

84.305

Sc N 84 / 91 A

DÉTECTION DE SURFACES PAR TRANSFORMÉE DE HOUGH EN VISION TRIDIMENSIONNELLE.



THESE

présentée et soutenue publiquement le 7 mai 1984

À L'UNIVERSITÉ DE NANCY I

pour l'obtention du grade de
DOCTEUR DE 3^{ème} CYCLE EN INFORMATIQUE

par

Yolande MULLER

devant la Commission d'Examen :

BIBLIOTHEQUE SCIENCES



D 095 113426 1

..... Jean-Paul HATON

..... Olivier FAUGERAS

Pierre MARCHAND

Alain MIRGAUX

Roger MOHR

**DÉTECTION DE SURFACES PAR TRANSFORMÉE
DE HOUGH EN VISION TRIDIMENSIONNELLE.**



THESE

présentée et soutenue publiquement le **7 mai 1984**

À L'UNIVERSITÉ DE NANCY I

pour l'obtention du grade de
DOCTEUR DE 3^{ème} CYCLE EN INFORMATIQUE

par

Yolande MULLER

devant la Commission d'Examen :

Président : Jean-Paul HATON
Examineurs : Olivier FAUGERAS
Pierre MARCHAND
Alain MIRGAUX
Roger MOHR

Je tiens à remercier ici :

Monsieur Jean-Paul HATON, Professeur à l'Université de Nancy I, de l'honneur qu'il me fait en présidant ce jury. Je lui suis très reconnaissante pour le soutien qu'il m'a apporté tout au long de ce travail et pour l'esprit amical qu'il a réussi à faire régner dans son équipe,

Monsieur Olivier FAUGERAS, Ingénieur de Recherche à l'INRIA, pour l'intérêt qu'il a porté à ce travail en acceptant de siéger à ce jury,

Monsieur Pierre MARCHAND, Professeur à l'Université de Nancy II, d'avoir eu la patience de lire mon manuscrit et pour l'ensemble des remarques et critiques qu'il m'a adressées. Son aide pour la résolution de certains problèmes géométriques m'a été d'un grand secours,

Monsieur Alain MIRGAUX, Maître-Assistant à l'Université de Nancy I, pour m'avoir aidée à résoudre les problèmes d'analyse numérique qui se sont posés au cours de ce travail,

Monsieur Roger MOHR, Professeur à l'Institut National Polytechnique de Lorraine, pour m'avoir proposé ce travail, pour l'intérêt qu'il a manifesté et les conseils qu'il m'a prodigués tout au long de cette thèse,

Enfin, je ne voudrais pas oublier de remercier l'Agence de l'Informatique pour son aide financière et surtout Madame Martine KUHLMANN qui a consacré tout son talent et son bon goût à la réalisation matérielle de cette thèse.

SOMMAIRE

INTRODUCTION	p.	1
CHAPITRE I : INTRODUCTION A LA VISION TRIDIMENSIONNELLE	p.	4
I .1 - LA TROISIEME DIMENSION DANS LE SYSTEME VISUEL HUMAIN	p.	7
I .2 - PERCEPTION AUTOMATIQUE DE LA TROISIEME DIMENSION	p.	10
a) Perspective et recouvrement	p.	10
b) Intensité lumineuse	p.	11
c) La stéréovision	p.	12
d) Les techniques de suréclairage	p.	13
e) Conclusion	p.	15
I .3 - LA VISION PAR ORDINATEUR	p.	16
I .4 - EXEMPLES DE SYSTEMES DE VISION TRIDIMENSIONNELLE	p.	18
I .5 - TRIDENT	p.	22
a) Le modèle	p.	22
b) Le fonctionnement du système TRIDENT	p.	25
I .6 - CONCLUSION	p.	28

