

UNIVERSITÉ DE NEUCHÂTEL
INSTITUT DE MICROTECHNIQUE

**Segmentation
et
traitements préliminaires
des
images de profondeur**

THÈSE

**PRÉSENTÉE À LA FACULTÉ DES SCIENCES
POUR OBTENIR LE GRADE DE DOCTEUR ÈS SCIENCES**

PAR

Gilbert Maître

1994

Segmentation
et
traitements préliminaires
des
images de profondeur

Gilbert Maître

imprimé par:

Service offset Brandt
Av. Léopold-Robert 65
2301 La Chaux-de-Fonds

A mon papa



René Maître 1928

A ma maman



Suzanne Maître, née Métrailler, 1930

A ma tante marraine



Odette Walther, née Maître, 1934

**«La thèse: un grand pas pour l'homme,
un petit pas pour l'humanité.»**

François Tièche

Résumé

Cette thèse s'inscrit dans le cadre d'une approche purement géométrique de la vision artificielle tridimensionnelle. Elle s'intéresse au traitement des images de profondeur, collections de positions tridimensionnelles mesurées à la surface d'objets. Plus précisément, son sujet est la segmentation des images de profondeur, c'est-à-dire le partage de l'ensemble des points mesurés en sous-ensembles de points conformément à un modèle géométrique de surface. La thèse défend une segmentation basée sur le modèle peu contraignant de continuité d'ordre 1 et propose une nouvelle méthode de détection des discontinuités d'ordre 0 et d'ordre 1 pour réaliser la segmentation. En plus, ce travail apporte une étude approfondie de certains traitements d'une image de profondeur liés à la méthode proposée.

La méthode de détection des discontinuités constitue la pièce maîtresse de ce travail. Elle se range parmi les méthodes qui procèdent à une détection locale des discontinuités. Trois éléments la caractérisent. Premièrement, elle utilise un unique opérateur pour détecter à la fois les discontinuités d'ordre 0 et d'ordre 1. Deuxièmement, elle combine la détection des discontinuités avec le lissage nécessaire à atténuer le bruit, auquel la détection des discontinuités est sensible. Troisièmement, elle fait dépendre le lissage et la détection des discontinuités d'une estimation du niveau du bruit, de sorte à ce que le signal ne soit pas lissé plus que nécessaire. La dénomination MDMS de la méthode, pour "minimal detector minimal smoothing", se veut être mnémonique de ces trois caractéristiques. Son originalité par rapport à d'autres méthodes

combinant lissage et détection des discontinuités repose sur deux points. D'une part, elle est la seule méthode à tenir compte d'une estimation du niveau de bruit pour contrôler le lissage et la détection des discontinuités. D'autre part, la détection des discontinuités d'ordre 0 et d'ordre 1 est combinée avec un lissage par diffusion, ce qui facilite une approche multirésolutionnelle.

En plus de la contribution pratique que représente la méthode MDMS, cette thèse fournit une analyse approfondie des problèmes suivants liés au traitement des images de profondeur: segmentation, détection des discontinuités, estimation et atténuation du bruit. Une partie importante est consacrée à décrire et à comparer les méthodes connues pour résoudre ces problèmes.

Zusammenfassung

Die vorliegende Dissertation plaziert sich im Umfeld der dreidimensionalen Computer Vision, wobei sich die Darstellung der Welt auf die rein geometrischen Aspekte beschränkt. Die Arbeit beschäftigt sich mit der Verarbeitung von Tiefenbildern, Mengen von 3D-Punkten, die an der Oberfläche von Gegenständen gemessen werden. Genauer ausgedrückt handelt es sich um die Segmentierung von Tiefenbildern, d.h. um die Aufteilung der Menge der gemessenen Punkte in Untermengen gemäss eines geometrischen Flächenmodells. Die Dissertation verteidigt eine Segmentierung basierend auf einem Kontinuitätsmodell erster Ordnung und schlägt eine neue Methode zur Detektion der Unstetigkeiten nullter und erster Ordnungen vor. Zusätzlich diskutiert diese Arbeit verschiedene andere Methoden zur Verarbeitung von Tiefenbildern im Zusammenhang mit der im Detail behandelten Methode.

Die Methode zur Detektion der Unstetigkeiten bildet das Kernstück dieser Arbeit. Sie kann unter diejenigen Methoden eingereiht werden, die sich auf eine lokale Detektion der Unstetigkeiten basieren. Diese Methode kann in folgenden drei Punkten charakterisiert werden: erstens in der Benutzung eines einzelnen Operators der gleichzeitig Unstetigkeiten nullter und erster Ordnung detektieren kann, zweitens in der Kombination der Detektion der Unstetigkeiten mit der zur Reduzierung des Rauschens unerlässlichen Glättung und drittens werden Glättung und Detektion der Unstetigkeiten in Abhängigkeit des geschätzten

Rauschens durchgeführt, so dass das Nutzsignal nicht übermässig geglättet wird.

Der Name MDMS der entwickelten Methode soll diese drei Charakteristika ausdrücken und steht für: "minimal detector minimal smoothing". Der spezielle Beitrag dieser Arbeit im Vergleich zu anderen Arbeiten, die Glättung und Detektion der Unstetigkeiten kombinieren, liegt in zwei Punkten. Einerseits ist sie die einzige Methode, bei der eine Schätzung des Rauschens zur Steuerung von Glättung und Detektion der Unstetigkeiten beigezogen wird. Andererseits wird die Detektion der Unstetigkeiten nullter und erster Ordnung mit einer Glättung durch Diffusion kombiniert, was ein Arbeiten mit variabler Auflösung ermöglicht.

Neben dem praktischen Beitrag in der Form der MDMS Methode liefert diese Arbeit eine vertiefte Analyse der folgenden Probleme im Zusammenhang mit Tiefenbildern: Segmentierung, Detektion der Unstetigkeiten, Schätzung und Abschwächung des Einflusses des Rauschens. Ein grosser Teil der Arbeit ist der Beschreibung und dem Vergleich der aus der Literatur bekannten Methoden zur Lösung dieser Probleme gewidmet.

Abstract

This thesis enters in the frame of a purely geometric approach of three-dimensional artificial vision. Its interest is in the processing of range images, which are collections of positions measured at object surfaces. More precisely, its subject is the segmentation of a range image, i.e. the sharing of the set of measured points into subsets according to a geometric model of surface. The thesis defends a segmentation approach based on the loosely constraining model of first order continuity and proposes a new detection method of 0th order and 1st order discontinuities to perform the segmentation. In addition, this work provides a deep study in several processing techniques which are linked to the proposed method.

The discontinuity detection method is the core contribution of this work. It is a local method characterized by the three following features. First, an unique detector is used for detecting both 0th order and 1st order discontinuities. Second, the noise, which disturbs the detection of discontinuities, is damped by combining smoothing with discontinuity detection. Third, discontinuity detection and smoothing are controlled according to an estimate of the noise level, so that the signal is not smoothed more than necessary. The method is called Minimal Detector Minimal Smoothing (MDMS). Compared to other methods which also combine discontinuity detection with smoothing, the MDMS method is original in regard of two aspects. First, it is the only one that takes into account an estimate of noise level to control discontinuity detection and smoothing. Second, it is

the only one that combines the detection of 0th order and 1st order discontinuities with the use of diffusion for smoothing.

The interest of this work goes beyond its practical contribution related to the MDMS method, since it makes a deep analysis of the following problems involved in the processing of range images: segmentation, discontinuity detection, noise estimation, and noise attenuation. An important part of this thesis is devoted to the description and the comparison of existing methods for solving these problems.

Table des matières

Chapitre 1

Introduction	1
1.1 Généralités	1
1.2 Contexte	1
1.2.1 Vision artificielle	1
1.2.2 Vision 3D géométrique	2
1.3 Sujet	3
1.3.1 Segmentation d'une image de profondeur	4
1.3.2 Choix de segmentation	5
1.3.3 Méthodes mises en oeuvre	5
1.4 Organisation du contenu	6

Chapitre 2

Image de profondeur, segmentation, discontinuités	9
2.1 Image de profondeur	10
2.1.1 Définitions et propriétés générales	10
2.1.2 Systèmes d'acquisition d'images de profondeur	12
2.1.3 Traitement d'image: définitions et notations	14
2.1.4 Types d'images de profondeur	17
2.1.5 Représentations graphiques d'une image de profondeur	19
2.2 Segmentation d'une image de profondeur	22
2.2.1 Définition de la segmentation	22
2.2.2 Comparaison des critères d'homogénéité	24
2.3 Discontinuités d'une image de profondeur	26
2.3.1 Définitions	26
2.3.2 Origine et propriétés des discontinuités	28
2.3.3 Importance d'une segmentation suivant C^1	29
2.4 Conclusion et discussion	31

Chapitre 3

Segmentation des images de profondeur: état de l'art	33
3.1 Méthodes élémentaires de segmentation	33
3.1.1 Introduction	33
3.1.2 Détection de frontière et étiquetage des composantes connexes (M1)	34
3.1.3 Seuillage et étiquetage des composantes connexes (M2)	35
3.1.4 "Clustering" et étiquetage des composantes connexes (M3)	36
3.1.5 Fusion de régions (M4a)	37
3.1.6 Division et fusion de régions (M4b)	38
3.1.7 Croissance indépendante de régions (M5)	39
3.1.8 Récapitulation	45
3.2 Revue des méthodes de segmentation	46
3.2.1 Type de courbure	46
3.2.2 Surface plane	48
3.2.3 Surface algébrique	49
3.2.4 Continuité d'ordre 1	52
3.2.5 Autres	54
3.3 Discussion et conclusion	55

Chapitre 4

Détection des discontinuités	59
4.1 Introduction	60
4.1.1 Détection des discontinuités	60
4.1.2 Détecteur de point de discontinuité	60
4.1.3 Détecteurs différentiels de point de discontinuité	62
4.2 Grandeurs différentielles utiles à la détection des points de discontinuité	65
4.2.1 Mesures de discontinuité	65
4.2.2 Estimation des grandeurs différentielles	67
4.3 Revue des détecteurs de point de discontinuité	74
4.3.1 Détecteurs sans condition d'extremum	74
4.3.2 Détecteurs avec condition d'extremum	75
4.3.3 Discussion	77
4.4 Détection conjointe des discontinuités à l'aide d'un unique détecteur	79
4.4.1 Mesures de discontinuité pour une IP cartésienne orthonormée	79
4.4.2 Extension aux autres types d'IP	84
4.4.3 Récapitulation	87
4.5 Conclusion	88

Chapitre 5

Estimation du bruit	89
5.1 Erreur de mesure et bruit	89
5.1.1 Erreur de mesure d'une image de profondeur	89

5.1.2	Bruit	91
5.2	Estimation du bruit.....	91
5.2.1	Introduction	91
5.2.2	Méthodes d'estimation du bruit	91
5.3	Estimation du bruit basée sur une mesure de la deuxième dérivée	93
5.3.1	Méthode élémentaire	93
5.3.2	Affinage de la méthode	95
5.3.3	Algorithme d'estimation robuste du bruit	101
5.3.4	Mise en application de l'estimation du bruit.....	103
5.3.5	Comparaison avec les autres méthodes d'estimation du bruit	109
5.3.6	Extension aux images de profondeur de type vectoriel	112
5.4	Conclusion	112

Chapitre 6

	Lissage adapté à la détection des discontinuités	115
6.1	Introduction.....	116
6.2	Méthodes de lissage linéaire à analogie physique: principe.....	116
6.2.1	Lissage par diffusion (D)	117
6.2.2	Lissage par minimisation d'énergie (E1, E2)	118
6.2.3	Comparaison	121
6.2.4	Application aux images de profondeur: problématique	123
6.2.5	Application aux images de profondeur: lissage optimal	125
6.2.6	Récapitulation et conclusion	130
6.3	Méthodes de lissage linéaire à analogie physique: mise en application	131
6.3.1	Formes discrètes du lissage	131
6.3.2	Mise en application du lissage	133
6.3.3	Méthodes itératives locales	134
6.3.4	Méthodes itératives de diffusion	135
6.3.5	Méthodes itératives de minimisation d'énergie	142
6.3.6	Récapitulation et conclusion	147
6.4	Méthodes de lissage préservant les discontinuités	148
6.4.1	Introduction	148
6.4.2	Diffusion non-linéaire.....	149
6.4.3	Minimisation d'énergie non-linéaire.....	152
6.4.4	Préservation conjointe des discontinuités \bar{C}^0 et \bar{C}^1	154
6.4.5	Discussion	155
6.5	Méthode MDMS	157
6.5.1	Principe	157
6.5.2	Algorithme MDMS	160
6.5.3	Extensions et modifications de la méthode	166
6.6	Conclusion	168

Chapitre 7

Résultats de segmentation par détection des discontinuités	169
7.1 Mise en place de l'expérimentation	169
7.1.1 Algorithme	169
7.1.2 Base de données	171
7.1.3 Plateforme	172
7.2 Exemples de segmentation	173
7.2.1 Choix des images de profondeur	173
7.2.2 Exécution de la segmentation et résultats	176
7.3 Analyse et évaluation des résultats	182
7.3.1 Images de profondeur à modèle de facettes	182
7.3.2 Autres	184
7.4 Discussion	184
7.5 Conclusion	185

Chapitre 8

Conclusions	187
8.1 Contributions	187
8.1.1 Apports théoriques	187
8.1.2 Apports pratiques	188
8.1.3 Perspectives	189
8.2 Conclusion	190

Remerciements	191
---------------	-----

Références	197
------------	-----

Annexes

- A. Convergence de la troncature itérative
- B. Stabilité de la méthode itérative de diffusion D^*
- C. Atténuation d'un signal aléatoire discret IID par lissage gaussien

Chapitre 1

Introduction

1.1 Généralités

Ce travail de thèse s'inscrit dans le cadre de la vision artificielle. Le développement important que connaît actuellement la vision artificielle est lié au progrès soutenu de l'automatisation et de la robotique. La demande d'une flexibilité accrue par rapport à un environnement changeant conduit à vouloir doter les machines de systèmes de perception de leur environnement physique. L'objectif convoité est de réaliser un système dont les performances soient comparables à celles de la vision humaine. Hélas, ou heureusement pour les ingénieurs qui veulent travailler dans ce domaine, le but est encore loin d'être atteint. Les systèmes existants de vision ne sont capables de fonctionner que dans des environnements spécifiques.

1.2 Contexte

1.2.1 Vision artificielle

L'objet de la vision artificielle peut être défini en ces termes: doter une machine de la capacité de déterminer quels objets se trouvent où dans son environnement ("what objects are where" [Marr82]). Le "what" pose la question de la reconnaissance des objets, le "where" celle de leur localisation.

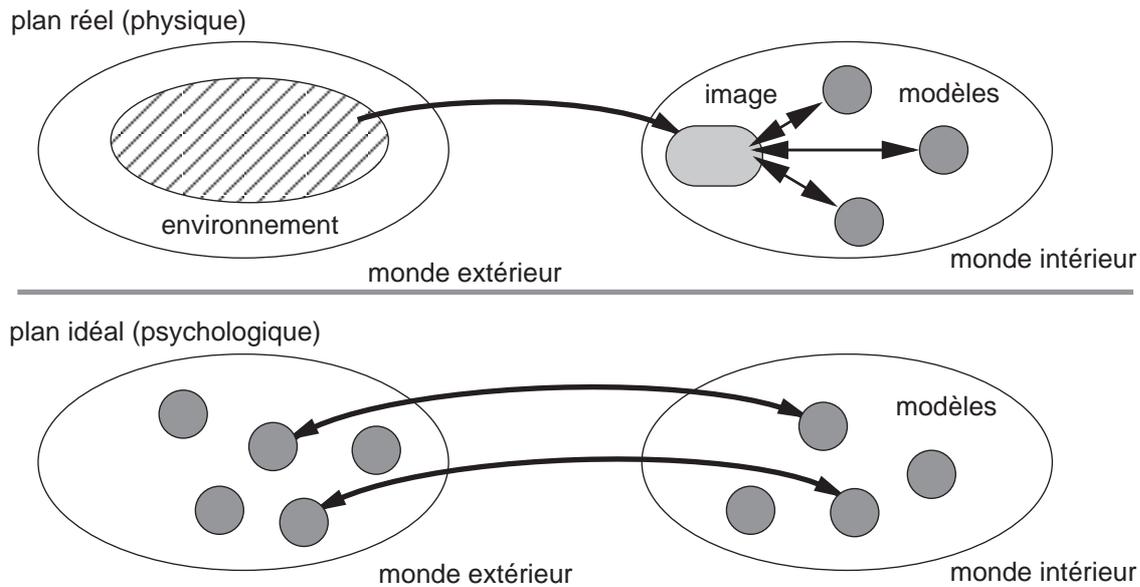


Figure 1.1: Phénomène de vision: relation entre monde extérieur et monde intérieur.

Postulats sur le phénomène de vision

Le phénomène de vision, c'est-à-dire la capacité pour un système naturel ou artificiel de reconnaître et localiser des objets, repose sur les principes suivants [figure 1.1]:

- Le système dispose d'une représentation des objets sous la forme de modèles. Le modèle est à la fois une représentation simplifiée de l'objet et son prototype: l'unité de l'objet provient de l'unité du modèle.
- Le système reçoit une représentation de la partie instantanément visible de son environnement sous la forme d'une image.
- La reconnaissance et la localisation des objets dans l'environnement a lieu par la reconnaissance et la localisation des modèles dans l'image.

1.2.2 Vision 3D géométrique

Elle constitue une approche particulière de la vision artificielle, basée d'une part sur une représentation du monde physique par la seule géométrie tridimensionnelle (3D) et, d'autre part, sur une mesure autant que possible complète de cette géométrie.

Motivation

La géométrie 3D constitue un mode de représentation établi du monde physique, comme en témoignent les nombreux exemples: cartes topographiques, plans de bâtiments, plans d'objets manufacturés.

En particulier, le modèle géométrique d'un objet rigide se caractérise par la séparabilité entre pose et forme. La pose (position et orientation) permet une localisation univoque. La forme est caractéristique de l'objet pour sa reconnaissance.

Cependant, une perception de l'environnement par images d'intensité ne peut obtenir qu'une mesure partielle de la géométrie 3D de l'environnement.

Acquisition d'images de profondeur

Pour la mesure géométrique, la vision 3D utilise des systèmes sensoriels capables de faire une mesure dense de positions à la surface des objets [Jarv83]. Dans certains cas, les positions sont organisées sous la forme d'un tableau appelé image de profondeur [Besl89].

Problème de la vision par images de profondeur

En considérant le problème de l'acquisition des images de profondeur comme résolu, la vision par images de profondeur pose les problèmes suivants [Aman90]. Il s'agit

- (1) de déterminer une forme de représentation appropriée pour les modèles d'objets
- (2) de trouver une méthode de mise en correspondance entre image et modèles de laquelle doit résulter la reconnaissance et la localisation des modèles dans l'image

1.3 Sujet

Cette thèse est l'aboutissement d'un travail de recherche sur la segmentation des images de profondeur.

1.3.1 Segmentation d'une image de profondeur

Celle-ci consiste en un partage de l'image de profondeur en régions de taille maximale conformément à un (des) modèle(s) de surface et peut

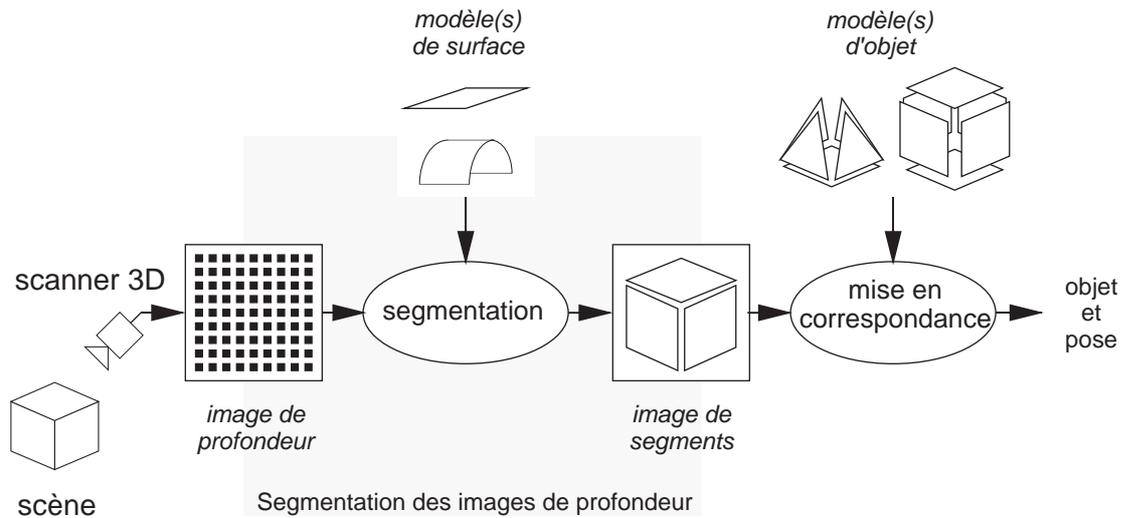


Figure 1.2: Place de la segmentation dans un système de vision géométrique 3D.

être interprétée comme la mise en correspondance de l'image avec ce(s) modèle(s) de surface.

Rôle de la segmentation

Dans un système de vision géométrique 3D, la segmentation joue le rôle d'une première étape de mise en correspondance entre modèles d'objets et image de profondeur [figure 1.2]. Mais pour qu'elle puisse jouer ce rôle, il faut que les modèles de surface soient adaptés aux modèles d'objet.

La part faite par la segmentation dans la mise en correspondance entre modèle d'objet et image de profondeur dépend du choix du modèle de surface. Par exemple, on peut utiliser comme modèle:

- une des surfaces d'un modèle d'objet ou
- un modèle de surface commun à différentes surfaces de différents objets.

Dans le premier cas, la segmentation réalise complètement la mise en correspondance, mais seulement pour cette surface de cet objet. Dans le deuxième cas, la segmentation met en correspondance plusieurs modèles d'objet avec l'image, mais la mise en correspondance entre modèle d'objet et image doit être achevée en déterminant quelle surface de quel objet correspond à quel segment.

Ainsi, la segmentation occupe une place différente dans la mise en correspondance suivant si le modèle est plutôt déterminé (particulier) et donc prescriptif (contraignant) ou s'il est plutôt vague (général) et donc descriptif.

1.3.2 Choix de segmentation

L'approche de segmentation poursuivie ici utilise un modèle de surface peu contraignant, soit celui d'une surface continue d'ordre 1. Ce modèle est adapté à un modèle d'objet à facettes.

Par rapport à un choix de segmentation utilisant un modèle plus prescriptif, notre approche se caractérise par une faible complexité de calcul, au prix, il est vrai, d'une robustesse moindre, puisqu'on n'utilise que peu ou pas du tout de connaissance a priori.

L'approche est donc à considérer comme une opération préliminaire permettant d'établir des hypothèses qu'il s'agit ensuite de vérifier.

1.3.3 Méthodes mises en oeuvre

La segmentation suivant le modèle de surface continue d'ordre 1 passe par une détection des discontinuités et utilise une estimation du bruit de l'image de profondeur.

Détection des discontinuités

La détection des discontinuités, surtout celle d'ordre 1, est connue pour être sensible au bruit et une atténuation du bruit doit être effectuée. Cependant, le lissage servant à atténuer le bruit atténue aussi les variations du signal correspondant aux discontinuités. Par conséquent, les deux opérations doivent être soigneusement combinées.

La méthode proposée repose sur l'argumentation suivante:

- Il y a un compromis à trouver pour le lissage. D'autre part, ce compromis dépend de la valeur locale du rapport signal sur bruit.
- Le seuil de détection des discontinuités doit être adapté au niveau du bruit résiduel (c'est-à-dire après atténuation)

La méthode, qui est itérative, effectue un lissage progressif et utilise une estimation du bruit résiduel en fonction du lissage.

A chaque itération le niveau du signal est mesuré et comparé avec le niveau estimé du bruit résiduel. Lorsque le niveau mesuré du signal dépasse sensiblement le niveau estimé du bruit, le lissage est considéré comme (localement) suffisant et est interrompu.

D'autre part, à chaque itération, le seuil de détection des discontinuités est abaissé suivant le niveau estimé du bruit. Les discontinuités sont ainsi détectées au fur et à mesure que le signal dépasse le niveau estimé du bruit.

L'estimation du bruit résiduel repose sur une estimation du bruit initial et sur la connaissance de l'effet du lissage sur un bruit de référence (identiquement indépendamment distribué).

Estimation du bruit

La méthode pour l'estimation du bruit initial se base sur une mesure de la deuxième dérivée du signal et est caractérisée par l'utilisation d'une méthode robuste pour en éliminer les valeurs aberrantes.

1.4 Organisation du contenu

Le contenu de la thèse est organisé de la manière suivante:

D'abord, les notions d'image de profondeur, de segmentation et de discontinuités d'une image de profondeur sont précisées [chapitre 2]. Vient ensuite un état de l'art des méthodes de segmentation [chapitre 3]. Puis, le chapitre 4 aborde le thème de la détection des discontinuités et montre qu'elle nécessite un traitement du bruit. Le chapitre 5 explique l'origine du bruit et traite de son estimation. Notre méthode d'estimation du bruit y est décrite. Le chapitre 6 est consacré au problème de l'atténuation du bruit et de la combinaison possible de cette atténuation avec la détection des discontinuités. La méthode originale de détection des discontinuités que nous proposons y est décrite. Finalement, nous présentons les résultats de segmentation obtenus en utilisant les méthodes d'estimation du bruit et de détection des discontinuités développées [chapitre 7].

Pour la personne qui ne s'intéresse qu'aux contributions pratiques de cette thèse, la lecture des passages suivants devrait lui fournir les renseignements utiles: le chapitre 2 pour les notions de base, le chapitre 5 pour la méthode d'estimation du bruit, le paragraphe 6.5

pour la méthode de détection des discontinuités, finalement le chapitre 7 pour les résultats de segmentation.

Chapitre 2

Image de profondeur, segmentation, discontinuités

Dans ce chapitre, nous introduisons tour à tour les trois notions de base: image de profondeur, segmentation d'une image de profondeur et discontinuités d'une image de profondeur.

D'abord, une image de profondeur est présentée comme une collection de positions 3D mesurées à la surface d'objet(s) et organisée en tableau. Deux systèmes d'acquisition d'images de profondeur sont décrits brièvement. Les images de profondeur sont classées en différents types. Des exemples des diverses représentations graphiques des images de profondeur en illustrent l'essence.

La segmentation d'une image de profondeur fait l'objet d'une définition formelle. L'introduction du concept d'homogénéité permet de rattacher la segmentation d'une image de profondeur à la segmentation au sens large. Les différents critères possibles d'homogénéité de surface sont évalués suivant plusieurs caractéristiques utiles en vue d'un choix.

Les discontinuités de profondeur et d'orientation d'une image de profondeur sont définies en relation avec un modèle de surface continue qui lui est associé. Le comportement d'une segmentation suivant un critère de continuité d'ordre 1 est étudié dans le cas d'une image de profondeur idéale. Ce cas idéal fixe les perspectives pour le cas réel.

2.1 Image de profondeur

2.1.1 Définitions et propriétés générales

Une image de profondeur constitue un ensemble particulier de données géométriques issues d'une mesure du monde physique. Nous précisons ici en quoi ces données géométriques sont particulières.

Définition

Une image de profondeur (IP) est un tableau bidimensionnel de positions 3D vérifiant la propriété de cohérence spatiale. Le terme anglais de *rangel* (range image element) désigne un élément du tableau.

Notation

Soit \mathfrak{R}^3 l'espace géométrique 3D, une IP est notée

$$\mathbf{P} = [\mathbf{p}_{i,j}], \quad \mathbf{p}_{i,j} \in (\mathfrak{R}^3 \cup \{NIL\}) \quad (2.1)$$

où la valeur *NIL* est utilisée pour désigner l'absence de valeur de position 3D.

Propriété de cohérence spatiale

La propriété de cohérence spatiale englobe plusieurs aspects que nous précisons maintenant. Tout d'abord, il faut faire la distinction entre propriété extrinsèque et intrinsèque. La cohérence spatiale est en effet en premier lieu une propriété extrinsèque à l'IP, liée aux conditions de son acquisition. Elle est seulement en deuxième lieu une propriété intrinsèque aux données que constitue l'IP, conséquence des conditions d'acquisition.

Aspect extrinsèque

Le système de mesure fixe un faisceau de droites dans \mathfrak{R}^3 , une droite par position du tableau. Les rangels sont interprétés comme l'intersection de ces droites avec une surface continue qui est l'objet de la mesure.

La propriété de cohérence spatiale pose les conditions suivantes sur la géométrie d'acquisition de l'IP [figure 2.1]:

- (e1) Le faisceau de droites fixé par le système de mesure est quasi-parallèle. De plus, des droites du faisceau qui sont voisines dans

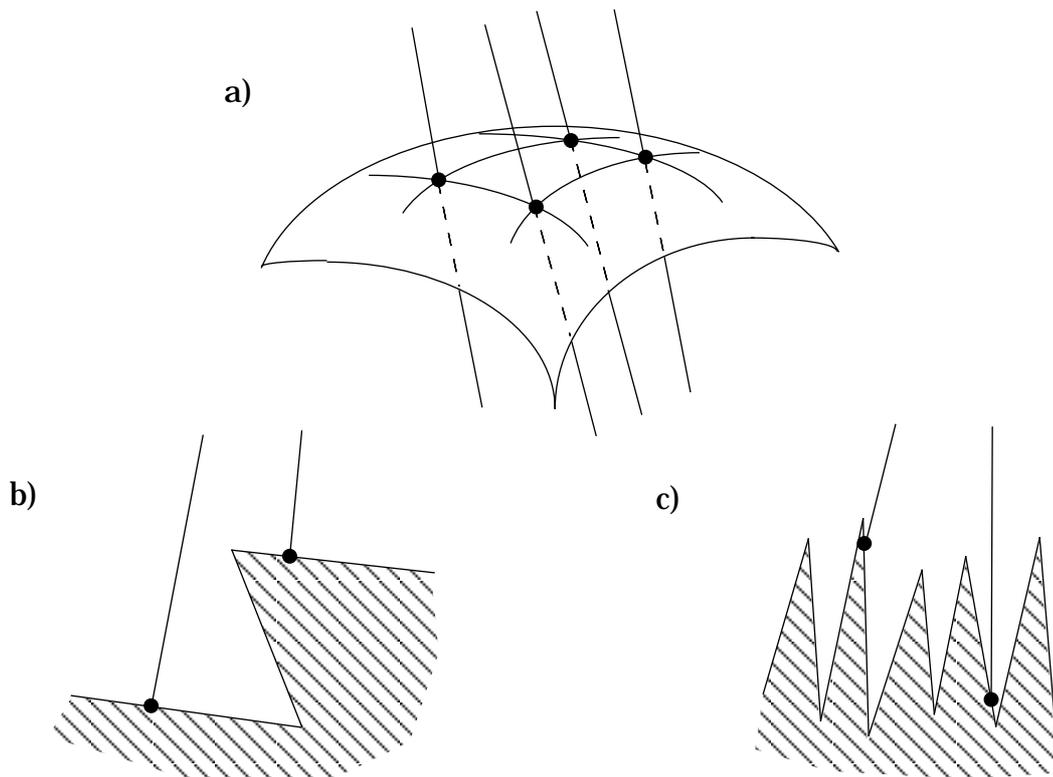


Figure 2.1: Cohérence spatiale. (a) localement vrai, (b) et (c) localement faux.

l'espace 3D correspondent à des positions voisines dans le tableau et vice-versa.

(e2) La majeure partie de la surface mesurée varie doucement par rapport à la densité du faisceau de droites.

Aspect intrinsèque

Au strict niveau des données de l'IP, la propriété de cohérence spatiale signifie:

(i1) Le rangel qui est le plus proche d'un autre dans l'espace 3D (en distance euclidienne) est un de ses voisins immédiats dans le tableau.

(i2) Généralement, le rangel voisin d'un autre dans l'image de profondeur est parmi les 9 plus proches dans l'espace 3D.

La propriété (i1) est une conséquence de (e1). (i2) est une conséquence de (e1) et de (e2).

Importance de la cohérence spatiale

Sur cette propriété reposent plusieurs procédés appliqués aux IP:

- (1) Interpolation: On construit une surface représentative de la surface mesurée à partir de l'IP simplement en "remplissant" les mailles du treillis dont les noeuds sont les rangels. Ce procédé n'est pas possible pour une liste quelconque de points, dans ce cas il faut trouver les points les plus proches entre eux dans l'espace 3D pour construire un maillage. Ce problème peut être résolu, par exemple, par la triangulation de Delaunay [Bois84].
- (2) Représentation par une surface d'approximation: Un ensemble de rangels voisins sont représentés par une surface algébrique d'approximation.
- (3) Estimation de l'erreur de mesure: La partie du signal non compatible avec la cohérence spatiale est assimilée à l'erreur de mesure [chap.5].

Remarque

S'ils sont appliqués de manière systématique, les procédés (1) et (2) négligent les cas d'exception que comporte la propriété (e2).

2.1.2 Systèmes d'acquisition d'images de profondeur

Il existe une grande variété de systèmes d'acquisition d'IP. Ceux-ci se distinguent par le procédé de mesure, la géométrie du dispositif, la performance en précision et en rapidité, etc. Jarvis [Jarv83] et, plus récemment, Besl [Besl89] ont présenté une revue des méthodes d'acquisition d'IP et des systèmes réalisés. Ici, seulement deux exemples sont présentés à fin d'illustration.

Le premier est un système expérimental de laboratoire, réalisé avec du matériel non spécialisé [HüMa89, MaHü88]. Il applique le principe de triangulation et fonctionne de la manière suivante [figure 2.2].

Un faisceau laser en forme de plan illumine la scène. Le faisceau dessine une trace visible sur la surface des objets. La trace est repérée dans l'image d'une caméra et amincie à une largeur d'un pixel. Pour chacun des pixels de la trace dans l'image, le point 3D, défini par l'intersection du plan de lumière et de la ligne de vue correspondant au pixel, est calculé. Ce point est retenu comme

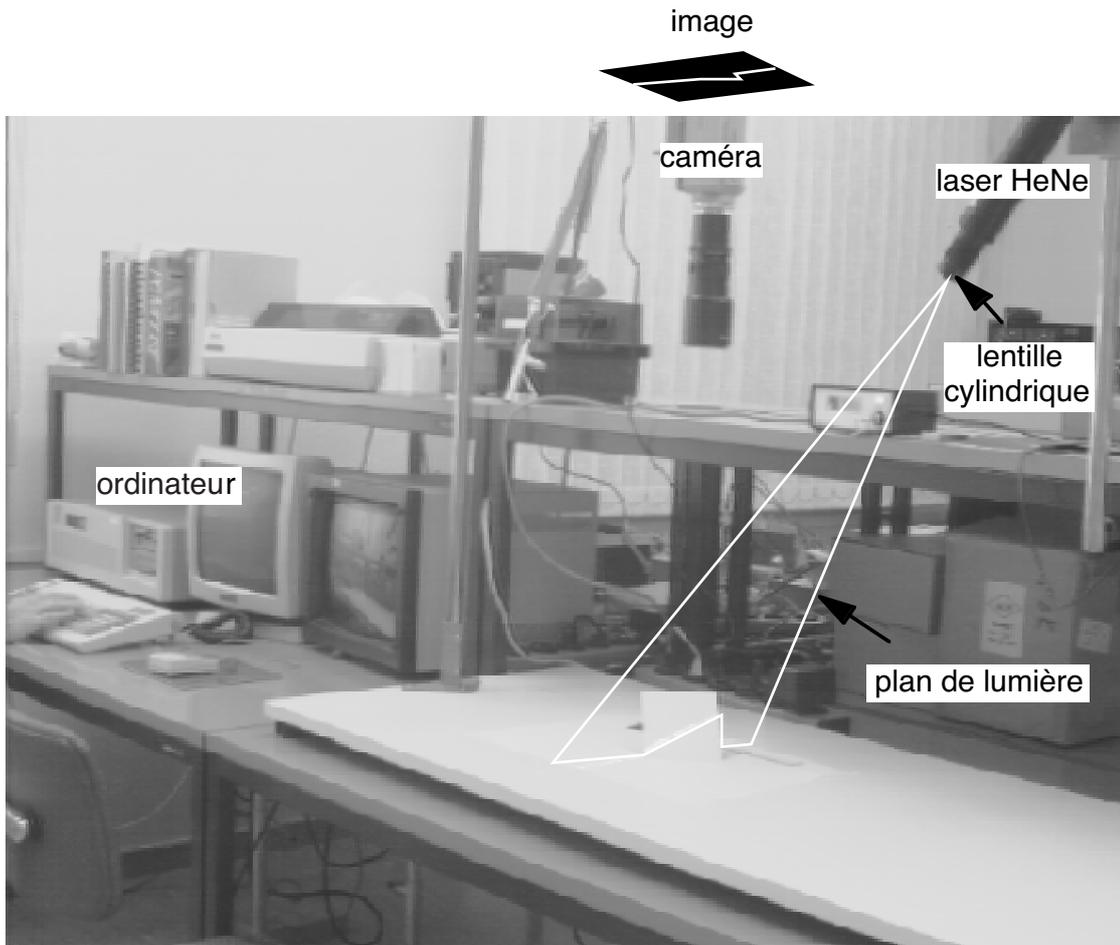


Figure 2.2: Système expérimental d'acquisition d'IP. Spécifications dans la configuration utilisée [MaHü88]. Domaine de mesure: 10x10 cm. Résolution: 0.2 mm. Précision: ± 0.4 mm. Vitesse maximale d'acquisition: 500 rangels/s (pour un PC 286). Prix: env. 5 kFS.

rangel. En traitant de la sorte une image de la caméra, on mesure un profil de la surface des objets correspondant à une section plane et on génère une ligne de l'IP. L'IP est construite profil par profil en translatant l'objet entre deux mesures de profil.

Le deuxième exemple est un système commercialisé, fonctionnant sur le principe de temps de vol [figure 2.3]. Un faisceau laser modulé en amplitude est projeté dans une direction, un capteur photosensible mesure la lumière revenant à la source. Grâce à la modulation en amplitude le décalage temporel entre signal lumineux émis et signal lumineux reçu peut être déterminé (à condition qu'il ne

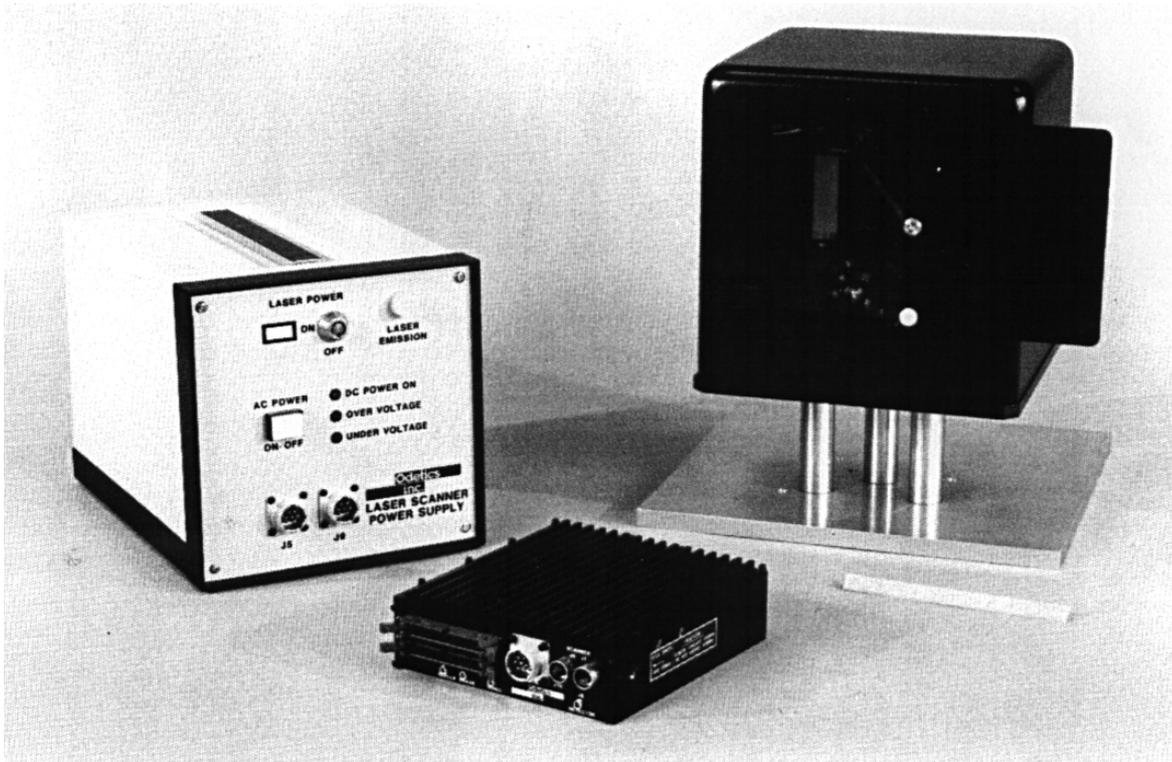


Figure 2.3: Système commercialisé d'acquisition d'IP. 3D Laser Imaging System, modèle GMS, Odetics 1989. Champs de vue: 60x60 degrés. Dimension d'image: 128x128 rangels. Temps d'acquisition: 0.835 s/image. Résolution maximale en profondeur: 18 mm. Intervalle d'ambiguïté: 9.37 m. Distance minimale: 45 cm. Puissance laser: 25 mW. Prix: env. 150 kFS.

dépasse pas la valeur d'une période de modulation). Le décalage temporel est une mesure directe de la distance entre système d'acquisition et surface rencontrée par le faisceau laser. Une mesure de distance dans une direction définit un point de l'IP. L'IP est construite par balayage du faisceau laser.

2.1.3 Traitement d'image: définitions et notations

Dans ce paragraphe, nous introduisons quelques définitions et notations empruntées au domaine du traitement d'images

- Une image est un tableau bidimensionnel de vecteurs à N composantes. i étant l'indice de la ligne et j l'indice de la colonne, deux notations sont utilisées pour désigner le tableau, respectivement un élément du tableau

$$\mathbf{V} = [\mathbf{v}_{i,j}] \quad (2.2.a)$$

$$\mathbf{V}(i, j) = \mathbf{v}_{i,j} \quad (2.2.b)$$

- Une position est un couple d'indices (i, j) du tableau.
- Un élément de l'image est défini par sa position et sa valeur. Pour être exact, il faudrait représenter l'élément par le triplet $(i, j, \mathbf{v}_{i,j})$ composé des indices i et j ainsi que du vecteur $\mathbf{v}_{i,j}$. Cependant, par soucis de simplification et aussi puisque c'est l'usage, l'élément est représenté uniquement par le symbole $\mathbf{v}_{i,j}$, ce qui rend celui-ci ambivalent. V désigne l'ensemble de tous les pixels de l'image \mathbf{V} . A noter que V représente une structure mathématique différente du tableau \mathbf{V} , permettant l'écriture de la relation

$$\mathbf{v}_{i,j} \in V \quad (2.3)$$

- Une région est un ensemble d'éléments de l'image.
- Un voisinage de la position (i, j) est un ensemble, noté $W_{i,j}$, de positions voisines de (i, j) . Un voisinage invariant en translation, noté $W(i, j)$, vérifie la relation

$$((k, l) \in W(i, j)) \Leftrightarrow ((k - i + i', l - j + j') \in W(i', j')) \quad (2.4)$$

- (u, v) sont les coordonnées d'image: les positions (i, j) correspondent aux valeurs entières de u et de v

Opérations de voisinage linéaires récursive et non-récursive

Soit une image \mathbf{V} et soit le tableau de coefficients

$$\mathbf{C} = \begin{bmatrix} c_{-M_1, -N_1} & \cdots & c_{-M_1, 0} & \cdots & c_{-M_1, N_2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{0, -N_1} & \cdots & \boxed{c_{0, 0}} & \cdots & c_{0, N_2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{M_2, -N_1} & \cdots & c_{M_2, 0} & \cdots & c_{M_2, N_2} \end{bmatrix} \quad (2.5)$$

où le rectangle autour de l'élément $c_{0,0}$ sert à désigner le coefficient d'indices $(0,0)$ dans le cas où les coefficients sont représentés par leur valeur numérique. Soit l'opération linéaire appliquée sur le voisinage de la position (i, j)

$$\mathbf{q} = \sum_{\substack{-M_1 \leq i' \leq M_2 \\ -N_1 \leq j' \leq N_2}} c_{i', j'} \cdot \mathbf{v}_{i+i', j+j'} \quad (2.6)$$

On considère deux manières de mettre dans une image les résultats de l'opération (2.6) appliquée à chaque position (i, j) : utiliser la même image

$$\mathbf{v}_{i,j} = \mathbf{q} \quad (2.7)$$

ou utiliser une nouvelle image $\tilde{\mathbf{V}}$

$$\tilde{\mathbf{v}}_{i,j} = \mathbf{q} \quad (2.8)$$

Dans le premier cas, le résultat de l'opération (2.6) à une position (i, j) dépend du résultat de l'opération à des positions voisines et donc de l'ordre dans lequel les opérations sont exécutées. L'opération de voisinage linéaire est dite réursive. Pour que l'opération soit entièrement définie, il faut fixer l'ordre dans lequel elle est exécutée.

Dans le deuxième cas, les résultats de l'opération (2.6) à des positions différentes sont indépendants les uns des autres. L'opération de voisinage linéaire est dite non-réursive. Il s'agit de plus d'une opération de voisinage invariante en translation, elle correspond à la corrélation croisée entre les tableaux \mathbf{C} et \mathbf{V} .

Définitions et notations

- Le tableau \mathbf{C} est appelé masque.
- L'opération de voisinage non-réursive définie par les équations (2.6) et (2.8), appelée aussi corrélation, s'écrit

$$\tilde{\mathbf{V}} = \mathbf{C} \otimes \mathbf{V} \quad (2.9)$$

- L'ordre naturel pour une opération de voisinage réursive est l'ordre naturel d'écriture dans un tableau: d'abord la première ligne en commençant par la première colonne, puis la deuxième ligne en commençant par la première colonne, etc.
- L'opération de voisinage réursive définie par les opérations (2.6) et (2.10)

$$\mathbf{v}_{i,j} = \mathbf{q}, \tilde{\mathbf{v}}_{i,j} = \mathbf{v}_{i,j} \quad (2.10)$$

exécutées dans l'ordre naturel est appelée, par extension, corrélation réursive. Notation:

$$\tilde{\mathbf{V}} = \mathbf{C} \otimes \mathbf{V} \quad (2.11)$$

Remarque:

La corrélation avec \mathbf{C} est équivalente à la convolution avec le masque symétrique de \mathbf{C} par rapport à $(0,0)$.

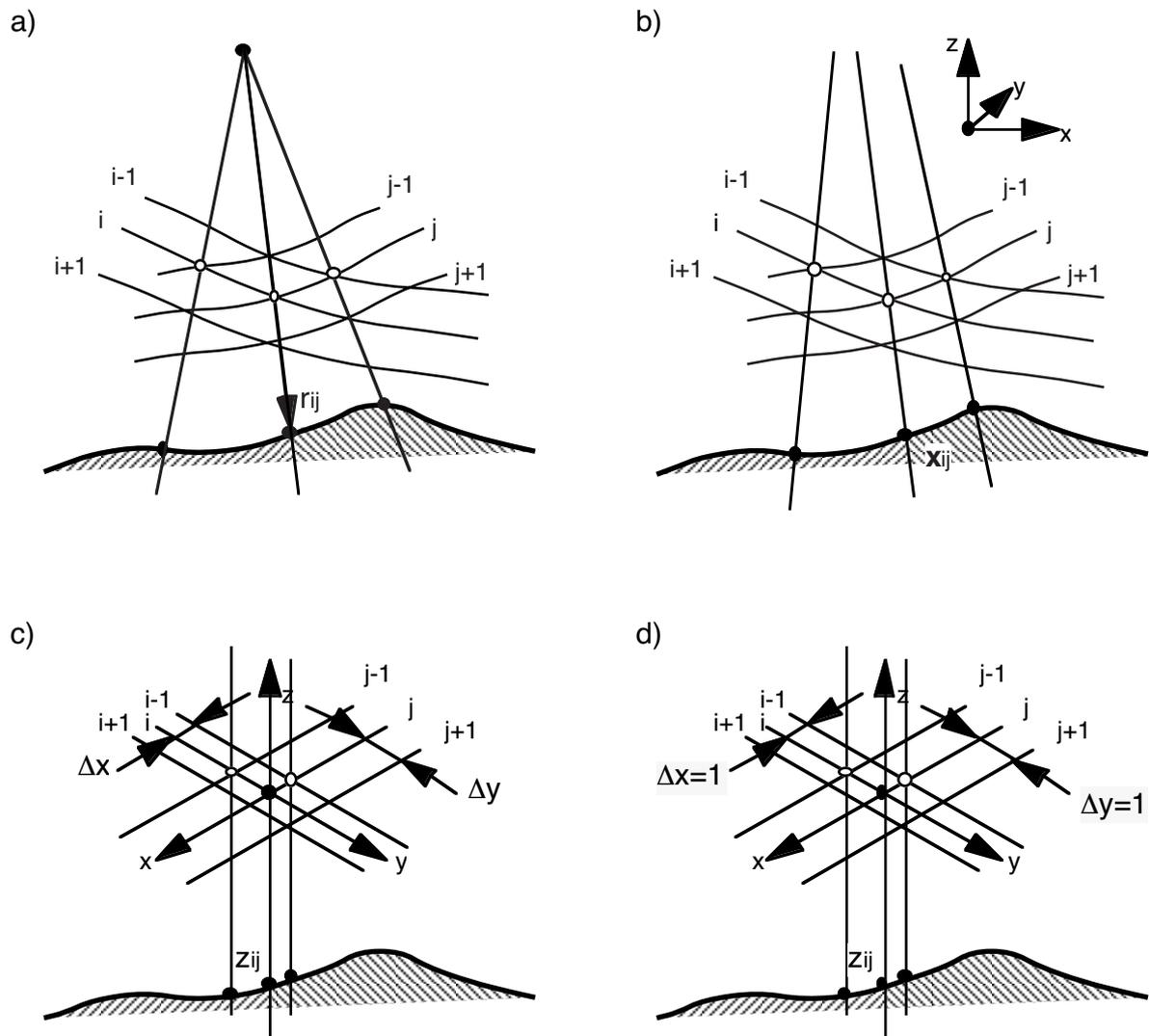


Figure 2.4: Types d'IP. (a) IP brute. (b) IP cartésienne. (c) IP cartésienne orthographique. (d) IP orthonormée.

2.1.4 Types d'images de profondeur

Parmi les IP couramment utilisées dans les milieux de la recherche en vision artificielle, on retient les quatre types suivants [figure 2.4]:

1. L'IP brute $[r_{i,j}]$

Les droites correspondant aux positions du tableau sont organisées en un faisceau régulier convergeant en un point. Les rangels sont des valeurs de distance prise de ce point de référence. En connaissant la position du point de référence et la direction des droites, on pourrait reconstruire la position 3D des points mesurés. Cependant origine et direction du faisceau ne sont pas explicites.

C'est à ce type d'IP que le terme général est le plus approprié, puisqu'il s'agit littéralement d'un tableau de profondeurs.

2. L'IP cartésienne $[x_{i,j}]$

Un point de mesure est représenté par un vecteur \mathbf{x} de coordonnées cartésiennes. Le fait que les éléments du tableau soient des vecteurs de dimension 3 et non des scalaires constitue la particularité de ce type d'image.

3. L'IP cartésienne orthographique $([z_{i,j}], \Delta x, \Delta y)$

Les positions 3D sont des coordonnées cartésiennes. Le tableau ne contient cependant que la coordonnée en z. Les coordonnées en x et en y sont proportionnelles aux indices du tableau i et j, respectivement:

$$x_{i,j} = \Delta x \cdot i \quad y_{i,j} = \Delta y \cdot j \quad (2.12.a-b)$$

Les facteurs de proportionnalité Δx et Δy sont les intervalles d'échantillonnage. Ils doivent être associés au tableau $[z_{i,j}]$ pour que l'information de position soit complète, d'où la notation utilisée.

4. L'IP cartésienne orthonormée $[z_{i,j}]$

Les IP de ce type forment un sous-ensemble des IP cartésiennes orthographiques, leur particularité étant que les intervalles d'échantillonnage en x et en y sont égaux et valent 1[†]:

$$\Delta x = \Delta y = 1 \quad (2.13.a-b)$$

On remarque que les IP peuvent être réunies en deux groupes [table 2.1]:

- les IP vectorielles, notées IP 3D. Ce sont les IP cartésiennes.
- les IP scalaires, notées IP 2.5D. Elles englobent tous les autres types d'IP.

Notations particulières:

- $[p_{i,j}]$ est utilisé pour souligner qu'il s'agit d'un tableau de scalaires
- P désigne l'ensemble de tous les rangels ayant une valeur différente de NIL

[†] le nombre est exempt d'unité, l'unité physique est associée au repère cartésien

Type d'IP	Dimension des rangels	Transformation des données de l'image en coordonnées cartésiennes 3D
IP brute $[r_{i,j}]$	1 (scalaire)	nécessite la connaissance de la géométrie du système d'acquisition
IP cartésienne $[\mathbf{x}_{i,j}]$	3 (vecteur)	$x_{i,j} = \mathbf{x}_{i,j,1}$ $y_{i,j} = \mathbf{x}_{i,j,2}$ $z_{i,j} = \mathbf{x}_{i,j,3}$
IP cartésienne orthographique $([z_{i,j}], \Delta x, \Delta y)$	1 (scalaire)	$x_{i,j} = \Delta x i$ $y_{i,j} = \Delta y j$ $z_{i,j} = z_{i,j}$
IP cartésienne orthonormée $[z_{i,j}]$	1 (scalaire)	$x_{i,j} = i$ $y_{i,j} = j$ $z_{i,j} = z_{i,j}$

Table 2.1: Types d'IP et interprétation des données.

2.1.5 Représentations graphiques d'une image de profondeur

Il existe de nombreuses manières de représenter graphiquement une IP. Voici la liste des principales formes de représentation, accompagnée d'exemples d'illustration. L'IP représentée est de type cartésien orthographique¹.

représentations en perspective (IP 2.5D & IP 3D)

- points 3D: projection des points [figure 2.5.a]
- fil de fer: projection du maillage [figure 2.5.b]
- profils: projection des profils [figure 2.5.c]
- surface neutre: projection du maillage visible [figure 2.5.d]
- rendu réaliste: projection des facettes visibles avec couleur et ombrage [figure 2.6]

¹ l'IP, intitulée MASK01, provient du Conseil National de Recherches du Canada (CNRC)

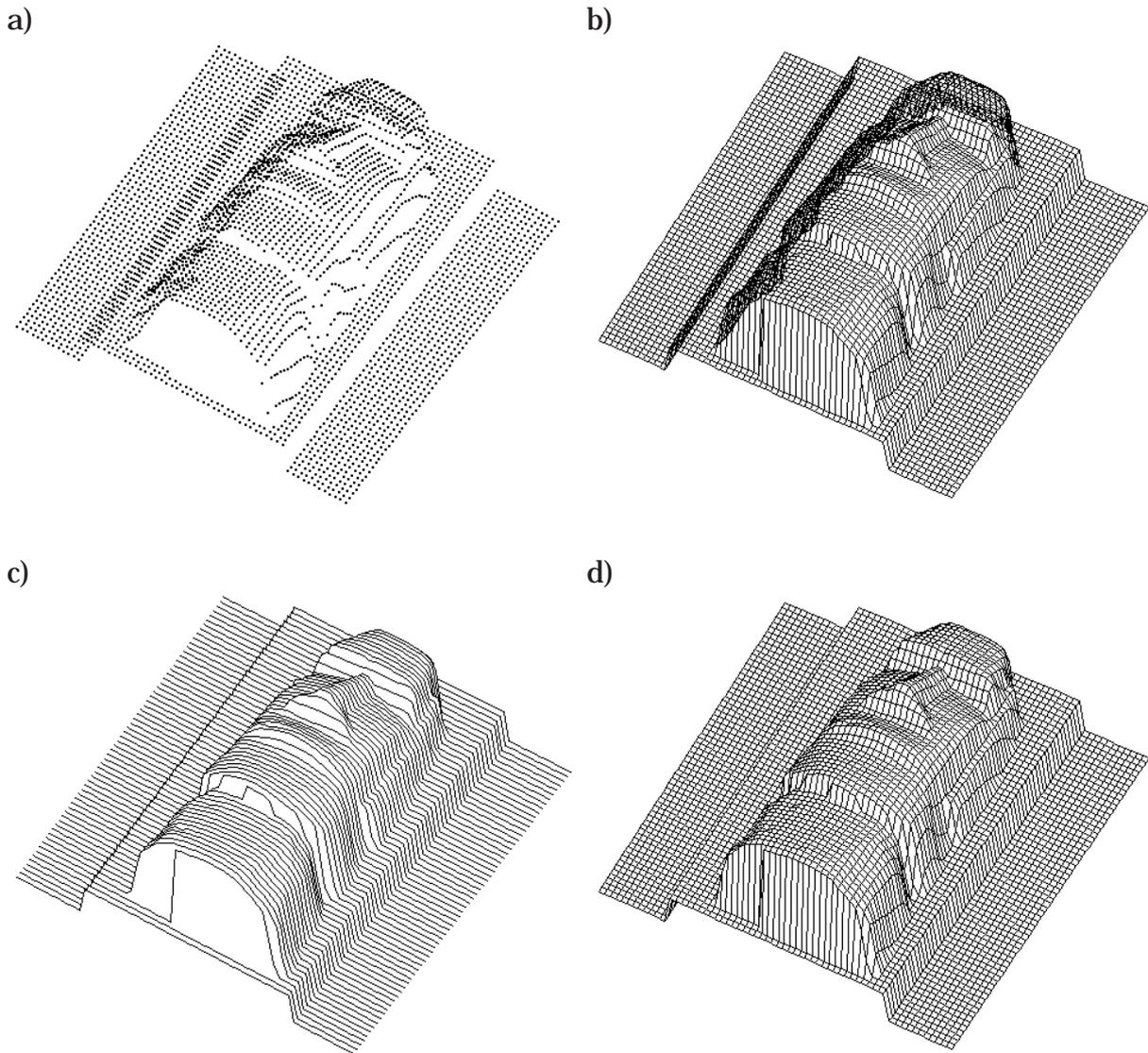


Figure 2.5: Représentations en perspective d'une IP. (a) points 3D (b) fil de fer (c) profils (d) surface neutre.

représentations sous forme d'image (IP 2.5D)

- image des courbes de niveaux [figure 2.7]
- image des profondeurs: niveaux de gris suivant une fonction monotone des profondeurs [figure 2.8.a]
- image pseudo-ombrée: niveaux de gris suivant une fonction monotone de la dérivée dans une direction [figure 2.8.b]



Figure 2.6: Représentation d'une IP par rendu réaliste.

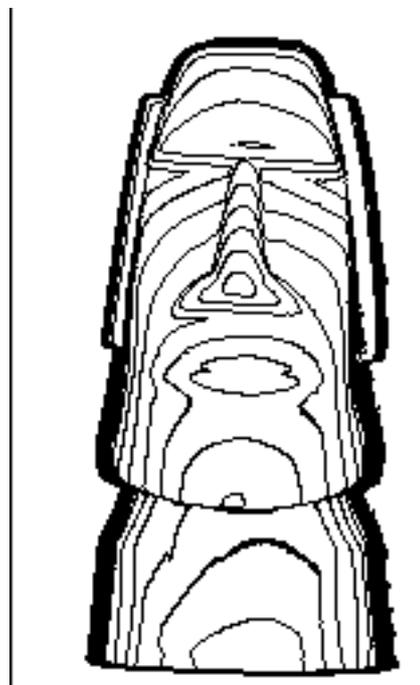


Figure 2.7: Représentation d'une IP par courbes de niveaux.

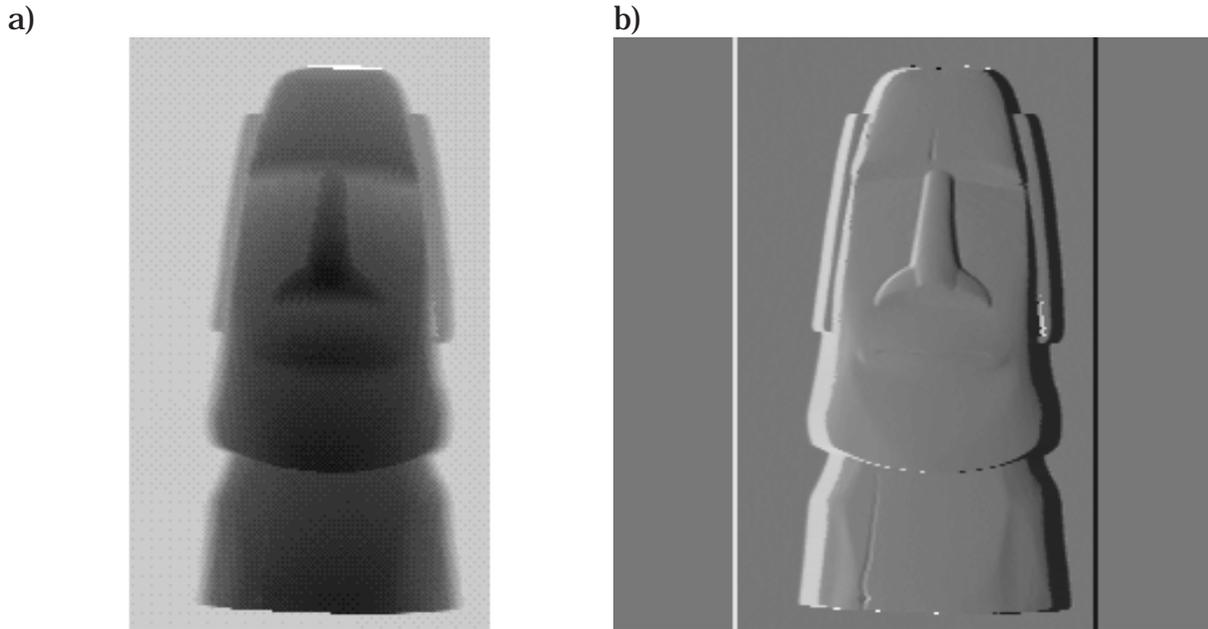


Figure 2.8: Représentation d'une IP par des images d'intensité (a) Image de profondeur. Le blanc code la valeur NIL (cf. pixels au sommet de la tête) (b) Image pseudo-ombrée par corrélation avec le masque $[-1 \ 0 \ 1]$, méthode proposée par [RiCo88].

2.2 Segmentation d'une image de profondeur

2.2.1 Définition de la segmentation

Nous donnons ici une définition formelle et générale de la segmentation d'une image de profondeur (IP). Nous précisons ensuite des concepts liés à cette définition générale.

Définition

La segmentation d'une IP $[\mathbf{p}_{i,j}]$ est un partage de l'ensemble P des rangels de valeur non *NIL* en régions (sous-ensembles) S_m homogènes et de taille maximale. Un sous-ensemble S_m est appelé segment.

Partition

La segmentation est une partition si le partage est en sous-ensembles deux à deux disjoints dont la réunion égale l'ensemble P [figure 2.9.c-d]. Dans certaines approches, la propriété de partition est exigée, dans d'autres pas.

