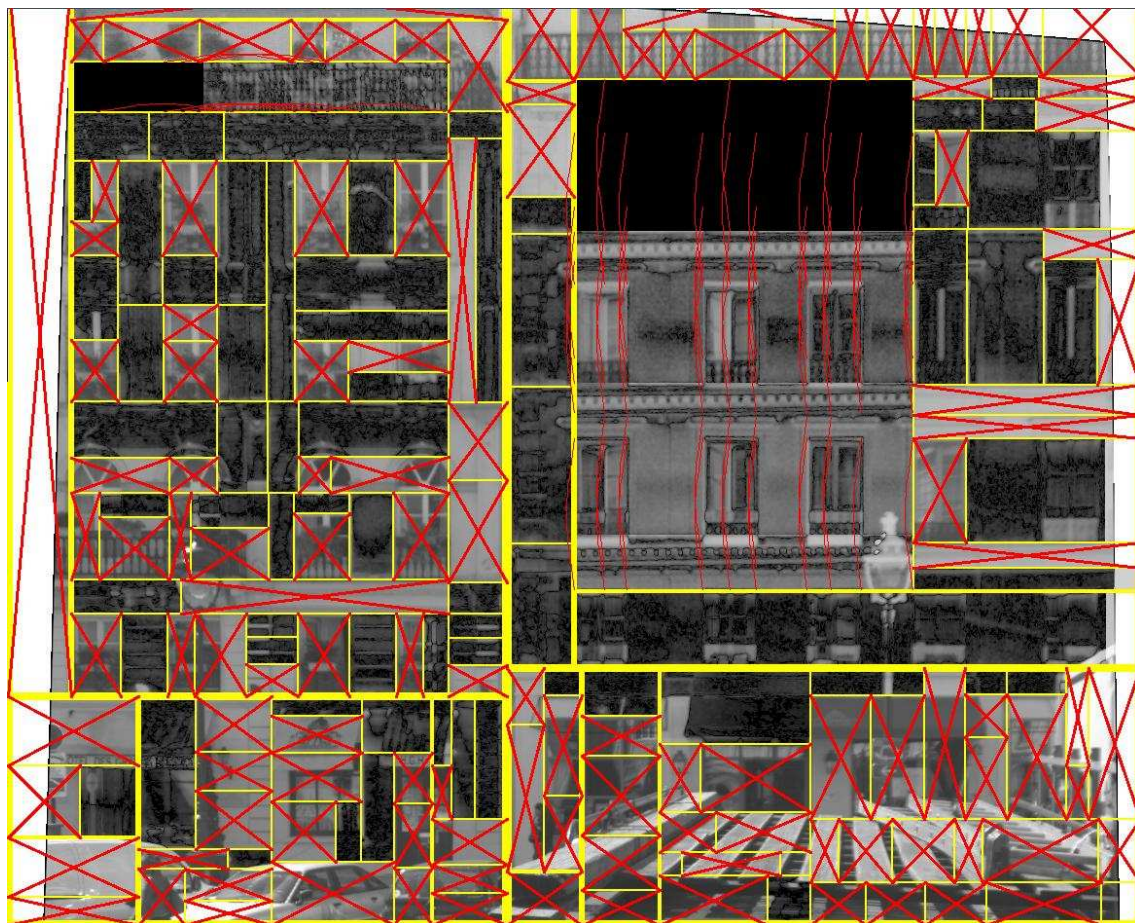


FIGURE 3.20 – Image du résidu entre l'image de la figure 3.24.a et les modèles (noir : résidu nul, blanc : résidu maximal). Les régions non modélisées sont barrées en rouge. Les découpes de la segmentation sont tracées en jaune.

De nouveaux modèles pourraient enfin être introduits à partir d'autres critères (orientation du gradient, bancs de filtres, etc.) pour les structures de type *appareil régulier* (mur de briques, textures en mosaïques, etc.) ou de type *motifs denses et arrondis* (fer forgé).

3.4.1.2 Perfectionnement de la classification des arêtes

Actuellement, le seuillage du poids des arêtes s'effectue autour de la médiane de ces poids, ce qui suppose qu'il y a autant d'arêtes de contour que d'arêtes de fond. Bien que cette approximation se vérifie grossièrement dans beaucoup de situations, il serait intéressant de calculer le seuil d'une manière plus réaliste. Les modes de l'histogramme des poids constituent, par exemple, un outil d'analyse tout à fait approprié à cette problématique. La figure 3.21 illustre ce point en présentant un histogramme de poids d'arêtes horizontales qui est séparé manuellement en trois parties. Les arêtes de contour sont immédiatement détectés.

D'autre part, actuellement, le poids des arêtes ne dépend que du gradient de l'image. Il serait intéressant d'observer le gain de robustesse avec une définition plus fine basée également, par exemple, sur l'orientation du gradient, ou bien sur

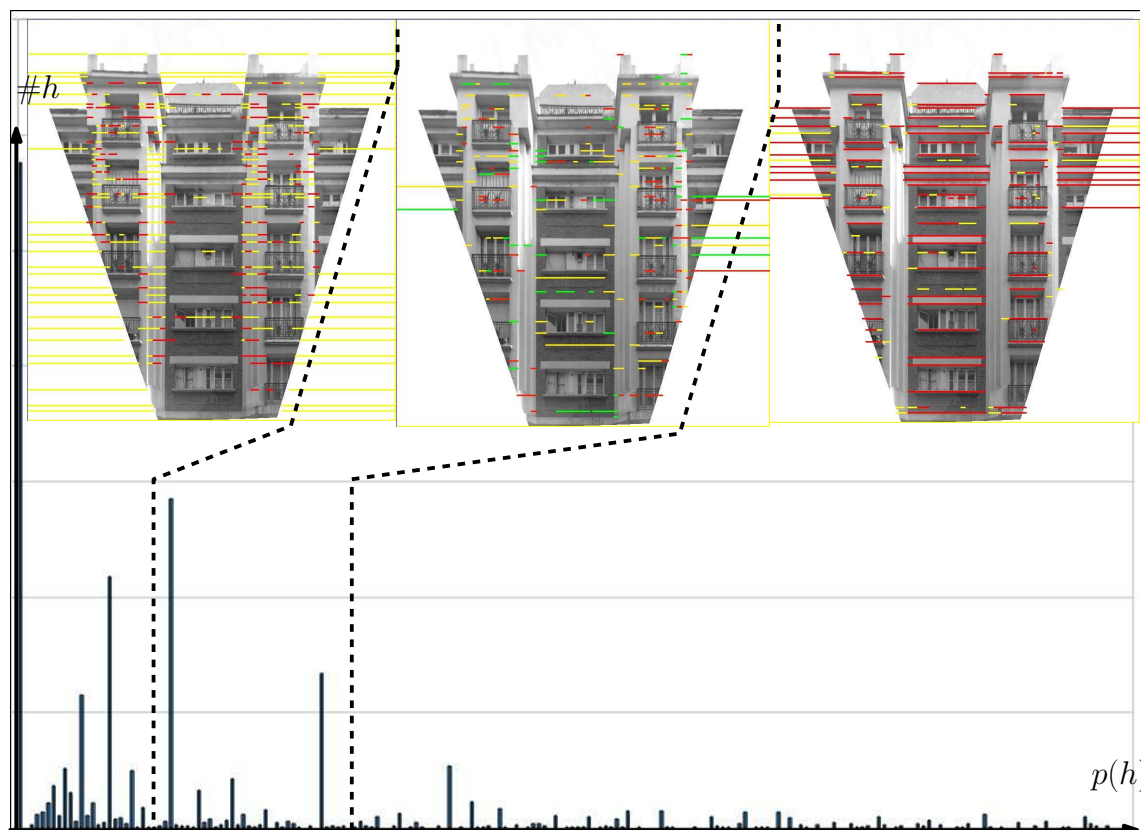


FIGURE 3.21 – Histogramme du poids des arêtes. Trois parties ont été définies manuellement et les arêtes correspondantes sont indiquées dans l'image pour chacune. Dans chacun de ces trois intervalles, les arêtes rouges ont les poids maximaux, les arêtes jaunes, les poids intermédiaires et les arêtes vertes ont les poids les plus bas (séparation entre ces valeurs en fonction des modes de l'histogramme).

le résultat d'une détection de contours effectuée en pré-traitement. La carte des scores de Harris et Stephens [1988] apparaît notamment tout à fait pertinente pour la pondération des arêtes.

3.4.1.3 Perfectionnement de la distance

En plus du gain de robustesse que peut apporter un perfectionnement de la classification des arêtes au calcul de la distance entre les deux sous-régions, nous envisageons d'autres pistes d'amélioration.

Concernant la distance globale, l'hypothèse d'équidistance des alignements dominants permet de garantir une compensation adéquate des problèmes de bords. En pratique, il arrive fréquemment que cette hypothèse ne soit pas vérifiée et qu'il existe de grands espaces entre plusieurs groupes d'alignements plus resserrés entre eux. Il serait envisageable d'ajouter artificiellement des alignements dans ces grands espaces afin de respecter l'hypothèse. Cet ajout ne perturberait pas le processus et ne le ralentirait pas de manière significative étant donné que les calculs sont effectués séparément dans chacune des deux dimensions.

Concernant la distance somme, nous avons noté l'importance du paramètre α

pour la pondération de la distance locale et de la distance globale entre les deux sous-régions. Une proposition a alors été faite de ne plus se contenter d'un choix manuel de sa valeur mais de la déterminer en fonction des propriétés géométriques de la région à découper.

D'autre part, les deux distances utilisées ne sont définies qu'à partir de la distribution des contours de l'image. Il pourrait être judicieux d'ajouter de nouveaux termes à la distance somme concernant des informations radiométriques ou de profondeur. L'utilisation de la teinte de l'image semble par exemple intéressante au vu des travaux de Liu et Gagalowicz [2010] qui en démontrent la pertinence pour extraire les vitres. Une simple différence de moyenne radiométrique entre la région mère et les sous-régions pourrait également être intéressante.

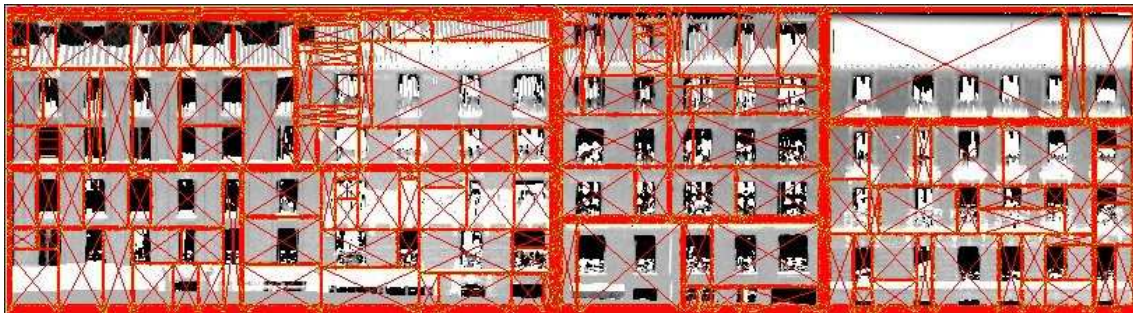


FIGURE 3.22 – Segmentation hiérarchique d'une carte de profondeurs.

La figure 3.22 présente un résultat de segmentation hiérarchique d'une carte de profondeurs construite à partir d'un nuage de points laser géo-référencés. Beaucoup d'auteurs utilisent cette information conjointement à celles de l'image (à commencer par les travaux de Schindler et Joachim [2003] qui utilisent la radiométrie et les points 3D générés à partir d'une séquence d'images).

Enfin, nous avons vu que la distance entre les deux régions manque encore radicalement d'information de périodicité des structures. Les problèmes comme ceux illustrés en figure 3.19 surviennent fréquemment. Il faudrait empêcher le processus de séparer les structures répétitives avant de les encadrer et de les apparier au modèle périodique. Il est ainsi possible d'utiliser une mesure de pénalité relative à la fréquence de paires de points bien corrélés par lesquelles passerait l'hypothèse de découpe ou le long de ceux-ci. Une telle mesure est illustrée à la figure 3.23 par le profil jaune. Wendel et al. [2010] également ont proposé une formalisation de ce type, comme nous l'avons vu dans le chapitre d'état de l'art (figure 2.6.b). Ils n'ont pas travaillé directement sur des points d'intérêt mais sur des profils radiométriques entre ces points. Nous n'avons pas poussé davantage nos recherches dans cette direction. Nous nous sommes orientés vers une détection de répétitivités dans les façades plus générale que le cadre de la segmentation hiérarchique. Ces dernières avancées sont expliquées en détail dans le chapitre 4.

3.4.1.4 Perfectionnement général de la segmentation

Il serait envisageable d'apporter de la robustesse à la segmentation des structures semblables dans l'image et de réduire la tendance du processus à sur-segmenter certaines autres structures en procédant à une phase de fusion des régions générées.



FIGURE 3.23 – Tentative de pénalisation de la découpe par des informations de périodicité (profil en jaune).

Cette fusion pourrait s'effectuer, soit à chaque itération, soit lors d'une deuxième passe indépendante. Vallet et al. [2009] proposent par exemple une segmentation dans cet esprit sur des empreintes d'immeubles à partir d'images aériennes et de modèles numériques de surface.

Une possibilité pour amener le processus à encadrer directement les structures répétitives sans les couper serait de lui permettre d'extraire une sous-région rectangulaire en une seule itération. Non plus partager la région en deux parties côte à côte mais en un rectangle et une région complémentaire concave.

La sélection du modèle s'effectue actuellement suivant un arbre de décision binaire. Il pourrait être envisageable de concevoir une distance normalisée entre l'image et les modèles permettant de guider cette sélection. La distance en question pourrait se baser sur un critère de moindre longueur de description (Minimum Description Length, *MDL*, Rissanen [1983]). Ainsi, la segmentation serait explicitement axée vers une compression de l'image, ce qui permettrait de s'affranchir de certains seuils (notamment le nombre de radiométries trop éloignées du modèle par exemple). Gorelick et al. [2011] proposent une segmentation de façade dans cet esprit en utilisant les répétitivités de structures.

3.4.2 Applications

La segmentation que nous venons de présenter permet de hiérarchiser les structures de la façade de manière asémantique. Le processus ne connaît pas la signification des éléments qu'il manipule. À partir d'un post-traitement, les différentes distances entre ces éléments pourraient fournir certains indices sur le style d'archi-

tecture de la façade. Ces indices pourraient alors par exemple guider par la suite le choix de grammaires de formes.

Compression Nous avons vu que la segmentation pouvait s'orienter vers des problématiques de compression. Cela fournirait au processus un objectif qui lui permettrait de s'affranchir de certains seuils. Schématiquement, la taille d'un modèle planaire serait l'unité. Nous pourrions utiliser une régression du profil des modèles cylindriques ainsi qu'une transformée fréquentielle des noyaux périodiques. Il serait alors envisageable de concevoir une distance normalisée dont nous avons parlé dans les perspectives de perfectionnement général de la segmentation. La figure 3.24 présente un résultat de compression.



FIGURE 3.24 – a) Image originale ; b) Segmentation hiérarchique (région supérieure de la figure 3.16 avec une représentation synthétique des modèles détectés ; régions trop petites barrées en rouge ; périodes signalées par des arcs rouges.

Appariement Une technique classique d'appariement de façades consiste en l'utilisation de points *SIFT* (par exemple Bris et Paparoditis [2010], figure 2.8.a). Mais leur emploi rencontre les mêmes difficultés relevées dans la segmentation qui vient d'être présentée. Leur appariement est perturbé par la répétition des structures, les changements d'éclairage entre les points de vue et les petits effets de parallaxe qui délocalisent les points. Une fenêtre située face à l'appareil d'acquisition peut ainsi tout à fait comporter des *SIFT* très similaires à ceux d'une fenêtre différente mais située elle aussi face à l'appareil sur une autre image. Or les *SIFT* de cette même fenêtre peuvent être beaucoup moins bien corrélés avec ceux d'une autre fenêtre, identique cette fois, située sur la même image mais pas dans l'axe du capteur et sujet à quelques effets de perspective. Ce phénomène peut embrouiller le processus d'appariement. Ceci sans parler des autres perturbations liées aux sources de mouvement de la façade (contrevents, pots de fleurs, reflets, etc.).

Il serait intéressant d'utiliser les premières découpes générées par la segmentation comme primitive d'indexation pour apparier deux images de façades (idée schématisée figure 1.1.c). Leur stabilité assurerait la robustesse du processus. Il est également envisageable de coupler cette approche descendante (*top-down*) à des indices plus locaux comme des ensembles de points d'intérêt (de type *SIFT* par exemple) contenus dans les premières régions. L'appariement de deux images porterait non plus entre des points isolés mais entre des regroupements cohérents de points.

Ajout de réalisme dans les modèles 3D urbains La partition de l'image, à laquelle aboutit la segmentation, se compose de régions cohérentes avec les structures de la façade. Elle peut apporter un réalisme très intéressant aux modèles 3D déjà existants. Il est par exemple tout à fait possible d'extraire, à partir d'un tel modèle et des régions de la segmentation hiérarchique, des parallélépipèdes rectangles dont la profondeur est la distance de la médiane des points 3D au plan de la façade (figures 3.25, 3.26, 3.27 et 3.28).



FIGURE 3.25 – Segmentation hiérarchique. Découpes en jaune. Modèles en rouge.

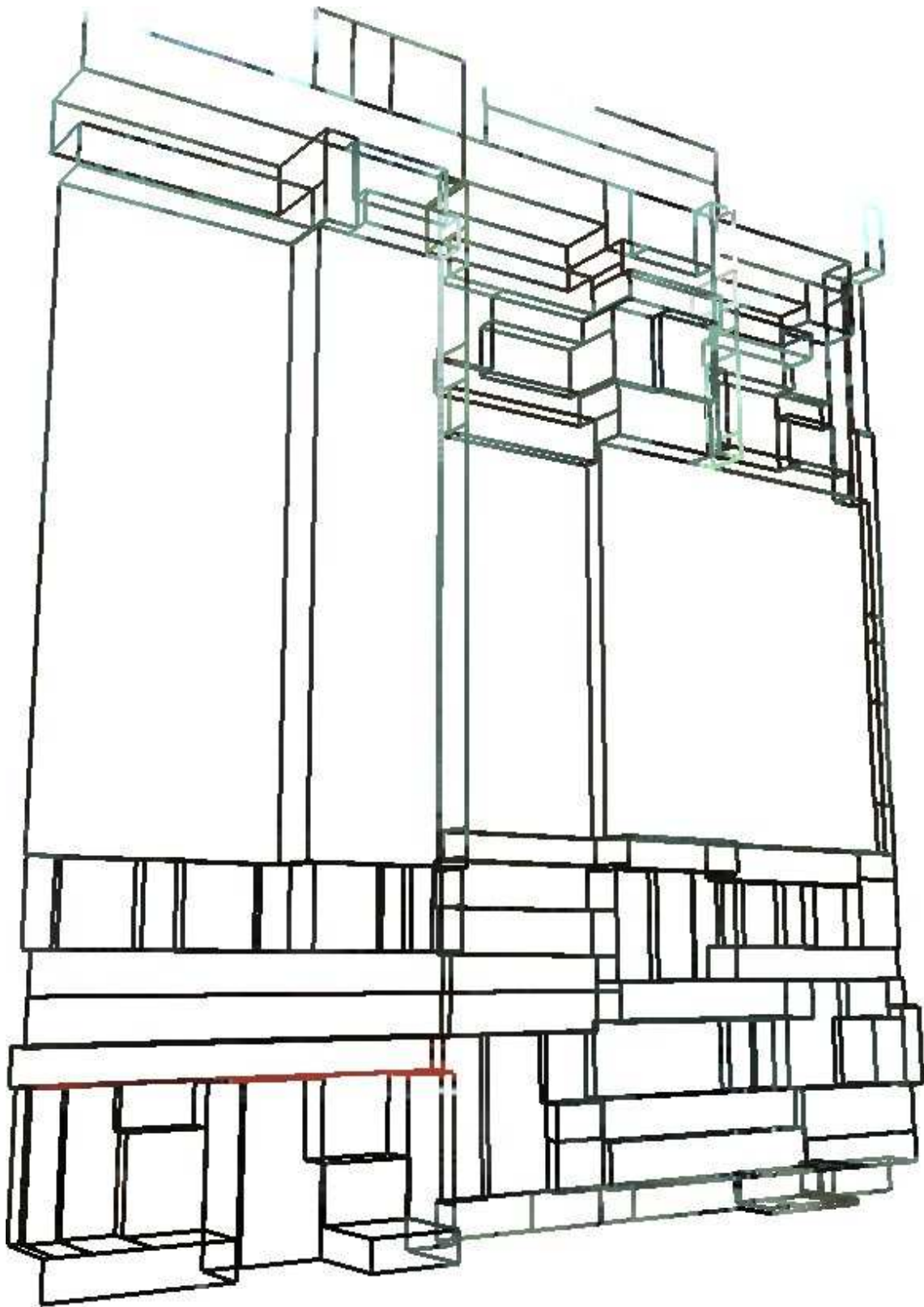


FIGURE 3.26 – Modélisation de la façade en facettes 3D à partir de la segmentation hiérarchique de la figure 3.25 et d'un nuage de points géoréférencés.



FIGURE 3.27 – Remplissage des facettes de la figure 3.26 par la radiométrie des régions de la segmentation hiérarchique de la figure 3.25.



FIGURE 3.28 – Vue de profil du résultat de la figure 3.27.

Grilles de structures répétitives

Nous avons présenté, au chapitre 3, une segmentation hiérarchique en modèles élémentaires 2D de la façade. Il s'agit d'une partition récursive de toute l'image. Elle se heurte au manque d'information globale de périodicité des structures de la façade et elle a une certaine tendance à séparer celles qui sont semblables avant de les comparer. La propriété de répétitivité des structures de la façade est cependant une propriété largement répandue qui constitue un enjeu crucial dans la littérature (Wenzel et al. [2008], Zeng et Gool [2008], Wendel et al. [2010], Müller et al. [2007]).

Il est très difficile de localiser un élément conceptuel donné, isolément, sur une façade mais il est possible de localiser des régions d'intérêt en utilisant la redondance d'occurrences de ces éléments. Ces régions d'intérêt pourraient alors être utilisées pour guider notre première approche afin d'éviter de séparer les périodicités dans la segmentation.

Nous proposons dans ce chapitre une détection des structures répétitives principales en posant certaines hypothèses sur leur localisation dans la façade et sur leur topologie. La section 4.1 définit ces objets : il ne s'agit pas nécessairement de fenêtres ou d'éléments identifiés *a priori*, mais juste de structures répétitives. Nous étudions alors une mesure de pertinence de leur localisation dans les images (section 4.2) : les structures doivent être similaires entre elles et contraster avec un fond de mur de radiométrie relativement uniforme. La section 4.3 évalue ensuite cette mesure de pertinence. Nous donnons enfin une conclusion, en section 4.4, pour synthétiser les points forts et les limites de l'approche et pour proposer quelques perspectives. Toute cette étude sert de base à la segmentation stochastique présentée dans le chapitre suivant (chapitre 5).

4.1 Définition des grilles de structures répétitives

Pour détecter les structures répétitives de la façade, nous les contraignons en des alignements réguliers selon la définition 4.1 (figure 4.1).

Définition 4.1 Une grille $\omega_{(m,n)}$ de m lignes et n colonnes de structures répétitives se compose des paramètres suivants :

- Les coordonnées (x, y) du coin supérieur gauche de la structure supérieure gauche de la grille ;

- Le gabarit (ν_l, ν_h) de la largeur et la hauteur des structures ;
- Le gabarit (e_l, e_h) de la largeur et la hauteur de l'espace entre les structures.

$$\omega_{(m,n)} = (x, y, \nu_l, \nu_h, e_l, e_h)$$

L'ensemble de ces grilles inscrites dans l'image forme l'espace $\Omega_{(m,n)}$.

Nous appelons nœud (ν) , les structures répétitives, et fond $(\psi_{\omega_{(m,n)}})$, la région complémentaire de ces nœuds sur la grille (définition 4.2).