CHAPITRE II : LA TRANSFORMÉE DE HOUGH p. 29

II.1 - PRINCIPE DE LA TRANSFORMÉE DE HOUGH p. 31

a) Exemple p. 31

b) Définition p. 33

c) Applications p. 34

II.2 - CHOIX DES PARAMETRES p. 36

II.3 - LE PROBLEME DES GRANDS ESPACES p. 39

II.4 - CONCLUSION p. 47

CHAPITRE III : RECHERCHE DE SURFACES p. 48

III.1 - EXEMPLES DE SYSTEMES EXISTANTS p. 50

III.2 - LA RECHERCHE DE PLANS p. 53

a) Première méthode p. 53

b) Deuxième méthode p. 54

III.3 - LA RECHERCHE DE QUADRIQUES p. 56

a) Principe p. 56

b) Mise en oeuvre p. 59

III.4 - CONCLUSION p. 63

CHAPITRE IV : RÉALISATION PRATIQUE ET RÉSULTATS p. 64

IV.1 - SYSTEME DE CREATION ET D'AFFICHAGE
D'IMAGES TRIDIMENSIONNELLES p. 66

a) Le modèle p. 66

b) La perspective p. 68

c) Elimination des parties cachées p. 70

d) L'intensité lumineuse p. 71

e) Mise en oeuvre p. 73

f) Conclusion p. 75

IV.2 - RESULTATS p. 76

a) Sur des images synthétiques p. 77

b) Sur des images réelles p. 87

IV.3 - CONCLUSION p. 93

CHAPITRE V : INTÉGRATION DANS UN SYSTÈME D'INTERPRÉTATION
D'IMAGES p. 94

V.1 - LES UTILISATIONS POSSIBLES ET LES
GAINS CORRESPONDANTS p. 96

a) Recherche de toutes les surfaces d'une
certaine nature p. 97

b) Recherche avec contraintes sur les
paramètres des surfaces p. 98

c) Recherche dans une fenêtre de l'image
p. 99

d) Segmentation en régions et recherche
dans une région p. 100

e) Recherche à grands pas p. 103

f) Recherche de surfaces évidentes	p. 103
g) Problème des bords des surfaces	p. 104
h) Recherche des points appartenant à une surface connue par ses paramètres	p. 106
i) Conclusion	p. 111
V .2 - UN EXEMPLE DE STRATEGIE D'UTILISATION	p. 112
V .3 - CONCLUSION	p. 116
CONCLUSION	p. 117
BIBLIOGRAPHIE	p. 120

INTRODUCTION

L'existence de capteurs d'informations tridimensionnelles a ouvert de nouveaux domaines d'investigation en interprétation de scènes par ordinateur [DUD - 76]. Ces capteurs ont, par ailleurs, posé de nouveaux problèmes : tel le choix des primitives. Dans le cadre du projet TRIDENT, qui a pour objectif d'analyser et d'interpréter des scènes naturelles, ce choix s'est porté sur des surfaces telles que des plans, des sphères, des cylindres de révolution, des cônes de révolution et des ellipsoïdes de révolution. Nous nous sommes donc intéressés à la détection de surfaces dans des images de distances.

Différentes approches ont été suggérées pour résoudre ce problème ; la voie a été ouverte par [DUD - 76]. La détection de plans et de quadriques a ensuite été explorée [OSH - 79], [DAN - 81]. Notre approche consiste à utiliser la transformée de Hough. Cette technique a déjà été utilisée pour rechercher des surfaces [BAL - 81b], [HEB - 82] mais nous proposons une nouvelle méthode : elle a l'avantage de détecter toute surface ayant une équation cartésienne polynomiale et elle permet à l'utilisateur de diriger sa recherche vers un type particulier de surfaces (exemple : recherche de cônes verticaux) en rejetant les points inintéressants. Toutes sortes de contraintes sur les paramètres des surfaces recherchées (exemple : rechercher des sphères de rayon 5) peuvent être prises en compte. Elles réduisent l'ensemble des solutions possibles et accélèrent donc le processus de recherche.

Le premier chapitre introduit la vision tridimensionnelle. Il expose différentes méthodes d'acquisition de la troisième dimension et présente quelques systèmes de vision dont TRIDENT.

Le deuxième chapitre est consacré à l'étude de la transformée de Hough et de ses principaux sous-problèmes : le choix des espaces des paramètres et la recherche de points d'accumulation.