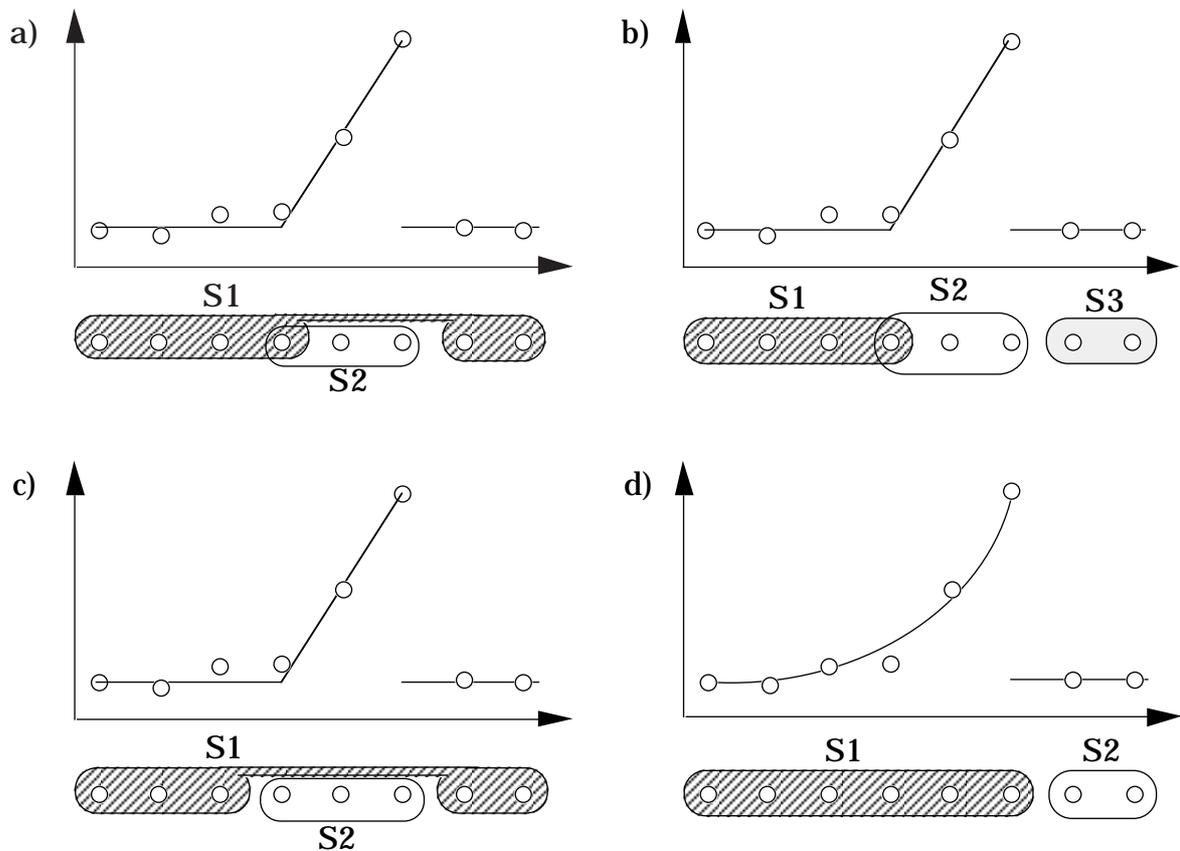


Figure 2.9: Segmentation d'une IP: illustration des concepts de base. (a) partage (b) partage en régions connexes (c) partition (d) partition en régions connexes

Régions connexes

Dans la plupart des cas, on exige que les régions soient connexes [figure 2.9.b-d] (pour une définition de la connexité, voir [BaBr82, p. 150]).

Homogénéité

Le critère d'homogénéité attribue au segment une propriété de surface. Par exemple une propriété différentielle (qualitative ou quantitative) ou une équation algébrique décrivant la surface. Cette propriété définit un modèle de surface.

Taille maximale

Il existe deux manières de définir le critère de taille maximale:

critère d'homogénéité	local/ global	intrin- sèque	contrainte de modèle			
			variété de forme des segments	fixation d'un segment		
				forme	orient.	pos.
continuité	local	oui	∞	non	0	0
type de courbure	local	oui	∞	non	0	0
orientation (normale)	local	non	1	oui	2 (max.)	0
courbures principales	local	oui	3 ^{a)}	oui	0	0
équation de plan	global	non	1	oui	2 (max.)	1 (max.)
équation surface 2.5D ^{b)}	global	non	–	oui	–	–
équation surface 3D	global	non	dépend équ. ^{c)}	oui	max.	max.

Table 2.2: Critères d'homogénéité et contraintes de forme. (a) plan, cylindre et sphère. (b) seulement pour les IP scalaires (c) voir table 2.3.

- (1) Un segment est de taille maximale si la réunion avec un quelconque autre segment forme une région non homogène (ce critère est en général lié à la condition de partition).
- (2) Un segment est de taille maximale si l'on ne peut pas lui rajouter de rangél sans le rendre inhomogène.

2.2.2 Comparaison des critères d'homogénéité

Les propriétés de surface qui se prêtent à être utilisées comme critère d'homogénéité sont relativement peu nombreuses. La liste de la table 2.2 en mentionne les principales. Nous allons les comparer sur la base de trois caractéristiques significatives qui sont à considérer pour faire un choix de segmentation: local-global, caractère intrinsèque et contrainte de modèle.

Local-global

Les propriétés différentielles sont des propriétés de surface locales. Les équations algébriques sont des propriétés de surface globales.

Caractère intrinsèque

Intrinsèque signifie que la propriété est indépendante du système de coordonnées dans lequel la surface est représentée. A ce sujet plusieurs remarques peuvent être faites:

- Les équations de surface 3D peuvent être transformées pour décrire la même surface dans deux repères cartésiens différents. Pour les équations de surface 2.5D cela n'est généralement pas possible.
- La plupart des critères d'homogénéité perdent leur caractère intrinsèque avec les IP brutes, car le système de coordonnées qui leur est associé n'est pas clairement défini.
- La continuité est le seul critère d'homogénéité intrinsèque valable pour tous les types d'IP.

Contrainte de modèle

Elle englobe deux aspects:

- la variété de formes que permet le critère d'homogénéité, c'est-à-dire le nombre de formes de surface différentes que peuvent avoir les divers segments d'une IP
- l'étendue avec laquelle le critère d'homogénéité détermine la surface, c'est-à-dire combien chaque segment d'une IP est déterminé dans sa forme, taille et pose.

Voici un exemple illustrant ce que sont ces caractéristiques. Nous comparons deux critères d'homogénéité similaires: "même orientation" et "même équation de plan". Les deux propriétés définissent une seule et même forme: la forme plane. Par conséquent, l'IP segmentée n'est décrite que par des plans. Chaque plan est déterminé de manière complète dans le cas du deuxième critère d'homogénéité alors qu'il manque la position dans le cas du premier.

Dans la table 2.2, on constate que le critère de "continuité" présente la contrainte de forme la plus petite, tandis que le critère de "équation de plan" présente la contrainte de forme la plus grande. Plus particulièrement, les critères d'homogénéité possède une structure hiérarchique, par exemple pour les critères de type différentiel: même orientation implique même courbure, qui implique même type de courbure, qui implique continu.

Primitive	position	orientation	taille	élongation	forme	
					effilée	courbe
plan	1	2	0	0	0	0
sphère	3	0	1	0	0	0
cylindre	2	2	1	0	0	0
ellipsoïde	3	3	1	2	0	0
superellipsoïde	3	3	1	2	1	1

Table 2.3: Principales primitives surfaciques 3D: nombre de paramètres déterminant pose, taille et forme. Commentaire: la sphère, l'ellipsoïde et le superellipsoïde en étant des surfaces fermées forment des primitives volumiques.

2.3 Discontinuités d'une image de profondeur

2.3.1 Définitions

Les discontinuités d'une IP sont définies par rapport à une surface qui est associée à l'IP en tant que modèle. Pour une définition formelle, on introduit les éléments suivants:

- (1) la fonction $\mathbf{p}(u, v)$, définie sur le domaine D_p du plan des coordonnées d'image (u, v) et représentant la surface
- (2) la fonction $\mathbf{n}(u, v)$ représentant la normale (unitaire) à la surface et définie sur D_p par

$$\mathbf{n} = \frac{\mathbf{p}_u \times \mathbf{p}_v}{\|\mathbf{p}_u \times \mathbf{p}_v\|} \quad (2.14)$$

où \mathbf{p}_u et \mathbf{p}_v désignent les dérivées partielles de \mathbf{p} par rapport à u et à v , respectivement

- (3) le système de coordonnées polaires locales (r, φ) dont l'origine est le point (u, v)

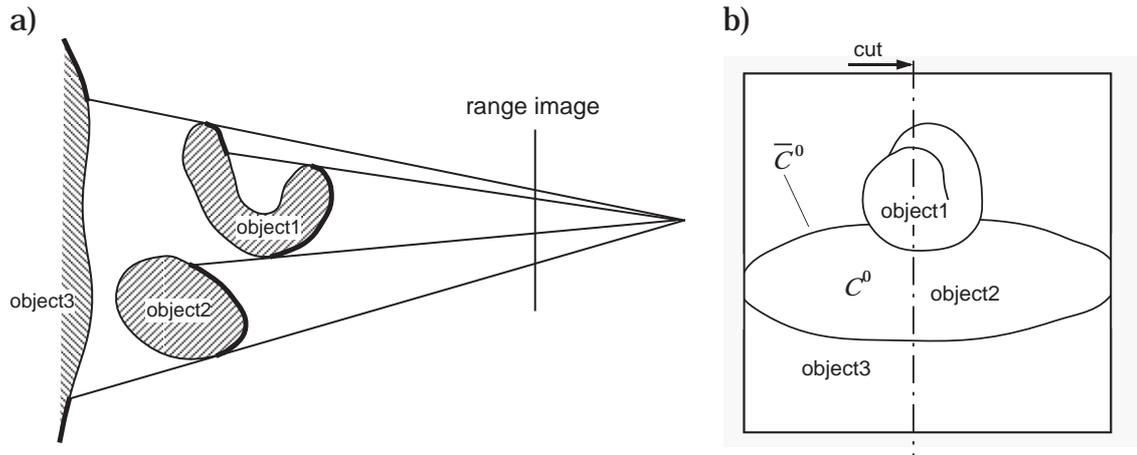


Figure 2.10: Domaine de définition d'une IP et discontinuités de profondeur. (a) vue en coupe de la scène (b) plan de l'IP, blanc: domaine de définition D_P , traits: courbes \bar{C}^0 .

Discontinuité de profondeur (discontinuité d'ordre 0)

Le point (u, v) de D_P est un point de discontinuité de profondeur si la fonction \mathbf{p} est discontinue (d'ordre 0) au point (u, v) , c'est-à-dire s'il existe des directions φ_1 et φ_2 pour lesquelles les limites de \mathbf{p} lorsque r tend vers 0 sont différentes. Formellement,

$$\exists \varphi_1, \varphi_2 \mid \lim_{r \rightarrow 0} \mathbf{p}(r, \varphi_1) \neq \lim_{r \rightarrow 0} \mathbf{p}(r, \varphi_2) \quad (2.15)$$

La propriété de discontinuité de profondeur est notée \bar{C}^0 . $D_{\bar{C}^0}$ désigne le domaine des points de discontinuité de profondeur.

Discontinuité d'orientation (discontinuité d'ordre 1)

Le point (u, v) du domaine $(D_P - D_{\bar{C}^0})$ est un point de discontinuité d'orientation si la fonction \mathbf{n} est discontinue au point (u, v) , c'est-à-dire s'il existe deux directions φ_1 et φ_2 pour lesquelles les limites de \mathbf{n} lorsque r tend vers 0 sont différentes. Formellement,

$$\exists \varphi_1, \varphi_2 \mid \lim_{r \rightarrow 0} \mathbf{n}(r, \varphi_1) \neq \lim_{r \rightarrow 0} \mathbf{n}(r, \varphi_2) \quad (2.16)$$

La discontinuité d'orientation est une discontinuité d'ordre 1 par rapport à la fonction $\mathbf{p}(u, v)$, elle est notée \bar{C}^1 .

Aux points pour lesquels il n'y a pas de discontinuité d'ordre 0 ou 1, la surface $\mathbf{p}(u, v)$ est continue d'ordre 1. Cette propriété est notée C^1 .

Les discontinuités de profondeur et d'orientation définies par rapport à la surface $\mathbf{p}(u, v)$ ont un rôle de modélisation.

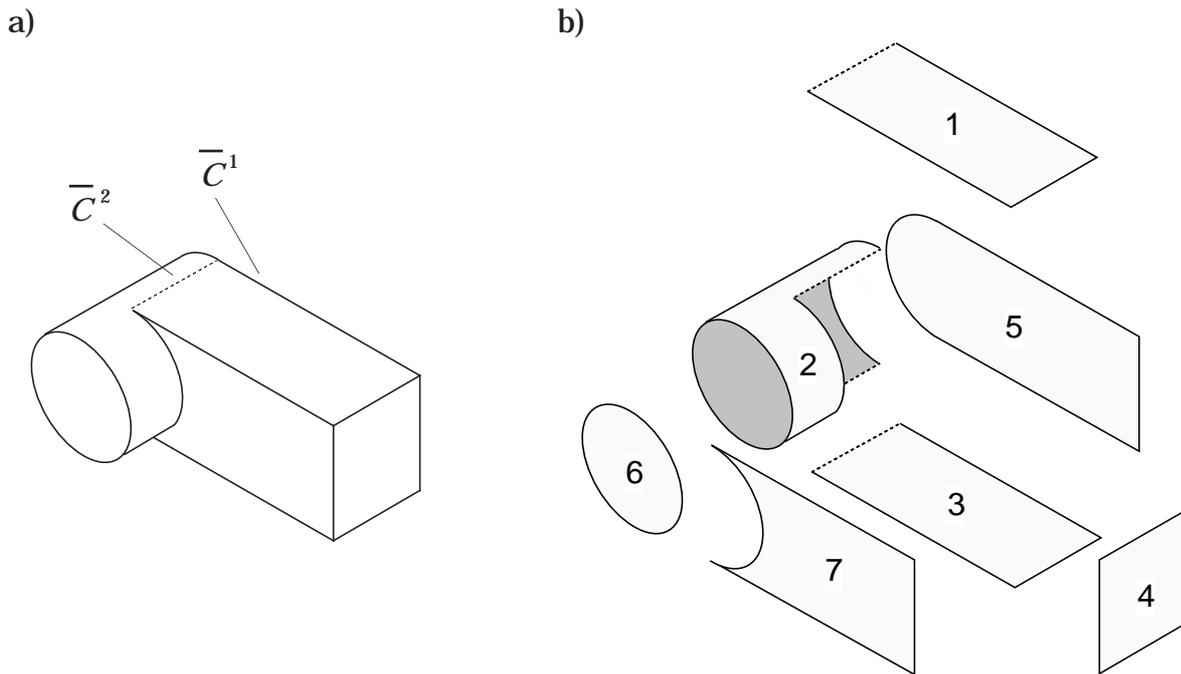


Figure 2.11: Modèle d'objet à facettes. [BaBr82: p. 266] (a) modèle (b) décomposition en facettes. Les facettes sont généralement séparées par une courbe \bar{C}^1 ; par contre, les interfaces 1-2 et 2-3 sont \bar{C}^2 .

2.3.2 Origine et propriétés des discontinuités

Discontinuités de profondeur

Le modèle de point \bar{C}^0 est une conséquence de la géométrie d'acquisition de l'IP et du modèle de surface mesurée continue.

Les points \bar{C}^0 forment des courbes. Les courbes sont généralement fermées, mais pas toujours [figure 2.10].

Discontinuités d'orientation

Le modèle de point \bar{C}^1 repose sur la géométrie particulière de certains objets répondant au modèle de facettes.

Modèle d'objet à facettes

Le modèle d'objet à facettes est caractérisé par une surface composée de morceaux juxtaposés, chaque morceau ayant une forme particulière. Le succès du modèle d'objet à facettes repose sur deux raisons pratiques qui concourent:

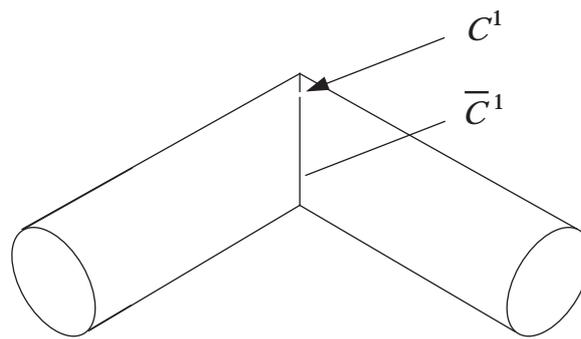


Figure 2.12: Modèle d'objet à facettes: point singulier le long d'une courbe \bar{C}^1 .

- chaque facette du modèle peut être décrite par une équation algébrique
- chaque facette de l'objet peut être manufacturée par un mouvement simple d'un outil (ex. facettes planes, cylindriques, etc.)

Le polyèdre constitue un cas particulier où toutes les facettes sont planes.

Discontinuités d'un modèle d'objet à facettes

Le cas ordinaire de la jonction de deux facettes définit (à la surface de l'objet) une courbe \bar{C}^1 [figure 2.11]. Pour des facettes planes, c'est toujours le cas. Cependant, il existe des jonctions de facettes pour lesquelles les points de la courbe ne sont pas tous \bar{C}^1 [figure 2.12], il existe aussi des cas où aucun point de la courbe n'est \bar{C}^1 [figure 2.11].

Discontinuités d'orientation d'une IP d'objets à facettes

Étant donné la géométrie d'acquisition (cohérence spatiale), les points \bar{C}^1 forment des courbes. Les courbes sont généralement fermées, mais pas toujours [figure 2.13].

2.3.3 Importance d'une segmentation suivant C^1

On étudie ici le comportement d'une telle segmentation dans le cas idéal d'une IP de synthèse continue, construite à partir de modèles d'objet à facettes. Ce cas idéal fixe les perspectives pour le cas réel.

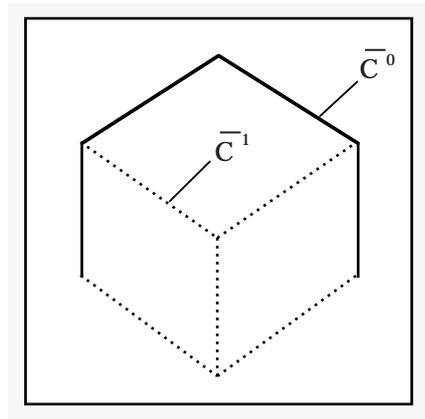


Figure 2.13: Discontinuités de profondeur et discontinuités d'orientation. IP représentant un cube sur un plan (exemple de juxtaposition due à la pesanteur); blanc et pointillé: domaine de définition D_p , traits pleins: courbes \bar{C}^0 , pointillé: courbes \bar{C}^1

IP idéale de modèles d'objet à facettes

L'IP établit une bijection entre la surface mesurée des objets et le domaine D_p du plan de l'IP. Les correspondances entre objets et facettes, d'un côté, et domaines de D_p connexes C^0 et C^1 , de l'autre, sont les suivantes:

- A un objet correspond généralement un domaine connexe C^0 de D_p . En cas d'occlusion (objet situé derrière un autre), il peut y avoir plusieurs domaines [figure 2.10].
- A une facette correspond généralement un domaine connexe C^1 de D_p [figure 2.13]. Mais, il peut y avoir plusieurs domaines en cas d'occlusion.

Dans l'autre sens:

- Un domaine connexe C^0 est issu généralement d'un seul objet, mais pas toujours (objets juxtaposés) [figure 2.13].
- Un domaine connexe C^1 est issu généralement d'une seule facette, mais il y a des exceptions (facettes non entièrement séparées par des points \bar{C}^0 ou \bar{C}^1)

De ces propriétés découlent directement celles de la segmentation en régions connexes C^1 .

Segmentation en régions connexes C^1

Elle vérifie les propriétés suivantes:

- Les segments correspondent souvent exactement aux facettes. Pour les objets polyédriques, la séparation des facettes est en principe garantie.
- Pour des objets plus généraux, un segment regroupe parfois plusieurs facettes, car les facettes ne sont pas toujours entièrement séparées par des points \bar{C}^0 ou \bar{C}^1 .
- Un segment regroupe des facettes appartenant à différents objets, seulement s'ils se touchent et ont même orientation.

2.4 Conclusion et discussion

Ce chapitre a été essentiellement consacré à la définition des notions de base: image de profondeur, segmentation et discontinuités. A côté de cela, nous avons fait un inventaire des différentes propriétés de surface pouvant servir de critère d'homogénéité pour la segmentation et avons examiné plus particulièrement le critère de continuité d'ordre 1 (C^1). Ce critère s'avère intéressant à mains égards:

- La continuité d'ordre 1 est une condition nécessaire implicite à tous les autres critères d'homogénéité considérés. Une segmentation basée sur le critère C^1 délivre en principe une segmentation grossière par rapport à d'autres critères d'homogénéité. Elle est la plus élémentaire dans une hiérarchie de segmentations utilisant différents critères d'homogénéité.
- Une segmentation basée sur le critère C^1 permet des segments de forme différente.
- Une segmentation basée sur le seul critère C^1 peut être suffisante dans bien des cas: séparation d'objets, séparation de faces d'un polyèdre.
- Comme la continuité est une propriété locale, la segmentation est potentiellement réalisable par une méthode locale.

Le critère C^1 est retenu dans la suite de ce travail.

«On ne trouve que ce qu'on cherche.»

Heinz Hügli

Chapitre 3

Segmentation des images de profondeur: état de l'art

L'étude des méthodes de segmentation est réalisée en deux temps.

D'abord la notion de méthode de segmentation est prise dans un sens strict: procédé qui à partir d'une image au sens général et d'un critère d'homogénéité définit des segments. Les méthodes correspondant à cette définition sont appelées élémentaires. Les méthodes élémentaires utilisées pour la segmentation des images de profondeur (IP) font l'objet d'une présentation brève et générale.

Dans une deuxième partie, le concept de méthode de segmentation d'IP est pris dans un sens plus large. Une revue de la littérature révèle l'existence de prétraitements et de post-traitements associés aux méthodes de segmentation. Parfois les méthodes élémentaires sont combinées. L'état de l'art du domaine est présenté.

3.1 Méthodes élémentaires de segmentation

3.1.1 Introduction

La présentation des méthodes élémentaires est faite de manière générale. La méthode est appliquée à une image V . Dans le cas de l'utilisation de la méthode pour les IP, V n'est pas forcément l'IP elle-même, ce peut être une image qui en est dérivée comme l'image des normales ou l'image des courbures.

Les méthodes élémentaires sont groupées en deux catégories: les méthodes directes et les méthodes basées sur la mise à jour de

régions. Dans les méthodes de la première catégorie, les régions sont formées en fin de procédé par étiquetage des composantes connexes. Dans les méthodes de la deuxième catégorie, les régions sont formées en début de procédé et sont mises à jour petit à petit jusqu'à ce qu'un critère de fin soit satisfait.

méthodes directes:

Ce groupe comprend trois méthodes:

- Détection de frontière et étiquetage des composantes connexes [§3.1.2]
- Seuillage¹ et étiquetage des composantes connexes [§3.1.3]
- "Clustering" et étiquetage des composantes connexes [§3.1.4]

méthodes basées sur la mise à jour de régions:

Dans ce cas, une description sous la forme d'un vecteur est associée à une région. Il existe une méthode M qui permet de déterminer le vecteur de description \mathbf{g}_k à partir de la région R_k

$$\mathbf{g}_k = M(R_k) \quad (3.1)$$

La description par \mathbf{g}_k des éléments de R_k est approximative. Il existe une mesure d'erreur. La région est homogène si l'erreur globale δ est inférieure à un seuil

$$H(R_k) = \left(\delta \sum_{\mathbf{v}_{i,j} \in R_k} (\mathbf{v}_{i,j}, \mathbf{g}_k) < T_H \right) \quad (3.2)$$

Les méthodes de cette catégorie sont:

- Fusion de régions [§3.1.5]
- Division et fusion de régions [§3.1.6]
- Croissance indépendante de régions [§3.1.7]

3.1.2 Détection de frontière et étiquetage des composantes connexes (M1)

Principe

La méthode se base sur une mesure (plus ou moins) locale de la dissimilitude entre éléments de l'image. Si cette dissimilitude est plus

¹Terme désignant l'action d'appliquer un seuil

élevée qu'un certain seuil, on décide qu'il y a, à cet endroit, un point de frontière. Le résultat du test est inscrit dans une image binaire. C'est l'image des points de frontière. Formellement

$$b_{i,j} = (\exists (k,l) \in W_{i,j} | \Delta(\mathbf{v}_{i,j}, \mathbf{v}_{k,l}) > T_H) \quad (3.3)$$

où $b_{i,j}$ représente un élément de l'image des points de frontière et $W_{i,j}$ un voisinage de la position (i,j) .

Pour obtenir les segments de $[\mathbf{v}_{i,j}]$, on procède à un étiquetage des composantes connexes de l'image des régions $[\bar{b}_{i,j}]$. Chaque ensemble d'éléments $\mathbf{v}_{i,j}$ de même étiquette forme un segment.

Traitement supplémentaire

Entre les deux parties de la méthode que sont la détection de points de frontière et l'étiquetage des composantes connexes, on peut intercaler différents traitements. Leur objet est d'obtenir à partir des points de frontière des courbes, respectivement des courbes fermées. Dans le cas d'image d'intensité ces traitements sont groupés sous le thème "détection de contours" [BaBr82, chap. 4].

Application aux IP

Dans le cas des IP, les points de frontière sont les points de discontinuité de type \bar{C}^0 et \bar{C}^1 détectés de l'une des manières décrites au chapitre 4.

3.1.3 Seuillage et étiquetage des composantes connexes (M2)

Principe

Un ou plusieurs seuillages sont effectués sur une ou plusieurs des N composantes du vecteur \mathbf{v} , découpant l'espace \mathfrak{R}^N en domaines disjoints D_l . Chaque élément $\mathbf{v}_{i,j}$ appartient donc à un et un seul domaine D_l . On lui attribue l'étiquette correspondante l .

Les segments sont obtenus par étiquetage des composantes connexes de l'image formée par les étiquettes l .

Application aux IP

Cette méthode est utilisée couramment dans le cas des IP pour faire une segmentation selon le type de courbure. Les composantes de \mathbf{v}

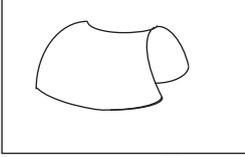
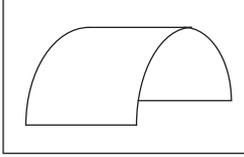
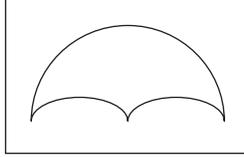
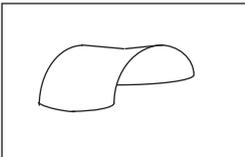
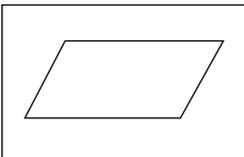
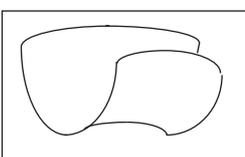
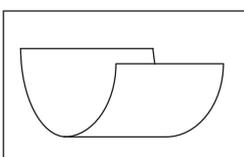
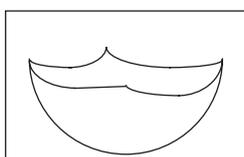
	$K < -\varepsilon_K$	$-\varepsilon_K \leq K \leq \varepsilon_K$	$K > \varepsilon_K$
$H < -\varepsilon_H$	 arête selle	 arête	 sommet
$-\varepsilon_H \leq H \leq \varepsilon_H$	 surface minimale	 plan	-
$H > \varepsilon_H$	 vallée selle	 vallée	 creux

Table 3.1: Types de surface définis par les signes des courbures gaussienne et moyenne. ε_K et ε_H définissent des seuils de tolérance. A noter que cette manière de faire définit 9 catégories, alors qu'il n'existe que 8 types de surface.

sont la courbure gaussienne K et la courbure moyenne H . Des seuils ε et $-\varepsilon$ sont appliqués à chacune des courbures [table 3.1].

3.1.4 "Clustering" et étiquetage des composantes connexes (M3)

Principe

Le procédé cherche à regrouper des éléments $\mathbf{v}_{i,j}$ ayant même valeur. Idéalement, c'est-à-dire si la valeur est strictement la même, ces éléments sont représentés par un point unique dans l'espace \mathfrak{R}^N . Dans un cas pratique, il faut laisser une marge de tolérance. La méthode consiste donc à rechercher des points de \mathfrak{R}^N formant un groupement rapproché ("cluster"). Les clusters C_m résultant de l'opération vérifient

$$\|\mathbf{v}_{i,j} - \mathbf{v}_{k,l}\| < T_m, \forall \mathbf{v}_{i,j}, \mathbf{v}_{k,l} \in C_m \quad (3.4)$$

On attribue à chaque élément $\mathbf{v}_{i,j}$ l'étiquette m du cluster correspondant.

Les segments sont obtenus par étiquetage des composantes connexes de l'image formée par les étiquettes m .

Application aux IP

Cette méthode est essentiellement utilisée pour obtenir des segments correspondant à des morceaux de surface (approximativement) plane. Dans ce cas, elle est appliquée à l'image des vecteurs normaux ou, éventuellement, à une image de paramètres de plans d'approximation locale.

3.1.5 Fusion de régions (M4a)

Principe

A l'initialisation, la méthode définit une partition d'image $P^{(0)}$ constituée de petites régions homogènes, le critère d'homogénéité étant défini par l'équation (3.2). La méthode fonctionne ensuite de manière itérative. A l'itération n , deux régions voisines $R_k^{(n)}$ et $R_l^{(n)}$ dont la réunion forme une région homogène sont fusionnées et la partition est remise à jour [figure 3.1.b]. Le processus s'arrête lorsqu'il n'y a plus aucune paire de régions voisines pouvant être fusionnées.

Le critère d'homogénéité (3.2) n'est pas favorable à la méthode de fusion. En effet, il faut évaluer quel serait le résultat de la fusion (réunion des deux ensembles, estimation du vecteur de description) pour savoir si elle doit être exécutée. Ceci représente une dépense lourde en calcul qu'il faut faire aussi lorsque la fusion n'est pas effectuée. La méthode de fusion n'a d'intérêt pratique que s'il existe un moyen de déterminer la mesure d'erreur accumulée δ sur la région $R_k \cup R_l$ à partir des valeurs de mesure sur les régions R_k et R_l . L'alternative est d'utiliser un autre critère d'homogénéité basé par exemple sur une mesure de dissimilitude entre les vecteurs \mathbf{g}_k et \mathbf{g}_l

$$H(R_k \cup R_l) = (\Delta(\mathbf{g}_k, \mathbf{g}_l)) < T_H \quad (3.5)$$

Ordre de fusion

L'ordre dans lequel ont lieu les fusions joue un rôle important sur la qualité des résultats. La solution qui paraît la meilleure consiste à fusionner à chaque fois les deux régions dont les caractéristiques sont

les plus semblables parmi toutes les paires de régions voisines dans l'image [BeGo89, FaHe87, Leon87]. Cependant cette manière de faire est gourmande en calcul et de plus elle ne se laisse pas paralléliser. En pratique, on utilise donc souvent une solution de compromis entre résultats de qualité maximale et temps de calcul minimal [Chan91].

Application aux IP

On trouve deux formes d'application de cette technique aux IP:

- segmentation de l'IP. Le vecteur de description associé à la région est constitué par les paramètres d'une surface algébrique approximant l'ensemble des rangels. La mesure sur laquelle se base le critère d'homogénéité est l'erreur globale avec laquelle la surface algébrique approche les rangels faisant partie de la région.
- segmentation d'une image dérivée de l'IP. Ce peut être l'image des vecteurs normaux, l'image des courbures principales ou encore une image constituée par les paramètres d'une surface algébrique d'approximation locale. Le vecteur de description associé à la région est alors la valeur moyenne des éléments de la région et la mesure d'homogénéité en est l'écart-type.

3.1.6 Division et fusion de régions (M4b)

Cette méthode commence par diviser l'image. Tout d'abord l'image est partagée en plusieurs (en général 4) régions de même taille et de même forme (carrée). Puis, chacune des régions, pour laquelle le critère d'homogénéité (3.2) n'est pas rempli, est à son tour divisée en plusieurs sous-régions. Et ainsi de suite... La figure 3.1.a donne l'organigramme d'un pas de division.

A cause du découpage régulier, le processus de division sépare généralement une région homogène en plusieurs morceaux. La technique de fusion décrite ci-dessus est appliquée, une fois la division terminée, pour regrouper les morceaux.

Application aux IP

Comme la méthode de fusion.

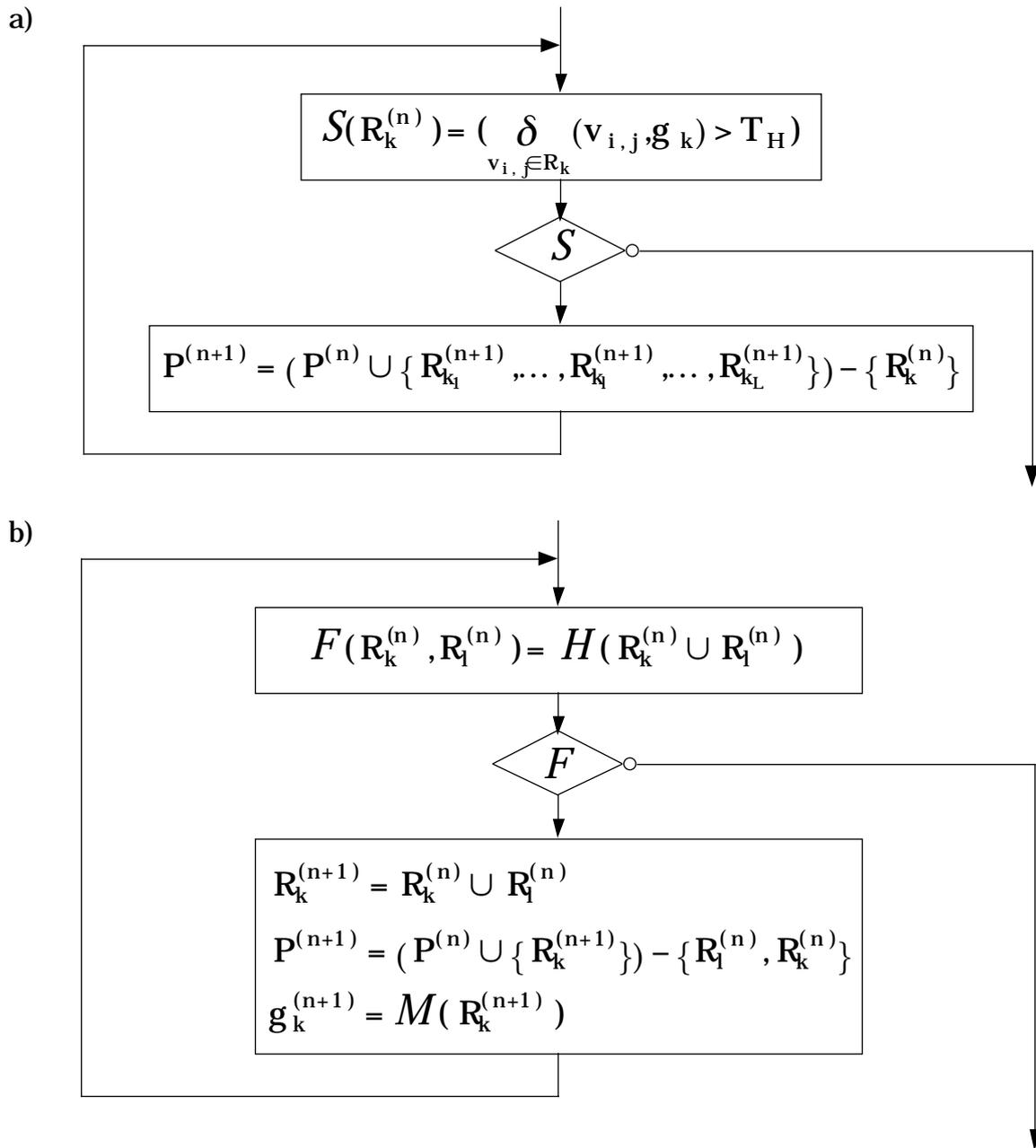


Figure 3.1: Division et fusion. (a) un pas de division (b) un pas de fusion

3.1.7 Croissance indépendante de régions (M5)

Plutôt qu'une méthode, il s'agit d'une catégorie regroupant plusieurs méthodes de segmentation qui se distinguent de celles présentées précédemment par le fait qu'elles procèdent par croissance indépendante des régions. Cette manière de faire conduisant généralement à un recouvrement non souhaité entre régions, les méthodes comportent en fait deux phases:

- (1) croissance des régions de manière indépendante
- (2) élimination des recouvrements entre régions

Croissance d'une région par ajout d'éléments compatibles

Il s'agit d'un processus itératif qui fait croître une région à partir d'un noyau.

Les éléments $\mathbf{v}_{i,j}$ candidats à faire croître la région $R_k^{(n)}$ au pas d'itération n doivent remplir les conditions suivantes: (1) leurs valeurs doivent être compatibles avec celles de la région, décrites par le vecteur $\mathbf{g}_k^{(n)}$ et (2) si la propriété de connexité est requise pour l'homogénéité, ils doivent former un ensemble connecté avec $R_k^{(n)}$. S'il existe des éléments $\mathbf{v}_{i,j}$ remplissant ces conditions, ils sont ajoutés à la région $R_k^{(n)}$ et le vecteur $\mathbf{g}_k^{(n)}$ est mis à jour. Sinon la croissance s'arrête.

Ensuite, la région est testée quant à son homogénéité (compatibilité globale des éléments avec le vecteur de description). Si la région est homogène, la croissance continue. Sinon elle s'arrête, à moins que le procédé soit appliqué avec différents degrés (ordres) de description. Si c'est le cas et qu'il existe une description d'ordre plus élevé que celle qui est présentement utilisée, l'ordre est augmenté (en augmentant l'ordre, on augmente la taille et avec elle les capacités de description du vecteur \mathbf{g}). Le nouveau vecteur \mathbf{g} est alors calculé et un nouveau test d'homogénéité est effectué.

L'algorithme de croissance qui vient d'être décrit est représenté à la figure 3.2 par son organigramme.

Élimination des recouvrements entre régions

Il y a deux mécanismes pour éliminer les recouvrements entre régions

1. Attribution d'éléments Ce mécanisme est utilisé lorsque deux ou plusieurs régions se superposent légèrement. Dans cette situation, l'élément est attribué à la région dont la description le représente le mieux:

$$\left(\delta(\mathbf{v}_{i,j}, \mathbf{g}_m^{(n)}) = \min_k (\delta(\mathbf{v}_{i,j}, \mathbf{g}_k^{(n)})) \right) \Rightarrow \left(R_k^{(n+1)} = R_k^{(n)} - \{\mathbf{v}_{i,j}\}, \forall k \neq m \right) \quad (3.6)$$

2. Choix de région Lorsque deux ou plusieurs régions se recouvrent fortement, l'élimination du recouvrement passe par une élimination de région(s). Sont écartées les régions dont le vecteur de description \mathbf{g}

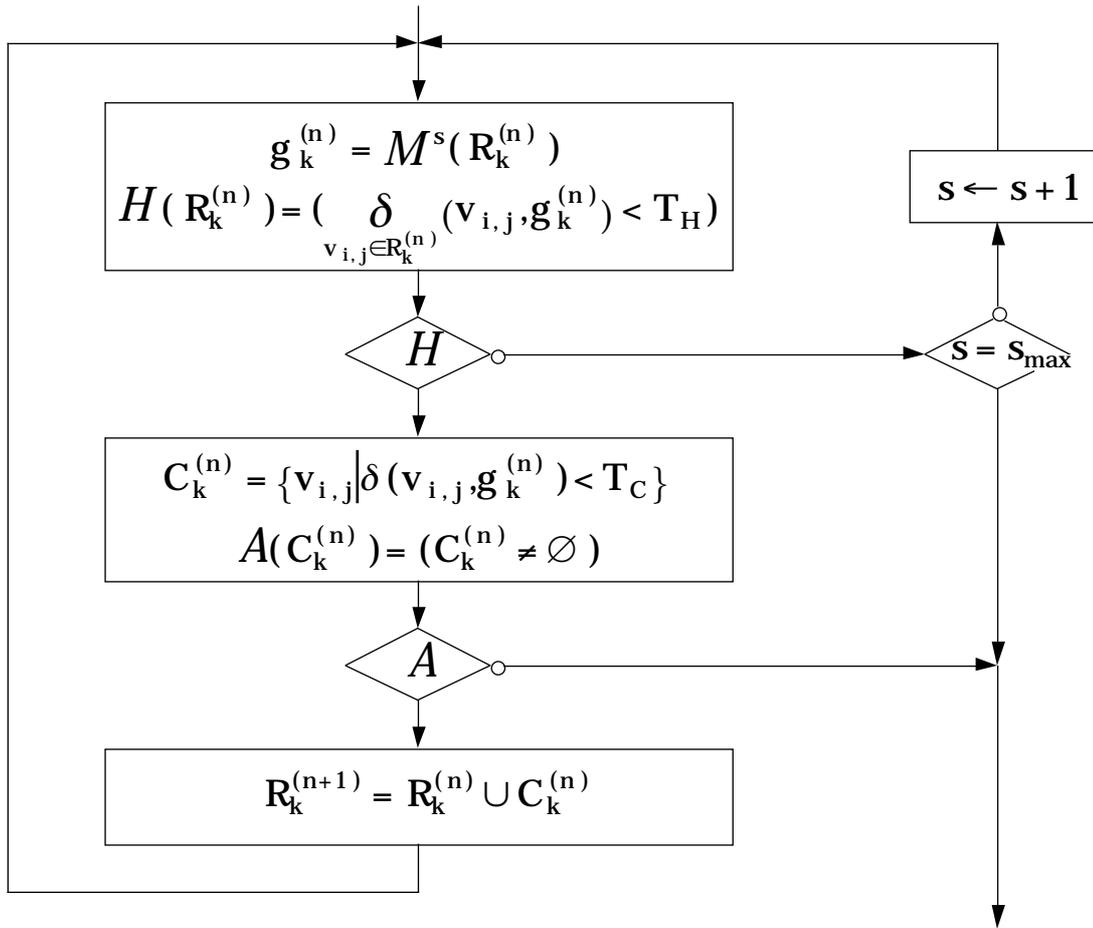


Figure 3.2: Croissance de région par ajout d'éléments compatibles. (sans condition de connexité)

représente le moins bien les données. Dans ce but, une mesure de qualité de description est définie. Cette mesure tient compte de la concurrence entre régions pour les zones de recouvrement. Ainsi, si la description \mathbf{g}_m de la région R_m est bonne, une mesure de qualité élevée est attribuée à \mathbf{g}_m . Mais si la description des régions voisines R_k est bonne pour les éléments qu'elles ont en commun avec R_m , la qualité de la description Q de \mathbf{g}_m est dévalorisée. Formellement:

$$Q(\mathbf{g}_m) = \delta_{\mathbf{v}_{i,j} \in R_m}(\mathbf{v}_{i,j}, \mathbf{g}_m) - \sum_{R_k | (k \neq m)} \left(\delta_{\mathbf{v}_{i,j} \in (R_k \cap R_m)}(\mathbf{v}_{i,j}, \mathbf{g}_k) \right) \quad (3.7)$$

Sont éliminées les régions pour lesquelles la mesure Q de qualité est faible.

Application aux IP

Le vecteur de description \mathbf{g} associé à la région est constitué par les paramètres d'une surface algébrique approximant l'ensemble des rangels. Les différentes mesures sont à interpréter en termes d'erreur d'approximation. Quatre méthodes de segmentation par croissance indépendante de régions sont présentées ci-après, ce sont celles de Besl, Taubin, Leonardis et Darrell.

La méthode de Besl [BeJa88, Besl88] fait croître les régions par ajout d'éléments compatibles avec condition de connexité. La croissance a lieu de manière itérative à partir de noyaux bien choisis [pour le choix voir §3.2.1]. La description de la région est faite par une surface 2.5D de type polynomial, le degré du polynôme pouvant varier de un à quatre. La mesure de compatibilité contrôlant la croissance est composée de l'erreur de position et de l'erreur d'orientation que forme l'estimation du modèle de surface par rapport aux valeurs issues de l'IP. La croissance de chaque région se fait entièrement de manière indépendante, de sa taille minimale à sa taille maximale. Comme les noyaux de départ ont été bien choisis, le recouvrement entre régions est le plus souvent mineur. Pour l'éliminer, les rangels communs sont simplement attribués à la région de taille maximale.

Taubin [Taub91] applique la méthode proposée par Besl avec des modèles de surfaces 3D. Cependant, il ne présente des résultats que pour un modèle de surface 3D plane. A noter que dans ce cas l'erreur de position intervenant dans la mesure de compatibilité est la distance perpendiculaire au plan et non la distance suivant le troisième axe de coordonnées (modèle de plan 2.5D). Remarquons finalement que les résultats de segmentation sont fournis avec un chevauchement de régions.

Leonardis et al. [LeGB90, GuBa92, GuBa93] mettent en principe en oeuvre la technique de Besl avec une description de la région par une surface 2.5D polynomiale de degré 1 ou 2. Néanmoins, leur approche comporte deux différences importantes par rapport à celle de Besl:

- Premièrement, les graines de croissance ne sont pas des régions bien choisies. Ce sont toutes les régions d'une partition fine de l'image pour lesquelles l'erreur d'approximation par un plan est au-dessous d'un certain seuil. Avec une telle initialisation du procédé de croissance, les régions sont amenées à se recouvrir

		[BeJa88]	[Taub91]	[LeGB90] [Leon93]	[DaSP90]
sélection des graines		•	•	–	–
contrainte de compatibilité	durant la croissance	•	•	•	–
	après la croissance	–	–	–	•
	d'orientation	•	•	–	–
sélection des régions après croissance		–	–	•	•
type de surfaces		2.5D	3D	2.5D, 3D	2.5D

Table 3.2: Croissance indépendante de régions (M5) appliquée aux IP.

fortement et un procédé d'élimination de régions est nécessaire. La croissance a donc lieu comme suit. (1) Chaque région croît de manière indépendante par ajout de tous les éléments connexes et compatibles avec la description courante [selon figure 3.2]. (2) Une mesure de qualité de description (3.7), exprimée en termes d'erreur d'approximation par la surface 2.5D, est calculée. (3) Les régions ayant les moins bons résultats sont éliminées. Ensuite, un nouveau pas de croissance a lieu.

- La deuxième différence réside dans le critère de compatibilité utilisé lors de la croissance d'une région. Chez Besl comme chez Taubin, l'erreur d'orientation que la surface d'approximation manifeste par rapport aux valeurs extraites de l'IP est prise en compte. Ici, en ce qui concerne l'approximation, seulement l'erreur de position est considérée. Par contre une autre condition est introduite dans le critère de compatibilité. Une limite de distance (dans le plan de l'image) par rapport à la région courante est imposée.

Plus récemment, Leonardis a proposé une généralisation de la méthode par rapport au modèle de surface pouvant être utilisé [Leon93].

Darrell et al. [DaSP90] ont une approche différente de celles présentées jusqu'ici: dans leur méthode, il n'y a pas de contrôle de croissance des régions. Le processus commence par créer de petites

méthode de segmentation	propriété d'homogénéité	décision	modèle de surface
détection de frontière et étiquetage des CC (M1)	locale	locale	continuité
seuillage et étiquetage des CC (M2)	locale	locale	type de courbure
"clustering" et étiquetage des CC (M3)	locale	globale	même orientation, – courbure surface algébrique (locale)
fusion, division-fusion (M4)	locale ou globale	globale	même orientation, – courbure surface algébrique
croissance indépendante (M5)	globale	globale et locale	surface algébrique

Table 3.3: Méthodes élémentaires de segmentation et modèles de surface.

régions centrées sur chaque rangel. Les paramètres du modèle de surface représentant la région sont calculés. Suit une élimination partielle des recouvrements par partage des rangels (au lieu d'attribuer un rangel à une seule région (3.6), il est attribué à plusieurs régions avec la description desquelles il est compatible). Une élimination des recouvrements par choix de région (3.7) est ensuite effectuée. Puis, les régions choisies sont étendues à toute l'image. Finalement, une élimination des recouvrements a lieu par attribution des rangels (3.6) à la région fournissant la meilleure description. Les surfaces utilisées pour décrire une région sont de type 2.5D polynomial de degré 1 et 2. Cependant aucune précision n'est donnée sur la manière de choisir un modèle de surface plutôt que l'autre. La mesure de qualité de description (3.7) utilisée est le taux de compression apporté par l'approximation au lieu d'être l'erreur d'approximation.

Ainsi se termine la présentation des méthodes de croissance indépendante de régions appliquée aux IP. La table 3.2 donne une vue d'ensemble comparative de ces techniques. Un aspect mérite d'être souligné: les méthodes ne procédant pas par sélection des régions après croissance utilisent durant la croissance, en plus de

méthode de segmentation	segments disjoints	segmentation complète	connexité
M1	oui	non	oui
M2	oui	oui	oui
M3	oui	oui	oui
M4	oui	oui	oui
M5	oui/non	non	oui/non

Table 3.4: Propriétés relatives au partage de l'image.

méthode de segmentation	complexité
M1	+
M2	+
M3	++
M4	++
M5	++

Table 3.5: Complexité des méthodes de segmentation.

l'erreur d'approximation, une contrainte de compatibilité basée sur l'orientation de la surface.

3.1.8 Récapitulation

Les tables 3.3 à 3.5 fournissent une vue d'ensemble des méthodes élémentaires de segmentation.

La table 3.3 met en relation méthodes de segmentation et modèles de surface. On constate que trois des cinq méthodes sont propres à un seul modèle de surface. Les deux autres méthodes sont plus générales, surtout la méthode M4

Ensuite, la table 3.4 compare les méthodes quant aux propriétés du partage de l'image. Il s'agit là de considérations purement formelles.

Finalement, la table 3.5 présente une classification grossière des méthodes de segmentation quant à leur complexité algorithmique. Les méthodes M1 et M2, par le fait que la décision est locale, sont sensiblement moins complexes que les autres.

3.2 Revue des méthodes de segmentation

Cette revue de la littérature range les méthodes de segmentation en quatre classes suivant le critère d'homogénéité auquel obéit la segmentation: type de courbure, surface plane, surface algébrique, continuité C^1 . Une cinquième classe regroupe les autres méthodes.

3.2.1 Type de courbure

Tous les auteurs utilisent la méthode de seuillage sur les courbures moyenne et gaussienne [§3.1.3].

L'estimation des courbures pose problème. Elle est perturbée à la fois par la présence du bruit (erreurs de mesure) et par la présence des discontinuités [figure 3.3]. Pour que l'estimation des courbures soit suffisamment bonne, il faut que le bruit soit faible. D'autre part, il faut éliminer les perturbations dues aux discontinuités par un traitement adéquat.

Besl et Jain [BeJa88] utilisent cette segmentation uniquement pour déterminer les graines initiales pour une croissance indépendante de régions [§3.1.7] (commentaire: les images qu'ils traitent sont de relativement mauvaise qualité).

Yokoya et Levine [YoLe89] combinent la détection des discontinuités avec la segmentation en type de courbure. L'image des discontinuités est utilisée comme masque lors de l'étiquetage des composantes connexes de l'image de type de courbure (les discontinuités restreignent la connexité). Ensuite, une croissance de région par simple dilatation aboutit à un partage d'image qui est une partition.

Kasvand [Kasv88] procède à une érosion de 1 à 2 pixels sur l'image de type de courbure pour supprimer l'effet des discontinuités. Ensuite, a lieu l'étiquetage des composantes connexes de l'image de type de courbure. Finalement, une croissance des régions est effectuée sous contrôle de l'approximation de la région par une surface algébrique.

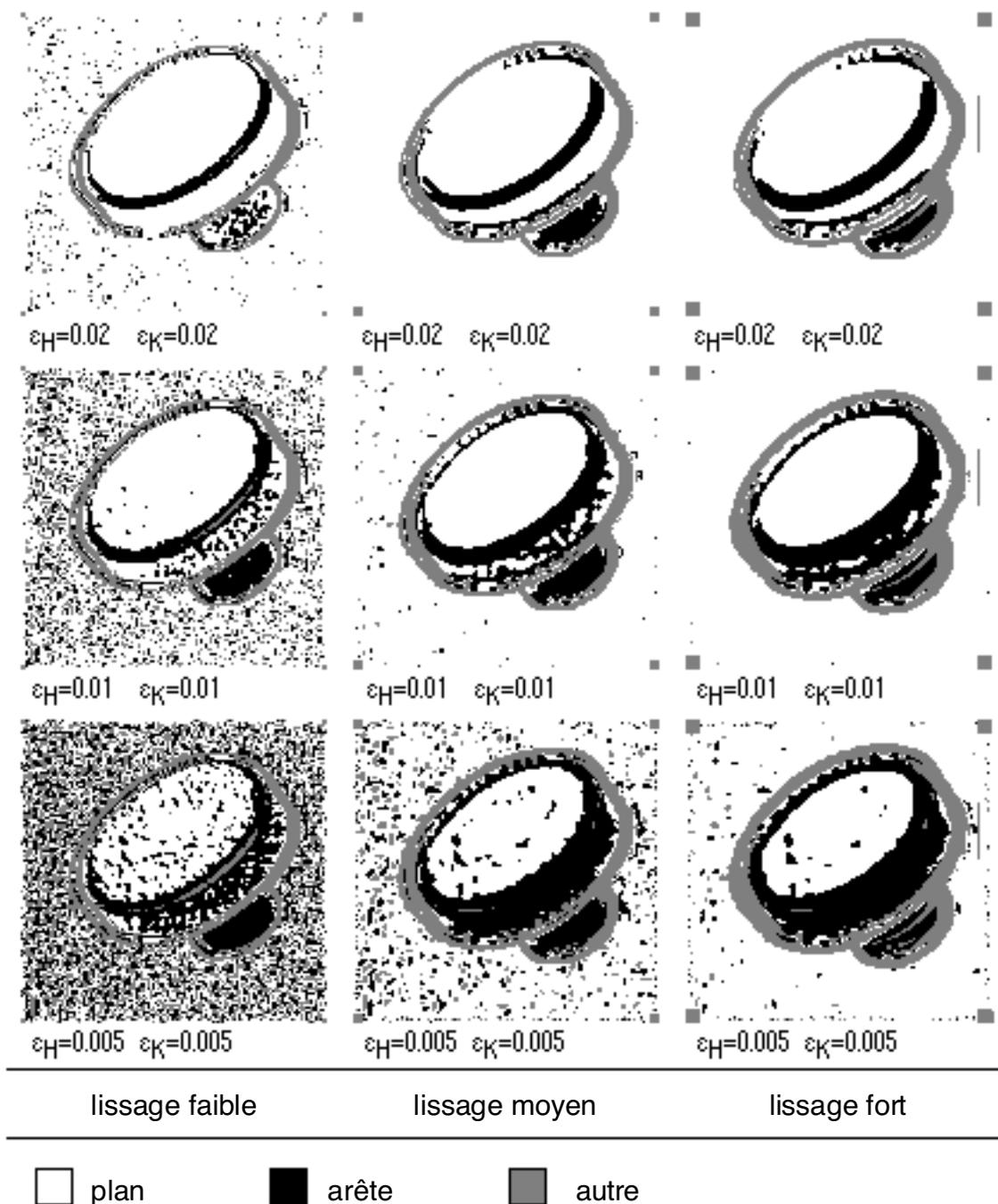


Figure 3.3: Types de courbure en fonction des seuils sur les courbures gaussienne et moyenne et de l'importance du lissage. Sont représentés uniquement les deux types de courbure correspondant aux surfaces mesurées (plan, cylindre). Commentaires: Pour chaque degré de lissage, il existe une valeur minimale pour les seuils ε_K et ε_H en dessous de laquelle on est sensible au bruit. Cette valeur dépend évidemment de la qualité de l'image originale. Pour pouvoir faire la distinction entre une surface légèrement courbe et une surface plane, il faut que le lissage soit suffisant. Notes: Les images de type de courbure proviennent de F. Tièche [Tièc90]. L'image originale MECA06 a été réduite à la taille 128x128 par sous-échantillonnage avant l'estimation des types de courbure.

3.2.2 Surface plane

Deux critères d'homogénéité définissent une segmentation en surfaces planes: "même orientation" et "même équation de plan". Les méthodes basées sur le premier critère doivent prendre en compte les discontinuités de profondeur pour éviter une fusion entre deux plans parallèles mais non colinéaires (par ex. face supérieure d'un cube et surface sur laquelle il est posé).

Orientation

La prise en compte des discontinuités d'ordre 0 est souvent faite de manière implicite: en estimant l'orientation au travers des discontinuités, on crée une zone de forte inclinaison faisant office de séparation. Le désavantage est l'apparition de régions (segments) parasites.

Maître et Hügli [MaHü89] ont une méthode de fusion basée sur la normale à la surface. Les normales sont calculées sur des facettes de taille minimale (carré de quatre rangels voisins). Ces facettes servent de partition initiale. La fusion est effectuée de manière séquentielle en fusionnant à chaque fois, parmi toute les paires de régions voisines, celles dont l'angle entre les normales est le plus petit [figure 3.4].

Parvin et Medioni [PaMe86] utilisent une méthode de division-fusion basée sur la normale à la surface ou, plus précisément, sur le gradient de l'IP (scalaire). Le gradient est estimé par la méthode conventionnelle utilisant un voisinage 3x3 invariant en translation. La méthode de division-fusion est ensuite appliquée à l'image de gradients. La mesure δ définissant le critère d'homogénéité pour la division est l'écart-type du gradient. L'homogénéité pour la fusion est définie sur la base de la valeur de l'angle entre vecteurs normaux associés aux régions. Étant donné le procédé d'estimation du gradient utilisé, cette méthode est soumise au même phénomène de régions parasites que celle de Maître et Hügli.

Taylor et al. [TaSR89] appliquent la méthode de division-fusion à une image de normales (gradient). L'estimation du gradient se fait sur un voisinage adapté. Le problème des régions parasites est ainsi éliminé. Lors de la division et de la fusion, un test de profondeur doit par contre être effectué lorsque les tests d'homogénéité sur les normales sont positifs. Ceci pour éviter que des plans parallèles

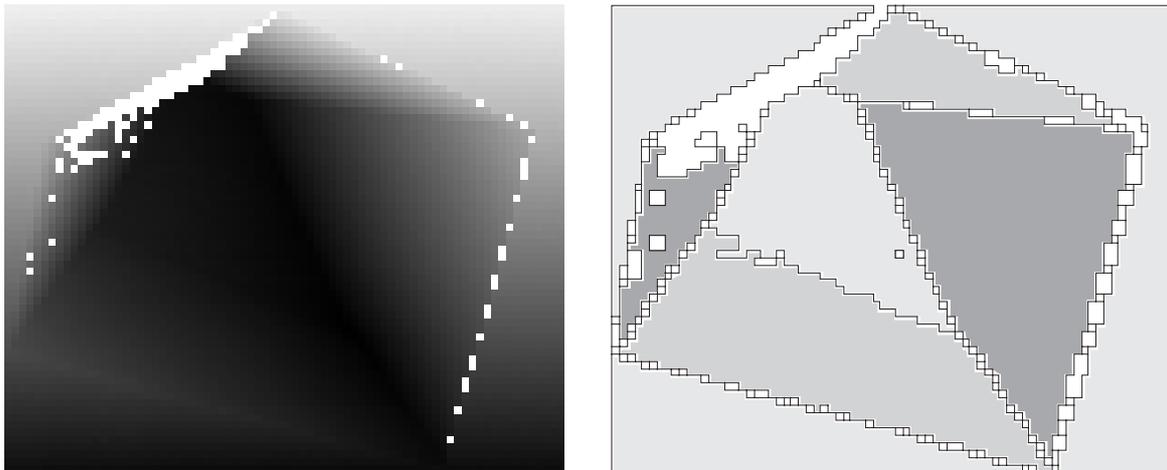


Figure 3.4: Segmentation en surfaces planes par une méthode de fusion basée sur la normale à la surface [MaHü89]. (a) IP (taille 64x76). (b) IP segmentée. Commentaires: Les surfaces planes (faces de la pyramide et fond) sont correctement extraites. On remarque cependant la présence de petites régions parasites situées à cheval sur les discontinuités (arêtes ou sauts en profondeur).

séparés par une discontinuité de profondeur forment un seul segment. Le test d'homogénéité sur les normales correspond à un seuillage sur l'angle entre les valeurs extrêmes des normales.

Équation algébrique

Jiang et Bunke [JiBu94] procèdent à une croissance de régions sous contrainte d'approximation par un plan (2.5D). La méthode de croissance est particulière car elle est basée sur des segments de droite: les régions sont des listes de segments et la croissance est effectuée segment par segment. Les segments de droite sont obtenus par division d'un profil (ligne ou colonne) de l'IP. La croissance des régions se fait de manière séquentielle, c'est-à-dire une région après l'autre. La graine avec laquelle est initialisée la croissance est constituée de trois segments de droite voisins issus de trois profils consécutifs. La croissance commence avec la graine qui satisfait au mieux un test de parallélisme entre segments de droite.

3.2.3 Surface algébrique

La segmentation en surfaces algébriques (non strictement planes) est divisée en deux catégories suivant la nature des surfaces: 2.5D et 3D.

Les surfaces 2.5D correspondent toutes à des fonctions polynomiales de deux variables. Elles s'appliquent évidemment uniquement aux images de type scalaire. Les surfaces 3D sont des quadriques ou des superquadriques. Leur traitement est d'une complexité supérieure à celui des surfaces 2.5D.

Surfaces 2.5D

Leonardi [Leon87] présente une méthode de division-fusion liée à un modèle de surface 2.5D de type polynomial. Le degré du polynôme peut varier de 0 à 3, mais il est fixe pour une segmentation donnée. La méthode est appliquée à des images d'intensité. Bien que certaines modifications pourraient éventuellement en améliorer les performances, elle peut néanmoins être utilisée telle quelle pour des IP de type scalaire. Le critère d'homogénéité pour la division et pour la fusion est basé sur l'erreur d'approximation des valeurs d'image par le polynôme de meilleure approximation. Lors de la fusion, erreur d'approximation et coefficients du polynôme sont mis à jour à partir de ceux des régions fusionnées: il n'est pas nécessaire de procéder à un calcul d'approximation sur la nouvelle région. On note finalement que Leonardi a jugé utile, voire nécessaire de diviser l'image en premier lieu suivant les contours (discontinuités de profondeur d'une image d'intensité).

A cette catégorie appartiennent la méthode de Leonardi et al. [LeGB90, GuBa92, GuBa93] [figure 3.5] et celle de Darrell et al. [DaSP90], décrites au paragraphe 3.1.7.

Surfaces 3D

L'approche proposée par Faugeras et Hébert [FaHe87] procède par fusion à partir d'une partition de l'IP en petites régions. Les surfaces algébriques associées sont soit des plans, soit des quadriques.

Taubin [Taub91] propose une méthode qui commence par une croissance indépendante de régions [§3.1.7], les régions étant décrites par des plans. A ce premier traitement succède une étape de fusion. Pour les cas où la fusion n'est pas possible, la description passe alors d'un plan à une quadrique.