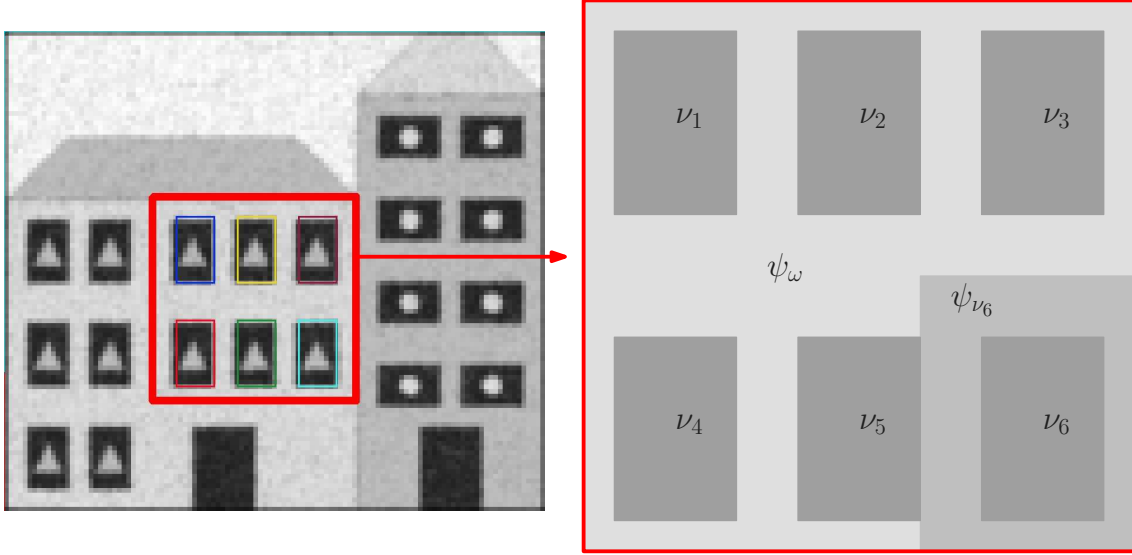


FIGURE 4.1 – Grille de structures répétitives sur une façade synthétique : ici, la grille $\omega_{(2,3)}$ possède six nœuds ν_i (en gris foncé) sur un fond de mur $\psi_{\omega_{(2,3)}}$ (en gris clair) ; la zone ψ_{ν_6} (en gris intermédiaire) correspond au voisinage du nœud ν_6 .

Notons que ces structures répétitives ne sont associées à aucun *a priori* sémantique : il peut s'agir d'éléments de fenêtre comme de motifs du mur suffisamment ressemblants et contrastés. Elles doivent se trouver inscrites dans une surface qui les englobe et que nous appelons *fond de mur*. Nous avons fixé empiriquement les limites de cette surface selon la définition 4.2.

Définition 4.2 *Le fond de mur ψ d'une grille ω de structures répétitives est la région complémentaire aux nœuds de celle-ci. Sa surface se limite à gauche et à droite par deux marges de largeur $\psi_l = \min(\frac{\nu_l}{2}, \frac{e_l}{2})$ ainsi, en haut et en bas, que par deux marges de hauteur $\psi_h = \min(\frac{\nu_h}{2}, \frac{e_h}{2})$.*

La définition 4.3 donne la structure de l'espace de toutes les grilles de structures répétitives.

Définition 4.3 *L'espace Ω est l'ensemble de toutes les grilles de structures répétitives inscrites dans l'image.*

$$\Omega = \bigcup_{m=m^{\min}}^{m^{\max}} \bigcup_{n=n^{\min}}^{n^{\max}} \Omega_{(m,n)}$$

où $(m^{\min}, m^{\max})$ et $(n^{\min}, n^{\max})$ sont les bornes respectives des nombres de lignes et de colonnes (entiers naturels non nul).

4.2 Critères de pertinence d'une grille de structures répétitives

Nous recherchons des grilles de structures répétitives dans la façade selon trois critères : la similarité des nœuds, leur contraste avec le fond de mur et l'uniformité de ce dernier. Cette formalisation est tout à fait empirique et vise à tenir compte d'un maximum d'éléments répétitifs de la façade en introduisant le moins possible d'*a priori*.

Cette section propose d'abord une définition des critères de pertinences (sous-section 4.2.1), puis expose la manière dont nous les utilisons conjointement (sous-section 4.2.2)) et se termine par les détails d'optimisation de nos calculs (sous-section 4.2.3).

4.2.1 Définition des critères de pertinence

Nous voulons mesurer l'adéquation $E_G(\omega)$ de la grille ω aux critères listés en définition 4.4.

Définition 4.4 *L'énergie $E_G(\omega)$ d'une grille ω de structures répétitives est la combinaison linéaire des termes suivants :*

- *Similarité, $E_{\text{similarité}}$, de la radiométrie i_ν des nœuds (définition 4.5) ;*
- *Distance, $E_{\text{contraste}}$, entre la radiométrie i_ν des nœuds et celle du fond de mur, i_ψ (définition 4.6) ;*
- *Uniformité, $E_{\text{uniformité}}$, du fond de mur (définition 4.7).*

$$E_G(\omega) = \frac{1}{3} [\alpha E_{\text{similarité}}(i_\nu) + \beta E_{\text{contraste}}(i_\nu, i_\psi) + \gamma E_{\text{uniformité}}(i_\psi)]$$

où α , β et γ sont les coefficients de pondération (cf sous-section 4.2.2).

Les coefficients α , β et γ doivent contraindre chaque terme dans le même intervalle de valeurs. Ce sont eux qui rendent possible l'addition des trois critères fondamentalement indépendants. Leur calcul est exposé dans la sous-section suivante (sous-section 4.2.2). La combinaison linéaire est une formalisation très simple à interpréter pour quantifier l'adéquation à ces mêmes critères.

4.2.1.1 Similarité des nœuds

La définition 4.5 donne le calcul de l'énergie $E_{\text{similarité}}$. Schématiquement, elle est égale au nombre de nœuds similaires (figure 4.2).

Définition 4.5 *L'énergie $E_{\text{similarité}}(i_\nu)$ de la radiométrie i_ν des nœuds est égale au coefficient de similarité multidimensionnel (CMI, décrit par Paparoditis [2007]).*

$$E_{\text{similarité}}(i_\nu) = \text{CMI}(i_\nu) = \frac{\sigma_{\sum_\nu i_\nu}^2}{\sum_\nu \sigma_{i_\nu}^2}$$

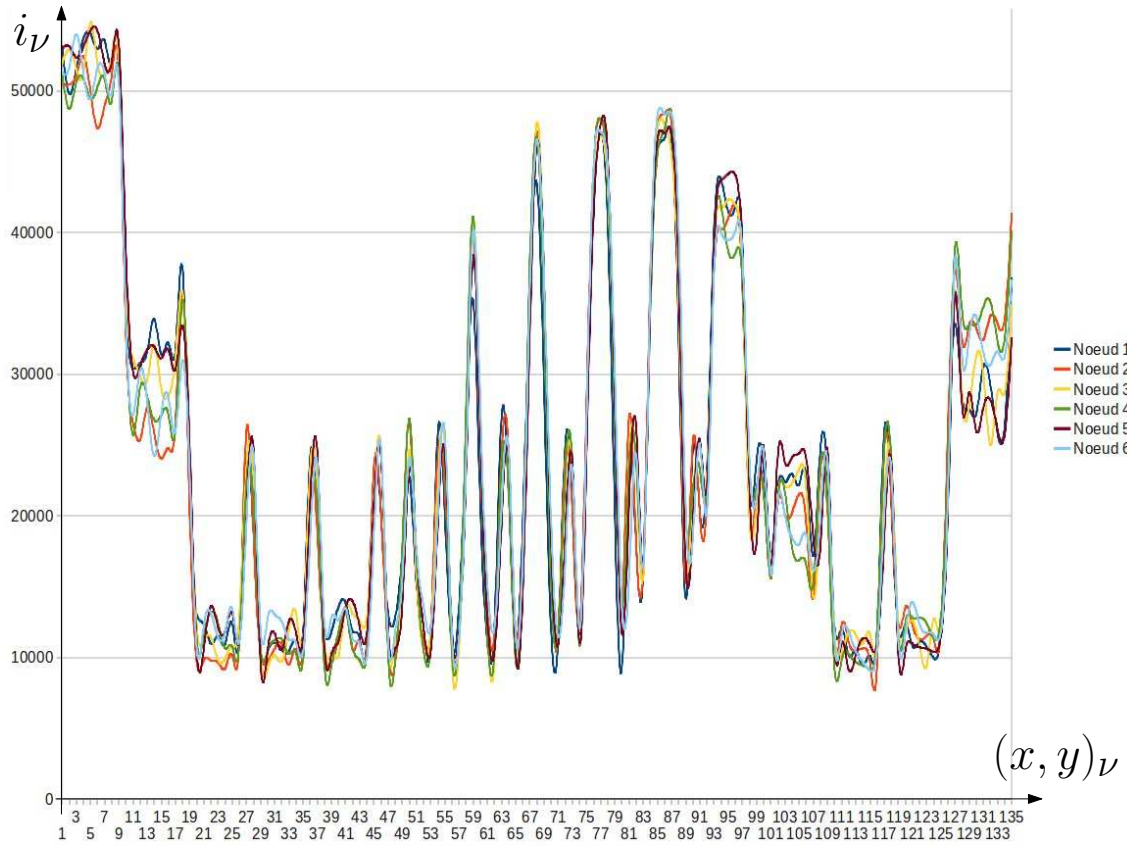


FIGURE 4.2 – Comparaison du profil radiométrique des six nœuds synthétiques de la figure 4.1. Le CMI vaut 5,89 ; soit une similarité de 98%.

Il peut arriver qu'un nœud ne soit pas éclairé d'une manière uniforme et que les variations de sa radiométrie soient beaucoup plus importantes que celle de ses voisins comme sur la figure 4.5.a. Pour homogénéiser les variations des nœuds, nous normalisons leur radiométrie avant de les sommer (variance de chacune égale à 1).

Notons quatre propriétés importantes de ce score concernant n nœuds :

Valeurs comprises entre 0 et n Le score tend vers n lorsque la radiométrie des nœuds suit, à une constante près, la même distribution. Il est plus bas autrement (et se rapproche de 0 lorsque les nœuds sont corrélés négativement).

Valeurs non significatives si $n < 3$ D'une manière générale, plus le nombre de nœuds comparés est important, plus le CMI est significatif. La figure 4.3 présente différents exemples de résultats pour $n = 3$ et $n = 5$.

Indépendance aux espérances Cette indépendance garantit une certaine stabilité aux conditions d'éclairage : pour n nœuds identiques, même si certains sont plus éclairés que d'autres et donc que leurs radiométries moyennes diffèrent, le score de similarité sera tout de même maximal. En revanche, un cas limite se produit lorsque les variances radiométriques sont négligeables : quelle que soit la radiométrie des nœuds (uniforme dans ce cas), le CMI sera élevé.

Pénalisation par les faibles variances Si m nœuds ont une variance radiométrique dominante par rapport à celle des autres, alors seules leurs radiométries ont une influence et la borne supérieure tend vers m .



FIGURE 4.3 – Comparaison de différentes imagerie en utilisant le coefficient de similarité multi-dimensionnelle (CMI). $CMI(F1, F2, F3) = 2,55$ (85%), $CMI(F1, F2, S1) = 1,99$ (66%), $CMI(F1, F2, F3, S1, S2) = 3,10$ (62%).

Afin de réduire l'influence des petites différences locales (dus aux problèmes de redressement ou simplement à la nature des nœuds), les signaux sont sous-échantillonnés. Cela permet également de réduire le temps de calcul.

4.2.1.2 Contraste des nœuds sur le mur

Pour calculer l'énergie $E_{\text{contraste}}$, nous utilisons une simple différence de moyennes radiométriques selon la définition 4.6.

Définition 4.6 L'énergie $E_{\text{contraste}}$, de contraste des nœuds sur le fond de mur, est la différence de la radiométrie moyenne $\bar{i}_\nu$ de nœuds et de la radiométrie moyenne $\bar{i}_\psi$ dans leur voisinage sur le fond de mur, i_ψ .

$$E_{\text{contraste}}(i_\nu, i_\psi) = |\bar{i}_\nu - \bar{i}_\psi|$$

Pour un nœud donné, nous ne considérons pas la radiométrie totale du fond de mur mais seulement celle comprise dans un voisinage. Ce voisinage est calculé de manière à couvrir une surface suffisamment représentative du fond de mur et en ayant une largeur identique dans les deux dimensions (la zone en gris intermédiaire sur la figure 4.1 pour le nœud ν_6). Cette dernière précaution permet de ne pas accorder trop de poids à la radiométrie du fond de mur dans une dimension par rapport à l'autre dans le cas de grilles comprenant de grands espacements entre des alignements de nœuds très rapprochés. La figure 4.4 illustre ce problème. Le score d'une grille peu contrastée est comparé à celui d'une grille très contrastée : la mesure de contraste est biaisée de manière significative si elle est normalisée.

Il arrive que certains objets soient plus éclairés que d'autres, ce qui implique une amplification de leurs contrastes. Il est courant dans la littérature de compenser ce phénomène. Liu et Gagalowicz [2010], par exemple, ajoutent un coefficient de pénalité en fonction de la hauteur sur la façade des objets considérés, le rez-de-chaussée étant souvent plus sombre que le toit, beaucoup plus exposé au soleil. Nous ne souhaitons pas ajouter un tel *a priori*.

4.2.1.3 Uniformité du fond de mur

La définition 4.7 propose une mesure de l'uniformité du fond de mur.

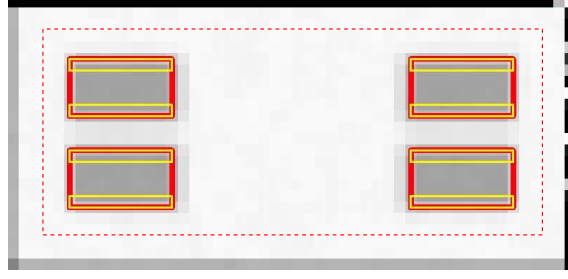


FIGURE 4.4 – Contraste des nœuds avec un voisinage ou directement le fond de mur global : en jaune une grille (4, 2) dont le contraste a une valeur unitaire de 24% s'il est calculé à partir du fond de mur global et 19% à partir du voisinage des nœuds (normalisés dans les deux dimensions) ; en rouge, la grille (2, 2) de référence dont le contraste vaut 29% (que le voisinage soit normalisé ou non).

Définition 4.7 L'énergie $E_{\text{uniformité}}$, d'uniformité du fond de mur ψ , mesure la dispersion de la radiométrie i_ψ de celui-ci.

$$E_{\text{uniformité}}(i_{\psi_\omega}) = \frac{1}{\sigma_{i_\psi}}$$

L'image d'entrée étant sous-échantillonnée, les détails fins comme les contours de briques ou les rainures du mur sont théoriquement supprimés, ce qui permet de supposer cette uniformité.

En pratique, nous utilisons plutôt deux variantes de cette énergie suivant les définitions 4.8 et 4.9.

D'autre part, pour évaluer l'uniformité ou le contraste, il est courant d'opérer une normalisation comme, par exemple la mesure de *Michelson* pour le contraste : $\frac{|i_\nu - i_\psi|}{i_\nu + i_\psi}$. Mais ce type de normalisation introduit un biais comme celui illustré sur la figure 4.5.b pour l'uniformité. Le contraste subit le même phénomène. Pour cette raison, nous ne normalisons pas ces deux mesures.



a)



b)

FIGURE 4.5 – Problème de variation d'éclairage. a) Grille de trois fenêtres identiques dont une située en plein soleil entre deux autres situées dans des zones d'ombre. Elles se ressemblent avec un *CMI* non normalisé de 51% et un *CMI* normalisé de 79%. b) Le score d'uniformité brut est deux fois plus élevé pour la région ν_1 que pour la région ν_2 mais il est plus faible pour la région ν_1 s'il est divisé par la moyenne.

4.2.2 Pondération des critères de pertinence

La comparaison entre les énergies de différentes grilles est rendue possible par l'ajustement des coefficients de pondération α , β et γ . Nous les estimons indépendamment par l'inverse des maxima des trois termes $E_{\text{similarité}}$, $E_{\text{contraste}}$ et $E_{\text{uniformité}}$ respectivement (équation 4.1).

$$\begin{cases} \alpha = \frac{1}{\max_{\omega \in \Omega} E_{\text{similarité}}(\omega)} \\ \beta = \frac{1}{\max_{\omega \in \Omega} E_{\text{contraste}}(\omega)} \\ \gamma = \frac{1}{\max_{\omega \in \Omega} E_{\text{uniformité}}(\omega)} \end{cases} \quad (4.1)$$

Les trois maxima globaux sont calculés dans l'espace Ω de toutes les grilles de structures répétitives lors d'une première passe (figure 4.6).

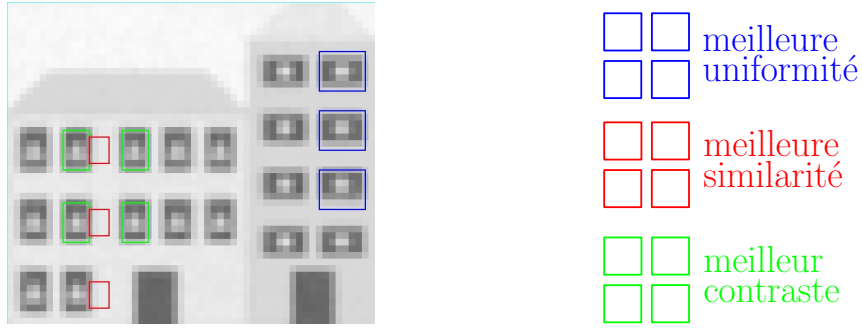


FIGURE 4.6 – Maximisations indépendantes des trois critères $E_{\text{similarité}}$ (en rouge), $E_{\text{contraste}}$ (en bleu) et $E_{\text{uniformité}}$ (en vert) sur l'image de la figure 4.1 sous-échantillonnée d'un facteur 10 dans les deux dimensions.

4.2.2.1 Choix du paramètre α

La figure 4.7 présente l'histogramme des mesures de similarité $E_{\text{similarité}}$ de l'image synthétique de la figure 4.6. Les valeurs de cette mesure s'étalent dans l'intervalle fixe $[0, 1]$. Les grilles de faible similarité sont les plus fréquentes. Celles de forte similarité sont peu nombreuses mais sont tout de même représentées dans l'image de manière significative (un certain nombre se situent évidemment dans les zones de radiométrie uniforme). C'est pour cette raison qu'en pratique nous ne recherchons pas à maximiser $E_{\text{similarité}}$: nous fixons directement *a priori* $\alpha = 1$.

4.2.2.2 Choix du paramètre β

L'intervalle de valeurs de l'énergie de contraste $E_{\text{contraste}}$ dépend de l'image et, contrairement au choix du paramètre α , il est nécessaire de rechercher le maximum pour obtenir sa normalisation dans l'intervalle $[0, 1]$. L'histogramme des mesures du contraste ainsi normalisé $\beta E_{\text{contraste}}$ est présenté sur la figure 4.8 pour la même image. La plupart des grilles ont un contraste nul ou faible et les plus forts contrastes se détachent très nettement des autres. La normalisation ne résout pas le risque que les grilles les plus nombreuses aient un score de contraste ainsi *écrasé* par un maximum si éloigné.

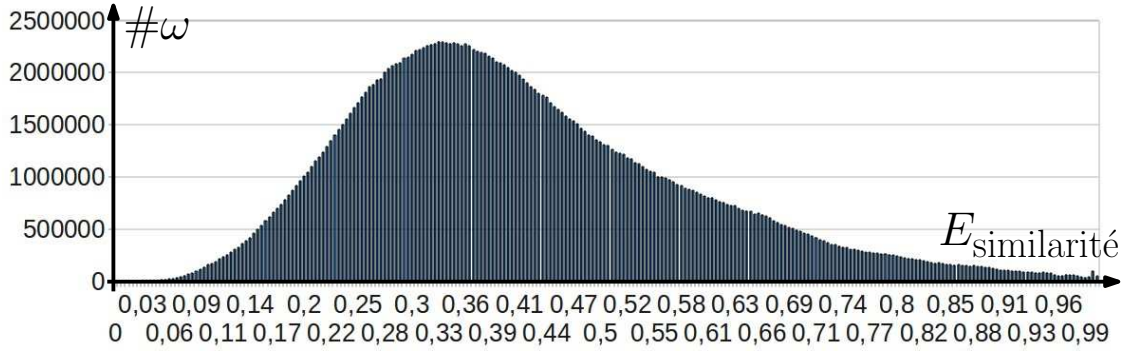


FIGURE 4.7 – Répartition des énergies brutes $E_{\text{similarité}}$ (définition 4.5) des grilles de structures répétitives sur toute l'image de la figure 4.6.

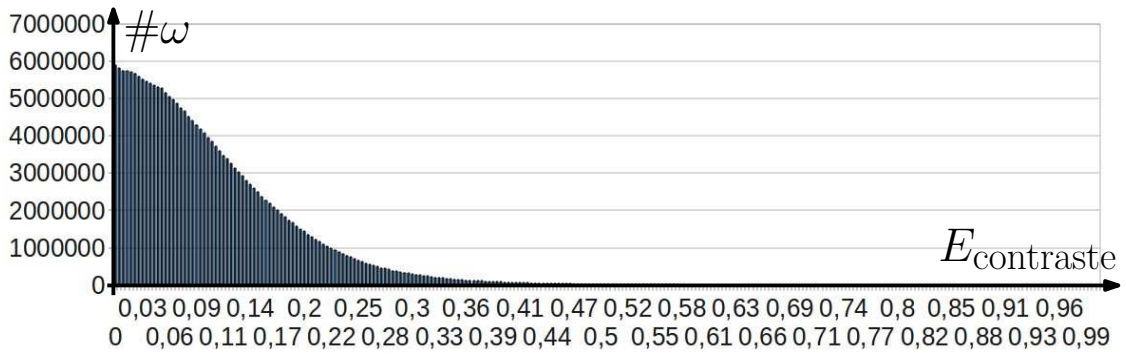


FIGURE 4.8 – Répartition des énergies $\beta E_{\text{contraste}}$ (définition 4.6) des grilles de structures répétitives, normalisées dans l'intervalle $[0, 1]$, sur toute l'image de la figure 4.6.

4.2.2.3 Choix du paramètre γ et adaptations de l'énergie d'uniformité

De même que pour l'énergie de contraste $E_{\text{contraste}}$, il est nécessaire de rechercher le maximum dans l'espace Ω pour obtenir la normalisation de l'énergie d'uniformité $E_{\text{uniformité}}$ du fond de mur dans l'intervalle $[0, 1]$. La figure 4.9 présente l'histogramme de ces mesures normalisées. Nous voyons que le phénomène d'écrasement des petites valeurs que nous constatons avec la répartition du contraste, se retrouve ici accentué à l'extrême. La plupart des grilles ont une uniformité très faible alors que certaines se détachent des autres de manière radicale.

Si nous laissons la normalisation telle quelle, cela signifie qu'à moins qu'elle soit très élevée, l'uniformité a une influence quasi-nulle sur l'énergie globale des grilles.

Pour compenser ce phénomène, il est possible d'utiliser une variante, $E_{\text{uniformité}}^{(1)}$, où un seuillage est effectué selon la définition 4.8.

Définition 4.8 L'énergie $E_{\text{uniformité}}^{(1)}$, d'uniformité du fond de mur ψ , mesure la dispersion de la radiométrie i_ψ de celui-ci en restant inférieure à un seuil $E_{\text{uniformité}}^{\max}$.

$$E_{\text{uniformité}}^{(1)}(\omega) = \min(E_{\text{uniformité}}(\omega), E_{\text{uniformité}}^{\max})$$

Le seuil $E_{\text{uniformité}}^{\max}$ peut être déterminé empiriquement ou bien grâce à une analyse de l'histogramme 4.6 (par exemple, 5% d'après cet histogramme).

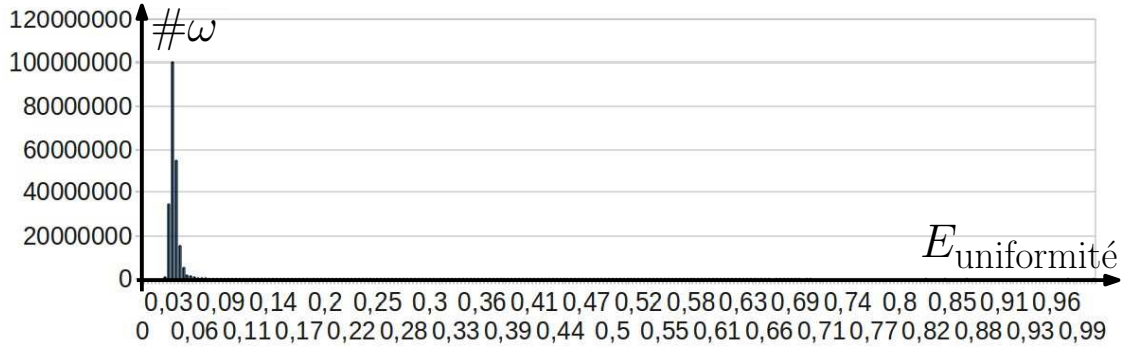


FIGURE 4.9 – Répartition des énergies $\gamma E_{\text{uniformité}}$ (définition 4.7) des grilles de structures répétitives, normalisées dans l'intervalle $[0, 1]$, sur toute l'image de la figure 4.6.

