UNIVERSITE DE MONTREAL

REPRESENTATION MULTIRESOLUTION D'IMAGES TRIDIMENSIONNELLES

par

Hong Minh CUNG
DEPARTEMENT DE GÉNIE ELECTRIQUE
ECOLE POLYTECHNIQUE

MEMOIRE PRESENTE EN VUE DE L'OBTENTION
DU GRADE DE MAITRE ES SCIENCES APPLIQUEES (M.Sc.A.)
(GÉNIE ELECTRIQUE)
AVRIL 1990

© Hong Minh CUNG 1990

ıal Library
nada

dian Theses Service

. Canada
N4

Bibliothèque nationale
du Canada

Service des thèses canadiennes

: author has granted an irrevocable non-
lusive licence allowing the National Library
Canada to reproduce, loan, distribute or sell
ies of his/her thesis by any means and in
form or format, making this thesis available
nterested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

: author retains ownership of the copyright
his/her thesis. Neither the thesis nor
stantial extracts from it may be printed or
erwise reproduced without his/her per-
sion.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-58113-1

nadä

UNIVERSITE DE MONTREAL

ECOLE POLYTECHNIQUE

Ce mémoire intitulé:

REPRESENTATION MULTIRESOLUTION D'IMAGES TRIDIMENSIONNELLES

présenté par: Hong Minh CUNG

en vue de l'obtention du grade de: M.Sc.A.

a été dûment accepté par le jury d'examen constitué de :

M. Richard HURTEAU D.Ing., Président

M. Paul COHEN Ph.D., Directeur

M. Pierre BOULANGER M.Sc.A., Membre

A mes chers parents pour leur soutien indéfectible.

SOMMAIRE

Nous présentons une approche multirésolution pour la segmentation des images 3-D. Cette méthode consiste en une combinaison de procédures de détection d'arêtes et de segmentation en primitives de surface et conduit à l'obtention d'une représentation en termes d'arêtes focalisées et de régions à signes de courbure constants. A un niveau de résolution donné, la procédure de détection d'arêtes extrait les discontinuités de profondeur et d'orientation des surfaces, et partitionne l'image en des régions intermédiaires à l'intérieur desquelles la profondeur et l'orientation de la surface demeurent continues. L'extraction est accomplie en détectant, en premier lieu, la présence d'arêtes significatives au niveau de résolution choisi, et en déterminant par la suite leur position exacte en effectuant une poursuite d'arêtes à travers l'espace d'échelles jusqu'au niveau de résolution le plus fin auquel les arêtes détectées sont le mieux localisées. A ce même niveau de résolution, la procédure de segmentation en primitives de surface consiste à appliquer un filtrage anisotropique sur les régions intermédiaires, délimitées par les arêtes extraites précédemment, et à segmenter par la suite ces régions en des primitives de surface à l'intérieur desquelles les signes des courbures Gaussienne et moyenne sont constants. Des résultats expérimentaux, obtenus à partir d'images synthétiques et réelles, sont présentés et une comparaison avec quelques techniques récemment publiées est effectuée.

ABSTRACT

We present a multiscale edge- and region-based segmentation scheme for range images, which leads to a rich representation in terms of focused edges and constant-curvature signs regions. At a given scale, the edge detection scheme extracts the discontinuities in surface depth and orientation, and partition the image into intermediate regions within which surface depth and orientation remain continuous. The extraction is accomplished by first detecting the presence of significant edges at that scale, and then determining their precise location by tracking them over decreasing scale, down to the finest scale at which they are most accurately positioned. At the same scale, the region-based segmentation scheme consists of applying anisotropic filtering to the intermediate regions, delimited by the previously detected edges, and segmenting these regions into surface primitives of constant Gaussian and mean curvature signs. Experimental results are presented for both synthetic and real images, and a comparison is done with some recently published techniques.

REMERCIEMENTS

Je tiens à exprimer ma gratitude et mes remerciements à toutes personnes qui ont contribué à la réalisation de cette thèse.

En premier lieu, j'exprime ma reconnaissance à mon directeur de thèse, le Dr. Paul Cohen, pour son esprit critique et ses conseils judicieux dispensés tout au long de ce travail. Je remercie également M. Pierre Boulanger pour son support technique et ses suggestions fertiles.

Je tiens à remercier particulièrement Mme Nathalie Plaziac pour son aide précieuse au moment de la rédaction de ce mémoire.

Je tiens à remercier M. Jacques Domey, M. Marc Rioux et M. Luc Cournoyer, du Conseil National de Recherche du Canada, pour les images tridimensionnelles.

Mes respects à tous les membres du Laboratoire de Vision artificielle du Département de Génie Electrique de l'école Polytechnique de Montréal. Merci enfin à mes collègues: Agnès, Ning, Sylvain, et François.

TABLE DES MATIERES

SOMMAIRE v

ABSTRACT vi

REMERCIEMENTS vii

LISTE DES FIGURES xiv

LISTE DES TABLEAUX xv

INTRODUCTION 1

CHAPITRE 1 Approche multirésolution pour l'extraction d'arêtes dans les images 3-D 7

1.1 Extraction d'arêtes: principes et revue bibliographique. 7

1.1.1 Principes. 7

1.1.2 Revue bibliographique. 10

1.1.2.1 Extraction d'arêtes dans les images 2-D. 10

1.1.2.2 Extraction d'arêtes dans les images 3-D. 12

1.1.3 Principe de l'approche proposée. 16

1.2 Détection d'arêtes potentielles. 19

1.3 Segmentation d'arêtes potentielles détectées. 21

1.4 Localisation de segments d'arête par focalisation. 24

1.5 Classification des segments d'arête focalisés. 27

1.5.1 Classification au niveau des pixels. 28

1.5.2 Classification au niveau des segments. 29

1.6 Résultats expérimentaux. **33**

1.6.1 Cas des images 3-D synthétiques. 34

1.6.2 Cas des images 3-D réelles. 35

CHAPITRE 2 Approche multirésolution pour la segmentation d'images 3-D en primitives de surface **52**

2.1 Segmentation d'images 3-D: principes et revue bibliographique. **52**

2.1.1 Principes. 52

2.1.2 Revue bibliographique. 57

2.1.3 Principe de l'approche proposée. 60

2.2 Filtrage des régions intermédiaires. **61**

2.3 Estimation des courbures K et H dans l'image filtrée. **63**

2.4 Calcul de la carte d'étiquettes initiale. **66**

2.5 Calcul de la carte d'étiquettes finale. **70**

2.6 Résultats expérimentaux. **70**

CHAPITRE 3 Proposition d'une représentation multirésolution. **80**

3.1 Principe fondamental. **81**

3.1.1 Représentation dans l'espace d'échelles: approche de Witkin. . . . 81

3.1.2 Proposition d'une représentation multirésolution. 82

3.2 Résultats expérimentaux. **86**

CHAPITRE 4 Conclusions et développements futurs **93**

4.1 Conclusions. **93**

4.1.1 Extraction d'arêtes. 93

4.1.2 Segmentation d'images 3-D. 94

4.1.3 Représentation multirésolution. 95

4.2 Développements futurs. **96**

BIBLIOGRAPHIE **98**

LISTE DES FIGURES

CHAPITRE 1 7

1.1 Arêtes d'ordre 0 détectées à $\sigma_g = 2.2$. (a) Image originale codée en niveaux de gris. (b) Image originale ombragée. (c) Arêtes détectées. (d) Arêtes qui subsistent après l'opération de seuillage. 22

1.2 Arêtes d'ordre 1 détectées à $\sigma_g = 2.2$. (a) Points d'arête en provenance de $\mathcal{R}_1(\sigma_g)$. (b) Points d'arête en provenance de $\mathcal{R}_2(\sigma_g)$. (c)-(d) Arêtes d'ordre 1 qui subsistent à l'opération de seuillage. . 23

1.3 Segmentation d'arêtes potentielles. (a) Image d'arêtes à segmenter. (b) Image de segments qui résultent de la segmentation. 37

1.4 Principe de la méthode de focalisation proposée. 37

1.5 Resultats de la focalisation d'arêtes ($\sigma_g = 2.2$ and $\sigma_f = 0.45$). (a) Segments d'arête initiaux. (b) Resultats obtenus avec notre approche. (c) Resultats obtenus avec l'approche de Bergholm. (d) Segment type, détecté à σ_g. (e) Resultats obtenus avec notre approche. (f) Resultats obtenus avec l'approche de Bergholm. . . . 38

1.6 Comportements caractéristiques des arêtes. (a) Image d'arêtes à $\sigma_g = 2.2$. (b) Image de segments focalisés à $\sigma_f = 0.45$. (c) Comportement caractéristique d'une vraie arête. (d) Comportement caractéristique d'une fausse arête. 39

1.7 Exemple de résultats produits par la méthode de classification d'arêtes proposée. (a) Segments détectés au niveau de résolution initial $\sigma_g = 1.7$. (b) Segments focalisés, obtenus au niveau de résolution final $\sigma_f = 0.45$. (c) Segments retenus après classification par la méthode proposée. 40

1.8 Comparaison de résultats produits par notre méthode et par celle de Bergholm et Sjöberg. (a) Résultats obtenus par la méthode de focalisation de Bergholm. (b) Arêtes retenues après la classification de Sjöberg. (c) Résultats obtenus par la méthode de focalisation proposée. (d) Arêtes retenues après notre classification. 41

1.9 Image synthétique d'un objet polyédrique. (a) Image originale. (b)-(d) Images bruitées avec rapports signal-à-bruit respectifs de 26.0 dB, 14.5 dB, et 4.5 dB. 42

1.10 Resultats obtenus à partir des images synthétiques montrées à la Figure 1.9. (a) Image originale. (b) SNR = 26.0 dB. (c) SNR = 14.5 dB. (d) SNR = 4.5 dB. 43

1.11 Images 3-D originales. (a) BLOC11. (b) BLOC9. (c) BLOC4. . . 44

1.12 Images 3-D originales. (a) BLOC15. (b) BLOC16. (c) BOLT6. . . 45

1.13 Arêtes de l'image "BLOC11" . (a) Arêtes d'ordre 0. (b) Arêtes d'ordre 1. (c) Superposition des deux types d'arêtes. (d) Les mêmes résultats que dans (c) avec arêtes dues aux effets d'ombre éliminés. 46

1.14 Arêtes de l'image "BLOC9". (a) Arêtes d'ordre 0. (b) Arêtes d'ordre 1. (c) Superposition des deux types d'arêtes. (d) Les mêmes résultats que dans (c) avec arêtes dues aux effets d'ombre éliminés. 47

1.15 Arêtes de l'image "BLOC4". (a) Arêtes d'ordre 0. (b) Arêtes d'ordre 1. (c) Superposition des deux types d'arêtes. (d) Les mêmes résultats que dans (c) avec arêtes dues aux effets d'ombre éliminés. 48

1.16 Arêtes de l'image "BLOC15". (a) Arêtes d'ordre 0. (b) Arêtes d'ordre 1. (c) Superposition des deux types d'arêtes. (d) Les mêmes résultats que dans (c) avec arêtes dues aux effets d'ombre éliminés. 49

1.17 Arêtes de l'image "BLOC16". (a) Arêtes d'ordre 0. (b) Arêtes d'ordre 1. (c) Superposition des deux types d'arêtes. (d) Les mêmes résultats que dans (c) avec arêtes dues aux effets d'ombre éliminés. 50

1.18 Arêtes de l'image "BOLT6". (a) Arêtes d'ordre 0. (b) Arêtes d'ordre 1. (c) Superposition des deux types d'arêtes. (d) Les mêmes résultats que dans (c) avec arêtes dues aux effets d'ombre éliminés. 51

CHAPITRE 2 **52**

2.1 Formes locales des primitives de surface. 55

2.2 Resultats d'extraction de plans. (a) Plans extraits avant filtrage morphologique. (b) Plans extraits après filtrage morphologique. (c) Image 3-D originale. 69

2.3 Segmentation de l'image "BLOC11". (a) Carte d'étiquettes initiale. (b) Carte d'étiquettes finale. (c) Carte d'arêtes. (d) Superposition des cartes d'arêtes et d'étiquettes finales. 75

2.4 Segmentation de l'image "BLOC9". (a) Carte d'étiquettes initiale. (b) Carte d'étiquettes finale. (c) Carte d'arêtes. (d) Superposition des cartes d'arêtes et d'étiquettes finales. 76

2.5 Segmentation de l'image "BLOC4". (a) Carte d'étiquettes initiale. (b) Carte d'étiquettes finale. (c) Carte d'arêtes. (d) Superposition des cartes d'arêtes et d'étiquettes finales. 77

2.6 Segmentation de l'image "BLOC15". (a) Carte d'étiquettes initiale. (b) Carte d'étiquettes finale. (c) Carte d'arêtes. (d) Superposition des cartes d'arêtes et d'étiquettes finales. 78

2.7 Segmentation de l'image "BLOC16". (a) Carte d'étiquettes initiale. (b) Carte d'étiquettes finale. (c) Carte d'arêtes. (d) Superposition des cartes d'arêtes et d'étiquettes finales. 79

CHAPITRE 3 **80**

3.1 Représentation dans l'espace d'échelles. (a) Exemple de signal 1-D. (b) Convolution du signal avec $g''(x)$ de grand support. (c) Carte d'échelles. Les passages par zéro du signal montré en (b) se trouvent le long de la ligne horizontale $\sigma = \sigma_l$. 83

3.2 Types de contours trouvés dans la carte d'échelles. Les contours de type (1) et (2) sont courants tandis que ceux de types (3) et (4) n'y apparaissent jamais. 84

3.3 Représentation multirésolution d'images 3D. 85

3.4 Image synthétique originale 86

3.5 Arêtes extraites à différentes échelles. (a) $\sigma = 0.0$. (b) $\sigma = 1.95$. (c) $\sigma = 3.7$. 88

3.6 Image filtrée à différentes échelles. (a) Image non filtrée ($\sigma = 0.0$). (b) $\sigma = 1.95$. (c) $\sigma = 3.7$. 90

3.7 Représentation à différentes échelles. (a) $\sigma = 0.0$. (b) $\sigma = 1.95$. (c) $\sigma = 3.7$. 91

CHAPITRE 4 **93**

LISTE DES TABLEAUX

CHAPITRE 1 **7**

CHAPITRE 2 **52**

2.1 Définition des huit primitives de surface. Le cas où $K > 0$ et $H = 0$ n'est pas défini. 74

CHAPITRE 3 **80**

CHAPITRE 4 **93**

INTRODUCTION

Toute considérations économiques mise à part, l'attrait que connait aujourd'hui la vision par ordinateur vient de la constatation que, de tous nos sens, le sens de la vue est celui qui offre la plus grande flexibilité et nous apporte la plus grande quantité d'information sur ce qui nous entoure. La perception de notre espace, la reconnaissance des objets qui le composent et la détection de leur mouvement sont des données essentielles pour l'humain et sont toutes transmises par le biais de la vision. A ces considérations s'ajoute un facteur non négligeable qu'est la séduction de réaliser un système de perception artificiel; l'histoire montre en effet le désir qu'a l'homme de créer un être à son image, depuis les premiers automates jusqu'aux robots capables de percevoir leur environnement et de réagir à celui-ci. L'arrivée sur le marché d'ordinateurs de plus en plus performants a rendu possible ce qui n'était qu'une vue de l'esprit il y a une vingtaine d'années.

Tout système de vision par ordinateur a comme première composante un capteur d'images. Son type détermine la nature de l'image captée et les étapes subséquentes de traitement à effectuer. Ceci implique entre autres une distinction faite entre la *vision 2-D* et la *vision 3-D*. En vision 2-D, l'image de la scène observée, captée et numérisée sous forme d'un tableau I de N lignes et M colonnes, représente une information proportionelle à la puissance lumineuse rayonnée par les différents points visibles de la scène. Cette *image 2-D* (ou image photométrique, ou encore image à niveaux de gris) véhicule donc de l'information relative à la géométrie des surfaces (leur forme, leur position, leur orientation), aux propriétés de réflectance et de texture de celles-ci, à l'éclairage, aux caractéristiques du milieu intermédiaire et à celles du capteur (résolution spatiale, plage dynamique, bruit, paramètres des caméras, ...). Cette quantité énorme de facteurs qui entrent en ligne de compte dans le processus de formation de

l'image fait de l'analyse et de l'interprétation des images 2-D un problème difficile à résoudre. Une étape intermédiaire de la vision 2-D consiste donc à inférer des images à niveaux de gris l'information tridimensionnelle sous-jacente à leur formation. Ceci constitue un problème fort complexe qui ne peut être résolu que pour des cas limités. Dans le cas de la vision 3-D, la valeur en chaque pixel de l'*image 3-D* (ou image télémétrique, ou encore image à données de profondeur) représente la distance séparant le point correspondant de la surface observée du plan focal du capteur. Ainsi, non seulement l'information de profondeur entre les différentes régions de l'image est disponible mais en plus la forme à l'intérieur de chaque région reflète celle de la surface correspondante. Cette caractéristique des images 3-D met la vision 3-D à l'abri de problèmes causés par les ombres et le marquage des surfaces.

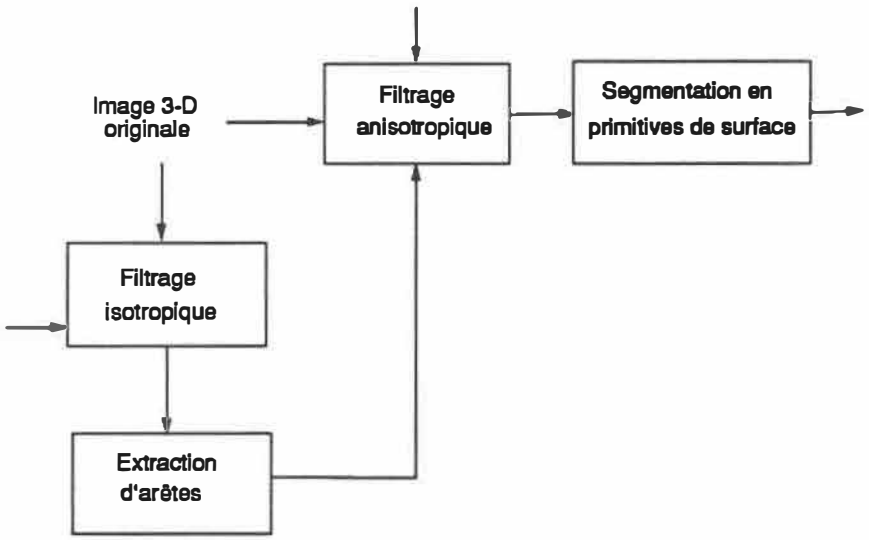
En général, les surfaces de la nature sont hiérarchiquement organisées en un ensemble discret de niveaux ([1], [3]). Au niveau le plus fin, un arbre est constitué de feuilles à l'intérieur de chacune desquelles se présente toute une structure complexe de veines. Au prochain niveau, chaque feuille est remplacée par une simple région, et au niveau le plus grossier il ne reste qu'une seule tâche à la place du feuillage. A chacun de ces niveaux de description est associé un intervalle de niveaux de résolution (échelles). De plus, les régions (feuilles, feuillage) à chaque niveau de description possèdent des frontières ou contours bien définies. Ces considérations suggèrent de concevoir, pour l'analyse des images 3-D, des représentations permettant la description des surfaces à différents niveaux de résolution, et requièrent donc la mise au point d'approches multirésolution pour la segmentation de ces surfaces.

En vision 3-D, les arêtes sont directement reliées à des structures significatives des surfaces ([4], [5], [6], [7]). Habituellement, elles consistent en des *arêtes d'ordre 0* et *d'ordre 1*, qui correspondent aux discontinuités de profondeur

et d'orientation des surfaces, respectivement. Les changements abrupts de profondeur se produisent en général à la frontière des objets tandis que la structure à l'intérieur des surfaces, délimitant ces objets, est décrite par les discontinuités d'orientation. Ainsi, l'extraction d'arêtes conserve l'information utile concernant la structure des surfaces visibles tout en réduisant la quantité de données à traiter par la suite. En plus, elle partitionne l'image en des régions à l'intérieur desquelles la profondeur et l'orientation des surfaces demeurent continues.

Contrairement à l'extraction d'arêtes, la segmentation d'images 3-D en primitives surfaciques est plus spécifique en ce qui concerne les régions qui en résultent, et met en évidence les propriétés géométriques des surfaces physiques correspondantes. La géométrie différentielle stipule que la forme locale d'une surface est uniquement déterminée par ses première et seconde formes fondamentales, et que la courbure Gaussienne K et la courbure moyenne H combinent respectivement de manière intrinsèque et extrinsèque ces formes fondamentales en des mesures scalaires qui sont invariantes aux rotations, aux translations, et aux changements de représentations paramétriques des surfaces ([8], [9]). En plus, Besl et Jain [10] ont indiqué que les différentes combinaisons de signes de ces courbures donnent naissance à un ensemble de huit primitives surfaciques, indépendantes de l'angle de vue, qui sont connues sous les noms *"surface pic"*, *"surface creuse"*, *"surface crête"*, *"surface plane"*, *"surface vallée"*, *"surface crête en selle"*, *"surface minimale"*, et *"surface vallée en selle"*, respectivement. Ainsi, nous pouvons associer à chaque point d'une surface une étiquette selon les signes des courbures calculées en ce point. Ceci constitue une approche intéressante pour la segmentation et la description de surfaces 3-D parce que les descripteurs possèdent des propriétés d'invariance, qui sont idéales pour l'identification d'objets à partir d'angles de vue arbitraires. Ces descripteurs conduisent également à une représentation locale, ce qui est désirable dans des cas d'occlusion sévère.

Nous proposons, dans ce mémoire une approche multirésolution pour la segmentation des surfaces 3-D, qui combine les techniques de segmentation que nous venons de mentionner, soient l'extraction d'arêtes et la segmentation en primitives de surface. Nous conjecturons qu'à un niveau de résolution donné, l'importance des arêtes et la forme des surfaces à l'intérieur des régions intermédiaires sont respectivement déterminées à l'aide d'un filtrage gaussien, appliqué isotropiquement et anisotropiquement sur l'image 3-D originale. En d'autres termes, l'image est d'abord filtrée isotropiquement, afin d'en extraire les arêtes; elle est ensuite filtrée de façon anisotropique, c'est-à-dire en tous points excepté aux points d'arêtes détectés, afin de permettre la segmentation des régions intermédiaires en primitives de surface. Ce principe est mieux illustré par le diagramme suivant:



Il est à noter que le support du filtre Gaussien est déterminé en fonction du niveau de résolution voulu.

L'extracteur d'arêtes proposé utilise des opérateurs différentiels du premier et second ordre pour extraire les arêtes d'ordre 0 et d'ordre 1, respectivement. Sa caractéristique principale consiste en l'utilisation d'une procédure de focalisation qui assure une localisation précise et une bonne classification des arêtes. A

un niveau de résolution donné, les arêtes sont détectées et leur position exacte est retracée par une technique de focalisation qui diffère de celle proposée par Bergholm [12] sur plusieurs aspects importants. Les arêtes détectées sont, en premier lieu, subdivisées en segments, ne présentant ni de changements rapides dans l'amplitude ni de variations brusques dans la direction. La procédure de focalisation est par la suite appliquée de manière indépendante sur chacun de ces segments, ne validant les pixels d'arêtes détectés au prochain niveau de résolution que lorsqu'ils sont consistents avec ceux obtenus au niveau courant. Une fois focalisés, les segments sont classifiés en arêtes significatives ou détails sans importance. La procédure de segmentation en primitives de surface utilise, de son côté, une méthode d'approximation de surface basée sur la technique standard des moindres carrées pour le calcul local des courbures Gaussienne K et moyenne H. La meilleure fenêtre d'approximation est automatiquement sélectionnée aux points se trouvant dans le voisinnage des discontinuités. Ensuite, à chaque point de l'image est associé, selon les signes des courbures calculées, un type de primitive parmi les huit qui sont possibles. La combinaison de la *carte d'arêtes*, obtenue par l'extracteur d'arêtes, et la *carte d'étiquettes*, produite par la procédure de segmentation en primitives de surface, constitue la description de l'image 3-D au niveau de résolution choisi.

Le chapitre 1 du présent mémoire décrit la méthode d'extraction d'arêtes proposée. La section 1.1 décrit les principes généraux du problème et présente une brève revue des méthodes les plus utilisées, tant en vision 2-D qu'en vision 3-D. De plus, elle indique les problèmes majeurs, rencontrés lors de l'extraction d'arêtes à différents niveaux de résolution, pour introduire finalement le principe de la méthode proposée ainsi qu'une esquisse de ses différentes étapes. Celles-ci sont par la suite détaillées dans les sections subséquentes (sections 1.2 à 1.5). Les résultats expérimentaux obtenus pour cette partie de l'approche sont donnés

dans la section 1.6. La méthode complète de segmentation en primitives de surface fait l'objet du chapitre 2. Celui-ci est organisé de la même manière que le chapitre précédent. La section 2.1 retrace les principes généraux du problème, effectue une brève revue bibliographique du domaine et esquisse les principales étapes de notre méthode. La section 2.2 traite par la suite la question de filtrage anisotropique et propose une technique simple qui tient compte de la disponibilité des arêtes extraites. Une méthode de calcul permettant l'estimation précise des valeurs locales de courbures Gaussienne K et moyenne H à partir des données de profondeur fait l'objet de la section 2.3. La classification des points de surface en différents types de primitive et la production de la description finale occupent les sections 2.4 et 2.5. Comme dans le chapitre précédent, nous présentons des résultats expérimentaux pour cette partie de l'approche dans la dernière section. Le chapitre 3 propose un exemple de représentation multirésolution. Nous terminons le mémoire avec le chapitre 4 qui donne une conclusion sur tout ce qui a été fait ainsi que les améliorations futures.

CHAPITRE 1

Approche multirésolution pour l'extraction d'arêtes dans les images 3-D

1.1 Extraction d'arêtes: principes et revue bibliographique.

Les arêtes constituent une partie importante de la représentation finale décrite dans ce mémoire. Leur extraction est depuis une vingtaine d'années le sujet principal d'une quantité non négligeable de travaux de recherche. La première section du présent chapitre décrit le principe général de l'extraction d'arêtes suivi d'une brève revue bibliographique du domaine.

L'extraction d'arêtes à différentes échelles doit cependant faire face à deux problèmes majeurs: premièrement, les arêtes détectées sont plus ou moins déplacées par l'opération de filtrage et deuxièmement, elles ne correspondent pas toutes aux arêtes physiques de la scène. Une méthode dont les différentes étapes sont présentées dans les sections restantes de ce chapitre est proposée pour extraire, tout en remédiant à ces problèmes, des arêtes à différents niveaux de résolution.

1.1.1 Principes.

En général les arêtes physiques d'une scène génèrent, par l'intermédiaire du processus de formation d'images, des changements brusques d'intensité ou de profondeur dans l'image 2-D ou 3-D correspondante . Cependant, à cause de la présence du bruit dans l'image et d'autres phénomènes physiques indésirables

(ombres, textures, ...), l'inverse n'est pas forcément vrai: un changement brusque d'intensité ou de profondeur ne correspond pas nécessairement à une arête physique de la scène. De ce fait, le but ultime de l'*extraction d'arêtes* est la caractérisation de ces changements en termes de propriétés des évènements physiques qui les ont générés: 1) arêtes d'objets, ombres, changements d'albedo, réflections spéculaires pour les images 2-D; 2) discontinuités de profondeur, arêtes, textures fines pour les images 3-D. En vision 2-D, ce but ne peut être atteint en une seule étape. L'extraction d'arêtes se divise souvent en deux sous-problèmes: le premier consiste à détecter et à caractériser les changements d'intensité de l'image, et le deuxième à faire l'usage de la description qui en résulte, ainsi que de connaissances de haut niveau, afin d'inférer des structures tridimensionnelles correspondantes. En vision 3-D, ce deuxième problème nécessite en général moins d'efforts en termes de traitements, puisque les arêtes extraites sont directement reliées à des structures significatives de la scène. La détection et la caractérisation de changements brusques de profondeur demeurent cependant un problème à résoudre.

La différentiation accentue en général les changements brusques d'une fonction, rendant ainsi explicites ses discontinuités. De ce fait, nous pouvons considérer l'extraction d'arêtes comme étant un problème de différentiation numérique suivi d'une étape de détection d'éléments mis en évidence. Afin de pouvoir caractériser les différentes catégories de changements, il est nécessaire d'utiliser des dérivées de différents types (dérivées directionnelles, dérivées rotationnellement invariantes, etc...) et de différents ordres (dérivées premières et secondes). Il est cependant bien connu que le fait de dériver numériquement une image s'avère être non robuste au bruit présent. Même si le bruit a une faible amplitude, il cause de grandes perturbations dans les résultats finaux. Par conséquent, il est souvent souhaitable de faire précéder la dérivation d'une opération de filtrage afin d'éliminer autant que possible les variations hautes-fréquences du signal. Torre

et Poggio [13] ont récemment donné une justification formelle à l'emploi de cette séquence filtrage-dérivation dans les approches d'extraction d'arêtes. En effet, la différentiation numérique est un problème mathématiquement mal posé (au sens de Hadamard) et le filtrage permet de régulariser ce problème. La régularisation, dans le cas d'images réelles, consiste à les convoluer avec la dérivée première d'une spline cubique qui est similaire à convoluer avec la dérivée première d'une Gaussienne.

Rappelons-nous cependant que les arêtes physiques dans la scène, et par conséquent les changements significatifs dans l'image, se produisent à différents niveaux de résolution. Habituellement, la détection à différentes échelles de ces changements peut être accomplie en effectuant différents degrés de filtrage sur l'image. Cependant, un filtre qui favorise la caractérisation de détails fins manque en général les structures grossières dans l'image, et celui qui permet l'extraction de traits grossiers estompe souvent les détails importants. De plus, à chaque niveau de résolution, certaines des caractéristiques extraites ne correspondent pas nécessairement aux changements physiques de la scène. Ainsi, il est probable que les fluctuations rapides détectées aux niveaux fins soient issues d'effets indésirables du bruit et que celles extraites aux niveaux grossiers correspondent aux fusionnements de discontinuités significatives avec des détails sans importance. Ces considérations suggèrent donc de définir une approche multirésolution, permettant d'une part de détecter les changements brusques à différents niveaux de résolution et d'autre part de mettre en relation des éléments détectés qui correspondent à un même évènement physique. Ce type d'approche est justifié aussi par certaines évidences biologiques qui révèlent la présence de filtres de différentes tailles dans le système visuel humain ([1], [2]).

L'utilisation d'un filtre Gaussien, bien que non optimale ([13]), offre toutefois les avantages suivants: 1) le simple ajustement du paramètre σ permet de choisir

le niveau de résolution auquel les arêtes sont à extraire; 2) aucun nouveau passage par zéro n'est introduit lorsque la taille du support augmente ([33]); 3) le filtre Gaussien bi-dimensionnel se décompose en un produit de deux filtres Gaussiens à une dimension, ce qui réduit considérablement la quantité de calculs à effectuer.

1.1.2 Revue bibliographique.

Dans cette section, nous n'avons ni l'intention d'effectuer une revue bibliographique complète des différentes approches d'extraction d'arêtes ni l'intention de décrire les méthodes dans leurs détails. Plutôt, nous présentons les idées clefs de certaines des méthodes les plus populaires, tant en vision 2-D qu'en vision 3-D.

1.1.2.1 Extraction d'arêtes dans les images 2-D.

Durant les vingt dernières années, des efforts considérables ont été investis dans la résolution du problème d'extraction d'arêtes des images 2-D. Ces efforts ont conduit à des résultats fort encourageants qui indiquent une certaine maturité du domaine. Les approches ont de plus en plus tendance à être analytiques. Des études approfondies ont été effectuées sur plusieurs filtres et opérateurs différentiels afin d'en déduire ceux qui sont les plus intéressants ([14], [13]).