Gupta et Bajcsy [GuBa92, GuBa93] présentent une méthode de segmentation dont le résultat est la description de l'IP par des superquadriques (plus précisément des superellipsoïdes). Pour aboutir au résultat, l'approche suivie par Gupta effectue un

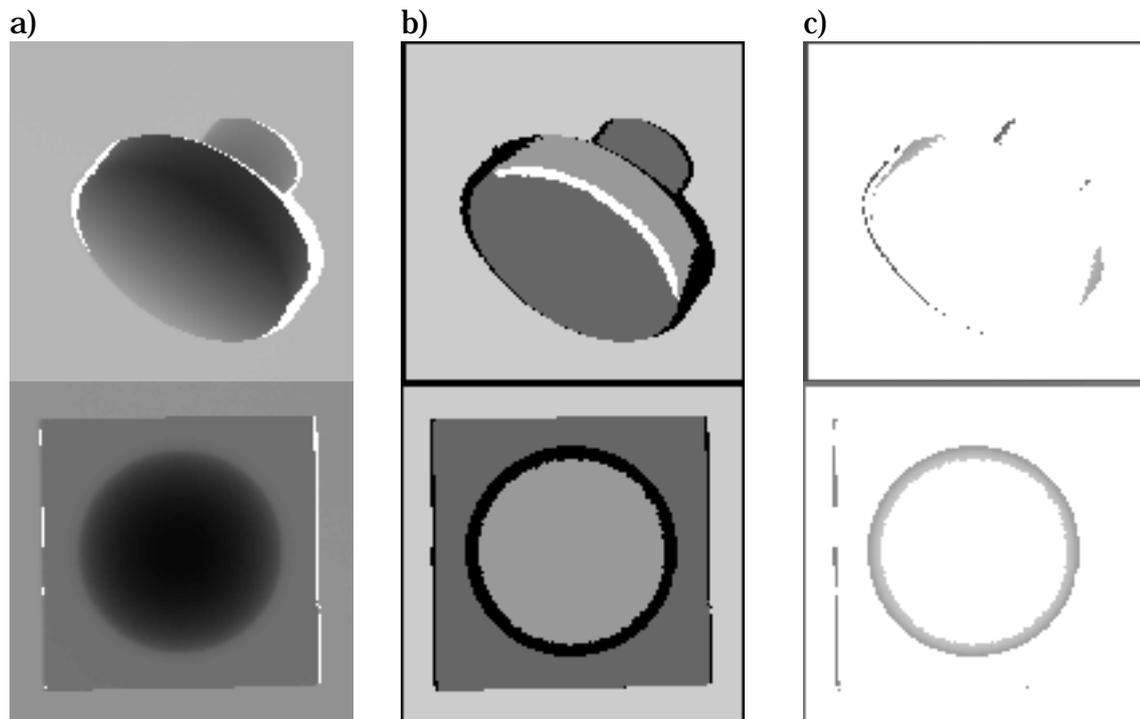


Figure 3.5: Segmentation en surfaces 2.5D polynomiales de degrés 1 et 2 par une méthode de croissance indépendante de régions [LeGB90]. (a) IP. (b) IP segmentée (blanc: appartient à plusieurs segments, noir: n'appartient à aucun segment) (c) Différence entre IP et surface algébrique. Commentaires: Le résultat de la segmentation correspond dans les deux cas à une décomposition naturelle de la scène. Une surface 2.5D polynomiale de degré 2 ne peut décrire qu'une portion de la partie visible d'une surface 3D de deuxième degré (sphère ou cylindre). Notes: Le programme implantant l'algorithme [LeGB90] est disponible sur le serveur public du laboratoire GRASP de l'Université de Pennsylvanie. Les images originales MECA06 (haut) et PYRA02 (bas) proviennent du CNRC. Elles ont été transformées dans des IP cartésiennes orthonormées de taille 128x128x8bits pour être traitées par le programme de segmentation.

"raisonnement géométrique" sur la base de deux autres segmentations: une segmentation préliminaire en surfaces 2.5D de type polynomial et une segmentation courante sur les résidus de l'approximation par le superellipsoïde. Il s'agit donc d'un procédé qui dépasse le cadre des traitements bas-niveau d'une IP.

Leonardis [Leon93] parvient à réaliser directement avec la méthode de croissance indépendante de régions décrite au paragraphe 3.1.7 une segmentation des IP suivant un modèle de superellipsoïde.

3.2.4 Continuité d'ordre 1

Plusieurs méthodes conduisent à une segmentation dont le critère d'homogénéité de surface peut être défini comme la continuité C^1 . Deux principes sont appliqués: (1) détection des discontinuités et étiquetage des composantes connexes ou (2) segmentation suivant un critère d'homogénéité plus contraignant et fusion des segments qui ne sont pas séparés par une discontinuité.

Détection des discontinuités et étiquetage des composantes connexes

Il s'agit du principe de détection de frontière décrit au paragraphe 3.1.2. Les points de frontière sont les points de discontinuité \bar{C}^0 et \bar{C}^1 .

La détection des points de discontinuités est traitée de manière approfondie au chapitre 4. Ici sont uniquement présentés les différents traitements supplémentaires dont le but est de fermer les frontières. Pour l'essentiel, il s'agit d'une croissance de la (des) région(s) constituée(s) des points de discontinuité.

Traitement supplémentaire

Fan et al. [FaMN87] utilisent une technique très rudimentaire qui comble les trous de petite taille dans la frontière.

Pour construire une courbe, si possible fermée, à partir des points de frontière détectés, Ferrie et al. [FeLW90] utilisent une méthode d'approximation par spline impliquant la formulation et la résolution d'un problème de minimisation d'énergie. Il s'agit d'une méthode relativement complexe.

Dans l'approche de Davignon [Davi92], la croissance du domaine des points de frontière détectés est effectuée de sorte à former toujours des frontières fermées. La croissance se fait en choisissant à chaque pas le pixel voisin dont la mesure de discontinuité est la plus proche (en dessous) du seuil (3.3). Une fois les frontières fermées, les régions connexes sont approchées par une surface algébrique (de type polynomial) de complexité croissante, jusqu'à ce que l'erreur d'approximation soit suffisamment petite. La position des discontinuités est ensuite réajustée à partir de cette représentation de surface.

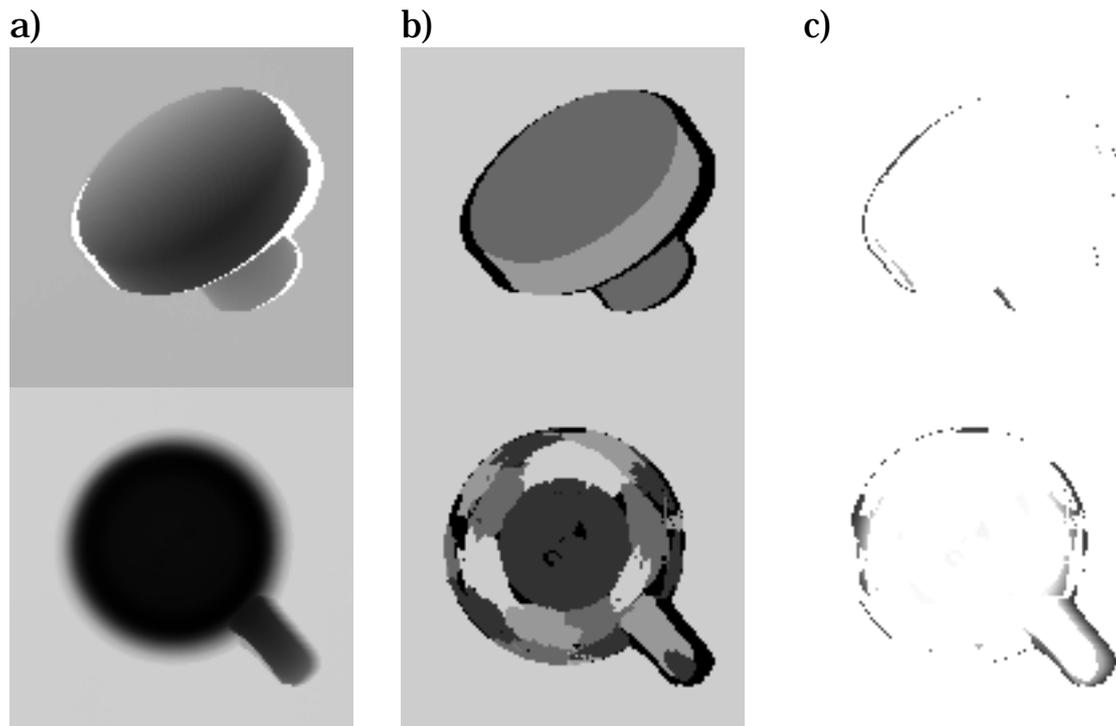


Figure 3.6: Segmentation en surfaces 2.5D polynomiales de degrés 1 à 4 par une méthode de croissance indépendante de régions [BeJa88]. (a) IP. (b) IP segmentée (c) Différence entre IP et surface algébrique. Haut: MECA06. Bas: CUP02. Commentaires: A l'aide d'un polynôme de degré 4 il a été possible de représenter de façon complète la partie visible du cylindre de grand diamètre dans MECA06. Le fond de la tasse dans CUP02 ne peut pas être représenté par une seule surface 2.5D de type polynomial. Le fond de la tasse sera représenté par un seul segment après l'étape de fusion. Notes: Les résultats proviennent de [Tièc90]. Les images originales du CNRC ont été réduites à la taille 128x128 par sous-échantillonnage avant la segmentation.

Fusion de segments issus d'une segmentation plus contrainte

La méthode de Besl [BeJa88, Besl88] commence par une segmentation selon les types de courbure [§3.2.1], les régions qui en sont issues sont ensuite érodées. A partir des graines ainsi obtenues, la méthode procède par croissance, comme décrit au paragraphe 3.1.7. Finalement a lieu une fusion des régions sur la base de l'existence ou non de discontinuité entre les régions.

Discussion

Pour certaines des méthodes que nous venons de décrire, le critère d'homogénéité comporte une condition supplémentaire à la pure

continuité C^1 . Cette condition provient de la fermeture systématique des courbes de discontinuité, alors que celles-ci ne sont pas forcément fermées [§2.3.2]. La fermeture systématique des discontinuités existe soit dans les méthodes procédant par détection des discontinuités, soit dans celles qui procèdent par fusion de segments obéissant à un critère d'homogénéité plus contraignant.

Dans les méthodes de détection des discontinuités de Fan et al. [FaMN87] et de Ferrie et al. [FeLW90], on peut admettre que l'accroissement du domaine des points de discontinuité est destiné à remédier à une détection incomplète des discontinuités. Par contre, dans la méthode de Davignon, la fermeture des discontinuités est systématique et rajoute ainsi une contrainte à la continuité.

Dans la méthode de fusion de segments plus contraints [BeJa88, Besl88], il suffit qu'une partie seulement de la frontière entre segments corresponde à une discontinuité pour que la fusion n'ait pas lieu. Ce procédé a donc un effet de contrainte de fermeture systématique des frontières associées aux discontinuités. Par conséquent, le critère d'homogénéité de la segmentation finale n'est pas uniquement la continuité C^1 .

3.2.5 Autres

Les méthodes décrites ci-après appliquent le même principe de fusion de segments plus contraints que la méthode de Besl. Néanmoins, la fusion repose sur des conditions plus restrictives que l'absence de discontinuités entre segments. La segmentation finale obéit donc à un critère d'homogénéité plus contraignant que la continuité C^1 .

La méthode utilisée par Ylä-Jääski et Ade [YJAd90] comporte plusieurs étapes. La première étape consiste en un "clustering" des vecteurs normaux et à un étiquetage des composantes connexes. Les régions de petites tailles (étroites) sont éliminées à la fin de cette première étape. La deuxième étape effectue une croissance des régions conservées. Cette croissance comporte trois pas : (1) un (seul) pas de croissance avec une contrainte d'approximation par un plan, (2) un (seul) pas avec une contrainte d'approximation par une surface 2.5D polynomiale de second degré, (3) dilatation jusqu'à ce que les régions se touchent. Finalement, a lieu une étape de fusion. Deux régions adjacentes sont fusionnées si leurs normales et leurs

courbures principales ont des valeurs proches et si elles ont une frontière commune suffisamment longue.

Hoffman et Jain [HoJa87] ont une méthode de segmentation constituée de deux étapes: "clustering" et fusion. Le "clustering" a lieu dans un espace à six dimensions. Aux composantes du vecteur normal sont ajoutées les coordonnées cartésiennes du rangel. Les coordonnées x et y imposent la condition de connexité et la coordonnée z contraint à une séparation des segments en présence de discontinuités de profondeur. L'algorithme de "clustering" procède par minimisation des distances intra-"cluster". Par conséquent, il existe un paramètre définissant une borne supérieure pour le nombre de clusters, sinon le résultat serait une segmentation en régions mono-élément. Les segments sont classés en régions planes, convexes et concaves. Sont finalement fusionnées des régions adjacentes appartenant à la même classe pour autant qu'elles ne soient pas séparées par des discontinuités.

3.3 Discussion et conclusion

Les méthodes de segmentation d'image de profondeur ont été passées en revue, d'abord les méthodes élémentaires, puis les méthodes formant l'état de l'art du domaine.

Méthodes élémentaires

Pour les cinq méthodes élémentaires recensées, nous avons relevé l'existence d'un lien étroit entre critère d'homogénéité et méthode. Si deux des cinq méthodes s'appliquent à plusieurs critères d'homogénéité (modèle de surface), les autres sont spécifiques à seul un critère d'homogénéité.

En particulier, nous avons relevé l'existence d'une méthode pour segmenter suivant le critère de continuité C^1 qui consiste à détecter les points de discontinuité de types \bar{C}^0 et \bar{C}^1 et à faire l'étiquetage des composantes connexes de l'image des frontières formées par les discontinuités. Cette méthode a en outre un avantage qu'elle partage avec la méthode pour la segmentation selon le type de courbure, celui d'être des plus simples.

Etat de l'art

La revue de littérature a permis de montrer comment les méthodes élémentaires étaient utilisées en pratique, formant parfois des méthodes combinées. Bien que le critère d'homogénéité de la segmentation résultante ne soit pas toujours explicite, on peut tout de même classer pratiquement toutes les méthodes suivant les quatre critères d'homogénéité: type de courbure, surface plane, surface algébrique, continuité C^1 . Voici une vue d'ensemble des méthodes, accompagnée de commentaires critiques.

type de courbure

Les méthodes sont confrontées au problème d'estimation de la courbure. Celle-ci est perturbée à la fois par le bruit et les discontinuités. La segmentation n'est représentative de la surface mesurée que si le bruit est faible et si les perturbations dues aux discontinuités sont éliminées par un traitement adéquat.

surface plane

Les méthodes sont nombreuses. La plupart utilisent des propriétés particulières à une surface plane. Nous sommes d'avis que la segmentation en surfaces planes devrait être traitée de la même manière que le cas plus général des surfaces algébriques.

surface algébrique

Il y a deux types de méthodes: division/fusion et croissance indépendante de région.

A notre avis les méthodes de division/fusion sont mal adaptées à la segmentation des images de profondeur suivant un modèle de surface algébrique, notamment à cause des discontinuités.

En ce qui concerne les méthodes de croissance indépendante de région, nous voyons deux approches qui se distinguent au niveau de leur principe fondamental. L'approche initiée par Besl cherche une représentation approximative minimale de l'image de profondeur. L'approche de Leonardis fait une recherche exhaustive de la meilleure localisation d'un ou de plusieurs modèles de surface dans l'image.

continuité C^1

A côté de la méthode simple de segmentation suivant C^1 , il existe des méthodes complexes. Tout l'effort investi par ces méthodes est toutefois consacré à forcer la fermeture des frontières définies par les discontinuités. Une fermeture systématique des frontières correspond à une modification du critère d'homogénéité le rendant plus contraignant que la pure continuité C^1 .

Une de ces méthodes procède par fusion de segments de surface algébrique. On note qu'elle a l'avantage de fournir, contrairement aux autres méthodes, une représentation simplifiée du segment.

Dans la suite, nous nous en tenons au critère de continuité C^1 et à la méthode simple constituée par la détection des points de discontinuité et l'étiquetage des composantes connexes de l'image des frontières.

Chapitre 4

Détection des discontinuités

La détection des discontinuités d'une image de profondeur est envisagée dans la perspective d'obtenir des frontières de surfaces. Parmi les différentes méthodes de détection des discontinuités, seulement la plus simple est considérée, celle qui consiste à appliquer une opération de voisinage pour obtenir des points de discontinuité. Plus précisément, la détection a lieu par seuillage d'une grandeur différentielle servant de mesure de discontinuité.

D'abord, les grandeurs différentielles appropriées à constituer des mesures de discontinuité (ordre 0 et 1) sont inventoriées. Puis les méthodes d'estimation de ces grandeurs sont passées en revue.

Ensuite, les méthodes existantes de détection de points de discontinuité sont présentées et analysées. Leurs caractéristiques sont considérées essentiellement en regard de l'utilisation des discontinuités comme frontières de surface.

Dans une dernière partie, sont proposés une série de détecteurs de discontinuités plus directement appropriés au but poursuivi de construction de frontières. Leur comportement est étudié.

4.1 Introduction

4.1.1 Détection des discontinuités

Objet

Les discontinuités (d'ordres 0 et 1) d'une image de profondeur sont définies par rapport à une surface que l'on associe à l'IP en tant que modèle [chap.2].

La détection des discontinuités consiste à décider où introduire des discontinuités dans le modèle de surface. Cette décision se base sur les données de l'IP.

Les figures 4.1 à 4.3 présentent des exemples d'IP auxquelles on souhaite associer un modèle de surface discontinue.

But

La détection des discontinuités d'une IP a souvent pour but d'extraire des primitives 1D liées aux courbes que forment les points de discontinuité.

Dans le présent travail, les discontinuités intéressent dans l'optique de créer des frontières de surface [chap.3].

Méthodes

A priori, différentes méthodes peuvent être envisagées. Le niveau de localité/globalité auquel est effectuée l'analyse du signal pour la détection constitue un critère essentiel de distinction entre les méthodes. Ici, on s'intéresse à des méthodes de détection basées sur une analyse purement locale du signal.

4.1.2 Détecteur de point de discontinuité

Soit une image de profondeur \mathbf{P} .

Un détecteur de point de discontinuité est une opération de voisinage sur \mathbf{P} dont le résultat est la décision concernant la présence d'un point de discontinuité à la position de référence. Formellement, pour la position (i, j) , la décision $b_{i,j}$ résulte de l'opération Φ sur le voisinage $W_{i,j}$

$$b_{i,j} = \Phi \left[\left(k, l, \mathbf{p}_{k,l} \right) : (k, l) \in W_{i,j} \right] \quad (4.1)$$

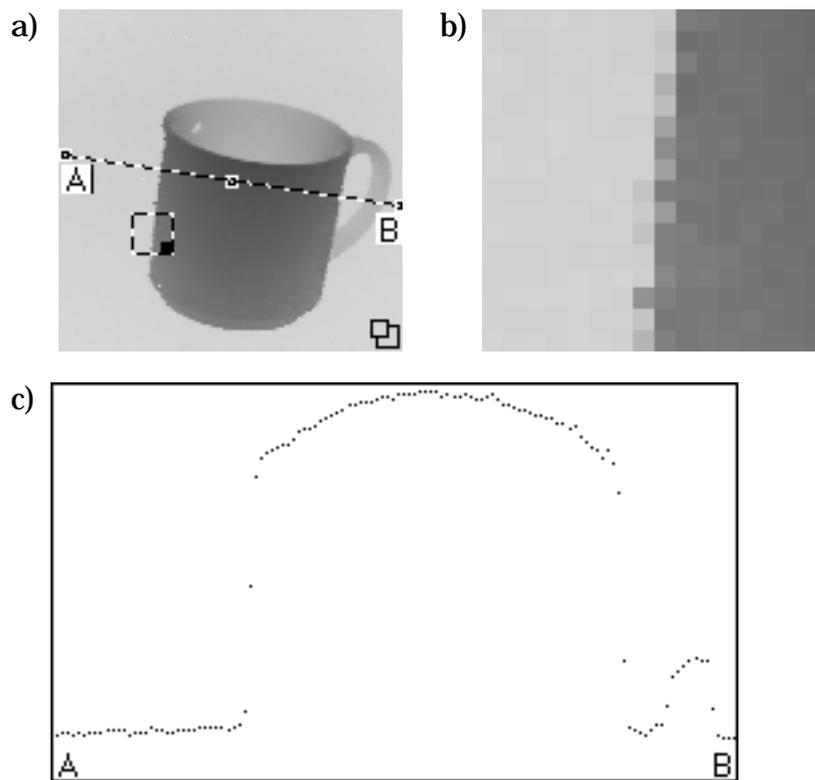


Figure 4.1: Exemple de discontinuité de profondeur. (a) Image de profondeur "Coffee Cup B" codée par niveaux de gris. Deux régions sont sélectionnées: un domaine carré et un profil. (b) Agrandissement de la région carrée. (c) Représentation orthonormée du profil. Note: L'image de profondeur (taille 128x128x8bits) provient d'un système d'acquisition de type "temps de vol" développé par l'ERIM (Environmental Research Institute of Michigan) [Besl88]. Il s'agit d'un capteur relativement ancien; depuis, la qualité de mesure des capteurs a été améliorée [FlJa89].

Un détecteur de point de discontinuité est discret et local:

Discret. La détection se limite à un ensemble fini de points. Ce sont soit les positions (i, j) de l'IP, comme décrit dans la formule (4.1), soit les positions intermédiaires, $(i + \frac{1}{2}, j)$ et $(i, j + \frac{1}{2})$, utilisées fréquemment pour décrire les frontières des régions. (A noter que cette deuxième solution permet d'obtenir une segmentation ayant la propriété de partition.)

Local. La décision est prise de manière indépendante pour chaque voisinage. (On peut par contre imaginer une méthode où la décision serait prise de manière coopérative.)

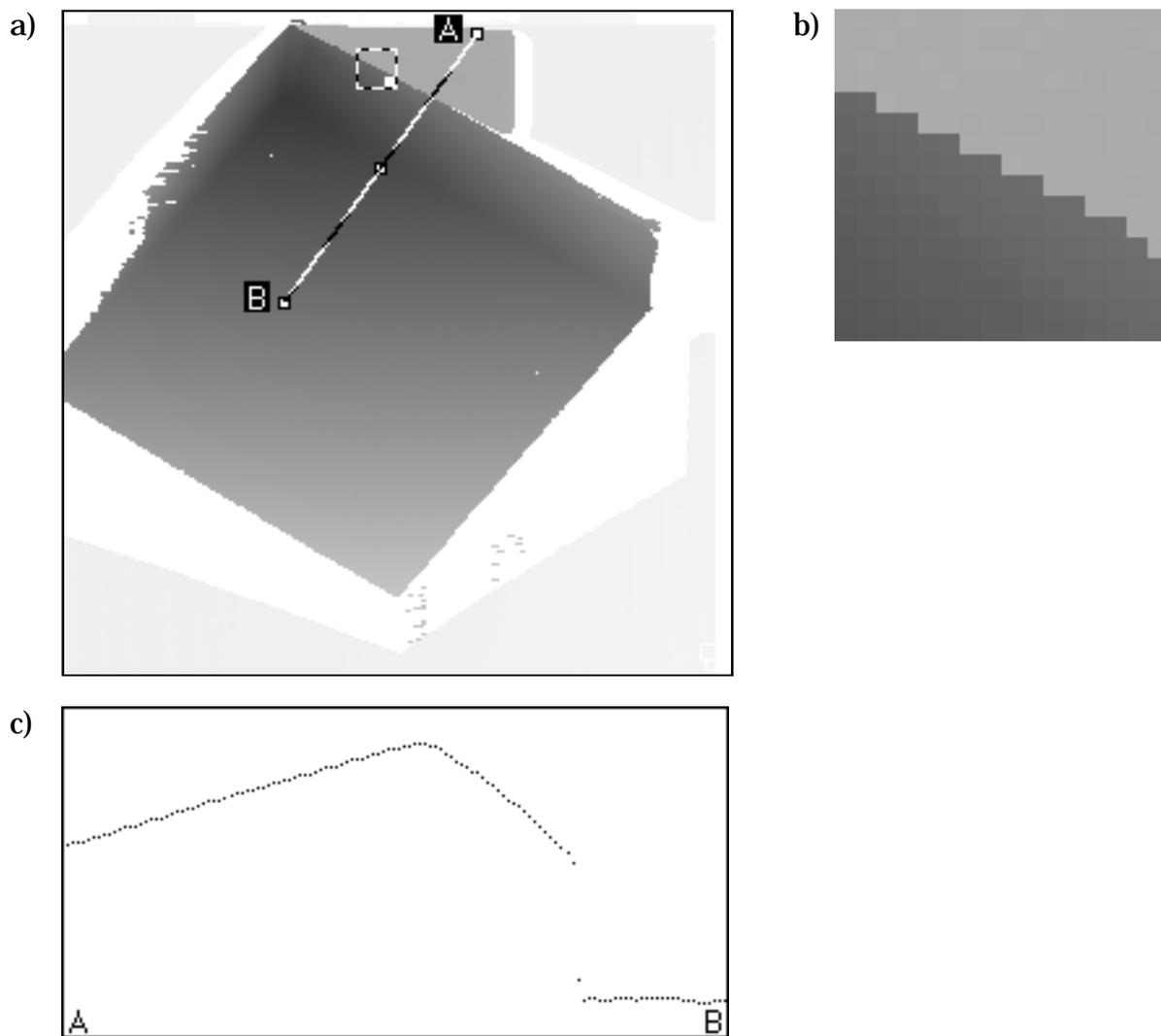


Figure 4.2: Exemples de discontinuités de profondeur et de discontinuités d'orientation. (a) Image de profondeur codée par niveaux de gris. Deux régions sont sélectionnées: un domaine carré et un profil. (b) Agrandissement de la région carrée. (c) Représentation orthonormée du profil. Note: L'image de profondeur BALISTO2ThrL (taille 256x256x8bits) provient du système d'acquisition expérimental de type "ligne laser" décrit au chapitre 2. La scène est constituée d'une boîte en carton inclinée.

4.1.3 Détecteurs différentiels de point de discontinuité

Il s'agit de détecteurs de point de discontinuité particuliers basés sur l'évaluation de grandeurs différentielles de l'IP. Dans ce paragraphe, le cas simple d'une discontinuité de profondeur et d'un signal discret unidimensionnel $[p_i]$ est présenté, il sert d'introduction.

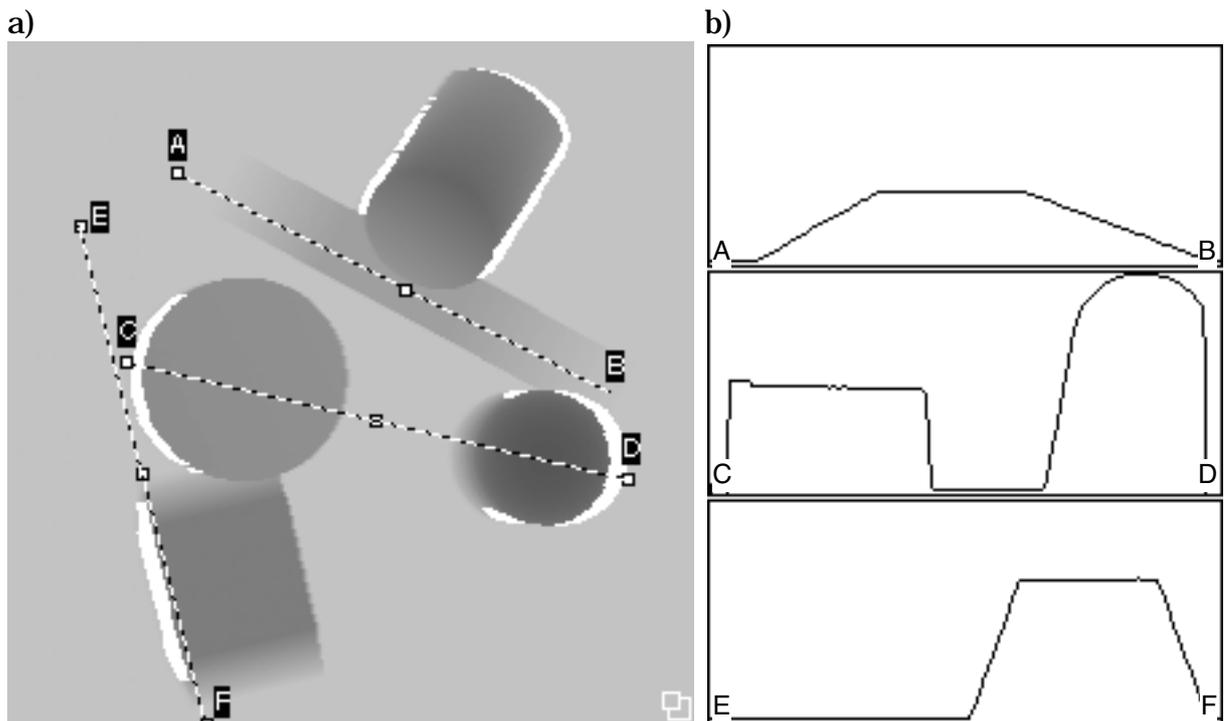


Figure 4.3: Exemples de discontinuités de profondeur et de discontinuités d'orientation. (a) Codage par niveaux de gris. Trois profils sont sélectionnés. (b) Représentation orthonormée des profils sélectionnés. Note: L'image de profondeur BLOC12 (taille 256x256x16bits) provient du CNRC (Conseil National de Recherches du Canada) [RiCo88].

Détecteur simple (sans condition d'extremum)

La figure 4.4 présente un exemple de signal auquel on associe le modèle de discontinuité.

De manière évidente, les positions proches de la discontinuité se distinguent des autres par une valeur plus élevée de la pente du signal. Un seuillage sur la valeur de la pente permet de mettre en évidence les positions proches de la discontinuité par rapport aux autres et réalise donc un détecteur simple de discontinuité.

Formellement, on attribue au signal discret les propriétés d'un signal continu dont on estime la dérivée. La détection des discontinuités se fait par seuillage sur la valeur absolue de la dérivée

$$b_i = (|p'_i| > T) \quad (4.2)$$

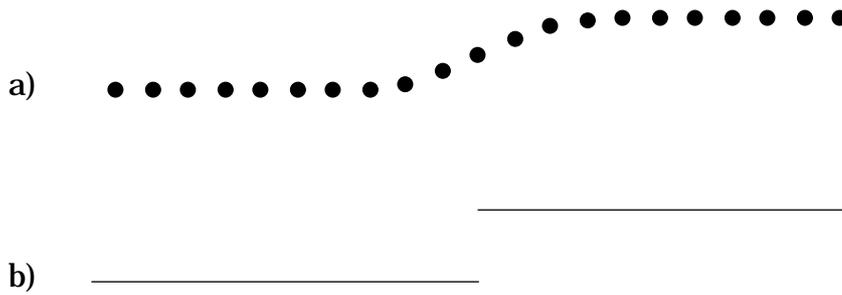


Figure 4.4: Signal et modèle de discontinuité de profondeur associé. (a) signal, (b) modèle.

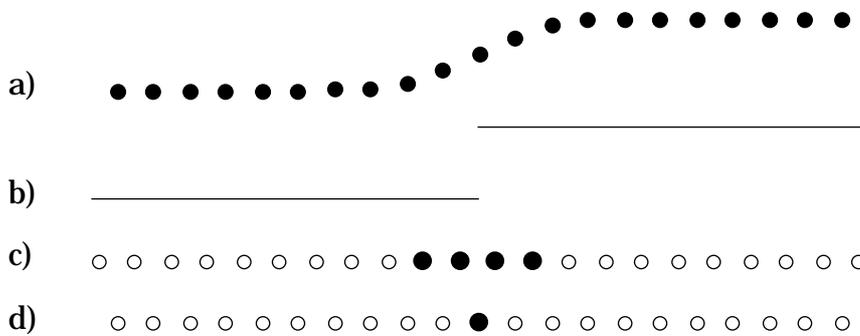


Figure 4.5: Détection et condition d'extremum. (a) signal, (b) modèle de discontinuité, (c) résultat de détection sans condition d'extremum (points de discontinuité en noir), (d) résultat de détection avec condition d'extremum (idem).

Terminologie

Nous disons qu'une pente élevée est l'indice d'une discontinuité et que la valeur absolue de la dérivée est une mesure de discontinuité notée Δ .

Détecteur avec condition d'extremum

En général, le détecteur différentiel simple de discontinuité (4.2) marque plusieurs positions contiguës pour une seule discontinuité [figure 4.5].

L'utilisation de la condition d'extremum, exprimée ici par le passage à zéro de la deuxième dérivée

$$b_i = (|p'_i| > T) \wedge ((p'_i)' = 0) \quad (4.3)$$

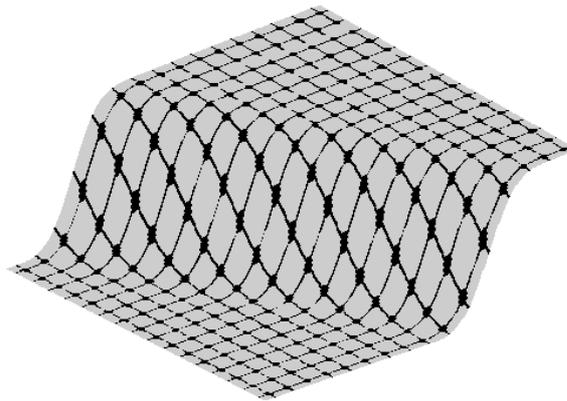


Figure 4.6: Signal auquel on associe un modèle de discontinuité de profondeur.

permet d'obtenir qu'une seule position soit marquée comme discontinuité (amincissement à 1 pixel).

On distingue donc les détecteurs avec et sans condition d'extremum.

Autre distinction

Pour une même grandeur différentielle, ici la dérivée, il existe plusieurs détecteurs de discontinuité qui se distinguent dans la manière avec laquelle la grandeur différentielle est estimée.

4.2 Grandeurs différentielles utiles à la détection des points de discontinuité

Nous présentons ici les grandeurs différentielles appropriées à constituer une mesure de discontinuité (ordre 0 ou 1) [§4.2.1] et traitons de leur estimation [§4.2.2].

4.2.1 Mesures de discontinuité

Nous attribuons à une IP les propriétés d'une surface continûment dérivable $\mathbf{p}(u, v)$ (ou $p(u, v)$ pour une image scalaire).

Discontinuité de profondeur

Une forte inclinaison de la surface $\mathbf{p}(u, v)$ par rapport à la direction de mesure est l'indice d'une discontinuité de profondeur. Dans le cas d'une IP scalaire, ceci correspond à une forte pente de la surface

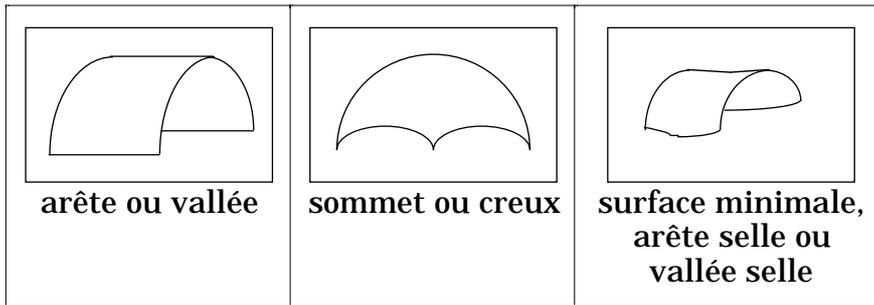


Figure 4.7: Signaux auxquels on associe un modèle de discontinuité d'orientation.

[figure 4.6]. Deux mesures de discontinuité de profondeur s'imposent: d'une part, le produit scalaire de la normale à la surface avec la direction de mesure et, d'autre part la norme ou le carré de la norme du gradient. La première exige de connaître la direction de mesure. La deuxième s'applique seulement à une IP scalaire [table 4.1].

Discontinuité d'orientation

Une valeur élevée de la courbure de la surface $\mathbf{p}(u, v)$ dans au moins une direction est l'indice d'une discontinuité d'orientation [figure 4.7]. Les mesures de discontinuité appropriées sont donc [table 4.1]:

- la courbure extrême en valeur absolue: maximum des valeurs absolues de la courbure minimale κ_m et de la courbure maximale κ_M
- la somme des carrés des courbures minimale κ_m et maximale κ_M . Exprimée à l'aide de la courbure moyenne H et de la courbure gaussienne K , cette grandeur vaut

$$\kappa_m^2 + \kappa_M^2 = 4H^2 - 2K \quad (4.4)$$

Pour une IP scalaire, une valeur élevée de la deuxième dérivée dans au moins une direction constitue aussi un indice pour une discontinuité d'orientation. Dans ce cas, il y a donc deux mesures de discontinuité supplémentaires:

- la deuxième dérivée extrême en valeur absolue: maximum des valeurs absolues de la deuxième dérivée minimale p''_m et de la deuxième dérivée maximale p''_M

	Δ^{C^0}	Δ^{C^1}
IP scalaires	$\ \mathbf{grad} p\ $ $p_u^2 + p_v^2$	$\max(p'_m , p'_M)$ $Q_{p'} = p_{uu}^2 + 2p_{uv}^2 + p_{vv}^2$
IP cartésiennes	$1 - \mathbf{n} \cdot \mathbf{e}^s$	$\max(\kappa_m , \kappa_M)$ $Q_\kappa = 2H^2 - K$

Table 4.1: Mesures de discontinuité dans un modèle continu de l'IP. Remarque: \mathbf{e}^s est un vecteur unitaire de \mathfrak{R}^3 représentant la direction du dispositif de mesure vu du point de mesure.

- la somme des carrés des deuxièmes dérivées minimale et maximale qui est égale à la variation quadratique

$$p_m'^2 + p_M'^2 = p_{uu}^2 + 2p_{uv}^2 + p_{vv}^2 \quad (4.5)$$

Discontinuités de profils

Pour une IP scalaire, les directions des axes de coordonnées (u, v) ainsi que les deux directions diagonales définissent en chaque position quatre profils plans de la surface $p(u, v)$. Les grandeurs différentielles de ces profils (dérivées d'ordre 1 et 2, courbure) constituent aussi des mesures de discontinuité appropriées pour la surface [table 4.2]. On remarque que, dans le cas des dérivées, les valeurs coïncident avec les valeurs des dérivées de $p(u, v)$ dans quatre directions séparées d'un angle $\pi/4$

$$p_n^{(m)} = p_{n\pi/4}^{(m)} \quad (4.6)$$

Ce n'est pas le cas pour la courbure, contrairement à ce qui est prétendu dans [FaMN87], puisque les directions doivent se mesurer dans le plan tangent à la surface. (Le lissage sera traité au chapitre 6.)

4.2.2 Estimation des grandeurs différentielles

Les grandeurs différentielles d'une image de profondeur sont estimées par des opérations de voisinage. Les opérateurs différentiels sont sensibles au bruit. La robustesse vis-à-vis du bruit s'obtient par

	Δ^{C^0}	Δ^{C^1}
IP scalaires	$\max_{n \in \{0,1,2,3\}} (p'_n)$	$\hat{p}'_n = \max_{n \in \{0,1,2,3\}} (p''_n)$ $\hat{k}'_n = \max_{n \in \{0,1,2,3\}} (k_n)$

Table 4.2: Mesures de discontinuité basées sur les profils dans les quatre directions discrètes. Elles s'appliquent seulement à des images de profondeur de type scalaire.

augmentation de la taille du voisinage. Celle-ci peut être effectuée soit de manière directe, soit de manière indirecte par une opération de lissage préalable. L'augmentation de la taille du voisinage a le désavantage d'une perte de résolution.

Notre approche consiste à séparer le lissage et l'estimation des propriétés différentielles. Pour cette raison, nous nous intéressons ici à des opérateurs différentiels de taille minimale.

Estimation des dérivées partielles (IP scalaire)

L'estimation de chacune des dérivées partielles d'une IP est réalisée par corrélation avec un masque approprié.

Les masques conventionnels pour l'estimation des dérivées partielles d'ordre 1 sont soit

$$\mathbf{D}_u^A = \frac{1}{6} \begin{bmatrix} -1 & -1 & -1 \\ & \square & \\ 1 & 1 & 1 \end{bmatrix}, \quad \mathbf{D}_v^A = \frac{1}{6} \begin{bmatrix} -1 & & 1 \\ -1 & \square & 1 \\ -1 & & 1 \end{bmatrix} \quad (4.7.a-b)$$

soit

$$\mathbf{D}_u^B = \frac{1}{8} \begin{bmatrix} -1 & -2 & -1 \\ & \square & \\ 1 & 2 & 1 \end{bmatrix}, \quad \mathbf{D}_v^B = \frac{1}{8} \begin{bmatrix} -1 & & 1 \\ -2 & \square & 2 \\ -1 & & 1 \end{bmatrix} \quad (4.8.a-b)$$

Les premiers sont connus en traitement d'image sous le nom d'opérateur de Prewitt, les derniers sous le nom d'opérateurs de Sobel [HaSh92, Kunt93].

Pour les dérivées partielles de deuxième ordre p_{uu} et p_{vv} , on utilise les masques

$$\mathbf{D}_{uu}^A = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ -2 & \boxed{-2} & -2 \\ 1 & 1 & 1 \end{bmatrix}, \quad \mathbf{D}_{vv}^A = \frac{1}{3} \begin{bmatrix} 1 & -2 & 1 \\ 1 & \boxed{-2} & 1 \\ 1 & -2 & 1 \end{bmatrix} \quad (4.9.a-b)$$

ou bien

$$\mathbf{D}_{uu}^B = \frac{1}{4} \begin{bmatrix} 1 & 2 & 1 \\ -2 & \boxed{-4} & -2 \\ 1 & 2 & 1 \end{bmatrix}, \quad \mathbf{D}_{vv}^B = \frac{1}{4} \begin{bmatrix} 1 & -2 & 1 \\ 2 & \boxed{-4} & 2 \\ 1 & -2 & 1 \end{bmatrix} \quad (4.10.a-b)$$

La dérivée partielle croisée p_{uv} est estimée par corrélation avec le masque

$$\mathbf{D}_{uv} = \frac{1}{4} \begin{bmatrix} 1 & & -1 \\ & \boxed{} & \\ -1 & & 1 \end{bmatrix} \quad (4.11)$$

On remarque que, simultanément à l'opération de dérivation, les masques conventionnels (4.7) à (4.10) réalisent une opération de lissage perpendiculairement à la direction de dérivation. Pour effectuer uniquement l'opération de dérivation, il est possible de réduire les masques de dérivation à leur forme minimale, soit

$$\mathbf{D}_u = \frac{1}{2} \begin{bmatrix} -1 \\ \boxed{} \\ 1 \end{bmatrix}, \quad \mathbf{D}_v = \frac{1}{2} \begin{bmatrix} -1 & \boxed{} & 1 \end{bmatrix} \quad (4.12.a-b)$$

pour les dérivées partielles d'ordre 1 et

$$\mathbf{D}_{uu} = \begin{bmatrix} 1 \\ \boxed{-2} \\ 1 \end{bmatrix}, \quad \mathbf{D}_{vv} = \begin{bmatrix} 1 & \boxed{-2} & 1 \end{bmatrix} \quad (4.13.a-b)$$

pour les dérivées p_{uu} et p_{vv} .

Estimation des dérivées dans les 4 directions discrètes (IP scalaire)

Au lieu d'être calculées à partir des dérivées partielles en utilisant les formules pour les dérivées directionnelles, les dérivées dans les quatre directions discrètes peuvent être estimées simplement par corrélation avec les masques appropriés.

Les masques symétriques de taille minimale pour les dérivées d'ordre 1 sont

$$\begin{aligned} \mathbf{D}_0^1 &= \frac{1}{2} \begin{bmatrix} -1 \\ \square \\ 1 \end{bmatrix} & \mathbf{D}_1^1 &= \frac{1}{2\sqrt{2}} \begin{bmatrix} -1 & & \\ & \square & \\ & & 1 \end{bmatrix} \\ \mathbf{D}_2^1 &= \frac{1}{2} \begin{bmatrix} -1 & \square & 1 \end{bmatrix} & \mathbf{D}_3^1 &= \frac{1}{2\sqrt{2}} \begin{bmatrix} & & 1 \\ & \square & \\ -1 & & \end{bmatrix} \end{aligned} \quad (4.14.a-d)$$

Les masques non symétriques

$$\begin{aligned} \mathbf{D}_{0-}^1 &= \begin{bmatrix} 1 \\ -1 \end{bmatrix} & \mathbf{D}_{0+}^1 &= \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ \mathbf{D}_{1-}^1 &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & \\ & \square \\ & -1 \end{bmatrix} & \mathbf{D}_{1+}^1 &= \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & \\ & \square \\ & 1 \end{bmatrix} \\ \mathbf{D}_{2-}^1 &= \begin{bmatrix} 1 & \square \\ -1 & \end{bmatrix} & \mathbf{D}_{2+}^1 &= \begin{bmatrix} -1 & 1 \\ \square & \end{bmatrix} \\ \mathbf{D}_{3-}^1 &= \frac{1}{\sqrt{2}} \begin{bmatrix} & \square \\ 1 & -1 \end{bmatrix} & \mathbf{D}_{3+}^1 &= \frac{1}{\sqrt{2}} \begin{bmatrix} & 1 \\ -1 & \square \end{bmatrix} \end{aligned} \quad (4.15.a-h)$$

estiment les dérivées à gauche (-) et à droite (+) de la position de référence dans les directions respectives.

Pour les dérivées de deuxième ordre, les masques de taille minimale sont

$$\begin{aligned} \mathbf{D}_0^2 &= \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} & \mathbf{D}_1^2 &= \frac{1}{2} \begin{bmatrix} 1 & & \\ & \square & \\ & & 1 \end{bmatrix} \\ \mathbf{D}_2^2 &= \begin{bmatrix} 1 & \square & 1 \\ -2 & & \end{bmatrix} & \mathbf{D}_3^2 &= \frac{1}{2} \begin{bmatrix} & & 1 \\ & \square & \\ 1 & & \end{bmatrix} \end{aligned} \quad (4.16.a-d)$$

Calcul du gradient et des dérivées deuxièmes minimale et maximale (IP scalaire)

Le calcul du gradient à partir des dérivées partielles p_u et p_v est immédiat. Sa norme vaut

$$\|\mathbf{grad} p\| = \sqrt{p_u^2 + p_v^2} \quad (4.17)$$

sa direction φ_{grad} par rapport au vecteur unitaire \mathbf{e}_u s'obtient à l'aide des équations

$$\begin{aligned} \cos(\varphi_{grad}) &= p_u / \|\mathbf{grad} p\| \\ \sin(\varphi_{grad}) &= p_v / \|\mathbf{grad} p\| \end{aligned} \quad (4.18)$$

Les dérivées directionnelles d'ordre 2 minimale et maximale se calculent à partir des dérivées partielles p_{uu} , p_{vv} et p_{uv} par la mise en application de la relation

$$p''_{M,m} = \frac{1}{2} (p_{uu} + p_{vv} \pm \Delta_{mM}) \quad (4.19)$$

où

$$\Delta_{mM} = \sqrt{(p_{uu} - p_{vv})^2 + 4 p_{uv}^2} \quad (4.20)$$

Leurs directions respectives φ_m et φ_M sont orthogonales. Elles s'obtiennent à partir des dérivées partielles d'ordre 2 à l'aide des équations

$$\begin{aligned} \cos(2\varphi_{m,M}) &= \pm (p_{uu} - p_{vv}) / \Delta_{mM} \\ \sin(2\varphi_{m,M}) &= \pm (-2 p_{uv}) / \Delta_{mM} \end{aligned} \quad (4.21)$$

le signe + correspondant à la valeur minimale.

Calcul de la normale et des courbures (IP cartésienne orthographique)

L'estimation de la normale et des courbures pour une IP cartésienne orthographique peut être réalisée par calcul à partir des dérivées partielles. La méthode de calcul est présentée dans la figure 4.8.

La courbure des profils correspondant aux directions horizontale, verticale et diagonales du tableau d'IP se calcule à partir des dérivées directionnelles p'_n et p''_n par les formules

$$\begin{aligned} z_0'^2 &= p'_0 / \Delta x^2 & z_0'' &= p''_0 / \Delta x^2 \\ z_1'^2 &= p'_1 / (\Delta x^2 + \Delta y^2) & z_1'' &= p''_1 / (\Delta x^2 + \Delta y^2) \\ z_2'^2 &= p'_2 / \Delta y^2 & z_2'' &= p''_2 / \Delta y^2 \\ z_3'^2 &= p'_3 / (\Delta x^2 + \Delta y^2) & z_3'' &= p''_3 / (\Delta x^2 + \Delta y^2) \end{aligned} \quad (4.22.a-h)$$

et

$$\kappa_n = \frac{z_n''}{(1 + z_n'^2)^{3/2}} \quad (4.23)$$

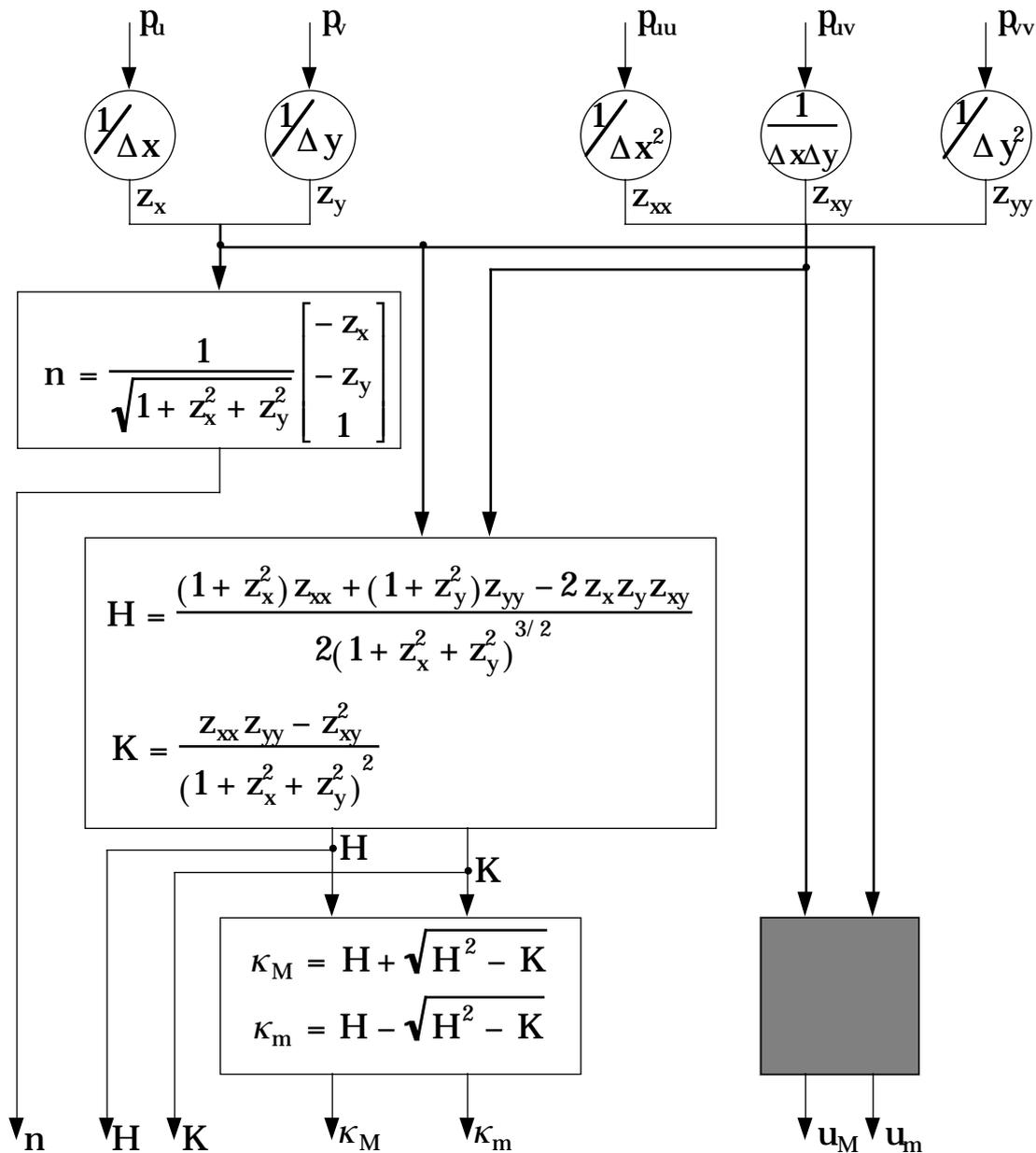


Figure 4.8: Calcul de la normale et des courbures d'une IP cartésienne orthographique. \mathbf{u}_m et \mathbf{u}_M sont les directions des courbures principales κ_m et κ_M dans le plan (u, v) . Les formules servant à leur calcul peuvent être trouvées par exemple dans [Besl90: (3.51-52) et (3.75-78)].

Estimation de la normale et des courbures (IP vectorielle)

L'estimation de la normale pour une IP vectorielle est réalisée par une méthode d'approximation de plan sur un voisinage de la position considérée. La valeur de la normale est déterminée par le plan qui approche au mieux les rangels du voisinage, au sens des moindres

carrés de la distance des points au plan. Du point de vue du calcul, la normale est donnée par le vecteur propre correspondant à la plus petite valeur propre d'une matrice 3x3. La matrice est construite à partir des valeurs des rangels [FaHe87, Besl90].

Pour l'estimation des courbures, trois méthodes sont possibles.

- La première méthode est basée sur une triangulation de la surface [Besl88: Annexe B]. La courbure gaussienne et d'autres caractéristiques intrinsèques de la surface en un point sont estimées par calcul à partir de la longueur des côtés des triangles ayant ce point pour sommet. Cette méthode n'utilise que les voisins immédiats de la position considérée.
- La deuxième méthode est basée sur une estimation de la courbure le long d'une courbe définie à partir de deux points \mathbf{p} et \mathbf{q} [HoJa87, FlJa89, Besl90]. La valeur de courbure est estimée par le quotient dont elle est la limite: l'angle entre les normales à la surface en ces points, divisé par la distance entre les points. La valeur de l'angle entre les normales est approchée par la norme de la différence des normales, d'où la formule

$$K_{p,q} = \begin{cases} \frac{\|\mathbf{n}_p - \mathbf{n}_q\|}{\|\mathbf{p} - \mathbf{q}\|}, & \text{si } \|\mathbf{p} - \mathbf{q}\| \leq \|(\mathbf{n}_p + \mathbf{p}) - (\mathbf{n}_q + \mathbf{q})\| \\ -\frac{\|\mathbf{n}_p - \mathbf{n}_q\|}{\|\mathbf{p} - \mathbf{q}\|}, & \text{sinon} \end{cases} \quad (4.24)$$

Un voisinage de $n+1$ points définit n valeurs de courbure en un point. Les valeurs extrêmes sont une estimation des courbures minimale et maximale en ce point. On peut en tirer une estimation des courbures moyenne et gaussienne en appliquant les formules correspondantes. L'éloignement des points tend à surestimer la courbure puisque la distance euclidienne $\|\mathbf{p} - \mathbf{q}\|$ sous-estime la distance entre les deux points le long de la surface. Pour cette raison, le voisinage doit rester de taille restreinte.

- La troisième méthode utilise un parabolôïde d'approximation dans un système de coordonnées locales. L'estimation de la normale fixe le vecteur $\mathbf{e}_{z'}$ du système de coordonnées locales $(\mathbf{o}, \mathbf{e}_{x'}, \mathbf{e}_{y'}, \mathbf{e}_{z'})$. Les coordonnées des rangels faisant partie d'un voisinage choisi sont calculées dans ce nouveau système. L'évaluation des paramètres du parabolôïde

$$z' = \frac{1}{2} (ax'^2 + 2bx'y' + cy'^2) \quad (4.25)$$

de meilleure approximation permet de déterminer les courbures minimale et maximale et leur direction respective dans le repère local. Ces dernières sont finalement transformées dans le repère global.

Les formules qui ont été présentées dans ce paragraphe permettent le calcul des grandeurs différentielles qui sont utilisées pour la détection des points de discontinuité [tables 4.1 et 4.2].

4.3 Revue des détecteurs de point de discontinuité.

Les détecteurs de point de discontinuité décrits dans la littérature sont de type différentiel. Nous les divisons en deux catégories, suivant s'ils appliquent ou non la condition d'extremum.

4.3.1 Détecteurs sans condition d'extremum

Chen et Kak [ChKa88, ChKa89] détectent les discontinuités de profondeur sur la base de la distance euclidienne entre rangels voisins

$$\Delta_{i+\frac{1}{2},j}^{(0)} = \|\mathbf{p}_{i+1,j} - \mathbf{p}_{i,j}\|, \quad \Delta_{i,j+\frac{1}{2}}^{(0)} = \|\mathbf{p}_{i,j+1} - \mathbf{p}_{i,j}\| \quad (4.26)$$

La mesure servant à la détection des discontinuités d'orientation est une estimation de la courbure de la surface le long d'une courbe reliant deux rangels voisins. La courbure est estimée par le quotient dont elle est la limite: variation d'angle de la normale à la surface divisée par le chemin parcouru

$$\Delta_{i+\frac{1}{2},j}^{(1)} = \frac{\cos^{-1}(\mathbf{n}_{i+1,j} \cdot \mathbf{n}_{i,j})}{\|\mathbf{p}_{i+1,j} - \mathbf{p}_{i,j}\|}, \quad \Delta_{i,j+\frac{1}{2}}^{(1)} = \frac{\cos^{-1}(\mathbf{n}_{i,j+1} \cdot \mathbf{n}_{i,j})}{\|\mathbf{p}_{i,j+1} - \mathbf{p}_{i,j}\|} \quad (4.27)$$

Pour l'estimation des vecteurs normaux, une technique de voisinage adapté est utilisée: la normale est celle du plan de meilleure approximation parmi les approximations sur les N^2 voisinages de taille $N \times N$ auxquels appartient la position (i, j) . La taille de voisinage utilisée dans [ChKa88] est 5×5 .

Les mesures de discontinuité définies par Chen et Kak s'appliquent avant tout à des images de type vectoriel. Pour des

images de type scalaire, on peut remplacer la mesure de discontinuité (4.26) par une forme simplifiée

$$\Delta_{i+\frac{1}{2},j}^{(0)} = |p_{i+1,j} - p_{i,j}|, \quad \Delta_{i,j+\frac{1}{2}}^{(0)} = |p_{i,j+1} - p_{i,j}| \quad (4.28)$$

Une simplification similaire peut être faite pour la mesure de discontinuité d'orientation (4.27).

On remarque que les différentielles servant à la détection des discontinuités ne sont estimées que suivant deux directions (axes de l'image de profondeur).

4.3.2 Détecteurs avec condition d'extremum

Davignon [Davi92] travaille avec des IP de type scalaire. Pour la détection des discontinuités de profondeur, la norme du gradient sert de mesure de discontinuité. L'amincissement de la région de discontinuité se fait en cherchant le maximum dans la direction du gradient, d'où la condition pour les points de discontinuité:

$$\left((\|\mathbf{grad} p\| > T) \wedge \left((\|\mathbf{grad} p\|)'_{\varphi_{grad}} = 0 \right) \right) \quad (4.29)$$

Les dérivées deuxièmes minimale et maximale sont utilisées comme mesure de discontinuité d'ordre 1. La condition d'extremum de la dérivée minimale et de la dérivée maximale dans leur direction respective sert à l'amincissement, d'où la formule

$$\left\{ (|p'_m| > T) \wedge \left((p''_m)'_{\varphi_m} = 0 \right) \right\} \vee \left\{ (|p'_M| > T) \wedge \left((p''_M)'_{\varphi_M} = 0 \right) \right\} \quad (4.30)$$

Fan et al. [FaMN87] traitent des images de profondeur de type scalaire. Deux méthodes de détection des discontinuités sont proposées. Toutes deux utilisent la courbure.

Dans la première méthode, la détection des discontinuités se fait sur la base de la courbure extrême κ_e (c'est-à-dire celle d'entre les courbures minimale et maximale qui est la plus grande en valeur absolue). Les points correspondant à un passage par zéro de la courbure extrême

$$\kappa_e = 0 \quad (4.31)$$

sont retenus comme points de discontinuité de profondeur. Il s'agit d'une condition similaire à celle qui impose un extremum du gradient. Elle est utile à l'amincissement. Cependant, il manque une autre condition (similaire au seuillage sur la norme du gradient) permettant d'éliminer les passages par zéro non significatifs. Bien

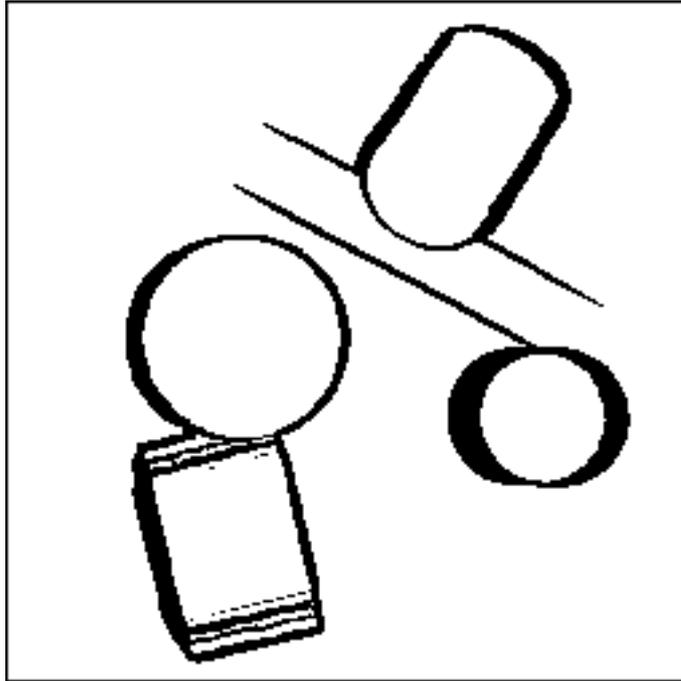


Figure 4.9: Détection des discontinuités de profondeur par seuillage sur la norme du gradient. Image BLOC12 [figure 4.3]. Les discontinuités et les valeurs de rangels NIL sont représentées en noir.

que l'utilisation d'une telle condition soit mentionnée par l'auteur, il n'est nulle part indiqué laquelle est utilisée.

La valeur absolue de la courbure extrême sert de mesure pour les discontinuités d'orientation. La condition d'extremum est utilisée pour réaliser l'amincissement, d'où la formule

$$\left(|\kappa_e| > T\right) \wedge \left(\left(\kappa_e\right)'_{\varphi_e} = 0\right) \quad (4.32)$$

La deuxième méthode est similaire à la première. Mais la courbure extrême est remplacée par une combinaison des courbures de profils dans les quatre directions discrètes correspondant aux axes et diagonales du tableau. Les points de discontinuité de profondeur sont déterminés par le passage à zéro de la courbure de l'un des profils. Formellement

$$\bigvee_{n \in \{0,1,2,3\}} \kappa_n = 0 \quad (4.33)$$

Les points de discontinuité d'orientation sont déterminés par le passage à une valeur extrême de la courbure de l'un des profils, à condition que cette valeur soit suffisamment élevée. Formellement

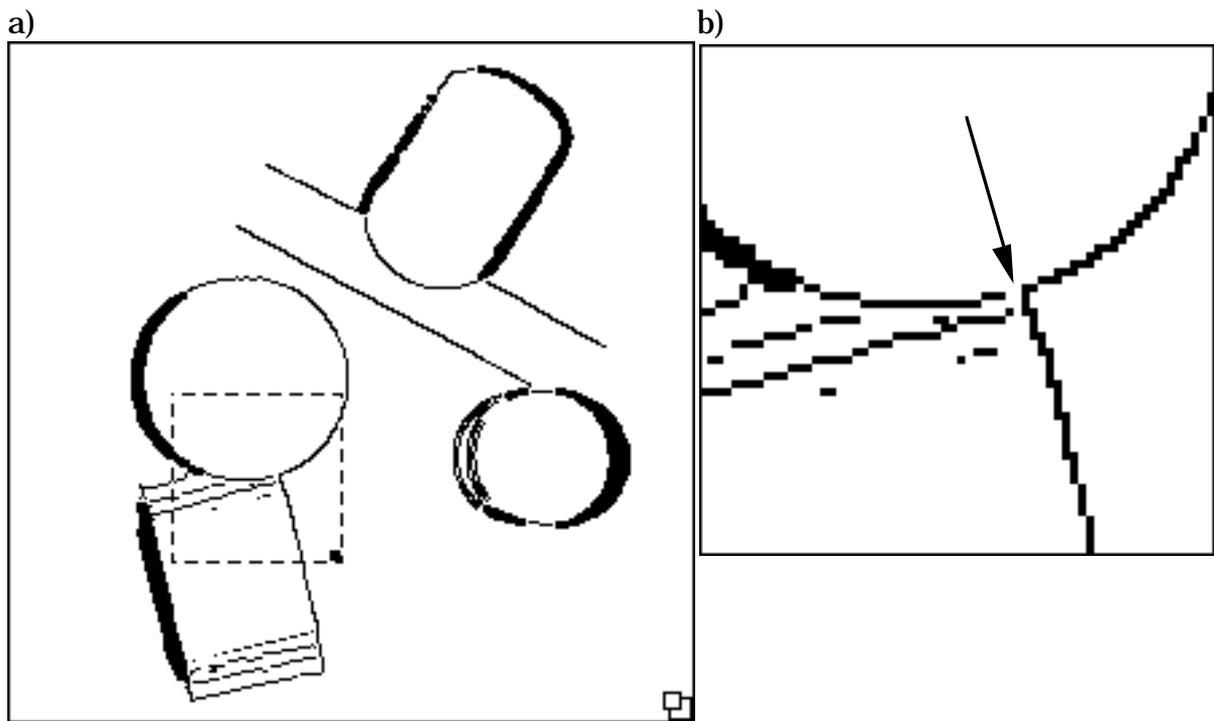


Figure 4.10: Disparition de connexité lors de l'amincissement d'une discontinuité de profondeur par condition d'extremum du gradient. (a) Image des points de discontinuité. Une région est sélectionnée. (b) Représentation agrandie de la région sélectionnée. Mise en évidence de la jonction en T. (Image de profondeur BLOC12 [figure 4.3]. Discontinuités et valeurs *NIL* sont représentées en noir.)

$$\bigvee_{n \in \{0,1,2,3\}} \left((|\kappa_n| > T) \wedge ((\kappa_n)'_n = 0) \right) \quad (4.34)$$

Remarque: Les auteurs des travaux présentés dans ce paragraphe ne font aucune mention des opérateurs utilisés pour estimer les mesures de discontinuité.