Il est également possible de calculer une variante $E_{\text{uniformité}}^{(2)}$ relativement à la dispersion de toute la grille de structures répétitives selon la définition 4.9.

Définition 4.9 L'énergie $E_{\text{uniformité}}^{(2)}$, d'uniformité du fond de mur ψ , mesure la dispersion de la radiométrie i_ψ de celui-ci relativement à celle, i_ω , de toute la grille ω .

$$E_{\text{uniformité}}^{(2)}(\omega) = \frac{\sigma_{i_\omega}}{\sigma_{i_\psi}}$$

La figure 4.10 présente l'histogramme de l'énergie $E_{\text{uniformité}}^{(2)}$. La répartition est différente : la quasi totalité des grilles ont un score très proche de $\gamma E_{\text{uniformité}}^{(2)}$ proche de 50%.

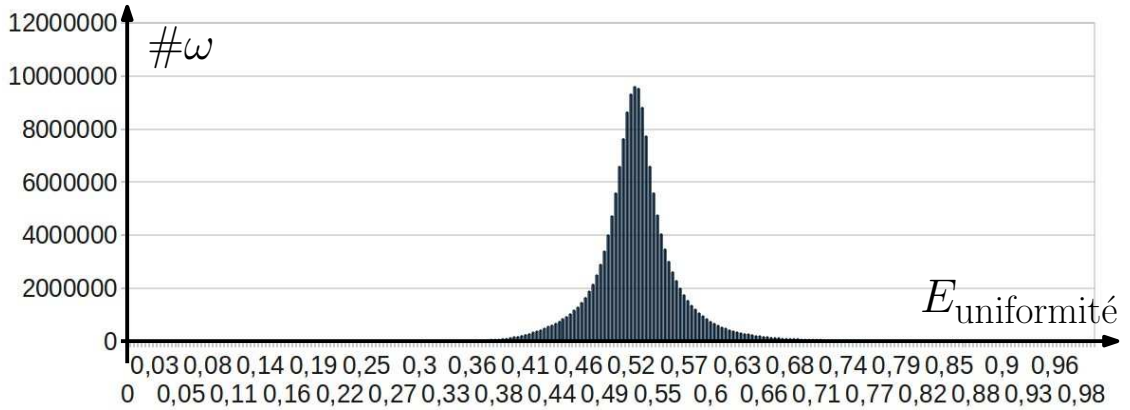


FIGURE 4.10 – Répartition des énergies $\gamma E_{\text{uniformité}}^{(2)}$ (définition 4.9) des grilles de structures répétitives, normalisées dans l'intervalle $[0, 1]$, sur toute l'image de la figure 4.6.

4.2.3 Optimisation des calculs

Le calcul des moyennes et variances radiométriques est effectué à partir d'images intégrales. Il s'agit d'une représentation de l'image de radiométrie permettant d'ob-

tenir ces informations extrêmement rapidement et indépendamment de la taille de la région rectangulaire considérée. L'usage de ce type d'outils est expliqué par Viola et Jones [2004]. Nous calculons les images intégrales en pré-traitement, ce qui est très rapide : leur dimension est la même que celle de l'image de radiométrie. Chaque pixel de l'image intégrale contient la somme des radiométries de tous les pixels au dessus et à gauche du pixel correspondant dans l'image d'entrée. Des opérations élémentaires d'addition et de soustraction de ces sommes permettent d'obtenir instantanément la moyenne et la variance radiométriques de n'importe quelle région rectangulaire de l'image. Ceci évite la lecture exhaustive de chaque pixel de la région.

De plus, nous utilisons les alignements dominants de l'image (définition 3.1), pour rechercher les maxima des trois critères lors de la première passe. L'espace de recherche $\Omega_{(m,n)}$ est alors réduit à l'espace $\hat{\Omega}_{(m,n)}$ décrit en définition 4.10 et nous obtenons trois approximations $\hat{\alpha}$, $\hat{\beta}$ et $\hat{\gamma}$ des coefficients de pondérations.

Définition 4.10 *La réduction $\hat{\Omega}_{(m,n)}$ de l'espace $\Omega_{(m,n)}$ des grilles de structures répétitives est obtenue à partir d'un quadrillage d'alignements dominants (définition 3.1) formant un ensemble Π de points d'intersection. Chaque triplet ordonné $((x_1, y_1), (x_2, y_2), (x_3, y_3))$ de ces points est associé à la grille $\omega_{(m,n)} = (x, y, \nu_l, \nu_h, e_l, e_h)$ selon le système 4.2.*

$$\begin{cases} x=x_1 \\ y=y_1 \\ \nu_l=x_2 - x_1 \\ e_h=y_2 - y_1 \\ e_l=\frac{x_3-x_1-n\nu_l}{n-1} \\ e_h=\frac{y_3-y_1-m\nu_h}{m-1} \end{cases} \quad (4.2)$$

Il s'agit d'une estimation assez grossière mais suffisante qui permet une réduction déterminante de la combinatoire du problème. Étant donné qu'il est possible que les maxima des énergies $E_{\text{similarité}}$, $E_{\text{contraste}}$ et $E_{\text{uniformité}}$ de l'espace $\Omega_{(m,n)}$ ne soient pas situés dans $\hat{\Omega}_{(m,n)}$, il faut utiliser l'approximation $\hat{E}_G$ de la compensation de l'énergie globale exprimée par la définition 4.11.

Définition 4.11 *L'énergie $\hat{E}_G$ d'une grille de structures répétitives est une estimation de l'énergie E_G à partir d'énergies $\hat{\alpha}E_{\text{similarité}}$, $\hat{\beta}E_{\text{contraste}}$ et $\hat{\gamma}E_{\text{uniformité}}$ normalisées dans l'espace $\hat{\Omega}_{(m,n)}$.*

$$\hat{E}_G = \frac{1}{3}[\min(1, \hat{\alpha} E_{\text{similarité}}) + \min(1, \hat{\beta} E_{\text{contraste}}) + \min(1, \hat{\gamma} E_{\text{uniformité}})]$$

4.3 Évaluation de l'énergie des grilles de structures répétitives

Cette section illustre à partir d'images synthétiques et réelles dans quelle mesure la détection de grilles de répétitions dans une image de façades est un problème bien posé :

- Existence d'une solution ;
- Unicité de la solution qui se *détache du lot* dans un voisinage donné ;
- Solutions dépendant de manière continue des données.

Nous proposons dans la suite le calcul de *cartes de scores optimaux des décalages* (définition 4.12) afin de déterminer dans quelle mesure notre problème est bien posé. Nous discutons alors sur les résultats obtenus.

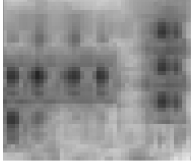
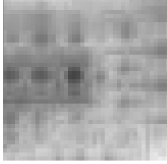
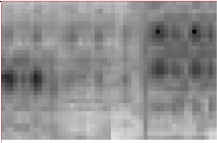
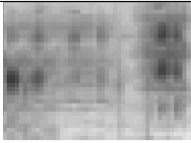
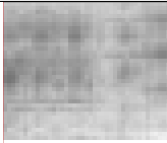
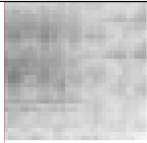
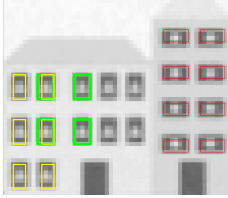
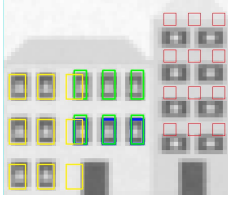
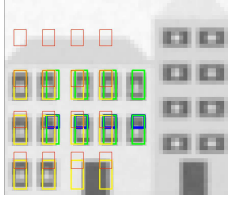
m	$n = 1$	$n = 2$	$n = 3$	$n = 4$
1			 81/93/4 → 59	 79/71/3 → 51
2		 89/100/12 → 67	 85/82/4 → 57	 72/73/3 → 49
3	 87/62/83 → 78	 82/94/5 → 60	 60/63/3 → 42	 51/63/3 → 39
4	 73/83/6 → 54	 76/86/6 → 56	 71/25/36 → 33	 61/28/3 → 31
				

Tableau 4.1 – Cartes des scores optimaux des décalages sur une image synthétique pour les espaces $(\Omega_{m,n})_{m,n \in [1,4]}$. En-dessous de chaque carte sont indiqués les scores optimaux sous le format $E_{\text{similarité}}/E_{\text{contraste}}/E_{\text{uniformité}} \rightarrow E_G$.

4.3.1 Cartes de scores optimaux des décalages

Pour visualiser le phénomène, nous projetons chaque espace $\Omega_{(m,n)}$ sur une carte $D_{(m,n)}^*$ des scores optimaux des décalages suivant la définition 4.12.

Définition 4.12 *Une carte des scores optimaux des décalages $D_{(m,n)}^*$ est une image de niveau de gris dont chaque pixel (x, y) a la valeur maximale de l'énergie E_G des grilles de décalage (x, y) .*

$$D_{(m,n)}^*(x, y) = \max_{(x, y, \nu_l, \nu_h, e_l, e_h) \in \Omega_{(m,n)}} E_G(x, y, \nu_l, \nu_h, e_l, e_h)$$

Le tableau 4.1 présente les cartes de scores optimaux des décalages pour les espaces $(\Omega_{m,n})_{m,n \in [1,4]}$. Les nœuds des grilles optimales correspondantes sont également représentés dans l'image au facteur d'échelle près (en dernière ligne du tableau). Sur les cartes, par soucis de clarté, le niveau de gris est représenté en négatif. Les pixels noirs correspondent alors aux énergies maximales. Les zones blanches périphériques correspondent aux zones où aucune grille ne peut être générée (en raison d'une taille minimale de gabarit).

4.3.2 Implémentation

Nous avons entrepris les calculs à partir de l'énergie $E_{\text{uniformité}}^2$ (définition 4.9) et de l'approximation $\hat{E}_G$ (définition 4.11). Il faut fixer la borne minimale des nœuds à une taille raisonnable (de l'ordre de la taille du plus petit objet possible). La taille maximale n'a généralement pas d'importance sur le résultat mais il faut veiller à ne pas la fixer trop grande pour éviter au processus des calculs inutiles.

Le tableau 4.1 a été obtenu par l'optimisation d'espaces $\Omega_{(m,n)}$. Le sous-échantillonnage étant fixé à un facteur 10 dans les deux dimensions, les cartes d'emplacements ont une résolution de (60, 50). Le temps d'exécution de leur calcul est de l'ordre de la minute sur un processeur de 2,83 GHz. Sur les images réelles, d'une résolution de 2 méga-pixels, nous avons effectué les calculs dans des espaces $\hat{\Omega}_{(m,n)}$ (définition 4.10) : durée de recherche de l'ordre de la seconde.

Les images de la figure 4.11 présentent quelques grilles de structures répétitives optimales (en vert) dans différents espaces $\hat{\Omega}_{(m,n)}$. Nous avons affiché (en rouge), dans ces même espaces, d'autres grilles, localement optimales, pour montrer la pertinence des résultats. Ces grilles localement optimales sont obtenus grâce aux cartes de scores optimaux des décalages : un opérateur clique sur le pixel de décalage et le processus affiche la grille optimale correspondante enregistrée en mémoire. Lorsque des structures (ici, principalement des fenêtres), se répètent de manière contrastée sur le fond de mur, elles sont correctement localisées par l'algorithme.

La figure 4.12 montre des grilles optimales de structures répétitives aux effets de perspective très marqués. Les fenêtres de la figure 4.12.a sont bien localisées mais les balcons de la figure 4.12.b le sont très mal à cause des effets d'ombre.

4.3.3 Pertinence de la formalisation

En considérant chaque carte $D_{(m,n)}^*$ (tableau 4.1), nous voyons qu'il existe à chaque fois une solution qui se démarque relativement nettement de son voisinage.

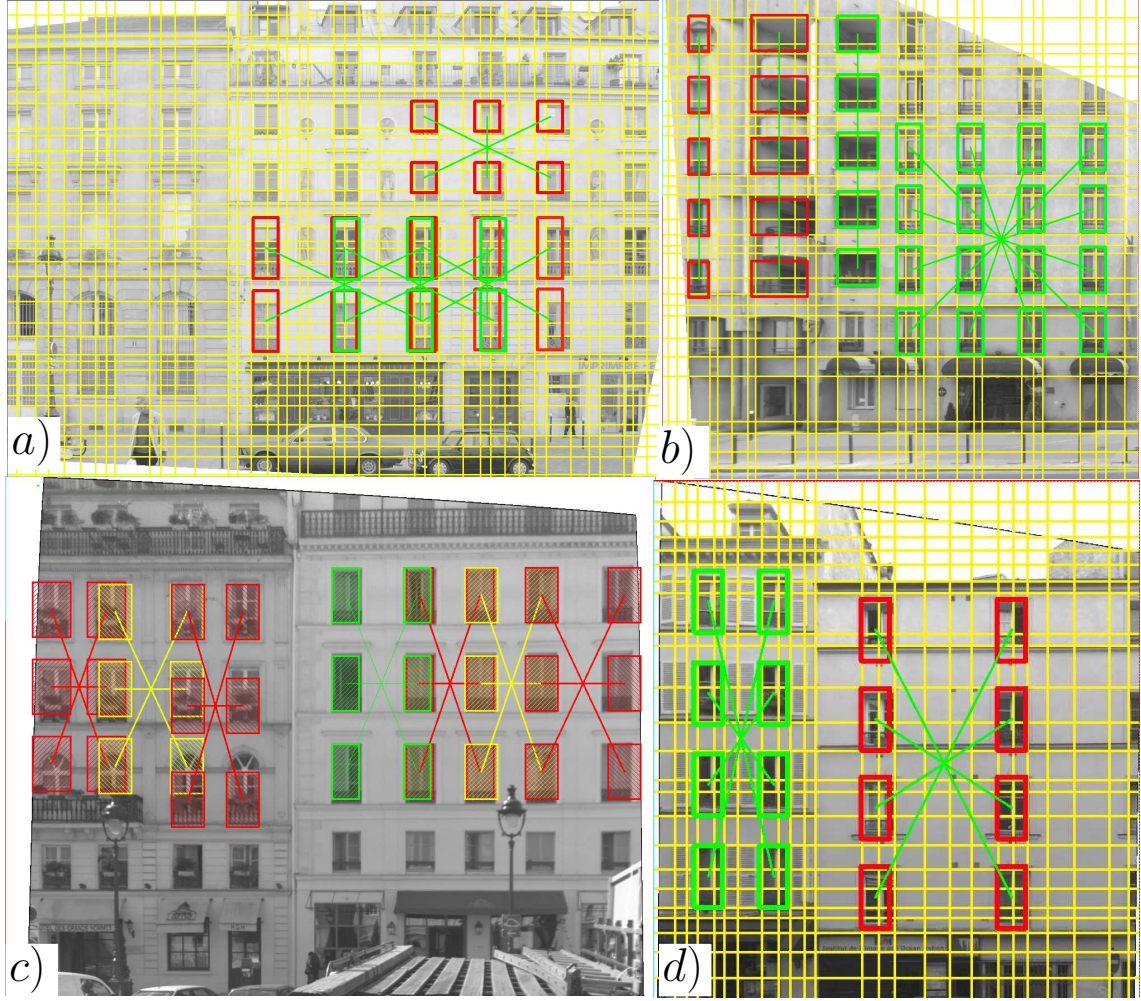


FIGURE 4.11 – Grilles optimales de structures répétitives (en vert : maxima absolus, en rouge : maxima locaux affichés manuellement) ; a) Optimisation de l'espace $\hat{\Omega}_{(2,3)}$; b) Optimisation des espaces $\hat{\Omega}_{(4,4)}$ et $\hat{\Omega}_{(5,1)}$; c) Optimisation des espaces $\hat{\Omega}_{(3,3)}$ et $\hat{\Omega}_{(3,2)}$; d) Optimisation de l'espace $\hat{\Omega}_{(4,2)}$.

Dans l'image considérée, nous voyons que les grilles optimales qui existent réellement obtiennent des scores nettement supérieurs aux autres : dans cet exemple, les 7 bonnes grilles ont un score moyen de 62% contre un score de 41% pour les 6 mauvaises grilles. De plus, ces cartes illustrent très explicitement que l'énergie E_G varie continûment avec le décalage (x, y) des grilles. Dans un espace $\Omega_{(m,n)}$ donné et sur l'image synthétique que nous avons utilisée, le problème est alors relativement bien posé.

Notons que l'optimisation n'est pas réalisable directement dans l'espace Ω (ou $\hat{\Omega}$) car nous avons constaté que les grilles d'un faible nombre de nœuds avaient tendance à avoir une énergie E_G (ou $\hat{E}_G$) supérieure à celles contenant davantage de nœuds.

D'une manière générale, l'approche présente deux limites principales :

Formalisation du contraste et de l'uniformité Ces deux critères ne reposent que sur des différences de moyennes radiométriques, pour le contraste $E_{\text{contraste}}$ (définition 4.6) et sur un rapport de variances pour l'uniformité $E_{\text{uniformité}}^{(2)}$

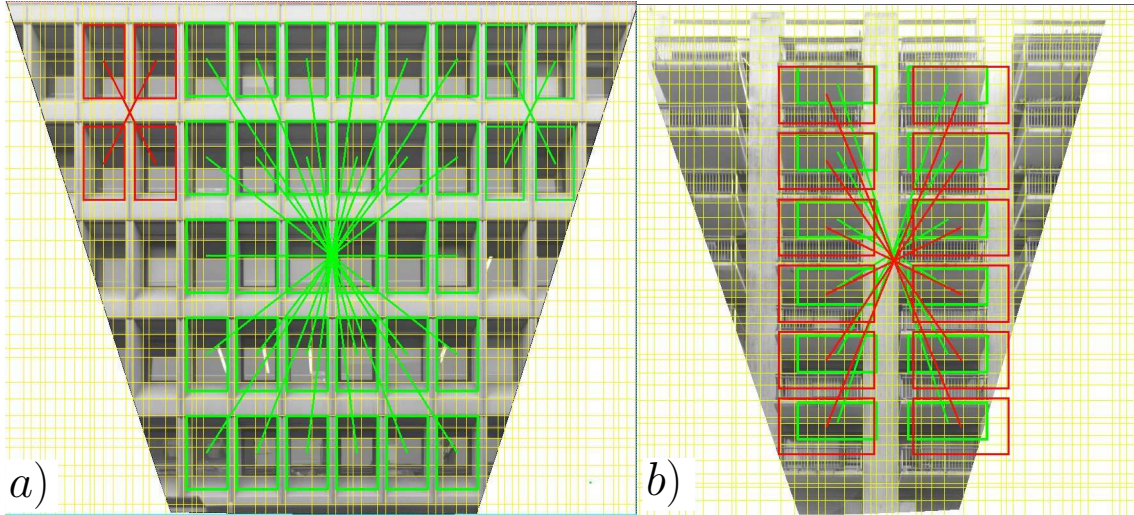


FIGURE 4.12 – Grilles optimales de structures répétitives aux effets de perspective très marqués ; a) Optimisation des espaces $\hat{\Omega}_{(5,6)}$ et $\hat{\Omega}_{(2,2)}$; b) Optimisation de l'espace $\hat{\Omega}_{(6,2)}$.

(définition 4.9). Ainsi, leur calcul, effectués via des images intégrales, ne dépend pas de la taille des nœuds et s'effectue de manière quasiment instantanée. Mais les résultats obtenus se heurtent fréquemment à cette simplification.

Échelle des alignements dominants L'échelle de l'image influence la localisation des alignements dominants. À petite échelle, des structures, même fortement redondantes, très contrastées et bien alignées, peuvent ne pas ressortir suffisamment dans le profil d'accumulation des gradients pour pouvoir être représentées dans les alignements dominants. Dans ce cas, l'espace $\hat{\Omega}_{(m,n)}$ n'est pas assez précis pour mener la recherche de toutes les structures.

La grille optimale de l'espace $\hat{\Omega}_{(3,2)}$, obtenue par une requête manuelle au niveau de la façade de gauche de l'image 4.11.c, n'est pas localisée autour des fenêtres. Cela est dû à la formalisation du contraste et de l'uniformité. De même pour la grille optimale de l'espace $\hat{\Omega}_{(5,1)}$, obtenue par requête manuelle sur l'image 4.13.d : les fenêtres ne sont pas localisées car les nœuds sont situés tout le long d'un très fort contraste. En revanche, sur la même image, le processus a placé la grille optimale de l'espace $\hat{\Omega}_{(2,2)}$ au niveau de quatre fenêtres identiques.

Les images 4.13.a et 4.13.b illustrent le problème des alignements dominants extraits à des échelles trop petites. S'il existe plusieurs façades dans l'image, le quadrillage des alignements dominants n'est pas suffisamment précis pour être un support adapté à un espace $\hat{\Omega}_{(m,n)}$. Beaucoup de structures répétitives ne sont alors même pas échantillonnées par le processus et évaluées dans des nœuds de grilles. Les images 4.13.c et 4.13.d, acquises à une échelle plus grande montrent que ce problème est soluble dans un espace $\hat{\Omega}_{(m,n)}$ plus précis.

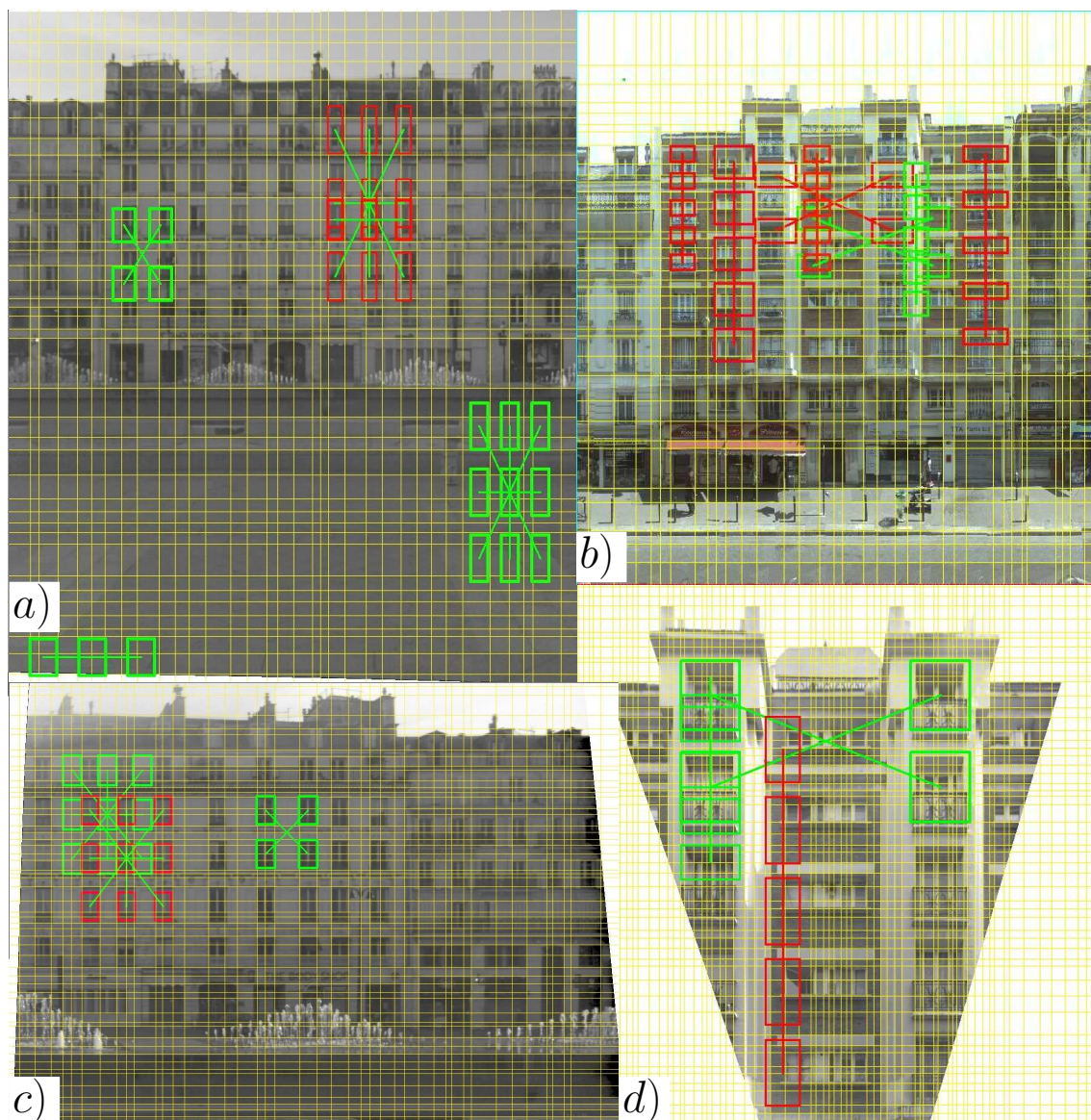


FIGURE 4.13 – Limite de l'espace $\hat{\Omega}_{(m,n)}$; a) et c) Comparaison à deux échelles différentes de l'optimisation des espaces $\hat{\Omega}_{(1,3)}$, $\hat{\Omega}_{(3,3)}$ et $\hat{\Omega}_{(2,2)}$; b) et d) Comparaison à deux échelles différentes de l'optimisation des espaces $\hat{\Omega}_{(2,2)}$ et $\hat{\Omega}_{(5,1)}$.