Marr et Hildreth [2] ont proposé un détecteur de changements d'intensité qui emploie un filtre de réponse impulsionnelle gaussienne et un Laplacien comme opérateur différentiel. Les éléments significatifs à détecter sont les passages par zéro du signal obtenu après différentiation. Les changements sont extraits de manière indépendante à différents niveaux de résolution, et leur mise en relation est basée sur l'hypothèse de *coïncidence spatiale*. Celle-ci stipule que si des passages par zéro sont localisés et orientés à peu près de la même manière dans l'image à plusieurs échelles consécutives, l'ensemble de ces passages par zéro indiquera la présence d'une arête physique dans la scène. Le choix du filtre Gaussien

a été motivé par des arguments plutôt heuristiques: une bonne localisation exige l'emploi d'un filtre à support spatial limité et la sélectivité d'échelle impose l'usage d'un filtre à support fréquentiel limité. Comme ces deux contraintes ne peuvent être simultanément satisfaites, la minimisation du produit conduit alors au filtre Gaussien.

Canny [22], [23] a employé des critères tels que bonne détection, bonne localisation et réponse unique pour trouver le détecteur d'arêtes optimal. Le premier critère impose la minimisation des probabilités de manque et de fausse détection. Le deuxième critère contraint la minimisation de la distance quadratique moyenne qui sépare les éléments détectés de la vraie arête. Le dernier critère impose des contraintes sur la distance moyenne entre les extrémas consécutifs de la réponse du filtre. Dans le cas 1-D, le détecteur optimal d'arêtes de type échelon a été trouvé comme étant une combinaison linéaire de quatre exponentielles, qui peut être bien approximée par la dérivée première d'une Gaussienne. Dans le cas 2-D, le détecteur optimal résulte de la convolution du détecteur 1-D correspondant, aligné orthogonalement à la direction de l'arête, avec une fonction de projection qui est parallèle à la direction de l'arête. La fonction de projection la plus simple et la moins coûteuse en termes d'efforts de calcul est la Gaussienne. Afin d'atteindre une meilleure performance et une meilleure localisation, l'extraction d'arêtes s'effectue à différents niveaux de résolution, en utilisant des opérateurs de différentes largeurs, longueurs et orientations et en faisant appel à un seuillage adaptatif pour éliminer les détails diffus ainsi qu'à la synthèse de caractéristiques pour combiner les réponses obtenues à différentes échelles. Cependant, dû à l'usage d'opérateurs directionnels, les arêtes extraites ne sont pas toutes fermées ([13])

Poggio, Voorhees, et Yuille [34] et Torre et Poggio [13] ont montré que la différentiation numérique est un problème mal posé, et que celui-ci nécessite d'être

régularisé afin d'être bien posé au sens de Hadamard. Dans le cas des images réelles, l'opérateur optimal de régularisation est une spline cubique qui est très semblable à une Gaussienne. Poggio, Voorhees, et Yuille [34] ont aussi analysé le rôle du paramètre de régularisation dans le contrôle de la résolution à laquelle les changements d'intensité dans l'image sont à extraire.

1.1.2.2 Extraction d'arêtes dans les images 3-D.

En vision 3-D, plusieurs solutions ont aussi été proposées pour résoudre le problème d'extraction d'arêtes. Cependant, les résultats obtenus sont encore loins d'être fiables, efficaces, et utilisables pour toutes applications. Ceci est dû aux problèmes suivants auxquels les algorithmes d'extraction d'arêtes doivent faire face:

1. Les arêtes significatives de l'image, c'est-à-dire ceux qui sont utiles à la représentation d'objets tridimensionnels, à leur reconnaissance et à leur positionnement dans l'espace, sont de nature très différente les uns des autres. Les *sauts de profondeur*, ou *arêtes d'ordre 0*, doivent leur existence à la présence de discontinuités de profondeur dans la scène. Celles-ci délimitent en général les surfaces d'objets isolés ou localisent, dans le cas d'objets multiples, les frontières des lieux de recouvrements. Les autres types d'arêtes résultent de changements brusques d'orientation (*arêtes d'ordre 1*) ou de courbure (*jointures douces* ou *arêtes d'ordre 2*) des surfaces. Ils révèlent en général les détails de la surface frontière de chacun des objets. Cette différence de nature des arêtes nécessite donc l'emploi de différents opérateurs pour la détection et rend ainsi plus difficile l'uniformisation des étapes subséquentes de traitement.

2. Comme les arêtes résultent de phénomènes physiques hautes-fréquences, leur extraction est facilement perturbée par la présence du bruit dans l'image

ou par la présence d'autres phénomènes physiques non significatifs dans la scène (ombrage, texture fine des surfaces, ...). Ces perturbations, entrainent en général des résultats indésirables dans la réponse des opérateurs utilisés (arêtes mal localisées, réponses non justifiées par des phénomènes significatifs de la scène, ...) et introduit en conséquence de l'ambiguité dans l'interprétation finale des arêtes extraites. Il est à noter que la régularisation par filtrage ne fait qu'atténuer les effets nuisibles du bruit sur les résultats obtenus mais ne les élimine pas complètement.

3. A chaque arête est associé un intervalle de résolutions à l'intérieur duquel elle demeure perceptible, et, en général, la scène observée est organisée de façon hiérarchique: elle est constituée de structures principales de large taille sur laquelle se superposent des structures de plus en plus détaillées. La détection d'arêtes doit donc refléter cette organisation hiérarchique.

Plusieurs opérateurs ont été proposés pour répondre à la diversité des types d'arêtes. Certains sont des emprunts à des techniques utilisées en vision 2-D, d'autres sont propres à la vision 3-D. Les premiers ont été repris parce que les arêtes d'ordre 0 en vision 3-D ont exactement le même profil que les arêtes rencontrées dans les images 2-D. L'existence d'opérateurs propres à la vision 3-D est rendue possible par le fait que les images traitées fournissent directement de l'information concernant la structure géométrique des surfaces observées. Ainsi, les opérateurs les plus utilisés pour mettre en évidence les arêtes d'ordre 0 sont ceux qui font appel aux dérivées premières ou qui permettent le calcul local de variations de profondeur (Inokuchi-Nevatia [15], Mitiche-Aggarwal [16], Tomita-Kanade [17], Herman [18]). Dans le cas des arêtes d'ordre 1 les opérateurs les plus typiques sont ceux qui utilisent les dérivées secondes (Bhanu-Lee-Ho-Henderson [19]), qui quantifient les changements d'orientation de la normale à la surface (Tomita-Kanade [17], Bhanu-Lee-Ho-Henderson [19]), ou qui permettent

le calcul de variations de la courbure dans une seule direction (Herman [18]), suivant plusieurs directions (Fan-Medioni-Nevatia [6], [7]), ou encore le long de lignes caractéristiques telles que les lignes de courbure (Brady-Ponce-Yuille-Asada [4]).

Une méthode souvent utilisée pour éliminer les effets indésirables causés par le bruit consiste à faire suivre l'étape de détection de points d'arête d'une opération de seuillage, accompagnée du chainage de points d'arête restants (Inokuchi-Nevatia [15], Tomita-Kanade [17], Herman [18], Bhanu-Lee-Ho-Henderson [19]). Il s'agit là d'une approche plus ou moins arbitraire. Récemment des techniques plus efficaces ont été proposées, au prix d'une plus grande complexité algorithmique et d'implantation. Elles consistent en des approches multirésolution qui détectent et combinent en une représentation unique les arêtes significatives à différentes échelles. Parmis ces techniques, citons celles proposées par Brady-Ponce-Yuille-Asada [4], Ponce-Brady [5], and Fan-Medioni-Nevatia [6], [7].

Brady, Ponce, Yuille, et Asada [4] ont proposé une approche à deux étapes pour détecter les discontinuités significatives de surface à différents niveaux de résolution. Elle se base sur le théorème des lignes de courbure et sur les travaux de Asada et Brady [35]. La première étape consiste à extraire les lignes de courbure des surfaces et la deuxième étape utilise l'algorithme de Asada et Brady [35], pour mettre en évidence les variations rapides de courbure le long des lignes extraites. Aucune mention n'a été faite concernant la possibilité de combiner en une seule représentation utile les descriptions obtenues à différents niveaux de résolution. Plus récemment, Ponce et Brady [5] ont identifié plusieurs problèmes dont souffre l'algorithme: inefficacité au point de vue calculs, ambiguité d'interprétation due aux différents niveaux de filtrage, etc....

Ponce et Brady [5] ont généralisé les travaux de Asada et Brady [35] et ceux de Brady, Ponce, Yuille, et Asada [4] à deux dimensions. En supposant que les

surfaces sont localement cylindriques, ils ont modélisé différents types de discontinuités de surface (discontinuités de profondeur, d'orientation, et de courbure) et ont effectué une étude mathématique de ces modèles afin de déduire les caractères distinctifs de chacun d'eux face à des variations de résolution. Ils ont proposé par la suite une approche à trois étapes pour détecter, localiser, et décrire symboliquement les discontinuités significatives de surface: les deux premières étapes consistent, pour un ensemble discret de valeurs de σ, à effectuer un filtrage gaussien sur l'image, à calculer les courbures et les directions principales en chaque point de l'image et à mettre en évidence les extrémums des courbures dominantes et les passages par zéro de la courbure gaussienne; la dernière étape extrait les différents types d'arêtes en établissant une correspondance, sur la base des modèles utilisés et de leur comportement caractéristique dans l'espace d'échelle, entre les points caractéristiques mis en évidence à chaque niveau σ par les deux étapes précédentes.

L'approche de Fan, Medioni, et Nevatia [6], [7] est similaire à celle proposée par Ponce et Brady [5]. Elle se distingue de celle-ci dans les détails d'implantation, notamment au niveau de la mise en évidence de points caractéristiques, qui s'effectue suivant les directions 0^o, 45^o, 90^o, et 135^o et qui nécessite en conséquence un algorithme de correspondance plus simple, appliqué sur les points détectés dans chacune des directions ainsi qu'une étape additionnelle de fusionnement, destinée à l'élimination de points redondants. L'identification des différents types d'arête à partir des points caractéristiques retenus est explicitement entreprise dans une étape subséquente. Il est à noter qu'aucune explication n'est fournie sur le choix des valeurs utilisées pour σ, et que l'ensemble des types d'arête à extraire est moins riche que celui proposé par Ponce et Brady [5]. En plus, pour ce qui est de la version simplifiée de l'approche, destinée à être utilisée à un niveau σ unique et lorsque le niveau de bruit dans l'image est négligeable, les auteurs ne

font aucune suggestion concernant la valeur de σ à utiliser et la limite au delà de laquelle le bruit ne sera plus considéré comme étant négligeable.

1.1.3 Principe de l'approche proposée.

Puisque les arêtes sont mises en évidence par le biais du filtrage gaussien, appliqué isotropiquement sur l'image originale, leur position exacte n'est pas directement disponible au niveau de résolution choisi. En effet, le filtrage estompe les discontinuités et en conséquence déplace les arêtes correspondantes, la gravité du déplacement dépendant du niveau de résolution choisi. La position exacte des arêtes doit donc être récupérée ultérieurement. De plus, tous les points d'arête détectés dans l'image filtrée ne correspondent pas nécessairement aux structures significatives des surfaces. En conséquence, il est nécessaire de mener une analyse plus approfondie sur les arêtes relocalisées pour être capable de distinguer les *vrais arêtes* (correspondant à des structures significatives de surface) des *fausses arêtes* (induites par les détails non significatifs tels que la texture, le bruit, ...).

Les algorithmes proposés par Bergholm [12] et Sjöberg [21] ont été conçus pour résoudre les deux problèmes que nous venons de mentionner en vision 2 - D. L'algorithme de Bergholm [12] résulte d'une étude mathématique de la manière dont se déplacent les arêtes dans l'espace d'échelle et permet la récupération de la position exacte de ceux-ci. Il consiste à appliquer un certain nombre d'*étapes de focalisation* sur les points d'arête potentiels, détectés à un niveau de résolution initial $\sigma = \sigma_g$ donné. Chacune de ces étapes consiste à effectuer un filtrage de niveau de résolution plus fin $\sigma - \Delta\sigma$ (où $\Delta\sigma$ dénote la variation de σ entre deux étapes consécutives) et à détecter la présence de points d'arêtes dans un voisinage 3×3 de ceux mis en évidence au niveau de résolution courant σ. La détection est assurée par une version simplifiée du détecteur d'arêtes de Canny [22], [23]. A la fin de chaque étape, seuls les points d'arêtes nouvellement détectés sont retenus.

Ceux obtenus à la fin de la dernière étape, au niveau de résolution final $\sigma=\sigma_f$, constituent les arêtes recherchées avec une précision de localisation adéquate. Sjöberg et Bergholm [21] ont observé que, lorsque σ diminue, les fausses arêtes se déplacent en général dans les deux directions, à partir de leur positions initiales, et qu'ils ont une forte tendance à se subdiviser en des segments plus petits. Ces observations les ont conduit à proposer un algorithme de classification constitué de trois étapes. Premièrement, les points d'arête détectés au niveau σ_g sont classifiés en utilisant comme critère de classification le nombre de points obtenus en excès à σ_f. Deuxièment, les points d'arête obtenus à σ_f sont classifiés en utilisant comme unique source d'information les résultats de l'étape précédente. Troisièmement, la classification proprement dite des arêtes est effectuée au moyen d'une uniformisation des étiquettes le long de chacun d'entre eux.

Ces deux algorithmes produisent en général des résultats satisfaisants. Ils souffrent cependant des problèmes suivants:

- La procédure de classification échoue en général lorsqu'il s'agit d'arêtes parallèles ou d'arêtes trop proches les uns des autres. Cet échec est dû d'une part au manque d'interaction entre la focalisation et la classification et d'autre part au fait qu'en réalité, la procédure de classification proposée n'exploite pas toutes les caractéristiques (déplacement aléatoire et haute tendance à la subdivision) qu'ont les fausses arêtes (seul le nombre de points d'arête obtenus en excès au niveau de résolution le plus fin est pris en considération).

- Comme la procédure de focalisation ne s'applique qu'au niveau des points d'arête, elle ne produit aucune information concernant le déplacement et la subdivision des arêtes à travers l'espace d'échelle. De telle informations se révèlent pourtant très utiles à la classification subséquente, effectuée sur les arêtes. De plus, des contraintes additionnelles, telles que la continuité des arêtes, ne peuvent être introduites dans la procédure pour réduire la tendance

du processus de focalisation à récolter des points d'arête non significatifs.

- Certaines valeurs de paramètres doivent être estimées par essais successifs avant que l'algorithme de classification puisse vraiment être utilisé.

Les problèmes que nous venons d'énumérer nous indiquent qu'il est incorrect d'appliquer la procédure de focalisation au niveau des points d'arête et que l'utilisation de l'information extraite seulement au voisinage immédiat des points d'arête n'est pas assez pertinente pour accomplir la classification, surtout lorsque les objets se chevauchent substantiellement. Ils suggèrent de prévoir une interaction entre le module de focalisation et celui de classification de telle sorte que l'information concernant le comportement des différentes arêtes soit disponible au moment où la classification est effectuée. Ces considérations nous ont conduit à proposer l'approche suivante pour extraire les arêtes à partir des images 3-D:

Etape 1: Détecter toutes les arêtes (arêtes d'ordre 0 et d'ordre 1) potentiels au niveau de résolution initial σ_g donné.

Etape 2: Segmenter ces arêtes en des composants (segments) plus petits de telle sorte que ceux-ci ne contiennent aucune discontinuité dans leur amplitude et leur orientation.

Etape 3: Effectuer la focalisation sur chacun de ces composants, en ne gardant à chaque étape que les points d'arête qui sont consistants avec ceux détectés à l'étape précédente. A la fin de la procédure de focalisation, calculer le nombre de segments générés ainsi que leur longueur respective.

Etape 4: Effectuer la classification sur les segments d'arête focalisés, retenir ceux classifiés comme étant des vraies arêtes et rejeter les autres.

Les détails concernant chacune de ces étapes sont donnés dans les sections qui suivent.

1.2 Détection d'arêtes potentielles.

La procédure d'extraction d'arêtes commence par une étape qui consiste à détecter, au niveau de résolution initial $\sigma = \sigma_g$, toutes les arêtes d'ordre 0 $\{J_{i\sigma} \quad i = 1, 2, \ldots, N_J\}$ et les arêtes d'ordre 1 $\{R_{i\sigma} \quad i=1, 2, \ldots, N_R\}$, qui sont potentiellement significatives.

Soit $s(x, y, \sigma)$ le résultat de la convolution de l'image 3-D originale $z(x, y)$ avec l'opérateur Gaussien:

$$g(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp(-\frac{x^2 + y^2}{2\sigma^2}) \tag{1.1}$$

et soient $s_x(x, y, \sigma)$ et $s_y(x, y, \sigma)$ les dérivées partielles d'ordre 1 de $s(x, y, \sigma)$ par rapport à x et à y, respectivement.

Les arêtes d'ordre 0 $\{J_{i\sigma}\}$ sont définis comme étant l'ensemble $\mathcal{J}(\sigma)$ de points (x, y) qui correspondent aux maximums locaux de l'amplitude

$$mj(x, y, \sigma) = \sqrt{s_x^2 + s_y^2} \tag{1.2}$$

du gradient de $s(x, y, \sigma)$, détectés dans la direction

$$dj(x, y, \sigma) = \tan^{-1}(\frac{s_y}{s_x}) \tag{1.3}$$

de celui-ci. La détection de ces maximums locaux est accomplie par une version simplifiée de la méthode de suppression des points non maximum de Canny [22], [23].

Bien que les arêtes d'ordre 1 correspondent à des discontinuités d'amplitude du gradient, la dérivation de celles-ci ne produit pas nécessairement des maximums locaux dans la direction orthogonale à celle des arêtes. En conséquence, l'approche suivante est adoptée pour la détection des arêtes d'ordre 1:

1. Trouver l'ensemble $\mathcal{R}_1(\sigma)$ de points (x, y) correspondant aux maximums locaux de l'amplitude

$$mr_1(x, y, \sigma) = \sqrt{s_{xx}^2 + s_{xy}^2} \tag{1.4}$$

du gradient de $s_x(x, y, \sigma)$, détectés dans la direction

$$dr_1(x, y, \sigma) = \tan^{-1}(\frac{s_{xy}}{s_{xx}}) \tag{1.5}$$

de celui-ci.

2. Trouver l'ensemble $\mathcal{R}_2(\sigma)$ de points (x, y) correspondant aux maximums locaux de l'amplitude

$$mr_2(x, y, \sigma) = \sqrt{s_{yx}^2 + s_{yy}^2} \tag{1.6}$$

du gradient de $s_y(x, y, \sigma)$, détectés dans la direction

$$dr_2(x, y, \sigma) = \tan^{-1}(\frac{s_{yy}}{s_{yx}}) \tag{1.7}$$

de celui-ci.

3. Valider un point (x, y) comme étant un point d'arête d'ordre 1 si et seulement si $(x, y) \in \mathcal{R}(\sigma)$=$\mathcal{R}_1(\sigma) \cup \mathcal{R}_2(\sigma)$

Une fois détectés, les arêtes d'ordre 0 $\{J_{i\sigma_g}\}$ et les arêtes d'ordre 1 $\{R_{i\sigma_g}\}$ sont soumis à une simple opération de seuillage qui consiste à éliminer, autant que possible, les arêtes non significatives, et donc non utiles aux traitements subséquents. Le seuil est sélectionné de façon automatique et correspond à la valeur fournie par le premier zéro de l'histogramme, de l'amplitude du gradient sur toute l'image. Il est à noter que l'introduction de cette opération de seuillage a pour seul but de réduire le temps total de calcul.

Les Figures 1.1 et 1.2 présentent les résultats typiques de cette étape de détection d'arêtes potentielles. L'image 3-D originale est illustrée aux Figures 1.1.a

and 1.1.b. Le niveau de résolution initial est choisi comme étant $\sigma_g = 2.2$. Les arêtes d'ordre 0 détectées à σ_g sont illustrées à la Figure 1.1.c et les Figures 1.2.a et 1.2.b illustrent les arêtes dont les éléments constituants appartiennent à $\mathcal{R}_1(\sigma_g)$ et $\mathcal{R}_2(\sigma_g)$, respectivement. Les arêtes qui subsistent à l'opération de seuillage et qui constituent en conséquence des arêtes potentielles pour les traitements subséquents sont reproduits à la Figure 1.1.d (cas des arêtes d'ordre 0) et aux Figures 1.2.c et 1.2.d (cas des arêtes d'ordre 1). Il est à noter que ceux qui ont été éliminés correspondent aux arêtes non significatives provenant du bruit.

Puisque les arêtes d'ordre 0 et d'ordre 1 sont extraites de la même manière, sauf en ce qui concernent les opérateurs utilisés pour la détection de leur présence respective dans l'image, nous allons utiliser $\{E_{i\sigma}\ \ i{=}1,\ldots,N_E\}$ pour dénoter, indépendamment de leur type, les arêtes détectées à un niveau de résolution σ donné, ainsi que $m(x,y,\sigma)$ et $d(x,y,\sigma)$ pour dénoter respectivement les fonctions d'amplitude et de direction du gradient. Des spécifications additionnelles seront données là où jugé nécessaire.

1.3 Segmentation d'arêtes potentielles détectées.

Cette étape constitue une phase de prétraitement qui précède la localisation des arêtes par focalisation, décrite dans la prochaine section. Elle consiste à segmenter les arêtes potentielles $\{E_{i\sigma_g}\ \ i{=}1,\ldots,N_E\}$, détectés précédemment, en des composants plus petits, ou segments, $\{S_{j\sigma_g}\ \ j{=}1,\ldots,N_S\}$, $(N_S \geq N_E)$, dont chacun satisfait les contraintes suivantes:

- Un segment ne présente aucune discontinuité dans la valeur de la fonction d'amplitude $m(x,y,\sigma_g)$ qui lui est associée.
- Un segment ne présente aucune discontinuité dans la valeur de la fonction de direction $d(x,y,\sigma_g)$ qui lui est associée.

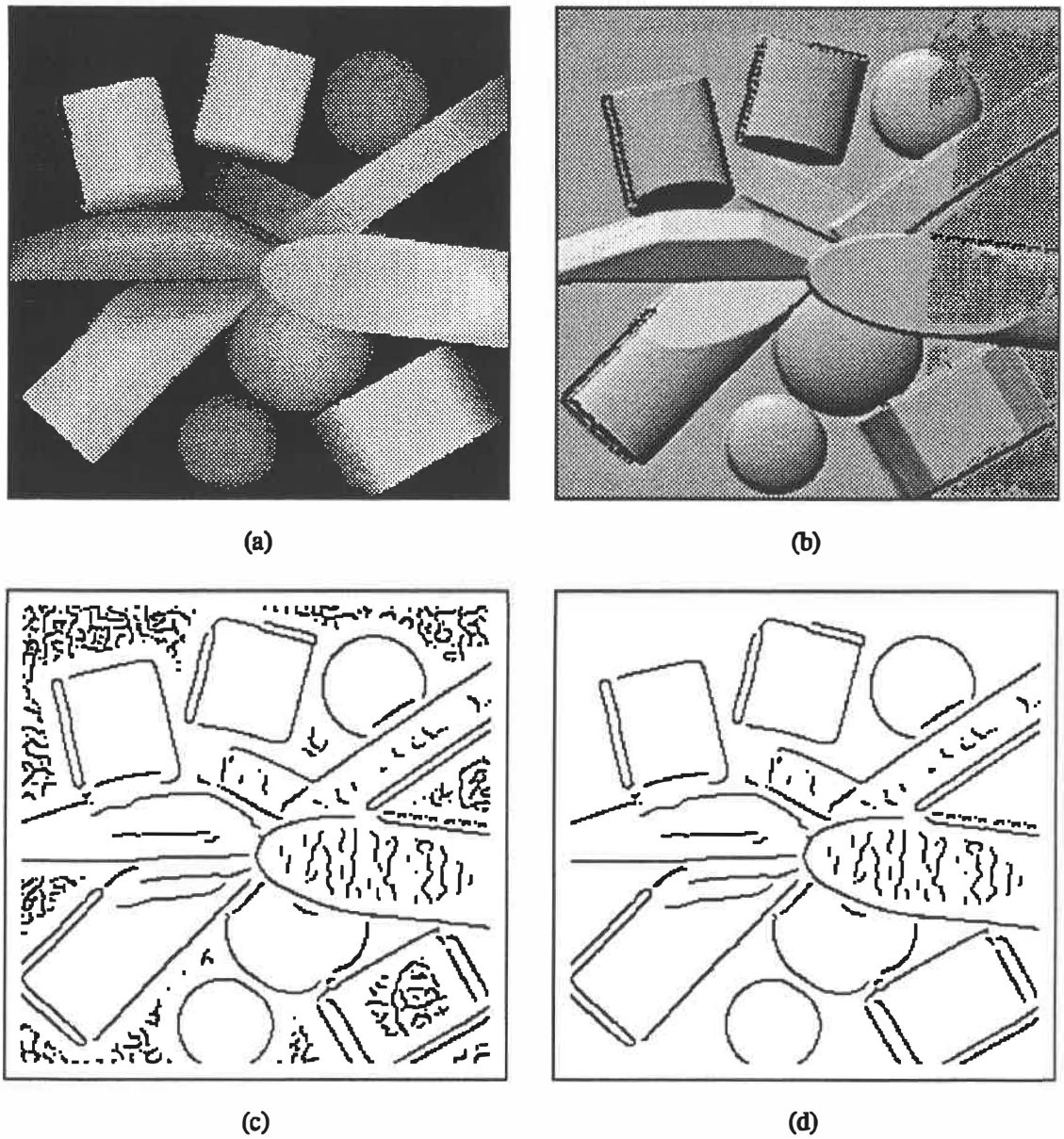


Figure 1.1 Arêtes d'ordre 0 détectées à $\sigma_g = 2.2$. (a) Image originale codée en niveaux de gris. (b) Image originale ombragée. (c) Arêtes détectées. (d) Arêtes qui subsistent après l'opération de seuillage.

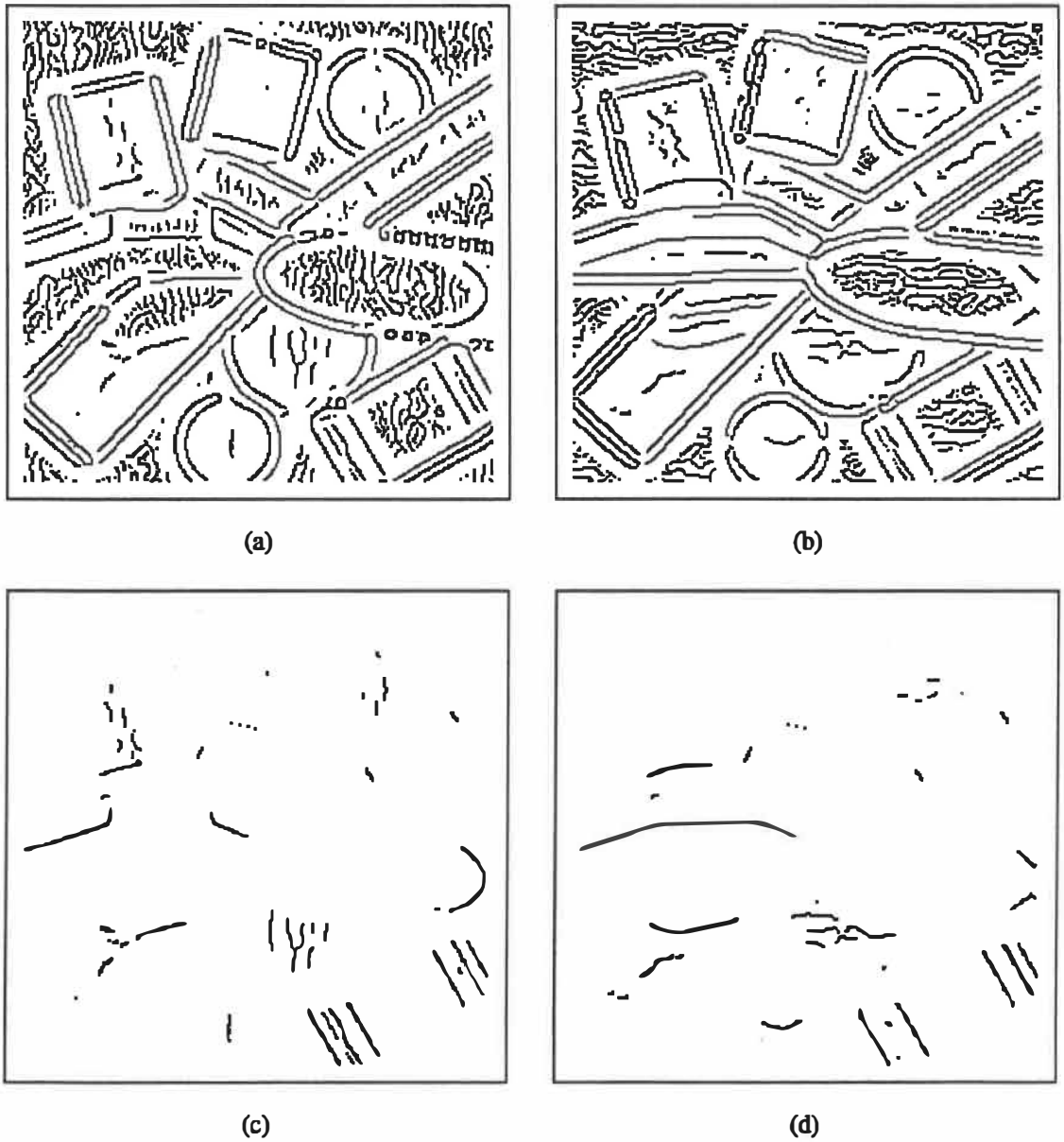


Figure 1.2 Arêtes d'ordre 1 détectées à $\sigma_g = 2.2$. (a) Points d'arête en provenance de $\mathcal{R}_1(\sigma_g)$. (b) Points d'arête en provenance de $\mathcal{R}_2(\sigma_g)$. (c)-(d) Arêtes d'ordre 1 qui subsistent à l'opération de seuillage.

Actuellement, la segmentation des arêtes $\{E_{i\sigma_g}\}$ en des segments $\{S_{j\sigma_g}\}$ est accomplie par un algorithme, initialement proposé par Saint-Marc et Medioni [24] pour effectuer l'extraction de points caractéristiques des courbes planaires. Il consiste à filtrer de façon adaptative la dérivée première de la fonction de courbure, calculée le long de la courbe en question. En chaque point du signal, le filtrage adaptatif est accompli en convoluant le signal original avec un filtre moyen de support très petit et modulé par une mesure de continuité, dictée par la valeur absolue de la dérivée du signal en ce point. Il est donc possible, par la suite, d'extraire les discontinuités de courbure (correspondant aux jointures douces) et les extremums de courbure (correspondant aux coins) qui constituent les points caractéristiques de la courbe. La même technique est appliquée ici sur les arêtes $\{E_{i\sigma_g}\}$ pour localiser les points où les fonctions $m(x,y,\sigma_g)$ et/ou $d(x,y,\sigma_g)$ présentent des variations brusques. En d'autres termes, le filtrage adaptatif est appliqué sur les dérivées premières de $m(x,y,\sigma_g)$ et $d(x,y,\sigma_g)$, respectivement, et les points de brise recherchés sont identifiés par les extrémums des signaux filtrés.

La Figure 1.3 montre des résultats types de cette étape de segmentation d'arêtes. La Figure 1.3.a illustre l'image des arêtes potentielles $\{E_{i\sigma_g}\}$, détectés au niveau de résolution σ_g=2.2. L'image des segments $\{S_{i\sigma_g}\}$ qui résultent de la segmentation est montrée à la Figure 1.3.b.

1.4 Localisation de segments d'arête par focalisation.

Cette troisième étape remplit deux fonctions distinctes. Premièrement, elle permet la récupération de la position exacte des segments $\{S_{j\sigma_g}\}$, et deuxièmement, elle fournit des informations concernant le comportement de chaque segment $S_{j\sigma_g}$ lorsque le paramètre σ varie. Elle utilise une technique de focalisation d'arêtes, ressemblant à celle proposée par Bergholm [12] dans le contexte des images 2-D, mais présentant plusieurs différences majeures.