4.3.3 Discussion

Les détecteurs de discontinuité qui viennent d'être présentés sont utilisés par les auteurs mentionnés pour construire des frontières de surfaces. Nous discutons ici l'utilité de ces détecteurs en regard du but poursuivi.

Tout d'abord, on constate qu'une détection séparée des discontinuités d'ordre différent est effectuée, alors que la connaissance de l'ordre de discontinuité n'est pas nécessaire dans le but poursuivi.

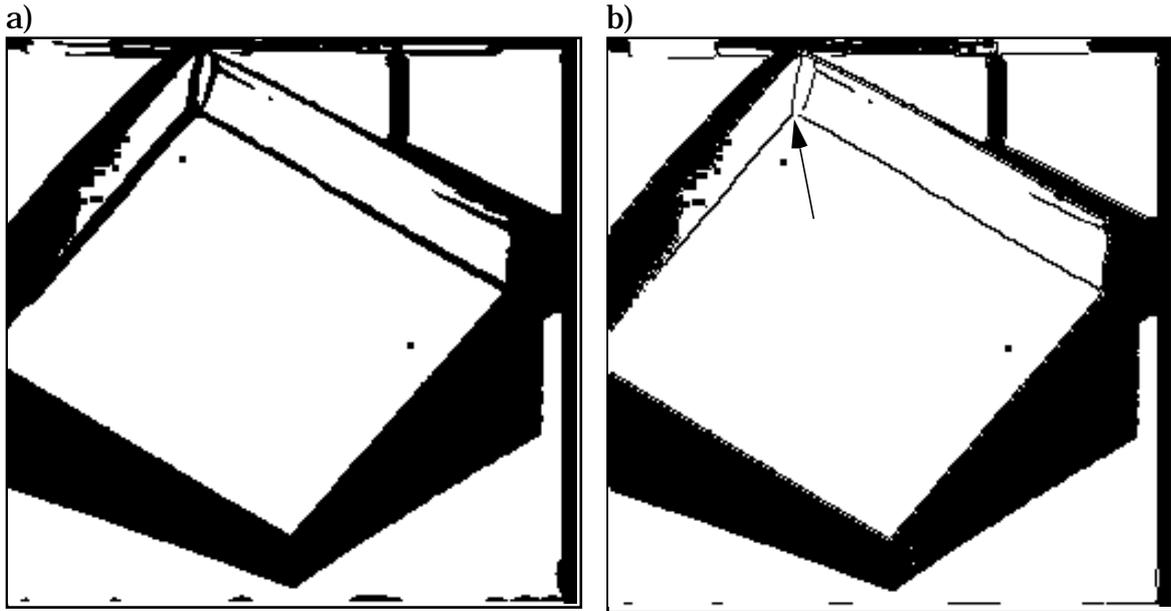


Figure 4.11: Disparition de connexité lors de l'amincissement d'une discontinuité d'orientation par condition d'extremum directionnel. Images des points de discontinuité d'ordre 1 obtenues (a) par simple seuillage sur la deuxième dérivée extrême en valeur absolue, (b) avec amincissement par condition d'extremum directionnel. Un cas de disparition de connexité est indiqué par la flèche. (Image de profondeur BALISTO2ThrL [figure 4.2]. Discontinuités et valeurs *NIL* sont représentées en noir.) Note: Avant de procéder à la détection des discontinuités, l'image originale a été lissée à l'aide de notre méthode adaptative présentée au chapitre 6.

Ensuite, les détecteurs avec condition d'extremum réalisent un amincissement qui n'est pas non plus nécessaire. De plus, la condition d'extremum génère des interruptions de frontières. La figure 4.10 montre cet effet dans le cas de la détection des discontinuités de profondeur avec un amincissement par condition d'extremum du gradient: il peut arriver que la connexité des courbes soit perdue en présence d'une jonction en T. (Il s'agit du même problème qui est rencontré dans la détection des contours d'une image d'intensité par passage à zéro du laplacien [Over92: pp.62-63]). La comparaison des figures 4.11.a et 4.11.b met en évidence le phénomène de disparition de jonction lors de la détection des discontinuités \bar{C}^1 par la condition (4.30).

En conclusion, la détection séparées des discontinuités \bar{C}^0 et \bar{C}^1 est inutile et les détecteurs avec condition d'extremum fonctionnent mal.

4.4 Détection conjointe des discontinuités à l'aide d'un unique détecteur

Ce paragraphe présente des détecteurs de discontinuités minimaux en vue de l'utilisation des discontinuités comme frontières de surface.

La continuité d'ordre 0 est une condition nécessaire à la continuité d'ordre 1. Par conséquent, les mesures de discontinuité d'orientation devraient aussi s'appliquer aux discontinuités de profondeur. Nous envisageons donc l'utilisation des mesures de discontinuité d'orientation pour détecter les deux types de discontinuité à la fois.

Comme mentionné en début de section 4.2.2, notre intérêt se porte sur les opérations de voisinage de taille minimale.

4.4.1 Mesures de discontinuité pour une IP cartésienne orthonormée

Ce sont les quatre mesures de discontinuité d'orientation inventoriées au paragraphe 4.2.1 [tables 4.1 et 4.2] qui sont estimées sur un voisinage de taille minimale (3x3):

- la variation quadratique sous la forme

$$\Delta_{Q_{p'}} = \left(\mathbf{D}_0^2 \otimes \mathbf{P}\right)^2 + \frac{1}{8} \left(4\mathbf{D}_{uv} \otimes \mathbf{P}\right)^2 + \left(\mathbf{D}_2^2 \otimes \mathbf{P}\right)^2 \quad (4.35)$$

- la somme des carrés des courbures minimale et maximale Δ_{Q_k} calculée à partir d'une estimation des dérivées partielles avec les masques \mathbf{D}_u^B , \mathbf{D}_v^B , \mathbf{D}_{uu}^B , \mathbf{D}_{vv}^B et \mathbf{D}_{uv} [§4.2.2]. (Le calcul utilise la relation avec les courbures moyenne et gaussienne (4.4).)
- la deuxième dérivée extrême parmi celles des quatre directions discrètes:

$$\Delta_{\hat{p}_n''} = \max_n \left| \mathbf{D}_n^2 \otimes \mathbf{P} \right| \quad (4.36)$$

- la courbure extrême de profil parmi celles des quatre directions discrètes $\Delta_{\hat{k}_n}$, ces dernières étant calculées à partir d'une

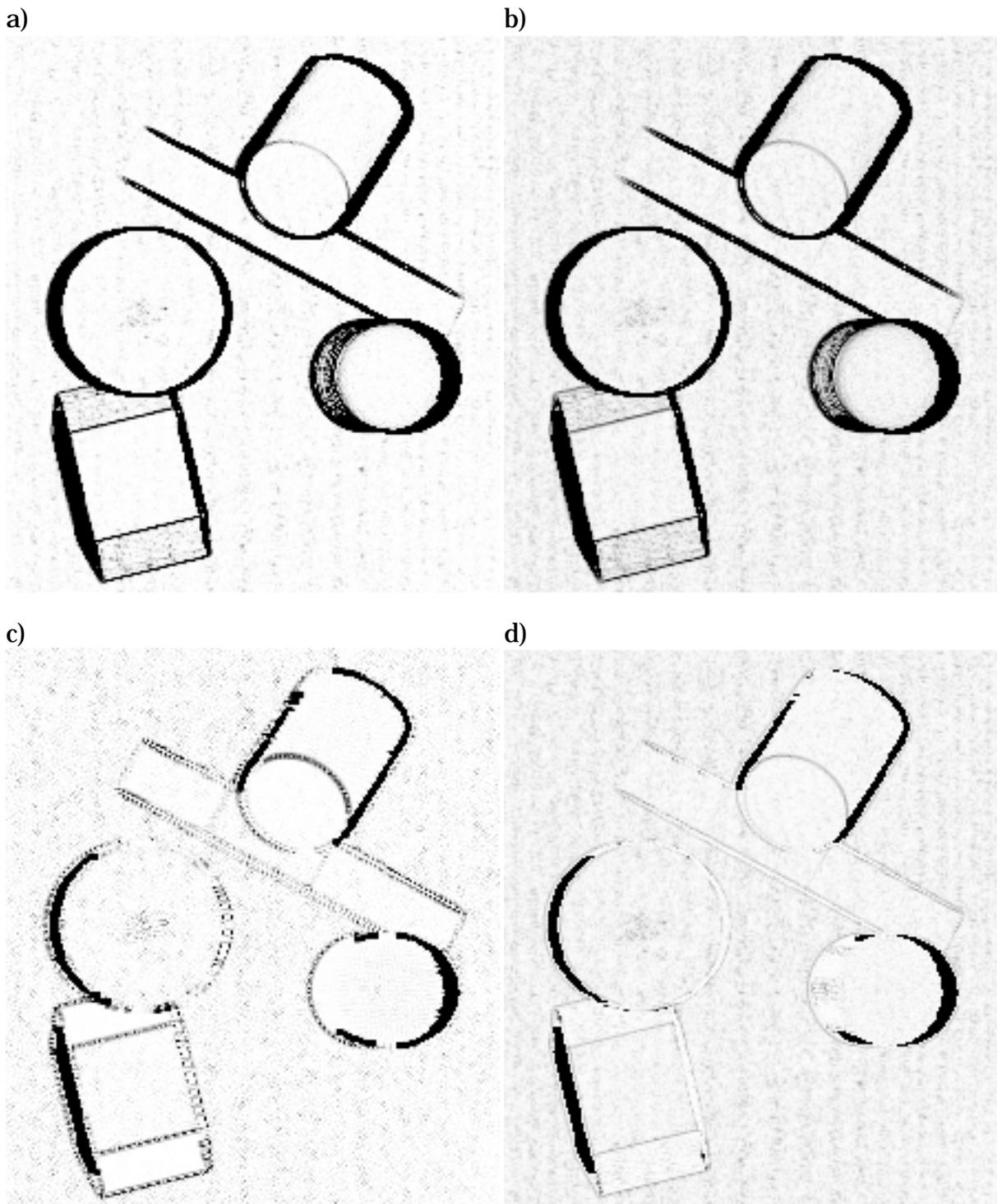


Figure 4.12: Mesures servant à la détection conjointe des discontinuités d'ordre 0 et 1. (Codage par niveaux de gris: plus le gris est foncé, plus la valeur de mesure est élevée.) (a) $\Delta_{Q_{p''}}$, (b) $\Delta_{\hat{p}''_n}$, (c) Δ_{Q_κ} , (d) $\Delta_{\hat{\kappa}_n}$.

estimation des dérivées directionnelles p'_n et p''_n à l'aide des masques de corrélation \mathbf{D}_n^1 et \mathbf{D}_n^2 .

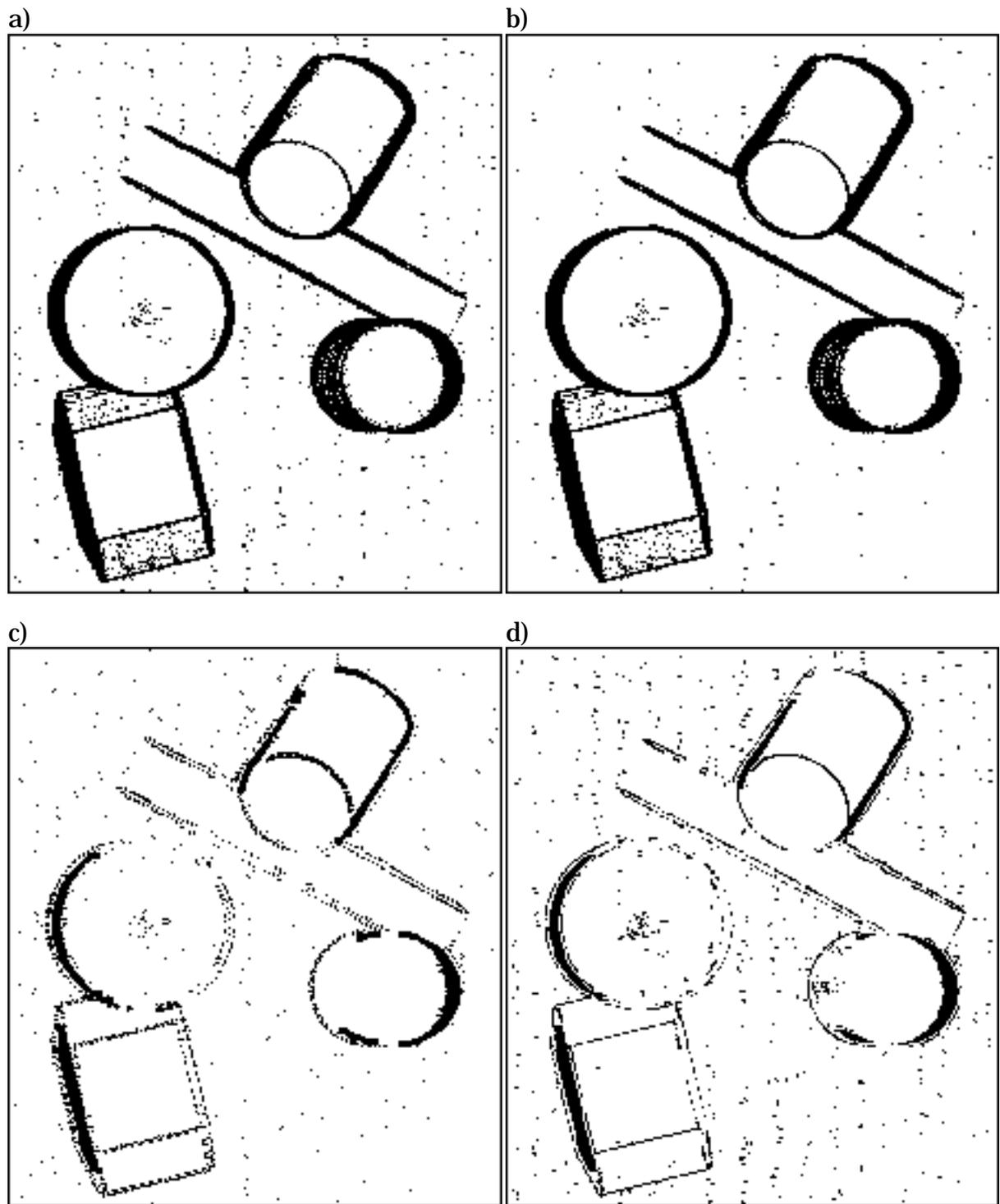


Figure 4.13: Détection conjointe des discontinuités d'ordre 0 et 1. Images des points de discontinuité pour une valeur de seuil élevée. (Convention: les points de discontinuité et les valeurs NIL sont représentés en noir.) (a) $\Delta_{Q_{p''}}$, (b) $\Delta_{\hat{p}''_n}$, (c) $\Delta_{Q_{\kappa}}$, (d) $\Delta_{\hat{\kappa}_n}$.

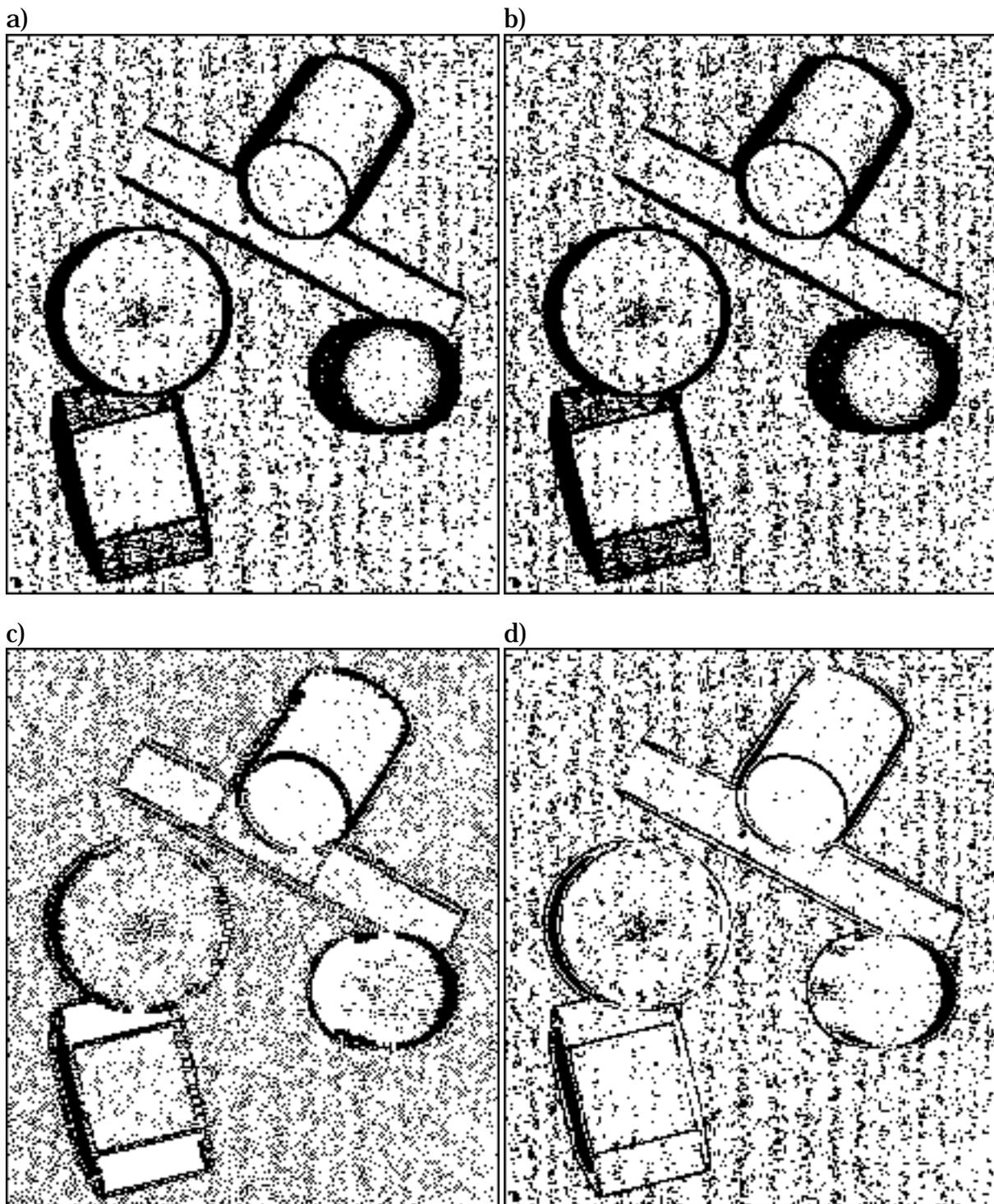


Figure 4.14: Détection conjointe des discontinuités d'ordre 0 et 1. Images des points de discontinuité pour une valeur de seuil basse. (Convention: les points de discontinuité et les valeurs NIL sont représentés en noir.) (a) $\Delta_{Q_{p''}}$, (b) $\Delta_{\hat{p}_n''}$, (c) Δ_{Q_κ} , (d) $\Delta_{\hat{\kappa}_n}$.

La figure 4.12 présente les mesures de discontinuité pour l'image de profondeur BLOC12 [figure 4.3]. Les figures 4.13 et 4.14 montrent deux résultats de seuillage sur ces mesures. On constate les faits suivants:

- Étonnamment, les mesures de discontinuité d'orientation basées sur la courbure donnent de mauvais résultats de détection des discontinuités de profondeur. La raison en est une mauvaise estimation de la courbure par les opérateurs utilisés [figure 4.15]. Pour obtenir le résultat escompté dans le cas de $\Delta_{\hat{\kappa}_n}$ [figure 4.16], il suffit d'évaluer le carré de la dérivée première par la formule

$$p_n'^2(i,j) = \min\left(\left(\mathbf{D}_{n-}^1 \otimes \mathbf{P}(i,j)\right)^2, \left(\mathbf{D}_{n+}^1 \otimes \mathbf{P}(i,j)\right)^2\right) \quad (4.37)$$

Par un moyen similaire, on peut rendre Δ_{Q_k} sensible aux discontinuités de profondeur. Les modifications apportées appliquent essentiellement le principe de voisinage adapté.

- Les mesures de discontinuité basées sur la deuxième dérivée réagissent aux deux types de discontinuité. Cependant, elles sont plus sensibles aux surfaces courbes fortement inclinées. Ce phénomène n'est pas lié à l'échantillonnage: la deuxième dérivée (extrême) d'une surface courbe inclinée est effectivement élevée [figure 4.17]. Ceci a comme inconvénient que de telles surfaces obtiennent l'étiquette de discontinuité suite à la décision par seuillage.
- Toutes les mesures de discontinuité sont fortement sensibles au bruit de l'image de profondeur [figure 4.14]. Il s'en suit que certaines discontinuités ne peuvent pas être détectées parce qu'on ne peut pas abaisser suffisamment le seuil de détection. Les arêtes du prisme de l'image BLOC12 [figures 4.13 et 4.14] en sont un exemple.

Au niveau de la complexité du calcul, les mesures de discontinuité basées sur la deuxième dérivée sont les plus simples. Pour le calcul de $\Delta_{Q_{p''}}$, 9 additions et 5 multiplications suffisent. Le calcul de $\Delta_{\hat{p}_n''}$ nécessite 8 additions, 3 multiplications et 10 autres opérations élémentaires.

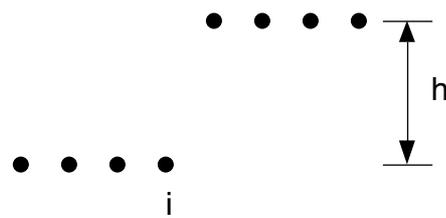


Figure 4.15: Discontinuité d'ordre 0 et estimation de la courbure. Les valeurs de courbure aux positions i et $i+1$ sont élevées, pourtant la grandeur $(p_{j+1} + p_{j-1} - 2p_j)/(p_{j+1} - p_{j-1})$ en ces positions est égale à 1, respectivement -1 , quelle que soit la valeur de h !

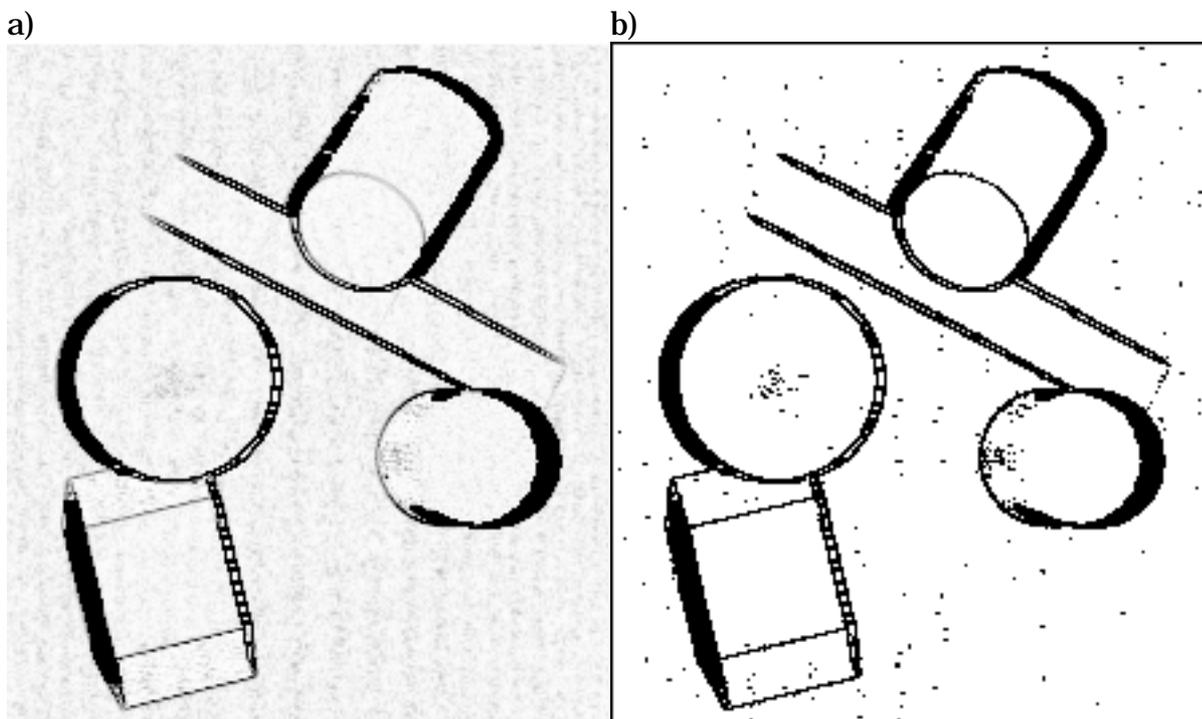


Figure 4.16: Détection conjointe des discontinuités d'ordre 0 et 1 avec $\Delta_{\hat{\kappa}_n}$ modifié. (a) Mesure de discontinuité codée par des niveaux de gris. (b) Image des points de discontinuité obtenue par seuillage sur la mesure de discontinuité.

4.4.2 Extension aux autres types d'IP

Les résultats présentés jusqu'à maintenant sont liés à un type d'image de profondeur particulier (cartésienne orthonormée). Voici

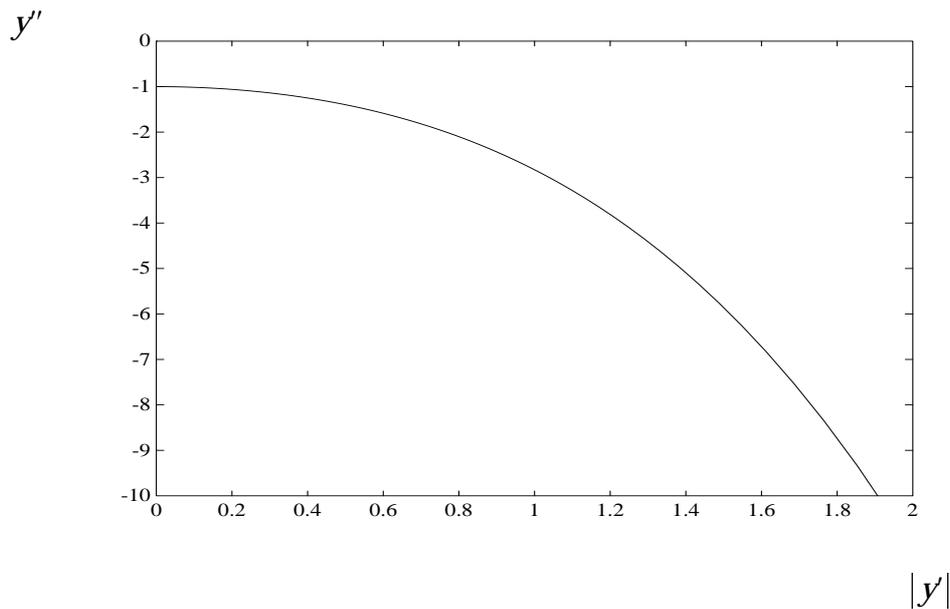


Figure 4.17: Deuxième dérivée en fonction de la pente pour l'arc de cercle $y = \sqrt{1 - x^2}$, $x \in [0, 0.9[$.

comment les mesures de discontinuité peuvent être étendues aux autres types d'IP.

IP scalaires

Formellement, toutes les mesures de discontinuité peuvent être appliquées à n'importe quelle IP scalaire. A ce moment-là, c'est la surface $p(u, v)$ plutôt que $z(x, y)$ qui est analysée. Qualitativement, les résultats devraient être similaires à ceux obtenus en transformant la surface dans le repère cartésien.

IP vectorielles

La mesure de discontinuité Δ_{Q_κ} peut être appliquée aux IP de type vectoriel, il suffit d'estimer les courbures avec les méthodes appropriées [§4.2.2].

La mesure de discontinuité de profil $\Delta_{\hat{\kappa}_n}$ peut aussi être appliquée à une image de type vectoriel. Le profil en question se trouve dans le plan passant par le rangel de position (i, j) et ses deux voisins dans la direction discrète de référence [figure 4.18]. La courbure est estimée à partir de la taille des segments consécutifs définis par les trois points et de l'angle que forment ces segments.

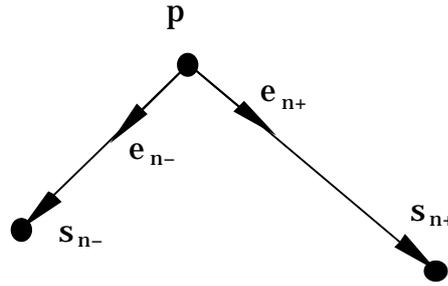


Figure 4.18: Grandeurs servant à l'estimation de la courbure d'un profil pour une IP vectorielle.

On en déduit la mesure de discontinuité

$$\Delta_{\hat{\kappa}_n}(i,j) = \max_n \frac{1 + \mathbf{e}_{n-}(i,j) \cdot \mathbf{e}_{n+}(i,j)}{\min(\|\mathbf{s}_{n-}(i,j)\|, \|\mathbf{s}_{n+}(i,j)\|)} \quad (4.38)$$

où l'angle entre segments consécutifs est approché par le produit scalaire de leurs vecteurs directeurs unitaires

$$\mathbf{e}_{n\pm}(i,j) = \mathbf{s}_{n\pm}(i,j) / \|\mathbf{s}_{n\pm}(i,j)\| \quad (4.39)$$

avec

$$\mathbf{s}_{n\pm}(i,j) = \mathbf{D}_{n\pm} \otimes \mathbf{P}(i,j) \quad (4.40)$$

Finalement, on peut aussi étendre l'utilisation des mesures de discontinuité basées sur la deuxième dérivée aux IP de type vectoriel. La variation quadratique devient

$$\Delta_{Q_{p'}} = \|\mathbf{D}_0^2 \otimes \mathbf{P}\|^2 + \frac{1}{8} \|4\mathbf{D}_{uv} \otimes \mathbf{P}\|^2 + \|\mathbf{D}_2^2 \otimes \mathbf{P}\|^2 \quad (4.41)$$

et la deuxième dérivée extrême parmi celles des quatre directions discrètes prend la forme

$$\Delta_{\hat{p}''_n} = \max_n \|\mathbf{D}_n^2 \otimes \mathbf{P}\| \quad (4.42)$$

On peut vérifier que dans le cas d'une image de profondeur cartésienne orthographique, c'est-à-dire

$$\mathbf{P}_{i,j} = (i\Delta x, j\Delta y, z) \quad (4.43)$$

les mesures de discontinuité (4.41) et (4.42) sont équivalentes à celles définies par les équations (4.35) et (4.36), respectivement. Les premières citées constituent donc une forme généralisée de celles citées en dernier.

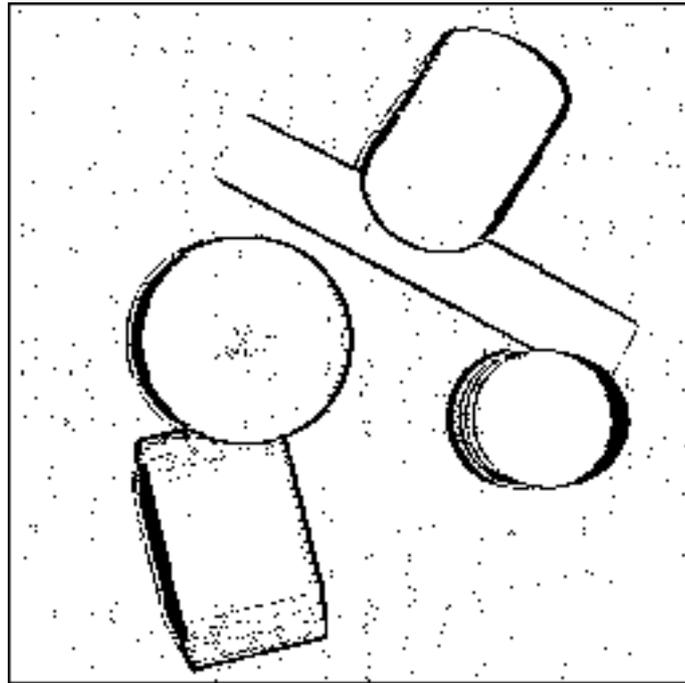


Figure 4.19: Détection des discontinuités convexes (2ème dérivée positive).

4.4.3 Récapitulation

Parmi les quatre mesures de discontinuité d'orientation proposées, trois se sont révélées utiles à la détection conjointe des discontinuités d'ordre 0 et d'ordre 1. Ce sont la variation quadratique ($\Delta_{Q_{p''}}$), la deuxième dérivée extrême parmi celles des quatre directions discrètes ($\Delta_{\hat{p}''_n}$) et la courbure extrême de profil parmi celles des quatre directions discrètes ($\Delta_{\hat{\kappa}_n}$). La quatrième mesure, soit la somme des carrés des courbures minimale et maximale (Δ_{Q_κ}), pourrait être modifiée afin d'être sensible aux discontinuités d'ordre 0. Mais elle correspondrait à un coût maximal de calcul sans offrir un avantage évident.

Chacune des trois mesures retenues pour la détection conjointe des discontinuités possède un intérêt particulier. $\Delta_{Q_{p''}}$ correspond au moindre coût en calcul. $\Delta_{\hat{\kappa}_n}$ est la seule mesure qui ne soit pas sensible aux surfaces courbes inclinées. Finalement, $\Delta_{\hat{p}''_n}$ offre la possibilité de fournir d'autres informations utiles: valeurs et signes (concavité ou convexité) des deuxièmes dérivées minimale et maximale [figure 4.19].

4.5 Conclusion

Ce chapitre a été consacré aux détecteurs différentiels de point de discontinuité et à leur utilisation pour la construction de frontières de surfaces.

En parcourant la littérature, nous avons constaté que tous les auteurs effectuent une détection séparée des discontinuités \bar{C}^0 et \bar{C}^1 , ce qui est inutile. D'autre part, ce sont souvent les détecteurs avec condition d'extremum qui sont utilisés alors qu'ils fonctionnent mal.

Suite à cette analyse, nous avons fait l'inventaire des détecteurs différentiels de taille minimale n'utilisant pas la condition d'extremum et capables de détecter à la fois les discontinuités d'ordre 0 et celles d'ordre 1. Nous en avons retenu trois qui remplissent les conditions. Leur performance est comparable. D'abord définis pour des IP cartésiennes orthographiques, ils ont été généralisés aux autres types d'IP.

Un de ces détecteurs, celui basé sur une mesure de la deuxième dérivée extrême, constitue un des éléments de base de notre méthode, présentée au chapitre 6, qui combine détection des discontinuités et atténuation du bruit.

Chapitre 5

Estimation du bruit

Ce chapitre est consacré à l'estimation du bruit d'une image de profondeur.

Dans une première partie, la relation entre bruit et erreur de mesure est précisée. Nous montrons comment l'erreur de mesure peut être considérée, sous certaines conditions, comme un bruit additif.

Une deuxième partie pose le problème de l'estimation du bruit et présente brièvement trois méthodes utilisées pour le résoudre.

Dans une dernière partie, nous proposons une nouvelle méthode d'estimation du bruit. Nous présentons les résultats obtenus avec cette méthode et terminons par une comparaison des méthodes.

5.1 Erreur de mesure et bruit

5.1.1 Erreur de mesure d'une image de profondeur

Une image de profondeur est, comme toute mesure, entachée d'erreur. Il existe différentes interprétations de l'erreur de mesure.

Erreur par rapport au point mesuré ou erreur absolue

A chaque rangel $\mathbf{p}_{i,j}$ on associe une position exacte $\overline{\mathbf{p}}_{i,j}$. L'erreur de mesure absolue est la différence

$$\overline{\mathbf{e}}_{i,j}^{\mathbf{p}} = \mathbf{p}_{i,j} - \overline{\mathbf{p}}_{i,j} \quad (5.1)$$

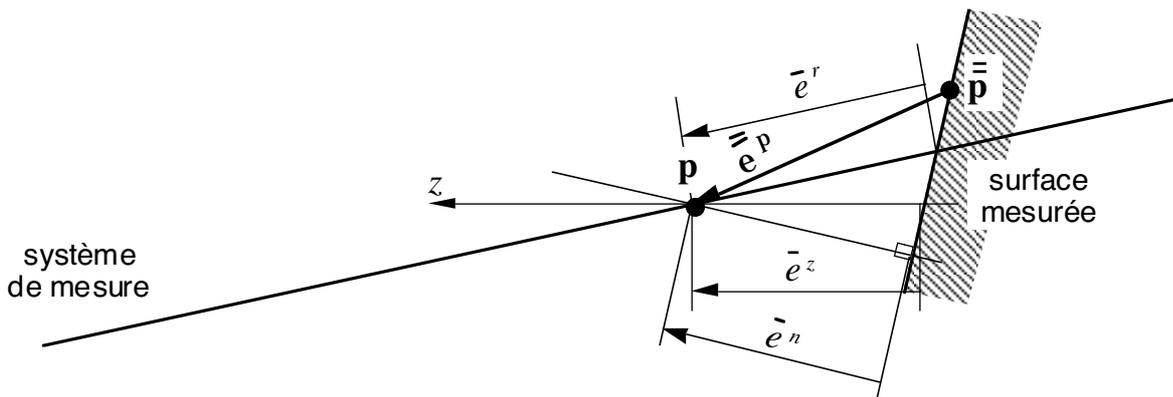


Figure 5.1: Différentes interprétations de l'erreur de mesure.

Erreur par rapport à la surface mesurée

Vu l'échantillonnage de mesures dense qui caractérise une IP, l'intérêt de la mesure d'un point se trouve plus dans le positionnement du point relativement à la surface que dans son positionnement en absolu. L'erreur significative est donc l'écart du point de mesure $\mathbf{p}_{i,j}$ par rapport à la surface échantillonnée. Cette grandeur est appelée ici erreur apparente.

Voici trois manières de définir l'erreur apparente. La première est référencée à la surface mesurée. Les deux autres sont référencées au système de mesure [figure 5.1].

- L'erreur apparente normale $\bar{e}_{i,j}^n$ est la distance (orthogonale) signée qui sépare le point de mesure $\mathbf{p}_{i,j}$ de la surface mesurée. Nous avons utilisé cette grandeur pour caractériser les performances du système d'acquisition présenté dans [HüMa89, MaHü88].
- L'erreur apparente de distance $\bar{e}_{i,j}^r$ est la distance signée qui sépare le point $\mathbf{p}_{i,j}$ de la surface mesurée, le long de la droite définie par la direction de mesure.
- L'erreur apparente dans la direction z $\bar{e}_{i,j}^z$ est la distance signée qui sépare $\mathbf{p}_{i,j}$ du point de la surface ayant mêmes coordonnées $x_{i,j}$ et $y_{i,j}$. C'est en général à cette grandeur que l'on fait allusion lorsqu'on parle d'erreur de mesure dans le cas d'IP orthographiques [BeJa88, Besl88, HäAM89].

5.1.2 Bruit

Dans la terminologie du traitement du signal, l'erreur de mesure d'une IP est un bruit: composante du signal non issue de la source, alors que la valeur exacte des rangels constitue le signal de source. Nous utilisons donc le terme de bruit comme synonyme d'erreur de mesure.

Modèles de bruit additif

Un modèle de bruit additif doit être interprété comme suit en termes d'erreur de mesure:

- bruit additif vectoriel

$$\mathbf{p}_{i,j} = \overline{\mathbf{p}}_{i,j} + \mathbf{e}_{i,j} \quad (5.2)$$

C'est l'erreur absolue pour une IP de type vectoriel.

- bruit additif scalaire

$$p_{i,j} = \overline{p}_{i,j} + e_{i,j} \quad (5.3)$$

C'est l'erreur apparente de distance, pour une IP brute, ou l'erreur de mesure apparente dans la direction z , pour une IP cartésienne orthographique.

5.2 Estimation du bruit

5.2.1 Introduction

Déterminer l'erreur de mesure d'une IP requiert l'utilisation d'un autre système de mesure plus précis pour indiquer la valeur de mesure exacte. Cependant, quand aucune donnée extérieure au signal mesuré n'est disponible, il faut estimer le bruit sur la base de l'IP seule. L'estimation du bruit repose alors sur des hypothèses concernant le bruit et le signal de source.

5.2.2 Méthodes d'estimation du bruit

Nous rapportons trois méthodes d'estimation du bruit tirées de la littérature. Les trois méthodes sont appliquées à des IP de type scalaire. Elles fournissent une estimée de l'erreur de mesure additive. Les hypothèses faites sur le signal de source et le bruit ne sont pas toujours explicites. Quoiqu'il en soit, nous nous contentons pour le moment de décrire les méthodes.

	Estimée locale $\varepsilon_{i,j}$	Condition d'élimination	Estimée globale E
M1	$p_{i,j} - \overline{p_{k,l}} _{W_{i,j}^{(5,9)}}$	-	$\sqrt{\overline{\varepsilon_{i,j}^2}}$
M2	$\sqrt{\overline{\delta_{k,l}^2}} _{W_{i,j}}$	$\delta_{i,j}^{max} - \delta_{i,j}^{min} < T$	$\sqrt{\overline{\varepsilon_{i,j}^2}}$
M3	$\sqrt{\overline{\delta_{k,l}^2}} _{W_{i,j}}$	$s_{i,j}^2 < T$	$\overline{\varepsilon_{i,j}}$

Table 5.1: Méthodes d'estimation du bruit, description formelle. Précisions concernant la notation: δ représente l'erreur d'approximation par le plan de meilleure approximation sur le voisinage V ; s est la pente du plan; T est une valeur de seuil.

Trois méthodes particulières

Pour un sommaire formel des méthodes, nous renvoyons le lecteur à la table 5.1.

(M1) Ylä-Jääski et Ade [YJAd90] utilisent l'opérateur 3x3 de moyennage sur les 5 plus proches voisins pour établir une estimée de la valeur exacte du rangel. Le bruit est estimé par la différence entre l'IP originale (valeur mesurée) et l'IP filtrée (estimée de la valeur exacte). La valeur efficace du signal de bruit est retenue comme estimée globale du bruit.

(M2) Jiang et Bunke [JiBu94] estiment localement la surface mesurée par le plan de meilleure approximation. Le niveau local de bruit est estimé par la racine carrée de l'erreur quadratique moyenne (EQM) d'approximation. Pour chaque position de l'IP sont retenues, à côté de l'EQM, les valeurs extrêmes (signées) des erreurs d'approximation. L'estimée globale du bruit est calculée à partir d'un moyennage (sur l'IP) des EQM locales, en écartant les positions pour lesquelles la différence entre valeurs d'erreurs extrêmes est au-dessus d'un seuil prédéfini. Plus précisément, l'estimée globale est la racine carrée de la moyenne des EQM locales.

(M3) Chez Besl [Besl88], c'est aussi l'erreur d'approximation locale par un plan qui sert de base à l'estimation de l'erreur de

mesure. A côté de l'erreur quadratique moyenne (EQM) d'approximation, sont retenues ici, à chaque position, les grandeurs déterminant la pente du plan. En éliminant les positions correspondant à un plan de forte pente, l'estimée global est déterminée soit par la racine carrée de la moyenne des EQM locales, soit par la moyenne des racines carrées des EQM locales. Dans les exemples présentés, c'est la deuxième solution qui est utilisée.

Remarque:

Dans les méthodes d'estimation du bruit par approximation de plan qui viennent d'être décrites, l'erreur d'approximation est définie comme la distance séparant le rangé du point du plan ayant mêmes coordonnées (u,v) . Une alternative consiste à définir l'erreur d'approximation comme la distance du point au plan. Dans ce cas, c'est l'erreur apparente normale qui est estimée. Cette manière de faire, plus coûteuse en calcul, a l'avantage d'être applicable aussi aux IP de type vectoriel.

5.3 Estimation du bruit basée sur une mesure de la deuxième dérivée

Cette section propose une nouvelle méthode d'estimation du bruit. Comme les méthodes présentées à la section précédente, elle est dédiée aux IP de type scalaire et fournit une estimée de l'erreur de mesure additive.

D'abord, la méthode est présentée sous une forme élémentaire, puis sous une forme perfectionnée. Certains choix nécessaires à fixer l'algorithme pour son implémentation sont ensuite précisés. Des résultats et une comparaison avec les méthodes d'estimation du bruit présentées précédemment terminent la section.

5.3.1 Méthode élémentaire

Hypothèses

Soit le signal de source $\bar{\mathbf{P}}$ et le bruit \mathbf{E} dont l'image de profondeur \mathbf{P} est la somme

$$\mathbf{P} = \bar{\mathbf{P}} + \mathbf{E} \quad (5.4)$$

$\bar{\mathbf{P}}$ et \mathbf{E} vérifient les propriétés suivantes

- $\bar{\mathbf{P}}$ est un plan échantillonné non quantifié (la scène est constituée d'une seule surface plane).
- \mathbf{E} est un signal stochastique identiquement indépendamment distribué (IID) de valeur moyenne nulle et d'écart-type σ_E .

Propriété

Soit \mathbf{Q} la deuxième dérivée de l'image de profondeur \mathbf{P} , calculée à l'aide d'un des masques de taille minimale \mathbf{D}_l^2 définis en (4.16)

$$\mathbf{Q} = \mathbf{D}_l^2 \otimes \mathbf{P} \quad (5.5)$$

L'écart-type σ_E du bruit \mathbf{E} est lié à Q_{eff} , la valeur efficace de \mathbf{Q} , par

$$\sigma_E \cong \frac{1}{\sqrt{6}} Q_{eff} \quad (5.6)$$

Démonstration

Soit $\chi_{i,j}$ la variable aléatoire associée à la position (i, j) du signal stochastique \mathbf{E} . Sans restreindre la généralité, nous considérons le cas de la deuxième dérivée selon les lignes ($l=0$ dans (5.5)). La démonstration de (5.6) est conduite en trois étapes:

- (1) Le signal de source étant un plan échantillonné, sa deuxième dérivée est nulle

$$\mathbf{D}_0^2 \otimes \bar{\mathbf{P}} = 0 \quad (5.7)$$

et la deuxième dérivée de \mathbf{P} est celle du bruit

$$\mathbf{Q} = \mathbf{D}_0^2 \otimes \mathbf{P} = \mathbf{D}_0^2 \otimes \mathbf{E} \quad (5.8)$$

Donc \mathbf{Q} est un signal stochastique et on peut associer à la position (i, j) de \mathbf{Q} une variable aléatoire $\gamma_{i,j}$ qui vérifie

$$\gamma_{i,j} = \chi_{i-1,j} - 2 \cdot \chi_{i,j} + \chi_{i+1,j} \quad (5.9)$$

- (2) Etant donné la propriété de linéarité de la valeur moyenne et comme les variables aléatoires $\chi_{i,j}$ sont de valeur moyenne nulle, on déduit de (5.9) que $\gamma_{i,j}$ possède une valeur moyenne nulle

$$E[\gamma_{i,j}] = 0 \quad (5.10)$$

D'autre part, comme les variables aléatoires $\chi_{i,j}$ sont statistiquement indépendantes, la variance d'une des variables $\gamma_{i,j}$ vérifie la relation

$$V[\gamma_{i,j}] = V[\chi_{i-1,j}] + 4V[\chi_{i,j}] + V[\chi_{i+1,j}] \quad (5.11)$$

et, comme les variables aléatoires $\chi_{i,j}$ sont identiquement distribuées, la relation (5.11) se simplifie sous la forme

$$V[\gamma_{i,j}] = 6 V[\chi_{i,j}] \quad (5.12)$$

- (3) Comme les variables $\chi_{i,j}$ sont identiquement distribuées, on déduit de la relation (5.9) que les variables $\gamma_{i,j}$ et \mathbf{Q} le sont également. D'autre part, comme les variables aléatoires $\gamma_{i,j}$ et \mathbf{Q} sont de valeur moyenne nulle (5.10), on en déduit que la valeur efficace Q_{eff} de \mathbf{Q} est une estimée de la variance d'une des variable $\gamma_{i,j}$. Finalement, en tenant compte de la relation (5.12), on aboutit à propriété (5.6).

qed

Remarque

L'opération de deuxième dérivée (5.5) est bien évidemment restreinte au domaine pour lequel les rangels intervenant dans l'opération ont une valeur différente de *NIL*.

Estimation élémentaire du bruit

La méthode élémentaire consiste donc à utiliser la valeur

$$E = \frac{1}{\sqrt{6}} Q_{eff} \quad (5.13)$$

comme estimée du bruit pour une IP quelconque.

5.3.2 Affinage de la méthode

La méthode d'estimation du bruit qui vient d'être présentée est modifiée par l'introduction d'un traitement supplémentaire.

Motif et objectif

L'hypothèse faite, suivant laquelle la scène n'est constituée que d'une surface plane, n'est quasiment jamais strictement vraie. Tout au plus peut-on considérer que la surface mesurée apparaisse localement plane. Ceci constituerait une condition suffisante pour valider l'utilisation de la méthode élémentaire. Cependant, dans beaucoup d'IP, il existe des discontinuités de profondeur, des discontinuités d'orientation et des surfaces relativement courbes qui engendrent des valeurs élevées de la deuxième dérivée [figure 5.2]. Celles-ci ont une grande influence sur la valeur efficace Q_{eff} et faussent l'estimation du

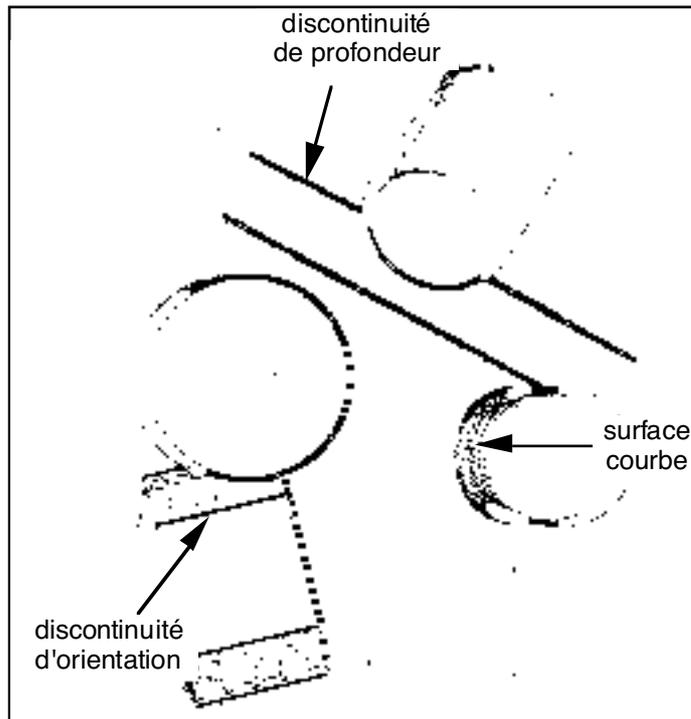


Figure 5.2: Valeurs élevées de la deuxième dérivée. Un seuillage sur la valeur absolue de la deuxième dérivée selon les lignes de l'IP BLOC12 [figure 4.3] fait apparaître discontinuités de profondeur, discontinuités d'orientation et certaines surfaces courbes.

bruit dans le sens d'une forte surestimation. Une modification de la méthode élémentaire est donc nécessaire.

Il s'agit d'éliminer du calcul de la valeur efficace ces valeurs élevées de deuxième dérivée.

Propriété

En vertu de l'hypothèse de cohérence spatiale [§2.1.1], les valeurs élevées de la deuxième dérivée qui ne correspondent pas à du bruit sont rares. L'histogramme de la figure 5.3, mis en relation avec l'image de la figure 5.2, confirme cet état de fait dans le cas de l'exemple considéré. En effet, si les valeurs supérieures à T ne sont pas visibles dans la représentation graphique de l'histogramme, à cause de l'effet de quantification, c'est qu'elles sont rares.

Dans la terminologie des statistiques, ce sont des observations aberrantes par rapport à la distribution que l'on associe au bruit [Hube81, BaLe84]. Une estimation qui rejette ces valeurs aberrantes est appelée robuste.

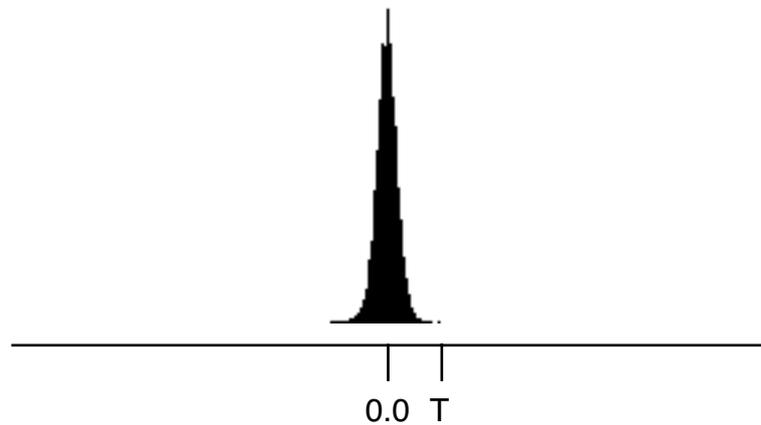


Figure 5.3: Histogramme de l'image de deuxième dérivée servant à l'estimation du bruit. T représente le seuil utilisé dans la figure 5.2. Remarque: Les valeurs extrêmes ne sont pas visibles car insuffisamment nombreuses.

Estimation robuste du bruit

Le fait que les valeurs élevées de la deuxième dérivée soient rares a comme conséquence que le rapport entre ces valeurs et la valeur efficace Q_{eff} est important. D'où la méthode itérative suivante pour éliminer ces valeurs élevées.

Un seuillage avec un seuil proportionnel à Q_{eff} d'un facteur $k > 1$ élimine les observations aberrantes les plus élevées. Ceci provoque une diminution de Q_{eff} pour les observations restantes. En appliquant à nouveau un seuil proportionnel à la nouvelle valeur de Q_{eff} (avec le même coefficient de proportionnalité k), de nouvelles observations aberrantes sont éliminées, et ainsi de suite. Comme la valeur des observations aberrantes éliminées est de plus en plus petite, la modification de Q_{eff} diminue à chaque itération. On s'attend à ce que le procédé converge vers une valeur non nulle pour un coefficient de proportionnalité suffisamment grand.

La méthode est réalisée avantageusement à partir de l'histogramme h de l'image de deuxième dérivée. Elle effectue une troncature itérative de l'histogramme.

Remarque

Dans notre réalisation de la méthode, la troncature est effectuée de manière symétrique par rapport à la valeur moyenne de

l'histogramme, au lieu d'être effectuée de manière symétrique par rapport à la valeur zéro. Ce décalage vis-à-vis du principe décrit ci-dessus repose sur une raison historique: la méthode a été implémentée et testée sous cette forme plus générale. Nous n'avons pas étudié quel était la différence de comportement entre ces deux versions.

Description formelle

Soit l'histogramme normalisé h de l'image de deuxième dérivée \mathbf{Q} défini par

$$h(\xi) = \sum_m \frac{\text{card}\left(\left\{(i,j) \in D_p^{l*} \mid (m-\frac{1}{2})\Delta\xi \leq \mathbf{Q}(i,j) < (m+\frac{1}{2})\Delta\xi\right\}\right)}{\text{card}(D_p^{l*})} \delta(\xi - m\Delta\xi) \quad (5.14)$$

où card est la fonction qui dénombre les éléments d'un ensemble, δ la distribution de Dirac, $\Delta\xi$ un intervalle de la droite des réels de taille arbitraire et D_p^{l*} le domaine pour lequel les rangels intervenant dans l'opération du calcul de la deuxième dérivée ont une valeur différente de NIL .

Soit les trois fonctions suivantes des deux variables a et b qui définissent des caractéristiques statistiques de la portion de l'histogramme comprise entre $a - b$ et $a + b$:

- $\pi(a, b)$ est la valeur intégrale de l'histogramme

$$\pi(a, b) = \int_{a-b}^{a+b} h(\xi) d\xi \quad (5.15)$$

- $\mu(a, b)$ est la valeur moyenne

$$\mu(a, b) = \frac{1}{\pi(a, b)} \int_{a-b}^{a+b} \xi \cdot h(\xi) d\xi \quad (5.16)$$

- et $\sigma(a, b)$ est l'écart-type

$$\sigma(a, b) = \sqrt{\frac{1}{\pi(a, b)} \int_{a-b}^{a+b} \xi^2 \cdot h(\xi) d\xi - \mu^2(a, b)} \quad (5.17)$$

Soit $\mu^{(n-1)}$ et $\sigma^{(n-1)}$, respectivement, la valeur moyenne et l'écart-type de l'histogramme tronqué après l'itération $(n-1)$.

A l'itération n , l'histogramme est tronqué à l'intervalle $[\mu^{(n-1)} - k\sigma^{(n-1)}, \mu^{(n-1)} + k\sigma^{(n-1)}]$ où k indique en quelle proportion la troncature a lieu par rapport à l'écart-type. La valeur moyenne et

l'écart-type de l'histogramme prennent des nouvelles valeurs $\mu^{(n)}$ et $\sigma^{(n)}$. On obtient ainsi les suites $\mu^{(n)}$ et $\sigma^{(n)}$ définies à l'aide des fonctions $\mu(\cdot, \cdot)$ et $\sigma(\cdot, \cdot)$ par

- les valeurs initiales: $\mu^{(0)}$ et $\sigma^{(0)}$ qui sont respectivement la valeur moyenne et l'écart-type de l'histogramme complet

$$\begin{cases} \mu^{(0)} = \mu(0, \infty) \\ \sigma^{(0)} = \sigma(0, \infty) \end{cases} \quad (5.18.a-b)$$

- les relations de récurrence: $\mu^{(n)}$ et $\sigma^{(n)}$ qui sont la valeur moyenne et l'écart-type de l'histogramme tronqué au pas d'itération n

$$\begin{cases} \mu^{(n)} = \mu(\mu^{(n-1)}, k\sigma^{(n-1)}) \\ \sigma^{(n)} = \sigma(\mu^{(n-1)}, k\sigma^{(n-1)}) \end{cases} \quad (5.19.a-b)$$

L'estimée du bruit est

$$E = \frac{1}{\sqrt{6}} \sigma^{(N)} \quad (5.20)$$

où N désigne le nombre d'itérations pour lequel la convergence est atteinte.

$$N = \min(n) \mid \pi(k\sigma^{(n)}) = \pi(k\sigma^{(n-1)}) \quad (5.21)$$

Comme un histogramme est discret, la convergence est atteinte en un nombre fini d'itérations.

Exemples analytiques

Afin d'illustrer l'effet de la troncature itérative de l'histogramme, trois exemples sont considérés, dans lesquels l'histogramme est remplacé par une fonction analytique. Deux d'entre eux correspondent à des distributions fréquemment utilisées pour des variables aléatoires: distribution uniforme et distribution gaussienne. Le troisième correspond à une distribution que nous appelons mixte. Il s'agit d'une distribution uniforme avec une probabilité non-nulle pour la valeur 0. On remarque que toutes les trois distributions présentent des valeurs moyennes nulles.

Pour chacun des cas, la valeur limite de la suite $\sigma^{(n)}$ a été calculée en fonction de k [Annexe A]. Les résultats sont rassemblés dans la table 5.2. La figure 5.4 complète les résultats pour la distribution gaussienne en représentant le rapport $\sigma^{(\infty)}/\sigma^{(0)}$ en fonction de k .

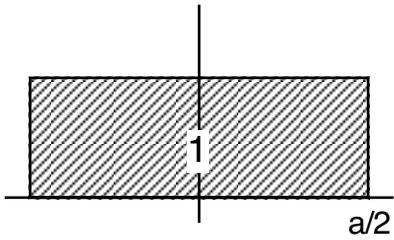
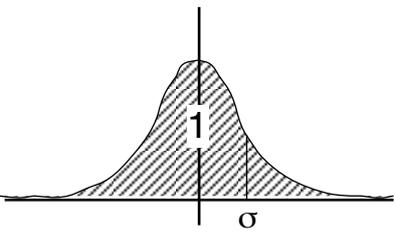
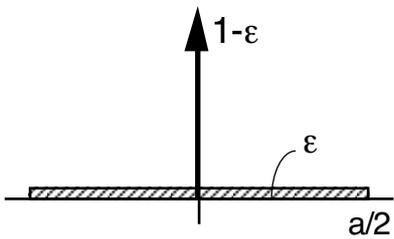
<p>a)</p>  $h(\xi) = \begin{cases} \frac{1}{a} & , \xi < \frac{a}{2} \\ 0 & , \xi > \frac{a}{2} \end{cases}$	$\lim_{n \rightarrow \infty} \sigma^{(n)} = \begin{cases} 0 & , k < \sqrt{3} \\ \sigma^{(0)} & , k \geq \sqrt{3} \end{cases}$
<p>b)</p>  $h(\xi) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\xi^2/2\sigma^2}$	$\lim_{n \rightarrow \infty} \sigma^{(n)} = 0 \quad , k \leq \sqrt{3}$ $0 < \lim_{n \rightarrow \infty} \sigma^{(n)} < \sigma^{(0)} \quad , k > \sqrt{3}$
<p>c)</p>  $h(\xi) = \begin{cases} (1 - \varepsilon)\delta(\xi) & , \xi = 0 \\ \frac{\varepsilon}{a} & , 0 < \xi < \frac{a}{2} \\ 0 & , \xi > \frac{a}{2} \end{cases}$	$\lim_{n \rightarrow \infty} \sigma^{(n)} = \begin{cases} 0 & , k < \sqrt{\frac{3}{\varepsilon}} \\ \sigma^{(0)} & , k \geq \sqrt{\frac{3}{\varepsilon}} \end{cases}$

Table 5.2: Effet de la troncature itérative sur quelques distributions standards. (a) distribution uniforme, (b) distribution gaussienne, (c) distribution mixte.

Ces résultats vont être utilisés pour choisir une valeur du paramètre k .