4.4 Conclusion

Nous venons d'analyser une énergie de pertinence de la localisation de grilles de structures répétitives dans les images de façade. Cette énergie est basée sur trois critères simples : la similarité des structures entre elles, leur contraste avec leur environnement et l'uniformité de ce dernier. Cette formalisation est tout à fait empirique pour tenir compte d'un maximum d'éléments répétitifs de la façade en introduisant le moins possible d'*a priori*. Nos résultats montrent que malgré la simplicité de la formalisation, il est possible de faire ressortir des structures intéressantes dans les images de manière semi-automatique.

Notre approche permet de générer des cartes de scores optimaux indiquant, pour chaque pixel de l'image, quelle est la grille de structures répétitives optimale, au sens de l'énergie de pertinence que nous avons définie. Une fois la carte générée, un opérateur, peut, instantanément, obtenir les boîtes englobantes d'éléments répétitifs de la façade où il le désire.

Les techniques d'optimisation que nous utilisons permettent d'accélérer considérablement le calcul de la-dite énergie. Les alignements dominants permettent de réduire la combinatoire de l'espace de recherche de manière conséquente. Les images intégrales permettent des calculs statistiques extrêmement rapides de la radiométrie des structures de la façade.

Dans un espace $\Omega_{(m,n)}$ de grilles de m lignes et n colonnes de structures répétitives, l'énergie E_G que nous analysons varie continûment avec le décalage (x,y) de ces grilles et les solutions localement optimales associées à de réelles structures répétitives se démarquent relativement nettement des autres.

La limite principale de cette approche est la longue configuration par un opérateur de l'espace de recherche : spécifier les bornes des paramètres à optimiser, décider dans quel espace $\Omega_{(m,n)}$ rechercher les objets (nombre de lignes et de colonnes) et naviguer dans ce dernier, via les cartes de scores optimaux, en cliquant aux emplacements qui lui semblent les plus pertinents.

Une autre faiblesse réside dans la formalisation des critères que nous utilisons. Le contraste est une différence de moyennes radiométriques et l'uniformité est un rapport de variances. Leur calcul est alors instantané mais les résultats obtenus se heurtent fréquemment à cette simplification. Une amélioration statistique du contraste pourrait être ajoutée à l'instar de ce que propose Lacoste et al. [2010] pour la détection de routes à partir d'images aériennes.

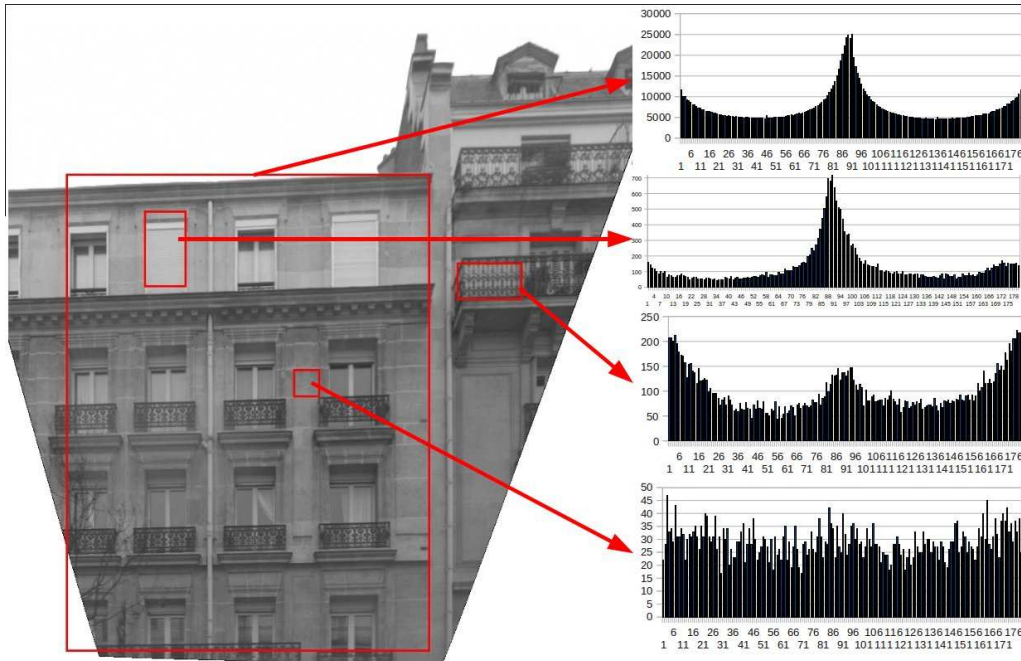


FIGURE 4.14 – Histogramme de l'orientation du gradient de certaines régions de la façade

Un autre mode de calcul de l'uniformité pourrait reposer sur des histogrammes d'orientations de gradient. Ces histogrammes pourraient être stockés dans des images intégrales (à 8 couches si nous restreignons le modèle à 8 orientations) (figure 4.14). L'approche, dont le calcul serait relativement rapide, serait certainement plus robuste.

Une troisième limite concerne l'approximation que nous faisons de restreindre la recherche aux grilles de structures répétitives dans les alignements dominants. Ces derniers, sur une image de trop faible échelle, ne sont pas assez précis et ne permettent pas d'échantillonner correctement l'espace de recherche.

D'autre part, en fonction du point de vue de l'acquisition, les éléments qui sortent du plan de la façade (en saillie comme les balcons ou en retrait comme les vitres) peuvent être délocalisés d'un nœud à l'autre. Pour compenser ces perturbations, nous pourrions ajouter un paramètre de profondeur p , aux grilles, pour estimer la profondeur des objets en arrière ou avant plan, comme illustré sur la figure 4.15. Une technique de ce type est utilisé par Van Gool et al. [2007]. Dans le même esprit que le concept d'*alignements dominants*, nous pourrions alors faire reposer la segmentation sur celui de *plans dominants*, parallèles entre eux. La façade se composerait d'un plan principal et d'autres plans annexes (pour les fenêtres en retrait et les balcons par exemple).

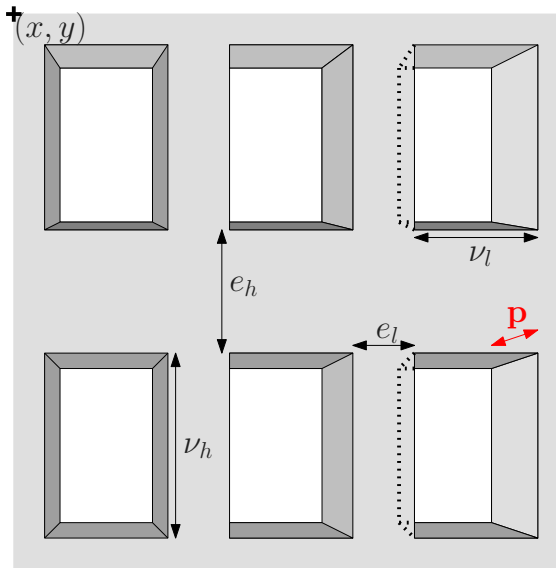


FIGURE 4.15 – Ajout du paramètre de profondeur p

Enfin, l'approche concerne des structures répétitives quelconques (leur seule contrainte est d'être contrastées sur un fond de mur uniforme). Il serait possible, dans une autre problématique, de l'appliquer à une détection de fenêtres proprement dite en ajoutant des *a priori* structurels dans la définition des nœuds. Le modèle pourrait par exemple supposer l'existence d'un cadre et de croisillons dont la surface serait supposée ne pas réfléchir la lumière ainsi que des vitres, dont la radiométrie ne serait alors pas à prendre en compte (pour éviter les reflets). Un critère d'apparence interne de fenêtre pourrait être défini à partir de l'aspect structuré des gradients de l'image (symétrie, forme en croix, etc.).

Pour conclure ce chapitre, rappelons que l'idée principale de notre approche est de détecter les structures répétitives de l'image. Nous venons de présenter une technique pour extraire isolément des grilles de ces structures. Il est également possible de les regrouper, de manière globale, dans un agencement optimal, de manière rapide et automatique. C'est ce que montre le chapitre suivant, à partir d'un *processus ponctuel marqué*.

Chapitre 5

Recherche du meilleur agencement de grilles de structures répétitives

Nous avons présenté et analysé une mesure de pertinence E_G (définition 4.4) d'une grille de structures répétitives (définition 4.1) au chapitre précédent (chapitre 4). Nous voulons à présent étendre cette mesure à tout un agencement de grilles. L'espace $\hat{\Omega}$ (définition 4.10) étant sujet à des problèmes d'échelle, nous ne l'utilisons pas pour optimiser l'agencement global.

Les définitions 5.1 et 5.2 donnent la structure des espaces de recherche $\mathcal{W}_k$ et $\mathcal{W}$ de ces agencements (le nombre de grilles k étant *a priori* inconnu).

Définition 5.1 *L'espace $\mathcal{W}_k$ est l'ensemble des agencements de k grilles de structures répétitives (définition 4.1).*

$$\mathcal{W}_k = \Omega^k$$

où k , placé en exposant, signifie le produit cartésien de k sous-espaces Ω .

L'espace Ω est l'ensemble des grilles de structures répétitives (définition 4.3).

Définition 5.2 *L'espace $\mathcal{W}$ est l'ensemble de tous les agencements de grilles de structures répétitives.*

$$\mathcal{W} = \bigcup_{k=0}^{\infty} \mathcal{W}_k$$

Ce chapitre propose la formalisation d'une énergie de coût associée à un agencement de grilles de structures répétitives (section 5.1). Pour trouver l'agencement optimal dans l'espace $\mathcal{W}$, cette énergie est minimisée (section 5.2) par le couplage d'un échantillonnage par *chaînes markoviennes de Monte Carlo à sauts réversibles* [Hastings, 1970; Green, 1995] et d'un *recuit simulé* (décrit par Salamon et al. [2002]). Enfin nous discutons sur les résultats (section 5.3) que nous avons déjà obtenus. Une conclusion présente en section 5.4 un bilan de l'avancement de ces recherches ainsi que les perspectives envisagées.

5.1 Énergie d'un agencement de grilles

La définition 5.3 propose le calcul d'une énergie $E_A(W)$ d'un agencement de grilles.

Définition 5.3 *L'énergie $E_A(W)$ d'un agencement W de grilles de structures répétitives est une combinaison linéaire d'une énergie d'attache aux données E_I (définition 5.4) et d'une énergie de régularisation E_R (définition 5.5).*

$$E_A(W) = E_I(W) + \xi E_R(W)$$

où $\xi \geq 0$ est un coefficient réglant l'influence relative des énergies d'attache aux données et de régularisation.

Cette section définit d'abord l'énergie d'attache aux données, E_I , relative à la pertinence de l'existence de chacune de ces grilles par rapport aux données de l'image. Elle définit ensuite l'énergie de régularisation, E_R , qui mesure la plausibilité *a priori* de la topologie des grilles de l'agencement, indépendamment des données de l'image. Elle se conclue par le calcul de la complexité de l'énergie E_A des agencements de grilles de structures répétitives.

5.1.1 Énergie d'attache aux données

L'énergie E_I d'attache aux données dépend de la somme des énergies de chaque grille (E_G , définition 4.4). Nous ne pouvons pas simplement la définir comme l'opposé de cette somme car, ainsi, elle ne permettrait pas au système de choisir de manière discriminante entre un agencement de grilles composées d'un grand nombre de nœuds et un agencement de plus petites grilles. La figure 5.1 présente ce problème. Neuf nœuds y sont représentés (en noir) et deux agencements A_0 (en gris) et A_1 (en rouge) sont proposés. Le premier se compose d'une unique grille ω_0 qui englobe les neuf nœuds. Le deuxième se compose de trois plus petites grilles ω_1 , ω_2 et ω_3 (en pointillés rouges). Dans les deux cas, le processus a regroupé chaque nœud de l'image dans une grille sans aucun chevauchement. Dans la situation idéale (les nœuds sont tous exactement identiques, ils sont contrastés de la même manière avec le fond de mur et ce dernier est parfaitement régulier), ces deux agencements possèdent la même énergie E_I .

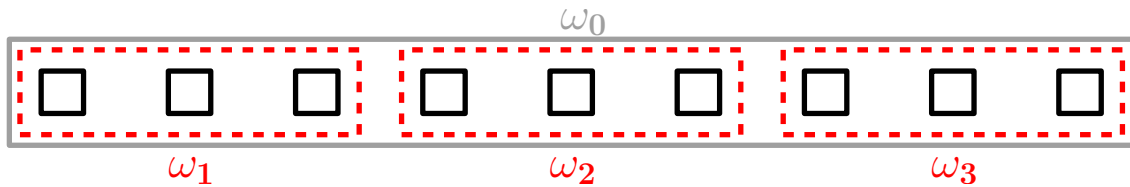


FIGURE 5.1 – Pénalisation de la sur-segmentation

Pour permettre au système de sélectionner l'un ou l'autre, nous avons choisi de privilégier les grandes grilles selon la définition 5.4.

Définition 5.4 *L'énergie d'attache aux données, E_I , d'un agencement W de grilles de structures répétitives, est l'opposé de la somme des énergies de chaque grille auxquelles sont ajoutées une constante de seuil $E_G^{\min} > 0$.*

$$E_I(W) = \sum_{\omega \in W} E_G^{\min} - E_G(\omega)$$

Il s'agit d'un critère de moindre longueur de description dans le même esprit que le critère *Minimum Description Length* (MDL, Rissanen [1983]).

Ici, $E_I(W_0) = E_G^{\min} - E_G(\omega_0)$ et $E_I(W_1) = 3E_G^{\min} - E_G(\omega_1) - E_G(\omega_2) - E_G(\omega_3)$. Dans l'hypothèse du cas idéal où $E_G(\omega_0) = E_G(\omega_1) + E_G(\omega_2) + E_G(\omega_3)$, c'est bien l'énergie $E_I(W_0)$ de l'agencement de moindre complexité qui est la plus faible. En pratique, le choix de $E_G^{\min}$ est effectué empiriquement.

5.1.2 Énergie de régularisation

Les grilles de structures répétitives sont échantillonnées dans l'espace de recherche indépendamment les unes des autres. Le terme de régularisation E_R permet de joindre à l'énergie totale une information liée à la politique d'agencement de ces grilles. Ce type d'information est indispensable, ne serait-ce que pour éviter au système de se bloquer en superposant indéfiniment des grilles au même endroit. La définition 5.5 propose le calcul de cette énergie en sommant des énergies Θ de régularisation binaire.

Définition 5.5 *L'énergie de régularisation, E_R , d'un agencement W de grilles de structures répétitives, est la somme de toutes les énergies de régularisation binaires $\Theta(\omega_1, \omega_2)$ entre les couples de grilles.*

$$E_R(W) = \sum_{\omega_1 \in W} \sum_{\omega_2 \in W - \{\omega_1\}} \Theta(\omega_1, \omega_2)$$

En fonction de l'application visée, nous proposons trois définitions de la régularisation binaire *Theta*.

5.1.2.1 Pénalisation de chevauchements des boîtes englobantes

La politique d'agencement la plus élémentaires est la pénalisation de la surface d'intersection des boîtes englobantes des grilles (figure 5.2) : définition 5.6.

Définition 5.6 *L'énergie de régularisation binaire $\Theta_1(\omega_1, \omega_2)$ entre les grilles de structures répétitives ω_1 et ω_2 est la surface d'intersection des grilles.*

$$\Theta_1(\omega_1, \omega_2) = \mathcal{A}(\omega_1 \cap \omega_2)$$

où $\mathcal{A}$ est l'aire d'une surface. Notons que la surface d'une grille comprend celle de son fond de mur (définition 4.2).

Ce type de régularisation permet une partition rigoureuse de l'image de façade et pourrait par exemple être orientée vers de la compression. Cette politique est appliquée par Tournaire et al. [2010] pour l'agencement d'empreintes de cadastres.

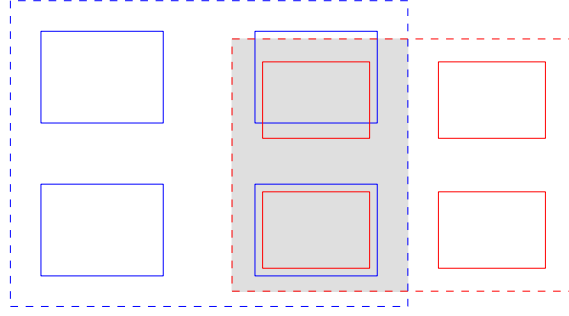


FIGURE 5.2 – Pénalisation des chevauchements de fonds de mur : la surface grisée est l'énergie E_{R_1} (définition 5.6).

5.1.2.2 Pénalisation de chevauchement des nœuds

Dans notre problématique, une politique un peu plus fine que celle de Θ_1 pourrait écarter de la pénalisation les éléments présents sur le fond de mur (décorations qui se répètent tout le long de la façade par exemple, et qui se retrouve sur le fond de mur de différentes grilles) 5.7. La figure 5.3 schématise cette mesure.

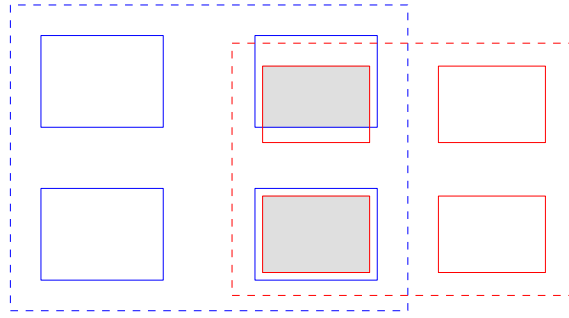


FIGURE 5.3 – Pénalisation de chevauchement des nœuds : la surface grisée est l'énergie E_{R_2} (définition 5.7).

Définition 5.7 L'énergie de régularisation binaire $\Theta_2(\omega_1, \omega_2)$ entre les grilles de structures répétitives ω_1 et ω_2 est la surface d'intersection de leurs nœuds.

$$\Theta_2(\omega_1, \omega_2) = \sum_{\nu_1 \in \omega_1} \sum_{\nu_2 \in \omega_2} \mathcal{A}(\nu_1 \cap \nu_2)$$

où $\mathcal{A}$ est l'aire d'une surface.

De même que pour la régularisation Θ_1 cette régularisation peut s'appliquer à de la compression.

5.1.2.3 Superpositions des nœuds

Il est également possible de permettre des *liaisons faibles* entre les grilles par la superposition de leurs nœuds (figure 5.4).

Dans ce cas, pour deux nœuds ν_1 et ν_2 donnés, il faut distinguer deux situations : soit ils sont *indépendants* soit ils sont *liés*. La définition 5.8 distingue ces deux concepts.

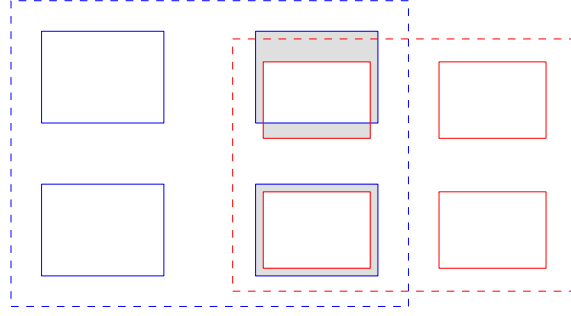


FIGURE 5.4 – Superposition de nœuds : la surface grisée est l'énergie E_{R_3} (définition 5.9).

Définition 5.8 Deux nœuds ν_1 et ν_2 sont dits liés si et seulement si leur différence symétrique est inférieure à leur intersection.

$$\mathcal{A}(\nu_1 \Delta \nu_2) < \mathcal{A}(\nu_1 \cap \nu_2)$$

où $\mathcal{A}$ est l'aire d'une surface.

Ils sont dits indépendants le cas échéant.

Nous parlons alors de *liaison faible* car, une fois deux grilles liées l'une à l'autre par un ou plusieurs de leurs nœuds, indépendamment de l'énergie d'attache aux données E_I , le système doit fournir une énergie plus forte pour les séparer que pour les laisser liées.

D'autre part, il est important de rejeter les chevauchements simultanés de tous les nœuds afin d'empêcher le système de placer toutes les grilles au même endroit. Cela se produit lorsque tous les nœuds ν_1 de la grille ω_1 chevauchent un nœud ν_2 de la grille ω_2 (aucune surface d'intersection $\max_{\nu_2 \in \omega_2} \mathcal{A}(\nu_1 \cap \nu_2)$ n'est nulle). Dans ce cas, nous fixons $\Theta_1 = \Theta_2$ pour pénaliser fortement ce chevauchement.

Regroupant ces deux considérations, la définition 5.9 donne le calcul de l'énergie de régularisation binaire Θ_3 .

Définition 5.9 L'énergie de régularisation binaire $\Theta_3(\omega_1, \omega_2)$ entre les grilles de structures répétitives ω_1 et ω_2 pénalise les nœuds indépendants qui se chevauchent et privilégie les nœuds liés (définition 5.8). Elle pénalise la superposition simultanée de tous les nœuds.

$$\Theta_3(\omega_1, \omega_2) = \begin{cases} \Theta_2(\omega_1, \omega_2) & \left| \min_{\nu_1 \in \omega_1} \max_{\nu_2 \in \omega_2} \mathcal{A}(\nu_1 \cap \nu_2) \right| > 0 \\ \sum_{\nu_1 \in \omega_1} \sum_{\nu_2 \in \omega_1} \min(\mathcal{A}(\nu_1 \cap \nu_2), \mathcal{A}(\nu_1 \Delta \nu_2)) & \left| \min_{\nu_1 \in \omega_1} \max_{\nu_2 \in \omega_2} \mathcal{A}(\nu_1 \cap \nu_2) \right| = 0 \end{cases}$$

où $\mathcal{A}$ est l'aire d'une surface.

Notons que la liaison faible entre deux grilles s'effectue de manière continue au sens de cette énergie de régularisation binaire. En effet, pour des nœuds de deux grilles différentes, s'ils sont indépendants, c'est la surface d'intersection $\mathcal{A}(\nu_1 \cap \nu_2)$ qui est considérée, s'ils deviennent liés, c'est la surface complémentaire $\mathcal{A}(\nu_1 \Delta \nu_2)$. Il existe une transition entre ces deux situations où $\mathcal{A}(\nu_1 \cap \nu_2) = \mathcal{A}(\nu_1 \Delta \nu_2)$.

En revanche, dans le cas d'un recouvrement total, l'énergie globale E_R de la régularisation de l'agencement ne suit plus une évolution continue : elle *saut* vers une énergie de pénalisation.

Ce type de régularisation, dans la mesure où elle encourage des agencements de grilles de structures recouvrantes, pourrait davantage se prêter à des applications de type indexation.

5.1.3 Complexité de l'énergie d'un agencement

L'énergie d'un agencement de grilles est très rapide à calculer. Elle se compose, en effet, d'une énergie d'attache aux données qui n'accède pas directement à l'image d'entrée mais à un espace beaucoup plus restreint et d'une énergie de régularisation basée sur des calculs très simples d'aires.