La focalisation est appliquée de manière indépendante sur chaque segment $S_{j\sigma_g}$. Elle consiste en un nombre fini d'étapes, appelées *étapes de focalisation*, déterminé à partir des valeurs des paramètres σ_g, σ_f, et $\Delta\sigma$ qui définissent respectivement le niveau de résolution initial, le niveau de résolution final et le pas de focalisation (l'écart de résolution entre deux étapes consécutives de focalisation). En partant du niveau de résolution initial σ_g, la résolution σ est diminuée de $\Delta\sigma$ à chaque étape de focalisation. Du niveau σ_g au niveau $\sigma_g - \Delta\sigma$, un segment $S_{j\sigma_g}$ peut éventuellement se subdiviser en plusieurs sous-segments dénotés par $\{S_{jk\sigma_g-\Delta\sigma} \quad k=1,\ldots,N_j\}$ (l'indice k est introduit ici dans le but d'indexer les sous-segments qui résultent de la subdivision, tout en conservant leur identité originale j). Pour chaque segment $S_{j\sigma_g}$, la détection des sous-segments $\{S_{jk\sigma_g-\Delta\sigma}\}$ correspondants est effectuée dans un voisinage incluant tous les pixels qui coincident avec $S_{j\sigma_g}$ ou qui lui sont adjacents (en assumant la connectivité-8). Les étapes de focalisation subséquentes sont effectuées de façon similaire et l'ensemble des sous-segments $\{S_{jm\sigma_f}\}$, détectés au niveau le plus fin σ_f, est considéré comme étant la version la mieux localisée de $S_{j\sigma_g}$. La Figure 1.4 illustre le principe que nous venons de décrire.

A chaque étape de focalisation, un seuillage adaptatif rejette les points d'arête inconsistants. Aucune connaissance *a priori* n'est requise pour la sélection du seuil, qui est basée uniquement sur les propriétés des sous-segments détectés à l'étape de focalisation précédente. Etant donné que l'amplitude et l'orientation de chaque segment $S_{j\sigma_g}$ ne présentent aucune discontinuité, nous pouvons caractériser ce dernier par la moyenne $\mu_{j\sigma_g}$ et l'écart-type $\alpha_{j\sigma_g}$ des valeurs locales d'orientation, distribuées le long de $S_{j\sigma_g}$. De façon analogue, nous introduisons la moyenne $\mu_{j\sigma}$ et l'écart-type $\alpha_{j\sigma}$ pour caractériser l'ensemble des sous-segments $\{S_{jk\sigma}\}$, associé à $S_{j\sigma_g}$. Ainsi, en allant du niveau σ au niveau $\sigma-\Delta\sigma$, un point d'arête ne sera validé que si et seulement si l'orientation en ce point satisfait les

inégalités suivantes:

$$\mu_{j\sigma} - \alpha_{j\sigma} \leq d(x,y,\sigma-\Delta\sigma) \leq \mu_{j\sigma} + \alpha_{j\sigma} \tag{1.8}$$

De cette façon, le seuillage empêche la procédure de focalisation de collecter des points d'arête induits par le bruit ou des points dont l'orientation s'avère trop différente de $\mu_{j\sigma}$.

En résumé, la procédure de focalisation suivante est appliquée sur chaque segment $S_{j\sigma_g}$:

Etape 0:

- Calculer la moyenne $\mu_{j\sigma_g}$ et l'écart-type $\alpha_{j\sigma_g}$ du segment $S_{j\sigma_g}$.
- $\sigma = \sigma_g - \Delta\sigma$.

Etape 1: Détecter les points d'arête au niveau σ.

Etape 2: Effectuer le seuillage sur les points d'arête détectés au niveau σ selon le test (1.8).

Etape 3: Calculer la moyenne $\mu_{j\sigma}$ et l'écart-type $\alpha_{j\sigma}$ pour l'ensemble $\{S_{jk\sigma}\}$.

Etape 4: Décrémenter σ de $\Delta\sigma$ et si $\sigma \geq \sigma_f$ retourner à l'étape 1.

A la fin de la procédure, nous calculons, pour le segment $S_{j\sigma_g}$ original, la longueur L_{jm} de chaque sous-segment focalisé $S_{jm\sigma_f}$ et le nombre total N_j de sous-segments associés à $S_{j\sigma_g}$. Celui-ci mesure le degré de subdivision de $S_{j\sigma_g}$.

Nous présentons à la Figure 1.5 des résultats obtenus par notre méthode de focalisation et ceux donnés par la méthode de Bergholm [12]. Les niveaux de résolution initial et final sont 2.2 et 0.45, respectivement. La Figure 1.5.a illustre les segments d'arête détectés au niveau 2.2. Ceux obtenus après focalisation, utilisant notre approche et celle de Bergholm, sont montrés aux Figures 1.5.b et 1.5.c, respectivement. Pour mieux montrer les effets du seuillage adaptatif

sur le résultat final de notre méthode de focalisation, nous avons appliqué cette dernière ainsi que la méthode de Bergholm à un segment type, celui illustré à la Figure 1.5.d et correspondant au segment pointé par une flèche dans la Figure 1.5.a. Les résultats de ces applications sont présentés aux Figures 1.5.e et 1.5.f, respectivement. Ils illustrent bien le fait que, sans seuillage, la technique de Bergholm laisse subsister plus de détails indésirables que la notre dans les résultats finaux.

1.5 Classification des segments d'arête focalisés.

La procédure de focalisation que nous venons de décrire détecte non seulement des vraies arêtes mais aussi certaines arêtes indésirables, induites par le bruit ou par des détails non significatifs. En effet, l'opération de seuillage appliquée tout au début sur les arêtes $\{E_{j\sigma_g}\}$, détectés au niveau de résolution initial, n'est pas apte à éliminer toutes les arêtes non significatives. Ceux qui subsistent après le seuillage ne sont pas éliminés du résultat final par les étapes de focalisation subséquentes mais sont au contraire raffinés de manière progressive jusqu'au niveau final σ_f. Ceci suggère donc l'introduction d'une procédure de classification afin de pouvoir discerner, parmi les segments focalisés $\{S_{jm\sigma_f}\}$, ceux qui sont significatifs (vraies arêtes) de ceux qui ne le sont pas (fausses arêtes). Cette section décrit notre méthode de classification d'arêtes. Elle se compose de deux étapes: la première concerne la classification des pixels de chaque segment focalisé $S_{jm\sigma_f}$ et la deuxième se charge de la classification de $S_{jm\sigma_f}$ lui-même, en uniformisant les étiquettes qui lui sont attribuées à la première étape. Une comparaison avec la méthode de classification de Sjöberg [21] est effectuée à la fin de la section.

1.5.1 Classification au niveau des pixels.

Cette première étape de classification consiste à attribuer à chaque pixel d'un segment focalisé $S_{jm\sigma_f}$ une étiquette qui le caractérise comme étant un point d'arête véritable (étiquette: L_{vrai}) ou erronné (étiquette: L_{faux}) . Elle utilise un ensemble de règles, dérivées d'une analyse mathématique du déplacement et de la subdivision des arêtes à travers l'espace d'échelles. L'information locale, nécessaire à cette classification, est obtenue par superposition des segments focalisés $\{S_{jm\sigma_f}\}$ sur le segment générateur $S_{j\sigma_g}$ correspondant.

Il a été observé que, durant la focalisation, une vraie arête a souvent tendance de se déplacer dans une direction unique, à partir de la position du segment initial détecté au niveau σ_g. Une fausse arête, par contre, se déplace en général de manière aléatoire et se subdivise souvent en plusieurs sous-segments. Ces deux comportements caractéristiques des arêtes sont illustrés à la Figure 1.6. Il s'agit des superpositions de $\{S_{jm\sigma_f}\}$, montrés en noir, sur leur $S_{j\sigma_g}$ correspondant, présenté en gris (Figures 1.6.c et 1.6.d). Ces segments correspondent à ceux mis en évidence (segments noirs) dans les Figures 1.6.b et 1.6.a, respectivement.

En conséquence, si $N_{ml\sigma_f}$ dénote le nombre de segments focalisés qui se trouvent dans le voisinage du l^{ieme} pixel de $S_{jm\sigma_f}$, alors les règles de classification seront les suivantes:

- **Règle 1**: Si $N_{ml\sigma_f}$=1 alors attribuer au l^{ieme} pixel de $S_{jm\sigma_f}$ l'étiquette L_{vrai}.
- **Règle 2**: Si $N_{ml\sigma_f}$=2 alors attribuer au l^{ieme} pixel de $S_{jm\sigma_f}$ l'étiquette L_{vrai} si les deux segments focalisés sont situés d'un même côté du pixel, sinon lui attribuer l'étiquette L_{faux}.
- **Règle 3**: Si $N_{ml\sigma_f}\geq$3 alors attribuer au l^{ieme} pixel de $S_{jm\sigma_f}$ l'étiquette L_{faux}.

Pour le l^{ieme} pixel de $S_{jm\sigma_f}$, la valeur de $N_{ml\sigma_f}$ est calculée en effectuant une recherche dans la direction orthogonale à $S_{jm\sigma_f}$ et en comptant le nombre de

segments focalisés croisés. La recherche est limitée à une distance (en pixels), définie comme étant le double du nombre d'étapes de focalisation, déterminé à partir de σ_g, σ_f, et $\Delta\sigma$. En résumé, cette première étape de classification est constituée des deux sous-étapes suivantes:

Etape 1: Superposer les segments focalisés $\{S_{jm\sigma_f}\}$ sur le segment générateur $S_{j\sigma_g}$ correspondant.

Etape 2: Traiter chaque segment focalisé $\{S_{jm\sigma_f}\}$. Pour chacun de ses pixels,

- calculer la valeur $N_{ml\sigma_f}$ qui lui est associée à partir de l'image contenant la superposition de $\{S_{jm\sigma_f}\}$ et de $S_{j\sigma_g}$.
- utiliser les règles définies précédemment pour étiqueter le pixel.

Comme la méthode s'applique à chaque segment $S_{j\sigma_g}$ et traite chacun des segments focalisés $\{S_{jm\sigma_f}\}$ qui lui sont associés, elle s'adapte parfaitement aux traitements parallèles. En plus, elle s'avère efficace dans des régions de l'image où le chevauchement des objets est dense, grâce au seuillage adaptatif appliqué durant la focalisation, et permettant d'assurer la correspondance entre les segments focalisés et le segment générateur correspondant.

1.5.2 Classification au niveau des segments.

A la suite de l'étape de classification précédente, chaque segment focalisé $S_{jm\sigma_f}$ est constitué de pixels étiquetés "L_{vrai}" et/ou de pixels étiquetés "L_{faux}". En général, pour un segment $S_{jm\sigma_f}$, le nombre de pixels portant l'étiquette L_{vrai} sera plus grand que le nombre de pixels portant l'étiquette L_{faux} si $S_{jm\sigma_f}$ est issu d'une discontinuité significative de la scène. L'inverse s'applique aux segments $S_{jm\sigma_f}$ induits par le bruit ou issus de détails non significatifs. Ainsi, si R_{jm} dénote, pour un segment focalisé $S_{jm\sigma_f}$, le rapport du nombre de pixels portant l'étiquette L_{vrai}

sur sa longueur L_{jm} en pixels, les inégalités suivantes seront vérifiées en général

$$\begin{cases} R_{jm} < 0.5 & \text{si } S_{jm\sigma_f} \text{ est une fausse arête.} \\ R_{jm} > 0.5 & \text{si } S_{jm\sigma_f} \text{ est une vraie arête.} \end{cases} \tag{1.9}$$

Ceci conduit à la procédure suivante (Sjöberg [21]), qui est simple et qui permet la sélection de la bonne étiquette pour $S_{jm\sigma_f}$.

$$\text{étiquette} = \begin{cases} L_{faux} & \text{si } R_{jm} \leq 0.5 \\ L_{vrai} & \text{autrement} \end{cases} \tag{1.10}$$

Cependant dans la pratique, les inégalités dans (1.9) ne sont pas toujours satisfaites par tous les segments, et ceci quelque soit le type (vrai ou faux) de ceux-ci. D'une part, le filtrage gaussien, appliqué tout au début sur l'image 3-D originale, peut entraîner la fusion de discontinuités significatives avec des détails non significatifs. En conséquence, les vraies arêtes qui en sont issus peuvent avoir plus que la moitié de leurs pixels étiquetés avec L_{faux}. D'autre part, il peut arriver que certaines fausses arêtes aient tendance à ne pas se subdiviser en composantes plus petites, et donc possèdent plus de pixels portant l'étiquette L_{vrai} que de pixels étiquetés L_{faux}. Ces considérations conduisent à utiliser une fonction de seuillage et non pas un seuil fixe dans (1.9). La fonction choisie doit être capable de s'ajuster automatiquement selon le type du segment traité, lui-même déterminé par son comportement caractéristique durant la focalisation. Nous avons choisi les paramètres suivants pour caractériser ce comportement:

- le nombre $N_{j\sigma_f}$ de segments focalisés $\{S_{jm\sigma_f}\}$, issus du même segment $S_{j\sigma_g}$
- la longueur L_{jm} en pixels de chacun de ces segments focalisés $S_{jm\sigma_f}$

Ce choix a été motivé par le fait que les vraies arêtes sont généralement plus longues que les fausses arêtes et que celles-ci se subdivisent plus facilement en composantes plus petites que les premières.

Soit $t(n,l)$ la fonction de seuillage désirée, où n et l sont des variables indépendantes qui représentent $N_{j\sigma_f}$ et L_{jm}, respectivement. L'idée de base consiste à faire de $t(n,l)$ une fonction qui soit proportionnelle à n et inversement proportionnelle à l. En effet, plus n augmente, plus $t(n,l)$doit être augmentée parce qu'une augmentation dans la valeur de n indique une augmentation de la probabilité que le segment correspondant soit faux. De la même manière, si l augmente, $t(n,l)$ doit être diminuée car une incrémentation de l équivaut à une augmentation de la probabilité que le segment correspondant soit vrai. Cette idée nous a guidé dans le choix de $t(n,l)$.

Suivant les inégalités (1.9), l'expression de $t(n,l)$ peut être initialement choisie comme étant:

$$t(n,l) = \exp(-\frac{R_{jm}^2}{2\sigma^2}) = \exp(-\frac{R_{jm}^2}{A}) \tag{1.11}$$

où A est une constante et est choisie de telle sorte que $t(n,l)$=0.5 lorsque que R_{jm}=0.5. Ceci implique $A = 0.3607$

Si n augmente, $t(n,l)$ doit être augmentée et ceci peut être accompli en substituant σ par $\sigma \ln(1+n)$. Ainsi, en tenant compte uniquement du degré de subdivision du segment $S_{j\sigma_g}$, nous obtenons l'expression suivante de $t(n,l)$:

$$t(n,l) = \exp(-\frac{R_{jm}^2}{A\ln^2(1+n)}) \tag{1.12}$$

Nous procédons de la même manière pour tenir compte des variations de l. Si l dépasse un certain nombre N_0, la fonction $t(n,l)$ doit être diminuée et ceci en substituant σ par $\sigma \log(N_0 - l + 10)$. Dans le cas contraire, c'est-à-dire si l est plus petit que N_0, nous augmentons $t(n,l)$ en remplacant σ par $\sigma/\log(N_0 - l + 10)$. Il est à noter que le choix de la fonction logarithmique en base 10 a été effectué dans le seul but de faire varier lentement σ, pour une variation rapide de l. Par conséquent, en ne tenant compte que de la longueur du segment focalisé $S_{jm\sigma_f}$,

nous aurons l'expression suivante de $t(n,l)$:

$$t(n,l) = \begin{cases} \exp(-\frac{R_{jm}^2}{A\log^2(N_0-l+10)}) & \text{si } l \leq N_0 \\ \exp(-\frac{\log^2(l-N_0+10)R_{jm}^2}{A}) & \text{autrement} \end{cases} \quad (1.13)$$

L'expression finale de $t(n,l)$, celle qui tient compte des changements de n et de l, sera la suivante:

$$t(n,l) = \begin{cases} \exp(-\frac{R_{jm}^2}{A\log^2(N_0-l+10)\ln^2(1+n)}) & \text{si } l \leq N_0 \\ \exp(-\frac{\log^2(l-N_0+10)R_{jm}^2}{A\ln^2(1+n)}) & \text{autrement} \end{cases} \quad (1.14)$$

où, encore une fois, $A = 0.3607$ est une constante et N_0 dénote la longueur en pixels au dessus de laquelle un segment aura la probabilité non négligeable d'être un vrai segment.

En résumé, nous proposons la méthode suivante pour achever la classification de chacun des segments focalisés $\{S_{jm\sigma_f}\}$.

Etape 1: Calculer le rapport R_{jm}, associé au segment $S_{jm\sigma_f}$ traité.

Etape 2: Calculer le seuil $t(N_{j\sigma_f}, L_{jm})$ correspondant.

Etape 3: Comparer R_{jm} à $t(N_{j\sigma_f}, L_{jm})$ et choisir la bonne étiquette pour $S_{jm\sigma_f}$ selon le test suivant:

- si $R_{jm} \leq t(N_{j\sigma_f}, L_{jm})$ alors étiquette$=L_{faux}$
- si $R_{jm} > t(N_{j\sigma_f}, L_{jm})$ alors étiquette$=L_{vrai}$

Nous présentons à la Figure 1.7 des résultats typiques de la méthode de classification d'arêtes proposée. L'image 3-D originale est synthétique et additionnée de bruit Gaussien (SNR = 4.5 dB). Elle est décrite avec plus de détails dans la section 1.6 traitant les résultats expérimentaux. Les Figures 1.7.a et 1.7.b illustrent les segments initiaux $\{S_{j\sigma_g}\}$ et focalisés $\{S_{jm\sigma_f}\}$, respectivement. Les vraies arêtes, retenus après classification, sont montrés à la Figure 1.7.c. Il est à noter que seuls les arêtes d'ordre 0 sont significatives, et par conséquent détectables

par le détecteur d'arêtes proposé. L'unique arête d'ordre 1 de l'image a été noyée dans le bruit additionnel. La méthode produit de bons résultats: tous les segments correspondant à la frontière de l'objet ont été bien identifiés. Elle échoue cependant dans la classification de certaines fausses arêtes, induites par le bruit additif. Cet échec est dû principalement au fait que ces segments se sont comportés plus comme des vraies arêtes que comme des fausses arêtes (c'est-à-dire qu'ils ne se subdivisent pas au niveau de résolution le plus fin).

Les arêtes d'ordre 0 détectés et classifiés par notre approche sont comparés, à la Figure 1.8, à ceux obtenus avec les méthodes de Bergholm [12] et Sjöberg [21], adaptées aux images 3-D. L'image 3-D originale correspond à celle montrée à la Figure 1.12.a de la prochaine section. Les valeurs des paramètres de focalisation, c'est-à-dire le niveau de résolution initial σ_g, le pas de focalisation $\Delta\sigma$, et le niveau de résolution final σ_f, ont été fixées à 2.2, 0.25, et 0.45, respectivement. La procédure de classification de Sjöberg [21] utilise une fenêtre de recherche de taille 15×15, jugée apte à produire les meilleurs résultats. La comparaison faite entre les deux approches montre qu'une plus grande proportion de détails non significatifs ont été récoltés par la technique de Bergholm et que la performance de notre technique de classification d'arêtes est supérieure, particulièrement lorsqu'il s'agit de parties complexes de l'image.

1.6 Résultats expérimentaux.

Des résultats expérimentaux ont été obtenus pour les images synthétiques aussi bien que pour les images réelles. Dans le cas des images synthétiques, différents rapports signal-à-bruit sont utilisés afin d'illustrer la robustesse au bruit de la méthode proposée. La performance de celle-ci est illustrée par les résultats obtenus à partir d'images réelles de niveau de complexité très diversifié, depuis

les images à objets isolés jusqu'à celles contenant des pièces multiples avec recouvrements.

La taille des images est de 256 x 256 pixels. Comme la taille des opérateurs de convolution est de 8σ, le niveau de résolution le plus fin que nous puissions utiliser pour localiser les arêtes avec la plus grande précision est $\sigma = 0.45$ et la plus petite valeur de $\Delta\sigma$ (paramètre denotant la variation de σ entre deux étapes de focalisation consécutives) qui correspond à une diminution de un pixel de la taille de l'opérateur à chaque étape est 0.25. L'utilisation de cette dernière valeur est justifiée par le désir d'accentuer la tendance à la subdivision des arêtes non significatifs. Le choix de la valeur de σ_c est arbitraire.

1.6.1 Cas des images 3-D synthétiques.

L'image synthétique d'un objet polyédrique est montrée à la Figure 1.9.a. La frontière qui délimite l'objet correspond à une discontinuité de profondeur d'amplitude 70 mm (par rapport à un fond à 0 mm) et les deux surfaces les plus hautes de l'objet donnent naissance à une discontinuité d'orientation de hauteur 100 mm. Un bruit gaussien de moyenne nulle et d'écart-type 0.25, 3.5, ou 35.0 mm est ajouté au signal de l'image (voir Figure 1.9.b-d). Ceci correspond respectivement à des rapports signal-à-bruit de valeur 26.0 dB, 14.5 dB, et 4.5 dB.

En premier lieu, les arêtes potentielles sont détectées à un niveau de résolution $\sigma_g = 1.7$. Elles sont soumises par la suite à une opération de focalisation et finalement à l'étape de classification pour produire des arêtes significatives et bien localisés. Les résultats sont montrés dans la Figure 1.10.a-d. Nous pouvons en déduire les observations suivantes:

- Dans le cas des arêtes correspondant à des discontinuités de profondeur, l'algorithme d'extraction d'arêtes est très robuste face à la présence du bruit dans l'image, même si le rapport signal-à-bruit est aussi bas que 4.5 dB (les

images 3-D réelles sont de qualité bien supérieure en général).

- En général, les arêtes en provenance de discontinuités de profondeur sont bien détectés et localisés, à l'exception du cas où le rapport SNR = 4.5 dB. Dans ce cas ci, les arêtes sont subdivisées en éléments plus petits, certains points d'arête sont manquants, et certains autres sont mal localisés. L'arête reste cependant "visible" au détecteur.

- Dans le cas des arêtes d'ordre 1 l'algorithme se montre plus sensible au bruit. Ceci est dû à l'emploi de détecteurs aux dérivées secondes. A 26 dB, les résultats sont bons bien que les arêtes commencent à zigzaguer comparés à ceux extraits de l'image originale (il est à noter que dans notre exemple les discontinuités d'orientation ne sont pas très prononcées). Les arêtes se perdent dans le bruit dans les cas SNR = 14.5 dB et SNR = 4.5 dB.

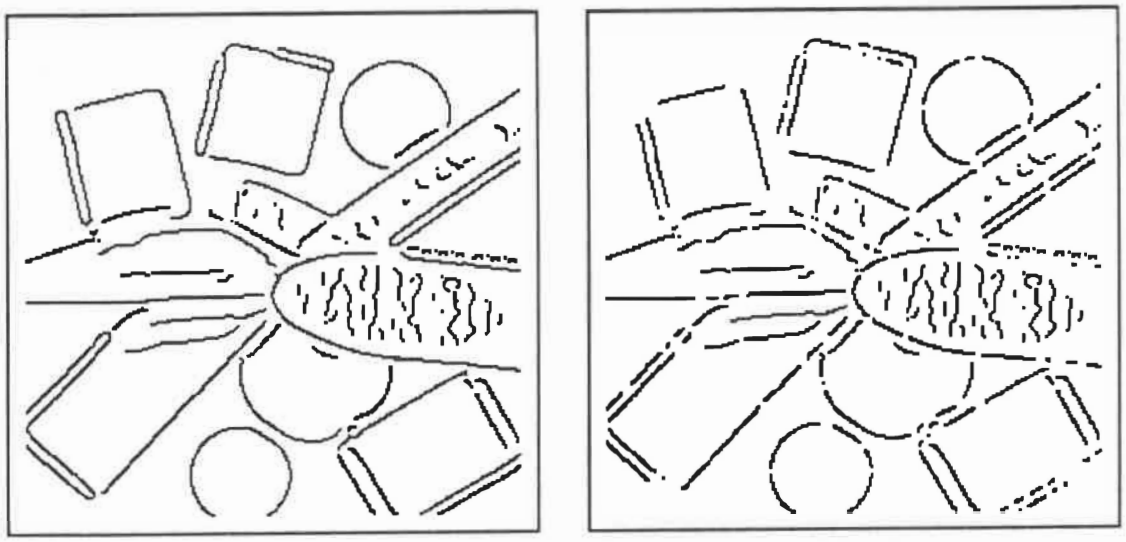
1.6.2 Cas des images 3-D réelles.

Les images 3-D réelles ont été fournies par la Division de Génie Electrique du Conseil National de Recherche du Canada. Les données sont à l'origine produites sous la forme d'une séquence de triplets (x, y, z), à l'aide d'un capteur laser basé sur la triangulation et l'autosynchronisation. Les coordonnées x et y représentent la grille d'échantillonnage originale qui correspond à la géométrie du mécanisme de balayage et qui se distinguent en général des coordonnées cartésiennes du plan image. A la valeur z est associée l'élévation du point observé de l'objet. Comme ces données subissent des distortions géométriques à l'affichage, une interpolation linéaire leur est appliquée par la suite, produisant ainsi les données contenues dans les images 3-D utilisées dans cette section.

Les Figures 1.11 et 1.12 montrent les versions codées en niveaux de gris des images 3-D réelles utilisées. Bien que ces images soient de qualité exceptionnelle

au point de vue bruit, elles souffrent d'artifices causés par des effets d'ombre. Ceci a pour conséquence l'introduction d'ambiguités dans l'interprétation subséquente des arêtes détectées, aux endroits atteints par le rayon laser mais cachés du capteur ou aux endroits observés par le capteur mais non accessibles au rayon laser. Nous proposons une méthode simple et efficace qui permet l'élimination des arêtes engendrés par les effets d'ombre. En effet, comme la valeur zéro est arbitrairement attribuée aux pixels des régions d'ombre, la méthode consiste tout simplement à extraire de la carte d'arêtes obtenue, ceux qui correspondent aux bordures de ces régions et à éliminer par la suite, pour chacune des bordures extraites, la partie qui ne fait pas partie de l'objet.

Les arêtes sont extraites au niveau de résolution $\sigma_c = 2.2$ et sont montrées dans les Figures 1.13 - 1.18. Les résultats illustrent bien la performance de l'algorithme, particulièrement dans le cas où les objets se touchent ou se chevauchent (Figure 1.16.d and Figure 1.17.d). Les arêtes extraites correspondant à des discontinuités d'orientation sont plus sensibles au bruit que ceux en provenance des discontinuités de profondeur; leur forme devient rapidement instable lorsque le rapport signal-à-bruit diminue (ce fait peut être observé en comparant les arêtes dues à des discontinuités d'orientation des Figures 1.14 et 1.18). D'autre part, le degré d'instabilité de ces arêtes dépend fortement de leur acuité (un exemple de cette propriété est illustré à la Figure 1.14.b par comparaison des arêtes du polyèdre à ceux du cylindre vertical). La méthode mentionnée précédemment pour l'élimination des effets d'ombrage s'est montrée efficace à cet égard car les arêtes dues à ces effets sont bien identifiés et éliminés. Cependant, celles conservées comme arêtes d'objets ne correspondent pas nécessairement aux contours physiques des surfaces. Ceci peut être vu dans la Figure 1.14.d.



(a) (b)

Figure 1.3 Segmentation d'arêtes potentielles. (a) Image d'arêtes à segmenter. (b) Image de segments qui résultent de la segmentation.

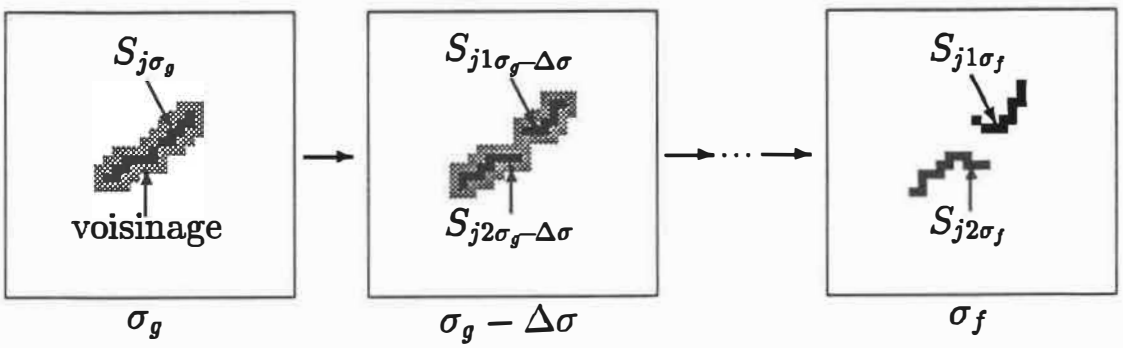


Figure 1.4 Principe de la méthode de focalisation proposée.

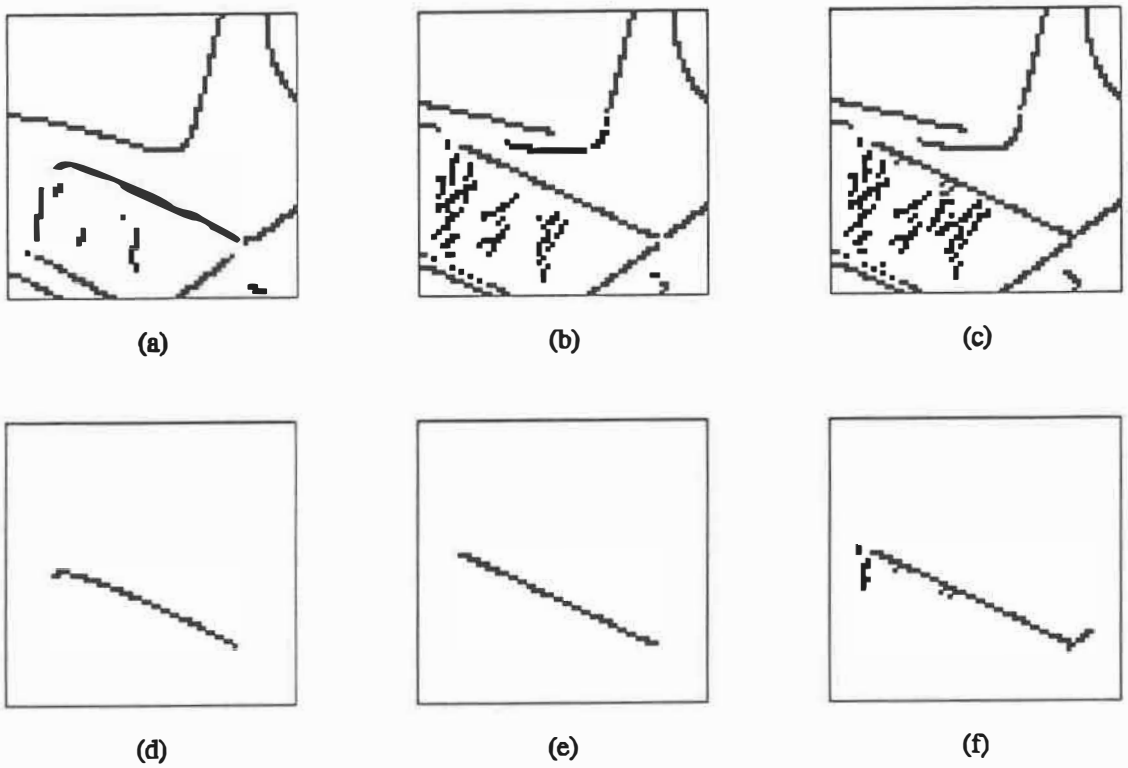


Figure 1.5 Resultats de la focalisation d'arêtes ($\sigma_g = 2.2$ and $\sigma_f = 0.45$). (a) Segments d'arête initiaux. (b) Resultats obtenus avec notre approche. (c) Resultats obtenus avec l'approche de Bergholm. (d) Segment type, détecté à σ_g. (e) Resultats obtenus avec notre approche. (f) Resultats obtenus avec l'approche de Bergholm.