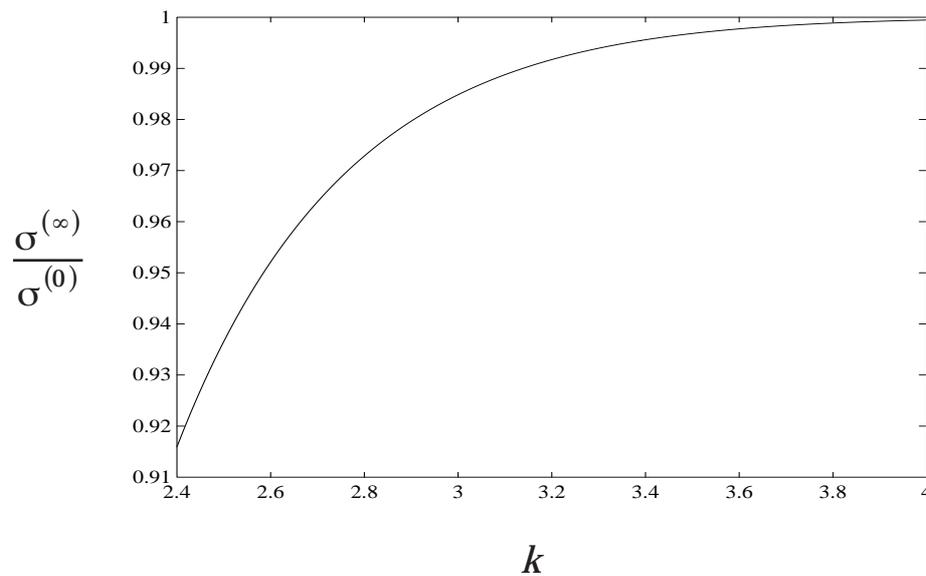


Figure 5.4: Effet de la troncature itérative sur l'écart-type dans le cas d'une distribution gaussienne.

5.3.3 Algorithme d'estimation robuste du bruit

La mise en oeuvre de la méthode d'estimation robuste du bruit nécessite deux choix: celui de k et celui de la direction de dérivation.

Choix de k

Globalement, l'histogramme h de l'image de deuxième dérivée \mathbf{Q} peut être modélisé par la distribution mixte de la table 5.2 où les valeurs non-nulles de l'histogramme représentent les observations aberrantes, la probabilité ε de l'apparition d'une observation aberrante étant très petite. Pour éliminer les valeurs aberrantes de \mathbf{Q} par troncature de l'histogramme, il faut que la suite $\sigma^{(n)}$ tende vers 0. Il faut donc choisir k plus petit que $\sqrt{\frac{3}{\varepsilon}}$ [table 5.2].

$$k < \sqrt{\frac{3}{\varepsilon}} \quad (5.22)$$

Considéré dans le voisinage de zéro, l'histogramme h présente des observations non-aberrantes qui ne sont pas toutes exactement nulles. Autour de zéro, elles obéissent à une distribution que l'on peut modéliser par exemple par une distribution uniforme ou par une distribution gaussienne. Comme on veut estimer le bruit par l'écart-

<p>Signal d'entrée: P image de profondeur</p> <p>Valeurs de sortie: E estimée du niveau de bruit</p> <p>Paramètres d'entrée: k seuil relatif pour tronquer l'histogramme k = 3 (défaut)</p>
<p>Pour chacune des deux directions, ligne et colonne ($l \in \{0, 2\}$) calculer l'image de deuxième dérivée Q = D_l² ⊗ P calculer l'histogramme de l'image de deuxième dérivée</p> $h(\xi) = \sum_m \frac{\text{card}\left\{\left\{(i, j) \in D_P^* \mid (m - \frac{1}{2})\Delta\xi \leq \mathbf{Q}(i, j) < (m + \frac{1}{2})\Delta\xi\right\}\right\}}{\text{card}(D_P^*)} \delta(\xi - m\Delta\xi)$ <p>calculer la valeur moyenne et l'écart-type de l'histogramme</p> $\begin{cases} \mu^{(0)} = \mu(0, \infty) \\ \sigma^{(0)} = \sigma(0, \infty) \end{cases}$ <p>répéter la procédure de troncature de l'histogramme</p> $\begin{cases} \mu^{(n)} = \mu\left(\mu^{(n-1)}, k\sigma^{(n-1)}\right) \\ \sigma^{(n)} = \sigma\left(\mu^{(n-1)}, k\sigma^{(n-1)}\right) \end{cases}$ <p>jusqu'à ce que l'histogramme ne soit plus modifié</p> $\pi\left(\mu^{(n)}, k\sigma^{(n)}\right) = \pi\left(\mu^{(n-1)}, k\sigma^{(n-1)}\right)$ <p>retenir l'écart-type de l'histogramme tronqué</p> $\sigma_l = \sigma^{(N)}$ <p>Déterminer l'estimée du bruit à partir de la valeur max. des écart-types des histogrammes tronqués</p> $E = \frac{1}{\sqrt{6}} \max(\sigma_0, \sigma_2)$

Figure 5.5: Algorithme d'estimation robuste du bruit.

type des observations non-aberrantes, il faut que la troncature itérative laisse quasiment inchangé cet écart-type.

Pour la distribution uniforme, la suite $\sigma^{(n)}$ est identiquement égale à $\sigma^{(0)}$ pour k plus grand que $\sqrt{3}$, sinon elle tend vers 0 [table 5.2]. Pour la distribution gaussienne, $\sigma^{(n)}$ converge vers une valeur

non-nulle pour k supérieur à $\sqrt{3}$ [table 5.2]. Le rapport $\sigma^{(\infty)}/\sigma^{(0)}$ est proche de 1 pour une valeur de k au-dessus de 2.5 [figure 5.4]

Par conséquent, pour éviter une sous-estimation du bruit, il faut donc choisir k vérifiant

$$k > 2.5 \quad (5.23)$$

Des contraintes ci-dessus résulte la condition globale pour k

$$2.5 < k < \sqrt{\frac{3}{\varepsilon}} \quad (5.24)$$

Le choix $k = 3$ constitue un bon compromis.

Choix de la direction de dérivation

Quel que soit le principe de mesure [§2.1.2], l'acquisition d'une IP possède toujours, semble-t-il, une asymétrie entre lignes et colonnes. C'est pourquoi, il est mieux d'effectuer une estimation du bruit pour chacune des deux directions de dérivation et d'utiliser comme estimation du bruit la valeur maximale des deux.

Description de l'algorithme

On aboutit finalement à l'algorithme décrit dans la figure 5.5.

5.3.4 Mise en application de l'estimation du bruit

Résultats

L'algorithme d'estimation du bruit a été appliqué aux cinq IP représentées à la figure 5.6. En dehors de la valeur de E qui constitue le résultat proprement dit de l'estimation du bruit, nous présentons pour chaque exemple

- l'histogramme de la deuxième dérivée selon les lignes [figure 5.7]
- l'écart-type de l'histogramme avant et après la troncature, ainsi que le nombre d'itérations de troncature [table 5.3]
- l'écart-type après troncature de l'histogramme de la deuxième dérivée selon les colonnes [table 5.3]
- l'image des valeurs tronquées, c'est-à-dire l'image binaire indiquant pour chaque position si la valeur de deuxième dérivée est au-dessus ou au-dessous du seuil final de troncature. Les images de la figure 5.8 présentent en fait la combinaison du résultat de troncature pour les deux directions de dérivation.

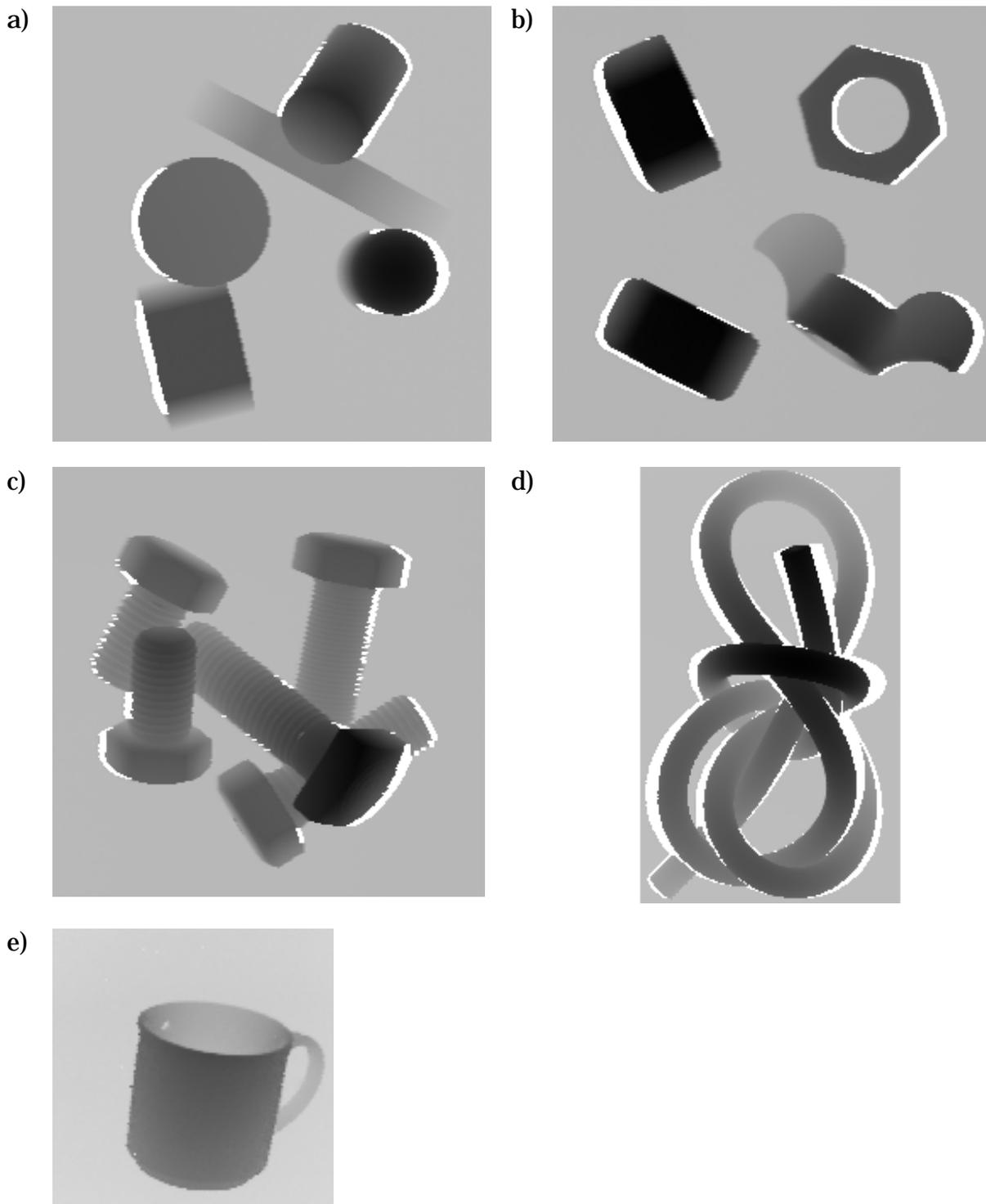


Figure 5.6: Images de profondeur. (a) BLOC12, (b) BOLT06, (c) BOLT07, (d) TUBE01, (e) "Coffee Cup B".

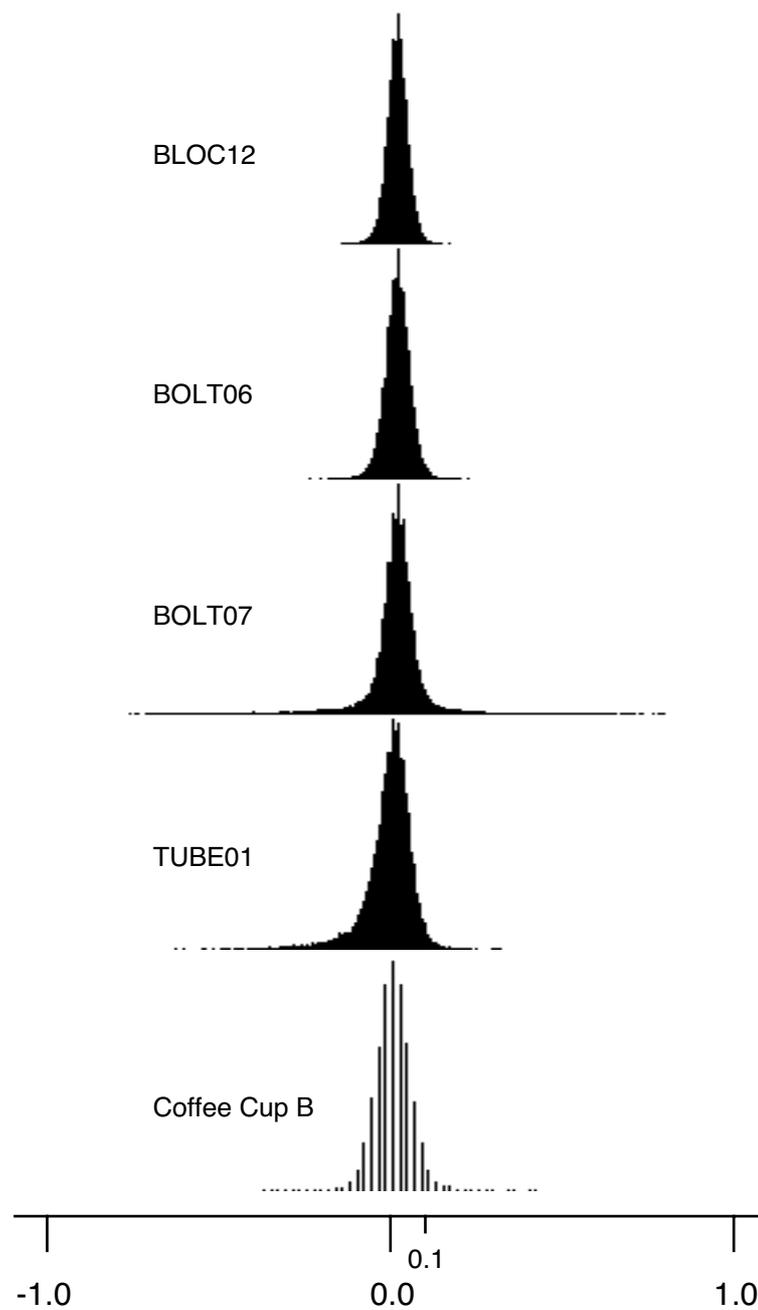


Figure 5.7: Histogrammes des images de deuxième dérivée selon les lignes. Remarque: Les valeurs extrêmes n'apparaissent pas car insuffisamment nombreuses.

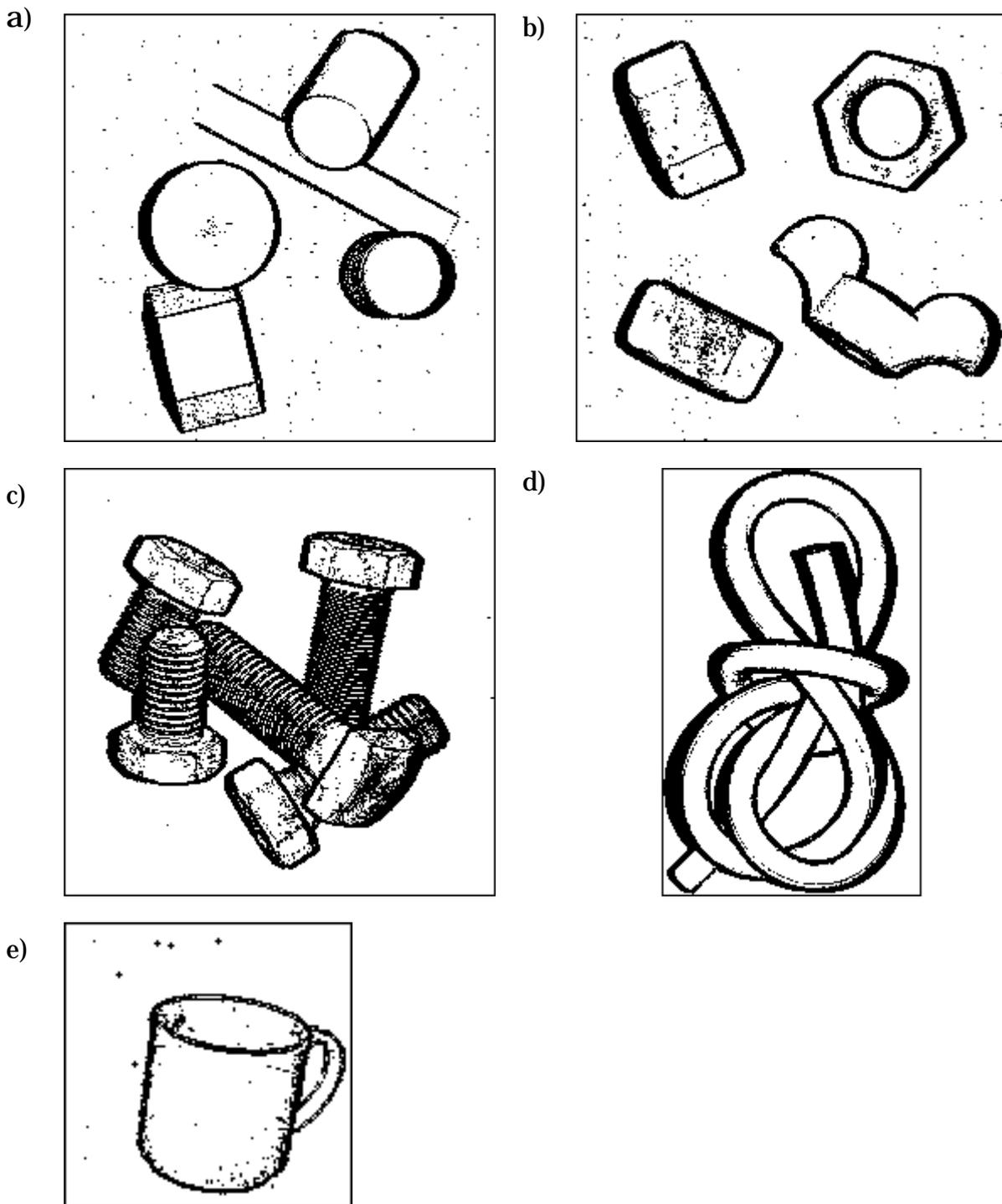


Figure 5.8: Images des valeurs tronquées. (a) BLOC12, (b) BOLT06, (c) BOLT07, (d) TUBE01, (e) "Coffee Cup B". Sont marquées en noir (1) les positions correspondant à des valeurs de la dérivée selon les lignes ou selon les colonnes situées au-dessus de leur seuil de troncature respectif, (2) les positions correspondant à des rangels de valeur *NIL*.

	$\mathbf{D}_0^2 \otimes \mathbf{P}$ (2ème dérivée selon les lignes)			$\mathbf{D}_2^2 \otimes \mathbf{P}$	E
	Convergence N	Estimée avant troncature $\frac{1}{\sqrt{6}} \sigma_0^{(0)}$	Estimée après troncature $\frac{1}{\sqrt{6}} \sigma_0$	Estimée après troncature $\frac{1}{\sqrt{6}} \sigma_2$	
BLOC12	5	0.542 mm	0.034 mm	0.032 mm	0.034 mm
BOLT06	7	0.772 mm	0.038 mm	0.035 mm	0.038 mm
BOLT07	8	0.866 mm	0.051 mm	0.046 mm	0.051 mm
TUBE01	6	3.174 mm	0.070 mm	0.071 mm	0.071 mm
Coffee Cup B	5	4.365 -	0.956 -	0.949 -	0.956 -

Table 5.3: Valeurs numériques relatives à l'estimation du bruit par notre méthode.

Commentaires sur le fonctionnement de la méthode

Les points suivants méritent d'être relevés:

- La troncature itérative de l'histogramme provoque une réduction de l'estimée du bruit d'un facteur allant jusqu'à 50 (TUBE01).
- La troncature complète est réalisée en un nombre très petit d'itérations.
- L'estimation du bruit suivant l'une ou l'autre direction ont des valeurs très voisines. Pour les systèmes ayant servi à l'acquisition des IP traitées ici, on pourrait donc se contenter d'une estimation du bruit suivant une seule direction de dérivation.

Evaluation des résultats d'estimation du bruit

Moyen d'évaluation

Sauf dans le cas d'exemples synthétiques où le bruit est connu, il n'est pas possible d'évaluer le résultat d'estimation du bruit à partir de sa valeur numérique. Toutefois, une évaluation qualitative est possible à partir de l'image des valeurs tronquées [figure 5.8]. En effet, dans le cas de l'image BLOC12, par exemple, la présence de

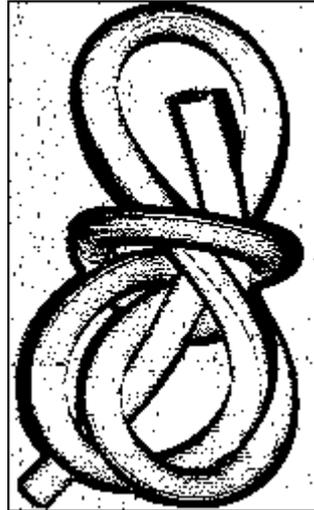


Figure 5.9: Niveau effectif du bruit de l'IP TUBE01.

points épars sur le fonds de l'image montre que le seuil de la troncature finale de l'histogramme correspond à des valeurs de bruit de grande amplitude. L'estimation du niveau du bruit est donc bonne. Dans le cas de l'image BOLT07, par contre, il n'existe pas de points épars, le niveau de bruit est surestimé.

Résultats de l'évaluation

En appliquant le moyen d'évaluation qui vient d'être cité, on obtient les résultats suivants:

- BLOC12 et "Coffee Cup B" sont des exemples de bonne estimation du bruit.
- BOLT07 et TUBE01 sont des exemples de surestimation du bruit. On constate en effet que le seuil de troncature finale n'est pas suffisamment bas. Cependant en considérant la table 5.3, on remarque qu'il y a environ qu'un facteur deux entre l'estimée de bruit maximale (celle de TUBE01) et l'estimée de BLOC12. En faisant l'hypothèse que le niveau de bruit est approximativement le même pour les deux IP, on en déduit que la surestimation ne dépasse pas un facteur deux. Un seuillage sur l'image de deuxième dérivée selon les lignes de TUBE01 avec une valeur de seuil égale à la moitié de la valeur finale de troncature confirme cette affirmation [figure 5.9]

- La particularité de BOLT06 réside dans le fait que le bruit est fortement non-stationnaire. Il y a donc localement sur- ou sous-estimation du bruit.

Conclusion

La méthode d'estimation robuste du bruit donne de bons résultats. Une légère surestimation du niveau de bruit intervient lorsqu'une grande partie de l'IP représente des surfaces de forte courbure. Dans ce cas, les valeurs élevées de la deuxième dérivée du signal de source sont trop nombreuses pour que la troncature itérative de l'histogramme soit pleinement efficace. Dans les exemples traités, la surestimation du bruit reste cependant dans des marges tolérables (facteur deux). Ce devrait être aussi le cas de toutes les IP pour lesquelles la cohérence spatiale est effective.

5.3.5 Comparaison avec les autres méthodes d'estimation du bruit

On considère les quatre méthodes regroupant les trois méthodes présentées à la section 5.2.2 ainsi que notre méthode:

- (M1) correction apportée par un moyennage des k plus proches voisins [YJAd90]
- (M2) approximation par un plan et annulation de l'erreur lorsque la différence entre erreur (signée) minimale et maximale est trop importante [JiBu94]
- (M3) approximation par un plan et annulation de l'erreur lorsque le plan est trop incliné, utilisation de la moyenne des racines des EQM locales [Besl88]
- (M4) mesure de la 2^{ème} dérivée et estimation robuste par troncature itérative de l'histogramme (notre méthode)

Principe des méthodes

Similitude

Pour l'essentiel, les méthodes se ressemblent. Leur principe commun s'exprime en trois points:

- estimation locale par mesure de la variation du signal relativement à une surface localement plane: l'estimée s'annule

	M1	M2	M3	M4
discontinuités \bar{C}_0	•	•	•	•
erreurs extraordinaires		•	•	•
discontinuités \bar{C}_1 , et/ou surface(s) de taille limitée, très courbe(s) ou avec des structures fines			•	•

Table 5.4: Robustesse des méthodes d'estimation du bruit vis-à-vis des influences parasites.

lorsque l'image de profondeur correspond à un plan échantillonné (non quantifié)

- élimination/atténuation de certaines valeurs locales de grande amplitude
- estimation globale par moyennage sur l'image

Autrement dit, toutes les méthodes sont basées sur un modèle (hypothèse) de surface mesurée plane et de bruit stationnaire et elles possèdent un moyen d'éliminer l'influence de certains écarts par rapport à ce modèle. C'est par ce moyen que les méthodes sont rendues robustes.

Différence

A notre avis, elle réside essentiellement dans la manière avec laquelle les méthodes éliminent ou atténuent les valeurs locales de grande amplitude pour le calcul de l'estimation globale. Chaque méthode utilise un procédé différent (voir la description en début de section). Relevons que l'utilisation de la moyenne des racines des EQM locales dans la méthode M3 constitue un moyen d'atténuer les valeurs de grandes amplitudes (norme L_1 au lieu de L_2).

Qualité d'estimation du bruit

Il n'est pas possible de comparer les méthodes au niveau du résultat numérique puisqu'il n'existe pas de valeur de référence (sauf dans le cas singulier d'IP synthétiques).

Notre comparaison se base sur une appréciation personnelle de l'efficacité des techniques qui sont utilisées pour rendre l'estimation

Image de profondeur	σ_χ	E (M3)	E (M4)
<u>IP acquises</u>			
Coffee Cup A	–	0.71	0.79
Coffee Cup B	–	0.84	0.96
Keyboard B	–	1.35	1.23
Polyhedron	–	1.38	1.52
Ring on Polyhedron	–	1.61	1.76
Curved Surface A	–	1.68	2.15
Road Scene A	–	0.72	0.71
Road Scene B	–	0.79	0.71
Auto Part	–	0.43	0.56
<u>IP synthétiques</u>			
Block A	2.3	1.82	2.27
Block B	7.5	5.93	7.78
Light Bulb	3.4	3.45	3.32

Table 5.5: Comparaison de résultats d'estimation du bruit. Les IP sont de P.J. Besl. Les IP synthétiques contiennent un bruit pseudo-aléatoire de distribution gaussienne dont l'écart-type est représenté dans la colonne correspondante. Les résultats de la méthode M3 sont tirés de [Besl88].

du bruit robuste. La table 5.4 montre le résultat de cette appréciation par rapport à 4 situations typiques où le modèle de surface plane et de bruit stationnaire n'est pas vérifié. Selon cette analyse, notre méthode (M4) possède avec la méthode M3 les meilleures caractéristiques.

La table 5.5 présente les résultats obtenus par les méthodes M3 et M4 sur un ensemble composé d'IP acquises et d'IP synthétiques. On note avec satisfaction que les valeurs d'estimation sont très proches.

Le cas des IP synthétiques permet de comparer l'estimée avec la valeur exacte. On remarque que l'estimation faite par notre méthode est excellente. Sans aucun doute, la raison principale en est que le générateur des IP synthétiques et notre méthode utilisent le même modèle de bruit. Toutefois, ce résultat ne serait pas obtenu si l'élimination des observations aberrantes mise en oeuvre par notre méthode n'était pas efficace.

	M1	M2	M3	M4
coût de calcul	**	***	***	*
nombre de paramètres	0	1	1	0

Table 5.6: Complexité des méthodes d'estimation du bruit. Remarque: La valeur nulle pour le nombre de paramètres de M4 (notre méthode) tient compte du fait que le choix $k = 3$ s'impose quasiment.

Algorithmique

La méthode M4 est sans doute la méthode la moins gourmande en puissance de calcul, surtout si l'estimation de la deuxième dérivée est faite dans une seule direction. D'autre part, pour le seul paramètre que la méthode utilise, il existe une valeur qui s'impose quasiment. Cette valeur est indépendante de la mise à l'échelle de l'IP, ce qui n'est pas toujours le cas pour les autres méthodes (par exemple, les seuils utilisés pour éliminer certaines erreurs d'approximation locale par un plan). Ces différents points donnent la faveur à la méthode M4 [table 5.6].

5.3.6 Extension aux images de profondeur de type vectoriel

Nous proposons d'étendre notre méthode d'estimation robuste du bruit aux IP de type vectoriel, en remplaçant simplement la deuxième dérivée par la norme du vecteur qui résulte de l'application de l'opérateur de deuxième dérivée à l'IP vectorielle

$$\mathbf{Q} = \|\mathbf{D}_j^2 \otimes \mathbf{P}\| \quad (5.25)$$

Cette proposition n'a pas encore été expérimentée.

5.4 Conclusion

Ce chapitre a été dédié à l'estimation du bruit d'une image de profondeur.

Estimation du bruit

Tout d'abord, nous avons souligné que l'erreur de mesure est interprétée comme un bruit dans une optique de traitement du signal. Nous avons d'autre part montré qu'un modèle de bruit additif correspond à une interprétation de l'erreur de mesure qui est différente pour chaque type d'IP. Dans le cas d'IP scalaires, il s'agit d'une erreur de distance par rapport à la surface mesurée.

Ensuite, nous avons mis en évidence que l'estimation du bruit repose sur des hypothèses concernant la surface mesurée et le bruit. Nous avons alors étudié plusieurs méthodes d'estimation du bruit. Toutes font essentiellement les mêmes hypothèses, représentées par le modèle d'une surface mesurée plane et d'un bruit additif stationnaire. D'autre part, toutes mettent en oeuvre des mécanismes d'estimation du bruit robuste vis-à-vis de certains écarts par rapport à ce modèle.

Méthode proposée

Nous avons proposé une nouvelle méthode d'estimation du bruit qui, dans le principe, ressemble aux autres. Dans cette méthode, l'estimée du bruit est la valeur efficace du signal de deuxième dérivée, calculée après une élimination adaptative des valeurs aberrantes. La méthode est simple, d'une part parce qu'elle ne présente pas de paramètre (seuil automatique) et, d'autre part, parce qu'elle est composée uniquement de quelques opérations élémentaires. La mise en application de la méthode sur plusieurs exemples a fourni de bons résultats.

Comparée aux autres méthodes, elle est des plus performantes tout en étant des plus simples.

Finalement, cette méthode d'estimation est bien adaptée à notre contexte. Nous allons montrer au chapitre 7 qu'elle se combine parfaitement avec notre méthode de détection des discontinuités présentée à la fin du chapitre suivant.

Chapitre 6

Lissage adapté à la détection des discontinuités

Ce chapitre est consacré à l'atténuation du bruit d'une image de profondeur par des méthodes de lissage capables de préserver les discontinuités. Ces méthodes de lissage correspondent à des méthodes numériques associées à la solution de problèmes physiques.

Dans une première partie, le lissage est envisagé sans la préservation des discontinuités. Les analogies physiques sont présentées et l'effet de lissage est démontré pour les trois méthodes de base: diffusion, minimisation d'énergie de membrane et minimisation d'énergie de plaque mince. La présentation est faite dans le cas continu.

Une deuxième partie présente le cas discret et les méthodes numériques associées à chacune des méthodes de lissage. L'attention se focalise sur les méthodes itératives locales. La complexité et les possibilités de parallélisation de ces méthodes sont discutées.

La troisième partie montre d'abord comment introduire un moyen de préserver les discontinuités dans le procédé de lissage. Elle met en évidence le lien étroit qui existe entre préservation et détection des discontinuités. Ensuite, les différentes approches utilisées pour préserver les discontinuités d'ordre 0 et d'ordre 1 d'une image de profondeur sont passées en revue.

Une dernière partie présente une nouvelle méthode de lissage avec détection des discontinuités. Ses avantages par rapport aux autres méthodes sont mis en évidence. Diverses extensions de la méthode sont proposées.

6.1 Introduction

Les opérateurs de voisinage servant à détecter les discontinuités d'une IP sont sensibles au bruit [chap.4]. Pour obtenir de bons résultats de détection, une atténuation du bruit doit être envisagée.

Le principe général appliqué pour l'atténuation du bruit est celui d'une opération locale de moyennage. L'importance de l'atténuation est dictée par la taille du voisinage. Plus le voisinage est grand, plus l'atténuation est importante. Pour garantir de bonnes caractéristiques de "douceur" du signal qui a été traité pour l'atténuation du bruit, le moyennage prend la forme d'un lissage décrit à la section suivante [§6.2].

Le lissage provoque une diminution de résolution du signal. Les discontinuités, qui correspondent à des variations locales du signal, ont tendance à disparaître, d'où l'intérêt pour des méthodes de lissage adaptées à la détection des discontinuités [§6.3-5].

6.2 Méthodes de lissage linéaire à analogie physique: principe

Dans cette section, trois méthodes de lissage linéaire sont exposées. Les mathématiques y relatives correspondent à des modèles de phénomènes physiques. D'où les dénominations utilisées: lissage par diffusion, par minimisation d'énergie de membrane et par minimisation d'énergie de plaque mince.

Pour simplifier la présentation, on considère le cas continu. La fonction $f(u,v)$, dont le domaine de définition est \mathfrak{R}^2 dans son intégralité, représente le signal qui doit être lissé. $\tilde{f}(u,v)$ est le signal lissé.

La section est organisée de la manière suivante. D'abord nous présentons le lissage par diffusion [§6.2.1], puis les méthodes de lissage par minimisation d'énergie [§6.2.2]. Après comparaison [§6.2.3], l'application des trois méthodes aux IP est discutée [§6.2.4-5]. Suit une conclusion [§6.2.6]

6.2.1 Lissage par diffusion (D)

La fonction \tilde{f} est déterminée à partir de la solution d'un problème aux valeurs initiales pour la fonction $F(u, v, t)$. Ce problème est défini par

(1) l'équation différentielle

$$\frac{\partial}{\partial t} F(u, v, t) = \Delta F(u, v, t) \quad (6.1)$$

où ΔF représente le laplacien de la fonction F

(2) la condition initiale d'égalité à f

$$F(u, v, 0) = f(u, v) \quad (6.2)$$

\tilde{f} est la valeur de la fonction F pour une valeur constante de la variable indépendante t égale à $\lambda_D^2/2$

$$\tilde{f}(u, v) = F(u, v, \lambda_D^2/2) \quad (6.3)$$

Paramètre de lissage

λ_D^2 détermine l'importance du lissage (cette affirmation est confirmée plus loin, lors de la présentation de la fonction de Green). Plus la valeur de λ_D^2 est élevée, plus l'éloignement (suivant l'axe t) de \tilde{f} par rapport à f est grand et plus le lissage est important.

Analogie physique

L'équation différentielle (6.1) modélise la diffusion de la chaleur dans un espace bidimensionnel pour lequel le coefficient de conduction thermique vaut 1. F représente la répartition de température en fonction du temps t . f et \tilde{f} sont les répartitions de température aux temps 0 et $\lambda_D^2/2$, respectivement.

Fonction de Green

La fonction de Green $w_D(u, v)$ de l'équation de diffusion de la chaleur dans \mathfrak{R}^2 amène une relation simple entre f et \tilde{f} [BrSe91]: \tilde{f} s'obtient par convolution de f avec w_D

$$\tilde{f}(u, v) = w_D(u, v) * f(u, v) \quad (6.4)$$

La fonction w_D est en l'occurrence une gaussienne

$$w_D(u, v) = \frac{1}{2\pi\lambda_D^2} e^{-\frac{u^2 + v^2}{2\lambda_D^2}} \quad (6.5)$$

dont l'écart-type vaut λ_D .

Visiblement, λ_D détermine l'importance du lissage, confirmant ainsi ce qui avait été indiqué précédemment.

Confinement à un domaine borné de \mathfrak{R}^2 : conditions de bord

La restriction de la diffusion à un domaine borné D_0 de \mathfrak{R}^2 contraint à la spécification de conditions de bord, sinon le problème n'est pas posé correctement et \tilde{f} n'est pas définie.

Spécifier les conditions de bord consiste à fixer la valeur de F et/ou d'une ou de plusieurs de ses différentielles sur la courbe Γ constituant le bord du domaine D_0 . Mais le problème de diffusion est correctement posé seulement pour certaines conditions de bord. En l'occurrence, il faut fixer sur Γ la valeur de F ou celle de $\frac{\partial}{\partial n} F$, ou encore une combinaison des deux (par exemple linéaire [BrSe91: p.480]).

Le fait de restreindre la diffusion à un domaine borné supprime l'équivalence entre diffusion et convolution avec une gaussienne. La fonction de Green du problème de diffusion dans le domaine borné varie en chaque point. Elle dépend de la forme du bord du domaine et des conditions qui y sont spécifiées [BrSe91].

6.2.2 Lissage par minimisation d'énergie (E1, E2)

La fonction \tilde{f} est la solution d'un problème de variation

$$J[\tilde{f}] = \min_f J[\bar{f}] \quad (6.6)$$

où la fonctionnelle $J[\bar{f}]$ est une somme pondérée de deux termes

$$J[\bar{f}] = E[\bar{f}, f] + \gamma S[\bar{f}] \quad (6.7)$$

le premier, $E[\bar{f}, f]$, mesurant l'écart de \bar{f} par rapport à f

$$E[\bar{f}, f] = \int_{\mathfrak{R}^2} (\bar{f} - f)^2 dudv \quad (6.8)$$

le deuxième, $S[\bar{f}]$, étant une mesure de la déformation de \bar{f}

$$S[\bar{f}] = \int_{\mathfrak{R}^2} \phi\langle \bar{f} \rangle dudv \quad (6.9)$$

où ϕ est un opérateur différentiel mesurant la déformation locale.

Mesures de déformation locale

Les opérateurs différentiels ϕ les plus fréquemment utilisés pour mesurer la déformation locale sont la norme du gradient

$$\phi_1\langle \bar{f} \rangle = \bar{f}_u^2 + \bar{f}_v^2 \quad (6.10.a)$$

et la variation quadratique

$$\phi_2\langle \bar{f} \rangle = \bar{f}_{uu}^2 + 2\bar{f}_{uv}^2 + \bar{f}_{vv}^2 \quad (6.10.b)$$

Ces mesures ont l'avantage d'être plus simples que d'autres mesures utilisables [BIZi87]. Nous nous limitons ici à ces deux cas. Chacun définit une méthode de lissage différente, auxquelles nous faisons référence par les symboles E1 et E2, les nombres 1 et 2 indiquant l'ordre de la différentielle correspondante.

Paramètre de lissage

En déterminant l'importance relative des deux fonctionnelles E et S dans l'équation (6.7), le facteur de pondération γ détermine l'importance du lissage. En effet, plus γ est grand, plus on accentue l'importance de la déformation relativement à l'erreur d'approximation et donc plus le lissage est effectif sur la solution \tilde{f} du problème de variation.

Afin d'obtenir des paramètres de lissage qui soient comparables pour les méthodes E1 et E2, γ est remplacé par les grandeurs respectives λ_{E1}

$$\gamma = \lambda_{E1}^2 \quad (6.11.a)$$

et λ_{E2}

$$\gamma = \lambda_{E2}^4 \quad (6.11.b)$$

λ_{E1} et λ_{E2} possèdent les mêmes unités que les variables indépendantes u et v .

Analogie physique

\tilde{f} représente la position d'équilibre d'un corps mince. f représente une forme rigide et fixe à laquelle est attachée le corps mince par un volume élastique. S est l'énergie de déformation du corps mince, E est l'énergie de déformation du volume élastique.

Le carré de la norme du gradient (6.10.a) représente l'énergie de déformation locale d'une membrane à condition que certaines conditions de déformation maximale soient remplies [CoHi68]. Sous

les mêmes réserves, la variation quadratique (6.10.b) représente l'énergie de déformation d'une plaque mince. Par conséquent, nous appelons les fonctionnelles respectives, fonctionnelle de membrane et fonctionnelle de plaque mince.

Équations d'Euler

L'équation d'Euler du problème de variation défini par les équations (6.6-9) est, dans le cas où la déformation est mesurée par la fonctionnelle de membrane,

$$\tilde{f} - \lambda_{E1}^2 \Delta \tilde{f} = f \quad (6.12)$$

où $\Delta \tilde{f}$ symbolise le laplacien de la fonction \tilde{f} [BlZi87]. L'équation d'Euler pour la fonctionnelle de plaque mince est

$$\tilde{f} + \lambda_{E2}^4 \Delta^2 \tilde{f} = f \quad (6.13)$$

où $\Delta^2 \tilde{f}$ représente le bilaplacien de \tilde{f} (laplacien du laplacien).

Fonctions de Green

Les fonctions de Green pour les équations différentielles (6.12) et (6.13) dans \mathfrak{R}^2 sont symétriques en rotation et invariantes en translation [BlZi87]. Le procédé de lissage par minimisation d'énergie peut donc être réalisé par convolution avec la fonction de Green.

Confinement à un domaine borné de \mathfrak{R}^2 : conditions de bords

L'application du problème de variation (6.6) à un domaine borné D_0 pose la question des conditions de bord.

Contrairement au cas de la diffusion, le problème est correctement posé sans conditions de bord. La solution \tilde{f} vérifie certaines propriétés sur le bord Γ du domaine, appelées conditions de bord naturelles. Il est nécessaire de joindre ces conditions aux équations d'Euler, pour obtenir un problème équivalent au problème de variation. Les conditions de bord naturelles pour la fonctionnelle de membrane fixent à zéro la valeur de la dérivée dans la direction normale au bord

$$\frac{\partial}{\partial n} \tilde{f} = 0 \quad (6.14)$$

Pour la fonctionnelle de plaque mince, les conditions de bord naturelles sont les suivantes

$$\frac{\partial^2}{\partial n^2} \tilde{f} = 0 \quad (6.15)$$

	Diffusion	Minimisation d'énergie (ordre 1)	Minimisation d'énergie (ordre 2)
a)	$\frac{\partial}{\partial t} F = \Delta F$ $f(u, v) = F(u, v, 0)$ $\tilde{f}(u, v) = F(u, v, \frac{\lambda_D^2}{2})$	$\tilde{f} - \lambda_{E1}^2 \Delta \tilde{f} = f$	$\tilde{f} + \lambda_{E2}^4 \Delta^2 \tilde{f} = f$
b)	$\left(F, \frac{\partial}{\partial n} F \right)$	$\left(\tilde{f}, \frac{\partial}{\partial n} \tilde{f} \right)$	(\dots)
c)	$w_d(\xi) = \frac{1}{\sqrt{2\pi\lambda_d}} e^{-\frac{\xi^2}{2\lambda_d^2}}$	$w_{e1}(\xi) = \frac{1}{2\lambda_{e1}} e^{-\frac{ \xi }{\lambda_{e1}}}$	$w_{e2}(\xi) = \frac{1}{2\lambda_{e2}} e^{-\frac{ \xi }{\sqrt{2\lambda_{e2}}}} \cos\left(\frac{ \xi }{\sqrt{2\lambda_{e2}}} - \frac{\pi}{4}\right)$

Table 6.1: Comparaison formelle des méthodes de lissage à analogie physique. (a) Équations différentielles 2D. (b) Conditions de bord 2D. (c) Fonctions de Green 1D.

et

$$\frac{\partial}{\partial n} \left(\Delta \tilde{f} + \frac{\partial^2}{\partial s^2} \tilde{f} \right) = 0 \quad (6.16)$$

où $\frac{\partial}{\partial n}$ et $\frac{\partial}{\partial s}$ désignent, respectivement, l'opération de dérivée dans la direction normale et dans la direction tangentielle au bord.

Il est aussi possible d'imposer des conditions de bord avec le problème de variation (6.6). Mais le problème qui en résulte est correctement posé seulement pour certaines conditions de bord. Dans le cas de la fonctionnelle de membrane, on ne peut que spécifier la valeur de la fonction \tilde{f} ou de sa dérivée $\frac{\partial}{\partial n} \tilde{f}$ dans la direction normale au bord. Pour la fonctionnelle de plaque mince, les possibilités sont plus nombreuses. On peut par exemple spécifier à la fois \tilde{f} et $\frac{\partial}{\partial n} \tilde{f}$.

Les fonctions de Green pour le problème de variation (6.6) sur le domaine borné D_0 varient en fonction de la localisation. Elles dépendent de la forme du bord du domaine et des conditions qui y sont spécifiées.

6.2.3 Comparaison

La table 6.1 et la figure 6.1 mettent en parallèle les méthodes que nous venons de présenter. La table 6.1 contient les équations

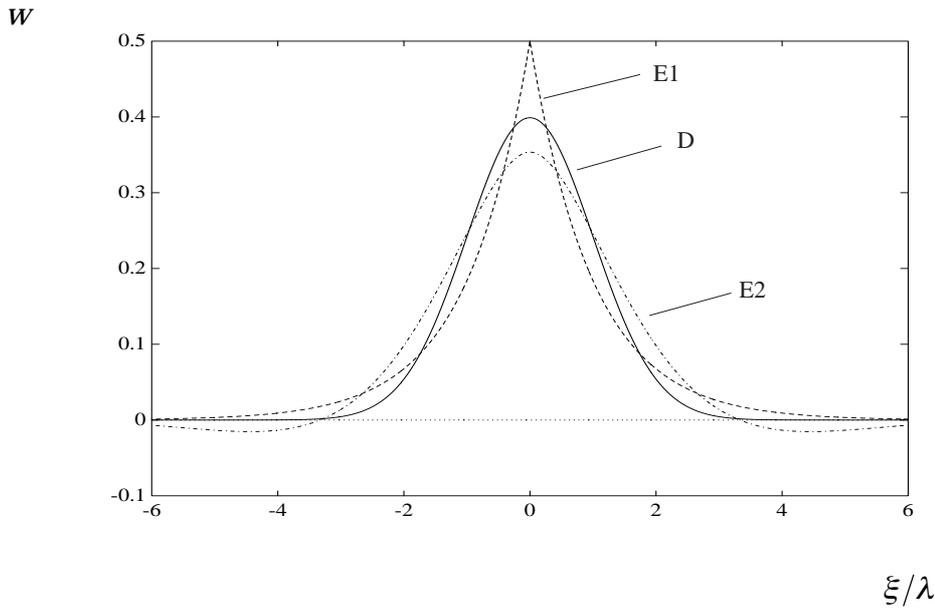


Figure 6.1: Méthodes de lissage unidimensionnel à analogie physique: fonctions de Green. Diffusion D et minimisation d'énergie E1 et E2.

différentielles et les conditions de bord pour le cas bidimensionnel, auxquelles s'ajoutent les fonctions de Green pour le cas d'un domaine infini unidimensionnel (\mathfrak{R}). La figure 6.1 offre une représentation graphique des trois fonctions de Green unidimensionnelles pour une même valeur de leur paramètre respectif λ .

La comparaison des méthodes amène les constatations suivantes:

- (1) Dans chaque méthode, l'importance du lissage est déterminée par un seul paramètre, pour lequel nous avons utilisé un symbole commun λ .
- (2) Pour le cas d'un domaine infini, le lissage peut être effectué par convolution avec la fonction de Green. La fonction de Green varie d'une méthode à l'autre et varie en fonction de λ . Pour une même valeur de λ , les fonctions sont néanmoins très similaires.
- (3) Pour le cas d'un domaine borné, les méthodes D et E1 offrent les mêmes possibilités concernant les conditions de bord: on peut fixer soit la valeur de la fonction lissée \tilde{f} , soit la valeur de sa dérivée dans la direction normale, ou encore une combinaison des deux, mais il n'y a pas d'autres possibilités. La méthode E2 laisse plus de liberté à la fonction \tilde{f} : il est possible de spécifier par exemple uniquement des dérivées d'ordre supérieur. On

remarque cependant que, quelle que soit la méthode, le comportement du lissage est le même dans un domaine borné que dans un domaine infini, lorsque l'éloignement par rapport aux bords est nettement supérieur à λ .

- (4) Que le domaine soit borné ou infini, toutes les méthodes atténuent les variations de signal qui interviennent dans un intervalle de taille inférieure à λ .

6.2.4 Application aux images de profondeur: problématique

Particularité d'une image de profondeur (IP)

Le signal correspondant à une IP possède un domaine de définition de taille finie, à cause de la taille limitée du tableau. De plus, ce domaine n'a pas toujours la forme régulière définie par le tableau, à cause des valeurs *NIL*.

Possibilités de mise en application des méthodes de lissage

Il y a deux alternatives principales:

- (1) Tout d'abord, on peut utiliser les méthodes de lissage pour un domaine infini après avoir étendu le domaine de définition de l'IP. L'extension du domaine de définition peut être faite en répétant périodiquement le signal ou en fixant une valeur arbitraire en dehors du domaine initial. En présence de valeurs *NIL*, il faut d'abord compléter le tableau, par exemple par interpolation.
- (2) L'alternative est d'utiliser les méthodes de lissage pour un domaine borné en fixant des conditions de bord.

Critère de lissage optimal

Afin d'évaluer les différentes possibilités de lissage, il faut définir un critère. Si l'on fait abstraction des problèmes de mise en oeuvre du lissage, le seul critère relevant est celui de qualité du résultat. Il est défini par le but du lissage: de manière idéale, le lissage devrait atténuer le bruit sans endommager le signal de source. Mais ce voeu, pris strictement, est en général irréalisable.

Pour un domaine infini, il existe au moins un cas où il est possible d'obtenir une atténuation du bruit sans déformation du

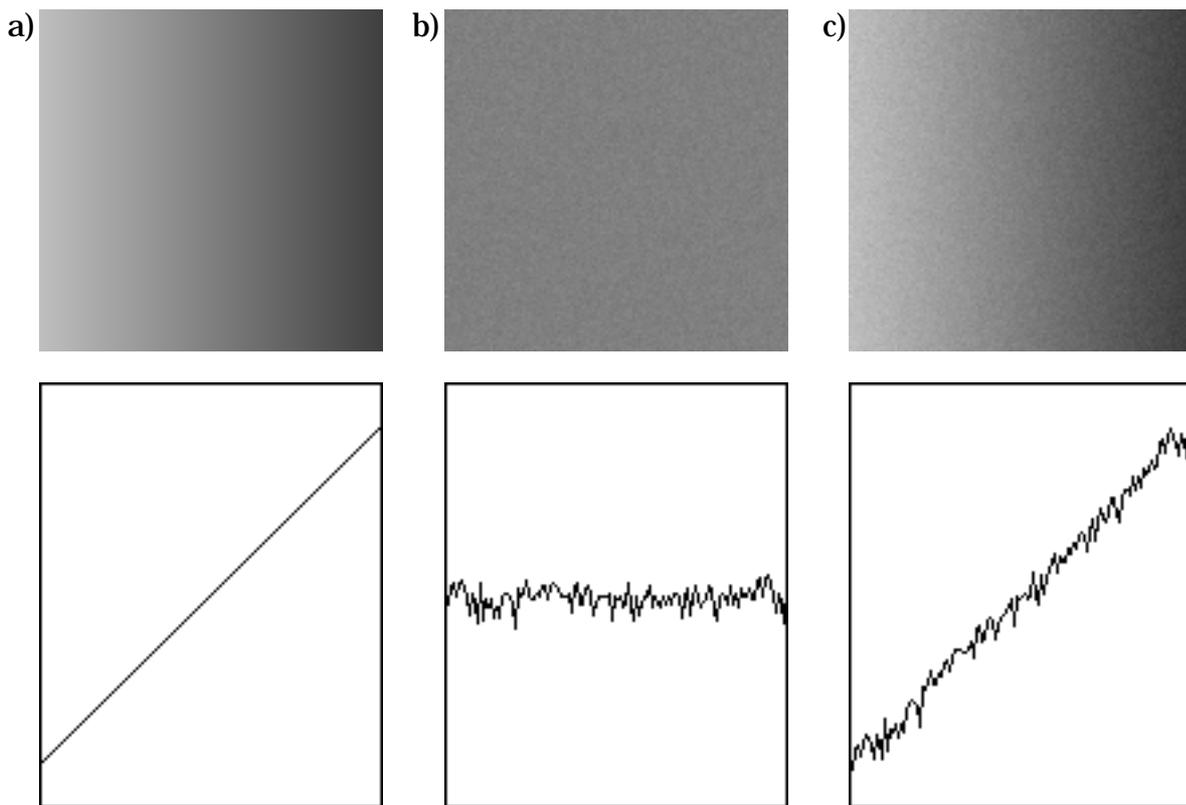


Figure 6.2: IP synthétique représentant un plan incliné bruité. Haut: image codée par des niveaux de gris. Bas: profil de l'image suivant une ligne. (a) Signal de source. (b) Bruit. (c) IP synthétique 128x128x8bits constituée par la somme des signaux a) et b).

signal. C'est la situation suivante: le signal de source est un plan et le bruit est additif et de valeur moyenne nulle. Dans ce cas, le signal lissé converge vers le signal de source lorsque le paramètre λ tend vers l'infini (pour s'en persuader, il suffit de considérer la convolution avec la fonction de Green).

Un lissage optimal de l'IP devrait garantir le même comportement pour une portion de plan (domaine borné), quel que soit le plan.

Illustration du comportement du lissage

Le comportement du lissage dans les différentes variantes possibles (domaine infini ou borné, conditions de bord) est illustré en appliquant la méthode D à l'IP synthétique représentée à la figure 6.2.

L'IP traitée est constituée d'un plan incliné, bruité de manière additive. Le plan est incliné à 45° afin d'éviter le bruit de

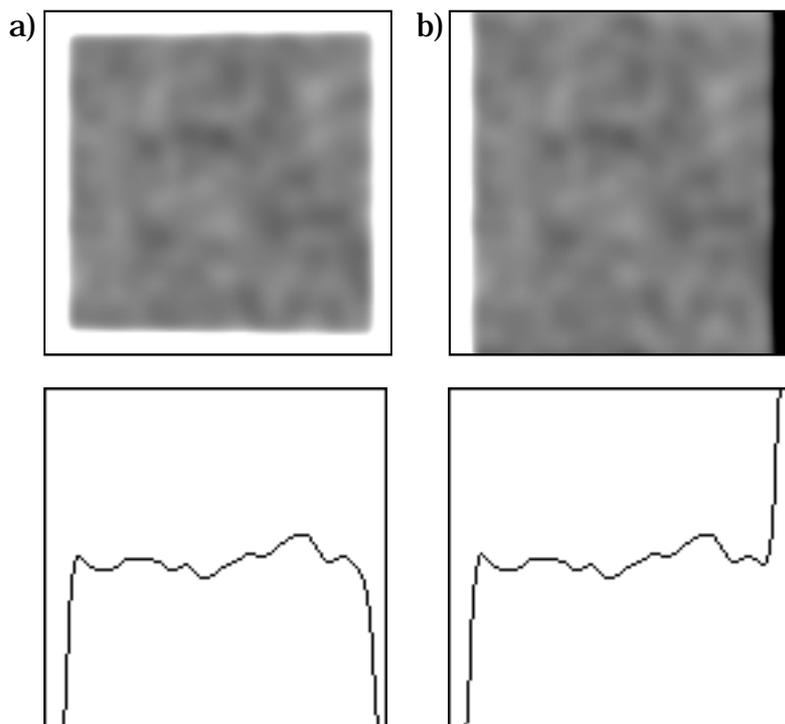


Figure 6.3: Bruit résiduel dans le cas d'une extension du signal. (a) Valeur fixe à l'extérieur du domaine initial. (b) Répétition périodique du signal.

quantification. Le bruit est pseudo-aléatoire de distribution gaussienne discrète dont l'écart-type vaut 4. La méthode D a été appliquée avec une valeur de paramètre $\lambda = 4$. Les résultats de lissage [figures 6.3-5] sont représentés par l'image de bruit résiduel, c'est-à-dire: signal lissé moins signal de source.

Remarque: Dans les figures 6.2 à 6.5, les images de bruit résiduel sont représentées par codage en niveaux de gris, la valeur 0 étant codée par un gris moyen. Un profil complète la représentation. Sauf autres indications, le profil correspond à la ligne du milieu.

6.2.5 Application aux images de profondeur: lissage optimal

Le critère de lissage optimal permet d'évaluer les différentes possibilités de mise en application des méthodes de lissage.

Lissage sur un domaine borné ou sur un domaine infini

De manière évidente, l'extension de l'IP à un domaine infini, quelle que soit la manière de la réaliser, ne permet pas de satisfaire le

critère de lissage optimal (on suppose bien sûr que le plan ne soit pas connu a priori). La figure 6.3 en apporte une illustration.

L'autre solution, c'est-à-dire l'application des méthodes de lissage pour un domaine borné, pose le problème de la fixation des conditions de bord. Celles-ci devraient être choisies pour satisfaire le critère de lissage optimal. Mais le choix n'est pas entièrement libre: il faut que le problème mathématique associé à la méthode de lissage soit posé correctement.

Conditions de bord pour E2

Dans le cas de la méthode E2, les conditions de bord peuvent être choisies de sorte à remplir parfaitement le critère d'optimalité qui a été défini. Il suffit d'utiliser les conditions de bord naturelles du problème de variation.

En effet, les conditions (6.15) et (6.16) sont vérifiées dans le cas où \tilde{f} représente un plan. Par conséquent, un signal plan est invariant au lissage et, dans le cas d'un plan bruité de manière additive, le bruit est atténué par le lissage sans que le plan soit distordu.

Conditions de bord pour E1 et D

Le cas des deux autres méthodes ne peut pas être résolu de manière satisfaisante. En effet, pour que le problème lié au lissage soit posé correctement, il faut fixer, sur le bord, les valeurs de la fonction \tilde{f} ou de sa dérivée dans la direction normale au bord. Par conséquent, un plan dont les valeurs au bord ne correspondent pas à ce qui a été fixé n'est pas invariant au lissage. Il n'existe donc pas des conditions de bord telles que le critère d'optimalité soit parfaitement rempli. Il faut se contenter d'une solution sous-optimale.

Solutions habituelles

Elles reposent sur une des deux alternatives suivantes, qui sont illustrées par des exemples pour le cas de la méthode D:

- fixer les conditions de bord à partir d'une valeur constante arbitraire. Par exemple, Ponce et Brady utilisent les conditions de bord adiabatiques (annulation de la dérivée dans la direction normale) [PoBr87]. Cai fixe le coefficient de conduction thermique à une valeur constante [Cai89]. La figure 6.4.b présente le cas des conditions de bord adiabatiques.

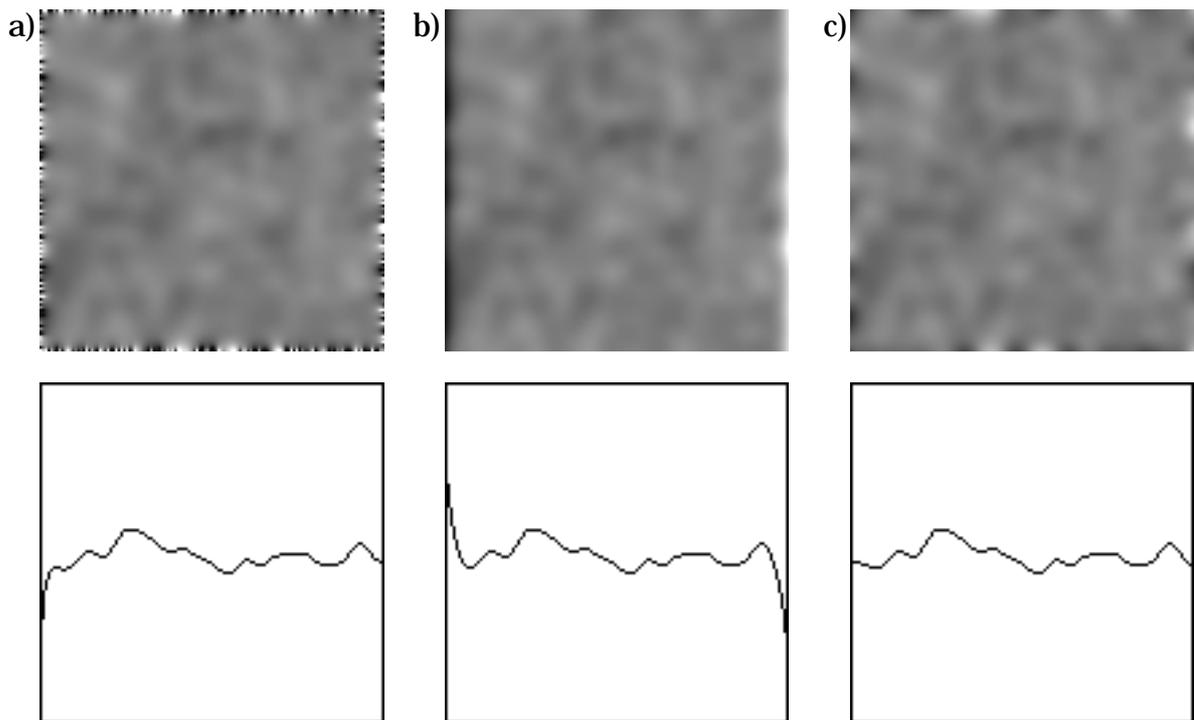


Figure 6.4: Bruit résiduel pour différentes conditions de bord: comportement dans la direction normale au bord. (a) Bord fixé aux valeurs du signal non lissé. (b) Conditions adiabatiques. (c) Bord lissé indépendamment.

- fixer les conditions de bord à partir des valeurs du signal non lissé f . Par exemple, Trucco et Fisher fixent le coefficient de conduction thermique en fonction de f [TrFi92]. La figure 6.4.a montre le cas de condition d'égalité avec les valeurs du signal non lissé.

Chacune des alternatives a son inconvénient. Dans le premier cas, on ne peut pas éviter de déformer le signal dans la direction normale au bord. La deuxième solution empêche le lissage dans la direction tangentielle au bord.

Autre solution

La meilleure réponse que nous ayons trouvée au problème consiste à découpler le lissage dans les directions tangentielle et normale au bord. Ce découplage est réalisé de la manière suivante: le bord est traité séparément en tant que domaine unidimensionnel et la fonction unidimensionnelle lissée résultante est imposée comme valeurs de bord pour le signal bidimensionnel lissé \tilde{f} . Voici, pour chacune des

méthodes D et E1, la formulation mathématique de ces nouvelles conditions de bord.

Conditions de bord pour D

Elles sont imposées par la valeur de la fonction $G(s, t)$

$$F(u, v, t) = G(s, t) \quad (6.17)$$

définie sur Γ par l'équation différentielle de diffusion à une dimension

$$\frac{\partial}{\partial t} G(s, t) = \frac{1}{2} \frac{\partial^2}{\partial s^2} G(s, t) \quad (6.18)$$

et la condition initiale d'égalité avec la fonction f

$$G(s, 0) = f(u, v) \quad (6.19)$$

Conditions de bord pour E1

Elles sont imposées par la valeur de la fonction \tilde{g}

$$\tilde{f}(u, v) = \tilde{g}(s) \quad (6.20)$$

définie sur Γ par le problème de variation

$$J_{\Gamma}[\tilde{g}] = \min_{\tilde{g}} J_{\Gamma}[\tilde{g}] \quad (6.21)$$

avec

$$J_{\Gamma}[\tilde{g}] = \int_{\Gamma} \left((\tilde{g} - g)^2 + \lambda_{E1}^2 \left(\frac{d}{ds} \tilde{g} \right)^2 \right) ds \quad (6.22)$$

où la fonction $g(s)$ est définie par l'égalité aux valeurs de f sur Γ

$$g(s) = f(u, v) \quad (6.23)$$

L'équation d'Euler du problème de variation (6.22)

$$\tilde{g} - \lambda_{E1}^2 \frac{d^2}{ds^2} \tilde{g} = g \quad (6.24)$$

définit un problème équivalent.

Comportement des nouvelles conditions de bord

Les figures 6.4.c et 6.5.c montrent que les conditions de bord que nous proposons permettent un lissage longitudinal au bord sans déformation transversale du signal de source (plan).