Complexité de l'énergie E_G d'une grille $C(E_G(\omega)) = O(mn\nu_l\nu_h)$ L'énergie d'attache aux données E_G se compose de trois termes. La complexité des termes $E_{\text{uniformité}}$, d'estimation de l'uniformité du fond de mur (définition 4.7), et $E_{\text{contraste}}$, d'estimation du contraste des nœuds (définition 4.6), est négligeable devant celle de l'estimation $E_{\text{similarité}}$ de la similarité de nœuds (définition 4.5). En effet, les calculs d'uniformité et de contraste font intervenir des images intégrales (sous-section 4.2.3), de complexité linéaire avec le nombre moyen $\bar{m}$ de lignes et le nombre moyen $\bar{n}$ de colonnes (quatre accès mémoire par région rectangulaire de surface quelconque). L'estimation de la similarité, en revanche, implique le parcours exhaustif de tous les pixels de chaque nœud. Ainsi la complexité $C(E_I(W)) = C(E_{\text{similarité}}(W))$ est de l'ordre de la surface de tous les nœuds de l'agencement. Nous utilisons alors une image sous-échantillonnée pour le calcul des similarités de nœud. Cette simplification ne perturbe pas beaucoup la mesure mais accélère les calculs d'une manière conséquente.

Complexité de l'énergie E_R d'un agencement $C(E_R(W)) = O(k^2)C(\Theta)$ L'énergie de régularisation est basée exclusivement sur des calculs d'aires d'intersections de rectangles orientés sur les axes verticaux et horizontaux. Sa complexité est quadratique avec le nombre k de grilles de l'agencement considéré. La complexité de la régularisation binaire Θ_1 est constante : $C_{\Theta_1}(\omega_1, \omega_2) = O(1)$. Celle des régularisations Θ_2 et Θ_3 est quadratique avec le nombre moyen de nœuds : $C_{\Theta_2}(\omega_1, \omega_2) = C_{\Theta_3}(\omega_1, \omega_2) = O(m^2.n^2)$ pour deux grilles de m lignes et n colonnes de nœuds.

5.2 Optimisation de l'agencement de grilles

Pour construire le meilleur agencement de grilles W^* au sens de l'énergie E_A que nous venons de définir (équation 5.3), nous ne pouvons pas parcourir l'espace $\mathcal{W}$ des agencements (définition 5.2) de manière exhaustive en un temps raisonnable. Cet espace est en effet très complexe : il se compose d'une union infinie de sous-ensembles de dimensions variables. L'énergie que nous voulons minimiser, E_A , ne possède pas de propriétés intéressantes comme la convexité. Il existe différentes

solutions d'optimisation à ce type de problème. Les méthodes déterministes ne sont pas réalisables dans un temps raisonnable. Des heuristiques d'inspiration biologique comme des algorithmes évolutionnistes (*cf* Jong [2006]) permettraient certainement de converger vers une solution intéressante mais c'est l'utilisation d'un *processus ponctuel marqué* couplé à un *recuit simulé* que nous avons choisi pour son formalisme stochastique rigoureux. Il s'agit d'une méthode de détection d'objets modélisés à partir de paramètres géométriques dans des espaces de très grandes dimensions. Leur utilisation s'est très répandue dans la littérature depuis la dernière décennie [Alegre et Dellaert, 2004; Ortner et al., 2007; Lafarge et al., 2010; Tournaire et al., 2010].

Nous utilisons l'implémentation générique d'optimisation stochastique présentée dans Brédif et Tournaire [2012] utilisée alors pour l'extraction d'empreintes de bâtiments dans des modèles numériques d'élévation. Sa paramétrisation est simple et son temps d'exécution est raisonnable (dans notre contexte des agencements de grilles de structures répétitives, de l'ordre de la minute pour la recherche d'une ou deux dizaines d'objets dans une image de 2 méga-pixels sur un processeur de 2,83 GHz).

Cette section donne un rapide aperçu du couplage d'un échantillonnage par *chaînes markoviennes de Monte Carlo à sauts réversibles* [Hastings, 1970; Green, 1995] et d'un *recuit simulé* (Salamon et al. [2002]) pour l'optimisation des agencements de grilles de structures répétitives. Elle ne prétend pas exposer l'intégralité du formalisme mathématique. Pour une explication complète, se référer à Casella et Robert [2004] ou Descombes [2011].

Nous montrons d'abord comment un processus ponctuel marqué nous permet d'échantillonner des agencements à partir d'une distribution de probabilité relative à l'énergie E_A . L'approche de Hastings [1970] permet de parcourir l'espace $\mathcal{W}_k$ des agencements d'un nombre fixe de grilles (sous-section 5.2.1). Nous expliquons ensuite l'extension de cette méthode (*cf* Green [1995]) pour échantillonner des agencements d'un nombre variable de grilles (sous-section 5.2.2). Puis nous détaillons notre utilisation du *recuit simulé* pour faire tendre cette distribution vers un échantillonnage des minima de l'énergie E_A (sous-section 5.2.3). Nous donnons ensuite une rapide explication sur notre choix de la distribution de probabilité de référence (sous-section 5.2.4). Nous terminons notre propos par le détail des noyaux que nous utilisons pour perturber itérativement les agencements de grilles de structures répétitives afin de les amener statistiquement vers l'agencement optimal (sous-section 5.2.5).

5.2.1 Échantillonnage par une chaîne de Markov de Monte Carlo

Pour un nombre fixé k de grilles de structures répétitives, il est possible d'appliquer la méthode de Hastings [1970] pour parcourir les agencements dans l'espace $\mathcal{W}_k$ par petites modifications successives des grilles. Une *chaîne de Markov de Monte Carlo* (MCMC) va échantillonner, à son régime stationnaire, une succession d'agencements à partir d'une distribution de probabilité cible à définir. Nous utilisons, pour cette distribution cible, une distribution de *Gibbs* (définition 5.10). D'une manière générale, cette dernière permet de construire une distribution de probabilité, π , à partir d'une fonction quelconque d'énergie, U , du type de celle que nous voulons mi-

nimiser (E_A). Selon cette distribution, plus l'énergie $U(W)$ relative à l'agencement W , est faible, plus la probabilité $\pi(W)$ est élevée.

Définition 5.10 *Les agencements de grilles de structures répétitives sont échantillonnés selon la distribution de probabilité π définie par l'équation 5.1.*

$$\pi(W) = \frac{\pi_{\text{référence}}(W)e^{-U(W)}}{\int_{V \in \mathcal{W}} \pi_{\text{référence}}(V)e^{-U(V)}dV} \quad (5.1)$$

où $\pi_{\text{référence}}(W)$ est la distribution de probabilité de référence de l'agencement W .

La *distribution de probabilité de référence*, $\pi_{\text{référence}}$ est la probabilité que l'agencement se trouve dans l'image pour une énergie U constante (section 5.2.4). Nous parlons d'une énergie $U(W)$ à la place de l'énergie E_A pour rester dans un cadre suffisamment général. La sous-section 5.2.3, relative à l'optimisation du processus, précise de quelle manière ces deux énergies dépendent l'une de l'autre. Pour suivre notre explication sur l'échantillonnage stochastique, le lecteur peut considérer U proportionnelle à E_A .

Notons qu'il n'est pas possible d'échantillonner directement le système en fonction de cette distribution de probabilité car le calcul de l'intégrale au dénominateur, qui assure la normalisation sur $\mathcal{W}$, est un concept théorique difficilement réalisable analytiquement en un temps acceptable.

Chaque état de la chaîne de Markov contient les paramètres de toutes les grilles d'un agencement. Les transitions possibles entre les états sont appelés *noyaux de proposition* du processus. Nous avons défini plusieurs noyaux relatifs à la modification des grilles (localisation, nombre, taille des nœuds et espacement entre eux). En partant d'un agencement initial, le processus choisit itérativement un noyau de proposition pour construire un nouvel agencement. Chaque noyau de proposition que nous avons implémenté est détaillé dans la suite de la section (sous-section 5.2.5).

À chaque itération, la probabilité du processus d'accepter ou de refuser le nouvel agencement W' est fixée en fonction du ratio des probabilités $\frac{\pi(W')}{\pi(W)}$ entre l'ancien agencement W et ce nouvel agencement (équation 5.2).

$$\frac{\pi(W')}{\pi(W)} = \frac{\pi_{\text{référence}}(W')}{\pi_{\text{référence}}(W)} e^{U(W)-U(W')} \quad (5.2)$$

Ainsi, plus la valeur de $\pi(W')$ est inférieure à celle de $\pi(W)$, moins le processus a de chance d'accepter le nouvel agencement. Notons que, cette fois-ci, le calcul, relatif aux deux distributions π , est réalisable sans connaître l'intégrale du dénominateur.

Chaque état de la chaîne ne dépend que de son prédécesseur. Au bout d'un certain nombre d'itérations, l'échantillonnage atteint un régime stationnaire et suit la distribution désirée (définition 5.10). Cette convergence est garantie par l'ergodicité de la chaîne (tout état est atteignable depuis tout autre état).

5.2.2 Processus ponctuel marqué

L'algorithme de Hastings [1970] ne permet de manipuler que des agencements d'un nombre fixe de grilles. Une extension proposée par Green [1995] permet d'échantillonner des agencements d'un nombre variable de grilles par l'utilisation de noyaux de type *sauts réversibles* (reversible jump MCMC).

Il s'agit d'un *processus ponctuel marqué* : un échantillonnage par chaîne de Markov sur des ensembles non ordonnés d'objets. Nous générons successivement des *points* sur toute l'image, que nous supprimons ou déplaçons par la suite. Chacun de ces *points* est le décalage d'une grille de structures répétitives. Nous associons à chacune d'elle, des *marques* : les six autres paramètres, à savoir la dimension des nœuds (ν_l, ν_h) , celle de l'espace entre les nœuds, (e_l, e_h) , et le nombre de lignes m et de colonnes n de ces derniers.

Un noyau de proposition K est réversible s'il définit, pour tout état $X = (k, W)$, relatif à l'agencement W de k grilles de structures répétitives, une distribution de probabilité $K(X, \cdot)$ sur $\mathcal{W}$ qui vérifie la condition 5.3.

$$\int_A \int_{A'} K(X, dX') \pi(X) dX = \int_{A'} \int_A K(X', dX) \pi(X') dX' \quad (5.3)$$

pour tout $A, A' \subset \mathcal{W}$ et pour une distribution de probabilité invariante π (Casella et Robert [2004]).

Pour réaliser cette condition, la définition 5.11 formalise la structure d'un noyau réversible selon la terminologie de Casella et Robert [2004].

Définition 5.11 *Un noyau réversible K se compose de la structure suivante :*

Une probabilité $\pi_{kk'}$ *pour le choix du noyau depuis l'état (k, W) vers l'état (k', W') .*

Soit $\pi_{k'k}$ la probabilité réciproque.

Une suite de variables $u_{kk'}$ *Si l'espace de l'agencement de départ et celui de l'agencement d'arrivée n'ont pas la même dimension, le principe de Green [1995] est d'ajouter des variables aléatoires complémentaires dans les deux espaces. Nous définissons ainsi une suite $u_{kk'}$ de telles variables qui s'ajoutent aux paramètres des k grilles de l'agencement W . Nous définissons, de même, une suite $v_{kk'}$ de variables aléatoires complémentaires des n grilles de l'agencement W' . Nous définissons d'autre part les densités de probabilité $\varphi_{kk'}(u)$ et $\varphi_{k'k}(v)$ respectivement associées à ces suites.*

Une bijection $T_{kk'}$ *entre l'état courant et l'état candidat telle que $T_{kk'}(W, u_{kk'}) = (W', v_{k'k})$. Nous définissons la bijection inverse $T_{k'k} = T_{kk'}^{-1}$.*

Le noyau K vérifie la condition de réversibilité 5.3 si l'agencement de grilles W' depuis l'agencement W est choisi en fonction du taux d'acceptation $\Gamma(W, W')$ donné par l'équation 5.4.

$$\Gamma(W, W') = \min \left(1, \frac{\pi(k', W')}{\pi(k, W)} \frac{\pi_{k'k} \varphi_{k'k}(v_{k'k})}{\pi_{kk'} \varphi_{kk'}(u_{kk'})} \left| \frac{\delta T_{kk'}(W, u_{kk'})}{\delta(W, u_{kk'})} \right| \right) \quad (5.4)$$

L'équation 5.5 en donne un développement par l'utilisation de l'équation 5.2.

$$\Gamma(W, W') = \min \left(1, \frac{\pi_{\text{référence}}(W')}{\pi_{\text{référence}}(W)} \frac{\pi_{k'k} \varphi_{k'k}(v_{k'k})}{\pi_{kk'} \varphi_{kk'}(u_{kk'})} \left| \frac{\delta T_{kk'}(W, u_{kk'})}{\delta(W, u_{kk'})} \right| e^{U(W) - U(W')} \right) \quad (5.5)$$

L'algorithme de Green [1995] est résumé par l'algorithme 1.

Nous expliquons, dans la sous-section 5.2.3, l'utilisation du *recuit simulé* qui permet de faire converger les échantillons de la distribution π vers les minima de l'énergie E_A .

Algorithme 1 : Pseudo-code de l'échantillonnage de Green [1995]

Initialiser l'agencement W à l'agencement vide ;

répéter

Choisir un noyau de proposition K_i suivant la probabilité $\pi_{kk'}$;

Générer la suite $u_{kk'}$ suivant la probabilité $\varphi_{kk'}(u)$.

Construire le nouvel agencement $(W', v_{k'k}) = T_{kk'}(W, u_{kk'})$;

Accepter W' suivant la probabilité $\Gamma(W, W')$;

jusqu'à atteindre la distribution stationnaire

5.2.3 Convergence vers l'agencement optimal

L'agencement optimal W^* est celui dont l'énergie $E_A(W^*)$ est globalement minimale sur l'espace $\mathcal{W}$. Nous utilisons le procédé de *recuit simulé* (cf Salamon et al. [2002]) pour minimiser l'énergie E_A . Le principe est la pénalisation progressive du taux d'acceptation Γ à un état d'énergie supérieure. Ce procédé permet d'explorer la plus large partie possible de l'espace de recherche pour détecter les minima globaux. Initialement, de grands écarts d'énergie entre deux états sont acceptés et le processus peut échantillonner des états d'énergies localement minimales sans y rester bloqué. Au fur et à mesure, il est de plus en plus contraint jusqu'à suivre une distribution de *Dirac* centrée sur le minimum global de l'énergie U (ou une mixture de *distributions de Dirac* si l'énergie U possède plusieurs minima).

Nous posons $U = \frac{E_A}{T}$ pour diviser l'énergie E_A par un paramètre T , dit *de température*, que nous initialisons à une très grande valeur et que nous faisons tendre vers zéro au fur et à mesure des itérations. La distribution suivant laquelle nous cherchons à échantillonner les agencements de grilles de structures répétitives est alors donnée par la définition 5.12.

Définition 5.12 À la température T , les agencements de grilles de structures répétitives sont échantillonnés selon la distribution de probabilité π_T définie par l'équation 5.6.

$$\pi_T(W) = \frac{\pi_{\text{référence}}(W)e^{-\frac{E_A(W)}{T}}}{\int_{V \in \mathcal{W}} \pi_{\text{référence}}(V)e^{-\frac{E_A(V)}{T}} dV} \quad (5.6)$$

À l'état initial du système, T est infinie. La distribution de l'échantillonnage est alors complètement indépendante de E_A et suit la distribution de référence $\pi_{\text{référence}}$. À la fin du processus, T est quasiment nulle et la distribution tend vers une *distribution de Diracs* centrée sur les minima globaux de l'énergie du système.

La figure 5.5 schématise l'évolution du régime de cet échantillonnage. L'énergie E_A à minimiser est symbolisée en rouge accompagnée des distributions de probabilités stationnaires guidant l'échantillonnage à quatre températures différentes : température infinie, température intermédiaire T_i , température intermédiaire T_j ($< T_i$) et température quasiment nulle. À température infinie, la probabilité de sélection des agencements de grilles $\pi_\infty(W)$ ne dépend pas de leur énergie mais suit la distribution de référence $\pi_{\text{référence}}$ (ici, uniforme). À températures intermédiaires, T_i et T_j , la distribution tend à privilégier progressivement les agencements d'énergie faible. À température très basse, la distribution $\pi_\epsilon(W)$ est focalisée sur l'optimum.

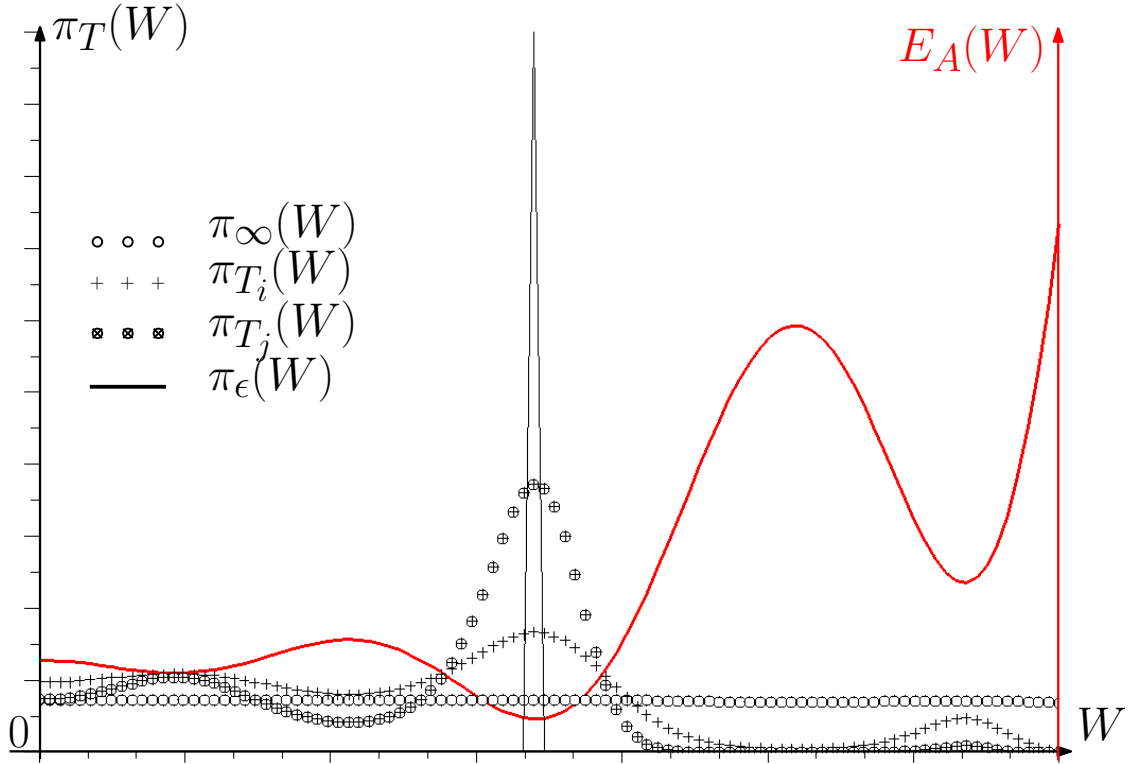


FIGURE 5.5 – Régimes stationnaires (définition 5.12) à différentes températures.

Le point critique du *recuit simulé* consiste en la décroissance de la température qui doit être suffisamment lente pour permettre la convergence de l'échantillonnage vers la distribution stationnaire mais qui doit s'annuler à un moment donné pour que le régime stationnaire échantillonne l'agencement optimal. Salamon et al. [2002] précisent que la convergence théorique du processus vers un optimum global est garantie par une décroissance logarithmique de la température ainsi que par une température initiale, T_0 , suffisamment élevée (pour laisser le processus explorer tout l'espace et ne pas rester coincé, dès le départ, dans un minimum local).

En pratique nous utilisons une décroissance géométrique de température : à chaque itération i , $T_i = \theta^i T_0$. Il s'agit d'un bon compromis entre la théorie et la pratique.

5.2.4 Distribution de référence

La distribution de probabilité π_T (définition 5.12), selon laquelle nous voulons échantillonner le processus, dépend d'une *distribution de référence*, $\pi_{\text{référence}}$. La seule contrainte pour le choix de cette dernière est qu'elle ne doit pas s'annuler au niveau des minima de l'énergie E_A dans l'espace $\mathcal{W}$. En effet, en fin de processus, à très basse température ($T = \epsilon \ll 1$), π_ϵ suit une *distribution de Diracs* dont toutes les valeurs sont nulles en dehors des minima globaux de l'énergie E_A .

Par définition, l'espace $\mathcal{W}$ est la triple union $\mathcal{W} = \bigcup_{k=0}^{\infty} \left(\bigcup_{m=m^{\min}}^{m^{\max}} \bigcup_{n=n^{\min}}^{n^{\max}} \Omega_{(m,n)} \right)^k$. Dans notre application, nous définissons une distribution $\rho(k)$ pour choisir le nombre k de grilles que doit contenir l'agencement W et une autre, $P_{(m,n)}(W[i])$, pour

construire, indépendamment entre elles, chacune de ses grilles $W[i]$. Les nombres de lignes et de colonnes m et n , entiers naturels non nuls, sont choisis selon une distribution de probabilité uniforme.

Échantillonnage du nombre de grilles Nous fixons le nombre k des grilles de l'agencement suivant une distribution de probabilité uniforme (équation 5.7).

$$\rho(k) = \begin{cases} \frac{1}{\lambda} & k \in [0, \lambda] \\ 0 & k \in [\lambda, +\infty] \end{cases} \quad (5.7)$$

Le paramètre λ est le nombre maximum de grilles que peut contenir l'agencement.

Échantillonnage des grilles Nous échantillonnons uniformément et indépendamment les deux triplets de paramètres (x, ν_l, e_l) et (y, ν_h, e_h) sur l'espace $\Omega_{m,n}$. Ce paragraphe détaille le calcul du premier cas, les paramètres verticaux se déduisant par une démarche analogue.

Les contraintes horizontales sont listées par l'équation 5.8.

$$\begin{cases} x^{min} < x < x^{max} \\ \nu_l^{min} < \nu_l < \nu_l^{max} \\ e_l^{min} < e_l < e_l^{max} \\ x^{min} < x + n\nu_l + (n-1)e_l < x^{max} \end{cases} \quad (5.8)$$

où x^{min} , x^{max} , ν_l^{min} , ν_l^{max} , e_l^{min} et e_l^{max} sont les bornes des paramètres.