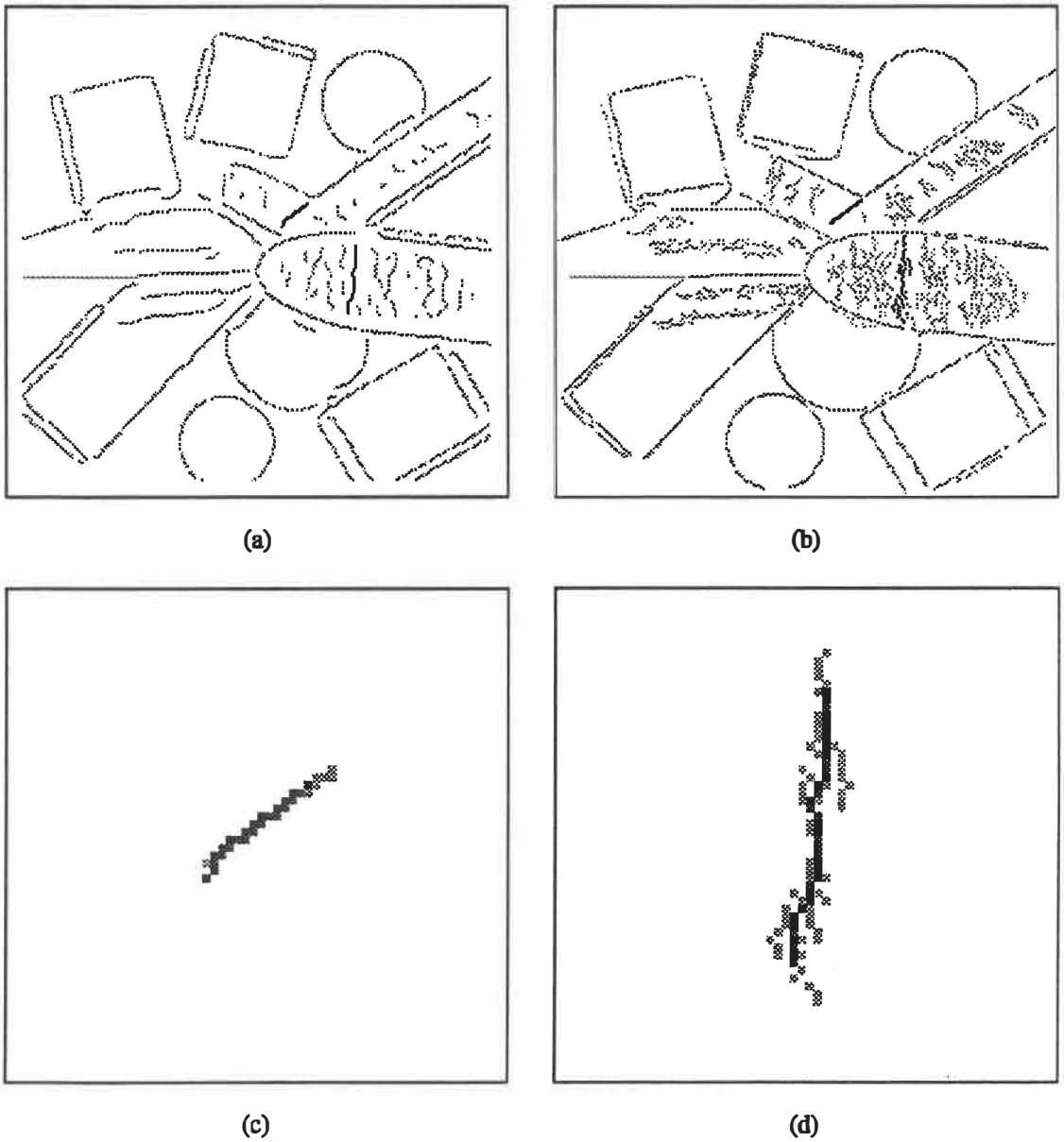


Figure 1.6 Comportements caractéristiques des arêtes. (a) Image d'arêtes à $\sigma_g = 2.2$. (b) Image de segments focalisés à $\sigma_f = 0.45$. (c) Comportement caractéristique d'une vraie arête. (d) Comportement caractéristique d'une fausse arête.

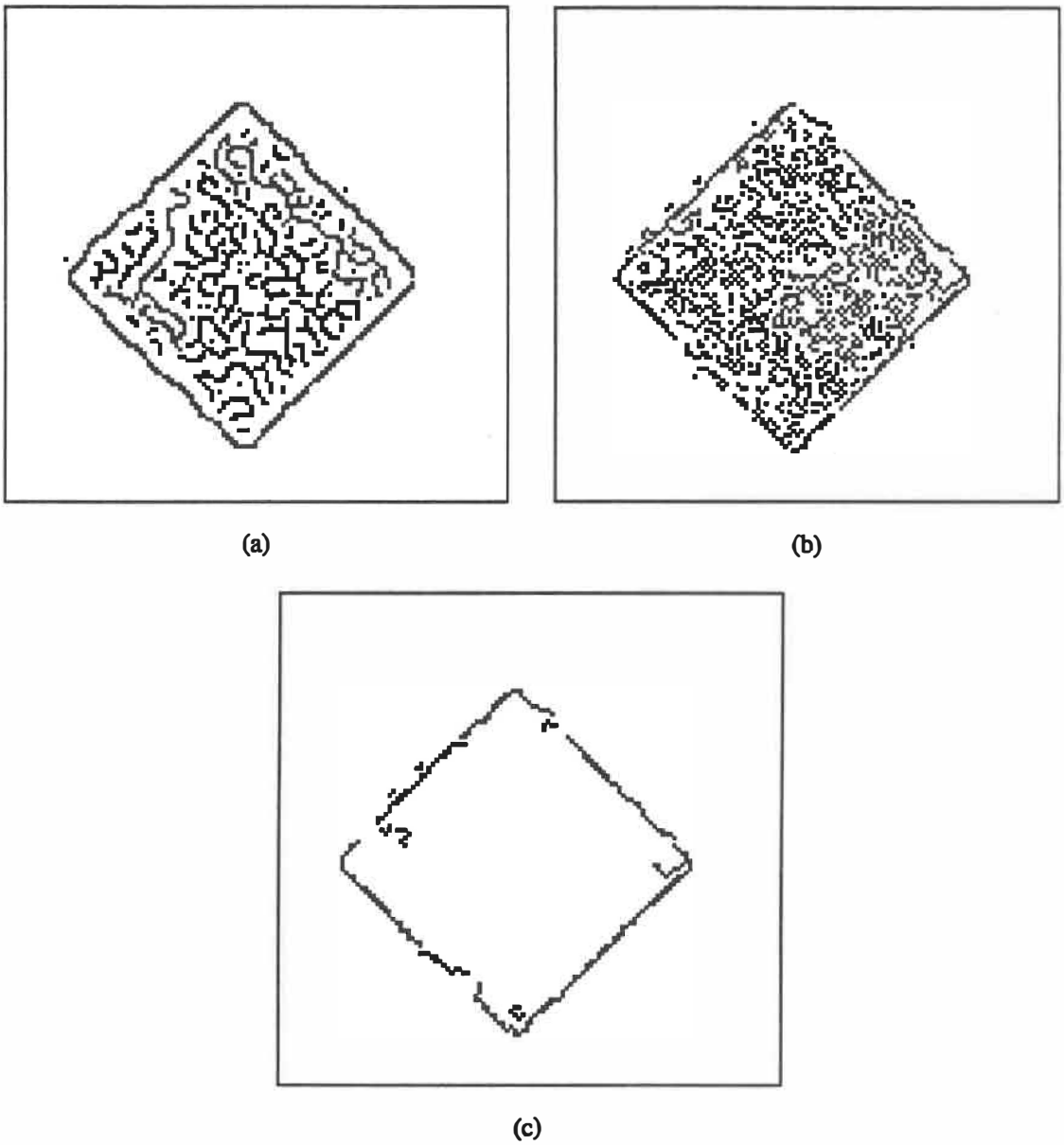


Figure 1.7 Exemple de résultats produits par la méthode de classification d'arêtes proposée. (a) Segments détectés au niveau de résolution initial $\sigma_g = 1.7$. (b) Segments focalisés, obtenus au niveau de résolution final $\sigma_f = 0.45$. (c) Segments retenus après classification par la méthode proposée.

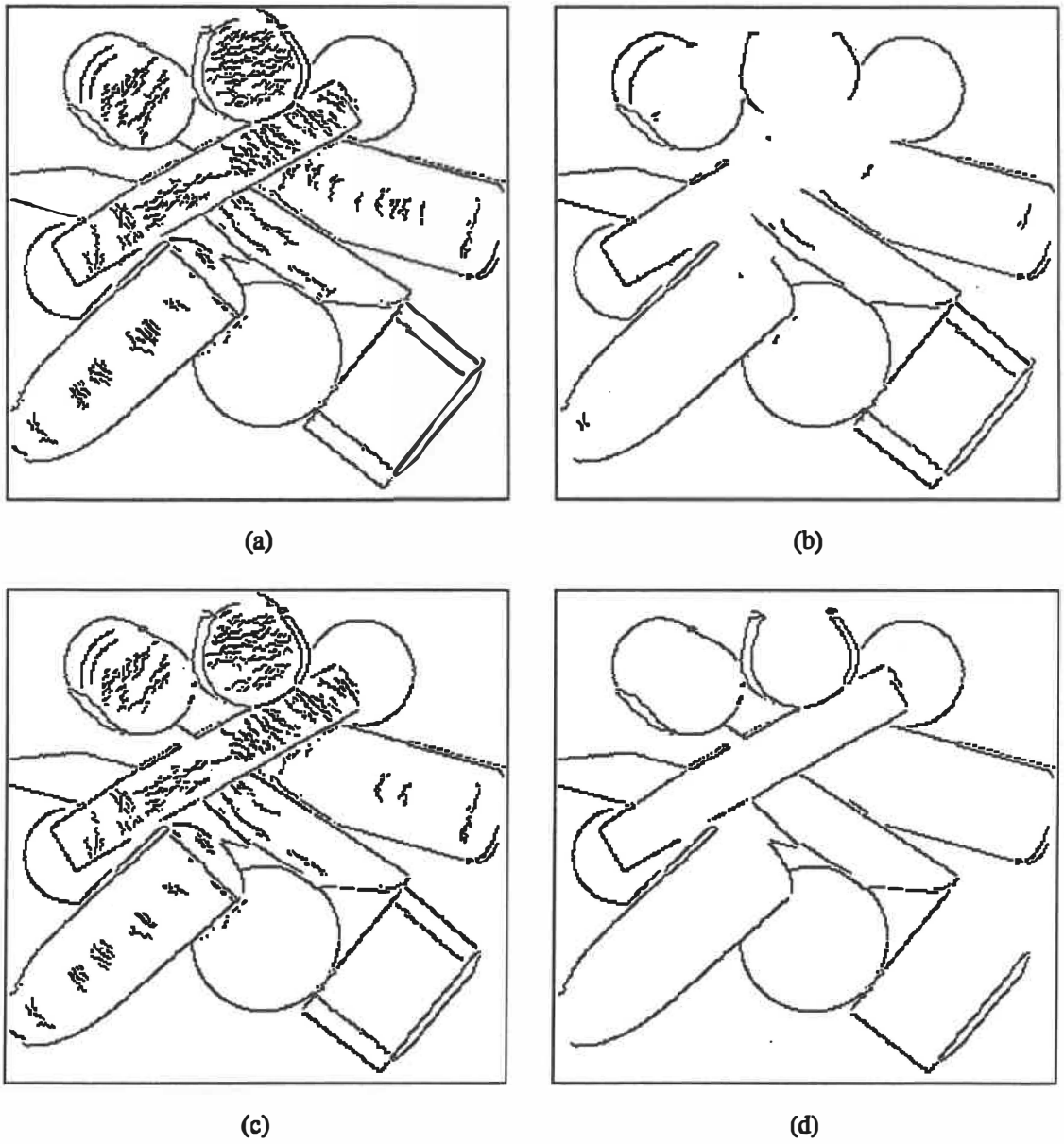


Figure 1.8 Comparaison de résultats produits par notre méthode et par celle de Bergholm et Sjöberg. (a) Résultats obtenus par la méthode de focalisation de Bergholm. (b) Arêtes retenues après la classification de Sjöberg. (c) Résultats obtenus par la méthode de focalisation proposée. (d) Arêtes retenues après notre classification.

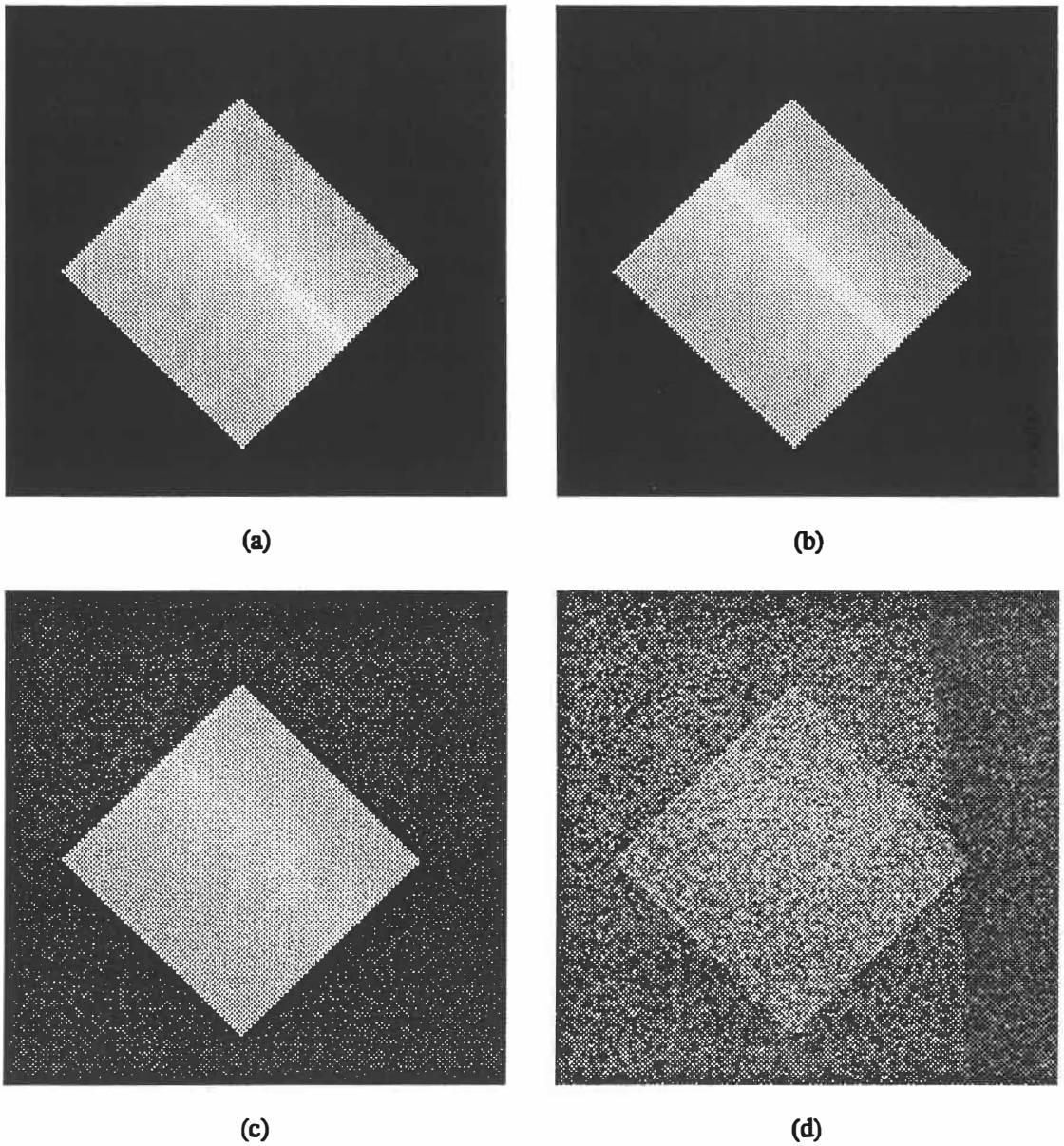


Figure 1.9 Image synthétique d'un objet polyédrique. (a) Image originale. (b)-(d) Images bruitées avec rapports signal-à-bruit respectifs de 26.0 dB, 14.5 dB, et 4.5 dB.

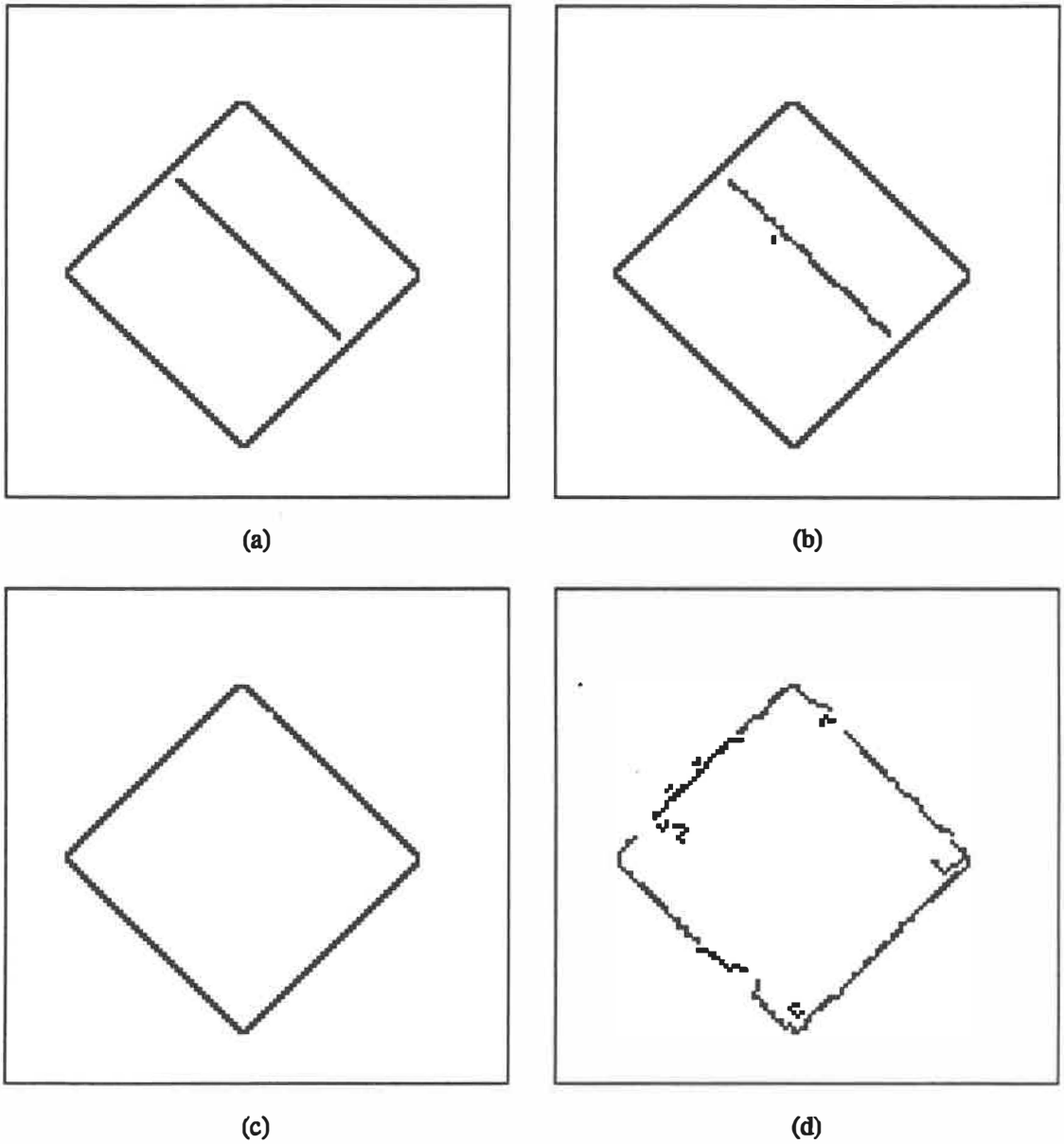


Figure 1.10 Resultats obtenus à partir des images synthétiques montrées à la Figure 1.9. (a) Image originale. (b) SNR = 26.0 dB. (c) SNR = 14.5 dB. (d) SNR = 4.5 dB.

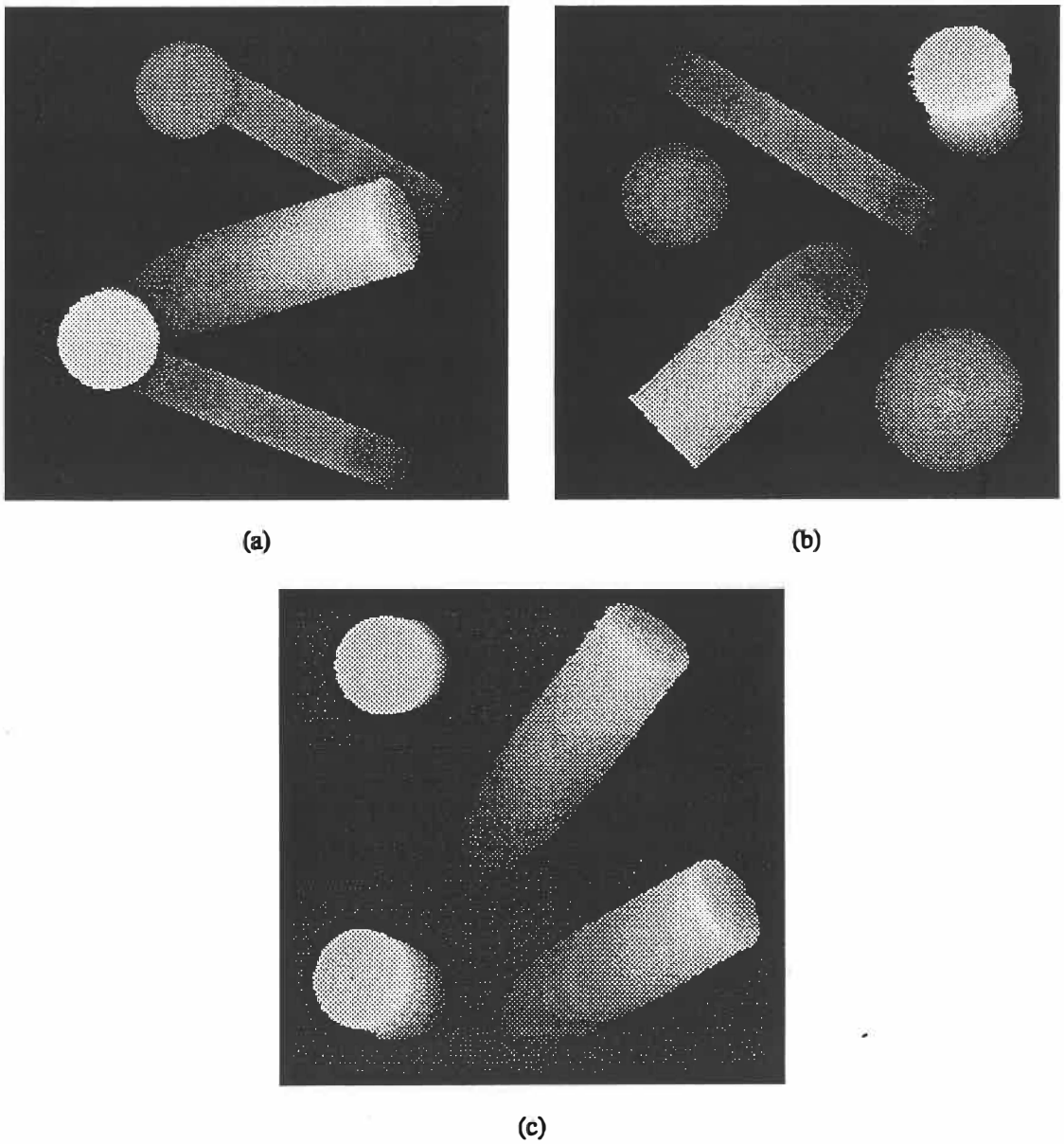
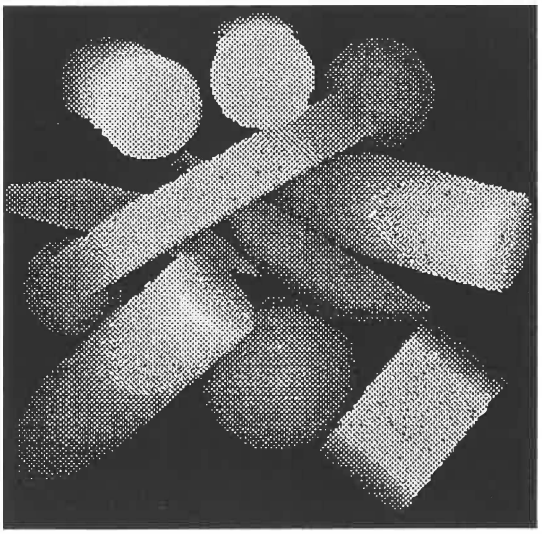
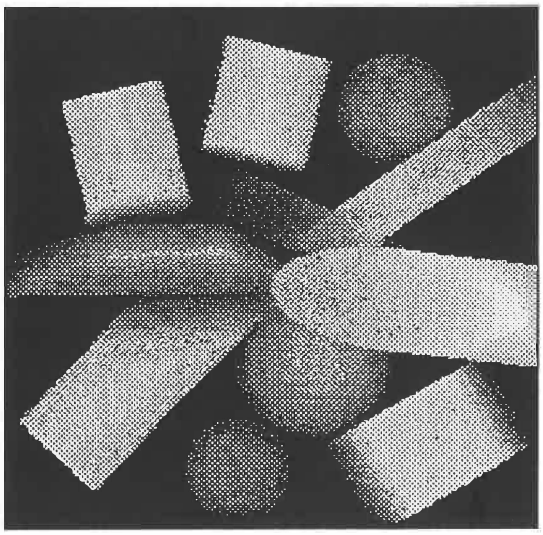


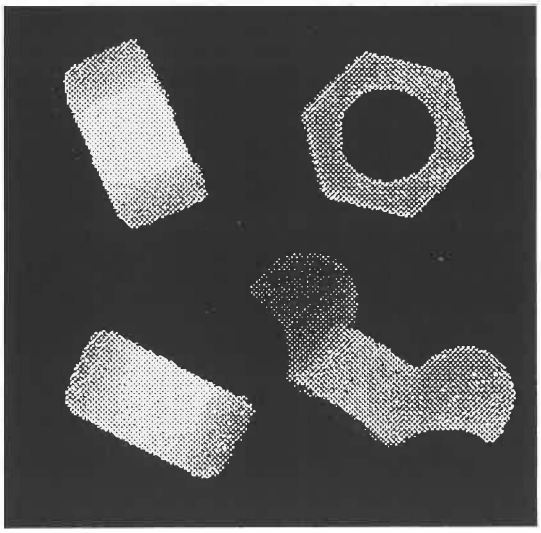
Figure 1.11 Images 3-D originales. (a) BLOC11. (b) BLOC9. (c) BLOC4.



(a)



(b)



(c)

Figure 1.12 Images 3-D originales. (a) BLOC15. (b) BLOC16. (c) BOLT6.

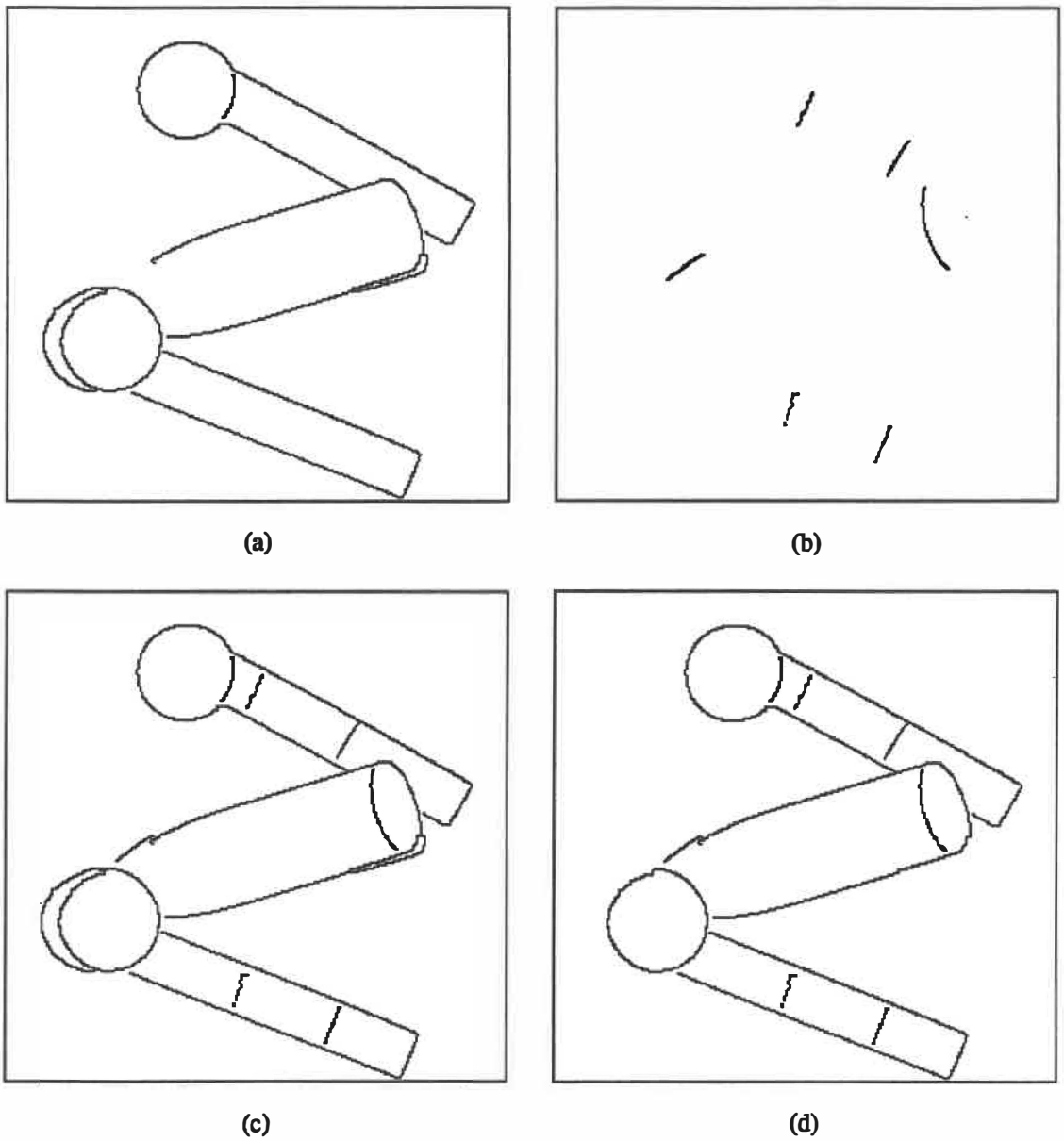


Figure 1.13 Arêtes de l'image "BLOC11". (a) Arêtes d'ordre 0. (b) Arêtes d'ordre 1. (c) Superposition des deux types d'arêtes. (d) Les mêmes résultats que dans (c) avec arêtes dues aux effets d'ombre éliminés.