On remarque cependant qu'une déformation du signal dans la direction transversale intervient localement, dans les coins du domaine. L'explication en est la suivante. Tant que le bord est rectiligne, le signal unidimensionnel, défini par la valeur du signal de source sur le bord, est une fonction affine. Il est donc invariant au

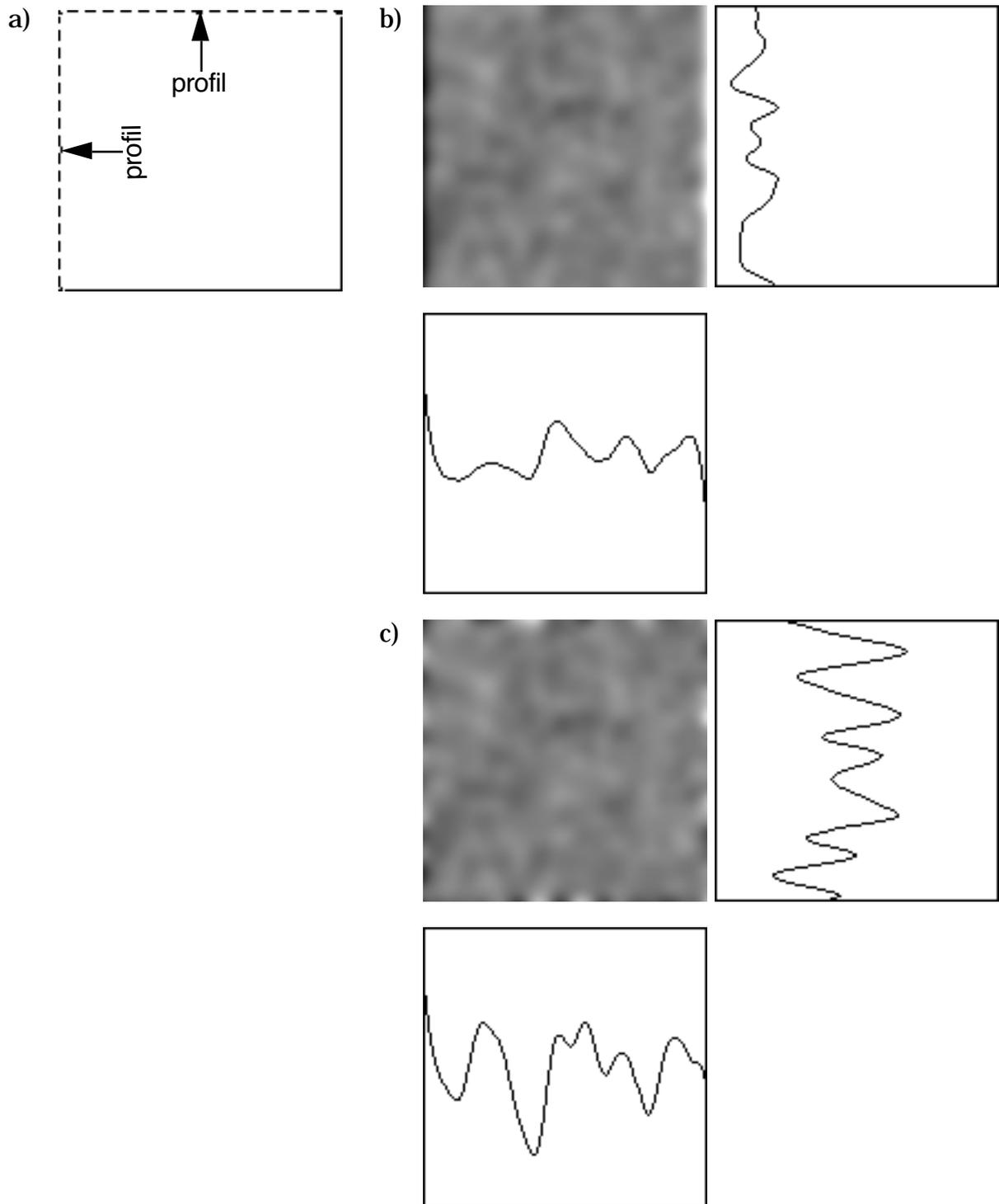


Figure 6.5: Bruit résiduel pour différentes conditions de bord: comportement dans la direction tangentielle au bord. (a) Choix des profils. (b) Conditions adiabatiques. On remarque l'erreur importante aux bords gauche et droite, là où la pente du signal dans la direction perpendiculaire au bord est importante. (c) Bord lissé indépendamment. On remarque une moins bonne atténuation du bruit sur les bords. On remarque aussi une erreur importante dans les coins, due au fait que le coin correspond à un changement de pente du signal unidimensionnel défini par le bord.

lissage. Cette propriété n'est plus vérifiée dès que le bord forme une courbe. La déformation du signal par le lissage est d'autant plus élevée que la courbe est prononcée.

On remarque d'autre part que le lissage est moins efficace sur le bord.

Nous concluons de l'analyse du comportement de ces nouvelles conditions de bord qu'elles offrent une solution plus proche de la solution optimale que le font les approches habituelles.

6.2.6 Récapitulation et conclusion

Dans cette section, trois méthodes de lissage ont été présentées: lissage par diffusion (D), par minimisation d'énergie de membrane (E1) et par minimisation d'énergie de plaque mince (E2).

Une comparaison des méthodes a mis en évidence qu'elles sont essentiellement similaires. Pour chaque méthode, l'importance du lissage est déterminée par un unique paramètre noté λ . Pour des valeurs identiques du paramètre λ , les méthodes réalisent un lissage comparable. Elles atténuent les variations du signal qui interviennent sur un intervalle de taille inférieure à λ .

Ensuite, la perspective d'application des méthodes aux IP a été étudiée. Un critère d'optimalité a été défini pour évaluer les différentes possibilités de mise en application. Il stipule qu'une IP constituée d'une portion de plan, quelle qu'elle soit, devrait être invariante au lissage. Avec ce critère d'optimalité, les possibilités d'application des méthodes de lissage aux IP ont été évaluées comme suit:

- La solution consistant à étendre l'IP au-delà des limites pour pouvoir effectuer un lissage sur un domaine de taille infinie ne permet pas un lissage optimal.
- La méthode E2 appliquée sur le domaine de définition de l'IP avec les conditions de bord naturelles vérifie le critère d'optimalité.
- Pour les méthodes E1 et D, appliquées sur le domaine de définition de l'IP, il n'existe pas de conditions de bord permettant de remplir le critère d'optimalité.

Finalement, de nouvelles conditions de bord pour les méthodes E1 et D ont été proposées. Elles fournissent, par rapport aux

solutions habituelles, un bon compromis vis-à-vis de l'optimalité recherchée.

6.3 Méthodes de lissage linéaire à analogie physique: mise en application

A la section précédente, nous avons considéré le cas continu pour simplifier la présentation des méthodes de lissage, notamment en ce qui concerne les conditions de bord. Nous passons maintenant au cas discret.

6.3.1 Formes discrètes du lissage

\mathbf{F} représente le signal original et $\tilde{\mathbf{F}}$ le signal lissé.

Diffusion (D)

La relation entre $\tilde{\mathbf{F}}$ et \mathbf{F} est déterminée par l'intermédiaire du signal $\mathbf{F}(t)$ de la manière suivante. D'abord

$$\tilde{\mathbf{F}} = \mathbf{F}(\lambda_D^2/2) \quad (6.25)$$

ensuite, $\mathbf{F}(t)$ satisfait l'équation différentielle

$$\frac{\partial}{\partial t} \mathbf{F}(t) = \frac{1}{2} \Delta \mathbf{F}(t) \quad (6.26)$$

avec les conditions initiales

$$\mathbf{F}(0) = \mathbf{F} \quad (6.27)$$

Approximation du laplacien

$\Delta \mathbf{F}$ est une approximation discrète du laplacien réalisée par corrélation de \mathbf{F} avec un masque \mathbf{L} approprié

$$\Delta \mathbf{F} = \mathbf{L} \otimes \mathbf{F} \quad (6.28)$$

En voici trois:

- le laplacien discret axial

$$\mathbf{L}^+ = \begin{bmatrix} & 1 & \\ 1 & \boxed{-4} & 1 \\ & 1 & \end{bmatrix} \quad (6.29)$$

Il est égal à la somme des opérateurs approximant la deuxième dérivée dans les directions horizontale et verticale

$$\mathbf{L}^+ = \mathbf{D}_0^2 + \mathbf{D}_2^2 \quad (6.30)$$

- le laplacien discret diagonal

$$\mathbf{L}^\times = \frac{1}{2} \begin{bmatrix} 1 & & 1 \\ & \boxed{-4} & \\ 1 & & 1 \end{bmatrix} \quad (6.31)$$

- et la forme générale du laplacien discret sur un voisinage 3×3 [HaSh92: p. 347] consistant en une combinaison linéaire de \mathbf{L}^+ et \mathbf{L}^\times

$$\mathbf{L}^* = b\mathbf{L}^+ + (1-b)\mathbf{L}^\times \quad (6.32)$$

où b est un nombre réel quelconque.

Remarque: Il s'agit de l'approximation du laplacien dans le système de coordonnées (u, v) . Les intervalles d'échantillonnage n'apparaissent pas car ils sont égaux à 1.

Minimisation d'énergie (E1 et E2)

La relation entre $\tilde{\mathbf{F}}$ et \mathbf{F} est déterminée par la condition de minimum de la fonction J

$$J[\tilde{\mathbf{F}}] = \min_{\tilde{\mathbf{F}}} J[\tilde{\mathbf{F}}] \quad (6.33)$$

définie par

$$J[\tilde{\mathbf{F}}] = \sum \left[(\mathbf{F} - \tilde{\mathbf{F}})^2 + \gamma \Phi\langle \tilde{\mathbf{F}} \rangle \right] \quad (6.34)$$

où Φ est un opérateur de voisinage (non-linéaire) réalisant une approximation de la norme du gradient (E1) ou une approximation de la variation quadratique (E2).

Dans le cas des opérateurs qui viennent d'être mentionnés, la condition de minimum aboutit à un système d'équations linéaires pour les éléments de $\tilde{\mathbf{F}}$ en fonction de ceux de \mathbf{F} . Pour E1, le système d'équations est défini par la relation

$$\tilde{\mathbf{F}} - \lambda_{E1}^2 \Delta \tilde{\mathbf{F}} = \mathbf{F} \quad (6.35)$$

où $\Delta \tilde{\mathbf{F}}$ représente l'approximation discrète du laplacien. Celle-ci est réalisée par corrélation avec le masque \mathbf{L}^+ , par exemple. Pour E2, le système d'équations est défini par

$$\tilde{\mathbf{F}} + \lambda_{E2}^4 \Delta^2 \tilde{\mathbf{F}} = \mathbf{F} \quad (6.36)$$

où $\Delta^2 \tilde{\mathbf{F}}$ représente l'approximation discrète du bilaplacien. Le bilaplacien est estimé par corrélation, par exemple avec le masque

$$\mathbf{L}^2 = \begin{bmatrix} & & 1 & & & \\ & 2 & -8 & 2 & & \\ 1 & -8 & \boxed{20} & -8 & 1 & \\ & 2 & -8 & 2 & & \\ & & 1 & & & \end{bmatrix} \quad (6.37)$$

6.3.2 Mise en application du lissage

La réalisation du lissage consiste à résoudre le problème posé par la méthode de lissage correspondante [§6.3.1]. Pour chacun des problèmes, différentes méthodes numériques peuvent être utilisées. Elles sont plus ou moins bien adaptées suivant les propriétés (taille et forme) du domaine sur lequel est appliqué le lissage.

La table 6.2 présente les principales méthodes numériques et indique pour quels cas elles sont bien adaptées. Voici, pour chacune des méthodes, quelques précisions:

- convolution avec une forme discrète de la fonction de Green
Elle s'applique uniquement dans le cas d'un domaine infini (extension de l'IP). L'utilisation de la fonction de Green pour la diffusion (gaussienne) est très répandue.
- résolution directe du système d'équations
Elle s'applique seulement aux méthodes E1 et E2. Il est en principe possible de résoudre directement le système d'équations (6.35-36). Pratiquement, cette solution se limite aux cas de domaines de petite dimension, à cause de la quantité importante de mémoire requise.
- méthodes itératives locales
Ce sont les plus universelles.
- méthodes itératives à différents niveaux de résolution
par exemple: techniques multigrilles [Terz86, Terz88], techniques de fonctions de base hiérarchiques [Szel90]. A notre connaissance, elles s'appliquent uniquement aux méthodes E1 et E2. Leur utilisation pour le cas d'un domaine de forme irrégulière exige un soin particulier (la description du domaine à différents niveaux de résolution peut engendrer des recouvrements).

	taille infinie	taille finie, forme rectangulaire	taille finie, forme irrégulière
Fonction de Green	✓	–	–
Solution directe du système d'équations ^{1,2)}	–	✓	✓
Itératives locales	✓	✓	✓
Itératives à différents niveaux de résolution ¹⁾	✓	✓	?

1) méthodes numériques pour E1 et E2 uniquement

2) uniquement pour des domaines de petite dimension

Table 6.2: Vue d'ensemble des méthodes numériques de lissage.

En conclusion, comme une IP est caractérisée par un domaine de taille finie et souvent de forme irrégulière et que, d'autre part, on ne souhaite pas être limité quant aux dimensions du domaine, l'utilisation de méthodes itératives locales s'impose. Dans la suite de cette section, nous traitons uniquement les méthodes itératives locales.

6.3.3 Méthodes itératives locales

Nous expliquons ici leur principe. Afin de rester général, le traitement du bord, qui varie avec les conditions qui y sont imposées, n'est pas mentionné.

Formellement, une méthode itérative locale génère une suite de signaux discrets $\mathbf{F}^{(n)}$, le premier étant initialisé à la valeur du signal original \mathbf{F}

$$\mathbf{F}^{(0)} = \mathbf{F} \quad (6.38)$$

et le dernier déterminant la valeur du signal lissé $\tilde{\mathbf{F}}$

$$\tilde{\mathbf{F}} = \mathbf{F}^{(N)} \quad (6.39)$$

Le calcul correspondant à un pas d'itération, soit le passage du signal $\mathbf{F}^{(n-1)}$ au signal $\mathbf{F}^{(n)}$, est réalisé par une opération de voisinage

$$\mathbf{F}^{(n)} = I\langle \mathbf{F}^{(n-1)} \rangle, \quad n \leftarrow n + 1 \quad (6.40)$$

Nous présentons en détail dans les deux sections suivantes les méthodes itératives locales pour le problème de diffusion [§6.3.4] et celles pour le problème de minimisation d'énergie [§6.3.5].

6.3.4 Méthodes itératives de diffusion

Introduction

La formule de récurrence pour une itération (6.40) s'obtient à partir d'une forme discrète de l'équation différentielle (6.26).

On distingue deux formes de l'équation discrète de diffusion:

- la forme explicite

$$\mathbf{F}^{(n)} = \mathbf{F}^{(n-1)} + \alpha \Delta \mathbf{F}^{(n-1)} \quad (6.41)$$

- la forme implicite

$$\mathbf{F}^{(n)} = \mathbf{F}^{(n-1)} + \alpha \Delta \mathbf{F}^{(n)} \quad (6.42)$$

Le paramètre α correspond à la taille de l'intervalle d'échantillonnage suivant la variable t .

Trois méthodes itératives de diffusion sont présentées ci-après. Elles se distinguent par la forme de l'équation de diffusion qu'elles appliquent (explicite ou implicite) et par la forme discrète du laplacien qui est utilisée.

Méthodes explicites et implicite

Méthode D^+

La forme explicite de diffusion et le choix du laplacien \mathbf{L}^+ conduisent à la formule itérative suivante

$$\mathbf{F}^{(n)} = \mathbf{F}^{(n-1)} + \alpha \mathbf{L}^+ \otimes \mathbf{F}^{(n-1)} \quad (6.43)$$

qui peut s'écrire sous la forme d'une opération de corrélation

$$\mathbf{F}^{(n)} = \left[\begin{array}{c} \alpha \\ \alpha \boxed{1 - 4\alpha} \alpha \\ \alpha \end{array} \right] \otimes \mathbf{F}^{(n-1)} \quad (6.44)$$

La méthode est stable si et seulement si

$$\alpha \leq \frac{1}{4} \quad (6.45)$$

selon [PFTV86: §17.2]

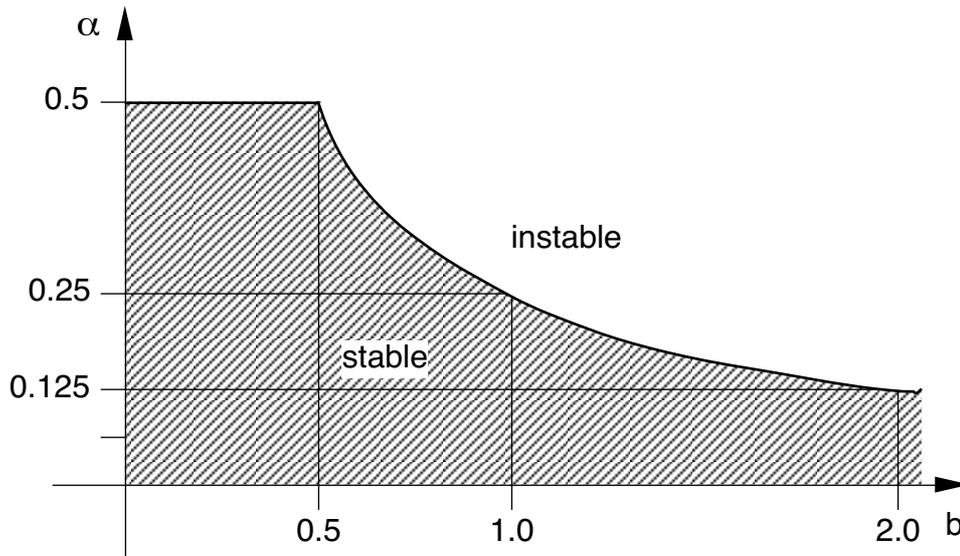


Figure 6.6: Domaine de stabilité pour la méthode D^* .

Méthode D^*

La forme explicite de diffusion avec le laplacien \mathbf{L}^* amène la formule itérative

$$\mathbf{F}^{(n)} = \mathbf{F}^{(n-1)} + \alpha \mathbf{L}^* \otimes \mathbf{F}^{(n-1)} \quad (6.46)$$

Celle-ci peut s'écrire sous la forme d'une corrélation avec un masque 3×3 approprié.

Une analyse de stabilité suivant la méthode de von Neumann [PFTV86: p.625] permet d'établir les conditions que les paramètres b (de \mathbf{L}^*) et α doivent vérifier pour que le procédé itératif soit stable [Annexe B]. Formellement, ces conditions s'expriment par la proposition

$$(b \geq 0) \wedge (\alpha \geq 0) \wedge (b \leq \frac{1}{2} \Rightarrow \alpha \leq \frac{1}{2}) \wedge (b \geq \frac{1}{2} \Rightarrow \alpha \leq \frac{1}{4b}) \quad (6.47)$$

La figure 6.6 représente géométriquement le domaine de stabilité dans le plan (b, α) . On remarque que la limite supérieure pour α vaut $\frac{1}{2}$.

Méthode D^+

L'équation de diffusion implicite combinée au laplacien \mathbf{L}^+ conduit à un procédé itératif qui a l'avantage d'être stable quelle que soit la valeur de α [PFTV86: p.638]. Le procédé ainsi obtenu a cependant les désavantages d'une précision moindre [PFTV86: p.636] et d'un calcul

$$\begin{array}{ll}
\text{a)} & \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & \boxed{1} & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \alpha = \frac{1}{3} \\
& & b = \frac{1}{3} \\
\text{b)} & \frac{1}{25} \begin{bmatrix} 4 & 2 & 4 \\ 2 & \boxed{1} & 2 \\ 4 & 2 & 4 \end{bmatrix} \quad \alpha = \frac{2}{5} \\
& & b = \frac{1}{5} \\
\text{c)} & \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & \boxed{4} & 2 \\ 1 & 2 & 1 \end{bmatrix} \quad \alpha = \frac{1}{4} \\
& & b = \frac{1}{2} \\
\text{d)} & \frac{1}{24} \begin{bmatrix} 1 & 2 & 1 \\ 2 & \boxed{12} & 2 \\ 1 & 2 & 1 \end{bmatrix} \quad \alpha = \frac{1}{6} \\
& & b = \frac{1}{2}
\end{array}$$

Figure 6.7: Masques de moyennage et leurs paramètres de diffusion correspondants. (a) "3×3 box filter" permet une atténuation maximale en une itération d'un bruit identiquement indépendamment distribué [HaSh92: p.310]. (b) Pour atténuer un bruit corrélé positivement. (c) Pour l'atténuation d'un bruit corrélé négativement. (d) Utilisé de manière itérative par Brady et al. [BPYA85] pour réaliser un filtrage gaussien.

difficile pour effectuer une itération [PFTV86: p.645]. La méthode implicite à direction alternée ("alternating-direction implicit method (ADI)") pallie à ces deux inconvénients [PFTV86: p.645].

Dans la méthode ADI, que nous désignons ici par D^\perp , une itération est divisée en deux demi-pas et le laplacien est décomposé en deux termes correspondant aux deuxièmes dérivées dans les directions horizontale et verticale (6.30). La formule pour chaque demi-pas d'itération est implicite suivant une direction et explicite suivant l'autre, ceci de manière alternée. Formellement

$$\begin{aligned}
\mathbf{F}^{(n-1/2)} &= \mathbf{F}^{(n-1)} + \frac{1}{2} \alpha \left(\mathbf{D}_0^2 \otimes \mathbf{F}^{(n-1/2)} + \mathbf{D}_2^2 \otimes \mathbf{F}^{(n-1)} \right) \\
\mathbf{F}^{(n)} &= \mathbf{F}^{(n-1/2)} + \frac{1}{2} \alpha \left(\mathbf{D}_0^2 \otimes \mathbf{F}^{(n-1/2)} + \mathbf{D}_2^2 \otimes \mathbf{F}^{(n)} \right)
\end{aligned} \quad (6.48.a-b)$$

Ces deux systèmes d'équations forment des matrices tridiagonales. Leur résolution par l'algorithme TRIDAG [PFTV86: p.40] correspond à une opération de voisinage récursive non-linéaire unidimensionnelle suivant les lignes, respectivement suivant les colonnes.

Forme explicite de diffusion et moyennage répété

On peut vérifier facilement que l'opération de corrélation correspondant à une itération des méthodes D^+ ou D^* effectue une moyenne pondérée des éléments du voisinage associé.

L'inverse est vrai: un moyennage pondéré des éléments du voisinage peut être exprimé sous la forme d'une itération de diffusion, sauf dans des cas singuliers. La figure 6.7 présente plusieurs masques de moyennage, issus de la littérature, accompagnés des paramètres de diffusion correspondants.

La technique de lissage qui consiste à répéter une opération de moyennage sur un voisinage 3×3 [BPYA85, par ex.] est donc identique à la méthode de diffusion D^* .

Caractéristiques des méthodes D

Condition d'arrêt et nombre d'itérations

La condition d'arrêt se déduit des relations (6.25) et (6.41-42). Elle s'exprime par l'inégalité

$$\sum_n \alpha_n \geq \lambda_D^2 / 2 \quad (6.49)$$

où α_n désigne l'incrément en t réalisé par l'itération n .

Cette condition détermine le nombre N d'itérations. Pour un paramètre de lissage λ_D donné, N est donc fixé par le choix de l'incrément α .

Choix de α et nombre minimal d'itérations

Dans le principe, la valeur de α doit être choisie telle que l'échantillonnage de la variable t soit fin relativement à λ_D^2 . Autrement dit

$$\alpha \ll \lambda_D^2 \quad (6.50)$$

A côté de cette règle qualitative, il existe deux bornes (supérieures) strictes pour la valeur de α . La première provient d'une condition sur α . La deuxième passe par une condition sur N . Toutes deux définissent des bornes (inférieures) pour le nombre d'itérations N :

- contrainte de stabilité

Pour les méthodes explicites (D^+ et D^*), le critère de stabilité impose une valeur maximale pour α , d'où une valeur minimale pour N

$$N > \lambda_D^2 / 2\alpha_{max} \quad (6.51)$$

- contrainte de propagation

Dans le cas des méthodes explicites (D^+ et D^*), une itération

	D^+	D^*	D^\perp
grandeur maximale de l'incrément (α_{max})	$\frac{1}{4}$	$\frac{1}{2}$	$+\infty$
complexité (1 élément, 1 itération)	4 additions 2 multiplications	8 additions 3 multiplications	10 additions 6 multiplications 4 divisions
complexité ¹⁾ (signal $L \times L$)	$O(\lambda_D^2 L^2)$	$O(\lambda_D^2 L^2)$	$O(\lambda_D L^2)$
parallélisation (signal $L \times L$)	L^2	L^2	L^2

1) pour la valeur minimale du nombre d'itérations

Table 6.3: Vue d'ensemble des méthodes numériques de diffusion.

effectue une opération de voisinage 3×3 non-réursive. Dans le cas de la méthode D^\perp , une itération correspond à une opération de voisinage 3×3 réursive, mais la récursivité est appliquée toujours dans la même direction. Par conséquent, quelle que soit la méthode, il existe une direction pour laquelle la distance d'influence d'un point augmente seulement de un à chaque itération. Il faut donc au moins $3\lambda_D$ itérations pour porter la distance d'influence à celle que génère une gaussienne d'écart-type λ_D (limiter le support de la gaussienne à un disque de rayon égal au triple de l'écart-type correspond à une erreur d'environ 1%). Cette relation pose une borne inférieure pour le nombre d'itérations

$$N > 3\lambda_D \quad (6.52)$$

Comparaison

Nous comparons les méthodes itératives de diffusion par rapport à la complexité et aux possibilités de parallélisation. La table 6.3 présente une vue d'ensemble, à laquelle nous apportons ci-après quelques précisions.

La complexité est de manière évidente proportionnelle au nombre L^2 d'éléments du signal et au nombre N d'itérations. Si l'on prend

comme référence le nombre minimal d'itérations, la complexité est proportionnelle à λ_D^2 pour les méthodes explicites (D^+ et D^*), à cause de la contrainte de stabilité, et proportionnelle à λ_D pour la méthode D^\perp , à cause de la contrainte de propagation.

En considérant quel est le nombre d'opérations par position et par itération [table 6.3], nous déduisons des réflexions que nous venons de faire que, pour une valeur élevée de λ_D , la méthode D^\perp est de complexité moindre que les deux autres. La situation est inverse pour une petite valeur de λ_D .

Les deux méthodes explicites sont de même complexité. En effet, avec une valeur α_{max} deux fois plus élevée, D^* nécessite deux fois moins d'itérations que D^+ , mais une itération compte environ deux fois plus d'opérations.

En ce qui concerne la parallélisation, les méthodes explicites sont évidemment parallélisables d'un facteur L^2 . Il existe aussi un algorithme parallèle en L^2 pour la méthode D^\perp [HoJe88: p.533].

En conclusion, la méthode D^\perp offre les meilleures performances pour une valeur élevée de λ_D . Mais les méthodes explicites l'emportent pour de petites valeurs de λ_D .

Atténuation du bruit

Nous terminons cette section consacrée aux méthodes itératives de diffusion en présentant des résultats expérimentaux concernant les performances d'atténuation du bruit de la méthode D^+ .

La méthode a été appliquée à un signal constitué par un bruit pseudo-aléatoire de distribution gaussienne discrète dont l'écart-type est 4.

La figure 6.8.a présente, pour différentes valeurs de α , l'évolution de la valeur efficace du signal lissé en fonction du nombre d'itérations. La figure 6.8.b en fait de même avec la deuxième dérivée du signal lissé, estimée par corrélation avec le masque \mathbf{D}_0^2 [§4.2.2]. La figure 6.9 présente, pour trois différentes valeurs de α , le signal lissé après 5 itérations.

Il ressort de ces résultats que $\alpha = 0.2$ constitue un choix optimal, si l'on décide de travailler avec une valeur unique de α quel que soit le nombre d'itérations.

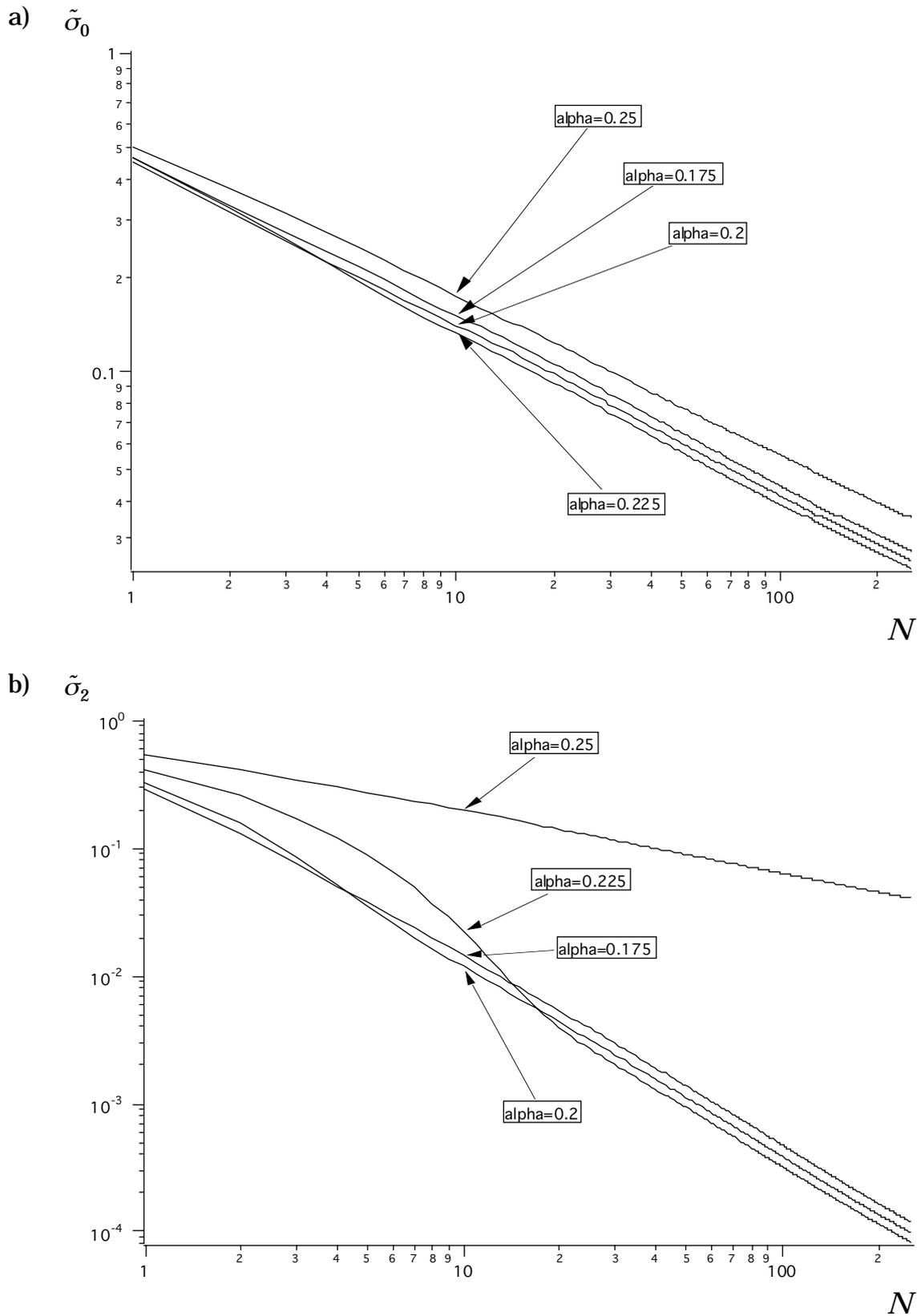


Figure 6.8: Atténuation de bruit par la méthode D^+ . Résidu de bruit pseudo-aléatoire de distribution gaussienne en fonction du nombre d'itérations (a) Valeur efficace du bruit résiduel. (b) Valeur efficace de la deuxième dérivée du bruit résiduel.

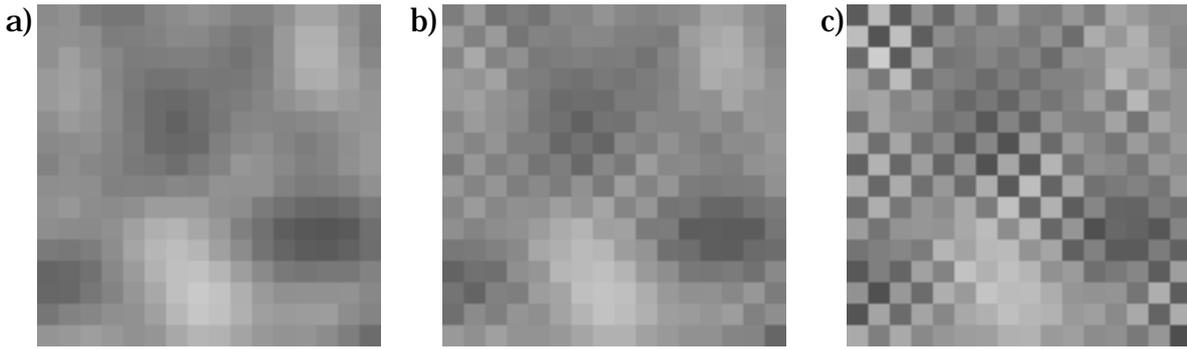


Figure 6.9: Bruit résiduel après lissage par la méthode D^+ . Résidu de bruit pseudo-aléatoire de distribution gaussienne après 5 itérations. Les images d'intensité représentent une portion du signal de taille 16x16. (a) $\alpha = 0.2$, (b) $\alpha = 0.225$, (c) $\alpha = 0.25$. La valeur de $\alpha = 0.2$ présente de manière évidente les meilleures caractéristiques d'atténuation, bien que la valeur efficace du bruit résiduel soit inférieure pour un choix de $\alpha = 0.225$.

6.3.5 Méthodes itératives de minimisation d'énergie

Introduction

Le problème mathématique associé aux méthodes de lissage E1 et E2 possède deux formes équivalentes: minimisation d'une fonction de plusieurs variables (6.34) ou résolution d'un système d'équations linéaires (6.35-36). Les solutions numériques de l'une ou l'autre forme du problème se rejoignent [HaYo81: §7.2]. Nous présentons ici les méthodes itératives locales par rapport à la résolution du système d'équations.

Principe

La forme générale du système d'équations à résoudre (6.35-36) est définie par

$$\mathbf{C} \otimes \tilde{\mathbf{F}} = \mathbf{F} \quad (6.53)$$

avec

$$\mathbf{C} = \mathbf{1} - \lambda_{E1}^2 \mathbf{L}^+ \quad (6.54.a)$$

pour E1 ou

$$\mathbf{C} = \mathbf{1} + \lambda_{E2}^4 \mathbf{L}^2 \quad (6.54.b)$$

pour E2.

Une équation correspond à l'opération locale

$$\sum_{m=0}^{M-1} c_m \cdot \tilde{f}_m = f_0 \quad (6.55)$$

où l'indice 0 indique la position de référence et l'indice m indique une position dans son voisinage (y compris la position de référence).

On considère l'approximation de $\tilde{\mathbf{F}}$ après l'itération $(n-1)$, soit le signal $\mathbf{F}^{(n-1)}$. Celui-ci ne vérifie pas exactement l'égalité (6.55). C'est-à-dire

$$\sum_{m=0}^{M-1} c_m \cdot f_m^{(n-1)} \neq f_0 \quad (6.56)$$

Dans le principe, l'itération n effectue localement une correction de $\mathbf{F}^{(n-1)}$ de sorte à ce que l'égalité (6.56) soit vérifiée en utilisant la valeur corrigée à la position de référence et l'ancienne valeur aux autres positions. La nouvelle valeur temporaire $f_0^{(n-1)*}$ est donc déterminée par la condition

$$c_0 \cdot f_0^{(n-1)*} + \sum_{m=1}^{M-1} c_m \cdot f_m^{(n-1)} = f_0 \quad (6.57)$$

ou, de manière explicite

$$f_0^{(n-1)*} = \left(f_0 - \sum_{m=1}^{M-1} c_m \cdot f_m^{(n-1)} \right) / c_0 \quad (6.58)$$

La modification de $f_0^{(n-1)}$ peut être exprimée comme l'addition d'un terme correctif

$$f_0^{(n-1)*} = f_0^{(n-1)} + \delta f_0^{(n-1)*} \quad (6.59.a)$$

qui est

$$\delta f_0^{(n-1)*} = \left(f_0 - \sum_{m=0}^{M-1} c_m \cdot f_m^{(n-1)} \right) / c_0 \quad (6.59.b)$$

La grandeur $\delta f_0^{(n-1)*}$ apparaît comme l'erreur par rapport à la solution du système d'équation. Elle est appelée pseudo-résidu.

Les différentes méthodes itératives se distinguent dans la manière de construire le signal $\mathbf{F}^{(n)}$ à partir des valeurs $f_0^{(n-1)*}$, respectivement le signal de pseudo-résidu $\delta \mathbf{F}^{(n-1)}$ à partir des valeurs $\delta f_0^{(n-1)*}$. Nous présentons ci-après les trois méthodes de base: Jacobi, Gauss-Seidel et SOR ("successive over-relaxation").

Méthodes de base

Méthode de Jacobi

L'opération locale (6.58) est effectuée de manière synchrone, d'où la formule de récurrence

$$\mathbf{F}^{(n)} = \mathbf{F}^{(n-1)} + \frac{1}{c_0} (\mathbf{F} - \mathbf{C} \otimes \mathbf{F}^{(n-1)}) \quad (6.60)$$

Une itération correspond donc, pour l'essentiel, à une opération de corrélation.

Méthode de Gauss-Seidel

Le signal $\mathbf{F}^{(n-1)}$ est mis à jour de manière asynchrone par l'opération locale (6.58). Par conséquent, l'ordre de mise à jour joue un rôle. Les ordonnancements usuels sont l'ordonnement naturel et l'ordonnement pair-impair (jeu d'échec).

Dans le cas d'un ordonnancement naturel, une itération peut être décrite à l'aide d'une opération de corrélation récursive avec le masque \mathbf{C}

$$\mathbf{F}^{(n)} = \mathbf{F}^{(n-1)} + \frac{1}{c_0} (\mathbf{F} - \mathbf{C} \tilde{\otimes} \mathbf{F}^{(n-1)}) \quad (6.61)$$

Méthode de sur-relaxation successive (SOR)

Elle peut être considérée comme une modification de la méthode de Gauss-Seidel. Au lieu de corriger le signal simplement par addition du signal de pseudo-résidu, celui-ci est multiplié par une constante ω avant d'être additionné. Formellement

$$\mathbf{F}^{(n)} = \mathbf{F}^{(n-1)} + \omega \delta \mathbf{F}^{(n-1)} \Big|_{GS} \quad (6.62)$$

La méthode converge pour

$$0 < \omega < 2 \quad (6.63)$$

Il existe une valeur du coefficient ω qui fournit les meilleures propriétés de convergence. Cette valeur ω_{opt} est comprise entre 1 et 2. Elle dépend de l'ordonnement.

Dans le cas des systèmes d'équations pour E1 et E2, il existe une formule exprimant ω_{opt} en fonction de λ_E pour un ordonnancement naturel [BlZi87: §7.6.1, table 6.4].

Caractéristiques des méthodes E

Taux de convergence

Le taux de convergence moyen $r^{(n)}$ est l'atténuation moyenne du résidu sur n itérations. Formellement

$$r^{(n)} = -\frac{1}{n} \log\left(\frac{\|\tilde{\mathbf{F}} - \mathbf{F}^{(n)}\|}{\|\tilde{\mathbf{F}} - \mathbf{F}^{(0)}\|}\right) \quad (6.64)$$

où $\|\cdot\|$ désigne une fonction de norme. $r^{(n)}$ mesure la rapidité de convergence sur les n itérations. Plus la valeur de $r^{(n)}$ est élevée, plus la convergence est rapide.

Pour caractériser la convergence d'une méthode, on utilise généralement le taux de convergence asymptotique, c'est-à-dire $r^{(\infty)}$. Par exemple, dans la méthode SOR, ω_{opt} est la valeur de ω qui maximise $r^{(\infty)}$.

Dans le cas des systèmes d'équations pour E1 et E2, il existe, pour chacune des méthodes de base, une formule exprimant $r^{(\infty)}$ en fonction de λ_E [BlZi87: §7.6.1, table 6.4].

Condition d'arrêt et nombre d'itérations

La condition d'arrêt classique pour la résolution itérative du système d'équations (6.53) est formée par la comparaison d'une évaluation du résidu avec une valeur de seuil ε [BlZi87: §7.6.1, HaYo81: §7.2]. Le résidu est évalué à partir du pseudo-résidu de l'itération en cours, d'où la formule

$$\kappa \|\delta \mathbf{F}^{(n)}\| < \varepsilon \quad (6.65)$$

pour le test de fin.

Avec une telle condition d'arrêt, le nombre d'itérations N varie en fonction de la valeur du signal \mathbf{F} et donc la transformation de signal résultant du procédé itératif n'est pas linéaire. D'autre part, le test d'arrêt nécessite l'évaluation de la norme $\|\cdot\|$, ce qui est une opération globale et donc seulement partiellement parallélisable.

A notre avis, dans le cas qui nous intéresse de lissage avec les méthodes E1 et E2, il est préférable de travailler avec un nombre fixe d'itérations (qui dépend bien sûr de λ_E). Pour déterminer le nombre d'itérations, on peut par exemple appliquer la condition (6.65) à un signal de référence. Une autre possibilité est de choisir une valeur p d'atténuation du résidu et d'estimer à partir du taux asymptotique de

meilleur algorithme séquentiel	SOR ($\omega_{opt} = \frac{2}{1+1/\sqrt{2\lambda}}$)	SOR ($\omega_{opt} = \frac{2}{1+1/\sqrt{10\lambda^2}}$)
• complexité ¹⁾ (signal $L \times L$)	$\mathcal{O}(p\lambda_E L^2)$	$\mathcal{O}(p\lambda_E^2 L^2)$
meilleur algorithme parallèle	SOR ou Gauss-Seidel avec accélération de Chebyshev ²⁾ , ordonnancement pair-impair	Jacobi ³⁾
• complexité ¹⁾ (signal $L \times L$)	$\mathcal{O}(p\lambda_E L^2)$	$\mathcal{O}(p\lambda_E^4 L^2)$
• parallélisation (signal $L \times L$)	$L^2/2$	L^2

1) nombre d'opérations nécessaires pour diminuer l'erreur résiduelle d'un facteur 10^P

2) L'accélération de Chebyshev effectue un changement de la valeur de ω à chaque pas d'itération. La convergence asymptotique est aussi bonne que celle de la méthode SOR. La convergence dans les premières itérations est meilleure [PFTV86: p.658]

3) appliquée non pas à l'équation d'Euler formulée directement, mais à l'équation d'Euler sous la forme d'équations différentielles couplées dans lesquelles interviennent les composantes du gradient [Harr87]

Table 6.4: Meilleures méthodes itératives de minimisation d'énergie.

convergence $r^{(\infty)}$ le nombre d'itérations nécessaires à réaliser cette atténuation. Le nombre estimé est $p/r^{(\infty)}$.

Meilleures méthodes

En ce qui concerne la complexité, les méthodes se distinguent seulement au niveau du nombre d'itérations. Si l'on prend comme référence pour le nombre d'itérations la valeur $p/r^{(\infty)}$, les meilleures méthodes, c'est-à-dire celles de complexité minimale, sont celles dont le taux de convergence asymptotique $r^{(\infty)}$ est maximal. La meilleure méthode est donc la méthode SOR car elle possède le meilleur taux de convergence [HaYo81: §2.4]. Sa complexité est proportionnelle à λ_E pour E1 et λ_E^2 pour E2.

La méthode SOR avec un ordonnancement pair-impair est parallélisable pour E1. En effet, comme il n'y a pas d'interaction entre positions de même parité, les opérations relatives à des positions différentes mais de même parité peuvent être effectuées en parallèle.

Cette possibilité n'existe pas pour E2. Pour pouvoir paralléliser, il faut utiliser la méthode de Jacobi. Mais la méthode de Jacobi appliquée à l'équation d'Euler sous la forme directe (6.54.b, 6.60) n'est pas stable. Par contre, l'utilisation d'équations différentielles couplées dans lesquelles interviennent les composantes du gradient offre la possibilité de résoudre le problème avec la méthode de Jacobi et donc de paralléliser [Harr87]. La complexité de la méthode est proportionnelle à λ_E^4 .

La table 6.4 récapitule ce que nous venons de dire.

6.3.6 Récapitulation et conclusion

Dans cette section, nous avons passé en revue les méthodes numériques permettant de réaliser le lissage par diffusion (D) et par minimisation d'énergie de membrane (E1) et de plaque mince (E2). L'intérêt s'est focalisé sur les méthodes itératives locales.

Méthodes itératives de diffusion

Les méthodes D sont caractérisées par l'existence d'un paramètre noté α qui détermine l'importance de la modification du signal à chaque itération. Une fois que la valeur de α est choisie, λ fixe le nombre d'itérations.

Dans les méthodes explicites de diffusion (D^+ et D^*), la condition de stabilité impose une valeur maximale pour α et par conséquent un nombre minimal d'itérations proportionnel à λ^2 . La méthode implicite D^\perp permet un nombre minimal d'itérations proportionnel à λ . Etant donné les opérations élémentaires intervenant dans chacune des méthodes, les méthodes explicites sont plus avantageuses lorsque la valeur de λ est petite. La méthode D^\perp l'emporte pour des valeurs élevées de λ .

Méthodes itératives de minimisation d'énergie

Dans les méthodes E1 et E2, le paramètre de lissage λ intervient dans l'opération réalisée à chaque itération. Le nombre d'itérations est fixé par le choix d'une condition d'arrêt. La solution la plus probante consiste à travailler avec un nombre fixe d'itérations inversement proportionnel au taux asymptotique de convergence.

Dans ces conditions, la meilleure méthode pour E1 est de complexité proportionnelle à λ . Pour E2, la complexité de la

meilleure méthode séquentielle est proportionnelle à λ^2 , alors que celle de la meilleure méthode parallèle est proportionnelle à λ^4 .

Comparaison

Pour les cas D et E1, il existe des méthodes de complexité proportionnelle à λ . Pour E2, les meilleures méthodes sont de complexité proportionnelle à λ^2 et λ^4 .

Dans les méthodes de diffusion, pour réaliser un lissage correspondant à une valeur de λ plus élevée, il suffit d'augmenter le nombre d'itérations. Par contre, dans les méthodes de minimisation d'énergie, une modification de λ entraîne une modification des opérations intervenant dans chaque itération, il faut donc recommencer le processus itératif.

En conclusion, les méthodes D sont de complexité inférieure à celle des méthodes E2 et de complexité semblable à celle des méthodes E1, mais elles sont plus avantageuses car elles permettent de travailler facilement avec plusieurs valeurs de λ (multirésolution).

6.4 Méthodes de lissage préservant les discontinuités

Cette section présente le principe et fait l'inventaire des méthodes de lissage préservant les discontinuités. Ces méthodes constituent des formes non-linéaires des méthodes de diffusion et de minimisation d'énergie présentées précédemment.

6.4.1 Introduction

Le principe appliqué est essentiellement celui d'un traitement adaptatif, c'est-à-dire dépendant du signal. Au lieu d'être uniforme, le lissage est contrôlé en fonction d'une mesure du signal, de telle manière à ce que le lissage soit empêché/atténué en présence d'une discontinuité.

En général, le contrôle du lissage est en fonction d'une mesure de discontinuité telle que décrite au chapitre 4. L'alternative est d'utiliser une mesure de la modification du signal engendrée par le lissage. Cette dernière approche repose sur le postulat qu'en préservant le signal on préserve les discontinuités.

Ci-après, nous montrons d'abord comment les méthodes de diffusion et de minimisation d'énergie sont modifiées pour préserver les discontinuités [§6.4.2-3]. Ensuite, nous montrons comment ces méthodes sont utilisées pour préserver à la fois les discontinuités d'ordre 0 (\bar{C}^0) et celles d'ordre 1 (\bar{C}^1) d'une image de profondeur [§6.4.4].

6.4.2 Diffusion non-linéaire

Contrôle du flux de diffusion (méthode D_{1f}^1)

Il s'agit d'une forme non-linéaire de la méthode itérative de diffusion D^+ [PeMa90].

Soit le laplacien discret axial \mathbf{L}^+ représenté comme la somme des opérateurs de différence entre la position centrale et ses quatre voisins dans les directions axiales [§4.2.2]

$$\mathbf{L}^+ = \mathbf{D}_{0-}^1 + \mathbf{D}_{0+}^1 + \mathbf{D}_{2-}^1 + \mathbf{D}_{2+}^1 \quad (6.66)$$

Cette décomposition de \mathbf{L}^+ permet d'écrire une itération de diffusion D^+ sous la forme

$$\mathbf{F}^{(n)} = \mathbf{F}^{(n-1)} + \alpha \sum_{m \in \{0\pm, 2\pm\}} \mathbf{D}_m^1 \otimes \mathbf{F}^{(n-1)} \quad (6.67)$$

ce qui correspond à l'opération locale

$$f_0^{(n)} = f_0^{(n-1)} + \alpha \sum_{m=1}^4 \left(f_m^{(n-1)} - f_0^{(n-1)} \right) \quad (6.68)$$

où chacun des quatre termes additionnés à $f_0^{(n-1)}$ représente le flux de diffusion (intégré sur l'intervalle Δt) de la position m vers la position 0.

La méthode linéaire est modifiée par l'introduction d'un coefficient d'atténuation $a_m^{(n)}$ (valeur comprise entre 0 et 1) qui contrôle le flux

$$f_0^{(n)} = f_0^{(n-1)} + \alpha \sum_{m=1}^4 a_m^{(n)} \cdot \left(f_m^{(n-1)} - f_0^{(n-1)} \right) \quad (6.69)$$

en fonction de la mesure de discontinuité \bar{C}^0 [chap.4] que constitue la différence entre valeurs voisines $f_m^{(n-1)}$ et $f_0^{(n-1)}$

$$a_m^{(n)} = \Theta_T \left(\left| f_m^{(n-1)} - f_0^{(n-1)} \right| \right) \quad (6.70)$$

La fonction Θ_T est un fonction de seuil [figure 6.11].

Contrôle des coefficients de moyennage (méthode D_{1c}^1)

Il s'agit d'une forme non-linéaire de diffusion explicite.

Soit l'itération de diffusion explicite représentée par la corrélation avec le masque de moyennage \mathbf{C}

$$\mathbf{F}^{(n+1)} = \mathbf{C} \otimes \mathbf{F}^{(n)} \quad (6.71)$$

La forme locale correspondante est

$$f_0^{(n)} = \sum_{m=0}^{M-1} c_m^{(n)} \cdot f_m^{(n-1)} \quad (6.72)$$

où l'indice 0 indique la position de référence et l'indice m indique une position dans son voisinage (y compris la position de référence).

La méthode linéaire est modifiée en faisant contrôler les coefficients c_m par une mesure de discontinuité d'ordre 0 Δ_m à la position correspondante. Pour cela, une fonction de seuil Θ_T [figure 6.11] est appliquée aux valeurs Δ_m et détermine des coefficient C_m

$$C_m = \Theta_T(\Delta_m) \quad (6.73)$$

Ceux-ci sont ensuite normalisés par leur somme pour obtenir les coefficients c_m

$$c_m = C_m / \sum_{m=0}^{M-1} C_m \quad (6.74)$$

Remarque

Cette méthode peut être considérée comme une extension de la méthode D_{1f}^1 .

Contrôle de la distance au signal original (méthode $D^{\delta_{max}}$)

Il s'agit d'une forme de diffusion explicite pour laquelle la distance au signal original est limitée à une valeur maximale T [MHTA90].

En limitant la distance au signal original, on obtient de manière simple un contrôle du lissage en fonction de la modification du signal engendrée par le lissage.

Dans cette méthode, une itération est décomposée en deux temps: d'abord, calcul de la nouvelle valeur sans tenir compte de la contrainte

$$\mathbf{F}^{(n-1/2)} = \mathbf{C} \otimes \mathbf{F}^{(n-1)} \quad (6.75)$$

a)

$$\Theta_T^b(\xi) = \begin{cases} 1, & |\xi| < T \\ 0, & |\xi| > T \end{cases}$$

$$\Theta_T^{exp}(\xi) = e^{-\xi^2/\sqrt{2}T^2} \quad \Theta_T^{sqr}(\xi) = \frac{1}{1 + (\xi/T)^2}$$

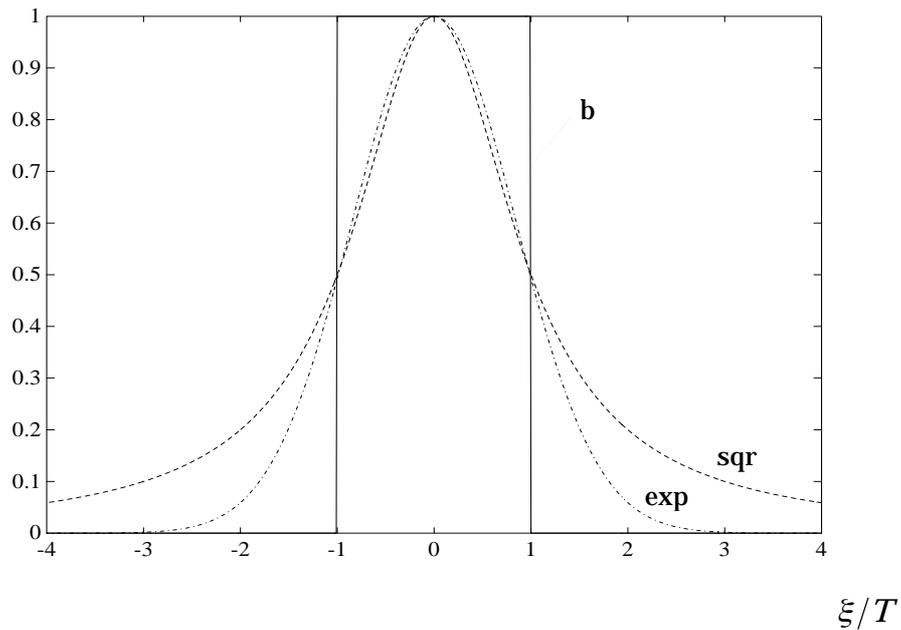
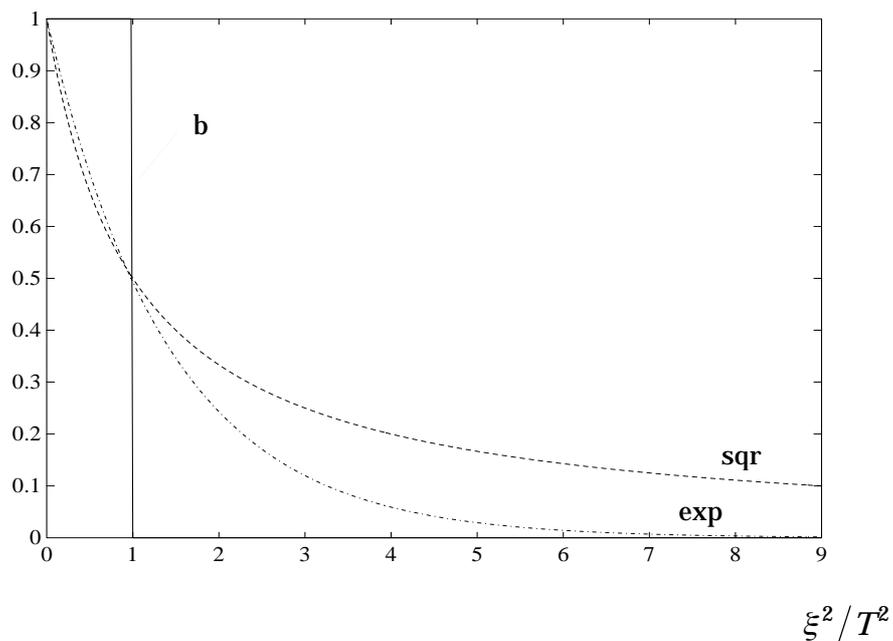
b) Θ_T c) Θ_T 

Figure 6.11: Fonctions de seuil utilisées pour atténuer les coefficients de diffusion. (a) Formules. (b) Représentation graphique en fonction de ξ/T . (c) Représentation graphique en fonction de ξ^2/T^2 .

puis, si la distance maximale T au signal original est dépassée, réajustement de la valeur du signal à la valeur limite

$$\mathbf{F}^{(n)} = \begin{cases} \mathbf{F}^{(n-1)} - \mathbf{T}, & \text{si } (\mathbf{F}^{(n-1/2)} < \mathbf{F}^{(0)} - \mathbf{T}) \\ \mathbf{F}^{(n-1)} + \mathbf{T}, & \text{si } (\mathbf{F}^{(n-1/2)} > \mathbf{F}^{(0)} + \mathbf{T}) \\ \mathbf{F}^{(n-1/2)}, & \text{sinon} \end{cases} \quad (6.76)$$

6.4.3 Minimisation d'énergie non-linéaire

Contrôle de l'énergie de déformation (méthodes E_1^l et E_2^l)

Les méthodes linéaires de minimisation d'énergie sont modifiées par l'introduction d'une fonction de correction appliquée à l'énergie de déformation. On obtient ainsi un contrôle du lissage en fonction d'une mesure de discontinuité: \bar{C}^0 pour E1, \bar{C}^0 et \bar{C}^1 pour E2 [chap.4].

Formellement, la fonctionnelle à minimiser devient

$$\mathcal{J}[\bar{\mathbf{F}}] = \sum \left[(\mathbf{F} - \bar{\mathbf{F}})^2 + \gamma \Xi_T(\Phi\langle\bar{\mathbf{F}}\rangle) \right] \quad (6.77)$$

où Ξ_T est la fonction de correction. En général, Ξ_T correspond à la transformation de la norme L_2 , qui mesure la déformation dans le cas linéaire, en une autre norme. Ce sont par exemple:

- L_1 [BILe92]

$$L_T^1(\xi) = T|\xi| \quad (6.78)$$

- fonction Minimax de Huber [BILe92]

$$L_T^H(\xi) = \begin{cases} \xi^2, & |\xi| < T \\ -T^2 + 2T|\xi|, & |\xi| \geq T \end{cases} \quad (6.79)$$

- norme de Blake et Zisserman [BlZi87]

$$L_T^{BZ}(\xi) = \begin{cases} \xi^2, & |\xi| < T \\ T^2, & |\xi| \geq T \end{cases} \quad (6.80)$$

- norme de Chelberg et Yi [ChYi92]

$$L_T^{CY}(\xi) = \frac{1}{1 + T|\xi|} \xi^2 \quad (6.81)$$

La figure 6.12 représente ces différentes normes et les fonction Ξ_T correspondantes. On remarque un effet de seuil plus ou moins prononcé par rapport au cas linéaire.

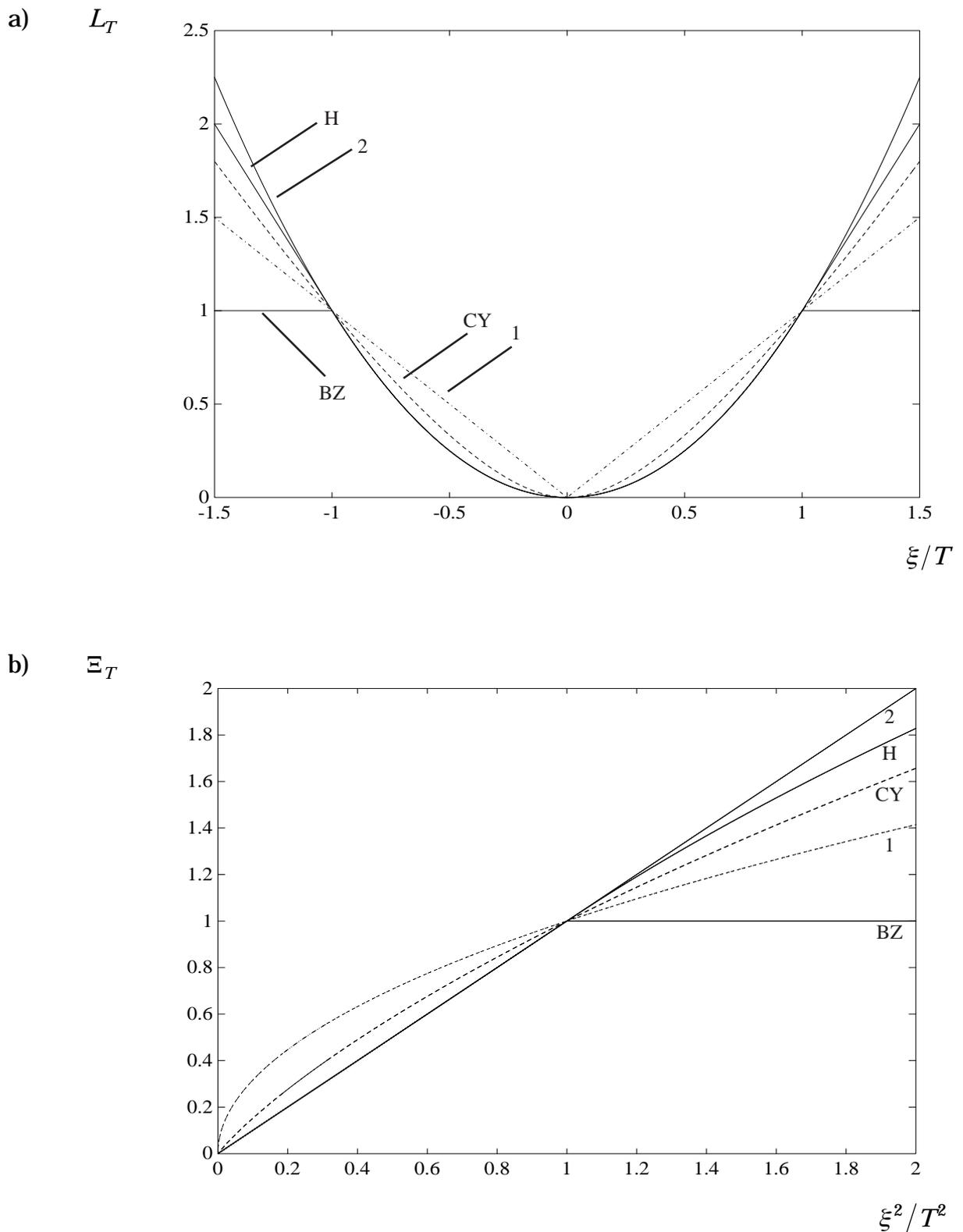


Figure 6.12: Fonctions pour la modification de l'énergie de déformation. (a) Norme L_T en fonction de ξ/T . (b) Correction Ξ_T en fonction de ξ^2/T^2 .

Contrôle de l'énergie d'approximation (méthodes $E^{\delta_{max}}$)

Les méthodes linéaires de minimisation d'énergie sont modifiées par l'utilisation d'une fonction à effet de seuil pour l'énergie d'approximation [MaHü92]. On obtient ainsi un contrôle du lissage en fonction d'une mesure de la modification du signal engendrée par le lissage.

Formellement, la fonctionnelle à minimiser devient

$$\mathcal{J}[\bar{\mathbf{F}}] = \sum [\Psi_T(\mathbf{F} - \bar{\mathbf{F}}) + \gamma \Phi\langle \bar{\mathbf{F}} \rangle] \quad (6.82)$$

où la fonction Ψ_T est définie par

$$\Psi_T(\xi) = \begin{cases} 0 & , |\xi| \leq T \\ \infty & , |\xi| > T \end{cases} \quad (6.83)$$

Ce qui revient à minimiser la fonctionnelle

$$\mathcal{J}[\bar{\mathbf{F}}] = \sum [\Phi\langle \bar{\mathbf{F}} \rangle] \quad (6.84)$$

sous la condition de distance maximale T par rapport au signal original

$$|\bar{\mathbf{F}} - \mathbf{F}| < \mathbf{T} \quad (6.85)$$

On remarque qu'il n'y a plus de coefficient de lissage dans la fonctionnelle à minimiser (6.84). Si la contrainte de distance maximale par rapport au signal original est vérifiée, le paramètre de lissage est égal à l'infini.

6.4.4 Préservation conjointe des discontinuités \bar{C}^0 et \bar{C}^1

Il existe quatre manières d'appliquer les méthodes décrites dans les deux sections précédentes pour préserver/détecter à la fois les discontinuités \bar{C}^0 et \bar{C}^1 d'une image de profondeur.