La probabilité $P_n(x, \nu_l, e_l)$ de tirer le triplet de paramètres horizontaux (équation 5.9) est alors l'inverse du volume $V_n(x, \nu_l, e_l)$ représenté par ces contraintes (figure 5.6).

$$P_n(x, \nu_l, e_l) = \frac{1}{V_n(x, \nu_l, e_l)} \quad (5.9)$$

L'équation 5.10 donne le calcul du volume $V(x, \nu_l, e_l)$.

$$V_n(x, \nu_l, e_l) = \int_{\nu_l^{min}}^{\nu_l^{max}} \int_{e_l^{min}}^{e_l^{max}} x_n(\nu, e) de d\nu - \int_{\nu_l^{min}}^{\nu_l^{max}} \int_{e_l^{min}}^{e_l^{max}} x^{min} de d\nu \quad (5.10)$$

où $x_n(\nu, e) = x^{max} - n\nu - (n-1)e$.

$$\begin{aligned} V_n(x, \nu_l, e_l) &= (x^{max} - x^{min}) [\nu]_{\nu_l^{min}}^{\nu_l^{max}} [e]_{e_l^{min}}^{e_l^{max}} - \frac{n}{2} [\nu^2]_{\nu_l^{min}}^{\nu_l^{max}} [e]_{e_l^{min}}^{e_l^{max}} - \frac{n-1}{2} [\nu]_{\nu_l^{min}}^{\nu_l^{max}} [e^2]_{e_l^{min}}^{e_l^{max}} \\ &= (\nu_l^{max} - \nu_l^{min})(e_l^{max} - e_l^{min}) \left[x^{max} - x^{min} - n \frac{\nu_l^{max} + \nu_l^{min}}{2} - (n-1) \frac{e_l^{max} + e_l^{min}}{2} \right] \end{aligned}$$

La probabilité $P_{(m,n)}(W[i])$ d'une grille $W[i]$ de structures répétitives est le produit des probabilités des paramètres horizontaux et verticaux (équation 5.11).

$$P_{(m,n)}(W[i]) = \frac{1}{m^{max} - m^{min}} \frac{1}{n^{max} - n^{min}} P_n(x, \nu_l, e_l) P_m(y, \nu_h, e_h) \quad (5.11)$$

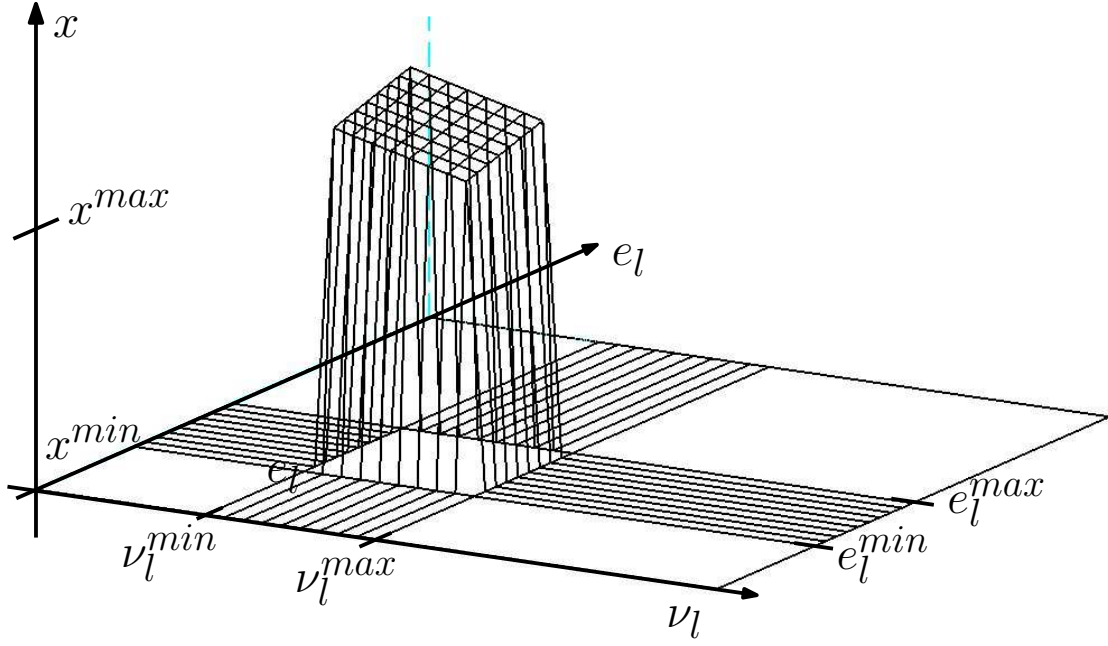


FIGURE 5.6 – Espace d'échantillonnage de volume $V_n(x, \nu_l, e_l)$ des paramètres horizontaux d'une grille de structures répétitives (suivant les contraintes 5.8).

Échantillonnage des agencements L'équation 5.12 donne alors l'expression de la distribution de référence $\pi_{\text{référence}}$ d'un agencement W de k grilles de structures répétitives.

$$\pi_{\text{référence}}(W) = \rho(k) \prod_{i=1}^k P_{(m,n)}(W[i]) \quad (5.12)$$

5.2.5 Implémentation des sauts réversibles

La dimension de l'espace à échantillonner est infinie car le nombre de grilles à détecter n'est pas connu d'avance. Des noyaux de transition, de type *naissance* et *mort*, permettent de faire évoluer sa dimension.

D'autres noyaux, que nous appelons *noyaux de perturbation*, permettent de modifier les paramètres des grilles sans modifier leur nombre afin de les approcher de l'optimum local. Théoriquement, le processus atteint le régime stationnaire sans leur utilisation avec uniquement les noyaux de *naissance* et *mort*. Mais le temps de convergence se réduit considérablement grâce aux noyaux de perturbation qui permettent des variations locales de grilles en une seule itération dans l'algorithme 1 (au lieu d'une mort suivie d'une naissance).

5.2.5.1 Noyaux de naissance et de mort

Ces noyaux permettent de modifier la dimension de l'agencement de grilles. Un noyau de type *naissance* est l'ajout d'une grille et un noyau de type *mort* en est la suppression.

Naissance Nous utilisons une série de noyaux de type *naissance*, chacun associé à un nombre m de lignes et un nombre n de colonnes donnés, pour ajouter une grille $\omega_{(m,n)}$ à l'agencement W comprenant k grilles. Le processus doit générer aléatoirement la suite transitoire de paramètres complémentaires $u_{kk'}$ de cette nouvelle grille suivant l'équation 5.13.

$$W[k+1] = \left(u_{kk'} = \begin{pmatrix} u_{kk'}[1] \in [x^{\min}, x^{\max}] \\ u_{kk'}[2] \in [y^{\min}, y^{\max}] \\ u_{kk'}[3] \in [\nu_l^{\min}, \nu_l^{\max}] \\ u_{kk'}[4] \in [\nu_h^{\min}, \nu_h^{\max}] \\ u_{kk'}[5] \in [e_l^{\min}, e_l^{\max}] \\ u_{kk'}[6] \in [e_h^{\min}, e_h^{\max}] \end{pmatrix} \right) \quad (5.13)$$

Un nouvel agencement W' de $k' = k + 1$ grilles est alors proposé suivant l'équation 5.14.

$$W'[k+1] = \left(\omega'_{(m,n)} = \begin{pmatrix} x = u_{kk'}[0] \\ y = u_{kk'}[1] \\ \nu_l = u_{kk'}[2] \\ \nu_h = u_{kk'}[3] \\ e_l = u_{kk'}[4] \\ e_h = u_{kk'}[5] \end{pmatrix} \right) \quad (5.14)$$

L'équation 5.15 donne le calcul des ratios des distributions de référence.

$$\frac{\pi_{\text{référence}}(W')}{\pi_{\text{référence}}(W)} = \frac{\frac{\lambda^{k+1}e^{-\lambda}}{(k+1)!} \prod_{i=1}^{k+1} P(W'[i])}{\frac{\lambda^k e^{-\lambda}}{k!} \prod_{i=1}^k P(W[i])} = \frac{\lambda}{k+1} P(W'[k+1]) \quad (5.15)$$

où le calcul de la probabilité $P(W'[k+1])$ est donné en équation 5.11.

La probabilité $\varphi_{kk'}(u_{kk'})$ correspond à la probabilité de la création d'une nouvelle grille, c'est à dire $P(W'[k+1])$. La probabilité réciproque, $\varphi_{k'k}(v_{k'k})$, correspond à la suppression de la grille ajoutée (le noyau de *mort* expliqué dans le paragraphe suivant). Elle vaut 1 car il s'agit d'une opération déterministe.

Le jacobien $\frac{\delta T_{kk'}(W, u_{kk'})}{\delta(W, u_{kk'})}$ est obtenu à partir des relations 5.13 et 5.14. Étant donné qu'il s'agit d'une association bijective entre les termes de la suite $u_{kk'}$ et les paramètres de la grille $\omega_{(m,n)}$, il vaut l'identité.

L'équation 5.16 donne alors l'expression du taux d'acceptation d'un agencement proposé par un noyau de type naissance.

$$\Gamma(W, W') = \min \left(1, \frac{\lambda}{(k+1)} P(W'[k+1]) \frac{\pi_{k'k}}{\pi_{kk'}} \frac{1}{P(W'[k+1])} (1) e^{\frac{E_A(W) - E_A(W')}{T}} \right)$$

$$\Gamma(W, W') = \min \left(1, \frac{\lambda}{k+1} \frac{\pi_{k'k}}{\pi_{kk'}} e^{\frac{E_A(W) - E_A(W')}{T}} \right) \quad (5.16)$$

Mort De même, nous utilisons une série de noyaux de type *mort* permettant de supprimer une certaine grille de l'agencement W de k grilles. Chacun est associé à un indice k de la-dite grille $W[k]$ à supprimer (équation 5.17).

$$W[k] = \left(\omega_{(m,n)} = \begin{pmatrix} x \\ y \\ \nu_l \\ \nu_h \\ e_l \\ e_h \end{pmatrix} \right) \quad (5.17)$$

Le nouvel agencement W' est alors généré avec une grille en moins. Le processus constitue une suite transitoire de paramètres complémentaires $v_{k'k}$ de la grille supprimée selon l'équation 5.18.

$$W'[k] = \left(v_{k'k} = \begin{pmatrix} v_{k'k}[0] = x \\ v_{k'k}[1] = y \\ v_{k'k}[2] = \nu_l \\ v_{k'k}[3] = \nu_h \\ v_{k'k}[4] = e_l \\ v_{k'k}[5] = e_h \end{pmatrix} \right) \quad (5.18)$$

Les calculs du taux d'acceptation $\Gamma(W, W')$ pour les noyaux de type *mort* sont analogues aux noyaux de type *naissance*. L'équation 5.19 fournit son résultat.

$$\Gamma(W, W') = \min \left(1, \frac{k+1}{\lambda} \frac{\pi_{k'k}}{\pi_{kk'}} e^{\frac{E_A(W) - E_A(W')}{T}} \right) \quad (5.19)$$

5.2.5.2 Noyaux de perturbation

Nous utilisons plusieurs séries de noyaux de perturbation de l'une des grilles de l'agencement W . De même que pour le noyau de mort, chacun de ces noyaux est associé à un indice k d'une grille $W[k] = \omega_{(m,n)}$ à modifier. Certains modifient le nombre de nœuds dans la grille. D'autres en font varier la taille et l'espacement.

L'équation 5.20 donne l'expression du taux d'acceptation pour tous ces types de noyaux.

$$\Gamma(W, W') = \min \left(1, e^{\frac{E_A(W) - E_A(W')}{T}} \right) \quad (5.20)$$

Quelques éléments de démonstration du résultat 5.20 sont proposés dans ce qui suit. Notons tout d'abord à ce sujet que, sous l'hypothèse de l'équiprobabilité des grilles, le nombre de ces dernières n'étant pas modifié, tous les rapports de distributions de référence valent 1.

Par soucis de synthèse, nous n'illustrons que le cas des opérations horizontales. Le résultat est différent selon qu'il s'effectue sur le bord droit ou le gauche. Les opérations verticales suivent les mêmes lois.

Variation du nombre de nœuds Pour modifier horizontalement le nombre de nœuds, deux types de noyaux de perturbation permettent l'ajout ou la suppression d'une séquence sur un bord de la grille (bord gauche ou bord droit). La figure 5.7 illustre ces variations.

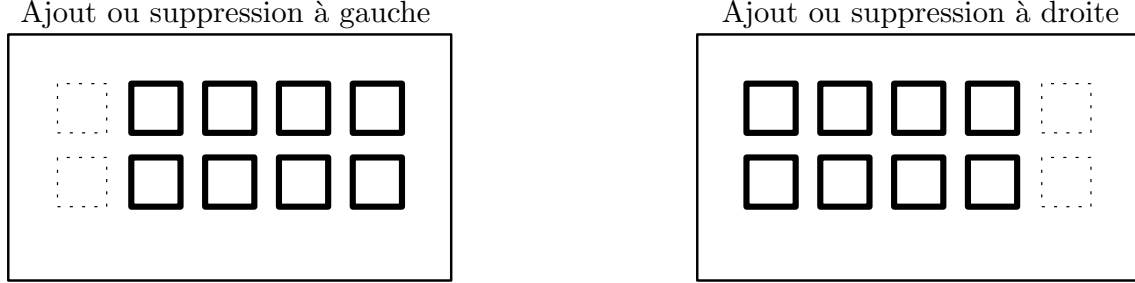


FIGURE 5.7 – Ajout ou suppression d'une colonne sur le bord gauche ou droit d'une grille.

Les noyaux étant déterministes, ils ne nécessitent pas de variable aléatoire supplémentaire : les suites $u_{kk'}$ et $v_{k'k}$ sont vides. Alors les probabilités $\varphi_{k'k}(v_{k'k})$ et $\varphi_{kk'}(u_{kk'})$ valent toutes les deux 1.

D'autre part, ajouter ou supprimer une séquence de nœuds sur le bord droit ne modifie pas les valeurs de la nouvelle grille $W'[k] = \omega_{(m,(n\pm 1))}$. En revanche, un ajout ou une suppression au bord gauche implique une variation de l'offset d'une valeur égale à $(\nu_l + e_l)$. Cette différence se manifeste dans le calcul du jacobien $\frac{\delta T_{kk'}(W, u_{kk'})}{\delta(W, u_{kk'})}$ dont le déterminant vaut 1 dans les deux cas. Dans le premier cas, le jacobien vaut l'identité. L'équation 5.21 donne son expression dans l'autre cas.

$$\frac{\delta T_{kk'}(W, u_{kk'})}{\delta(W, u_{kk'})} = \begin{pmatrix} 1 & 0 & \pm 1 & 0 & \pm 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (5.21)$$

Translation d'un côté des nœuds Deux premières séries de noyaux de perturbation horizontales de la taille des nœuds consistent en la translation d'un seul côté des nœuds (côtés gauches et côtés droits). La figure 5.8 illustre ces translations.

Ce type de noyau nécessite l'utilisation d'une variable aléatoire complémentaire $u_{kk'} \in [0, \nu_l + e_l]$ pour sélectionner la nouvelle taille de nœuds (équation 5.22).

$$W[k] = \begin{pmatrix} \omega_{(m,n)} = \begin{pmatrix} x \\ y \\ \nu_l \\ \nu_h \\ e_l \\ e_h \end{pmatrix} \\ u_{kk'} \in [0, \nu_l + e_l] \end{pmatrix} \quad (5.22)$$

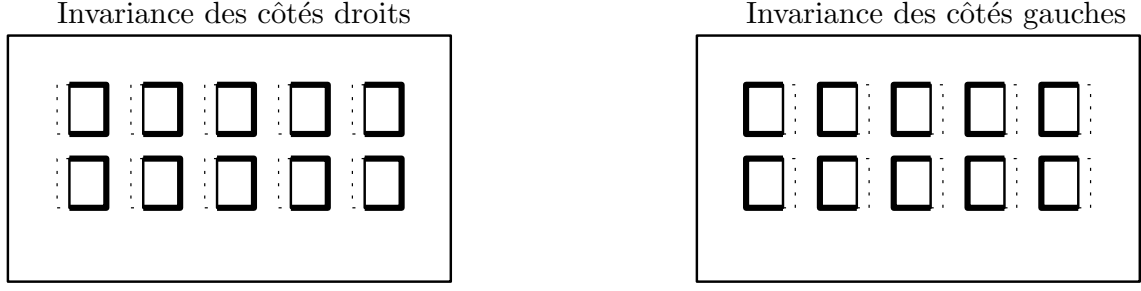


FIGURE 5.8 – Translation horizontale d'un côté des nœuds de la grille : configuration initiale en traits pointillés, variation en traits continus, segments invariants en gras.

L'équation 5.23 donne l'expression de l'état de la grille $W'[k]$ après une translation des côtés droits des nœuds ainsi que celle du jacobien $\frac{\delta T_{kk'}(W, u_{kk'})}{\delta(W, u_{kk'})}$.

$$W'[k] = \left(\omega'_{(m,n)} = \begin{pmatrix} x \\ y \\ u_{kk'} \\ \nu_h \\ e_l + \nu_l - u_{kk'} \\ e_h \\ v_{k'k} = \nu_l \end{pmatrix} \right) \quad \frac{\delta T_{kk'}(W, u_{kk'})}{\delta(W, u_{kk'})} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (5.23)$$

La translation des côtés gauches implique une variation de l'offset de la grille $W'[k]$. L'équation 5.24 en donne l'expression ainsi également que celle du jacobien $\frac{\delta T_{kk'}(W, u_{kk'})}{\delta(W, u_{kk'})}$.

$$W'[k] = \left(\omega'_{(m,n)} = \begin{pmatrix} x + \nu_l - u_{kk'} \\ y \\ u_{kk'} \\ \nu_h \\ e_l + \nu_l - u_{kk'} \\ e_h \\ v_{k'k} = \nu_l \end{pmatrix} \right) \quad \frac{\delta T_{kk'}(W, u_{kk'})}{\delta(W, u_{kk'})} = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (5.24)$$

D'autre part, pour les deux types de translation, la variable complémentaire $u_{kk'}$ étant choisie uniformément, nous avons $\varphi_{kk'}(u_{kk'}) = \varphi_{k'k}(v_{k'k}) = \frac{1}{e_l + \nu_l}$.

Compression et dilatation des grilles Une autre manière de modifier la taille des nœuds horizontalement consiste à tradater plusieurs côtés parallèles de nœuds en même temps. Nous utilisons deux séries de noyaux relatifs à des paires de segments invariants de types *gauche-droit* et *droit-gauche*. Les noyaux de type *gauche-droit* font varier la taille des nœuds en laissant invariants le côté gauche d'un premier nœud et le côté droit d'un deuxième à droite du premier (figure 5.9). Notons, pour cette série de noyaux, qu'une réduction de la taille des nœuds augmente leur espacement et entraîne donc une dilatation de la grille.

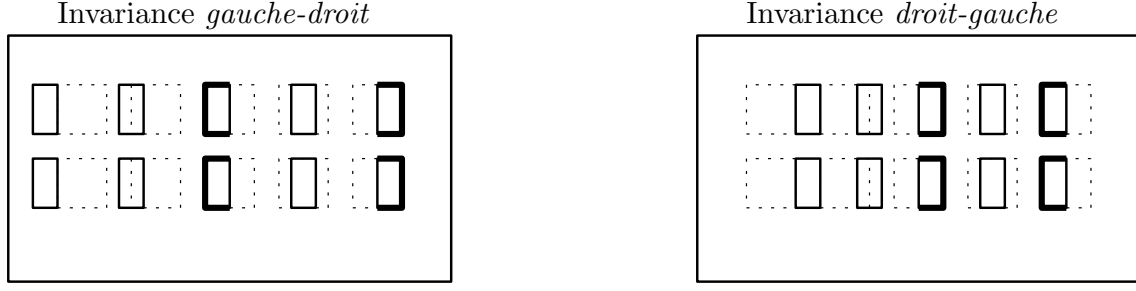


FIGURE 5.9 – Variations horizontales de la taille des nœuds relativement à une paire de segments invariants j_1 et j_2 de types *gauche-droit* et *droit-gauche* : configuration initiale en traits pointillés, variation en traits continus, segments invariants en gras.

De la même manière, les noyaux de type *droit-gauche* font varier la taille des nœuds en laissant invariants le côté droit d'un premier nœud et le côté gauche d'un deuxième à droite du premier (figure 5.9). Pour cette autre série de noyaux, à l'inverse de la précédente, une réduction de la taille des nœuds entraîne une contraction de la grille.

Il s'agit d'une variante des translations de côtés de nœuds, présentées précédemment, permettant de faire varier la grille de manière plus localisée sur une certaine partie. Il faut également utiliser une variable aléatoire complémentaire $u_{kk'} \in [0, \nu_l + \frac{e_l}{2}]$ pour sélectionner la nouvelle taille de nœuds (équation 5.25).

$$W[k] = \left(\begin{array}{c} \omega_{(m,n)} = \begin{pmatrix} x \\ y \\ \nu_l \\ \nu_h \\ e_l \\ e_h \end{pmatrix} \\ u_{kk'} \end{array} \right) \quad (5.25)$$

L'équation 5.26 donne l'expression de la grille $W'[k]$ dans le cas de l'invariance d'une paire de segments de type *gauche-droit* :

$$W'[k] = \left(\begin{array}{c} \omega'_{(m,n)} = \begin{pmatrix} x - \frac{j_1-1}{j_2-j_1}(\nu_l - u_{kk'}) \\ y \\ u_{kk'} \\ \nu_h \\ e_l + \frac{1+j_2-j_1}{j_2-j_1}(\nu_l - u_{kk'}) \\ e_h \\ \nu_l \end{pmatrix} \\ v_{k'k} = \nu_l \end{array} \right) \quad (5.26)$$

L'équation 5.27 donne l'expression du jacobien $\frac{\delta T_{kk'}(W, u_{kk'})}{\delta(W, u_{kk'})}$ dans le cas de l'inva-

riance d'une paire de segments de type *gauche-droit*. Son déterminant vaut 1.

$$\frac{\delta T_{kk'}(W, u_{kk'})}{\delta(W, u_{kk'})} = \begin{pmatrix} 1 & 0 & \frac{1-j_1}{j_2-j_1} & 0 & 0 & 0 & \frac{j_1-1}{j_2-j_1} \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{1+j_2-j_1}{j_2-j_1} & 0 & 1 & 0 & \frac{j_1-j_2-1}{j_2-j_1} \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (5.27)$$

L'équation 5.28 donne l'expression de la grille $W'[k]$ dans le cas de l'invariance d'une paire de segments de type *droit-gauche* :

$$W'[k] = \left(\omega'_{(m,n)} = \begin{pmatrix} x - \frac{j_2-1}{j_2-j_1}(\nu_l - u_{kk'}) \\ y \\ u_{kk'} \\ \nu_h \\ e_l + \frac{j_2-j_1-1}{j_2-j_1}(\nu_l - u_{kk'}) \\ e_h \\ \nu_l \\ v_{k'k} = \nu_l \end{pmatrix} \right) \quad (5.28)$$

L'équation 5.29 donne l'expression du jacobien $\frac{\delta T_{kk'}(W, u_{kk'})}{\delta(W, u_{kk'})}$ dans le cas de l'invariance d'une paire de segments de type *droit-gauche*. Son déterminant vaut également 1.

$$\frac{\delta T_{kk'}(W, u_{kk'})}{\delta(W, u_{kk'})} = \begin{pmatrix} 1 & 0 & \frac{1-j_2}{j_2-j_1} & 0 & 0 & 0 & \frac{j_2-1}{j_2-j_1} \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{j_2-j_1-1}{j_2-j_1} & 0 & 1 & 0 & \frac{1+j_1-j_2}{j_2-j_1} \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (5.29)$$

Pour l'invariance de type *gauche-droit*, si la variable complémentaire $u_{kk'}$ est choisie uniformément dans l'intervalle $[0, \nu_l + \frac{2}{3}e_l]$, la probabilité $\varphi_{kk'}(u_{kk'}) = \varphi_{k'k}(v_{k'k}) = \frac{1}{\nu_l + \frac{2}{3}e_l}$.

De la même manière, pour l'invariance de type *droit-droit*, si la variable complémentaire $u_{kk'}$ est choisie uniformément dans l'intervalle $[0, \nu_l + e_l]$, la probabilité $\varphi_{kk'}(u_{kk'}) = \varphi_{k'k}(v_{k'k}) = \frac{1}{\nu_l + e_l}$.

5.3 Résultats

De la même manière que pour l'extraction de grilles isolées de structures répétitives, nous maximisons les trois critères indépendamment dans l'espace $\hat{\Omega}_{\text{Omega}}$ (définition 4.10) pour les normaliser dans la combinaison linéaire de l'énergie $\hat{E}_G$ (définition 4.11). Nous optimisons un agencement de grilles de l'espace $\Omega_{(2,2)}$ et

nous utilisons l'énergie de régularisation binaire $\Theta^{(3)}$ autorisant les chevauchements de nœuds.

Les figures 5.10 et 5.11 présentent des agencements optimaux de grilles de structures répétitives. Comme pour l'extraction des grilles isolées, le problème de la simplicité de la formalisation des critères utilisé se retrouve : certaines grilles ont des nœuds disposés sur des régions de radiométrie uniforme encadré par un fort contraste sur un ou deux côtés.

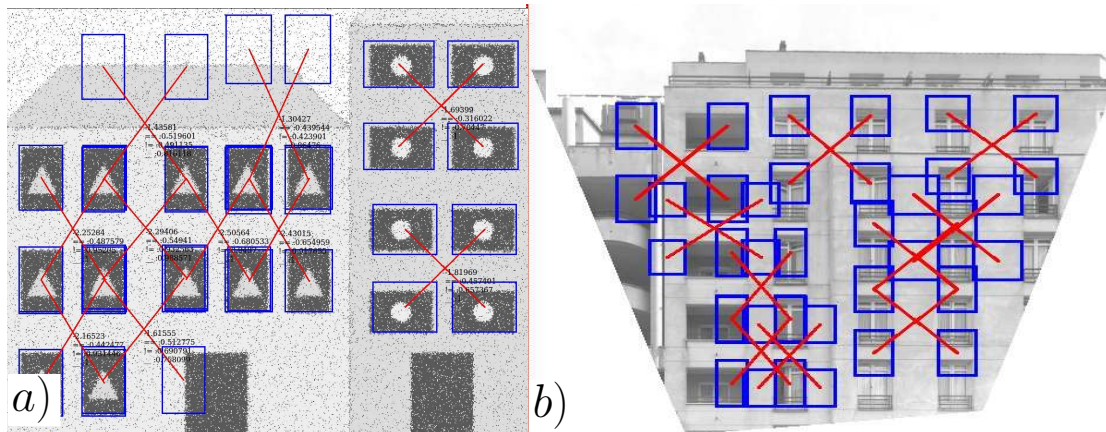


FIGURE 5.10 – Agencement de grilles de structures répétitives optimisé par processus ponctuel marqué. a) Sur une image de synthèse ; b) Sur une image réelle.