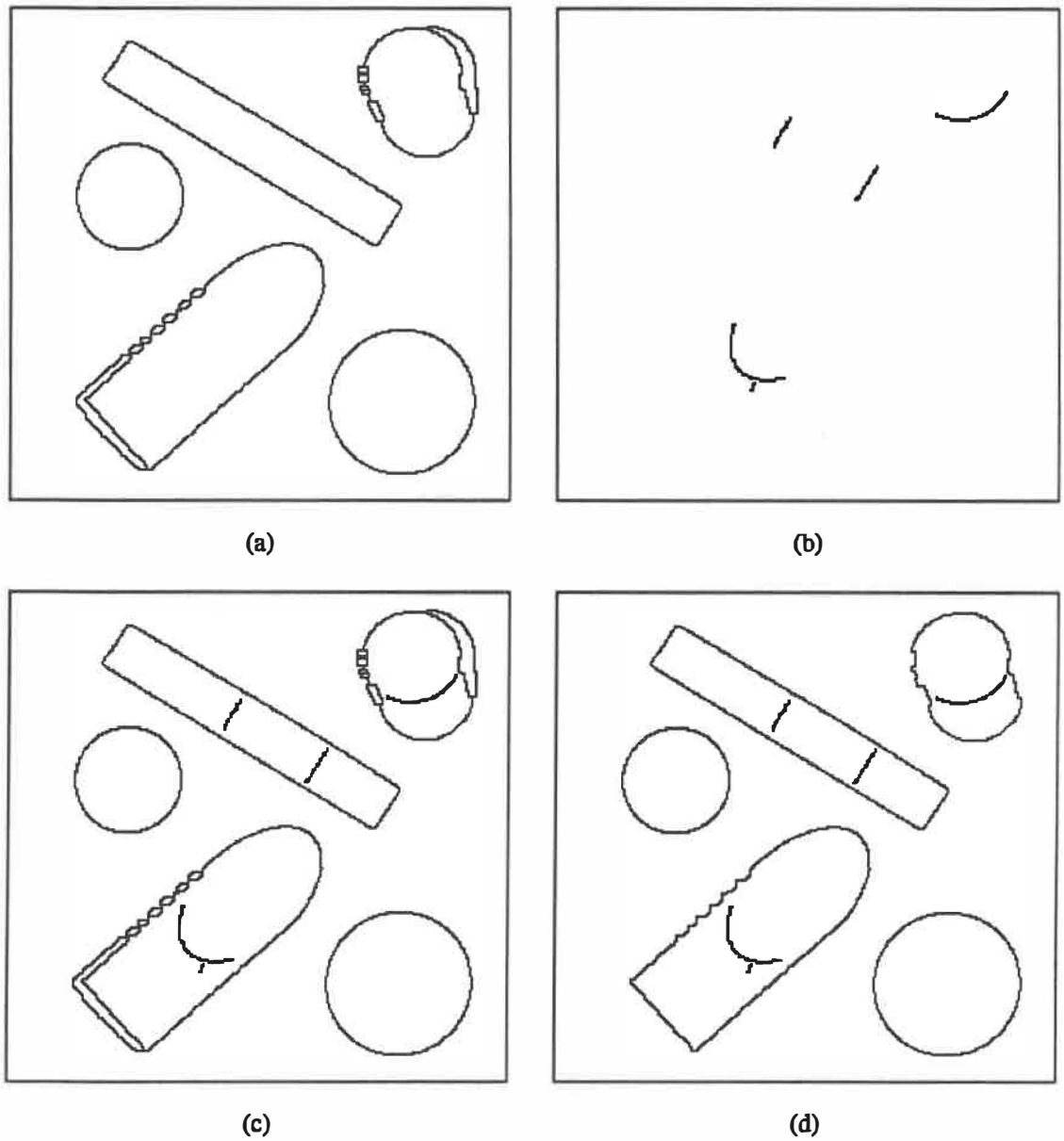


Figure 1.14 Arêtes de l'image "BLOC9". (a) Arêtes d'ordre 0. (b) Arêtes d'ordre 1. (c) Superposition des deux types d'arêtes. (d) Les mêmes résultats que dans (c) avec arêtes dues aux effets d'ombre éliminés.

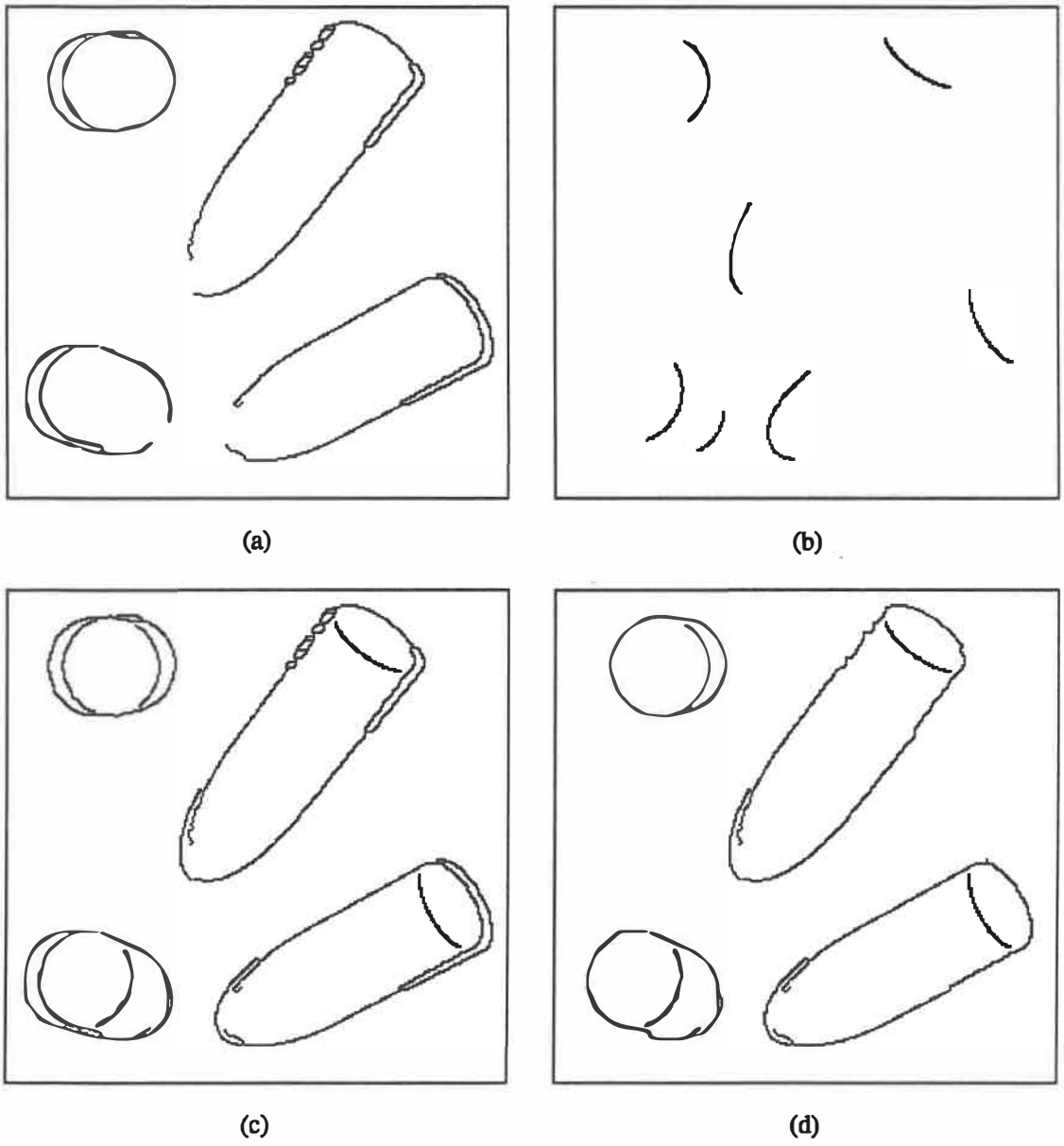


Figure 1.15 Arêtes de l'image "BLOC4". (a) Arêtes d'ordre 0. (b) Arêtes d'ordre 1. (c) Superposition des deux types d'arêtes. (d) Les mêmes résultats que dans (c) avec arêtes dues aux effets d'ombre éliminés.

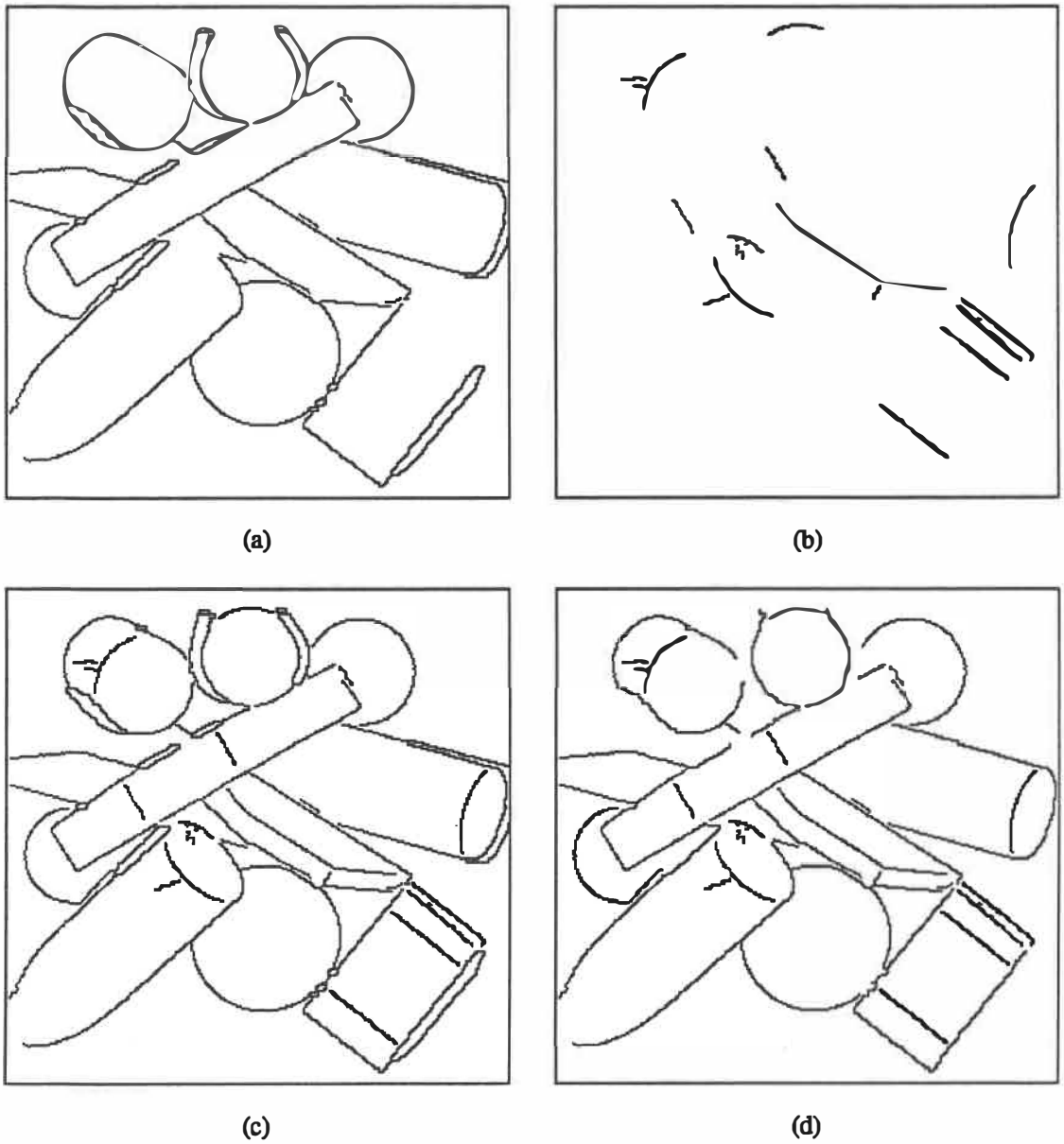


Figure 1.16 Arêtes de l'image "BLOC15". (a) Arêtes d'ordre 0. (b) Arêtes d'ordre 1. (c) Superposition des deux types d'arêtes. (d) Les mêmes résultats que dans (c) avec arêtes dues aux effets d'ombre éliminés.

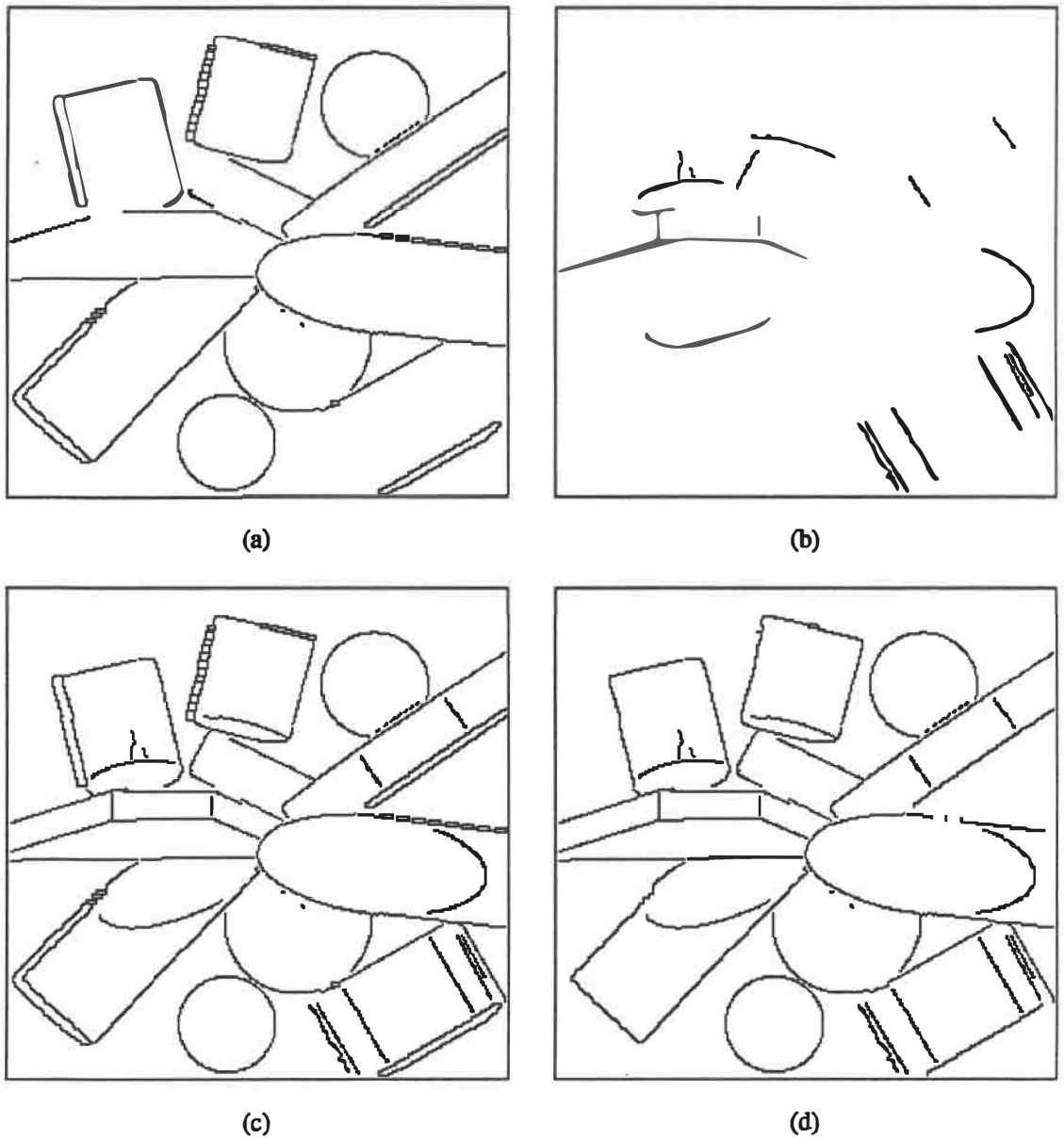


Figure 1.17 Arêtes de l'image "BLOC16". (a) Arêtes d'ordre 0. (b) Arêtes d'ordre 1. (c) Superposition des deux types d'arêtes. (d) Les mêmes résultats que dans (c) avec arêtes dues aux effets d'ombre éliminés.

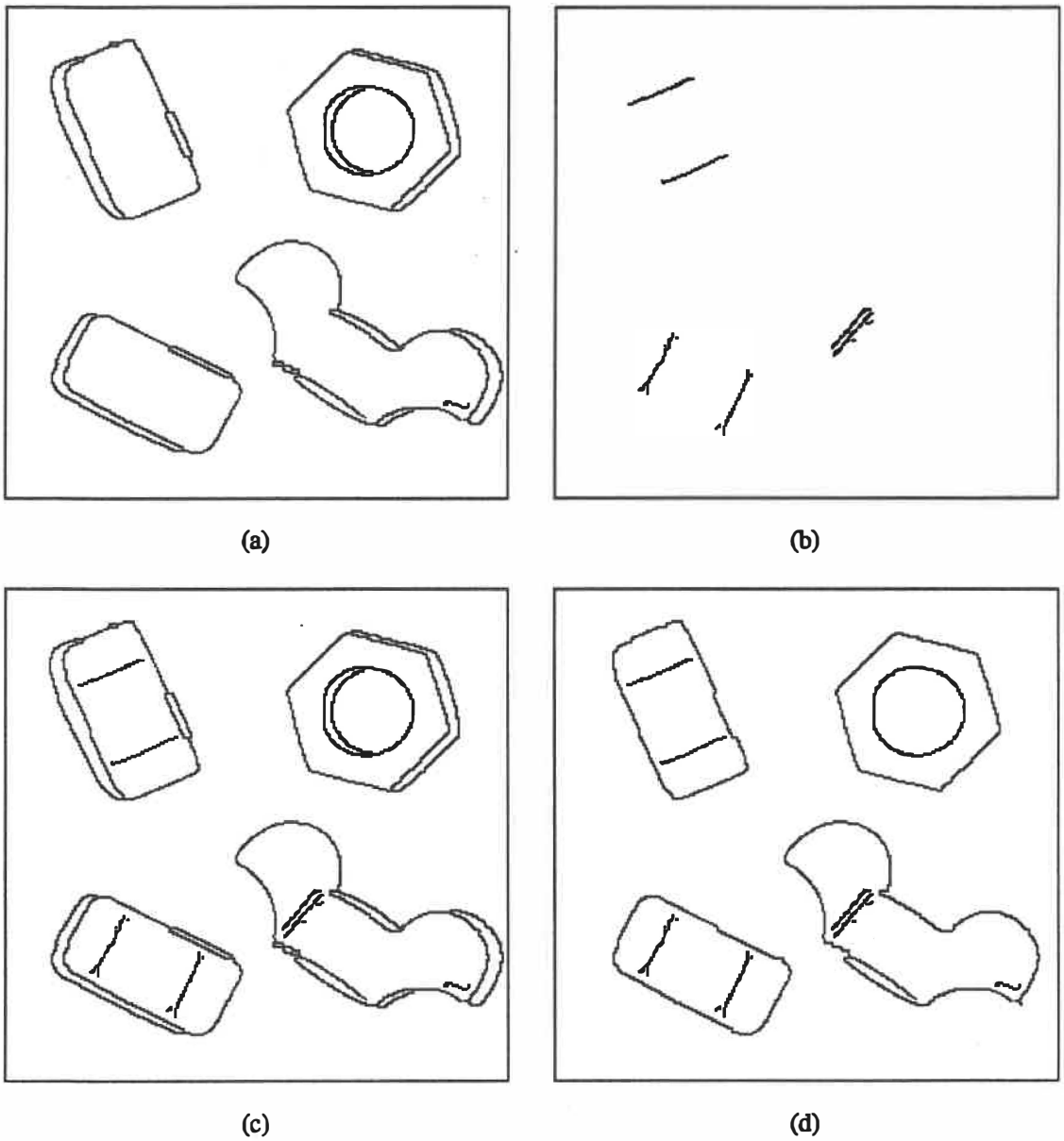


Figure 1.18 Arêtes de l'image "BOLT6". (a) Arêtes d'ordre 0. (b) Arêtes d'ordre 1. (c) Superposition des deux types d'arêtes. (d) Les mêmes résultats que dans (c) avec arêtes dues aux effets d'ombre éliminés.

CHAPITRE 2

Approche multirésolution pour la segmentation d'images 3-D en primitives de surface

Le chapitre précédent présentait une approche visant l'extraction à différentes échelles de l'information sur les discontinuités de profondeur et d'orientation de surface. Ce chapitre propose une méthode permettant l'extraction à différents niveaux de résolution de l'information sur les propriétés géométriques des surfaces à l'intérieur des régions délimitées par les arêtes correspondantes, détectées auparavant.

Similairement au chapitre précédent, le présent chapitre commence par une description générale du principe de segmentation des images tridimensionnelles en primitives de surface et par une revue non détaillée des méthodes existantes. Cette première section justifie également le choix des courbures Gaussienne K et moyenne H comme caractéristiques de surface, utiles pour la description en termes de primitives de surface des images tridimensionnelles. Les détails concernant chacune des étapes de la méthode proposée sont par la suite donnés dans les sections qui suivent.

2.1 Segmentation d'images 3-D: principes et revue bibliographique.

2.1.1 Principes.

Un bon système de reconnaissance d'objets tridimensionnels doit posséder les caractéristiques suivantes: 1) le système doit être capable de traiter des données

captées à partir d'un point d'observation arbitraire, et de fournir un résultat indépendant du point de vue; 2) le système doit être en mesure de reconnaître des objets naturels, d'un niveau de complexité structurale quelconque, sans être restreint à des classes d'objects trop limitées (objets polyédriques ou à surfaces courbes); 3) le système doit être capable de manipuler un nombre assez large d'objets orientés et localisés de façon arbitraire, en demeurant robuste aux problèmes causés par les occlusions. Un tel système est en général basé sur une procédure de caractérisation de surfaces, dont la qualité descriptive influence de façon primordiale la performance du système. Les caractéristiques de surface extraites par cette procédure doivent être *invariantes* par rapport à l'attitude, de telle sorte qu'un changement d'orientation de la scène observée ne modifie pas de manière radicale la description finale. Elles doivent aussi être *complètes* pour pouvoir décrire efficacement les différents types de surface et permettre l'identification d'objets similaires. Finalement, elles doivent être *locales* pour que la reconnaissance d'objets partiellement visibles soit possible.

Il est à noter que la notion d'invariance fait référence à une quantité qui ne varie pas lorsque soumise à un groupe de transformations tridimensionnelles donné (e.g. les translations, les rotations, etc...). En plus, cette notion d'invariance est intimement reliée à la notion de visibilité. En effet, il est assez courant dans la nature qu'un objet, vu à partir de deux points d'observation différents, peut paraître sous deux aspects différents (tel est le cas, par exemple, d'un cylindre qui se présente comme une surface plane munie d'une frontière circulaire lorsque le point d'observation se situe sur le prolongement de son axe, ou comme une surface courbe au frontière rectangulaire lorsque observé dans une direction perpendiculaire à son axe). Dans ce cas dégénéré, il n'existe aucune caractéristique qui soit commune aux deux vues de l'objet. Par conséquent, nous

définissons une caractéristique invariante de surface comme étant une particularité d'une région de surface qui reste inchangée vis-à-vis des changements de point d'observation qui n'affectent pas la visibilité de la région.

La courbure Gaussienne K et la courbure moyenne H, constituent des caractéristiques invariantes de surface qui sont complètes et locales. La géométrie différentielle ([8], [9]) stipule que la forme locale d'une surface est uniquement déterminée par ses première et seconde formes fondamentales, et que la courbure Gaussienne K et la courbure moyenne H combinent, de manière intrinsèque et extrinsèque respectivement, ces formes fondamentales en des mesures scalaires qui sont locales et invariantes aux rotations, aux translations, et aux changements de représentations paramétriques des surfaces. En plus, Besl et Jain [10] ont montré que les différentes combinaisons de signes de ces courbures donnent naissance à huit primitives de surface, indépendantes de la position dans l'espace du point d'observation, formant un ensemble assez riche pour caractériser chacun des points de la scène observée. Les combinaisons possibles des signes de courbure, ainsi que les différents types de primitive qui en résultent sont indiqués dans le tableau 2.1.1. Il est à noter que K ne peut être positive lorsque H est nulle. La Figure 2.1 illustre les formes locales de ces primitives de surface (reproduction faite à partir de Besl et Jain [10]).

Soit P un point quelconque d'une surface régulière S dans $\mathbf{R}^3$,

$$S = \{\mathbf{x}(u,v)=(x(u,v),y(u,v),z(u,v)) : \quad (u,v)\in D\subseteq\mathbf{R}^2\} \tag{2.1}$$

et soit $\mathbf{x}(u,v)=[x(u,v) \quad y(u,v) \quad z(u,v)]^T$ une facette de S, contenant P. Les *première* et *seconde formes fondamentales* de $\mathbf{x}(u,v)$, dénotées par **I** et **II** respectivement, sont définies par les formes quadratiques suivantes:

$$\mathbf{I} = d\mathbf{x}.d\mathbf{x} = \begin{bmatrix} du & dv \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} du \\ dv \end{bmatrix} = d\mathbf{u}^T[g]d\mathbf{u} \tag{2.2}$$

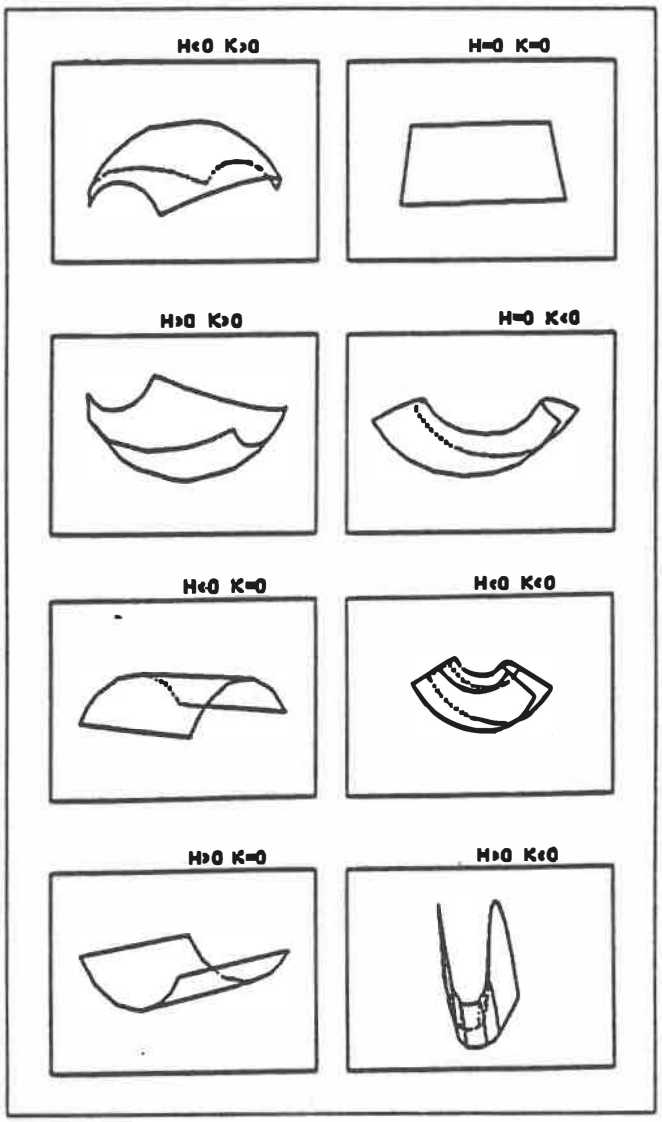


Figure 2.1 Formes locales des primitives de surface.

$$\mathbf{II} = -d\mathbf{x}.d\mathbf{n} = \begin{bmatrix} du & dv \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} du \\ dv \end{bmatrix} = d\mathbf{u}^T[b]d\mathbf{u} \tag{2.3}$$

où $\mathbf{n}=\frac{\mathbf{x}_u \times \mathbf{x}_v}{|\mathbf{x}_u \times \mathbf{x}_v|}$ est un vecteur normal unitaire au point P de S. Les éléments de la matrice $[g]$ de la première forme fondamentale et ceux de la matrice $[b]$ de la seconde forme fondamentale sont respectivement donnés par:

$$\begin{aligned} g_{11} &= E = \mathbf{x}_u.\mathbf{x}_u & b_{11} &= L = \mathbf{x}_{uu}.\mathbf{n} \\ g_{12} &= F = \mathbf{x}_u.\mathbf{x}_v = g_{21} & b_{12} &= M = \mathbf{x}_{uv}.\mathbf{n} = b_{21} \\ g_{22} &= G = \mathbf{x}_v.\mathbf{x}_v & b_{22} &= N = \mathbf{x}_{vv}.\mathbf{n} \end{aligned} \tag{2.4}$$

où les indices impliquent les dérivées partielles d'ordre 1 et d'ordre 2 de $\mathbf{x}(u,v)$

$$\begin{aligned} \mathbf{x}_u &= \frac{\partial \mathbf{x}}{\partial u} & \mathbf{x}_{uv} &= \frac{\partial^2 \mathbf{x}}{\partial u \partial v} & \mathbf{x}_{uu} &= \frac{\partial^2 \mathbf{x}}{\partial u \partial u} \\ \mathbf{x}_v &= \frac{\partial \mathbf{x}}{\partial v} & \mathbf{x}_{vu} &= \frac{\partial^2 \mathbf{x}}{\partial v \partial u} & \mathbf{x}_{vv} &= \frac{\partial^2 \mathbf{x}}{\partial v \partial u} \end{aligned} \tag{2.5}$$

A partir de ces matrices $[g]$ et $[b]$ des formes fondamentales, la courbure Gaussienne K et la courbure moyenne H sont définies de la manière suivante:

$$K = det([g^{-1}][b]) = \frac{LN - M^2}{EG - F^2} \tag{2.6}$$

$$H = \frac{1}{2} tr([g^{-1}][b]) = \frac{EN + GL - 2FM}{2(EG - F^2)} \tag{2.7}$$

Cependant, les surfaces dans les images 3-D sont habituellement décrites sous forme de surfaces graphes dont la représentation paramétrique est donnée sous la forme $\mathbf{x}(x,y)=[x \;\; y \;\; z(x,y)]^T$. Cette paramétrisation des surfaces permet la simplification des expressions de courbure (2.6) et (2.7) pour en arriver à:

$$K = \frac{z_{xx} z_{yy} - {z_{xy}}^2}{(1 + {z_x}^2 + {z_y}^2)^2} \tag{2.8}$$

$$H = \frac{z_{xx}(1 + {z_y}^2) + z_{yy}(1 + {z_x}^2) - 2 z_x z_y z_{xy}}{2(1 + {z_x}^2 + {z_y}^2)^{3/2}} \tag{2.9}$$

A partir des équations (2.8) et (2.9), il est évident que les valeurs estimées des dérivées partielles d'ordre 1 et d'ordre 2 de la carte de profondeur (image

3-D) sont nécessaires au calcul de courbures. Ceci suggère à son tour un filtrage de l'image avant que l'opération de dérivation y soit appliquée. Encore une fois, il s'agit de la nécessité de régulariser le problème de différentiation numérique afin de le rendre plus stable vis-à-vis du bruit. D'ailleurs, l'emploi du filtrage est aussi justifié par le fait que les structures de la scène peuvent être plus ou moins significatives suivant le niveau de résolution choisi. En effet, rappelons-nous que la scène se compose de structures grossières de large taille sur lesquelles se superposent des structures de plus en plus détaillées. Les changements physiques qui se produisent à différents niveaux de résolution fixent en général les limites de ces différentes structures. A une échelle donnée, un changement brusque de profondeur peut ne plus être significatif, bienqu'il le soit à la résolution précédente (niveau de résolution plus fin), et par conséquent disparait de la carte d'arêtes. Ceci implique à son tour le fusionnement à l'échelle donnée des deux régions adjacentes correspondantes en une région unique dont la structure intérieure est rendue plus grossière.

2.1.2 Revue bibliographique.

Différentes primitives de surface ont été proposées dans la littérature pour la description de surface.

1. *Facettes* comme primitives de surface ([36], [37], [38], [39], [40]).

   - les *facettes planes* ont été employées afin d'approximer de façon très grossière les surfaces courbes. Leur emploi est justifié par la simplicité de l'implantation pratique.
   - les *facettes quadriques* offrent une meilleure approximation et ceci pour une gamme plus grande de surfaces. L'utilisation de ces facettes se limite cependant à la représentation des surfaces possèdant un certain degré de

symétrie structurale. Aussi, ces facettes sont reconnues comme sensibles au bruit.