E_2^1 appliquée à l'IP

Elle préserve à la fois les discontinuités \bar{C}^0 et \bar{C}^1 [BlZi87].

Blauer et Levine [BlLe92] utilise une méthode similaire, la différence étant que le carré du laplacien remplace la variation quadratique comme énergie de déformation (pour le cas linéaire). Ce choix permet le passage de la norme L_2 à la norme L_1 (pour le cas non-linéaire).

D_1^1 ou E_1^1 appliqué à l'IP et à son gradient

Appliquée à l'IP, D_1^1 ou E_1^1 préserve les discontinuités \bar{C}^0 .

Pour préserver les discontinuités \bar{C}^1 , D_1^1 ou E_1^1 est appliquée aux deux images des dérivées partielles d'ordre 1 et le lissage est contrôlé par une mesure de discontinuité \bar{C}^1 exprimée en fonction des dérivées partielles. Les opérations de lissage doivent être exécutées simultanément sur les deux images car elles sont couplées par la mesure de discontinuité.

Le lissage de l'IP et celui de ses dérivées partielles est exécuté soit de manière séquentielle, soit en parallèle:

- Lisser l'IP, puis lisser les images des dérivées partielles d'ordre 1 en tenant compte des discontinuités \bar{C}^0 détectées [BlZi87]. Le problème avec cette approche est que les discontinuités \bar{C}^1 subissent une détérioration lors du lissage de l'IP.
- Lisser séparément l'IP et les images des dérivées partielles d'ordre 1 [SMCM91, ChYi92]. Dans ce cas, se pose la question de l'effet des discontinuités \bar{C}^0 lors du lissage des images des dérivées partielles.

D_1^1 appliquée à une structure composée de l'IP et de ses propriétés différentielles

Les discontinuités \bar{C}^0 et \bar{C}^1 de l'IP apparaissent comme des discontinuités \bar{C}^0 de ce signal composite, d'où l'utilisation possible de la méthode D_1^1 [FeMS93].

$D^{\delta_{max}}$ ou $E^{\delta_{max}}$ appliquée à l'IP

Cette manière de faire s'est avérée utile à préserver les deux types de discontinuité [MHTA90, MaHü92].

6.4.5 Discussion

Principe

Toutes les méthodes appliquent le principe décrit à la section 6.4.1. Cependant, bien que la préservation des discontinuités en fait une détection implicite, celle-ci n'est que rarement explicitée [BlZi87]. Dans la plupart des cas, la détection est réalisée par traitement ultérieur, en appliquant au signal lissé une des méthodes décrite au chapitre 4.

Toutes les méthodes sont paramétrisées par les grandeurs λ et T , qui ont partout une signification semblable: λ indique l'importance du lissage et T définit un seuil sur la mesure de discontinuité au-dessus duquel le lissage est atténué.

Les méthodes de diffusion se distinguent des méthodes de minimisation d'énergie par le fait qu'une variation du seuil T en fonction des itérations peut être envisagée, ce qui accroît fortement l'espace des paramètres. Dans la plupart des mises en application de la méthode, le seuil est pourtant maintenu constant.

Les paramètres λ et T sont généralement libres. Deux propositions de seuil T adapté ont été faites. Dans les deux cas, il s'agit d'une méthode de diffusion:

- Dans l'approche de Ferrie et al. [FeMS93], le seuil est déterminé à chaque itération et à chaque position à partir d'une statistique locale. Il est fixé à une valeur proportionnelle à l'écart-type des valeurs voisines de la position considérée. Le coefficient de proportionnalité est constant. Dans les résultats présentés, il a été fixé à $1/8$.
- Dans une des versions de la méthode de Perona et Malik [PeMa90], le seuil est déterminé à chaque itération sur la base d'une statistique globale des valeurs de la mesure de discontinuité (gradient). Il est fixé à 90 % de la valeur intégrale de l'histogramme.

Le lissage est rarement mis en relation avec l'atténuation du bruit. Le travail de Blake et Zisserman [BlZi87] est le seul où une telle relation est considérée. En imposant qu'une discontinuité ne soit pas détectée en présence d'un signal constant bruité, ils établissent une condition pour le seuil T sous la forme d'une inégalité entre sa valeur et celle de l'écart-type du bruit.

Préservation/détection des deux types de discontinuité

Les principes utilisés pour préserver/détecter à la fois les discontinuités d'ordre 0 et celles d'ordre 1 ont été présentés à la section 6.4.4. On souligne ici les inconvénients de chaque méthode. La table 6.5 en fait la récapitulation.

méthode	inconvénients
E_2^1 appliquée à l'IP	coût de calcul très élevé
D_1^1 ou E_1^1 appliquée à l'IP et à son gradient	non disponibilité d'une IP lissée sinon coût de calcul très élevé*
D_1^1 appliquée à une structure composée de l'IP et de ses propriétés différentielles	coût de calcul très élevé
$D^{\delta_{max}}$ ou $E^{\delta_{max}}$ appliquée à l'IP	sensibilité au bruit

*on peut reconstruire l'IP si on impose la condition d'intégrabilité. Le coût de calcul devient alors malheureusement exorbitant [BIZi87, SMCM91]

Table 6.5: Méthodes de lissage préservant les discontinuités \bar{C}^0 et \bar{C}^1 .

6.5 Méthode MDMS

La nouvelle méthode que nous proposons dans cette section fait partie des méthodes de diffusion non-linéaire. Elle se distingue des autres méthodes de diffusion présentées auparavant à deux points de vue. Premièrement, elle permet de préserver/détecter à la fois les discontinuités \bar{C}^0 et \bar{C}^1 par application à un seul signal, qui est constitué par l'IP uniquement. Deuxièmement, elle fait varier le seuil de détection en fonction d'une estimation du bruit résiduel.

6.5.1 Principe

Diffusion explicite D^*

Le laplacien discret sur un voisinage 3×3 \mathbf{L}^* est représenté comme la somme pondérée des opérateurs de deuxième dérivée dans les directions axiales et diagonales

$$\mathbf{L}^* = b(\mathbf{D}_0^2 + \mathbf{D}_2^2) + (1 - b)(\mathbf{D}_1^2 + \mathbf{D}_3^2) \quad (6.86)$$

ce qui permet d'écrire un pas de diffusion sous la forme

$$\mathbf{F}^{(n)} = \mathbf{F}^{(n-1)} + \alpha \left\{ b \sum_{m \in \{0,2\}} \mathbf{D}_m^2 + (1 - b) \sum_{m \in \{1,3\}} \mathbf{D}_m^2 \right\} \otimes \mathbf{F}^{(n-1)} \quad (6.87)$$

De cette manière, la correction du signal au pas d'itération n est exprimée comme une somme pondérée des deuxièmes différences $\mathbf{D}_m^2 \otimes \mathbf{F}^{(n-1)}$ dans les directions axiales et diagonales.

Limitation symétrique du flux

La diffusion est contrôlée par l'introduction d'un coefficient d'atténuation pour la correction apportée par chacune des deuxièmes différences. Formellement, les coefficients d'atténuation sont représentés par le signal $\mathbf{A}_m^{(n)}$ tel que

$$\mathbf{A}_m^{(n)}(i, j) \in [0, 1], \quad \forall i, j \quad (6.88)$$

d'où la notation

$$\mathbf{F}^{(n)} = \mathbf{F}^{(n-1)} + \alpha \left\{ b \sum_{m \in \{0, 2\}} \mathbf{A}_m^{(n)} \mathbf{D}_m^2 + (1 - b) \sum_{m \in \{1, 3\}} \mathbf{A}_m^{(n)} \mathbf{D}_m^2 \right\} \otimes \mathbf{F}^{(n-1)} \quad (6.89)$$

Le lien qui existe entre les opérateurs de deuxième dérivée et ceux de première dérivée à gauche et à droite

$$\begin{aligned} \mathbf{D}_m^2 &= \mathbf{D}_{m-}^1 + \mathbf{D}_{m+}^1, & m \in \{0, 2\} \\ \mathbf{D}_m^2 &= \frac{1}{\sqrt{2}} (\mathbf{D}_{m-}^1 + \mathbf{D}_{m+}^1), & m \in \{1, 3\} \end{aligned} \quad (6.90.a-b)$$

met en évidence que le coefficient d'atténuation qui a été introduit limite le flux de diffusion de manière symétrique.

Contrôle du flux par la deuxième différence

La deuxième différence dans une des directions axiales ou diagonales est une mesure pour les discontinuités \bar{C}^0 et \bar{C}^1 dans cette direction [§4.4.1]. Dans le but de préserver les discontinuités des deux ordres, le flux de diffusion est donc limité par une fonction de seuil Θ_T appliquée à la deuxième différence, soit

$$\mathbf{A}_m^{(n)} = \Theta_T \left(\left| \mathbf{D}_m^2 \otimes \mathbf{F}^{(n-1)} \right| \right) \quad (6.91)$$

Seuil adapté à l'atténuation du bruit

Le but du lissage est d'atténuer le niveau du bruit afin de pouvoir abaisser le seuil de détection des discontinuités [§4.4.1]. Le seuil T_n est donc abaissé à chaque itération de sorte à ce qu'il soit proportionnel (et supérieur) au niveau estimé du bruit résiduel. C'est-à-dire

$$T_n = k \cdot \sigma_2^{(n-1)} \quad (6.92)$$

où $\sigma_2^{(n-1)}$ représente le niveau estimé du bruit sur la deuxième différence après exécution de l'itération $n-1$. k est une constante arbitraire, mais dont la valeur s'impose quasiment: une valeur de 3 à 4 constitue un choix sensé (probabilité de dépassement du seuil inférieure à 3‰ pour une distribution gaussienne du bruit).

Pour estimer le niveau de bruit $\sigma_2^{(n)}$, on utilise un modèle comprenant: premièrement, un bruit additif identiquement et indépendamment distribué et, deuxièmement, un lissage gaussien (convolution avec une gaussienne).

Dans ce modèle, si σ_0 est l'écart-type du bruit IID et si λ représente l'écart-type de la gaussienne, l'écart-type $\tilde{\sigma}_2$ du bruit sur la deuxième différence du signal lissé $\mathbf{D}_m^2 \otimes \tilde{\mathbf{F}}$ est donné par la relation

$$\tilde{\sigma}_2 = \frac{\sqrt{3}}{4\sqrt{\pi}} \sigma_0 \lambda^{-3} \quad (6.93)$$

(voir [Annexe C]).

En considérant que le lissage effectué par N itérations de diffusion équivaut à celui d'une gaussienne dont l'écart-type λ vérifie la relation

$$\lambda^2 = 2\alpha N \quad (6.94)$$

(voir [§6.3.4]) et en considérant que le bruit dont on possède une estimé σ_0 de l'écart-type est atténué comme un bruit IID, on obtient une formule pour $\tilde{\sigma}_2$ qui se traduit directement dans la formule suivante pour le seuil T_n

$$\begin{aligned} T_1 &= k\sqrt{6} \sigma_0 \\ T_n &= k \frac{\sqrt{3}}{4\sqrt{\pi}} \sigma_0 (2\alpha(n-1))^{-1.5}, \quad n > 1 \end{aligned} \quad (6.95.a-b)$$

Détection explicite

La détection des discontinuités à chaque itération est rendue explicite en marquant les points de discontinuité dans une image booléenne

$$\mathbf{B}^{(n)} = \mathbf{B}^{(n-1)} \vee \left(\left| \mathbf{D}_m^2 \otimes \mathbf{F}^{(n-1)} \right| > \mathbf{T}_n \right) \quad (6.96)$$

Remarques

La méthode de détection des discontinuités qui vient d'être définie est appelée méthode MDMS (Minimal Detector Minimal Smoothing) en

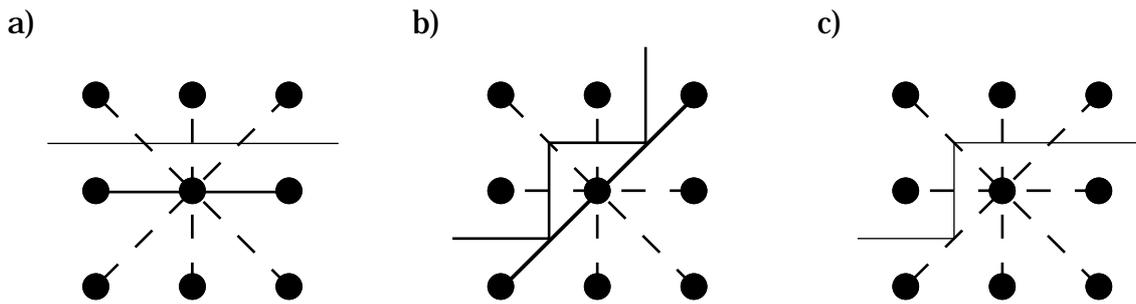


Figure 6.13: Effets de bord pour la méthode MDMS: principe. La contribution des deuxièmes différences dans les directions représentées en pointillé est annulée. Par conséquent, on obtient dans les situations (a) et (b) un lissage unidimensionnel et dans la situation (c) un point fixe.

référence à ses deux caractéristiques essentielles:

- l'utilisation d'un détecteur unique pour les discontinuités d'ordre 0 et 1
- la réalisation d'un lissage minimal adapté au niveau du bruit

L'image des points de discontinuité obtenue est multirésolutionnelle. En effet, les points sont détectés et marqués à différentes itérations et donc à différents degrés de lissage ou de résolution. Cet aspect peut être mis en évidence en codant les points de discontinuité par le numéro de l'itération à laquelle ils ont été détectés.

6.5.2 Algorithme MDMS

Afin d'obtenir un algorithme à partir de la méthode décrite précédemment, certains aspects doivent être encore précisés: il faut choisir les conditions de bord, la fonction de seuil Θ_T et les paramètres de diffusion. Finalement, l'algorithme diffère aussi de la méthode à deux points de vue: premièrement, une compensation de l'atténuation globale du flux est introduite, deuxièmement, la variation du seuil en fonction des itérations est faite différemment.

Conditions de bord

Comme la méthode MDMS est conçue pour préserver les discontinuités, un traitement automatique du bord peut être réalisé par l'introduction d'une condition de bord sous la forme d'une discontinuité. En l'occurrence, une discontinuité \bar{C}^0 est introduite en

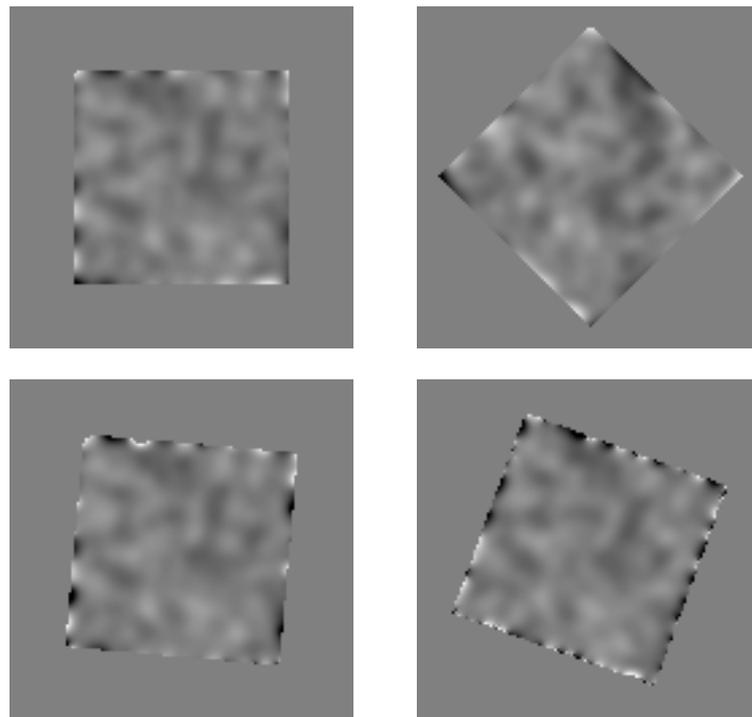


Figure 6.14: Effets de bord pour la méthode MDMS: exemples. Les images représentent, codé par des niveaux de gris, le bruit résiduel (signal bruité lissé moins signal non-bruité) après lissage avec la méthode MDMS. Le signal original est une fonction constante sur un domaine carré, à laquelle est ajouté un bruit pseudo-aléatoire de distribution gaussienne. La méthode MDMS a été appliquée avec un nombre d'itérations égal à 13 ($\lambda = 3$). A cause de la discrétisation, la forme des bords du carré varie en fonction de son orientation. On constate des qualités de lissage différentes.

attribuant aux positions à l'extérieur du bord (positions de valeur *NIL* ainsi que positions en bordure du tableau) un nombre de valeur extrême.

Conséquences pour la diffusion

La figure 6.13 montre trois formes typiques de bord. Le coefficient d'atténuation des deuxièmes différences pour lesquelles une des positions est en dehors du domaine est nul. Le nombre de deuxièmes différences dont la contribution est ainsi annulée dépend de la forme du bord. Si un point a plus de trois voisins qui n'appartiennent pas au domaine, aucune des deuxièmes différences n'est active et le point est fixe. Lorsque trois voisins n'appartiennent pas au domaine, une seule deuxième différence est active et on assiste à une diffusion unidimensionnelle. C'est le cas en présence d'un bord régulier de

direction 0 ou 45 degrés. La figure 6.14 montre l'influence de la forme du bord sur les caractéristiques du lissage dans la direction tangentielle au bord. Plus le bord est irrégulier, moins bon est le lissage, ceci à cause des points fixes.

Fonction de seuil

Pour des raisons de simplicité, la fonction de seuil binaire a été choisie.

Paramètres de diffusion

Le choix s'est porté sur les valeurs

$$\alpha = \frac{1}{3} \quad b = \frac{1}{3} \quad (6.97.a-b)$$

car elles correspondent, dans le cas de la diffusion linéaire, à un moyennage simple [figure 6.7], lequel fournit la meilleure atténuation d'un bruit IID à la première itération [HaSh92].

Avec ces valeurs de α et b , un pas de diffusion à flux limité symétriquement prend la forme

$$\mathbf{F}^{(n)} = \mathbf{F}^{(n-1)} + \frac{1}{9} \left\{ \sum_{m \in \{0,2\}} \mathbf{A}_m^{(n)} \mathbf{D}_m^2 + 2 \sum_{m \in \{1,3\}} \mathbf{A}_m^{(n)} \mathbf{D}_m^2 \right\} \otimes \mathbf{F}^{(n-1)} \quad (6.98)$$

Compensation de l'atténuation globale du flux

Les coefficients introduits pour atténuer la diffusion la ralentissent. Ils diminuent donc l'effet de lissage obtenu avec un même nombre d'itérations, ce qui n'est pas souhaité.

Comme c'est l'atténuation du flux dans une direction relativement aux autres directions qui est recherchée, on peut pallier à l'inconvénient du ralentissement de la diffusion en amplifiant α . Différentes solutions peuvent être envisagées. Il faut néanmoins faire attention de ne pas rendre le procédé itératif instable.

Dans l'algorithme tel qu'il a été implémenté, l'itération de diffusion (6.98) a été modifiée de la manière suivante:

$$\mathbf{F}^{(n)} = \mathbf{F}^{(n-1)} + \frac{1}{1+2\mathbf{S}^{(n)}} \left\{ \sum_{m \in \{0,2\}} \mathbf{A}_m^{(n)} \mathbf{D}_m^2 + 2 \sum_{m \in \{1,3\}} \mathbf{A}_m^{(n)} \mathbf{D}_m^2 \right\} \otimes \mathbf{F}^{(n-1)} \quad (6.99)$$

avec

$$\mathbf{S}^{(n)} = \sum_{m \in \{0,1,2,3\}} \mathbf{A}_m^{(n)} \quad (6.100)$$

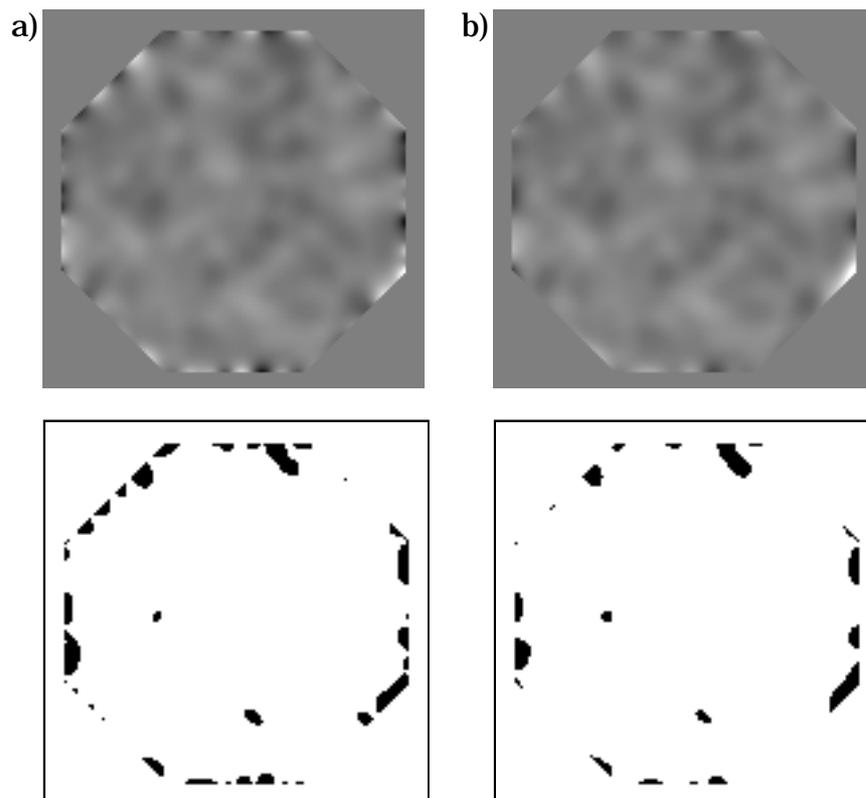


Figure 6.15: Accélération de la diffusion dans les bords. Bruit résiduel (signal bruité lissé moins signal non-bruité) après avoir appliqué la méthode MDMS avec 24 itérations ($\lambda = 4$), (a) sans, (b) avec accélération de la diffusion dans les bords. Haut: images de bruit résiduel codées par des niveaux de gris. Bas: images de bruit résiduel seuillées pour mettre en évidence les valeurs extrêmes. Le signal original est une fonction constante sur un domaine octogonal, à laquelle est ajouté un bruit pseudo-aléatoire de distribution gaussienne.

Autrement dit, on compense l'atténuation globale du flux.

Cette manière de faire a donné de bons résultats pratiques [figure 6.15]. En particulier, aucun problème d'instabilité n'a été rencontré dans les nombreux cas où l'algorithme a été appliqué (avec la fonction de seuil binaire).

Nous faisons remarquer que la manière avec laquelle la diffusion est accélérée est liée aux valeurs de α et b égales à $\frac{1}{3}$.

Variation du seuil en fonction du nombre d'itérations

Dans la méthode de base, la variation du seuil en fonction du nombre d'itérations a été déterminée en référence à un modèle comportant un bruit additif IID et un lissage gaussien. Pour une application prati-

<p>Signal d'entrée: F image de profondeur</p> <p>Signaux de sortie: $\tilde{\mathbf{F}}$ image de profondeur lissée B image des points de discontinuité</p> <p>Paramètres d'entrée: N nombre d'itérations σ_0 estimation du niveau du bruit pour le signal non-lissé k seuil relatif $k = 4$ (défaut) s vitesse d'abaissement du seuil $s = 1.5$ (défaut)</p> <p>Précautions: Vérifier que les valeurs <i>NIL</i> de F soient codées par un nombre de valeur extrême</p>
<p>Initialisation</p> <p>$\mathbf{F}^{(0)} = \mathbf{F}$ $\mathbf{B}^{(0)}(i, j) = \text{faux}, \forall i, j$</p>
<p>Itération</p> <p>$T_n = k\sqrt{6} \sigma_0 n^{-s}$</p> <p>$\mathbf{A}_m^{(n)} = \Theta_{T_n}^b \left(\left \mathbf{D}_m^2 \otimes \mathbf{F}^{(n-1)} \right \right), m \in \{0, 2\}$ $\mathbf{A}_m^{(n)} = \Theta_{2T_n}^b \left(\left 2\mathbf{D}_m^2 \otimes \mathbf{F}^{(n-1)} \right \right), m \in \{1, 3\}$</p> <p>$\mathbf{S}^{(n)} = \sum_{m \in \{0, 1, 2, 3\}} \mathbf{A}_m^{(n)}$</p> <p>$\mathbf{F}^{(n)} = \mathbf{F}^{(n-1)} + \frac{1}{1+2\mathbf{S}^{(n)}} \left\{ \sum_{m \in \{0, 2\}} \mathbf{A}_m^{(n)} \mathbf{D}_m^2 + \sum_{m \in \{1, 3\}} \mathbf{A}_m^{(n)} 2\mathbf{D}_m^2 \right\} \otimes \mathbf{F}^{(n-1)}$</p> <p>$\mathbf{B}^{(n)} = \mathbf{B}^{(n-1)} \vee \left(\mathbf{S}^{(n)} \neq \mathbf{4} \right)$</p>
<p>Arrêt</p> <p>$\tilde{\mathbf{F}} = \mathbf{F}^{(N)}$ $\mathbf{B} = \mathbf{B}^{(N)}$</p>

Figure 6.16: Algorithme MDMS.

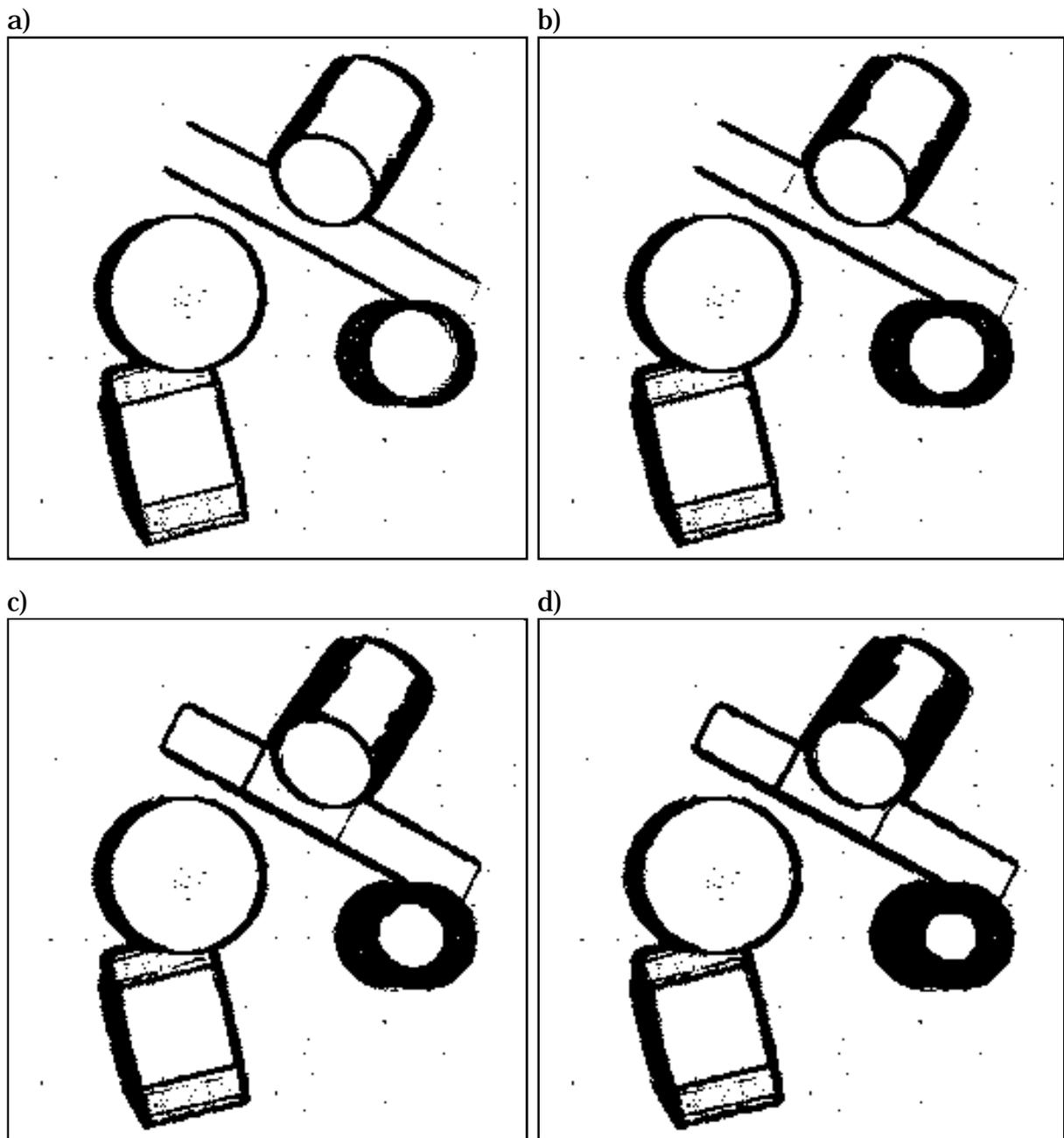


Figure 6.17: Exemple de résultats de l'algorithme MDMS. Image de profondeur BLOC12 (a) $n=2$, (b) $n=3$, (c) $n=4$, (d) $n=5$.

que de la méthode MDMS, il faut introduire une certaine tolérance par rapport à ce modèle idéal. Il faut aussi donner la possibilité à l'utilisateur de modifier la fonction $T(n)$.

La formule (6.95) est donc remplacée par l'expression

$$T_n = k\sqrt{6} \sigma_0 n^{-s} \quad (6.101)$$

k et s étant deux paramètres pouvant être modifiés par l'utilisateur avec

$$k = 4 \quad s = 1.5 \quad (6.102.a-b)$$

comme valeurs de défaut.

Conclusion

L'algorithme MDMS est maintenant défini dans tous les détails. La figure 6.16 en fait une description complète. La figure 6.17, comparée aux figures 4.13.b et 4.14.b, en illustre l'effet.

Pour terminer, nous indiquons le nombre d'opérations par itération et par rangel de l'algorithme MDMS. Ce nombre varie suivant si les coefficients des deuxièmes différences sont nuls ou non. Le nombre maximal d'opérations par rangel correspond à la situation où tous les coefficients d'atténuation sont égaux à 1. Ce coût de calcul maximal est: 18 additions, 9 comparaisons, 2 multiplications, 1 division, 1 opération booléenne.

6.5.3 Extensions et modifications de la méthode

Autre mesure de discontinuité

La méthode MDMS utilise la deuxième dérivée dans une des directions discrètes comme mesure de discontinuité. La courbure des profils dans les directions axiales et diagonales [§4.4.1] constitue un autre choix possible pour la mesure de discontinuité. Ce choix définit une méthode similaire.

IP vectorielles

La méthode MDMS peut facilement être adaptée aux IP de type vectoriel. En voici les éléments:

- Le lissage linéaire par diffusion est appliqué au signal discret de type vectoriel que représente l'IP. L'opération équivaut à appliquer la diffusion de manière indépendante aux signaux scalaires constitués par chacune des composantes du vecteur.
- La diffusion est limitée par l'introduction de coefficients d'atténuation scalaires.
- Les coefficients d'atténuation sont déterminés par une fonction de seuil appliquée à la mesure de discontinuité que constitue la grandeur $\|\mathbf{D}_m^2 \otimes \mathbf{P}\|$ [§4.4.2].

Formellement, rien ne change, si ce n'est la formule pour les coefficients d'atténuation

$$\mathbf{A}_m^{(n)} = \Theta_{T_n} \left(\left\| \mathbf{D}_m^2 \otimes \mathbf{F}^{(n-1)} \right\| \right) \quad (105)$$

où la fonction "valeur absolue" a été remplacée par la fonction "norme".

On peut vérifier facilement que dans le cas où l'image vectorielle est issue d'une IP cartésienne orthographe, c'est-à-dire

$$\mathbf{p}_{i,j} = (i\Delta x, j\Delta y, z) \quad (106)$$

l'algorithme est équivalent à celui qui a été défini pour les IP de type scalaire. La méthode que nous venons de définir constitue donc une forme généralisée de la méthode MDMS définie plus haut.

Changement de résolution

La méthode MDMS effectue un lissage progressif avec, à chaque itération, une détection des discontinuités. Le changement de résolution apporté par le lissage est celui de la méthode de diffusion explicite. Il est proportionnel à la racine carrée du nombre d'itérations. Rien n'empêche d'utiliser le même principe d'alternance entre lissage et détection des discontinuités avec une autre vitesse de changement de résolution (et év. un autre détecteur de discontinuité). Les moyens de réaliser un changement de résolution plus rapide sont: la méthode de diffusion D^\perp et le changement d'échelle par sous-échantillonnage.

Estimation multirésolutionnelle de propriétés différentielles

Le principe de lissage minimal, appliqué dans la méthode MDMS à une propriété différentielle (deuxième dérivée) dans le but de détection des discontinuités, peut être appliqué à la même propriété différentielle dans le seul but d'en faire l'estimation. Le principe peut aussi être appliqué à d'autres propriétés différentielles. Dans un tel cas, au lieu de mettre à jour une image des discontinuités, c'est une image de la propriété différentielle en question qui sera construite au fur et à mesure des itérations. Le résultat sera une image multirésolutionnelle de la propriété différentielle.

Il est évident que le principe de lissage minimal peut aussi être appliqué à la détection des discontinuités d'ordre 0 sur la base de la valeur du gradient.

6.6 Conclusion

Le lissage d'une image de profondeur a été considéré en relation avec la détection des discontinuités.

D'abord des méthodes de lissage particulières ont été présentées: lissage par diffusion, lissage par minimisation d'énergie de membrane, lissage par minimisation d'énergie de plaque mince. Les caractéristiques de lissage de chaque méthode ont été étudiées. Une attention particulière a été donnée à l'effet de bord. On a mis en évidence qu'à cause de l'effet de bord, la méthode de diffusion et la méthode de minimisation d'énergie de membrane engendrent une distorsion d'un signal correspondant à un plan incliné. Des conditions de bord ont été proposées pour éviter le plus possible cette distorsion.

L'étude comparative des trois types de méthodes donne la préférence aux méthodes de diffusion. Le lissage y est directement en rapport avec le nombre d'itérations, ce qui permet de réaliser un lissage ou, autrement dit, un changement de résolution progressif.

Dans un deuxième temps, on a montré comment les méthodes de lissage pouvaient être modifiées pour préserver les discontinuités. On a aussi mis en évidence que de telles méthodes effectuent une détection implicite des discontinuités. Ensuite, une revue des méthodes de lissage préservant à la fois les discontinuités d'ordre 0 et celles d'ordre 1 d'une image de profondeur a été effectuée.

Finalement, on a proposé une méthode conjointe de lissage et de détection des discontinuités, appelée méthode MDMS. Elle se distingue des autres essentiellement par sa simplicité et sa cohérence. Simplicité, parce qu'elle combine de manière propice une des méthodes numériques de lissage par diffusion avec un des détecteurs des discontinuités d'ordre 0 et 1. Cohérence, parce qu'elle met en relation le lissage avec le niveau du bruit, pour l'atténuation duquel le lissage est utilisé.

«Dans la pratique, on ne peut utiliser que des méthodes simples.»

Olivier Besson

Chapitre 7

Résultats de segmentation par détection des discontinuités

Ce chapitre présente et analyse les performances pratiques de l'algorithme de segmentation par détection des discontinuités que nous proposons. L'algorithme est construit à partir des composantes qui ont été décrites dans les chapitres précédents.

Nous commençons par décrire ce nouvel algorithme, ainsi que la base de données et l'environnement informatique utilisés pour son expérimentation.

Ensuite, nous décrivons précisément les conditions d'expérimentation et présentons un ensemble représentatif de résultats obtenus.

Finalement, nous évaluons et discutons les résultats.

7.1 Mise en place de l'expérimentation

7.1.1 Algorithme

L'algorithme de segmentation par détection des discontinuités est réalisé à l'aide des composantes suivantes [figure 7.1]:

- (1) estimation du bruit par l'algorithme décrit au chapitre 5
- (2) détection des discontinuités par l'algorithme MDMS [chapitre 6]
- (3) étiquetage des composantes connexes de l'image des points de discontinuité (ceux-ci constituent les frontières des régions)

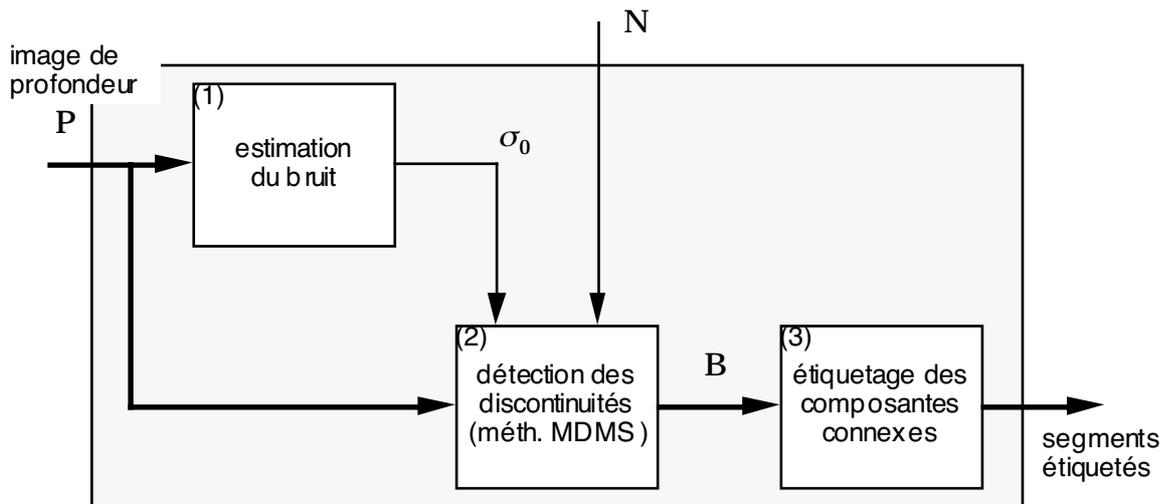


Figure 7.1: Algorithme de segmentation par détection des discontinuités.

Paramètres

Pour l'algorithme d'estimation du bruit, le seuil relatif servant à la troncature itérative de l'histogramme est fixé à sa valeur de défaut (3).

En ce qui concerne l'algorithme MDMS, on utilise aussi les valeurs de défaut pour le seuil de détection des discontinuités ($k = 4$, $s = 1.5$).

L'algorithme de segmentation possède donc un seul paramètre. C'est le nombre d'itérations de l'algorithme MDMS.

Variantes

Il existe par conséquent deux utilisations possibles: soit le nombre d'itérations est fixé et le résultat est une segmentation unique, soit il n'est pas fixé et le résultat est une suite hiérarchique de segmentations. (La suite est hiérarchique parce que l'algorithme MDMS garantit une croissance monotone de l'ensemble des points de frontière [chap.6].)

Remarques:

- L'estimation du bruit n'est pas une composante nécessaire de l'algorithme de segmentation. Par exemple, lorsqu'on traite plusieurs IP acquises dans des conditions similaires, la meilleure solution est certainement d'effectuer une seule estimation du bruit sur une IP bien choisie.

	type	taille	quantification	$\Delta x \times \Delta y$	Δz
Besl	brut/orthonormée ¹⁾	128x128	8 bits	–	–
CNRC	orthographique	256x256	16 bits	var.	0.01mm
GRASP	orthographique	var.	8 bits	1.0x1.0mm	1.5mm
EPFZ	brut/orthonormée ¹⁾	256x256	32 bits (float)	–	–
IMT	orthonormée	256x256	8 bits	0.75x0.75mm	0.75mm

1) léger effet de perspective

Table 7.1: Caractéristiques des IP constituant la base de données.

- L'image des points de frontières qui est le résultat de la première itération de l'algorithme MDMS correspond approximativement à l'image des valeurs tronquées [chap.5] (puisque les valeurs respectives 3 et 4, pour le seuil relatif de l'algorithme d'estimation du bruit et celui de l'algorithme MDMS, sont proches).

7.1.2 Base de données

Un total d'environ 260 IP ont été rassemblées. Elles proviennent de cinq sources différentes: P.J. Besl (25 IP), CNRC¹ (205), GRASP²(10), EPFZ³ (10), IMT⁴ (10). Leurs caractéristiques respectives sont représentées dans la table 7.1. Voici quelques informations supplémentaires concernant notamment les systèmes de mesure utilisés et la qualité des IP.

Besl Il y a deux types d'IP. Une partie sont des IP synthétiques bruitées, l'autre partie sont des IP acquises à l'aide d'un système à laser de type temps de vol provenant de l'ERIM⁵. Les IP synthétiques présentent différents niveaux de bruit. La qualité des IP acquises est considérée comme mauvaise.

¹Division de Génie Electrique, Conseil National de Recherches Canada, Ottawa, Canada

²GRASP Laboratory, Departement of Computer and Information Science, University of Pennsylvania, Philadelphia, USA

³Robot Vision Group, Institute for Communication Technology, Swiss Federal Institute of Technology, Zurich, Switzerland

⁴Groupe d'Electronique et de Reconnaissance des Formes, Université de Neuchâtel, Suisse

⁵Environmental Research Institute of Michigan, Ann Arbor, Michigan, USA

CNRC Les IP sont issues d'un système perfectionné de type triangulation laser [RBTD87, RiCo88]. Elles sont d'excellente qualité.

GRASP Nous n'avons pas d'information concernant le système d'acquisition. Les IP sont de qualité moyenne.

EPFZ Le système d'acquisition utilisé est de type triangulation, suivant le principe de lignes codées par une séquence temporelle binaire [PoAl82]. Un bruit important, lié au motif de lignes projeté, est présent dans les IP. La qualité des IP est moyenne.

IMT L'acquisition a été faite avec notre système de triangulation présenté au chapitre 2. Les IP sont de bonne qualité.

7.1.3 Plateforme

Les algorithmes d'estimation du bruit et de détection des discontinuités ont été implantés dans deux environnements différents: (1) sur Macintosh™, sous la forme d'une extension du programme de traitement d'image "NIH Image" version 1.44, (2) sur SUN™, en tant que tâches dans l'environnement KBVision™.

"NIH Image" est un programme du domaine public provenant du "National Institutes of Health" aux États-Unis. Il constitue un excellent outil pour le traitement des images à 256 niveaux de gris (8 bits). Les sources sont écrites en Think Pascal™. Elles contiennent un module "User" destiné à faciliter l'extension du programme. L'auteur de cette thèse a utilisé cette opportunité pour introduire dans le programme le traitement des images de format flottant. Cette extension lui a fourni un outil de développement et de test pour son travail de recherche sur les traitements préliminaires des images de profondeur. Entre autres, différents algorithmes d'estimation du bruit et de détection des discontinuités ont pu être testés.

KBVision est un environnement logiciel commercial d'analyse et d'interprétation d'images. Pour chacun des deux algorithmes d'estimation du bruit et de détection des discontinuités, une tâche a été créée sous KBVision à partir de routines écrites en C [Clau93].

Pour les résultats de segmentation présentés dans ce chapitre, l'étiquetage des composantes connexes de l'image des points de discontinuité a été effectué à l'aide d'une tâche standard de KBVision.

7.2 Exemples de segmentation

7.2.1 Choix des images de profondeur

Sur le total de 260 IP à disposition, l'algorithme de segmentation a été appliqué à une quarantaine d'IP environ. Nous présentons ici les résultats de segmentation pour 24 IP, soit

- 4 de chacune des 5 sources, choisies de sorte à offrir une palette diversifiée au niveau du contenu de la scène
- 4 IP supplémentaires sur lesquelles d'autres méthodes de segmentation ont été appliquées [chap.3].

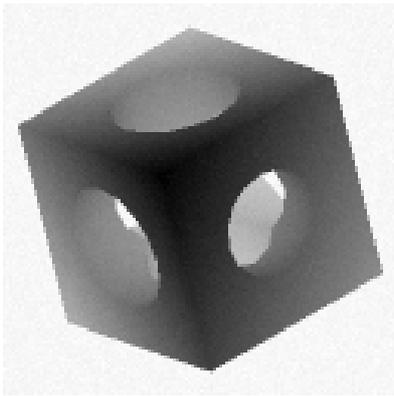
Les figures 7.2 et 7.3 représentent, par codage en niveaux de gris, l'ensemble des IP choisies. Voici quelques précisions concernant les scènes représentées par ces IP.

Besl Blcka est une IP synthétique bruitée. La scène est composée uniquement d'un cube percé de trois trous. Le bruit est IID (identiquement indépendamment distribué). La distribution est gaussienne d'écart-type 2.3. Officer et Ringbloc sont des IP issues du système d'acquisition de l'ERIM⁶. Officer est une vue d'un bureau. Ringbloc est l'IP d'une bague à faces planes et cylindriques posée sur un polyèdre en forme de L dont un coin est tronqué. Les données originales de Rpt7 qui se présentent sous la forme d'une liste de coordonnées (x,y,z) proviennent de l'INRIA⁷. Elles ont été converties en une IP 64x64, qui a été étendue à une IP 128x128. Rpt7 représente une pièce de voiture.

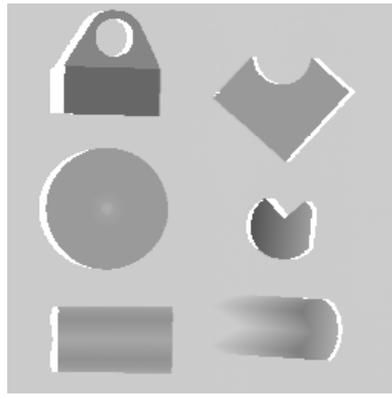
CNRC Toutes les IP sélectionnées ont le même intervalle d'échantillonnage, soit 0.8mm. Dans le cas des IP BLOC19, BLOC24 et BLOC30, la scène est constituée d'objets manufacturés ayant des faces planes ainsi que des faces courbes simples. BLOC19 et BLOC24 ne présentent pas de cas d'occlusion, si ce n'est par rapport à la surface sur laquelle sont posés les objets. Dans BLOC30, les cas d'occlusion sont nombreux. HELI01 représente une hélice.

⁶Environmental Research Institute of Michigan, Ann Arbor, Michigan, USA

⁷Institut National de Recherches en Intelligence Artificielle, Sophia-Antipolis, Valbonne, France



Blcka



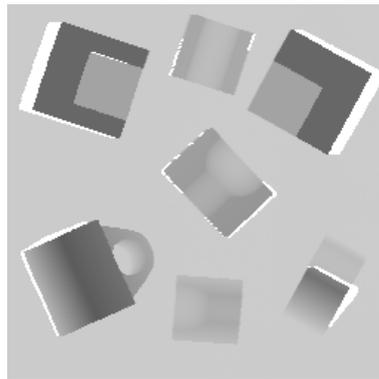
BLOC19



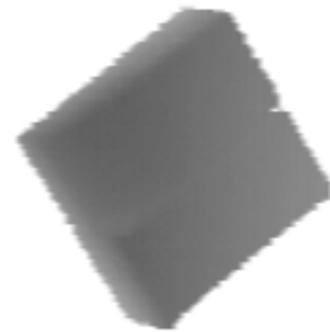
Cup



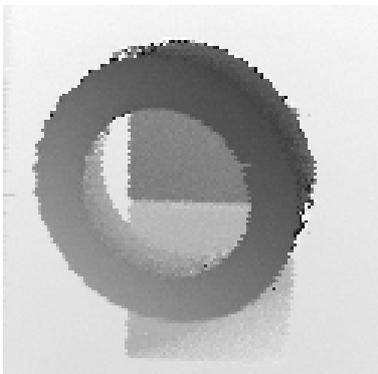
Officer



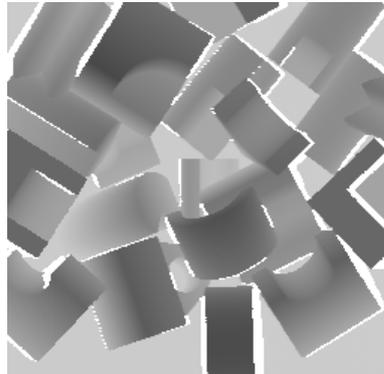
BLOC24



DiffPoly



Ringbloc



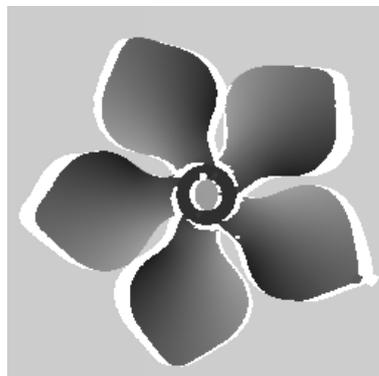
BLOC30



Image1



Rpt7



HELI01



NBS

Figure 7.2: Images de profondeur (1ère partie).

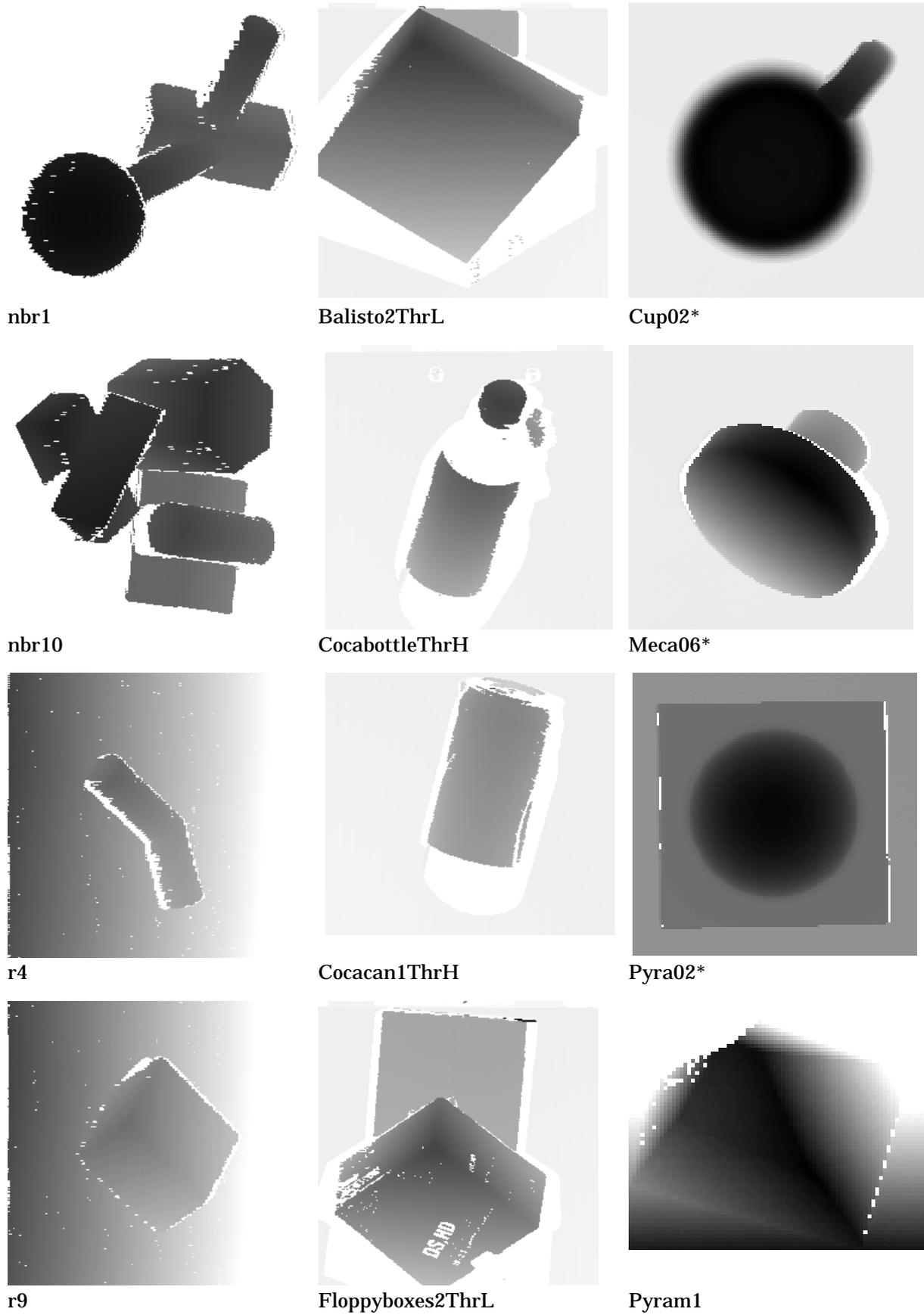


Figure 7.3: Images de profondeur (2ème partie).

GRASP Cup est l'IP d'une tasse. DiffPoly représente un objet à facettes dont deux des facettes se joignent sans former de discontinuité \bar{C}^1 . La scène de Image1 est constituée d'un prisme triangulaire et de deux demi-cylindres, l'ensemble est posé sur une boîte. NBS représente un objet provenant du Bureau National Américain des Standards. L'objet présente deux trous, diverses faces planes et courbes.

EPFZ Nbr1 représente un objet composé de plusieurs primitives volumiques simples: un parallélépipède, deux cylindres et une sphère. Les rangels représentant la surface sur laquelle repose l'objet ont été supprimés. Il en est de même pour nbr10. Cette deuxième IP représente plusieurs objets empilés, ce sont des polyèdres et un cylindre. La scène de r4 est constituée d'un cylindre coudé. R9 représente un polyèdre.

IMT Les IP représentent des emballages de divers produits de consommation. La scène de Balisto2ThrL est constituée d'une boîte de barres de céréales Balisto™, posée de manière oblique sur une autre boîte. CocaBottleThrH représente une bouteille en PET de Coca-Cola Light™. CocaCan1ThrH représente une boîte en alu de Coca-Cola™. FloppyBoxes2ThrL représente 2 boîtes de disquettes de 3M™.

Divers Pyram1 est une IP cartésienne orthographique issue d'une ancienne série d'IP de l'IMT. Cup02* est une IP issue de la collection du CNRC, réduite à une taille 128x128 par sous-échantillonnage. Il en est de même pour Meca06* et Pyra02*, ces IP ont été de plus réduites à une quantification sur 8 bits.

7.2.2 Exécution de la segmentation et résultats

L'algorithme de segmentation décrit à la section 7.1.1 a été appliqué aux IP choisies. Le nombre d'itérations de l'algorithme MDMS a été fixé par un opérateur a posteriori [figure 7.4], de sorte à ce que ce nombre soit optimal (la question de l'optimalité sera discutée à la section 7.3).

Les résultats obtenus sont décrits par la table 7.2, ainsi que par les figures 7.7 et 7.8. Les figures 7.5 et 7.6 présentent les résultats intermédiaires constitués par les images des points de discontinuité.

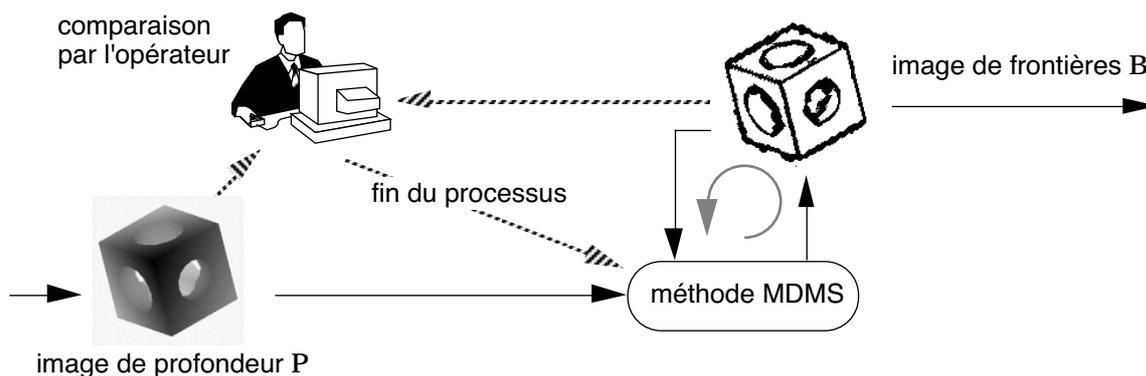


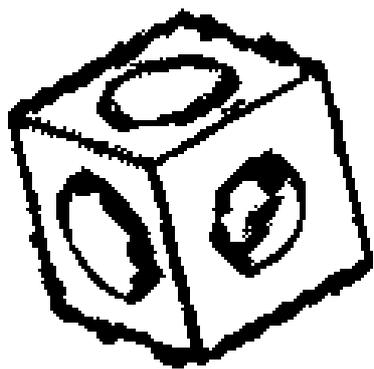
Figure 7.4: Exécution de la segmentation

Image de profondeur	bruit ¹⁾ σ_0	bruit ²⁾ σ_0 [mm]	#iter. MDMS N
Blcka	2.27	-	13
Officer	0.54	-	16
Ringbloc	1.76	-	11
Rpt7	0.56	-	9
BLOC19 (0.8mm)	3.87	0.039	3
BLOC24 (0.8mm)	4.11	0.041	2
BLOC30 (0.8mm)	4.23	0.042	2
HELI01 (0.8mm)	4.94	0.049	2
Cup (150x150)	0.446	0.669	8
DiffPoly (140x150)	0.434	0.658	8
Image1 (256x80)	0.519	0.779	17
NBS (180x180)	0.521	0.781	11
nbr1	0.354	-	8
nbr10	0.395	-	14
r4	0.538	-	9
r9	0.473	-	15
Balisto2ThrL	0.272	0.204	6
CocabottleThrH	0.176	0.132	6
Cocacan1ThrH	0.203	0.152	2
Floppyboxes2ThrL	0.279	0.209	6
Cup02* (128x128x16bits)	4.75	0.048	3
Meca06* (128x128x8bits)	0.190	0.190	4
Pyra02* (128x128x8bits)	0.200	0.200	3
Pyram1 (63x78x16bits)	13.9	0.139	8

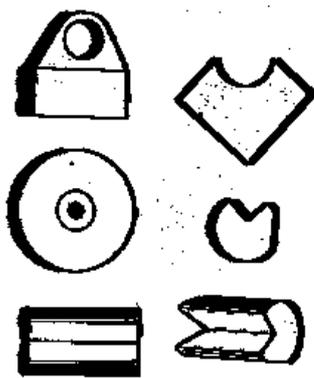
1) en unités de quantification

2) en unités absolues

Table 7.2: Résultats numériques



Blcka



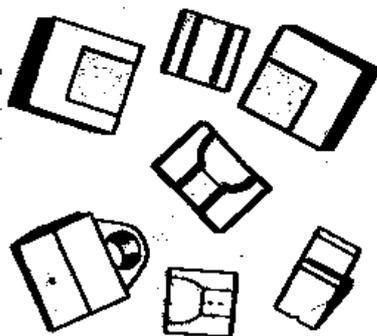
BLOC19



Cup



Officer



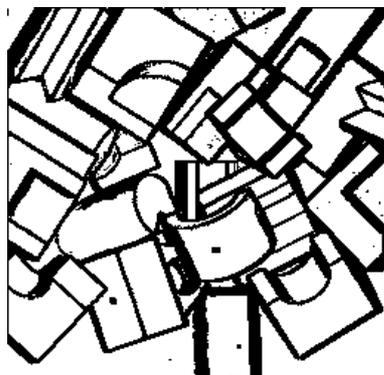
BLOC24



DiffPoly



Ringbloc



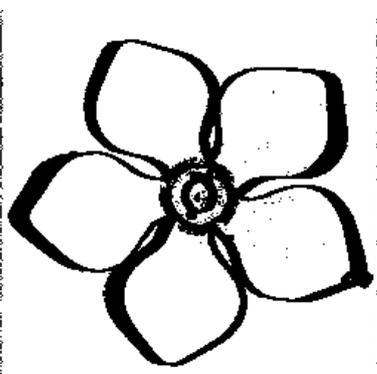
BLOC30



Image1



Rpt7

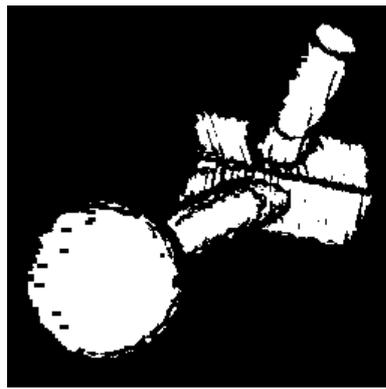


HELI01

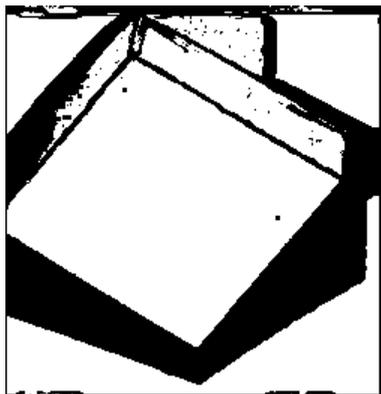


NBS

Figure 7.5: Images des points de discontinuité (1ère partie).



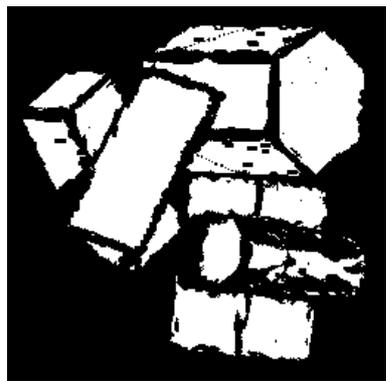
nbr1



Balisto2ThrL



Cup02*



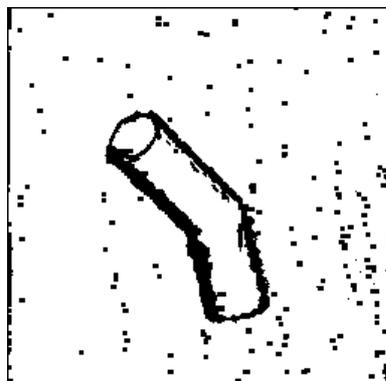
nbr10



CocabottleThrH



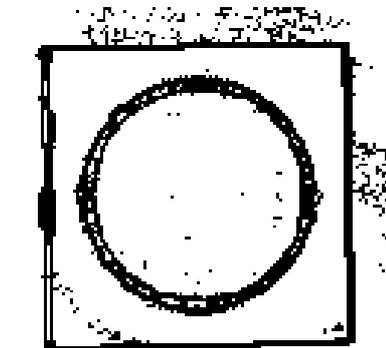
Meca06*



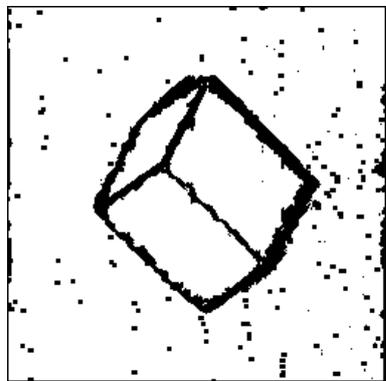
r4



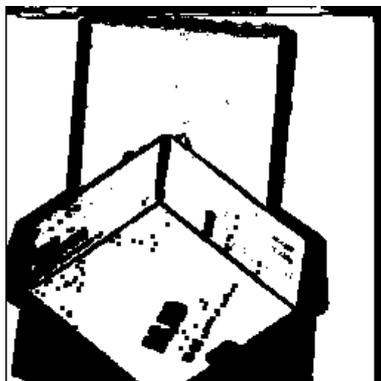
Cocacan1ThrH



Pyra02*



r9

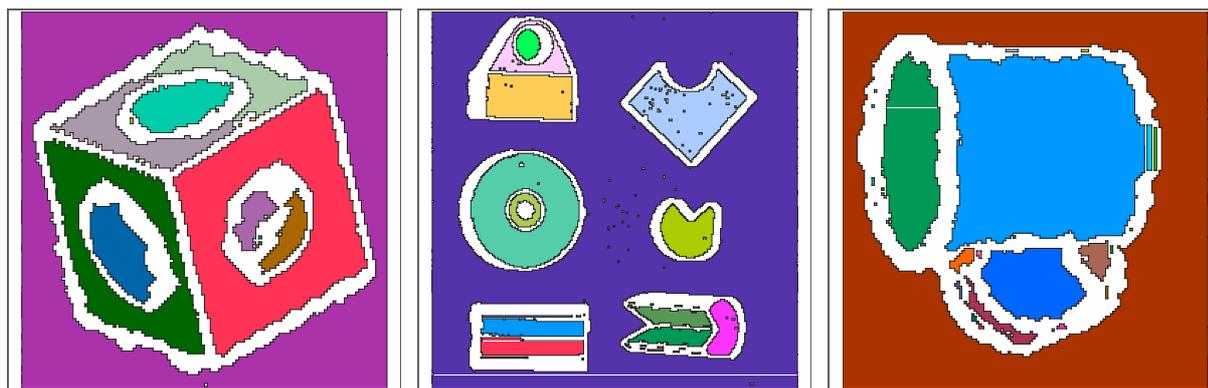


Floppyboxes2ThrL



Pyram1

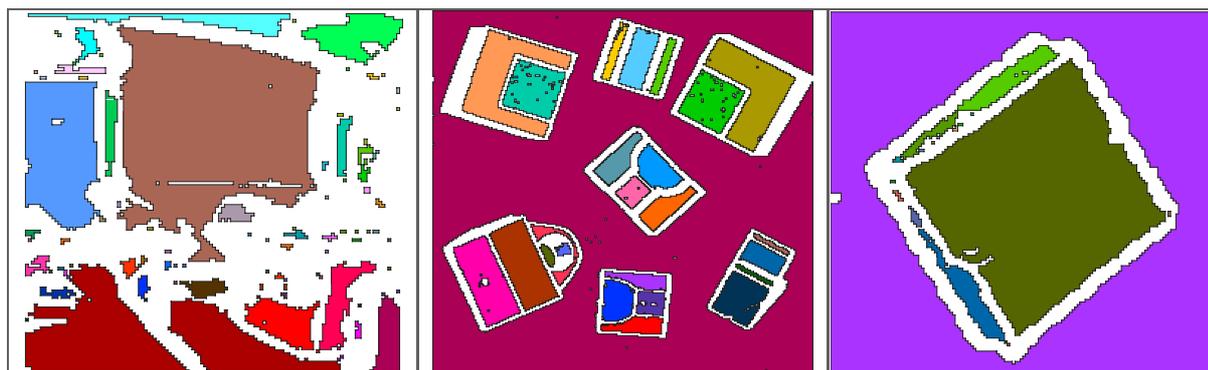
Figure 7.6: Images des points de discontinuité (2ème partie).