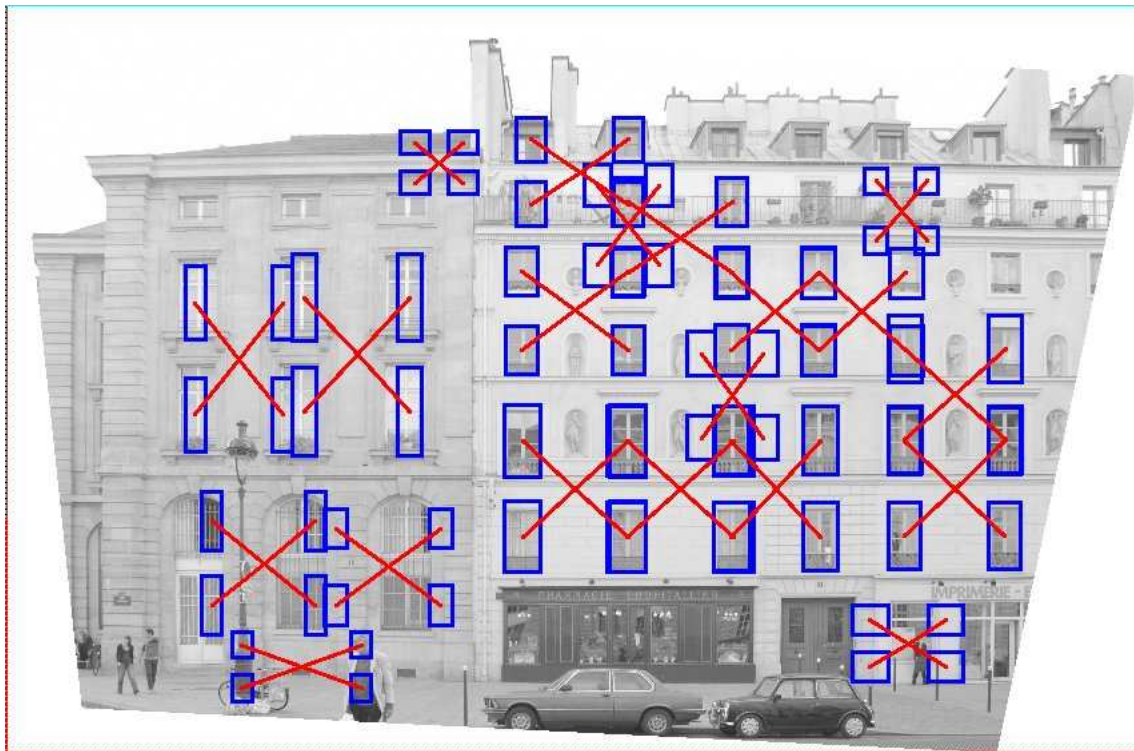


FIGURE 5.11 – Agencement de structures répétitives optimisé par processus ponctuel marqué.

La figure 5.12 présente un résultat représentatif de l'état actuel de notre mise en

œuvre du processus. La courbe en bleu représente l'énergie E_A que nous minimisons en fonction du nombre d'itération. Elle se compose des valeurs maximales et minimales enregistrées entre deux affichages (un nombre fixe d'itérations). La courbe rouge représente la décroissance de la température. Initialement, cette dernière est très élevée pour laisser le processus échantillonner les grilles de structures répétitives indépendamment de l'énergie E_A . À la fin, elle est très basse pour contraindre l'échantillonnage sur une distribution autour des minima de l'énergie. Nous constatons que beaucoup de fenêtres identiques sont localisées. En revanche, la formalisation actuelle du contraste laisse certains nœuds de grilles situés sur le fond de mur entre deux fenêtres (au lieu de justement se positionner sur ces mêmes fenêtres). Notre implémentation souffre encore d'un petit problème technique relatif aux bords de l'image (où le contraste est maximal) : les grilles qui sortent du masque de redressement de l'image ne devraient pas être échantillonnées.

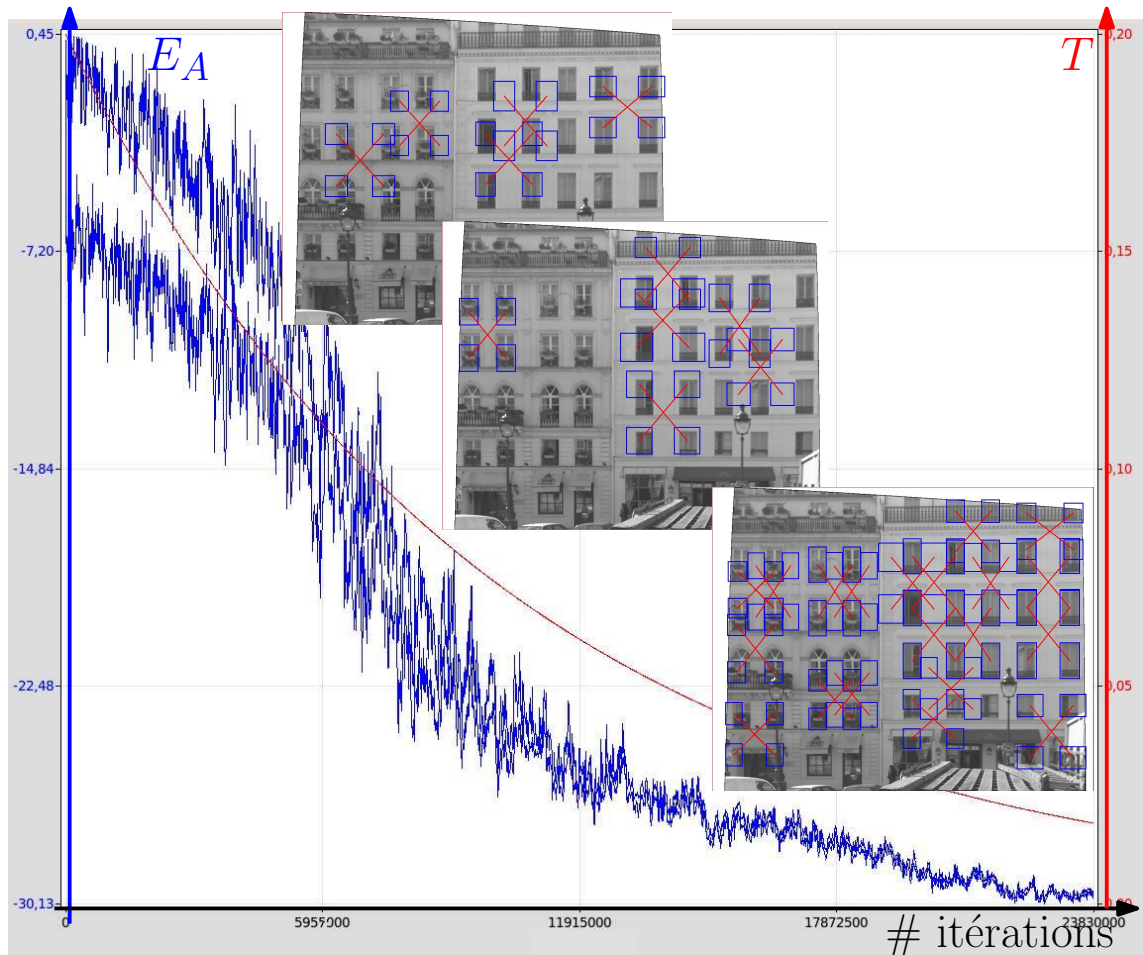


FIGURE 5.12 – Processus d'optimisation de l'agencement des structures répétitives : courbe (en bleu) de l'énergie E_A à minimiser ; courbe (en rouge) de la température ; Trois images de résultats de l'agencement optimal

5.4 Conclusion

Nous venons de présenter une solution globale à l'agencement de grilles de structures répétitives dans les images de façades. Notre méthode est automatique. Son temps de calcul est de l'ordre de la minute pour un processeur de 2,83 GHz. Son caractère innovant dans l'état de l'art est qu'elle autorise les objets recherchés à se superposer. Nous avons fait le choix de préférer combiner des grilles d'un petit nombre de structures répétitives plutôt que de détecter de plus grandes directement.

Les résultats que nous obtenons peuvent aboutir à la mise au point d'une mesure de la répétitivité des structures pouvant guider la segmentation hiérarchique (chapitre 3). Ils pourraient également servir de base pour l'appariement d'images à partir d'un graphe de grilles de structures.

Nous pourrions améliorer le processus en adaptant la distribution de probabilité de référence à notre problématique. Nous l'avons choisie uniforme sur l'espace des grilles mais elle pourrait être tout à fait être constituée de gaussiennes centrées sur les alignements dominants de l'image pour accélérer le traitement.

Enfin, le choix que nous avons fait de contraindre les structures dans les alignements réguliers pourrait être adapté. Il serait envisageable d'utiliser une autre contrainte pour rechercher les structures : dans des grilles incomplètes, par exemple (grilles dont certains nœuds ne seraient pas traités).

Conclusion

Nous venons de présenter les avancées réalisées durant la thèse sur la segmentation d'images de façades de bâtiments acquises du point de vue de la rue. Les traitements que nous proposons s'appliquent à des images isolées pouvant comporter une ou plusieurs façades, en partie ou bien dans leur intégralité. Le réglage des paramètres de nos algorithmes peut s'effectuer en une fois pour s'appliquer à un grand volume d'images comportant des façades de styles d'architecture différents. En cela, nos travaux proposent des solutions pertinentes vis à vis de la problématique de la thèse.

Nous avons détaillé deux approches asémantiques reposant sur les propriétés radiométriques et géométriques des façades comme l'alignement et la répétitivité de leurs structures. La première consiste en une recherche récursive de modèles élémentaires *2D* suivant un critère de séparation optimale des contours de l'image. Nous appelons ce procédé *segmentation hiérarchique*. La deuxième tend à détecter les principales grilles de structures similaires et contrastées sur le fond du mur. Nous parlons d'approches asémantiques car le processus ne fournit pas d'information sémantique sur les structures localisées. Notre objectif, en effet, n'est pas de reconnaître un étage, une fenêtre ou bien une porte. Il s'agit de générer des indices architecturaux à partir d'heuristiques reposant sur le moins d'*a priori* possible.

Nos travaux comportent un caractère innovant dans de nombreux domaines :

La nature des données Nos outils de segmentation automatique s'appliquent à de gros volumes d'images présentant une variété architecturale importante. Le nombre et la position des façades sont inconnus. Nous avons vu que l'état de l'art ne s'est attaché que très faiblement à cette problématique qui se situe en amont de la plupart des approches sémantiques.

L'approche contours La distance que nous avons formalisée à partir des densités de contours de l'image pour guider la segmentation hiérarchique est une avancée notable dans le domaine de la séparation de façades et de ses structures. Elle localise leurs changements de topologie et repose sur une compensation statistique du problème des bords dont le calcul est très rapide.

Le coût d'un agencement de structures répétitives La localisation globale des structures répétitives de l'image à partir de petits quadrillages recouvrants est une innovation dans l'optimisation stochastique de ce type de structures.

Nous relevons dans cette conclusion les points forts des éléments de solution que nous proposons pour répondre à la problématique. Nous montrons également dans quelle mesure ils restent incomplets et synthétisons alors quelques perspectives pour leur approfondissement.

6.1 Points forts de nos travaux

La rapidité des calculs et la cohérence des structures détectées par nos approches constituent le point fort de nos travaux.

6.1.1 Rapidité des calculs

Un gros atout de nos approches est leur faible complexité en temps de calcul. Les étapes les plus longues sont les pré-traitements pour le redressement de l'image, le sous-échantillonnage et l'extraction du gradient. Les étapes suivantes sont plus rapides.

Nous utilisons ce que nous appelons les *alignements dominants* de la façade qui permettent de travailler dans un espace restreint de l'image.

Nous nous sommes aussi attachés, lorsque l'occasion se présentait, à traiter le problème indépendamment dans chacune des deux dimensions principales de l'image. Ces traitements indépendants concernent la détection des alignements dominants, l'analyse des hypothèses de découpe et des modèles de la segmentation hiérarchique ainsi que les calculs d'intersection d'aires des structures répétitives.

Nous utilisons des images intégrales pour rendre les calculs d'espérances et de variances radiométriques indépendants de la surface des régions rectangulaires concernées. Il est également possible de travailler sur des images sous-échantillonnées.

6.1.2 Localisation de structures cohérentes

L'ordre des découpes de la segmentation hiérarchique peut varier selon l'échelle et la localisation de la façade dans l'image mais chacun des modèles élémentaires encadre une structure cohérente. La façade est séparée de ses éventuelles voisines et de son environnement.

La localisation des structures répétitives peut s'effectuer de manière semi-automatique à partir de cartes de scores optimaux indiquant, pour chaque pixel de l'image, quelle est la grille de structures optimale au sens de l'énergie de pertinence que nous avons définie.

L'énergie de coût d'un agencement de ces structures répétitives fournit une solution globale, rapide et automatique dans toute l'image.

6.2 Limites de nos travaux

Le manque d'évaluation quantitative de nos travaux reste un des points principaux à approfondir dans la suite. Il existe également quelques faiblesses dans la formalisation de la segmentation.

6.2.1 Manque d'évaluation des résultats

D'une manière générale, il manque encore une évaluation quantitative à partir de vérités terrain sur les structures localisées. Cela est difficile pour la segmentation hiérarchique mais tout à fait envisageable pour l'optimisation d'agencement des grilles de structures répétitives.

La robustesse au point de vue d'acquisition reste également à démontrer de manière quantitative.

6.2.2 Faiblesses des critères de localisation des structures

Il existe une tendance à la sur-segmentation dans les découpes profondes de la segmentation hiérarchique (au delà du troisième niveau de profondeur). Nous avons proposé quelques pistes pour atténuer ce phénomène. D'autre part, pour cette même approche, il peut arriver que des structures de même type soient séparées différemment sur une même image. L'approche manque d'information globale de périodicité des structures de la façade pour pouvoir séparer celles qui sont semblables avant de les comparer.

La formalisation des critères de pertinence de la localisation des grilles de structures répétitives permet des calculs très rapides mais aboutit encore à de fausses détections. Nous avons proposé des solutions de perfectionnement de cette formalisation.

6.3 Perspectives

Nous synthétisons dans cette section les pistes de perfectionnement des outils de segmentation ainsi que quelques propositions d'évolution de notre stratégie.

6.3.1 Perfectionnement général de la modélisation

La formalisation des modèles élémentaires de la segmentation hiérarchique pourrait être affinée. Nous avons proposé différentes solutions (ajout d'autres modèles comme des symétries, analyse d'histogramme des radiométries, utilisation d'autres primitives bas niveau, etc.). D'une manière générale, les modèles de structures que le processus recherche sont rectangulaires : il pourrait être intéressant d'utiliser aussi des courbes pour tenir compte des porches, oculus, arabesques, frontons flamands, etc. Nous pourrions également utiliser l'image de résidus entre les données et les modèles pour accroître encore davantage le niveau de détail de la modélisation.

Concernant le mode de découpe de cette même approche, nous avons proposé quelques pistes de perfectionnement de la classification des arêtes de contours et de la pondération des distances locale et globale entre les régions à séparer. Nous avons également discuté du choix de la stratégie générale à appliquer. Une fois l'image découpée, ne faudrait-il pas effectuer une deuxième passe de fusion des régions pour éviter la sur-segmentation ? Cette deuxième passe devrait-elle s'effectuer indépendamment de la première ou bien à chaque itération de l'algorithme ?

Concernant la construction de l'agencement optimal des structures répétitives, nous avons proposé diverses améliorations. Il serait ainsi envisageable d'accélérer

la convergence de l'échantillonnage en adaptant la distribution de probabilité de référence à l'hypothèse d'alignement des structures. Nous insistons également sur l'aspect provisoire de la régularisation de ces structures sous forme de quadrillage : il est tout à fait envisageable de relaxer cette contrainte assez rigide en recherchant d'autres types de topologies.

6.3.2 Utilisation conjointe des deux approches

La segmentation hiérarchique découpe l'image de manière dichotomique indépendamment des informations de répétitivité des structures. L'appariement au modèle périodique ne s'effectue qu'une fois la région découpée. Il serait très intéressant d'utiliser l'agencement optimal des grilles de structures répétitives pour concevoir une mesure de périodicité qui guiderait la découpe.

6.3.3 Utilisation d'informations d'autres natures

Nous n'avons travaillé jusqu'à présent que sur le niveau de gris des images. Nous avons vu qu'il était tout à fait possible d'utiliser l'information de la couleur ou de la profondeur des structures.

L'adaptation, en particulier, des algorithmes en $2, 5D$ (pour le modèle périodique de la segmentation hiérarchique ou la détection des grilles de structures répétitives) pourrait passer par l'utilisation de ce que nous appellerions des *plans dominants*, dans la même idée que les *alignements dominants*. Il s'agirait de plans secondaires (celui de la balustrade des balcons, des vitres des fenêtres, etc.), parallèles au plan principal de la façade.

6.3.4 Application à l'indexation

Les résultats obtenus par les deux approches peuvent être utilisés comme indices géométriques en vue d'un appariement de façades.

Ces indices géométriques pourraient se composer des premières découpes générées par la segmentation hiérarchique ou bien d'un graphe des grilles de structures répétitives. Ils pourraient d'autre part être couplés à des indices plus locaux comme des ensembles de points d'intérêt (de type *SIFT* par exemple). Il s'agirait alors de comparer des regroupements cohérents de points entre deux images.

Nous pourrions également envisager de travailler non plus sur une seule image mais sur plusieurs en même temps pour rendre le processus plus robuste au point de vue d'acquisition.

6.3.5 Application à la compression

Il serait intéressant d'adapter nos approches à une problématique de compression de la texture de façade.

Concernant la segmentation hiérarchique, il serait envisageable de concevoir une distance normalisée entre l'image et les modèles permettant de guider la compression. Nous avons vu qu'il était possible d'adapter l'algorithme de découpe pour minimiser la taille des modèles générés.

Parallèlement, la détection des structures répétitives pourrait permettre de factoriser l'information de l'image.

6.3.6 Application à la sémantisation

Enfin, la problématique de cette thèse concerne des traitements à faibles *a priori*. Il est tout à fait envisageable de l'orienter vers la reconnaissance de structures de plus haut niveau en leur assignant une *classe* d'élément architectural (*étage, fenêtre, mur de façade*, etc.).

Plus particulièrement, l'arborescence de régions que fournit la segmentation hiérarchique suit le même principe que les approches de type *grammaire*. Il serait intéressant de concevoir un processus d'apprentissage automatique de règles pour ces grammaires (hauteur des éléments, distance entre eux, leur topologie, etc.).

Calculs combinatoires

Cette section fournit la démonstration du calcul des sommes de type $H^i(n)$ et $S_p^i(n)$, définies par l'équation A.1, respectivement utilisées dans les calculs 3.2 et 3.3 de l'espérance des distances. Ils ont gracieusement été entrepris par Bruno Vallet, chargé de recherche au laboratoire *MATIS* de l'*IGN*.

$$\begin{cases} H^i(n) = \sum_{k=0}^{\lfloor n/2 \rfloor} k^i C_n^k \\ S_p^i(n) = \sum_{k=0}^n k^i C_n^k (1-p)^k p^{n-k} \end{cases} \quad (\text{A.1})$$

L'équation A.2 fournit le résultat des premiers termes.

$$\begin{cases} H^0(2n) = \frac{1}{2}(2^{2n} + C_{2n}^n) \\ H^0(2n+1) = 2^{2n} \\ H^1(2n) = n2^{2n-1} \\ H^1(2n+1) = \frac{2n+1}{2}(2^{2n} - C_{2n}^n) \end{cases} \quad \begin{cases} S_p^0(n) = 1 \\ S_p^1(n) = np \\ S_p^2(n) = np(1-p+np) \end{cases} \quad (\text{A.2})$$

A.1 Démonstration du premier terme de H

Nous proposons une démonstration détaillée du calcul de $H^0(n)$ dans le cas pair. Le calcul dans le cas impair peut se retrouver suivant la même démarche.

$$\begin{aligned} H^0(2n) &= \sum_{k=0}^n C_{2n}^k = \frac{1}{2} \left[\left(\sum_{k=0}^n C_{2n}^k \right) + \left(\sum_{k=0}^n C_{2n}^k \right) \right] \\ H^0(2n) &= \frac{1}{2} \left[\left(\sum_{k=0}^n C_{2n}^k \right) + \left(\sum_{k=0}^n C_{2n}^{2n-k} \right) \right] = \frac{1}{2} \left[\left(\sum_{k=0}^n C_{2n}^k \right) + \left(\sum_{k=n}^{2n} C_{2n}^k \right) \right] \\ H^0(2n) &= \frac{1}{2} \left[\left(\sum_{k=0}^{2n} C_{2n}^k \right) + C_{2n}^n \right] = \frac{1}{2} \left[\left(\sum_{k=0}^{2n} C_{2n}^k 1^k 1^{2n-k} \right) + C_{2n}^n \right] \end{aligned}$$

Le résultat de l'équation A.2 s'obtient alors immédiatement par la factorisation de la somme $\sum_{k=0}^{2n} C_{2n}^k 1^k 1^{2n-k}$ en un binôme de Newton $(1+1)^{2n}$. De même, dans le cas impair : $H^0(2n+1) = \sum_{k=0}^n C_{2n+1}^k = \frac{1}{2} [\sum_{k=0}^{2n+1} C_{2n+1}^k 1^k 1^{n-k}] = \frac{1}{2} [(1+1)^{2n+1}]$.

A.2 Démonstration du deuxième terme de H

Par définition, $H^1(n) = \sum_{k=0}^{\lfloor n/2 \rfloor} kC_n^k$. Commençons par éliminer l'indice k en facteur dans cette somme. Cela est réalisable quelle que soit la borne maximale de cet indice. Appelons m , cette borne.

$$\begin{aligned}
 \sum_{k=0}^m kC_n^k &= \sum_{k=1}^m \frac{n!}{(k-1)!(n-k)!} = \sum_{k=0}^{m-1} \frac{n!}{k!(n-k)!} (n-k) = \sum_{k=0}^{m-1} C_n^k (n-k) \\
 \sum_{k=0}^m kC_n^k &= n \left[\left(\sum_{k=0}^m C_n^k \right) - C_n^m \right] - \left[\left(\sum_{k=0}^m kC_n^k \right) - mC_n^m \right] \\
 \sum_{k=0}^m kC_n^k &= n \left(\sum_{k=0}^m C_n^k \right) - \left(\sum_{k=0}^m kC_n^k \right) + (m-n)C_n^m \\
 \sum_{k=0}^m kC_n^k &= \frac{1}{2} \left[n \left(\sum_{k=0}^m C_n^k \right) + (m-n)C_n^m \right] \tag{A.3}
 \end{aligned}$$

Alors, d'après l'équation A.3 :

$$H^1(2n) = \frac{1}{2} \left[2n \left(\sum_{k=0}^n C_{2n}^k \right) + (n-2n)C_{2n}^n \right] = \frac{1}{2} [2nH^0(2n) - nC_{2n}^n]$$

Le développement du terme $H^0(2n)$ aboutit alors immédiatement au résultat A.2. La même démarche permet de retrouver le résultat dans le cas impair :

$$\begin{aligned}
 H^1(2n+1) &= \frac{1}{2} \left[(2n+1) \left(\sum_{k=0}^n C_{2n+1}^k \right) + (n-(2n+1))C_{2n+1}^n \right] \\
 H^1(2n+1) &= \frac{1}{2} [(2n+1)H^0(2n+1) - (n+1)C_{2n+1}^n]
 \end{aligned}$$

A.3 Démonstration du calcul de S

Les termes de la suite S^i peuvent se calculer à partir des dérivations successives du développement en somme de C_n^k du binôme de Newton $(x+y)^n$ exprimées dans l'équation A.4. À chaque étape, les deux membres de l'équation sont dérivés par rapport à la variable x puis multipliés par cette dernière pour compléter la somme.

$$\begin{cases} (x+y)^n &= \sum_{k=0}^n C_n^k x^k y^{n-k} \\ nx(x+y)^{n-1} &= \sum_{k=0}^n kC_n^k x^k y^{n-k} \\ nx(nx+y)(x+y)^{n-2} &= \sum_{k=0}^n k^2 C_n^k x^k y^{n-k} \end{cases} \tag{A.4}$$

Posons $x = p$ et $y = 1-p$. Les relations A.4 donnent immédiatement le résultat :

$$\begin{cases} S_p^0(n) = \sum_{k=0}^n C_n^k (1-p)^k p^{n-k} = (p+1-p)^n \\ S_p^1(n) = \sum_{k=0}^n kC_n^k p^k (1-p)^{n-k} = np(p+1-p)^{n-1} \\ S_p^2(n) = \sum_{k=0}^n k^2 C_n^k p^k (1-p)^{n-k} = np(np+1-p)(p+1-p)^{n-2} \end{cases}$$

Publications de l’auteur

- Burochin, J.-P., Tournaire, O., and Paparoditis, N. (2009). An Unsupervised Hierarchical Segmentation of a Facade Building Image in Elementary 2D-Models. In *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XXXVIII (Part 3/W4), pages 223–228, Paris, France.
- Burochin, J.-P., Vallet, B., Tournaire, O., and Paparoditis, N. (2010). A Formulation for Unsupervised Hierarchical Segmentation of Facade Images with Periodic Models. In *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XXXVIII (Part 3 A), pages 227–232, Saint-Mandé, France.