- les *facettes cubiques* approximent de manière plus précise les surfaces et étendent encore plus la gamme de surfaces qu'elles permettent de décrire. Malheureusement, elles ne sont utilisées que pour effectuer la classification des surfaces en des catégories plan, cylindre, et surface courbe.

2. *Arêtes* comme primitives de surface ([6], [7], [5], [41]): bien qu'elles soient populaires, leur défaut majeur réside dans le fait qu'elles ne permettent pas la description des surfaces intérieures.

3. *Courbes surfaciques* ([4]): les lignes de courbure, les lignes asymptôtiques, les courbes paraboliques, les contours et les intersections de surfaces constituent un ensemble de primitives qui permet la description des surfaces lisses. Certaines contraintes telles que la planarité des courbes, l'invariance de certaines courbures,...sont imposées, mais aucune n'est explicitement reliée à des significations perceptuelles. Aucune discussion n'a été faite sur l'utilisation possible de ces courbes intrinsèques dans la description des surfaces plus générales et plus complexes. L'approche proposée ne s'applique qu'aux surfaces de révolution, surfaces développées, et surfaces réglées.

Une approche différente pour la description des surfaces consiste à les approximer par des équations polynômiales afin d'en déduire des caractéristiques de surface utiles. Il est à noter cependant que les coefficients de ces équations varient en fonction des rotations, des translations, des changements de forme, du degré de visibilité des surfaces, et du niveau de bruit dans l'image. Ceci fait échouer toute tentative d'effectuer la reconnaissance par simple comparaison des

coefficients. Par contre, ceux qui sont non nuls peuvent être utilisés pour caractériser les surfaces. C'est dans cet ordre d'idée que Jain, Sripradisvarakul, et O'Brien [?] ont proposé leur approche. Les caractéristiques de surface suggérées consistent en les valeurs propres λ_1 et λ_2 de la matrice quadratique, associée aux termes de deuxième ordre d'une biquadratique, et la classification des points de surface d'une image 3-D est formulée de la manière suivante, en supposant que $\lambda_1 \geq \lambda_2$

$$\begin{aligned}
&\lambda_1 > 0 \text{ et } \lambda_2 > 0 : \text{ Paraboloïde élliptique convexe} \\
&\lambda_1 < 0 \text{ et } \lambda_2 < 0 : \text{ Paraboloïde élliptique concave} \\
&\lambda_1 > 0 \text{ et } \lambda_2 = 0 : \text{ Surface cylindrique convexe} \\
&\lambda_1 = 0 \text{ et } \lambda_2 < 0 : \text{ Surface cylindrique concave} \\
&\lambda_1 > 0 \text{ et } \lambda_2 < 0 : \text{ Paraboloïde hyperbolique}
\end{aligned} \tag{2.10}$$

Besl et Jain [10] [11] ont proposé K et H comme caractéristiques de surface. Ils ont adopté le filtrage uniforme et l'approximation de surface, utilisant des fenêtres d'approximation de large taille, dans le calcul des dérivées partielles. Cette technique estompe malheureusement les discontinuités de surface et par conséquent alterne la forme de l'objet original. En donnant comme argument que K et H ne sont pas définies aux endroits où la surface exhibe des variations rapides (la géométrie différentielle n'est définie que pour les surfaces continues), ils ont effectué une érosion sévère sur les surfaces intermédiaires pour en extraire des régions semence, utiles à la segmentation subséquente par croissance des régions. Il est à noter que l'ensemble des primitives de surface engendré par K et H est plus riche que celui issu des valeurs propres λ_1 et λ_2 (Jain et al. [?]). En effet, ce dernier ne permet pas la caractérisation des surfaces planes et englobe les trois types de surface en selle en un seul type de surface (paraboloïde hyperbolique).

2.1.3 Principe de l'approche proposée.

Nous proposons dans ce chapitre une approche multirésolution pour la segmentation d'images 3-D en primitives de surface. Elle permet de créer une représentation symbolique de l'image en termes de primitives surfaciques qui sont invariantes par rapport à la position d'observation. De plus, cette représentation doit être complète et locale. L'ensemble de ces primitives est déterminé à partir des différentes combinaisons possibles de signes des courbures Gaussienne K et moyenne H. Chaque représentation correspond à un niveau de résolution donné et décrit le type de surface à l'intérieur de chacune des régions significatives.

En premier lieu, un filtrage Gaussien est appliqué de manière anisotropique sur l'image 3-D originale. Il établit par conséquent le niveau de résolution auquel va avoir lieu la segmentation. La caractéristique principale du filtre réside dans le fait qu'il conserve l'acuité des arêtes significatives tout en filtrant l'intérieur des régions délimitées par celles-ci. En chaque point de l'image filtrée, les courbures Gaussienne K et moyenne H sont par la suite calculées, utilisant une méthode d'approximation de surface qui se base sur une technique standard des moindres carrés. La meilleur fenêtre d'approximation est automatiquement sélectionnée aux points localisés dans le voisinage des arêtes. A la fin du calcul de courbures, une carte d'étiquettes initiale est produite en associant à chaque point de l'image, selon les signes des courbures calculées, un type de primitive parmi les huit qui sont possibles. Finalement, une procédure stochastique de relaxation d'étiquettes ([29]) est appliquée sur la carte d'étiquettes initiale afin de réduire au maximum toute inconsistance locale. Le résultat de cette procédure constitue la carte d'étiquettes finale qui représente symboliquement l'image. En résumé, l'approche proposée fait appel aux étapes suivantes:

Etape 1: Filtrer les régions intermédiaires en utilisant un filtrage gaussien anisotropique.

Etape 2: Calculer les courbures Gaussienne et moyenne des surfaces à partir de l'image filtrée.

Etape 3: Calculer la carte d'étiquettes initiale en se basant sur les signes des courbures de surface calculées.

Etape 4: Calculer la carte d'étiquettes finale de l'image en appliquant sur la carte initiale une procédure de relaxation d'étiquettes.

2.2 Filtrage des régions intermédiaires.

Le filtre Gaussien est appliqué anisotropiquement sur les régions intermédiaires, délimitées par les contours extraits. En général, un filtre isotropique estompe les discontinuités. Un filtre anisotropique se distingue d'un filtre isotropique par le fait qu'il favorise le lissage à l'intérieur des régions et empêche le filtrage à travers des frontières qui les séparent. De cette façon, la netteté des discontinuités correspondantes est conservée. Habituellement, le filtrage anisotropique consiste à effectuer un filtrage isotropique aux endroits situés loin des discontinuités et inhibe progressivement, l'effet de celui-ci au fur et à mesure que l'on approche des discontinuités.

En chaque point (x, y) de l'image, une fenêtre de convolution $W_m(x, y)$ de taille $(2m+1) \times (2m+1)$ et centrée en (x, y) est extraite de l'image originale. La valeur de m qui définit la taille de la fenêtre à extraire est établie par le support du filtre Gaussien utilisé qui, à son tour, dépend de la résolution à laquelle le filtrage doit être effectué. Aux points (x, y) qui se trouvent loin à l'intérieur de chaque région, le filtrage consiste tout simplement à convoluer les données de profondeur originales dans la fenêtre $W_m(x, y)$ correspondante avec le filtre Gaussien. Dans le voisinage des discontinuités, cependant, il doit être effectué différemment afin de conserver l'acuité des arêtes correspondantes tout en filtrant

les points voisins. Comme, à ce stade, la carte d'arêtes a été rendue disponible , la tâche est considérablement simplifiée et la technique qui en résulte consiste premièrement à effectuer une extrapolation de surface aux points se trouvant à l'intérieur de la fenêtre de convolution $W_m(x,y)$ mais n'appartenant pas à la surface traitée (celle déterminée par le pixel central), et deuxièmement à effectuer la convolution en utilisant le même filtre Gaussien. L'opération d'extrapolation consiste à rendre continues les données de profondeur qui se trouvent dans la fenêtre de convolution $W_m(x,y)$, en étendant la surface contenant le point à filtrer au delà de sa frontière. De cette façon, nous pouvons par la suite effectuer un simple filtrage isotropique. Habituellement, chacune des fenêtres de convolution, extraites dans le voisinage des discontinuités, chevauche deux régions adjacentes ou plus (ce dernier cas ne se produit qu'à des endroits où deux contours ou plus se croisent). La région qui correspond à la surface traitée peut être facilement identifiée par l'usage d'un algorithme d'étiquetage de régions (Pavlidis [42]) et d'un point de semence qui correspond au pixel central. Son complément par rapport à la fenêtre de convolution constitue l'ensemble de toutes les positions, incluant celles occupées par les pixels de contour, auxquelles les données sont à extrapoler. L'extrapolation de surface est accomplie par la suite en approximant la surface traitée par la méthode proposée dans [25] et en calculant les valeurs extrapolées par l'usage de la fonction approximante.

Le filtre Gaussien est synthétisé à partir d'une cascade de filtres uniformes ([26]). Le choix de cette technique en cascade a été motivé par le fait qu'il permet l'utilisation de supports de petite taille, typiquement ceux à 3×3, tout en produisant les mêmes résultats. Ceci contribue à réduire les temps de calcul, surtout durant le processus d'extrapolation, qui sinon serait excessivement longs.

2.3 Estimation des courbures K et H dans l'image filtrée.

Il est reconnu que la géométrie différentielle est une théorie pour les surfaces continues et différentiables. Cependant, les objets présents dans la scène donnent naissance, dans les images 3-D correspondantes, à des surfaces qui ne sont que continues par parties. Par conséquent, le calcul des valeurs exactes de courbures Gaussienne K et moyenne H est souvent difficile sinon impossible, aux abords des frontières entre surfaces. Cette section propose une méthode de calcul qui permet une estimation précise de ces valeurs locales de courbure.

Rappelons-nous que le calcul des courbures de surface nécessite l'évaluation des dérivées partielles d'ordre 1 et d'ordre 2 (expressions (2.8) et (2.9)). Or l'estimation de ces dérivées partielles par la méthode des différences finies a un inconvénient majeur qui est la sensibilité à la moindre variation irrégulière des données. Il est donc préférable d'approximer localement les données de profondeur par une fonction continue et différentiable et d'effectuer par la suite la différentiation analytique de la fonction approximante pour obtenir les dérivées désirées. Dans ce cas une dérivabilité d'ordre deux est requise pour la fonction approximante, comme le suggèrent les équations (2.8) et (2.9). Il est habituel de croire que des fonctions de plus haut degré permettront de meilleures approximations. Cependant, ces fonctions ont tendance à s'ajuster aux détails hautes-fréquences tels que le bruit, la texture fine, etc .... Ces considérations nous a donc mené à choisir la fonction biquadratique suivante comme fonction approximante:

$$\hat{z}(x,y) = a + bx + cy + dxy + ex^2 + fy^2 \tag{2.11}$$

L'approximation est effectuée à l'intérieur d'une fenêtre $W_n(x,y)$ de taille $(2n+1) \times (2n+1)$ et centrée sur le point (x,y). La méthode standard des moindres carrés est utilisée et consiste à calculer les coefficients $a, \ldots, f$ de la

fonction biquadratique en minimisant l'erreur quadratique suivante:

$$E^2(x,y) = \sum_{i=-n}^{n} \sum_{j=-n}^{n} \{\hat{z}(x+i, y+j) - z(x+i, y+j)\}^2 \qquad (2.12)$$

où $z(x+i, y+j)$ est la valeur de profondeur originale au point $(x+i, y+j)$ se trouvant dans la fenêtre d'approximation $W_n(x,y)$. Une fois que les coefficients $a, \ldots, f$ sont calculés, les dérivées partielles d'ordre 1 et d'ordre 2 au point (x,y) peuvent être calculées à partir des équations suivantes:

$$\frac{\partial z}{\partial x} = b + dy + 2ex \qquad (2.13)$$

$$\frac{\partial z}{\partial y} = c + dx + 2fy \qquad (2.14)$$

$$\frac{\partial^2 z}{\partial x \partial y} = d = \frac{\partial^2 z}{\partial y \partial x} \qquad (2.15)$$

$$\frac{\partial^2 z}{\partial x^2} = 2e \qquad (2.16)$$

$$\frac{\partial^2 z}{\partial y^2} = 2f \qquad (2.17)$$

obtenues en dérivant analytiquement (2.11). Les valeurs des courbures Gaussienne K et moyenne H au point (x,y) sont finalement évaluées en substituant ces dérivées partielles dans (2.8) et (2.9), respectivement.

On doit se rappeler que les images 3-D contiennent des discontinuités de différentes catégories (discontinuités de profondeur, d'orientation). Effectuer l'approximation des surfaces dans le voisinage d'une discontinuité estompera inévitablement celle-ci et par conséquent, déformera la surface originale qui se trouve aux alentours de la discontinuité. Ceci introduit à son tour de l'imprécision dans le calcul des valeurs estimées de courbures. Ces considérations indiquent que la fenêtre d'approximation pour le calcul des dérivées partielles ne doit pas chevaucher les discontinuités. Nous proposons ici une méthode pour sélectionner la meilleure fenêtre dans laquelle les données de profondeur (et leur gradient) ne présentent pas de changements brusques. L'idée de base est de commencer avec la

taille maximum permise $(2N_{max}+1)\times(2N_{max}+1)$ pour la fenêtre d'approximation et de diminuer itérativement n de 1 jusqu'à ce qu'aucun pixel d'arête ne soit présent dans la fenêtre ou que n atteigne sa valeur minimale 1. La méthode suivante, est appliquée à chaque point (x,y):

Etape 1: $n = N_{max}$

Etape 2: Si n=1 alors aller à l'étape 4. Sinon, extraire une fenêtre de taille $(2n+1)\times(2n+1)$ à partir de la carte d'arêtes obtenue auparavant, et vérifier la présence de pixels d'arête dans cette fenêtre.

Etape 3: Si un ou plusieurs pixels d'arête sont présents dans la fenêtre, alors décrémenter n de 1 et retourner à l'étape 2.

Etape 4: Extraire la fenêtre d'approximation à partir de l'image 3-D et effectuer l'approximation de surfaces, l'estimation de dérivées partielles et le calcul de courbures.

Les courbures de surface sont calculées en chaque point (x,y) de l'image filtrée, en utilisant la méthode que nous venons de décrire. Les résultats sont placés dans des images appellées respectivement *carte de K* et *carte de H*. Il est à noter que les valeurs de courbure aux points d'arête ne sont pas définies. De plus, plus l'on s'approche des discontinuités, plus la taille des fenêtres d'approximation devient petite. Ainsi, les valeurs estimées des courbures seront plus précises aux points trouvés loin à l'intérieur des régions délimités par les arêtes qu'aux points voisins de ces derniers.

2.4 Calcul de la carte d'étiquettes initiale.

La classification locale des surfaces est effectuée à l'aide des primitives suivantes:

$$
\begin{array}{lll}
K>0 \text{ et } H<0: & \text{Surface pic} & \longrightarrow \text{ étiquette} = 1 \\
K>0 \text{ et } H>0: & \text{Surface creuse} & \longrightarrow \text{ étiquette} = 2 \\
K=0 \text{ et } H<0: & \text{Surface crête} & \longrightarrow \text{ étiquette} = 3 \\
K=0 \text{ et } H=0: & \text{Surface plane} & \longrightarrow \text{ étiquette} = 4 \\
K=0 \text{ et } H>0: & \text{Surface vallée} & \longrightarrow \text{ étiquette} = 5 \\
K<0 \text{ et } H<0: & \text{Surface crête en selle} & \longrightarrow \text{ étiquette} = 6 \\
K<0 \text{ et } H=0: & \text{Surface minimale} & \longrightarrow \text{ étiquette} = 7 \\
K<0 \text{ et } H>0: & \text{Surface vallée en selle} & \longrightarrow \text{ étiquette} = 8
\end{array}
\tag{2.18}
$$

et résulte en une *carte d'étiquettes* dans laquelle chaque point est codé par une valeur de 1 à 8.

Cependant, dans la pratique (2.18) constitue un dictionnaire de primitives irréaliste. Les variations dues au bruit dans les valeurs de profondeur et les effets causés par d'autres phénomènes hautes-fréquences empêchent K et H d'être rigoureusement nulles dans les régions où elles devraient l'être. En plus, les valeurs observées de K et H peuvent avoir des signes opposés à ceux des valeurs réelles, surtout lorsque la valeur absolue de celles-ci est très faible ([27]). En conséquence, les courbures Gaussienne K et moyenne H nécessitent d'être seuillées avant que les règles (2.18) puissent être appliquées sur elles.

Habituellement, un seuil est sélectionné symétriquement par rapport à zéro et pour chaque courbure. Soient ε_K et ε_H les seuils de K et de H, respectivement. En général, leur sélection n'est pas un problème trivial parce qu'ils dépendent fortement du niveau de bruit dans l'image, dont l'estimation demeure encore un problème difficile à résoudre. Yokoya et Levine [28] ont proposé une méthode qui permet la sélection des valeurs de ces seuils. Leur technique se base sur des hypothèses qui supposent que le bruit dans l'image ne dépend pas des objets

observés dans la scène et a été causé uniquement par le capteur et par la procédure de calibration utilisée. La méthode consiste à calculer les courbures K et H d'une surface de référence plane, fournie par le capteur courant, et à attribuer à ε_K et ε_H des valeurs minimales de courbures nécessaires à classifier la surface de référence comme étant plane. Afin d'être moins dépendant du capteur utilisé (les sorties des différents types de capteur n'ont pas le même niveau de bruit quand elles sont comparées entre elles), et vu que nos images 3-D ont déjà été filtrées avant d'être segmentées, nous proposons une méthode qui permet l'estimation de ε_K et ε_H à partir de l'image filtrée elle-même.

La technique proposée fait intervenir les étapes suivantes: 1) extraire les régions planes de l'image filtrée; 2) calculer les courbures K et H des plans extraits; 3) sélectionner ε_K et ε_H comme étant proportionnels aux écart-types α_K et α_H des courbures K et H respectivement.

En général, les surfaces planes dans une image 3-D génèrent des pics dans un histogramme, établi à partir de la distribution sur toute l'image des normes $||\vec{n}||$ des vecteurs normaux $\vec{n}$ aux surfaces. L'extraction des régions planes de l'image filtrée consiste donc à établir cet histogramme à partir de l'image, à détecter par la suite les différents pics (ou modes) dans l'histogramme et finalement à en déduire les points correspondants dans l'image. Cette opération résulte en une image binaire dans laquelle les points appartenant aux différentes surfaces planes de l'image filtrée ont une valeur 0 et le restant des points, correspondant à d'autres types de surface, possèdent la valeur 1.

Habituellement, la détection des modes dans un histogramme n'est pas triviale, surtout lorsque ce dernier est établi à partir d'une image qui résulte de la différentiation de l'image originale. La détection par seuillage est souvent difficile et non optimale parce que la sélection du seuil approprié, permettant la détection de tous les modes significatifs, est impossible. Cependant, s'il existe un moyen

quelconque qui permette l'accentuation de la caractéristique modale de tous les pics, rendant évidente la séparation entre les modes et le reste de l'histogramme (les vallées), alors le problème sera résolu. Cette idée, initialement indiquée par Touzani et Postaire [43], nous a poussé à faire usage de la relaxation probabiliste (Peleg [44]) dans le problème de détection de modes.

Soit $\{a_1, a_2, \ldots, a_n\}$ l'ensemble des n points de l'histogramme et soit Λ l'ensemble des étiquettes possibles, λ_{mod} et λ_{val}, en chaque point a_i. Si $P_i^{(0)}(\lambda)$ dénote la probabilité initiale assignée à l'étiquette λ, $\lambda \in \Lambda$, au point a_i alors les valeurs estimées des probabilités *a priori* $P(\lambda)$ seront données par:

$$\hat{P}(\lambda) = \frac{1}{n} \sum_{i=1}^{n} P_i^{(0)}(\lambda) \tag{2.19}$$

et celles des probabilités conjointes par

$$\hat{P}_{i,i+\delta}(\lambda, \lambda') = \tfrac{1}{n} \textstyle\sum_{i=1}^{n} P_i^{(0)}(\lambda) P_{i+\delta}^{(0)}(\lambda'), \quad \delta \in \Delta \tag{2.20}$$

où Δ est un voisinage du point a_i et l'indice $i + \delta$ identifie le voisin $a_{i+\delta}$ de a_i. La probabilité initiale $P_i^{(0)}(\lambda)$ est mise à jour à chaque itération par la formule suivante:

$$P_i^{k+1}(\lambda) = \tfrac{1}{n_\Delta} \textstyle\sum_{\delta \in \Delta} q_{i,i+\delta}^{(k+1)}(\lambda), \quad \lambda \in \Lambda \tag{2.21}$$

où n_Δ dénote le nombre de voisins de a_i et

$$q_{i,i+\delta}^{(k+1)}(\lambda) = \frac{S_{i,i+\delta}^{(k+1)}(\lambda)}{\sum_{\lambda' \in \Lambda} S_{i,i+\delta}^{(k+1)}(\lambda')}, \quad \delta \in \Delta \tag{2.22}$$

$$S_{i,i+\delta}^{(k+1)}(\lambda) = \sum_{\lambda' \in \Lambda} \hat{P}_i^k(\lambda) \hat{P}_{i+\delta}^k(\lambda') r_{i,i+\delta}(\lambda, \lambda') \tag{2.23}$$

Les coefficients de compatibilité $r_{i,i+\delta}(\lambda, \lambda')$, qui représentent les mesures de compatibilité entre les étiquettes λ et λ', sont donnés par

$$r_{i,i+\delta}(\lambda, \lambda') = \frac{\hat{P}_{i,i+\delta}(\lambda, \lambda')}{\hat{P}(\lambda)\hat{P}(\lambda')}, \quad \delta \in \Delta \tag{2.24}$$

A la fin de la N^{ieme} itération (la dernière), l'étiquette du point a_i est définie comme étant celle ayant la probabilité $P_i^{(N)}(\lambda)$ maximale.

Il est à noter que cette procédure de détection extrait en même temps certains points en provenance de surfaces d'ordre deux ou plus (e.g. points sphériques qui constituent des cercles concentriques dans l'image binaire, ou des points cylindriques qui donnent naissance à des lignes parallèles dans l'image binaire, etc ...). Ces points peuvent, par contre, facilement être éliminés en appliquant du filtrage morphologique (une érosion suivie d'une expansion, par exemple) sur l'image binaire des points marqués.

La Figure 2.2 illustre les résultats types obtenus à la fin de la procédure d'extraction de plans. L'image originale, montrée à la Figure 2.2.c, a été filtrée avec σ=2.2. Les images binaires, obtenues avant et après filtrage morphologique, sont présentées aux Figures 2.2.a et 2.2.b, respectivement. Les régions illustrées en noir dans ces images correspondent aux plans extraits. Il est à noter que les points et régions en provenance de surfaces d'ordre supérieur (Figure 2.2.a) sont bien nettoyés par le filtrage morphologique.



(a) (b) (c)

Figure 2.2 Resultats d'extraction de plans. (a) Plans extraits avant filtrage morphologique. (b) Plans extraits après filtrage morphologique. (c) Image 3-D originale.

2.5 Calcul de la carte d'étiquettes finale.

La carte d'étiquettes produite précédemment nécessite d'être traitée de nouveau et celà est dû aux faits suivants:

- les pixels adjacents (connectivité 8) aux discontinuités ont des valeurs de courbure indéfinies (parce qu'à l'endroit de ces pixels, même une fenêtre d'approximation de taille 3×3 va empièter les discontinuités). Par conséquent ils ne sont pas étiquetés.
- comme nous nous y attendions, les valeurs de courbure calculées dans le voisinage des discontinuités sont généralement imprécises et ceci est dû à l'usage d'une petite taille pour la fenêtre d'approximation. Les pixels correspondants ont ainsi tendance à être étiquetés de manière inconsistente par rapport à la région qui les entoure.

Une procédure stochastique de relaxation d'étiquettes ([29]) est appliquée à la carte d'étiquettes initiale afin d'améliorer sa consistence locale. Chaque pixel dont l'étiquette n'est pas consistent avec celle de ses voisins verra son étiquette substituée par une autre, sélectionnée en concordance avec une fonction de mérite qui lui est associée. Par la suite, pour chaque pixel non étiqueté, une étiquette est choisie sur la base de la majorité des étiquettes trouvées dans son voisinage immédiat 3×3.

2.6 Résultats expérimentaux.

Des résultats expérimentaux ont été obtenus pour la série d'images 3-D utilisée précédemment dans le chapitre 1. Pour tous les résultats, une taille maximum de 15×15 a été choisie pour la fenêtre d'approximation, utilisée lors du calcul local des courbures Gaussienne K et moyenne H. Pour chacune des images, les seuils

ϵ_K et ϵ_H ont été automatiquement sélectionnés comme étant les écart-types α_K et α_H de K et H du plan de référence (le fond), respectivement. Il est à mentionner qu'une expansion morphologique a été appliquée deux fois sur les arêtes extraites avant que la segmentation ne commence. Ceci est dû au fait que ces arêtes ne sont pas toutes fermées en général (il s'agit d'une conséquence inhérente à l'usage de la dérivée première d'une Gaussienne comme détecteur d'arêtes [13]). Habituellement, cette opération ferme les arêtes puisque les interstices entre les extrémités de celles-ci ne sont en général que d'un ou deux pixels de large. Cependant, l'opération peut s'avérer nuisible pour les petites régions. En effet, appliquer deux fois l'expansion morphologique sur les arêtes qui délimitent ces dernières peut éventuellement les remplir (voir, par exemple, une des faces du polyèdre central de la Figure 2.6.d).

Les résultats sont montrés aux Figures 2.3 à 2.7, respectivement. Chacune d'elles illustre les résultats obtenus à partir d'une image particuliere et est présentée de la manière suivante: a) la carte d'étiquettes initiale; b) la carte d'étiquettes finales; c) la carte d'arêtes; d) la superposition des cartes d'arêtes et d'étiquettes finales. Les régions en noir dans les cartes d'étiquettes correspondent aux artifices causés par des effets d'ombre. Les pixels à l'intérieur de ces régions n'ont pas été classifiés lors du calcul de la carte d'étiquettes finale (ils ne le sont pas plus lors des étapes précédentes vu que l'expansion morphologique remplit ces régions en général). Les couleurs gris pâle, gris, et gris foncé correspondent respectivement aux primitives de surface de type surface "pic", "surface plane", et "surface crête". A partir de ces résultats, nous pouvons formuler les observations suivantes:

1. Les cartes d'étiquettes initiale (Figures 2.3, 2.5, et 2.6) ne contiennent presque pas d'artifices causés par le bruit. Il en existe quand même mais très peu et seulement dans le voisinage des discontinuités. Ceci montre que

notre méthode de sélection de seuils ε_K et ε_H est efficace et que les valeurs sélectionnées sont bonnes. Ceci montre aussi que les valeurs calculées de K et de H sont moins bonnes aux alentours des arêtes qu'à l'intérieur des régions. Il est à noter aussi que ces artifices sont en général de taille négligeable, et peuvent facilement être éliminées en seuillant l'image sur la base de la taille des régions. Nous préférons cependant la méthode de relaxation d'étiquettes vu qu'elle permet la récupération d'étiquettes exactes en ces pixels.

2. Nous pouvons faire les mêmes remarques pour les cartes d'étiquettes initiales des Figures 2.4 et 2.7, à l'exception des surfaces cylindriques, orientées horizontalement. En général, la mauvaise classification de ces dernières est due au fait que le niveau de bruit est plus important sur les surfaces orientées horizontalement que sur celles orientées suivant n'importe quelle autre direction (comparer les cylindres de la Figure 2.4, par exemple). Nous suggèrerons, dans le dernier chapitre consacré à la conclusion et aux améliorations futures, une méthode que nous pensons bonne pour résoudre ce problème.

3. En général, les différents types de surface (surface sphérique, polyédrique, et cylindrique) ont été bien identifiés. Les sphères ont cependant tendance à être classifiées en surface sphérique dans la région centrale et en surface cylindrique dans la région périphérique (comparer la sphère située au bas de l'image de la Figure 2.6 aux autres sphères de cette même Figure. Cette tendance est autant plus accentuée lorsque le rayon de la sphère est grand. Ce phénomène peut être expliqué par le fait qu'à la périphérie une des deux courbures devient vite négligeable (celle associée à la direction perpendiculaire à la direction longitudinale) et par conséquent la valeur de K ne subsiste pas au seuillage par ε_K.

4. La comparaison faite entre les cartes d'étiquettes initiales et les cartes d'étiquettes finales montre que la procédure de relaxation d'étiquettes utilisée est performante. La plupart des étiquettes inconsistantes par rapport à leur voisinage ont été éliminées et remplacées par des bonnes étiquettes.

5. La présence d'arêtes dans les cartes d'étiquettes finales permet la segmentation des régions adjacentes de même type de primitive mais séparées par des discontinuités de profondeur et/ou d'orientation. Ainsi, les surfaces polyédriques ont été bien partitionnées en leurs différentes faces (voir le polyèdre de la Figure 2.6) et nettement séparés du plan de référence (Figure 2.3).

H	+	0	-
-	Surface pic	Surface crête	Surface crête en selle
0	rien	Surface plane	Surface minimale
+	Surface creuse	Surface vallée	Surface vallée en selle

Tableau 2.1 Définition des huit primitives de surface. Le cas où $K > 0$ et $H = 0$ n'est pas défini.

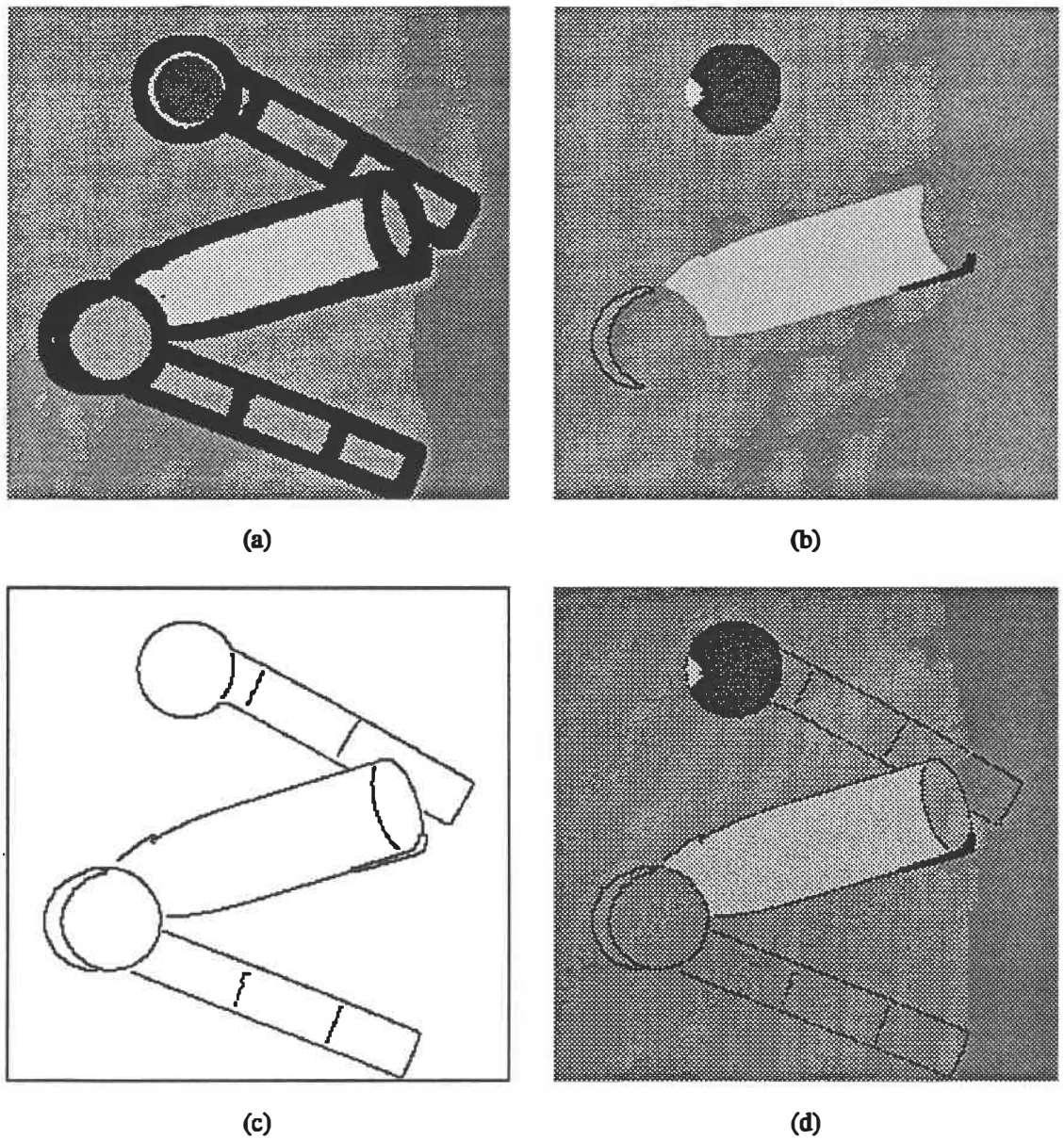


Figure 2.3 Segmentation de l'image "BLOC11". (a) Carte d'étiquettes initiale. (b) Carte d'étiquettes finale. (c) Carte d'arêtes. (d) Superposition des cartes d'arêtes et d'étiquettes finales.