Blcka

BLOC19

Cup



Officer

BLOC24

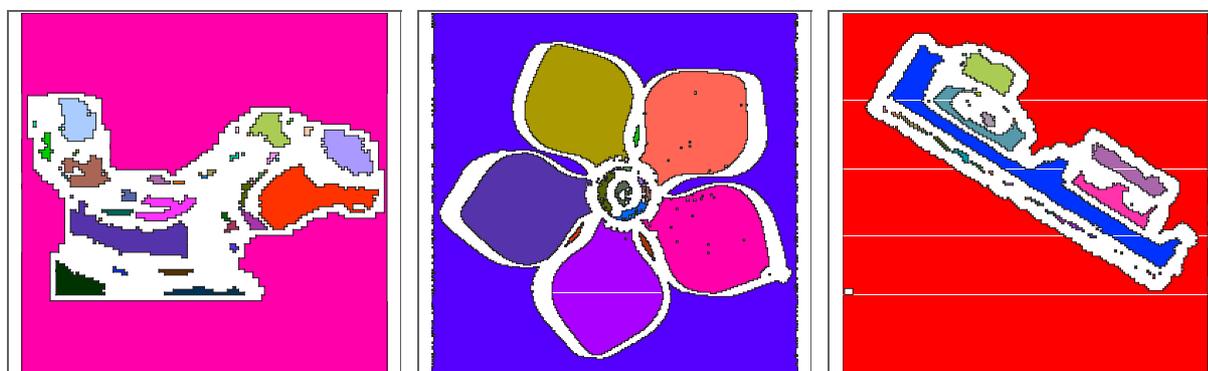
DiffPoly



Ringbloc

BLOC30

Image1



Rpt7

HELI01

NBS

Figure 7.7: Images des segments (1ère partie).

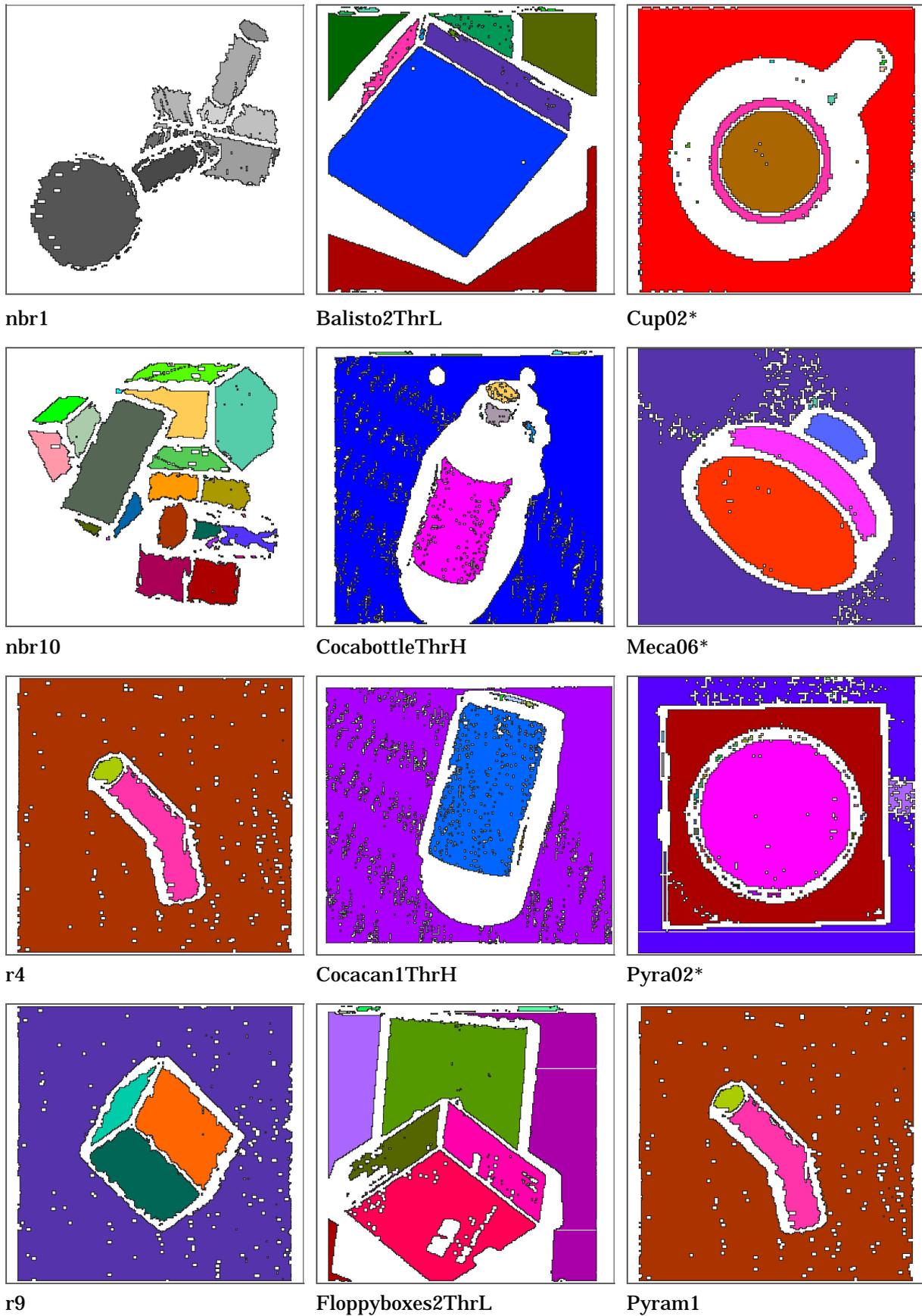


Figure 7.8: Images des segments (2ème partie).

7.3 Analyse et évaluation des résultats

L'évaluation des résultats nécessite une référence. A notre avis cette référence ne peut être définie de manière objective que s'il existe pour l'IP un modèle a priori de facettes algébriques.

Nous divisons donc l'ensemble des IP en deux catégories: avec et sans modèle de facettes [table 7.3].

7.3.1 Images de profondeur à modèle de facettes

Dans ce cas, la segmentation de référence est la suivante: un et un seul segment par facette, s'il n'y a pas d'occlusions; pour une facette occlue, autant de segments que de régions connexes dans la facette.

Erreurs de segmentation

Il y a trois types (non-exclusifs) de segmentation incorrecte: sur-segmentation (+), sous-segmentation (-) et disparition de facette (×).

La table 7.3 présente notre appréciation des résultats selon les règles qui ont été établies plus haut. On constate de bons résultats d'ensemble. Il y a cependant neuf IP qui présentent une segmentation partiellement incorrecte.

La face supérieure du cube de Blcka est séparée en deux (sur-segmentation). Ringbloc présente un cas unique de sous-segmentation: biais du polyèdre par rapport au plan de base. Dans l'exemple BLOC19, il y a disparition de quatre petites facettes dans le bas de l'image. DiffPoly est un exemple de sous-segmentation et de sur-segmentation conjointes. NBS présente deux cas de disparitions de facettes (facettes planes devant et à gauche). nbr1 et nbr10 sont deux cas de sur-segmentation. r4 est un cas de sous-segmentation due à l'absence de discontinuités entre les facettes cylindriques qui forment l'objet représenté. Finalement, on constate la disparition d'une des faces de la pyramide Pyram1.

Nombre optimal d'itérations

La sous-segmentation constitue l'erreur la plus grave. En effet, il est plus facile de réunir deux régions appartenant à la même facette que de diviser correctement une région chevauchant deux facettes (dans le deuxième cas, il s'agit de résoudre un nouveau problème de segmentation).

Image	modèle de facettes	Sgm.	Disc.
Blcka	o	+	+
Officer	n		
Ringbloc	o	-	-
Rpt7	n		
BLOC19	o	×	+
BLOC24	o	✓	✓
BLOC30	o	✓	✓
HELI01	n		
Cup	n		
DiffPoly	o	- +	✓
Image1	o	✓	✓
NBS	o	×	+
nbr1	o	+	+
nbr10	o	+	+
r4	o	-	✓
r9	o	✓	✓
Balisto2ThrL	o	✓	✓
CocabottleThrH	n		
CocanThrH	o	✓	✓
Floppyboxes2ThrL	o	✓	✓
Cup02*	n		
Meca06*	o	✓	✓
Pyra02*	o	✓	✓
Pyram1	o	×	+

Table 7.3: Evaluation des résultats.

Pour cette raison, en tant qu'opérateur de la segmentation, nous avons considéré que le nombre optimal d'itérations était le nombre minimal pour lequel il n'y avait pas de sous-segmentation (à condition que le nombre d'itérations reste dans une marge raisonnable).

Causes d'erreur

On recense trois causes d'erreur de segmentation liées à la détection des discontinuités:

absence de discontinuités (au niveau du modèle). Cette erreur (sous-segmentation) est prévue [chap.2]. DiffPoly et r4 sont des exemples d'une telle situation.

fausse détection positive: détection d'une discontinuité qui n'existe pas dans le modèle (+). C'est la cause d'une sur-segmentation ou d'une disparition de facette. Cet effet apparaît soit en présence d'une facette courbe, soit dans le cas d'une facette plane de petite taille. Les exemples sont nombreux.

fausse détection négative: non-détection d'une discontinuité du modèle (-). Elle est responsable d'une sous-segmentation. Cette situation est rare puisqu'on a cherché à éviter autant que possible l'erreur de sous-segmentation. Le seul exemple est celui de Ringbloc, il correspond à un niveau de bruit élevé.

7.3.2 Autres

La segmentation de référence a été définie de manière subjective par l'opérateur, en relation avec le contenu de l'image. Le nombre optimal d'itérations est donc indirectement fixé par la condition de proximité du résultat par rapport à cette segmentation de référence.

Par exemple pour l'image Cup, la référence est la séparation de la poignée par rapport au corps de la tasse. Le nombre d'itérations a donc été choisi de sorte à ce que ce résultat soit atteint sans que la poignée de la tasse disparaisse. Dans le cas de HELI01, la référence est que les pales de l'hélice soient séparées du moyeu et que le segment représentant chaque pale soit complet, d'où le nombre d'itérations et le résultat.

En général, lorsqu'en tant qu'utilisateur on est confronté à la décision d'arrêt de l'algorithme MDMS, on a tendance à attendre jusqu'à ce que la plupart des régions aient une forme plus ou moins convexe.

7.4 Discussion

Les résultats de segmentation que nous avons présentés ont été réalisés en fixant a posteriori le nombre d'itérations de l'algorithme MDMS. Dans cette expérimentation, l'algorithme de segmentation s'est donc trouvé être une composante d'un système de vision dans lequel l'opérateur jouait le rôle d'une autre composante. Nous discutons maintenant le rôle que peut jouer notre algorithme ou, d'un point de vue plus général, notre méthode à l'intérieur d'un système de vision artificielle. Il y en a deux.

Mise en correspondance d'une IP avec un modèle d'objets à facettes

Les résultats obtenus montrent que la méthode de segmentation peut être utile à la mise en correspondance d'une IP avec des modèles d'objets à facettes. Elle est un prétraitement qui facilite le problème, à condition bien sûr que la correspondance de 1 à 1 entre segment et facette soit remise en cause par les traitements ultérieurs.

Dans cette perspective d'utilisation, la hiérarchie de segmentations que peut fournir l'algorithme MDMS est attractive.

Analyse à différents niveaux de détails

L'algorithme MDMS présente la caractéristique de masquer à chaque pas d'itération les plus petits détails. Il réalise donc pour ainsi dire une analyse d'une IP à différents niveaux de détails.

La combinaison de la segmentation par l'algorithme MDMS avec une approximation des segments par des surfaces algébriques permet de réaliser une méthode simple de représentation approximative (compression) de l'IP.

7.5 Conclusion

Dans ce chapitre, nous avons présenté les résultats de segmentation obtenus avec notre méthode basée sur la détection des discontinuités avec l'algorithme MDMS.

Couplé avec notre algorithme d'estimation robuste du bruit, l'algorithme MDMS a démontré un bon comportement et une capacité d'adaptation remarquable à différents types d'IP.

Les résultats obtenus permettent d'affirmer que l'utilisation de notre méthode de segmentation comme composante d'un système de vision artificielle s'avère intéressante à deux points de vue:

- Elle peut constituer un prétraitement utile à la mise en correspondance d'une IP avec des modèles d'objets à facettes.
- Elle offre une méthode itérative d'analyse d'une IP qui fait disparaître à chaque pas les détails les plus petits.

Chapitre 8

Conclusions

Ce travail de thèse a été consacré au problème de la segmentation des images de profondeur par détection des discontinuités. Nous en dressons ici le bilan.

8.1 Contributions

On peut les classer en deux groupes: apports théoriques, apports pratiques. Nous y rajoutons quelques perspectives d'extension.

8.1.1 Apports théoriques

Ce sont les développements utilisés pour mettre en perspective et justifier les méthodes proposées au niveau pratique. Les éléments traités ne sont pas fondamentalement nouveaux mais, exposés sous un éclairage différent et peut-être plus approfondi par rapport à ce qu'on trouve ailleurs dans la littérature, ils sont d'un intérêt certain. Nous en récapitulons l'essentiel.

Tout d'abord, nous avons mis en évidence le fait que la segmentation des images de profondeur comporte deux aspects essentiels: le critère d'homogénéité (modèle de surface), d'une part, et la méthode réalisant la segmentation, d'autre part. Dans les méthodes rencontrées dans la littérature, le critère d'homogénéité n'est pas toujours explicite. Nous avons cependant pu classer la quasi-totalité des méthodes suivant quatre critères d'homogénéité: type de courbure, surface plane, surface algébrique, et continuité d'ordre 1.

Le modèle de continuité d'ordre 1 (C^1) constitue un critère de segmentation minimale. Dans certains cas, il offre une légère sous-segmentation par rapport à ce qu'on souhaiterait obtenir. Il existe une méthode simple pour réaliser la segmentation suivant un critère de continuité C^1 . Elle procède par détection locale des discontinuités et étiquetage des composantes connexes. On rencontre dans la littérature des méthodes beaucoup plus complexes, dont on peut considérer que tout l'effort supplémentaire est consacré à forcer la fermeture des courbes de discontinuité, ce qui revient à renforcer le critère de continuité C^1 .

Nous avons présenté les différents détecteurs de discontinuité, avec leurs caractéristiques propres. Notre attention s'est portée sur les détecteurs de taille minimale servant à la détection conjointe des discontinuités d'ordre 0 et de celles d'ordre 1.

Finalement, nous avons fait une présentation approfondie de différentes méthodes de lissage et de leurs possibilités d'être modifiées pour préserver les discontinuités. Nous avons mis en évidence le principe commun à toutes les méthodes et avons souligné que rares sont les méthodes qui mettent en relation le lissage avec le niveau du bruit. Nous avons montré que la méthode de lissage par diffusion possède un avantage par rapport aux méthodes de lissage par minimisation d'énergie parce qu'elle facilite une approche multirésolutionnelle.

8.1.2 Apports pratiques

Ce sont essentiellement les deux méthodes d'estimation du bruit et de détection des discontinuités. Nous en récapitulons les points forts.

La méthode d'estimation du bruit est simple et efficace. Son succès est dû avant tout à la technique utilisée pour éliminer les observations aberrantes de deuxième dérivée.

La méthode de détection des discontinuités (MDMS) est originale à deux points de vue. Premièrement, elle combine de manière judicieuse le lissage par diffusion avec la détection conjointe des discontinuités d'ordre 0 et celles d'ordre 1. Deuxièmement, elle utilise une estimation du niveau de bruit résiduel pour contrôler le lissage et la détection des discontinuités. Ces deux éléments garantissent, d'une part, la simplicité de la méthode et, d'autre part, son succès en

permettant une bonne détection à la fois des discontinuités très marquées et de celles qui le sont moins.

La combinaison des deux méthodes et d'un étiquetage des composantes connexes s'est révélé être une bonne méthode pour la segmentation des images de profondeur suivant le critère de continuité C^1 .

8.1.3 Perspectives

Nous formulons ici quelques idées relatives à l'extension et à la mise en application des méthodes que nous avons développées.

Les méthodes proposées sont destinées au traitement des images de profondeur scalaires. La plupart du temps, une extension pour les images de profondeur vectorielles a été envisagée. Elle reste à faire expérimentalement.

Le principe de lissage minimal mis en oeuvre par la méthode MDMS peut être utilisé à d'autres fins que la détection des discontinuités d'ordre 0 et 1 d'une image de profondeur. Par exemple, pour obtenir une estimée robuste de la courbure de la surface. Il peut aussi être appliqué pour la détection des discontinuités d'ordre 0 d'un signal quelconque (par exemple contours d'une image d'intensité). En fait, ce principe peut être utilisé dans toutes les situations où un lissage est nécessaire.

Les itérations de la méthode MDMS établissent une relation d'ordre sur les rangels qui est une fonction de la "résolution locale" du signal, ou autrement dit, de la "densité d'information". A notre avis, l'ordre défini par la méthode MDMS pourrait fournir: un ordre de sélection des noeuds pour l'approximation des rangels par une surface de type spline et aussi un ordre de sélection des rangels pour une mise en correspondance partielle entre deux images représentant le même objet.

Finalement, nous remarquons que les méthodes simples que nous avons développées présentent des caractéristiques similaires à certains phénomènes qui ont été identifiés dans la vision humaine. Tout d'abord, la méthode d'estimation du bruit en détectant les valeurs aberrantes réalise une mise en évidence d'exceptions par rapport à un contexte. On peut y voir un comportement de focalisation d'attention. D'autre part, le procédé itératif de la méthode MDMS comporte l'aspect dynamique qui a été identifié dans

les traitements préliminaires du signal visuel par le cerveau humain [Burg92].

8.2 Conclusion

En traitant du problème de la segmentation des images de profondeur par détection des discontinuités, cette thèse s'est confinée dans un domaine très pointu de la vision artificielle. A l'intérieur de ce domaine, pourtant, un travail important a été effectué. L'intérêt de cette thèse repose sur trois éléments.

Tout d'abord, elle réalise une étude approfondie des problèmes que constituent, dans le traitement des images de profondeur, la segmentation, la détection des discontinuités, l'estimation et l'atténuation du bruit. Ensuite, elle fait une présentation à la fois synthétique et précise des méthodes existantes pour résoudre ces problèmes. Finalement, elle propose des méthodes de détection des discontinuités et de traitement du bruit, dont le mérite est de réduire le traitement à ce qui semble être le strict minimum nécessaire.

Aussi, si nous sommes conscient que la contribution de cette thèse est modeste au niveau du progrès de la vision artificielle, nous osons néanmoins prétendre qu'elle peut être utile aux personnes s'intéressant au traitement des images de profondeur.

«Moi, en tant qu'homme, je n'existe pas seulement en tant que créature individuelle, mais je me découvre membre d'une grande communauté humaine.»

Albert Einstein

Remerciements

Une thèse est un travail personnel. Pourtant, beaucoup de personnes y contribuent. On doit déjà admettre cet état de fait lorsqu'on ne considère que le travail de thèse proprement dit, mais à plus forte raison lorsqu'on prend en compte la formation et le travail de recherche qui l'ont précédé.

J'ai fouillé dans mes souvenirs pour essayer de faire une synthèse objective des contributions dont j'ai bénéficié. J'ai forcément oublié certains éléments et j'en ai déformé d'autres. Que les personnes concernées me pardonnent! Merci à tous.

Ont contribué à mon travail de thèse

Mon directeur de thèse, Heinz Hügli, a veillé à ce que mon travail soit complet et cohérent. Il m'a poussé à mettre en perspective les méthodes que j'ai développées en faisant un état de l'art approfondi du domaine et il m'a encouragé à les illustrer par des résultats expérimentaux suffisamment nombreux et divers. Il a investi beaucoup de temps et de patience pour réviser les nombreuses versions du rapport de thèse qui ont précédé celle-ci.

Messieurs les experts, Frank Ade et Olivier Besson, ont examiné mon rapport de thèse avec beaucoup d'attention. Ils l'ont commenté de manière constructive et ont proposé des corrections qui m'ont permis de l'améliorer. D'autre part, tous deux ont contribué à ce travail bien avant de jouer le rôle d'expert.

Frank Ade a attiré mon attention sur des publications de Perona et Malik et de Saint-Marc, Chen et Medioni, qui ont eu une influence sur mon travail.

Olivier Besson a répondu à certaines questions de mathématique relatives à la continuité de surface et aussi concernant les méthodes de calcul numérique que j'ai utilisées.

Jean-Pierre Amann a contribué de manière significative à ce travail. Conscient de mes points faibles, il m'a aidé à planifier mon travail de rédaction et à me procurer le matériel informatique nécessaire à sa réalisation. D'autre part, le Docteur s'est toujours montré disponible à réaliser les travaux qui pouvaient m'être utiles, quels qu'ils soient. Ainsi, il a mis en forme plusieurs figures, il a réalisé la mise en page de la majeure partie du document et s'est occupé des problèmes d'impression. Dans un tout autre domaine, mais qui a aussi son importance, il m'a offert l'hospitalité lors de mes nombreux séjours à Neuchâtel après la fin de mon engagement à l'IMT.

François Tièche m'a accompagné et soutenu dans mon travail. Ayant lui-même travaillé sur le sujet lors de son diplôme, il a été pour moi un interlocuteur privilégié. Nous avons échangé nos points de vue sur de nombreux aspects du traitement des images de profondeur et de la vision artificielle en général. Nous avons entre autres discuté de certains concepts fondamentaux tels que représentation et modèle. Je me souviens par exemple d'un échange de courrier électronique que nous avons eu sur le sujet au courant du mois d'avril 94 et qui a été pour moi très fructueux. François a aussi contribué aux travaux de détail. Il a mis en forme plusieurs figures. Il a vérifié de nombreuses formules mathématiques, qui, pour une part importante, n'apparaissent malheureusement plus dans cette version finale du rapport de thèse!

Philippe Gingins s'est intéressé à mon travail pour en utiliser les résultats. Les questions qu'il m'a posées et les discussions qui ont suivi m'ont permis de progresser dans la compréhension de certains aspects du problème de la vision artificielle et d'entrevoir des applications autres que la segmentation pour les méthodes que j'ai développées.

François Corthay et Claudio Facchinetti ont révisé plusieurs chapitres. Ils ont vérifié l'orthographe et la numérotation des

équations et des figures. Ils m'ont aussi proposé des améliorations quant au style. François Corthay a contribué à améliorer radicalement l'introduction dans sa structure.

Stéphane Brunet et Christian Schütz ont participé à la production et à la mise en forme des résultats expérimentaux présentés dans ce rapport.

Laurent Claud a implémenté dans un environnement logiciel commercial les algorithmes que j'ai développés dans ma thèse. Il a ainsi vérifié mon travail.

Louisa Grisoni, Alex Heubi, Philippe Thévenaz et Ken Weible ont aussi apporté leur contribution, l'un en vérifiant quelques formules mathématiques, un autre en mettant en forme une figure, un troisième en me prêtant le matériel informatique dont j'avais besoin...

Jean-William Brandt a veillé à la bonne qualité de la reproduction du document.

Plusieurs autres personnes, issues du cercle de la famille, de celui des amis ou encore de celui des collègues de travail, m'ont proposé leur aide. Si je ne les ai pas sollicités c'est qu'il m'était impossible de déléguer certains travaux. D'autres personnes encore m'ont supporté par leurs encouragements.

Daniel Osherson et Thomas Breuel, responsables du projet sur lequel je travaille à l'IDIAP, m'ont laissé consacrer du temps à l'achèvement de ma thèse.

Paul J. Besl, Alok Gupta et Marjan Trobina ont mis à ma disposition les bases de données (images de profondeur) qu'ils ont utilisées pour leurs travaux et sur lesquelles j'ai pu expérimenter mes méthodes.

En publiant leurs travaux, de nombreux chercheurs m'ont fait profiter de leurs connaissances. Je pense notamment à plusieurs membres des groupes suivants:

- Fachgruppe fur Bildwissenschaft, Institut fur Kommunikations-technik, ETH Zürich, Dr. Ade
- Fachgruppe fur Künstliche Intelligenz, IAM, Universität Bern, Prof. H. Bunke
- Groupe de Vision, CUI, Université de Genève, Prof. Thierry Pun

- GRASP Laboratory, University of Pennsylvania, Prof. Ruzena Bajcsy
- Computer Vision Laboratory, CS Department, University of Southern California, Prof. Gerard Medioni,
- EECS Department, University of Michigan, Prof. Ramesh C. Jain
- Center for Intelligent Machines, McGill University, Prof. Frank Ferrie

D'autres chercheurs m'ont prodigué leurs conseils lors d'une rencontre dans une conférence ou à une autre occasion. Je me souviens par exemple d'une rencontre avec Ricardo Leonardi à Swiss Vision'92, lors de laquelle il m'a fait remarquer qu'une segmentation basée sur les frontières pouvait jouer un rôle important dans le cas des images de profondeur.

Mon travail de recherche a été financé en partie par un poste d'assistant et en partie par des projets de recherche du Fonds National Suisse pour la Recherche Scientifique (FN 2588, FN 2265, FN 20-25572, PNR23 - 4023 037027). Par l'intermédiaire de ces fonds, ce sont les contribuables suisses, et neuchâtelois en particulier, qui ont financé mon travail.

Ont contribué à ma formation

Mes parents m'ont donné une très bonne éducation. Je retiendrai en premier lieu l'exemple qu'ils m'ont donné: mon papa, par sa volonté et sa persévérance, ma maman, par son courage et sa patience. En grandissant à mes côtés, mon frère m'a beaucoup apporté, tout simplement parce qu'il était très différent de moi.

Beaucoup d'enseignants ont participé à ma formation. Je me permets de ne mentionner que mon institutrice et mon instituteur d'école primaire, Lucie Fauchère et Roger Pralong.

Tout au long des années d'école et d'études, j'ai côtoyé de nombreux camarades. Nous avons cheminé ensemble. Eric Vuignier est la personne avec laquelle j'ai eu le plus de contact puisque nous avons été compagnons pendant 18 ans. Sa présence à mes côtés m'a réconforté à maintes reprises.

Je garderai un excellent souvenir de la période que j'ai passée à l'IMT, en particulier des premières années, avant que mon travail de thèse m'accapare. J'affectionne particulièrement cette période de ma

vie pour plusieurs raisons. Tout d'abord, elle coïncide avec le début tant attendu de ma vie professionnelle. Ensuite, j'ai trouvé à l'IMT une excellente ambiance de travail. Mon chef, Heinz Hügli m'a laissé beaucoup de liberté dans l'orientation de ma recherche. D'autre part, il régnait entre collègues une entente cordiale, favorable à la collaboration. Je me souviens comment les anciens m'ont fait bénéficier généreusement de leur expérience lors de mon arrivée à l'IMT. Jean-Pierre Amann, François Corthay et Philippe Thévenaz ont notamment fait progresser rapidement mes connaissances pratiques de l'informatique. Daniel Glauser a assuré ma formation pratique dans le domaine de l'électronique.

Sans doute la plus grande chance de ma vie est d'avoir reçu une éducation chrétienne, dans laquelle l'honnêteté et l'égard pour l'autre passent avant le succès personnel. Si au premier abord, une telle règle parait conduire à un sacrifice total de sa vie, en l'expérimentant, on s'aperçoit que la vie nous sourit finalement.

Valais, janvier 1995

Références

- [Aman90] H. P. Amann
3D object recognition based on surface representations
Thèse de doctorat
Institut de Microtechnique, Université de Neuchâtel, Octobre 1990
- [BaBr82] D. H. Ballard, C. M. Brown
Computer vision
Prentice-Hall, Englewood Cliffs, New Jersey, 1982
- [BaLe84] V. Barnett, T. Lewis
Outliers in statistical data
Wiley, Chichester, etc., 1984
- [BeGo89] J. M. Beaulieu, M. Goldberg
"Hierarchy in picture segmentation: A stepwise optimization approach"
IEEE Trans. Pattern Anal. Machine Intell.
vol. PAMI-11, no. 2, pp. 150-163, February 1989
- [BeJa88] P. J. Besl and R. C. Jain
"Segmentation through variable-order surface fitting"
IEEE Trans. Pattern Anal. Machine Intell.
vol. PAMI-10, no. 2, pp. 167-192, March 1988
- [Besl88] P. J. Besl
Surfaces in range image understanding
Springer-Verlag, New York, etc., 1988
- [Besl89] P. J. Besl
"Active optical range imaging sensors"
Advances in machine vision
J. L. C. Sanz, Editor
Springer-Verlag, New York, etc., pp.1-63, 1989

- [Besl90] P. J. Besl
"Geometric signal processing"
Analysis and interpretation of range images
R. C. Jain, A. K. Jain, Editors
Springer-Verlag, New York, etc., pp.141-205, 1990
- [BILe92] M. Blauer, M. D. Levine
"Image smoothing with shape invariance and L1 curvature constraints"
Curves and Surfaces in Computer Vision and Graphics II
M. J. Silberman, H. D. Tagare, Editors
Proc.SPIE 1610, pp.299-310, 1992
- [BLZi87] A. Blake, A. Zisserman
Visual Reconstruction
MIT Press, Cambridge, Massachusetts, 1987
- [Bois84] J.-D. Boissonnat
"Geometric structures for three-dimensional shape representation"
ACM Transactions on Graphics
vol. 3, no. 4, pp. 266-286, October 1984
- [BPYA85] M. Brady, J. Ponce, A. Yuille, H. Asada
"Describing surfaces"
Proc. 2nd Int. Symp. on Robotics Research
H. Hanafusa, H. Inoue, Eds.
MIT Press, Cambridge, Massachusetts, 1985
- [BrSe91] I. N. Bronstein, K. A. Semendjajew
Taschenbuch der Mathematik
G. Grosche, V. Ziegler, D. Ziegler Eds.
Verlag Nauka, Moskau,
Teubner, Stuttgart · Leipzig,
Harri Deutsch, Thun und Frankfurt/Main, 1991
- [Burg92] P.-Y. Burgi
Understanding the early human visual system through modeling
and temporal analysis of neuronal structures
Thèse de doctorat No. 2536
CUI, Université de Genève, 1992
- [Cai89] L.-D. Cai
"A "small leakage" model for diffusion smoothing of image data"
Proc. 6th Scandinavian Conference on Image Analysis
pp. 278-285, Oulu, Finland, June 1989
- [Chan91] F. Chantemargue
Segmentation d'images par approche de type division-fusion:
schémas de parallélisation
Thèse de doctorat
Université Blaise-Pascal de Clermont-Ferrand, France, Décembre
1991

- [ChKa88] C. H. Chen, A. C. Kak
"3D-POLY: A robot vision system for recognizing objects in occluded environments"
Tech. Rep. 88-48
Dept. Elec. Eng., Purdue Univ., December 1988
- [ChKa89] C. H. Chen, A. C. Kak
"A robot vision system for recognizing 3-D objects in low-order polynomial time"
IEEE Trans. Systems Man Cybernet.
vol. 19, no. 6, pp. 1535-1563, December 1989
- [ChYi92] D. M. Chelberg, J. H. Yi
"Range image segmentation using regularization"
Intelligent Robots and Computer Vision X: Algorithms and Techniques
D. P. Casasent, Ed.
Proc.SPIE 1607, pp.??, 1992,
- [Clau93] L. Claud
Tâches de segmentation sous KBVision
Rapport 356HU11/93
Institut de Microtechnique, Université de Neuchâtel, Novembre 1993
- [CoHi68] R. Courant, D. Hilbert
Methoden der Mathematischen Physik I
Springer-Verlag, Berlin, 1968
- [DaSP90] T. Darrell, S. Sclaroff, A. Pentland
"Segmentation by minimal description"
3rd Int. Conf. Computer Vision,
pp. 112-116, Osaka, Japan, December 1990
- [Davi92] A. Davignon
"Contribution of edges and regions to range image segmentation"
Applications of Artificial Intelligence X: Machine Vision and Robotics
K. W. Bowyer, Editor
Proc.SPIE 1708, pp.228-239, 1992
- [FaHe87] O. D. Faugeras, M. Hebert
"The representation, recognition, and positioning of 3-D shapes from range data"
Three-Dimensional Machine Vision,
T. Kanade Ed.
Kluwer Academic Publishers, Norwell, MA, pp. 301-353, 1987
- [FaMN87] T.-J. Fan, G. Medioni, R. Nevatia
"Segmented descriptions of 3-D surfaces"
IEEE J. Robotics Automation
vol. RA-3, no. 6, pp. 527-538, December 1987

- [FeLW90] F. P. Ferrie, J. Lagarde, P. Whaite
"Recovery of volumetric object descriptions from laser rangefinder images"
Proc. 1st European Conf. on Computer Vision
pp. 387-396, Antibbes, France, April 1990
Lecture Notes in Computer Science, vol. 427
Springer Verlag, New York
- [FeMS93] F. P. Ferrie, S. Mathur, G. Soucy
"Feature extraction for 3-D model building and object recognition"
Three-Dimensional Object Recognition Systems
A. K. Jain, P. J. Flynn, Editors
Elsevier, pp.57-88, 1993
- [FlJa89] P. J. Flynn, A. K. Jain
"On reliable curvature estimation"
Proc. IEEE Comput. Soc. Conf. Comput. Vision Patt. Recognition
pp. 110-116, June 1989
- [GuBa92] A. Gupta, R. R. Bajcsy
"Integrated approach for surface and volumetric segmentation of range images using biquadrics and superquadrics"
Applications of Artificial Intelligence X: Machine Vision and Robotics
K. W. Bowyer, Editor
Proc.SPIE 1708, pp.210-227, 1992
- [GuBa93] A. Gupta, R. R. Bajcsy
"Recognition of superquadric models in dense range data"
Three-Dimensional Object Recognition Systems
A. K. Jain, P. J. Flynn, Editors
Elsevier, pp.285-310, 1993
- [Harr87] J.G. Harris
"A new approach to surface reconstruction: the coupled depth/slope model"
Proc. IEEE 1st Int.. Conf. on Comput. Vision
pp. 277-283, London, England, June 8-11, 1987
- [HaSh92] R. M. Haralick, L. G. Shapiro
Computer and robot vision vol.1
Addison-Wesley, 1992
- [HaYo81] L.A. Hageman, D.M. Young
Applied Iterative Methods
Academic Press, Orlando, 1981.
- [HäAM89] A. Härkönen, H. Ailisto, I. Moring
"Noise analysis and filtering of range images produced by a scanning laser range finder"
Proc. 6th Scandinavian Conference on Image Analysis
pp. 481-491, Oulu, Finland, June 1989

- [HoJa87] R. Hoffman, A. K. Jain
"Segmentation and classification of range images"
IEEE Trans. Pattern Anal. Machine Intell.
vol. PAMI-9, no. 5, pp. 608-620, September 1987
- [HoJe88] R. W. Hockney, C. R. Jesshope
Parallel computers 2: architecture, programming and algorithms,
2nd ed.
Adam Hilger, Bristol and Philadelphia, 1988
- [Hube81] P. J. Huber
Robust statistics
John Wiley & Sons, New York, etc., 1981
- [HüMa89] H. Hügli, G. Maître
"3D by structured light: implementation and evaluation of a vision
system for small parts"
Optical 3-D measurement techniques
A. Gruen and H. Kahmen, Editors
pp. 153-163
Wichmann, Karlsruhe, Germany, 1989
- [Jarv83] R. A. Jarvis
"A perspective on range finding techniques for computer vision"
IEEE Trans. Pattern Anal. Machine Intell.
vol. PAMI-5, no. 2, pp. 122-139, March 1983
- [JiBu94] X. Y. Jiang, H. Bunke
"Fast segmentation of range images into planar regions by scan line
grouping"
à paraître
- [Kasv88] T. Kasvand
"The k1k2 space in range image analysis"
Proc.9th Int. Conf. on Pattern Recognition
14-17 November 1988, Rome, Italy
pp.923-926
- [Kunt93] M. Kunt (directeur de publication)
Traitement numérique des images
Collection électricité, Traitement de l'information: volume 2
Presses polytechniques et universitaires romandes, Lausanne, 1993
- [LeGB90] A. Leonardis, A. Gupta, R. R. Bajcsy
"Segmentation as the search for the best description of the image in
terms of primitives"
3rd Int. Conf. Computer Vision, pp. 121-125, Osaka, Japan,
December 1990
- [Leon87] R. Leonardi
Segmentation adaptive pour le codage d'images
Thèse de doctorat, no. 691
Département d'Electricité, Ecole Polytechnique Fédérale de
Lausanne, 1987

- [Leon93] A. Leonardis
Image analysis using parametric models: model-recovery and model-selection paradigm
Ph. D. thesis
Department of Computer and Information Science, University of Ljubljana, 1993
- [MaHü88] G. Maître, H. Hügli
Réalisation d'un dispositif de mesure 3D utilisant de la lumière structurée
Rapport 250EC12/88
Institut de Microtechnique, Université de Neuchâtel , Décembre 1988
- [MaHü89] G. Maître, H. Hügli
Extraction de surfaces par croissance de régions sur un modèle d'objet 2.5D
Rapport 259EC05/89
Institut de Microtechnique, Université de Neuchâtel , Mai 1989
- [MaHü92] G. Maître, H. Hügli
"Range image segmentation by controlled-continuity spline approximation for parallel computation"
Curves and Surfaces in Computer Vision and Graphics II
M. J. Silberman, H. D. Tagare, Editors
Proc.SPIE 1610, pp.238-249, 1992
- [MHTA90] G. Maître, H. Hügli, F. Tièche, J. P. Amann
"Range image segmentation based on function approximation"
Close-Range Photogrammetry Meets Machine Vision
A. Gruen, E. P. Baltsavias, Editors
Proc.SPIE 1395, pp.275-282, 1990
- [Over92] I. Overington
Computer Vision, a unified, biologically-inspired approach
Elsevier, Amsterdam, etc., 1992
- [PaMe86] B. Parvin, G. Medioni
"Segmentation of range images into planar surfaces by split and merge"
Proc. Comput. Vision Patt. Recognition
pp. 415-417, 1986
- [PeMa90] P. Perona, J. Malik
"Scale-space and edge detection using anisotropic diffusion"
IEEE Trans. Pattern Anal. Machine Intell.
vol. PAMI-12, no. 7, pp. 629-639, July 1990
- [PFTV86] W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling
Numerical recipes, the art of scientific computing
Cambridge University Press, Cambridge, etc. , 1986

- [PoAl82] J. L. Posdamer, M. D. Altschuler
"Surface measurement by space-encoded projected beam systems"
Computer Graphics and Image Processing
vol. 18, pp. 1-17, 1982
- [PoBr87] J. Ponce, M. Brady
"Toward a surface primal sketch"
Three-Dimensional Machine Vision,
T. Kanade Ed.
Kluwer Academic Publishers, Norwell, MA, pp. 195-240, 1987
- [RBTD87] M. Rioux, G. Bechthold, D. Taylor, M. Duggan
"Design of a large depth of view three-dimensional camera for robot vision"
Applied Optics
vol. 26, no. 12, pp. 1245-1250, Dec. 1987
- [RiCo88] M. Rioux, L. Cournoyer
The NRCC three-dimensional image data files
Tech. rep. CNRC-29077, Division Elec. Eng.,
National Research Council Canada, Ottawa, Canada, June 1988
- [SMCM91] P. Saint-Marc, J.-S. Chen, G. Medioni
"Adaptive smoothing: a general tool for early vision"
IEEE Trans. Pattern Anal. Machine Intell.
vol. PAMI-13, no. 6, pp. 514-529, June 1991
- [Szel90] R. Szeliski
"Fast surface interpolation using hierarchical basis functions"
IEEE Trans. Pattern Anal. Machine Intell.
vol. PAMI-12, no.6, pp.513-528, June 1990
- [TaSR89] R. W. Taylor, M. Savini, A. P. Reeves
"Fast segmentation of range imagery into planar regions"
Computer, Vision Graphics and Image Processing
vol. 45, pp. 42-60, 1989
- [Taub91] G. Taubin
"Estimation of planar curves, surfaces, and nonplanar space curves defined by implicit equations with applications to edge and range image segmentation"
IEEE Trans. Patt. Anal. Mach. Intell.
vol. PAMI-13, no.11, pp.1115-1138, November 1991
- [Terz86] D. Terzopoulos
"Regularization of inverse visual problems involving discontinuities"
IEEE Trans. Patt. Anal. Mach. Intell.
vol. PAMI-8, no.4, pp.413-424, July 1986
- [Terz88] D. Terzopoulos
"The computation of visible-surface representations"
IEEE Trans. Patt. Anal. Mach. Intell.
vol. PAMI-10, no.4, pp.417-438, July 1988

- [Tièc90] F. Tièche
Segmentation et modélisation d'objets 3D selon la courbure de la surface
Travail de licence
Institut de Microtechnique, Université de Neuchâtel, Mai 1990
- [TrFi92] E. Trucco, R. Fisher
"Computing surface-based representations from range images"
Proc. IEEE Int. Symposium on Intelligent Control ISIC-92
pp. 275-280, Glasgow, UK, 1992
- [YJAd90] A. Ylä-Jääski, F. Ade
Segmentation and symbolic description of range images
Tech. Rep. BIWI-TR-113
Inst. for Comm. Techn., Swiss Federal Institute for Technology,
Zürich, April 1990
- [YoLe89] N. Yokoya, M. D. Levine
"Range image segmentation based on differential geometry: a hybrid approach"
IEEE Trans. Patt. Anal. Mach. Intell.
vol. PAMI-11, no.6, pp.643-649, June 1989

Annexe A

Convergence de la troncature itérative

Nous démontrons les propriétés de convergence de la procédure de troncature itérative qui sont affirmées à la table 5.4 et qui concernent les trois distributions suivantes: uniforme, gaussienne et mixte. Dans cette démarche, nous utilisons, sans y faire explicitement référence, les définitions et notations du chapitre 5 .

Nous commençons par démontrer des propriétés vérifiées par toute distribution symétrique et donc par les trois distributions mentionnées ci-dessus.

A.1 Distribution symétrique

Assertion

Soit la distribution $h(\xi)$ vérifiant la propriété de symétrie

$$h(-\xi) = -h(\xi) \tag{A.1}$$

Les suites $\mu^{(n)}$ et $\sigma^{(n)}$ vérifient alors les propriétés suivantes:

$$\begin{aligned} \mu^{(n)} &\equiv 0 \\ \sigma^{(n)} &= \sigma(0, k\sigma^{(n-1)}) \end{aligned} \tag{A.2.a-b}$$

Démonstration

Comme $h(\xi)$ est symétrique, sa valeur moyenne est nulle

$$\mu^{(0)} = 0 \tag{A.3}$$

et la valeur moyenne résultant d'une troncature symétrique par rap-

port à 0 est aussi nulle

$$\mu(0, b) = 0, \quad \forall b \in \mathfrak{R}^+ \quad (\text{A.4})$$

De cette dernière propriété, résulte la propriété de récurrence suivante pour $\mu^{(n)}$

$$\left(\mu^{(n-1)} = 0\right) \Rightarrow \left(\mu^{(n)} = 0\right) \quad (\text{A.5})$$

Finalement, de (A.3) et (A.5) découle (A.2.a) et de (A.2.a) découle directement (A.2.b).

qed

A.2 Distribution uniforme

Assertion

Soit la distribution uniforme définie par

$$h(\xi) = \begin{cases} \frac{1}{a} & , |\xi| < \frac{a}{2} \\ 0 & , |\xi| > \frac{a}{2} \end{cases} \quad (\text{A.6.a-b})$$

la valeur limite de la suite $\sigma^{(n)}$ vaut

$$\lim_{n \rightarrow \infty} \sigma^{(n)} = \begin{cases} 0 & , k < \sqrt{3} \\ \sigma^{(0)} & , k \geq \sqrt{3} \end{cases} \quad (\text{A.7.a-b})$$

Démonstration

Le calcul de $\sigma(0, b)$, où b est un nombre réel positif, conduit aux relations suivantes

$$\sigma(0, b) = \begin{cases} \frac{1}{\sqrt{3}} b & , b < \frac{a}{2} \\ \frac{1}{\sqrt{3}} \frac{a}{2} & , b \geq \frac{a}{2} \end{cases} \quad (\text{A.8.a-b})$$

On en déduit, pour la suite $\sigma^{(n)}$, l'expression de la valeur initiale

$$\sigma^{(0)} = \frac{1}{\sqrt{3}} \frac{a}{2} \quad (\text{A.9})$$

et, grâce à la propriété (A.2.b), la relation de récurrence

$$\sigma^{(n)} = \begin{cases} \frac{1}{\sqrt{3}} k \sigma^{(n-1)} & , k \sigma^{(n-1)} < \frac{a}{2} \\ \frac{1}{\sqrt{3}} \frac{a}{2} & , k \sigma^{(n-1)} \geq \frac{a}{2} \end{cases} \quad (\text{A.10.a-b})$$

De (A.9) et (A.10), on tire la formule explicite pour $\sigma^{(n)}$

$$\sigma^{(n)} = \begin{cases} \left(\frac{k}{\sqrt{3}}\right)^n \sigma^{(0)} & , k < \sqrt{3} \\ \sigma^{(0)} & , k \geq \sqrt{3} \end{cases} \quad (\text{A.11.a-b})$$

d'où découle directement la propriété (A.7).

qed

A.3 Distribution gaussienne

Assertion

Soit la distribution gaussienne définie par

$$h(\xi) = \frac{1}{\sqrt{2\pi}} e^{-\xi^2/2\sigma_0^2} \quad (\text{A.12})$$

la valeur limite de la suite $\sigma^{(n)}$ vérifie les relation suivantes

$$\left\{ \begin{array}{l} \lim_{n \rightarrow \infty} \sigma^{(n)} = 0 \quad , k \leq \sqrt{3} \\ 0 < \lim_{n \rightarrow \infty} \sigma^{(n)} \leq \sigma^{(0)} \quad , k > \sqrt{3} \end{array} \right. \quad (\text{A.13.a-b})$$

Démonstration

Le calcul de $\sigma(0, b)$, où b est un nombre réel positif, conduit à la relation

$$\sigma(0, b) = \sigma_0 \sqrt{1 - \Psi(b/(\sqrt{2}\sigma_0))} \quad (\text{A.14})$$

avec

$$\Psi(\xi) = \xi \cdot e^{-\xi^2} \left/ \int_0^\xi e^{-\vartheta^2} d\vartheta \right. \quad (\text{A.15})$$

On en déduit, pour la suite $\sigma^{(n)}$, l'expression de la valeur initiale

$$\sigma^{(0)} = \sigma_0 \quad (\text{A.16})$$

et, grâce à la propriété (A.2.b), la relation de récurrence

$$\sigma^{(n)} = \sigma_0 \sqrt{1 - \Psi(k\sigma^{(n-1)}/(\sqrt{2}\sigma_0))} \quad (\text{A.17})$$

En utilisant le développement en série de son numérateur et de son dénominateur (A.15), on obtient l'approximation suivante de $\Psi(\xi)$ pour $\xi \ll 1$

$$\Psi(\xi) = \frac{1 - \xi^2 + O(\xi^4)}{1 - \xi^2/3 + O(\xi^4)} \quad (\text{A.18})$$

d'où

$$1 - \Psi(\xi) = 2\xi^2/3 + O(\xi^4) \quad (\text{A.19})$$

Si l'on suppose que la limite de la suite $\sigma^{(n)}$ est nulle, on peut utiliser

l'approximation (A.19) pour calculer, à partir de (A.17), la limite du rapport $\sigma^{(n)}/\sigma^{(n-1)}$ lorsque n tend vers l'infini. Le résultat est

$$\lim_{n \rightarrow \infty} \frac{\sigma^{(n)}}{\sigma^{(n-1)}} = \frac{k}{\sqrt{3}} \quad (\text{A.20})$$

Dans ces conditions, la suite $\sigma^{(n)}$ est décroissante si et seulement si $k \leq \sqrt{3}$, ce qui prouve (A.13.a) et l'inégalité de gauche de (A.13.b). D'autre part, le fait que $\Psi(\xi)$ soit positif pour ξ positif, mis en relation avec (A.17), prouve l'inégalité de droite de (A.13.b).

qed

A.4 Distribution mixte

Assertion

Soit la distribution mixte définie par

$$h(\xi) = \begin{cases} (1-\varepsilon)\delta(\xi) & , \xi = 0 \\ \frac{\varepsilon}{a} & , 0 < |\xi| < \frac{a}{2} \\ 0 & , |\xi| > \frac{a}{2} \end{cases} \quad (\text{A.21.a-c})$$

la valeur limite de la suite $\sigma^{(n)}$ vaut

$$\lim_{n \rightarrow \infty} \sigma^{(n)} = \begin{cases} 0 & , k < \sqrt{\frac{3}{\varepsilon}} \\ \sigma^{(0)} & , k \geq \sqrt{\frac{3}{\varepsilon}} \end{cases} \quad (\text{A.22.a-b})$$

Démonstration

Le calcul de $\sigma(0, b)$, où b est un nombre réel positif, conduit aux relations suivantes

$$\sigma(0, b) = \begin{cases} \Phi\left(\frac{b}{a/2}\right)\sqrt{\frac{\varepsilon}{3}} b & , b < \frac{a}{2} \\ \sqrt{\frac{\varepsilon}{3}} \frac{a}{2} & , b \geq \frac{a}{2} \end{cases} \quad (\text{A.23.a-b})$$

avec

$$\Phi(\xi) = \sqrt{\frac{1}{\varepsilon + \frac{1}{\xi}(1-\varepsilon)}} \quad (\text{A.24})$$

On en déduit, pour la suite $\sigma^{(n)}$, l'expression de la valeur initiale

$$\sigma^{(0)} = \sqrt{\frac{\varepsilon}{3}} \frac{a}{2} \quad (\text{A.25})$$

et, grâce à la propriété (A.2.b), la relation de récurrence

$$\sigma^{(n)} = \begin{cases} \Phi\left(\frac{k\sigma^{(n-1)}}{a/2}\right)\sqrt{\frac{\varepsilon}{3}} k\sigma^{(n-1)} & , k\sigma^{(n-1)} < \frac{a}{2} \\ \sqrt{\frac{\varepsilon}{3}} \frac{a}{2} & , k\sigma^{(n-1)} \geq \frac{a}{2} \end{cases} \quad (\text{A.26.a-b})$$

De (A.25) et de (A.26), où on remarque que $\Phi < 1$, on déduit les relations suivantes pour $\sigma^{(n)}$

$$\begin{cases} \sigma^{(n)} < \left(\sqrt{\frac{\varepsilon}{3}}\right)^n \sigma^{(0)} & , k < \sqrt{\frac{3}{\varepsilon}} \\ \sigma^{(n)} = \sigma^{(0)} & , k \geq \sqrt{\frac{3}{\varepsilon}} \end{cases} \quad (\text{A.27.a-b})$$

De ces propositions découle directement la propriété (A.22).

qed

Annexe B

Stabilité de la méthode itérative de diffusion D^*

Nous appliquons la méthode d'analyse de von Neumann [PFTV86: p.625] pour établir les conditions de stabilité de la méthode itérative D^* [§6.3.4]. Nous montrons que la stabilité impose des conditions sur les valeurs de b et α .

Analyse de von Neumann

La manière de procéder est la suivante:

(1) on part de la formule de récurrence (6.46)

$$f_{i,j}^{(n)} = f_{i,j}^{(n-1)} + \alpha \left(b \left(f_{i-1,j}^{(n-1)} + f_{i+1,j}^{(n-1)} + f_{i,j-1}^{(n-1)} + f_{i,j+1}^{(n-1)} - 4 f_{i,j}^{(n-1)} \right) + \frac{1-b}{2} \left(f_{i-1,j-1}^{(n-1)} + f_{i+1,j+1}^{(n-1)} + f_{i-1,j+1}^{(n-1)} + f_{i+1,j-1}^{(n-1)} - 4 f_{i,j}^{(n-1)} \right) \right) \quad (\text{B.1})$$

(2) on procède à la substitution définie par

$$f_{k,l}^{(m)} = \xi^m e^{i\omega_u k} e^{i\omega_v l} \quad (\text{B.2})$$

où \mathbf{i} désigne le nombre complexe, et on simplifie par $\xi^{n-1} e^{i\omega_u i} e^{i\omega_v j}$ les deux membres de l'égalité ainsi obtenue. En utilisant les relations entre fonction exponentielle d'une variable complexe et fonctions trigonométriques, on aboutit à l'égalité

$$\xi = 1 + \alpha \left(b \left(2 \cos(\omega_u) + 2 \cos(\omega_v) - 4 \right) + \frac{1-b}{2} \left(4 \cos(\omega_u) \cos(\omega_v) - 4 \right) \right) \quad (\text{B.3})$$

(3) on impose la condition

$$|\xi| \leq 1, \forall (\omega_u, \omega_v) \in \mathfrak{R}^2 \quad (\text{B.4})$$

qui définit le critère de stabilité de von Neumann.

Conditions sur b et α

Le critère de stabilité (B.4) peut être réécrit sous la forme

$$\frac{1}{\alpha} < h(\cos(\omega_u), \cos(\omega_v)) < 0 \quad , \forall (\omega_u, \omega_v) \in \mathfrak{R}^2 \quad (\text{B.5})$$

avec

$$h(\gamma, \eta) = b(\gamma + \eta - 2) + (1 - b)(\gamma \cdot \eta - 1) \quad (\text{B.6})$$

La recherche des valeurs extrêmes de $h(\gamma, \eta)$ dans le domaine

$$\begin{aligned} -1 &\leq \gamma \leq 1 \\ -1 &\leq \eta \leq 1 \end{aligned} \quad (\text{B.7.a-b})$$

permet d'éliminer ω_u et ω_v du critère de stabilité et d'obtenir finalement les conditions que b et α doivent remplir pour garantir la stabilité de la méthode itérative, soit

$$(b \geq 0) \wedge (\alpha \geq 0) \wedge (b \leq \frac{1}{2} \Rightarrow \alpha \leq \frac{1}{2}) \wedge (b \geq \frac{1}{2} \Rightarrow \alpha \leq \frac{1}{4b}) \quad (\text{B.8})$$

Annexe C

Atténuation d'un signal aléatoire discret IID par lissage gaussien

Nous nous intéressons à l'effet du lissage gaussien (convolution avec une gaussienne) sur un signal aléatoire discret identiquement indépendamment distribué (IID). Nous montrons que l'écart-type du signal lissé et celui de ses dérivées (d'ordre 1 et 2) peuvent, chacun, être exprimés comme le produit de l'écart-type du signal original avec un coefficient d'atténuation qui dépend de l'écart-type de la gaussienne.

Nous développons le sujet suivant la structure: problème, assertion et démonstration.

Problème

Soit \mathbf{F} un signal aléatoire discret IID d'écart-type σ_0 , auquel on applique consécutivement les opérations suivantes:

- (1) transformation vers l'espace continu à l'aide de la fonction de Dirac, le résultat étant la fonction

$$f(u, v) = \sum_{i, j} f_{i, j} \delta(u - i, v - j) \quad (\text{C.1})$$

(Les coordonnées u et v sont telles que les coordonnées de position i et j correspondent aux valeurs entières de u et v , respectivement.)

- (2) lissage par convolution avec une gaussienne puis, le cas échéant, dérivation d'ordre ν suivant la direction φ , pour aboutir à la

ν	0	1	2
$D_\nu(\lambda)$	$\frac{1}{2\sqrt{\pi}} \lambda^{-1}$	$\frac{1}{2\sqrt{2\pi}} \lambda^{-2}$	$\frac{\sqrt{3}}{4\sqrt{\pi}} \lambda^{-3}$

Table C.1: Rapport des écarts-type en fonction de l'ordre de la dérivée.

fonction

$$\tilde{f}_\varphi^{(\nu)}(u, \nu) = (g^* f)_\varphi^{(\nu)}(u, \nu) \quad (\text{C.2})$$

où g représente la gaussienne d'écart-type λ

$$g(u, \nu) = \frac{1}{2\pi\lambda^2} e^{-\frac{u^2 + \nu^2}{2\lambda^2}} \quad (\text{C.3})$$

(A noter que les indices " $^{(0)}$ " désignent le signal non dérivé.)

(3) échantillonnage suivant les valeurs entière de u et de ν , aboutissant au signal discret

$$\tilde{f}_{i,j}^{\nu,\varphi} = \tilde{f}_\varphi^{(\nu)}(i, j) \quad (\text{C.4})$$

On cherche à exprimer l'écart-type $\tilde{\sigma}_{\nu,\varphi}$ de $\tilde{\mathbf{F}}^{\nu,\varphi}$ en fonction de l'écart-type σ_0 de \mathbf{F} et de l'écart-type λ de la gaussienne g .

Assertion

L'écart-type $\tilde{\sigma}_{\nu,\varphi}$

- (1) est indépendant de la direction de dérivation φ , (d'où la notation $\tilde{\sigma}_\nu$ dorénavant)
- (2) est proportionnel à σ_0 , le facteur de proportionnalité étant dépendant de λ

$$\tilde{\sigma}_\nu = D_\nu(\lambda) \cdot \sigma_0 \quad (\text{C.5})$$

La valeur du facteur de proportionnalité est donnée dans la table C.1 pour les dérivées d'ordre 0 à 2.

Remarque

$D_\nu(\lambda)$ exprime une atténuation ($D_\nu(\lambda) < 1$) uniquement pour des valeurs de λ suffisamment distantes de 0.

Démonstration

Première partie: mise en évidence du coefficient d'atténuation

En utilisant les propriétés de permutation et d'associativité des opérations de dérivation et de convolution, on obtient, dans l'espace continu, l'expression suivante pour le signal lissé et dérivé $\tilde{f}_\varphi^{(v)}$ en fonction du signal original f

$$\tilde{f}_\varphi^{(v)}(u, v) = (g_\varphi^{(v)} * f)(u, v) \quad (\text{C.6})$$

d'où, pour les signaux discrets, l'expression suivante de $\tilde{\mathbf{F}}^{v, \varphi}$ en fonction de \mathbf{F}

$$\tilde{f}_{i,j}^{v, \varphi} = \sum_{k,l} g_\varphi^{(v)}(i-k, j-l) \cdot f_{k,l} \quad (\text{C.7})$$

A chaque position (i, j) d'un signal aléatoire discret est toujours associée conceptuellement une variable aléatoire, Comme le signal \mathbf{F} est IID, on peut écrire, indépendamment de la position, la relation suivante pour la variance de la variable $\tilde{\chi}^{v, \varphi}$ en fonction de la variance de la variable χ

$$V[\tilde{\chi}^{v, \varphi}] = V[\chi] \sum_{(k,l)} (g_\varphi^{(v)}(k, l))^2 \quad (\text{C.8})$$

$\tilde{\sigma}_{v, \varphi}$ est donc proportionnel à σ_0 , le facteur de proportionnalité étant

$$D_{v, \varphi}(\lambda) = \sqrt{\sum_{(k,l)} (g_\varphi^{(v)}(k, l))^2} \quad (\text{C.9})$$

Deuxième partie: calcul du coefficient d'atténuation

En utilisant l'approximation de la somme par l'intégrale, $D_{v, \varphi}^2$ devient

$$D_{v, \varphi}^2(\lambda) = \int_{\mathfrak{R}^2} (g_\varphi^{(v)}(u, v))^2 dudv \quad (\text{C.10})$$

Ensuite, en utilisant les coordonnées polaires r et θ , $D_{v, \varphi}^2$ s'exprime sous la forme

$$D_{v, \varphi}^2(\lambda) = \int_{r=0}^{+\infty} \int_{\theta=0}^{2\pi} (g_\varphi^{(v)}(r \cos(\theta), r \sin(\theta)))^2 d\theta r dr \quad (\text{C.11})$$

Etant donné la symétrie de rotation de g , la dérivée dans la direction φ au point de coordonnées polaires (r, θ) est égale à la dérivée dans

la direction $\varphi + \theta$ au point de coordonnées polaires $(r, 0)$, ceci quels que soient r et θ . Formellement

$$g_{\varphi}^{(v)}(r \cos(\theta), r \sin(\theta)) = g_{\varphi+\theta}^{(v)}(r, 0) \quad , \forall r, \theta \quad (\text{C.12})$$

(A noter que cette relation est aussi vraie pour le cas trivial de la dérivée d'ordre 0.)

Par conséquent, au lieu d'intégrer $(w_{\varphi}^{(v)})^2$ sur un cercle de rayon r par rapport à la coordonnée θ , on peut intégrer cette fonction en un point fixe de distance r (par exemple celui de coordonnée $\theta = 0$) par rapport à la direction de la dérivée. D'où la nouvelle formule pour $D_{v,\varphi}^2$

$$D_{v,\varphi}^2(\lambda) = \int_{r=0}^{+\infty} \int_{\varphi=0}^{2\pi} (w_{\varphi}^{(v)}(r, 0))^2 d\varphi r dr \quad (\text{C.13})$$

On en déduit immédiatement l'indépendance de D_v et donc de $\tilde{\sigma}_v$ par rapport à φ (point (1) de l'assertion)

Le calcul des dérivées directionnelles d'ordre 1 et 2 à partir des dérivées partielles du même ordre fournit les résultats suivants

$$w'_{\varphi}(r, 0) = -\lambda^{-2} r w(r, 0) \cos(\varphi) \quad (\text{C.14})$$

et

$$w''_{\varphi}(r, 0) = \lambda^{-2} w(r, 0) \left((\lambda^{-2} r^2 - 1) \cos^2(\varphi) - \sin^2(\varphi) \right) \quad (\text{C.15})$$

En calculant l'intégrale (C.13) pour chacun des trois cas (dérivées d'ordre 0 à 2), on aboutit aux résultats de la table C.1.

qed