Bibliographie

- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. Dans *Proc. IEEE International Symposium on Information Theory*, pages 267–281, San Antonio, Texas, États Unis.
- Alegre, F. et Dellaert, F. (2004). A Probabilistic Approach to the Semantic Interpretation of Building Facades. Dans *Proc. International Workshop on Vision Techniques Applied to the Rehabilitation of City Centres*, Lisbon, Portugal.
- Ali, H., Seifert, C., Jindal, N., Paletta, L., et Paar, G. (2007). Window Detection in Facades. Dans *Proc. IAPR International Conference on Image Analysis and Processing*, volume 14, pages 837–842, Modène, Italie. International Association on Pattern Recognition.
- Anagnostopoulos, C.-N., Koutsonas, A., Anagnostopoulos, L., Loumos, V., et Karyafas, E. (2005). Tile classification using the CIELAB color model. Dans *Proc. International Conference on Computational Science*, volume 3516, pages 695–702, Atlanta, États-Unis.
- Bochko, V. et Petrou, M. (2007). Recognition the structural parts of a building using support vector machines. Dans *Proc. International Conference on Pattern Recognition and Information Processing*, Minsk, Biélorussie.
- Bris, A. I. et Paparoditis, N. (2010). Matching terrestrial images captured by a nomad system to images of a reference database for pose estimation purpose. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII (Part 3A) :133–138.
- Brédif, M. et Tournaire, O. (2012). libjmmcmc : an open-source generic c++ library for stochastic optimization. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIX (part 3) :—.
- Burochin, J.-P., Tournaire, O., et Paparoditis, N. (2009). An Unsupervised Hierarchical Segmentation of a Facade Building Image in Elementary 2D-Models. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII (3/W4) :223–228.

- Burochin, J.-P., Vallet, B., Tournaire, O., et Paparoditis, N. (2010). A Formulation for Unsupervised Hierarchical Segmentation of Facade Images with Periodic Models. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII (Part 3A) :227–232.
- Casella, G. et Robert, C. P. (2004). *Monte Carlo Statistical Methods*. Springer Verlag, New York, États-Unis, 2e édition.
- Chazelle, B. M., Drysdale, R. L. S., et Lee, D. T. (1986). Computing the largest empty rectangle. *Journal On Computing*, 15(1) :300–315.
- Chomsky, N. (1956). Three Models for the description of Language. *IRE Transactions on Information Theory*, 2 :113–123.
- Descombes, X. (2011). *Stochastic geometry for image analysis*. Iste, Londres, Angleterre.
- Dick, A. R., Torr, P. H. S., et Cipolla, R. (2004). Modelling and interpretation of architecture from several images. *International Journal of Computer Vision*, 60(2) :111–134.
- Drauschke, M. (2008). Selecting appropriate features for detecting buildings and building parts. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVII (Part B3b-1) :447–452.
- Drauschke, M. (2009). An Irregular Pyramid for Multi-scale Analysis of Objects and their Parts. Dans *Proc. Workshop on Graph-based Representations in Pattern Recognition*, pages 293–303, Venise, Italie.
- Drauschke, M. et Mayer, H. (2010). Evaluation of texture energies for classification of facade images. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII (Part 3A) :257–262.
- Gorelick, L., DeLong, A., Veksler, O., et Boykov, Y. (2011). Recursive MDL via Graph Cuts : Application to Segmentation. Dans *Proc. IEEE International Conference on Computer Vision*, Barcelone, Espagne.
- Green, P. J. (1995). Reversible Jump Markov Chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82(4) :711–732.
- Gupta, A., Efros, A. A., et Hebert, M. (2010). Blocks world revisited : Image understanding using qualitative geometry and mechanics. Dans *Proc. European Conference on Computer Vision*, Heraklion, Grèce.
- Han, F. et Zhu, S.-C. (2005). Bottom-up/top-down image parsing by attribute graph grammar. Dans *Proc. IEEE International Conference on Computer Vision*, pages 1778–1785, Pékin, Chine.
- Harris, C. et Stephens, M. (1988). A combined corner and edge detector. Dans *Alvey Vision Conference*, volume 15, page 50, Manchester, Royaume-Uni.

- Hartz, J. et Neumann, B. (2009). Learning a Knowledge Base of Ontological Concepts for High-Level Scene Interpretation. Dans *Proc. International Conference on Machine Learning and Applications*, Miami Beach, Floride, États-Unis.
- Hastings, W. K. (1970). Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika*, 57(1) :97–109.
- Haugeard, J.-E. et Philipp-Foliguet, S. (2010). Recherche d’objets par appariement de graphes combinant contours et régions. Dans *Proc. congrès sur la Reconnaissance des Formes et Intelligence Artificielle*, Caen, France.
- Haugeard, J.-E., Philipp-Foliguet, S., et Precioso, F. (2009). Extraction of Windows in Facade Using Kernel on Graph of Contours. Dans *Proc. Scandinavian Conference on Image Analysis*, Oslo, Norvège.
- Heesch, D. et Petrou, M. (2007). Non-Gibbsian Markov random field models for contextual labelling of structured scenes. Dans *Proc. British Machine Vision Conference*, University of Warwick, Royaume-Uni.
- Hernández, J. et Marcotegui, B. (2009). Morphological Segmentation Of Building Facade Images. Dans *Proc. IEEE International Conference on Image Processing*, le Caire, Egypte.
- Hernández, J., Marcotegui, B., et Meyer, F. (2009). Morphological facade image analysis for building modeling. Dans *International Society of Stereology*, Paris, France.
- Jahangiri, M. et Petrou, M. (2008). Fully Bottom-Up Blob Extraction in Building Facades. Dans *Proc. International Conference on Pattern Recognition and Image Analysis*, Nizhny Novgorod, Russie.
- Jong, K. (2006). *Evolutionary Computation : A Unified Approach*. MIT Press, Cambridge, Royaume-Uni.
- Kalantari, M., Jung, F., Paparoditis, N., et Guédon, J.-P. (2008). Robust and Automatic Vanishing Points Detection with their Uncertainties from a Single Uncalibrated Image, by Planes Extraction on the Unit Sphere. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVII (Part3A) :203–208.
- Karantzas, K., Simon, L., Teboul, O., Koutsourakis, P., et Paragios, N. (2008). Une étude sur l’architecture de paris : styles, architectes et grandes constructions. Technical report, Laboratoire de Mathématiques appliquées aux systèmes, École Centrale de Paris.
- Korah, T. et Rasmussen, C. (2008). Analysis of building textures for reconstructing partially occluded facades. Dans *Proc. European Conference on Computer Vision*, number 14, pages 359–372, Marseille, France.

- Korč, F. et Förstner, W. (2008a). Approximate Parameter Learning in Conditional Random Fields : An Empirical Investigation. Dans *Proc. Annual Symposium of the German Association for Pattern Recognition (DAGM)*, Munich, Allemagne.
- Korč, F. et Förstner, W. (2008b). Finding Optimal Non-Overlapping Subset of Extracted Image Objects. Dans *Proc. International Workshop on Combinatorial Image Analysis*, Buffalo, États-Unis.
- Koutsourakis, P., Simon, L., Teboul, O., Tziritas, G., et Paragios, N. (2009). Single View Reconstruction Using Shape Grammars for Urban Environments. Dans *Proc. IEEE International Conference on Computer Vision*, Heraklion, Grèce.
- Lacoste, C., Descombes, X., et Zerubia, J. (2010). Unsupervised line network extraction in remote sensing using a polyline process. *Pattern Recognition*, 43(4) :1631–1641.
- Lafarge, F., Descombes, X., Zerubia, J., et Pierrot-Deseilligny, M. (2010). Structural approach for building reconstruction from a single DSM. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(1) :135–147.
- Lee, S. C. et Nevatia, R. (2004). Extraction and Integration of Window in a 3D Building Model from Ground View images. Dans *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 113–120, Washington DC., États-Unis.
- Leibe, B., Leonardis, A., et Schiele, B. (2004). Combined Object Categorization and Segmentation with an Implicit Shape Model. Dans *Proc. European Conference on Computer Vision*, Prague, République Tchèque.
- Leung, T. et Malik, J. (1996). Detecting, localizing and grouping repeated scene elements from an image. Dans *Proc. European Conference on Computer Vision*, Cambridge, Angleterre.
- Li, Y. et Shapiro, L. G. (2002). Consistent line clusters for building recognition in cbir. Dans *Proc. International Conference on Pattern Recognition*, Québec, Canada.
- Liu, C. et Gagalowicz, A. (2010). Image-based Modeling of Haussmannian Facades. *International Journal of Virtual Reality*, 9 :13–18.
- Liu, Y. et Collins, R. (2001). Skewed symmetry groups. Dans *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 872–879, Hawaï, États-Unis.
- Liu, Y., Lin, W.-C., et Hays, J. H. (2004). Near-regular texture analysis and manipulation. Dans *Proc. SIGGRAPH conference*, Los Angeles, Californie, États-Unis.
- Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2) :91–110.

- Mignot, C. (2004). *Grammaire des immeubles parisiens, Six siècles de façades du Moyen-Âge à nos jours*. Parigramme, Paris, France.
- Mičušík, B., Wildenauer, H., et Košecká, J. (2008). Detection and matching of rectilinear structures. Dans *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Anchorage, États-Unis.
- Müller, P., Zeng, G., Wonka, P., et Gool, Luc, V. (2007). Image-based Procedural Modeling of Facades. *ACM Transactions on Graphics*, 26(3) :85.
- Ortner, M., Descombes, X., et Zerubia, J. (2007). Building Outline Extraction from Digital Elevation Models using Marked Point Processes. *International Journal of Computer Vision*, 72(2) :107–132.
- Paparoditis, N. (2007). Systèmes de vision 3d en géomatique : Numérisation et modélisation 3d des environnements urbains construits. Habilitation à Diriger des Recherches, Université René Descartes - Paris V.
- Petrou, M. et Sevilla, P. G. (2006). *Image Processing : Dealing with Texture*. Wiley, New York, États-Unis.
- Recky, M. et Leberl, F. (2009). Semantic Segmentation of Street-Side Images. Dans *Proc. Annual Workshop of the Austrian Association for Pattern Recognition (OAGM/AAPR)*, Stainz, Australie.
- Recky, M. et Leberl, F. (2010). Detection Using K-means in CIE-Lab Color Space. Dans *Proc. International Conference on Pattern Recognition*, pages 356–360, Istanbul, Turquie.
- Recky, M., Wendel, A., et Leberl, F. (2011). Façade Segmentation in a Multi-View Scenario. Dans *Proc. TDPVT*, Hangzhou, Chine.
- Reznik, S. et Mayer, H. (2008). Implicit Shape Models, Self-Diagnosis, and Model Selection for 3D Facade Interpretation. Revised and extended version of (Reznik & Mayer 2007) presented at the ISPRS conference “Photogrammetric Image Analysis”, Munich, Germany, September 19–21, 2007.
- Ricard, J., Royan, J., et Aubault, O. (2008). Visualization of real cities based on procedural modeling. Dans *Proc. Workshop on Virtual Cityscapes*, Reno, États-Unis.
- Ripperda, N. (2008). Determination of Facade Attributes for Facade Reconstruction. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVII (Part3A) :285–290.
- Rissanen, J. (1983). A universal prior for integers and estimation by minimum description length. *The Annals of Statistics*, 11(2) :416–431.

- Robertson, D. et Cipolla, R. (2004). An image-based system for urban navigation. Dans *Proc. British Machine Vision Conference*, Kingston University, Royaume-Uni.
- Salamon, P., Sibani, P., et Frost, R. (2002). *Facts, Conjectures, and Improvements for Simulated Annealing*. Number 7. Society for Industrial and Applied Mathematics, Philadelphie, États-Unis.
- Schaffalitzky, F. et Zisserman, A. (1998). Geometric grouping of repeated elements within images. Dans *Proc. British Machine Vision Conference*, Southampton, Royaume-Uni.
- Schindler, G., Krishnamurthy, P., Lubliner, R., Liu, Y., et Dellaert, F. (2008). Detecting and matching repeated patterns for automatic geo-tagging in urban environments. Dans *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Anchorage, États-Unis.
- Schindler, K. et Joachim, B. (2003). A model-based method for building reconstruction. Dans *Proc. International Workshop on Higher-Level Knowledge in 3D Modeling and Motion Analysis*, Nice, France.
- Simon, L., Teboul, O., Koutsourakis, P., et Paragios, N. (2010). Random Exploration of the Procedural Space for Single-View 3D Modeling of Buildings. *International Journal of Computer Vision*, 93(2) :253–271.
- Sirmacek, B. (2011). Graph Theory and Mean Shift Segmentation Based Classification of Building Facades. Dans *Proc. Joint Urban Remote Sensing Event*, pages 409–412, Munich, Allemagne.
- Teboul, O. (2011). *Shape Grammar Parsing : Application to Image-based Modeling*. PhD thesis, École Centrale de Paris, MAS laboratory.
- Teboul, O., Simon, L., Koutsourakis, P., et Paragios, N. (2010). Segmentation of Building Facades Using Procedural Shape Priors. Dans *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, San Francisco, États-Unis.
- Tournaire, O., Brédif, M., Boldo, D., et Durupt, M. (2010). An efficient stochastic approach for building footprint extraction from digital elevation models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(4) :317–327.
- Vallet, B., Pierrot-Deseilligny, M., et Boldo, D. (2009). Building footprint database improvement for 3d reconstruction : a direction aware split and merge approach. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII (Part 3/W4) :139–144.
- Van Gool, L., Zeng, G., Borre, Filip, V. d., et Müller, P. (2007). Towards mass-produced building models. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVI (Part 3A) :209–220.

- Čech, J. et Sára, R. (2007). Windowpane Detection based on Maximum A Posteriori Probability Labeling. Technical report, Center for Machine Perception, Czech Technical University.
- Viola, P. et Jones, M. J. (2004). Robust Real-Time Face Detection. *International Journal of Computer Vision*, 57(2) :137–154.
- Wang, X., Totaro, S., Taillandier, F., Hanson, A. R., et Teller, S. (2002). Recovering Facade Texture and Microstructure from Real-World Images. Dans *Proc. International Workshop on Texture Analysis and Synthesis at ECCV*, pages 381–386, Copenhagen, Denmark.
- Wendel, A., Donoser, M., et Bischof, H. (2010). Unsupervised facade segmentation using repetitive patterns. Dans *Proc. Annual Symposium of the German Association for Pattern Recognition (DAGM)*, Darmstadt, Allemagne.
- Wenzel, S., Drauschke, M., et Förstner, W. (2008). Detection of repeated structures in facade images. *Pattern Recognition and Image Analysis*, 18(3) :406–411.
- Zeng, G. et Gool, Luc, V. (2008). Multi-label Image Segmentation via Point-wise Repetition. Dans *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Anchorage, États-Unis.
- Zhang, H., Xiao, J., et Quan, L. (2010). Supervised Label Transfer for Semantic Segmentation of Street Scenes. Dans *Proc. European Conference on Computer Vision*, pages 561–574, Heraklion, Grèce.
- Zhang, W. et Košecká, J. (2005). Localization based on building recognition. Dans *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, San Diego, États-Unis.
- Zhang, W. et Košecká, J. (2006). Hierarchical Building Recognition. Technical report, Department of Computer Science, George Mason University, États-Unis.
- Zhao, P., Fang, T., Xiao, J., Zhang, H., Zhao, Q., et Quan, L. (2010). Rectilinear Parsing of Architecture in Urban Environment. Dans *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, San Francisco, États-Unis.

Table des figures

1.1	Applications de la thèse	14
1.2	Niveaux et travées	15
1.3	Structures de la façade	16
1.4	Baies	16
1.5	Ornements	17
1.6	Diversité des styles de façades 1	18
1.7	Structures de consolidation	19
1.8	Variation des baies	20
1.9	Variation des contrevents	20
1.10	Diversité des styles de façades 2	21
1.11	Système de numérisation <i>Stéréopolis</i>	22
1.12	Image panoramique de la rue	23
1.13	Projection des contours dans la sphère de Gauss	25
1.14	Redressement multi-planaire	25
2.1	Sémantisation par la position	28
2.2	Hypothèses de fenêtres	29
2.3	Profils de projection d'informations locales	31
2.4	Grammaire	33
2.5	Apparence d'une fenêtre	34
2.6	Séparation de façades	35
2.7	Profondeur des structures de la façade	36
2.8	Appariement de façades	37
2.9	Segmentations génériques	40
3.1	Segmentation hiérarchique : algorithme	44
3.2	Alignements dominants	46
3.3	Exemple de modèles radiométriques	48
3.4	Distances entre alignements dominants	49
3.5	Appariement de points périodiques	50
3.6	Modèle périodique	51
3.7	Coin de fenêtre en perspective	52
3.8	Segmentation hiérarchique : distance globale	56

3.9	Distance globale : cas particuliers	57
3.10	Segmentation hiérarchique : paramètre α	59
3.11	Distance globale : allure générale de la correction	60
3.12	Estimation de l'occurrence des arêtes de contour	61
3.13	Distance globale : comparaison des espérances théoriques et empiriques	63
3.14	Distance globale : formalisation primitive	65
3.15	Segmentation hiérarchique : comparaison des distances globales	65
3.16	Segmentation hiérarchique : résultat 1	67
3.17	Segmentation hiérarchique : résultat 2	68
3.18	Segmentation hiérarchique : résultat 3	69
3.19	Segmentation hiérarchique : variation du point de vue	69
3.20	Segmentation hiérarchique : résidus des modèles	70
3.21	Poids des arêtes	71
3.22	Découpe selon la profondeur	72
3.23	Pénalisation des périodicités	73
3.24	Segmentation hiérarchique : résultat de compression	75
3.25	Segmentation hiérarchique pour un rendu 3D	77
3.26	Modélisation en facettes 3D	78
3.27	Modélisation en textures 3D : vue de face	79
3.28	Modélisation en textures 3D : vue de profil	80
4.1	Schéma d'une grille de structures répétitives	82
4.2	Profil radiométriques des structures répétitives	84
4.3	Exemple de résultats du <i>CFI</i>	85
4.4	Contraste local <i>vs</i> global	86
4.5	Problème de variation d'éclairage	86
4.6	Maximisations des trois énergies des grilles de répétitions	87
4.7	Grilles de structures répétitives : répartition de la similarité	88
4.8	Grilles de structures répétitives : répartition du contraste	88
4.9	Grilles de structures répétitives : répartition de l'uniformité 1	89
4.10	Grilles de structures répétitives : répartition de l'uniformité 2	89
4.11	Grilles optimales de structures répétitives 1	93
4.12	Grilles optimales de structures répétitives 2	94
4.13	Grilles optimales de structures répétitives 3	95
4.14	Histogrammes d'orientation	96
4.15	Utilisation de la profondeur	97
5.1	Pénalisation de la sur-segmentation	100
5.2	Pénalisation des chevauchements de boîtes englobantes	102
5.3	Pénalisation des chevauchements de nœuds	102
5.4	Liaisons faibles	103
5.5	Régimes stationnaires (définition 5.12) à différentes températures.	109
5.6	Espace d'échantillonnage d'une grille	111
5.7	Variation du nombre de nœuds	114
5.8	Translation d'un côté des nœuds	115

5.9	Compression et dilatation des grilles	116
5.10	Agencement de grilles de structures répétitives 1	118
5.11	Agencement de grille de structures répétitives 2	118
5.12	Agencement de grille de structures répétitives 3	119

Table des matières

1	Introduction	13
1.1	Applications	13
1.2	Architecture de la façade	14
1.2.1	Tendances architecturales fréquentes	14
1.2.2	Diversité des structures	15
1.2.3	Mouvements de la façade	19
1.2.4	Évolution permanente des styles d’architecture	21
1.3	Données	22
1.3.1	Données du système Stéréopolis	23
1.3.2	Sources d’instabilité liées à l’acquisition	23
1.3.3	Redressement multi-planaire	24
1.4	Problématique et plan du manuscrit	26
2	État de l’art	27
2.1	Étiquetage des éléments de la façade	28
2.1.1	La position	28
2.1.2	L’apparence	34
2.2	Séparation des façades	35
2.3	Estimation des profondeurs	36
2.4	Appariement de façades	37
2.5	Segmentation générique	38
2.5.1	Sélection de traitements génériques	38
2.5.2	Notre contribution	39
3	Segmentation hiérarchique	43
3.1	Alignements dominants	45
3.1.1	Mode de calcul	45
3.1.2	Propriétés intéressantes	45
3.1.3	Discussion sur la détection et la localisation	46
3.2	Reconnaissance de modèles	46
3.2.1	Modèles planaires et cylindriques	47
3.2.2	Modèle périodique	48

3.3	Processus de découpe dichotomique	53
3.3.1	Espace des arêtes	53
3.3.2	Classification des arêtes	54
3.3.3	Définition de la distance locale	54
3.3.4	Définition de la distance globale	55
3.3.5	Correction de la distance globale	56
3.3.6	Distance somme	58
3.3.7	Discussion sur la distance globale	58
3.4	Conclusion	64
3.4.1	Perspectives	66
3.4.2	Applications	73
4	Grilles de structures répétitives	81
4.1	Définition des grilles de structures répétitives	81
4.2	Critères de pertinence	83
4.2.1	Définition des critères de pertinence	83
4.2.2	Pondération des critères de pertinence	87
4.2.3	Optimisation des calculs	89
4.3	Évaluation de l'énergie des grilles	90
4.3.1	Cartes de scores optimaux des décalages	92
4.3.2	Implémentation	92
4.3.3	Pertinence de la formalisation	92
4.4	Conclusion	95
5	Optimisation de l'agencement de grilles	99
5.1	Énergie d'un agencement de grilles	100
5.1.1	Énergie d'attache aux données	100
5.1.2	Énergie de régularisation	101
5.1.3	Complexité de l'énergie d'un agencement	104
5.2	Optimisation de l'agencement de grilles	104
5.2.1	Échantillonnage MCMC	105
5.2.2	Processus ponctuel marqué	106
5.2.3	Convergence vers l'agencement optimal	108
5.2.4	Distribution de référence	109
5.2.5	Implémentation des sauts réversibles	111
5.3	Résultats	117
5.4	Conclusion	120
6	Conclusion	121
6.1	Points forts de nos travaux	122
6.1.1	Rapidité des calculs	122
6.1.2	Localisation de structures cohérentes	122
6.2	Limites de nos travaux	122
6.2.1	Manque d'évaluation des résultats	123
6.2.2	Faiblesses des critères de localisation des structures	123

6.3	Perspectives	123
6.3.1	Perfectionnement général de la modélisation	123
6.3.2	Utilisation conjointe des deux approches	124
6.3.3	Utilisation d'informations d'autres natures	124
6.3.4	Application à l'indexation	124
6.3.5	Application à la compression	124
6.3.6	Application à la sémantisation	125
A	Calculs combinatoires	127
A.1	Démonstration du premier terme de H	127
A.2	Démonstration du deuxième terme de H	128
A.3	Démonstration du calcul de S	128
	Publications de l'auteur	130
	Bibliographie	130
	Table des figures	137