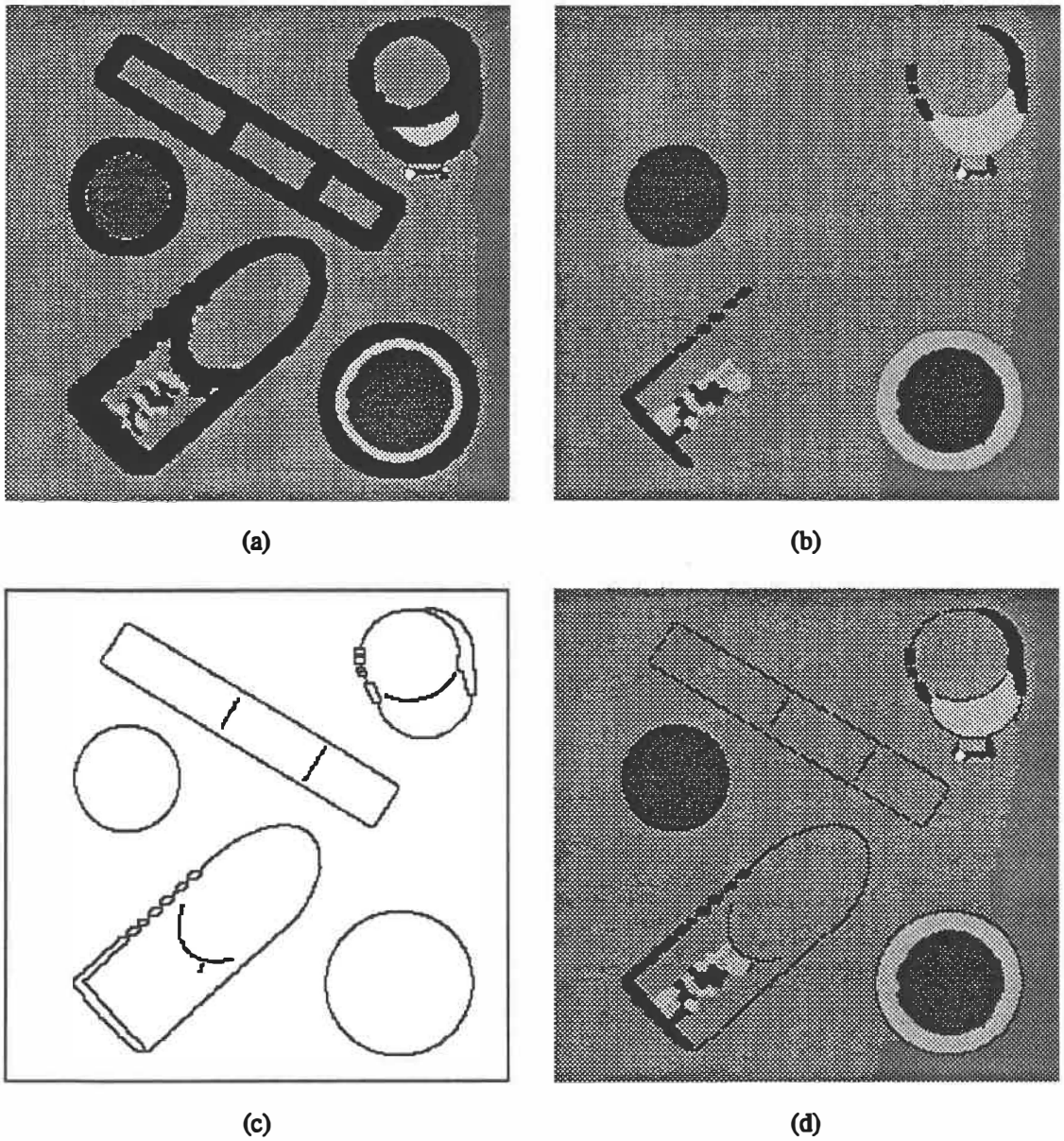


Figure 2.4 Segmentation de l'image "BLOC9". (a) Carte d'étiquettes initiale. (b) Carte d'étiquettes finale. (c) Carte d'arêtes. (d) Superposition des cartes d'arêtes et d'étiquettes finales.

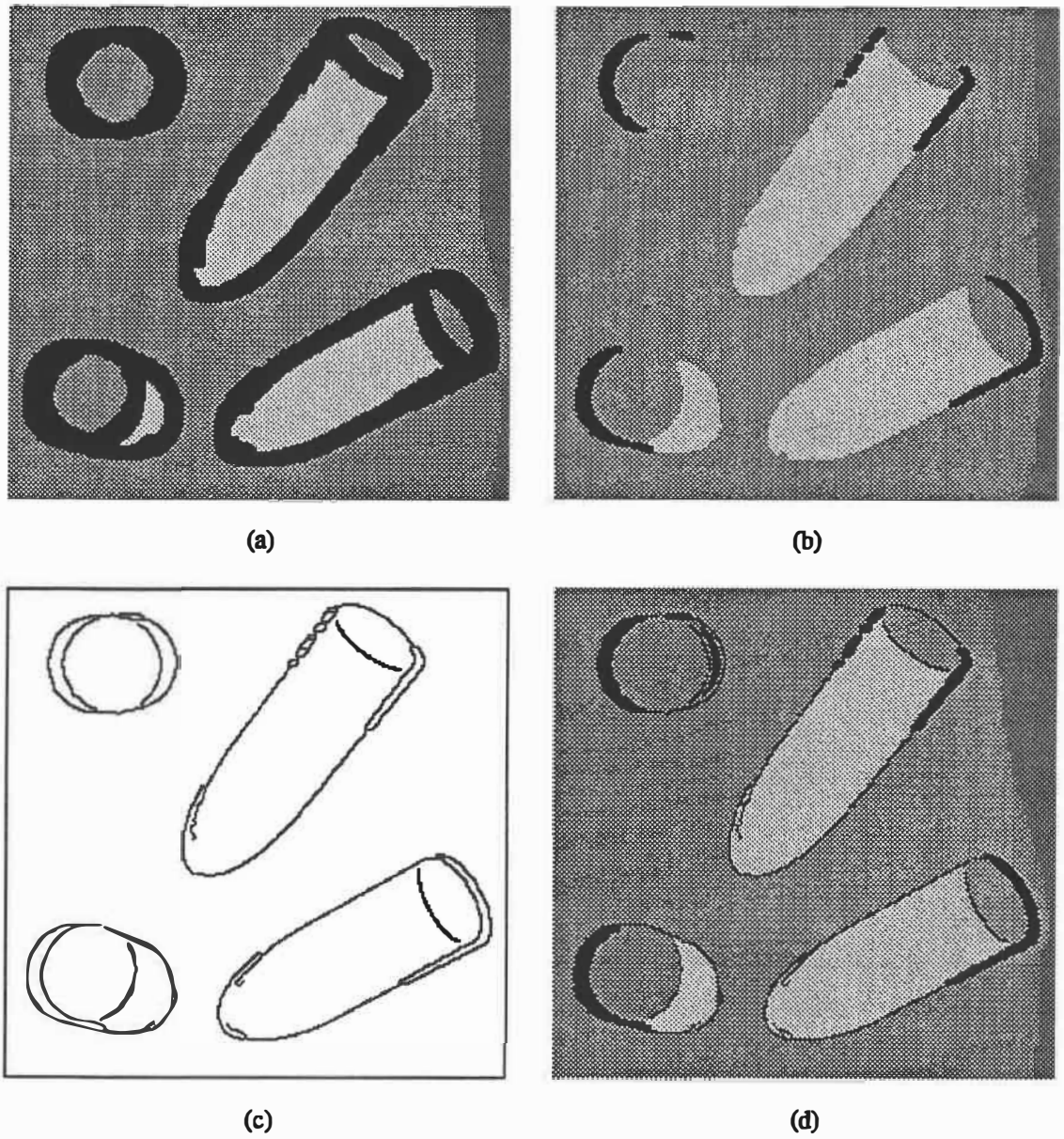


Figure 2.5 Segmentation de l'image "BLOC4". (a) Carte d'étiquettes initiale. (b) Carte d'étiquettes finale. (c) Carte d'arêtes. (d) Superposition des cartes d'arêtes et d'étiquettes finales.

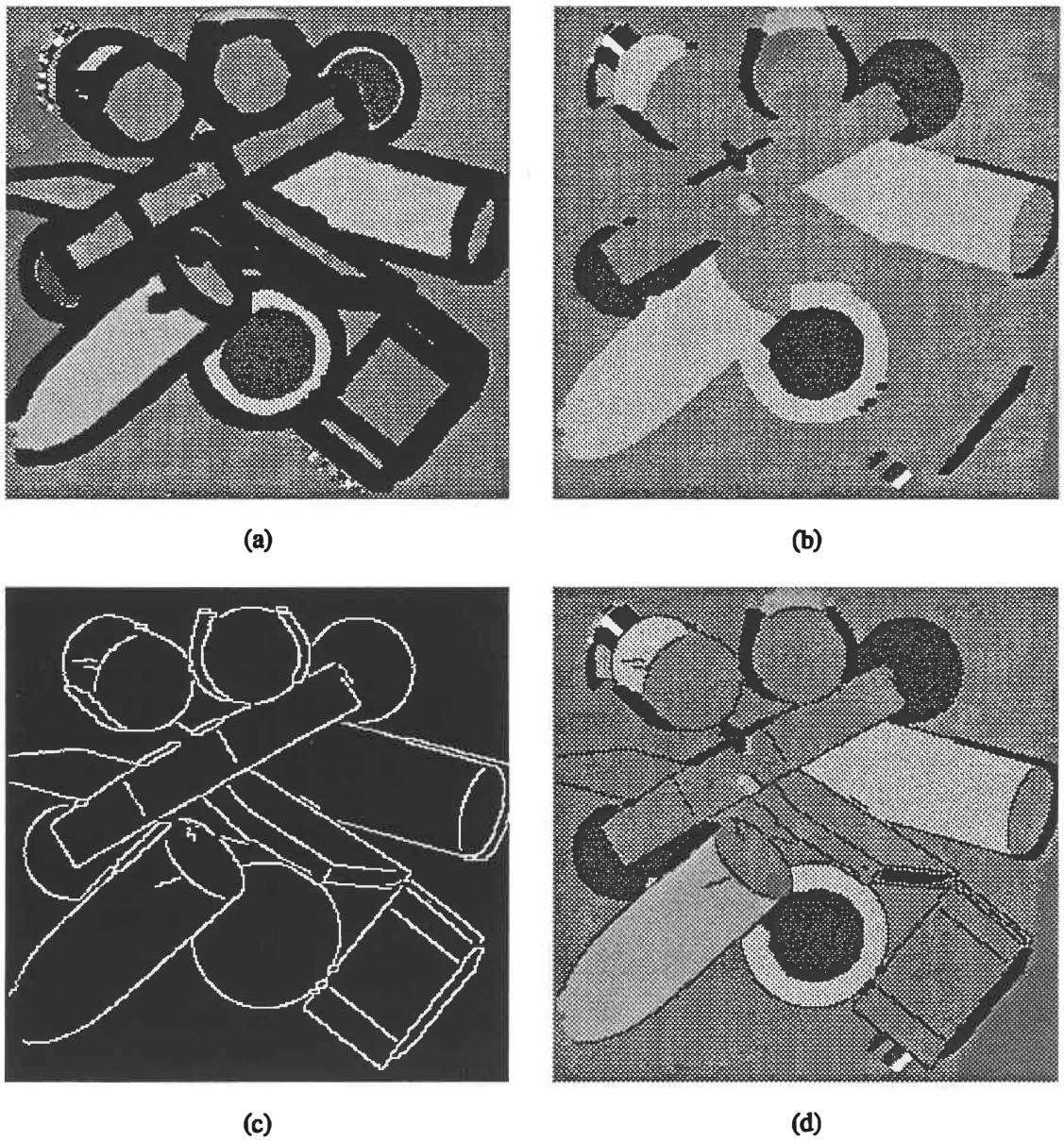


Figure 2.6 Segmentation de l'image "BLOC15". (a) Carte d'étiquettes initiale. (b) Carte d'étiquettes finale. (c) Carte d'arêtes. (d) Superposition des cartes d'arêtes et d'étiquettes finales.

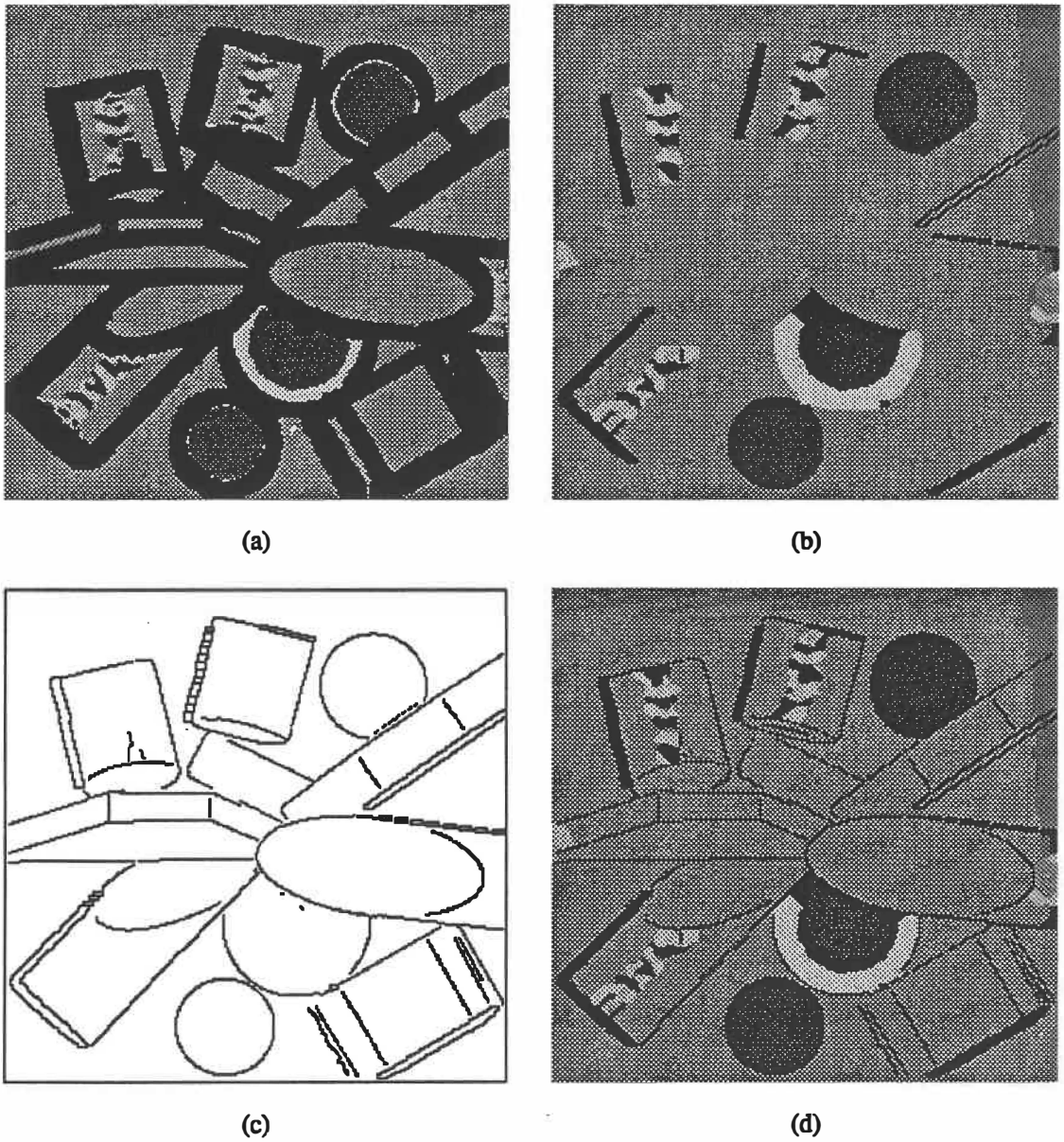


Figure 2.7 Segmentation de l'image "BLOC16". (a) Carte d'étiquettes initiale. (b) Carte d'étiquettes finale. (c) Carte d'arêtes. (d) Superposition des cartes d'arêtes et d'étiquettes finales.

CHAPITRE 3

Proposition d'une représentation multirésolution.

Afin de distinguer un objet polyédrique d'une pièce mécanique, nous n'avons pas besoin d'aller chercher tous les détails concernant leurs surfaces frontalières pour effectuer leur discrimination. Seul les contours de ces objets suffisent. Cependant, si nous avons à faire la distinction entre deux cubes de taille identique, chacun d'eux sculpté différemment sur ses faces, alors cette distinction ne peut être accomplie par simple comparaison des contours. D'autres détails, tels que ceux concernant la sculpture, sont nécessaires afin de trouver le bon cube. Ces exemples nous montrent que même l'observateur moyen applique inconsciemment un effet psychologique par lequel il observe chaque chose à différents niveaux de résolution suivant l'usage qu'il désire en faire. En plus, ils nous montrent qu'il existe dans la nature une hiérarchie pour la description. Le rang de visibilité détermine en général l'existence d'un objet donné. Un grain de blé ne peut exister (être visible) au niveau auquel le champs de blé est observé. Cependant, lorsqu'il est visible, il est bien localisé. Cette bonne localisation de l'objet s'étend sur tout son rang de visibilité. La forme intérieure des surfaces de l'objet est aussi organisée de manière hiérarchique. Cette hiérarchie est cependant intimement reliée à celle des changements physiques. A la disparition de l'un d'eux, les surfaces adjacentes correspondantes interagissent entre elles, changeant ainsi leur forme originale et introduisant plus de détails.

Il existe donc deux types d'espace d'échelles, l'un pour les arêtes, l'autre pour les surfaces. Ces deux types interagissent entre-eux et possèdent en général des comportements opposés. L'un ne sera cependant totalement complet sans

la présence de l'autre. Ce chapitre introduit la représentation, proposée par Witkin [20], dans l'espace d'échelles des signaux 1-D pour ensuite décrire la représentation multirésolution proposée qui concrétise les idées que nous venons d'introduire.

3.1 Principe fondamental.

3.1.1 Représentation dans l'espace d'échelles: approche de Witkin.

Une description qui caractérise un signal 1-D par ses extrémums et ceux de ses premières dérivées est qualitative. Habituellement, elle nécessite des calculs de dérivées qui à leur tour font intervenir la notion d'échelle. L'échelle influence en général les dérivées calculées et par conséquent, la description qualitative qui en découle. Ainsi, la description d'un signal n'aura aucune signification si l'échelle à laquelle elle a été extraite n'est pas précisée. De plus, différentes échelles produisent en général différentes descriptions.

Witkin [20] a introduit la représentation dans l'espace d'échelles des signaux 1-D. Pour un signal donné, celle-ci unifie en une *carte d'échelles* les descriptions indépendantes du signal à différentes échelles. Soit $g(x)$ une fonction Gaussienne à une dimension dont l'écart-type est σ et soit $g''(x)$ la dérivée seconde de $g(x)$. Les expressions de $g(x)$ et de $g''(x)$ sont respectivement données par:

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{x^2}{2\sigma^2}) \tag{3.1}$$

$$g''(x) = \frac{d^2 g(x)}{dx^2} = \frac{1}{\sigma^3\sqrt{2\pi}}(\frac{x^2}{\sigma^2} - 1) \exp(-\frac{x^2}{2\sigma^2}) \tag{3.2}$$

Si $i(x)$ dénote le signal 1-D à représenter dans l'espace d'échelles, alors la carte d'échelles correspondante sera établie de la manière suivante: $i(x)$ est d'abord convolué avec $g''(x)$ pour un intervalle continu de valeurs de σ, puis les positions des passages par zéro du signal convolué sont marquées dans la carte pour chaque

valeur de σ. L'écart-type de la Gaussienne joue donc le rôle de paramètre échelle. Un exemple de carte d'échelles est illustré à la Figure 3.1. La Figure 3.1.a montre le signal $i(x)$ et sa convolution avec une fonction $g''(x)$ de grand support est présentée à la Figure 3.1.b. Les axes vertical et horizontal dans la carte d'échelles (Figure 3.1.c) représentent l'axe des σ et l'axe des positions x dans le signal, respectivement. Pour chaque valeur de σ, les positions des passages par zéro de $i(x) * g''(x)$ sont indiquées par les points se trouvant le long d'une ligne horizontale dans la carte. Par exemple, les points se trouvant le long de la ligne en pointillé à $\sigma=\sigma_l$ indiquent les positions des passages par zéro du signal montré à la Figure 3.1.b. Ainsi, la carte d'échelles illustre le comportement des passages par zéro à travers l'espace d'échelles. Pour des petites valeurs de σ, les passages par zéro caractérisent tous les changements brusques de l'image originale. A des valeurs plus grossières de σ, seuls les changements majeurs sont caractérisés.

Outre le fait qu'elle détermine le signal de manière unique, à une constante multiplicative près (Yuille [33]), la carte d'échelles possède les caractéristiques suivantes:

- Les "contours" de la carte d'échelles consistent en des lignes sensiblement verticales (Figure 3.2(1)) ou formant des dômes (Figure 3.2(2)).

- Des contours additionnels, du genre de ceux montrés aux Figures 3.2(3) et 3.2(4), ne peuvent pas apparaître lorsque σ augmente. Cette caractéristique de la carte d'échelles est d'ailleurs justifiée par des études théoriques de certains auteurs ([31], [32], [33]).

3.1.2 Proposition d'une représentation multirésolution.

A chaque niveau de résolution, deux types d'information sont rendus explicites par la carte d'arêtes et la carte d'étiquettes finales. Il s'agit des types d'information

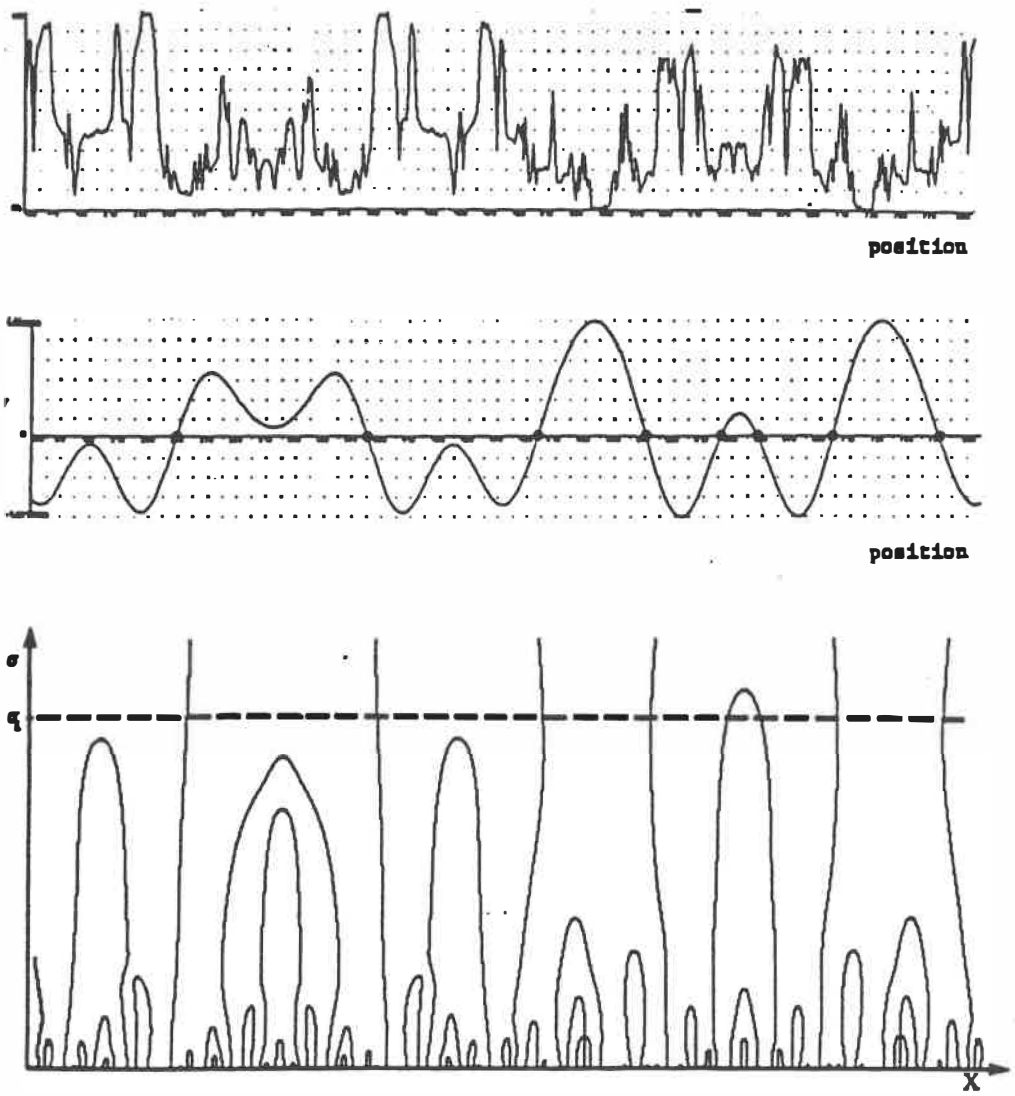


Figure 3.1 Représentation dans l'espace d'échelles. (a) Exemple de signal 1-D. (b) Convolution du signal avec $g''(x)$ de grand support. (c) Carte d'échelles. Les passages par zéro du signal montré en (b) se trouvent le long de la ligne horizontale $\sigma = \sigma_1$.

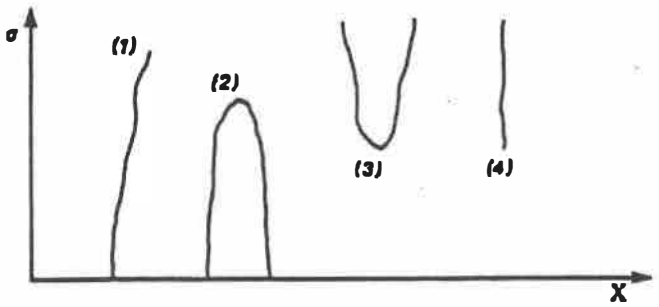


Figure 3.2 **Types de contours trouvés dans la carte d'échelles. Les contours de type (1) et (2) sont courants tandis que ceux de types (3) et (4) n'y apparaissent jamais.**

suivants:

1. Information concernant les régions de surfaces: Cette information est fournie par la carte d'étiquettes finale augmentée de la carte d'arêtes. Cette superposition est nécessaire pour partitionner correctement les surfaces adjacentes de même types de primitive mais séparées par des discontinuités physiques.

2. Information concernant les discontinuités de surface: Elle est fournie par la carte d'arêtes dans le cas des discontinuités de profondeur et d'orientation et par les frontières des régions de surface qui ne coincident pas aux arêtes d'ordre 0 et d'ordre 1, dans le cas des jointures douces.

Ces deux types d'information peuvent être représentés sous forme d'un graphe d'adjacence dans lequel chaque noeud représente une région de surface et à chaque arc, qui relie deux noeuds, est associé le type de discontinuité qui séparent les régions de surfaces correspondantes.

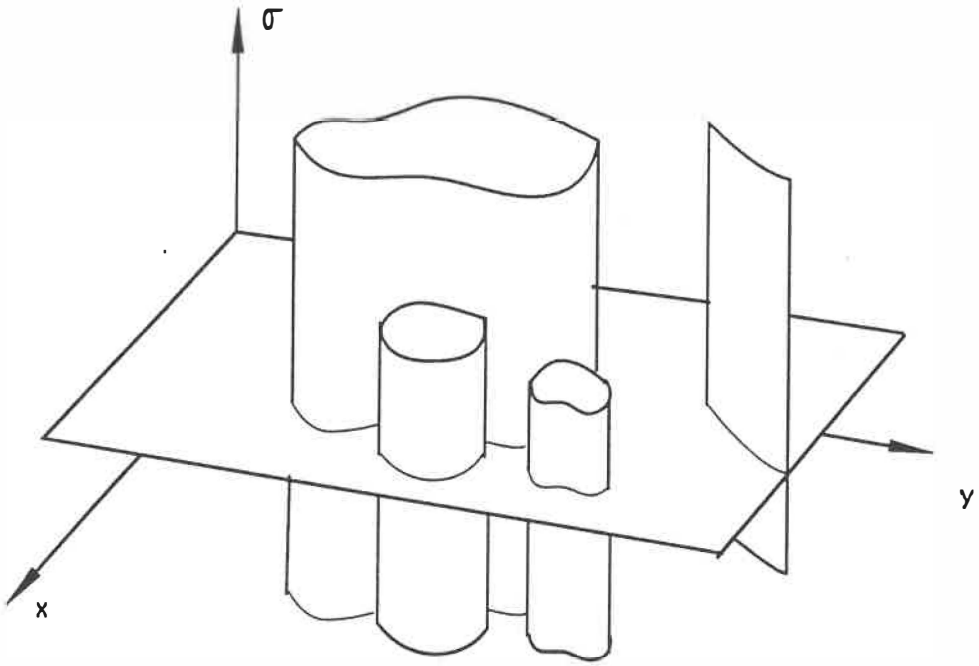


Figure 3.3 Représentation multirésolution d'images 3D.

Il est à noter que l'ensemble des cartes d'arêtes, détectées à différentes échelles, constitue une représentation multirésolution qui généralise, au cas des signaux image, la représentation dans l'espace d'échelles de Witkin [20]. Un exemple de cette représentation multirésolution est montré à la Figure 3.3. Le lieu des passages par zéro du signal image convolué constitue, dans l'espace d'échelles tridimensionnel (x, y, σ), des surfaces cylindriques dont l'axe est perpendiculaire au plan (x, y). Pour chaque valeur de σ, les positions des passages par zéro du signal image convolué sont indiquées par des points se trouvant le long des courbes planes qui résultent de l'intersection d'un plan horizontal avec la surface cylindrique. La caractéristique principale de cette représentation réside dans le fait qu'à chaque résolution, les contours bien localisés des régions de l'image sont immédiatement disponibles.

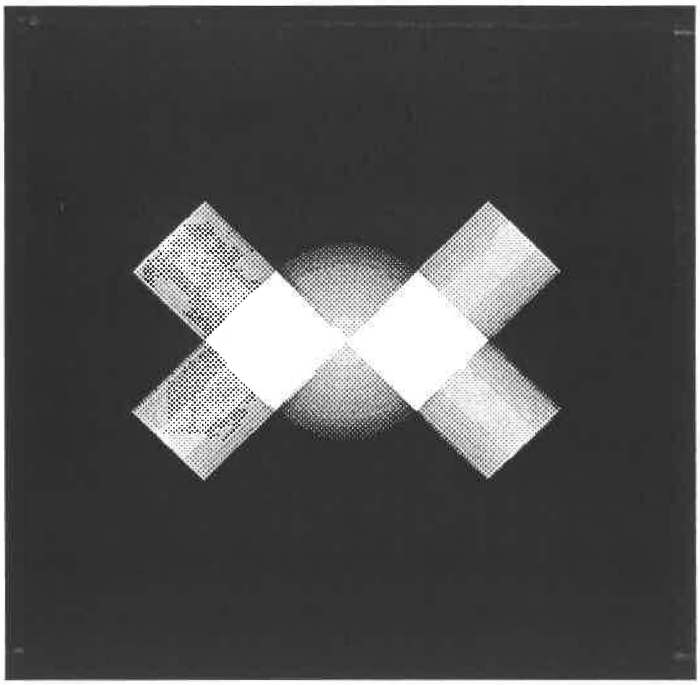


Figure 3.4 Image synthétique originale

3.2 Résultats expérimentaux.

Nous allons présenter dans cette section un exemple de représentation multirésolution de surfaces, conforme à l'approche que nous venons de décrire. L'image dont la représentation est à extraire est synthétique. La raison pour laquelle nous avons choisi d'éffectuer la représentation d'une image synthétique plutôt que celle d'une image réelle est uniquement la facilité qu'elle offre dans la sélection des échelles de description types permettant de bien illustrer les différents aspects de notre approche.

L'unique objet dans l'image a l'apparence d'une pièce industrielle et se compose de surfaces planes, cylindriques, et sphériques. La version codée en niveau de gris de l'image est montrée à la Figure 3.4. Les différentes échelles auxquelles l'objet va être décrit ont été sélectionnées comme étant 0.0, 1.95, et 3.7, respectivement.

A chacun des niveaux de résolution choisis, les arêtes sont extraites par

la méthode d'extraction d'arêtes qui fait l'objet du chapitre 1. Les résultats obtenus sont illustrés à la Figure 3.5. Nous pouvons observer que les arêtes de l'objet sont bien extraites et bien localisées. Au niveau de résolution $\sigma =$ 0.0, c'est-à-dire lorsque l'image n'est pas filtrée, les arêtes se présentent avec tous leurs détails (Figure 3.5.a). A l'échelle $\sigma = 1.95$, celles qui séparaient les surfaces cylindriques des surfaces planes de l'objet disparaissent (Figure 3.5.b). Ceci veut dire que le champ de visibilité de ces arêtes occupe l'intervalle de résolution qui va de 0.0 à 1.95. De la même manière, seul le contour de l'objet est extrait au niveau de résolution σ=3.7 parce qu'à ce niveau les discontinuités, créant les frontières entre les surfaces planes et sphériques de l'objet, ne sont plus significatives et par conséquent, sont devenues indétectables au détecteur d'arêtes utilisé (Figure 3.5.c).

La segmentation en primitives de surface des régions, délimitées par les arêtes extraites, est ensuite appliquée sur l'image, aux trois échelles choisies. L'étape initiale de filtrage Gaussien anisotropique emploie la méthode par extrapolation dont il est question à la section 2.2. Les résultats de cette étape sont présentés à la Figure 3.6. Les versions ombragées des images correspondantes ont été utilisées pour mieux illustrer les effets du filtrage. Il est à noter qu'il existe une certaine interaction entre la partie haute de la sphère et celle du fond qui lui est adjacente. Le degré d'interaction est d'autant plus accentué que le niveau de résolution est grossier (comparer les Figures 3.6.b et 3.6.c). D'après nous, ceci est dû au fait que certains pixels de l'arête, séparant les deux régions, sont mal localisés. La raison de cette mauvaise localisation demeure pour nous introuvable. Cependant, cet artifice a su montrer que la bonne localisation des arêtes est essentielle à la technique de filtrage proposée. La Figure 3.7 présente les images résultantes de la superposition des cartes d'arêtes et d'étiquettes finales. Les couleurs gris pâle, gris, et gris foncé sont encore une fois utilisées pour identifier les primitives de

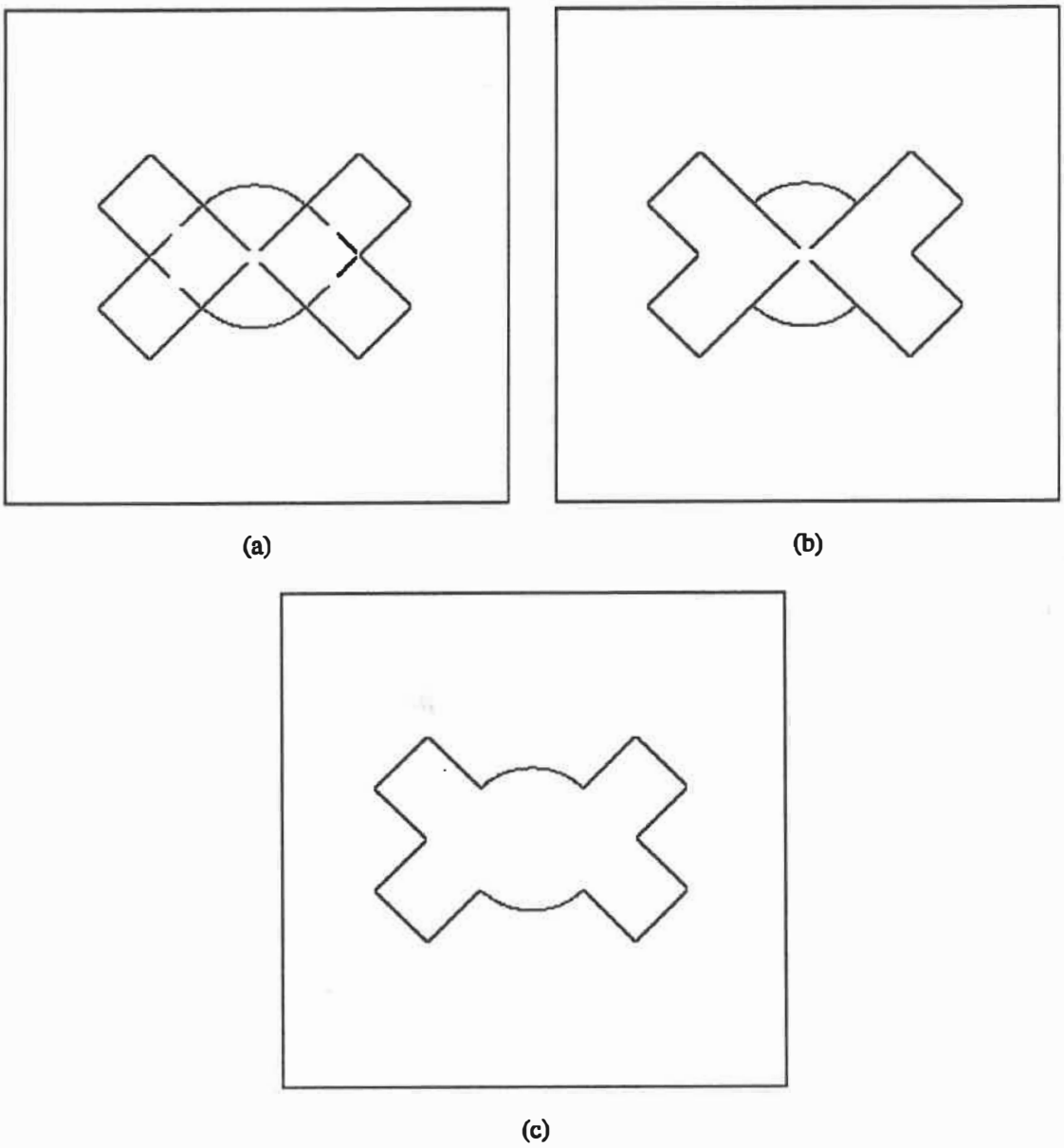


Figure 3.5 Arêtes extraites à différentes échelles. (a) $\sigma = 0.0$. (b) $\sigma = 1.95$. (c) $\sigma = 3.7$.

surface de type "surface pic", "surface plane", et "surface crête", respectivement. A partir de ces résultats, nous pouvons observer que les différentes surfaces de l'objet sont bien identifiées. Ce fait se présente mieux à la Figure 3.7.a où les surfaces n'interagissent pas encore entre-elles. De plus, à l'echelle $\sigma = 1.95$, la disparition des arêtes qui figuraient entre les surfaces cylindriques et planes donne naissance à l'estompage des discontinuités correspondantes et par conséquent, la création d'autres primitives de surface dans le voisinage des discontinuités originales (Figure 3.7.b). De façon similaire, l'éclipse des arêtes, qui séparaient les surfaces sphériques des surfaces planes, de la carte d'arêtes au niveau de résolution σ=3.7 permet le filtrage des discontinuités correspondantes et à nouveau génère différents types de primitives dans la carte d'étiquettes finale (Figure 3.7.c).

Notons que les résultats montrés à la Figure 3.7 constituent en même temps une version discrète de la représentation multirésolution de l'objet. Ainsi, en allant de $\sigma = 0.0$ à $\sigma = 3.7$, la représentation par arêtes de l'objet se simplifie graduellement. Par contre, la représentation par primitives de surface de l'objet devient de plus en plus complexe puisque les surfaces adjacentes se fusionnent aux endroits où les arêtes disparaissent. Ce fusionnement modifie la forme de ces surfaces et par conséquent, génère des primitives de surface additionnelles dans la représentation finale

De manière générale, lorsque σ augmente, la complexité au niveau des surfaces augmente tandis que celle des arêtes diminue. En effet, une diminution dans la complexité au niveau des arêtes implique la disparition d'un certain nombre d'arêtes et par conséquent, l'interaction entre les régions adjacentes correspondantes. Cette interaction déforme les surfaces créant ainsi un plus grand nombre de primitives de surface et par conséquent, conduit à une augmentation dans la complexité au niveau des surfaces. Globalement, cette complexité demeure constante, et seule sa localisation varie en fonction de l'échelle, c'est-à-dire lorsque σ

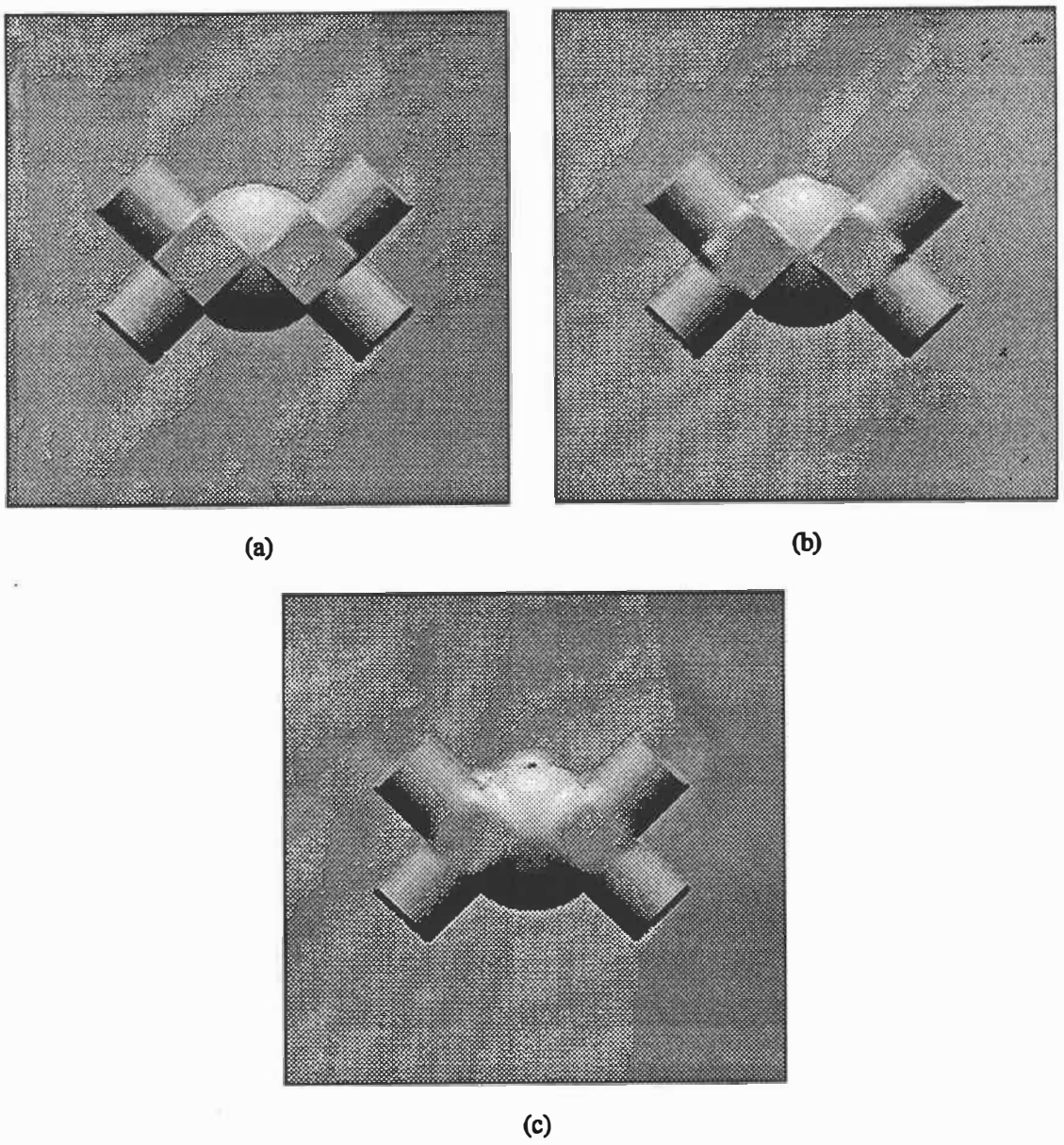
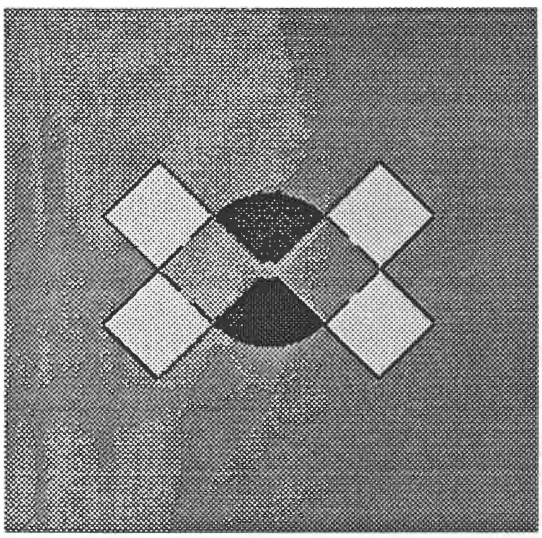
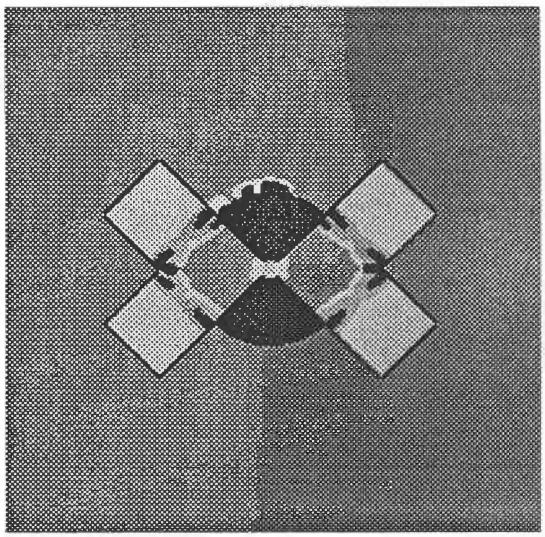


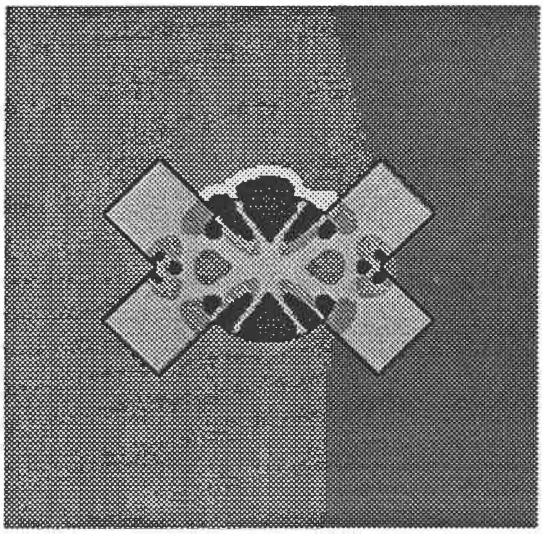
Figure 3.6 Image filtrée à différentes échelles. (a) Image non filtrée ($\sigma = 0.0$). (b) $\sigma = 1.95$. (c) $\sigma = 3.7$.



(a)



(b)



(c)

Figure 3.7 Représentation à différentes échelles. (a) $\sigma = 0.0$. (b) $\sigma = 1.95$. (c) $\sigma = 3.7$.

diminue elle se déplace vers les arêtes alors qu'une augmentation de σ la déplace vers les surfaces.

CHAPITRE 4

Conclusions et développements futurs

Nous avons présenté une approche multirésolution pour l'extraction d'arêtes et pour la segmentation d'images 3-D qui vise à produire une représentation multirésolution de l'image en termes d'arêtes bien localisées et de primitives de surface. Dans ce chapitre, nous dégageons des conclusions spécifiques sur les résultats fournis par cette approche. Nous suggérons par la suite quelques développements pouvant résulter de ce travail.

4.1 Conclusions.

4.1.1 Extraction d'arêtes.

La méthode d'extraction d'arêtes dans les images 3-D dont il est question au chapitre 1 constitue une approche multirésolution à double titre. D'une part, elle permet le choix de l'échelle à laquelle les arêtes significatives de l'image sont à extraire. Cette capacité de sélectivité d'échelle est offerte par l'usage de l'étape de filtrage Gaussien isotropique. D'autre part, elle constitue une approche multirésolution au sens conventionnel. En effet, à une échelle donnée elle utilise une technique de focalisation d'arêtes pour relocaliser les arêtes, détectées à cette échelle. Cette technique de focalisation constitue une approche multirésolution qui permet la mise en relation des éléments significatifs détectés aux différentes échelles.

La méthode donne des résultats très satisfaisants, surtout si l'on tient compte du fait que le niveau de complexité des images réelles est très diversifié et que le rapport signal-à-bruit des images synthétiques est très faible. Les arêtes fournies

par la méthode sont de nature différente (arêtes d'ordre 0 et d'ordre 1) et ont une largeur d'un pixel. En plus, elles sont bien localisées et correspondent bien aux arêtes physiques. Une bonne précision dans la localisation et la classification des arêtes est essentielle aux applications ultérieures (vision stéréoscopique, inférence de formes à partir des arêtes, ...). Toutes les étapes de la méthode, se prêtent bien à un traitement en parallèle. Ainsi, le temps calcul nécessaire pour extraire toutes les arêtes significatives à un niveau de résolution donné est la somme totale du temps calcul requis pour les détecter à ce niveau et du temps calcul indispensable à la localisation et la classification du segment d'arête le plus long.

Il est à noter cependant que les arêtes ne sont pas toutes fermées. Ceci est la conséquence de l'utilisation de la dérivée première d'une Gaussienne comme détecteur d'arêtes.

4.1.2 Segmentation d'images 3-D.

La méthode de segmentation d'images 3-D qui fait l'objet du chapitre 2 offre des performances très satisfaisantes. Tout d'abord, l'étape de filtrage Gaussien anisotropique permet l'élimination des variations haute fréquence des données de profondeur à l'intérieur des régions, tout en conservant l'acuité des arêtes interrégionales. Cette performance est essentielle à la conservation de la forme originale des surfaces situées dans le voisinage des discontinuités. De plus, l'emploi des filtres uniformes en cascade pour la synthèse du filtre Gaussien permet de réduire considérablement le temps calcul. Ensuite, la méthode de sélection de la meilleure fenêtre d'approximation, utilisée durant le calcul de courbures locales K et H, offre une bonne précision des valeurs de courbure calculées et, ultérieurement, une bonne identification des surfaces situées aux alentours des discontinuités. En effet, elle empêche le chevauchement de la fenêtre d'approximation sur ces dernières, rendant ainsi impossible l'interaction entre les différentes régions. Finalement,

la méthode de sélection automatique des seuils ε_K et ε_H rend la méthode de segmentation d'images 3-D indépendante de l'origine des données de profondeur et robuste au bruit. En effet, le choix difficile de ces seuils est crucial au fonctionnement adéquat de l'algorithme de segmentation. Comme le niveau de bruit n'est pas uniforme sur toute l'image (nous avons vu que le niveau de bruit sur les surfaces orientées horizontalement n'est pas le même que celui observé sur les surfaces fortement obliques) et n'est jamais le même dans des images provenant de caméras tridimensionnelles différentes, il est difficile de choisir à l'avance les valeurs de ces seuils. Un choix judicieux consiste alors à opter pour les valeurs qui produisent les meilleurs résultats de segmentation. Cette robustesse au bruit est aussi offerte par l'application initiale du filtrage Gaussien anisotropique sur l'image originale ainsi que par l'application finale de la procédure stochastique de relaxation d'étiquettes sur la carte d'étiquettes initiale.

La méthode est aussi multirésolution, c'est-à-dire qu'elle permet, comme dans le cas de l'extraction d'arêtes, le choix de l'échelle à laquelle les points de surface de l'image vont être classifiés. Encore une fois, cette capacité de sélectivité d'échelle est offerte par l'usage de l'étape de filtrage Gaussien anisotropique. De plus, la méthode pourrait être employée comme outil de segmentation dans un système de reconnaissance d'objets tridimensionnels. En effet, l'emploi des caractéristiques de surface riches, locales, et invariantes aux rotations, aux translations, et aux changements de paramétrisation des surfaces offre à la méthode la possibilité de décrire une large gamme de surfaces naturelles et d'être robuste aux problèmes causés par l'occlusion et aux changements du point d'observation.

4.1.3 Représentation multirésolution.

L'idée de combiner ensemble, à un niveau de résolution donné, les arêtes extraites et les primitives de surface identifiées est pertinente. En effet, différentes

régions adjacentes les unes aux autres peuvent avoir le même type de primitive. Ce cas se produit, par exemple, lorsque deux surfaces planes partagent la même discontinuité d'orientation ou lorsque deux surfaces de même type se chevauchent, créant ainsi une discontinuité de profondeur. Cette combinaison est donc d'abord essentielle à la partition adéquate des régions. Ensuite, le mariage rend la représentation finale riche en information utile: régions de l'image, information intérieure et frontalière de chacune d'elles. Finalement, combiner ensemble les arêtes et les primitives de surface est nécessaire pour une approche qui vise à produire une représentation multirésolution. Les arêtes physiques délimitent les différentes structures de la scène. Au niveau de résolution fin, la carte d'arêtes est dense et la quantité correspondante de régions est énorme. A un niveau de résolution grossier, la carte d'arêtes devient éparse, enlevant ainsi plusieurs contraintes frontalières et par conséquent rendant les régions correspondantes plus grossières.

4.2 Développements futurs.

La poursuite du travail pourrait se faire à trois niveaux:

1. Il est nécessaire de fermer les arêtes extraites. En effet, nous avons vu qu'il est essentiel que les arêtes soient fermées pour que l'algorithme de filtrage Gaussien anisotropique fonctionne. D'une part, l'acuité des discontinuités physiques sera conservée par le filtrage si les arêtes correspondantes dans l'image sont présentes. D'autre part, l'existence d'un espace vide entre deux parties d'une arête, même s'il est petit, causera l'estompage à cet endroit de la discontinuité physique correspondante. Le degré de sévérité de cet estompage croit avec l'augmentation de la résolution (l'échelle devient plus grossière).

2. Il est également nécessaire de résoudre le problème concernant la mauvaise classification des surfaces cylindriques orientées horizontalement. La façon de procéder suivante pourrait, d'après nous, remédier à ce défaut: d'abord, appliquer un seuillage à hystérésis sur les valeurs de courbure calculées afin d'en déduire, en chaque point de l'image, toutes les étiquettes $\{L_1, L_2, \ldots, L_8\}$ possibles; ensuite, dénotant par L l'étiquette en un point (x, y) de l'image ($L \in \{L_1, L_2, \ldots, L_8\}$), appliquer une procédure de relaxation d'étiquettes sur l'espace (x, y, L) pour établir la meilleure étiquette en chaque point.

3. Il serait intéressant de concevoir un modèle de la scène sous forme de graphe d'adjacence comme nous l'avons mentionné au chapitre 2.6, qui éventuellement pourrait servir à la reconnaissance d'objets tridimensionnels.

Bibliographie

[1] D.Marr, *Vision*, W.H.Freeman Co., San Francisco, 1982.

[2] D.Mar and Hildreth, "Theory of edge detection", *Proc. of the Royal Society of London* B 207, 1980, pp. 187–217.

[3] P.Perona and J.Malik, "Scale space and edge detection using anisotropic diffusion", *Proc. Workshop on Computer Vision*, Miami, 1987, pp. 16–22.

[4] M.Brady, J.Ponce, A.L.Yuille and H.Asada, "Describing surfaces", *MIT Artificial Intelligence Laboratory Memo* 822, 1985.

[5] J.Ponce and M.Brady, "Toward a surface primal sketch", *Proc. Int. Conf. on Robotics and Automation*, St-Louis, Missouri, 1985, pp. 420–425.

[6] T.J.Fan, G.Medioni and R.Nevatia, "Description of surfaces from range data using curvature properties", *Proc. Conf. on Computer Vision and Pattern Recognition*, Miami, 1986, pp. 86–91.

[7] T.J.Fan, G.Medioni and R.Nevatia, "Segmented descriptions of 3D surfaces", *IEEE Journal of Robotics and Automation*, RA-3, 1987, pp. 527–538.

[8] M.P.do Carmo, *Differential Geometry of Curves and Surfaces*, Prentice-Hall, Englewood Cliffs, New Jersey, 1976.

[9] M.M.Lipschutz, *Theory and Problems of Differential Geometry*, McGraw-Hill, New York, 1969.

[10] P.J.Besl and R.C.Jain, "Invariant surface characteristics for 3D object recognition in range images", *Computer Vision, Graphics, and Image Processing* 33, 1986, pp. 33–80.

[11] P.J.Besl and R.C.Jain, "Segmentation through symbolic surface descriptions", *Proc. Conf. on Computer Vision and Pattern Recognition*, 1986, pp. 77–85.

[12] F.Bergholm, "Edge focusing", *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-9, 1987, pp. 726–741.

[13] V.Torre and T.A.Poggio, "On edge detection", *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-8, 1986, pp. 147–163.

[14] E.C.Hildreth, "Edge detection", *MIT Artificial Intelligence Laboratory Memo* 858, 1985.

[15] S.Inokuchi and R.Nevatia, "Boundary detection in range pictures", *Proc. 5th Int. Conf. on Pattern Recognition*, Miami Beach, Florida, 1980, pp. 1301–1303.

[16] A.Mitiche and J.Aggarwal, "Detection of edges using range information", *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-5, 1983, pp. 174–178.

[17] F.Tomita and T.Kanade, "A 3D vision system: generating and matching shape descriptions in range images", *Proc. 1st Conf. on Artificial Intelligence Applications*, 1984, pp. 186–191.

[18] M.Herman, "Generating detailed scene descriptions from range images", *Proc. Int. Conf. on Robotics and Automation*, St-Louis, Missouri, 1985, pp. 426–431.

[19] B.Bhanu, S.Lee, C.C.Ho and T.Henderson, "Range data processing: representation of surfaces by edges", *Proc. 8th Int. Conf. on Pattern Recognition*, Paris, France, 1986, pp. 236–238.

[20] A.P.Witkin, "Scale-space filtering", *Proc. Int. Joint Conf. on Artificial Intelligence*, Karlsruhe, 1983, pp. 1019–1022.

[21] F.Sjoberg and F.Bergholm, "Extraction of diffuse edges by edge focusing", *Pattern Recognition Letters* 7, 1988, pp. 181–190.

[22] J.F.Canny, "Finding edges and lines in images", *MIT Artificial Intelligence Laboratory Technical Report* 720, 1983.

[23] J.F.Canny, "A computational approach to edge detection", *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-8, 1986, pp. 679–698.

[24] P.Saint-Marc and G.Medioni, "Adaptive smoothing for feature extraction", *Image Understanding Workshop*, 1988, pp. 1100–1113.

[25] H.Akima, "A method of bivariate interpolation and smooth surface fitting for irregularly distributed data points", *ACM Trans. Math Software* 4, 2, 1978, pp. 148–159.

[26] W.M.Wells, "Efficient synthesis of gaussian filter by cascaded uniform filters", *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-8, 1986, pp. 234–239.

[27] N.N.Abdelmalek and P.Boulanger, "Algebraic error analysis for surface curvatures of 3D range images", *Proceedings of Vision Interface '89*, London, Ont., 1989, pp. 29–36.

[28] N.Yokoya and M.D.Levine, "Range image segmentation based on differential geometry: a hybrid approach", *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-11, 1989, pp. 643–649.

[29] P.Boulanger and P.Cohen, "Stable estimation of a topographic primal sketch for range image interpretation", *IAPR Workshop on Computer Vision*, Tokyo, Japan, 1988, pp. 436–440.

[30] A.L.Yuille and T.Poggio, "Fingerprints theorems for zero-crossings", *MIT Artificial Intelligence Laboratory Memo* 730, 1984.

[31] J.Babaud, A.P.Witkin, M.Baudin, and R.O.Duda, "Uniqueness of the gaussian kernel for scale-space filtering", *Fairchild Technical Report* TR-645, 1984.

[32] J.J.Koenderink, "The structure of images", *Biological Cybernetics* 50, 1984, pp. 363–370.

[33] A.L.Yuille and T.Poggio, "Scaling theorems for zero-crossings", *MIT Artificial Intelligence Laboratory Memo* 772, 1984.

[34] T.Poggio, H.Voorhees, and A.L.Yuille, "A regularized solution to edge detection", *MIT Artificial Intelligence Laboratory Memo* 773, 1984.

[35] H.Asada and M.Brady, "The curvature primal sketch", *Proc. Workshop on Computer Vision*, New York, 1984, pp.8–17.

[36] P.Allen and R.Bajcsy, "Object recognition using vision and touch", *Proc. 9th Int. Joint Conf. on Artificial Intelligence*, 1985, pp. 1131–1137.

[37] P.K.Allen, "Sensing and describing 3D structure", *Proc. Int. Conf. on Robotics and Automation*, 1986, pp. 126–131.

[38] O.D.Faugeras and M.Hebert, "A 3D recognition and positioning algorithm using geometrical matching between primitive surfaces", *Proc. 8th Int. Joint Conf. on Artificial Intelligence*, 1983, pp. 992–1002.

[39] O.D.Faugeras, N.Ayache, and B.Faverjon, "A geometric matcher for recognition and positioning 3D rigid objects", *IEEE Comp. Soc. Conf. on Artificial Intelligence*, 1984, pp. 218–224.

[40] E.Hall, J.B.K.Tio, C.A.McPherson, C.S.Draper, and F.A.Sadjadi, "Measuring curved surfaces for robot vision", *IEEE Computer* 15, 12, 1982, pp. 42–54.

[41] M.Hebert and T.Kanade, "The 3D-Profile method for object recognition", *Proc. Conf. on Computer Vision and Pattern Recognition*, 1985, pp. 458–463.

[42] T.Pavlidis, *Algorithms for graphics and image processing*, Computer Science Press, 1982.

[43] A.Touzani and J.G.Postaire, "Mode detection by relaxation", *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-10, 1988, pp. 970–978.

[44] S.Peleg, "A new probabilistic relaxation scheme", *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-2, 1980, pp. 362–369.

ÉCOLE POLYTECHNIQUE DE MONTRÉAL
3 9334 00243288 6