L'utilisation de la transformée de Hough pour la recherche de plans et de quadriques est expliquée dans le troisième chapitre. Il présente également quelques exemples de systèmes de détection de surfaces existant.

La réalisation pratique et la présentation des résultats font l'objet du quatrième chapitre.

Enfin, dans le cinquième chapitre, nous étudions comment intégrer efficacement notre système de détection de surfaces dans un système d'interprétation d'images.

CHAPITRE I
INTRODUCTION A LA
VISION TRIDIMENSIONNELLE

La vision par ordinateur est une branche de l'Intelligence Artificielle qui prend de plus en plus d'importance. Son but est de comprendre une scène à partir d'une image.

On peut considérer une image comme une fonction f de \mathbb{R}^2 dans X . Si l'ensemble d'arrivée de cette fonction est un sous-ensemble à deux éléments de X , l'image est dite binaire. Dans une image de niveaux de gris, X correspond à \mathbb{R}^+ et $f(x,y)$ désigne la luminance du point (x,y) . Dans une image couleur, $f(x,y)$ est une combinaison de la teinte, de la saturation et de la luminance du point (x,y) . Enfin, dans une image des distances (ou image tridimensionnelle), $f(x,y)$ représente la distance de l'observateur au point (x,y) de la scène (figure I.1).

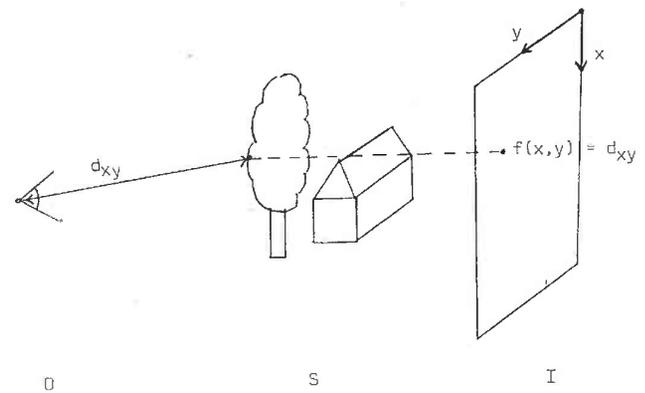


Figure I.1 : L'image des distances I d'une scène S vue par un observateur O .

La vision par ordinateur est un problème difficile et dont la résolution peut se résumer comme ceci : un certain nombre de caractéristiques sont extraites de l'image de la scène à analyser. Elles sont ensuite mises en correspondance avec l'univers des connaissances (il s'agit de connaissances fournies au système sur la scène à analyser). Ces caractéristiques ou primitives ont été pendant longtemps bidimensionnelles (contours, régions). L'avènement des capteurs de distance facilite l'introduction de primitives tridimensionnelles (surfaces, volumes). L'évolution de la deuxième dimension vers la troisième dimension a ouvert de nouveaux axes de recherche en traitement d'images.

Avant d'aborder le problème de la vision, nous essayerons de comprendre comment l'homme perçoit la troisième dimension. Nous verrons ensuite quelles sont les techniques qui nous permettent de la capter effectivement.

I.1 - LA TROISIEME DIMENSION DANS LE SYSTEME VISUEL HUMAIN

Plusieurs caractéristiques permettent au système visuel humain de percevoir la troisième dimension. La taille et la perspective sont deux de ces caractéristiques. Dans la figure I.2, c'est la taille qui nous permet de dire que l'arbre de droite est plus éloigné que l'arbre de gauche.

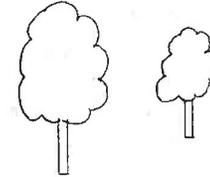


Figure I.2 : L'importance de la taille dans l'évaluation des distances.

La notion de perspective nous indique que la figure I.3 représente un cube.

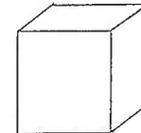


Figure I.3 : L'importance de la perspective dans l'évaluation des distances.

La variation des intensités lumineuses, les ombres et la texture nous apportent également des indications sur le relief. La figure I.4 nous montre l'importance de l'intensité lumineuse.

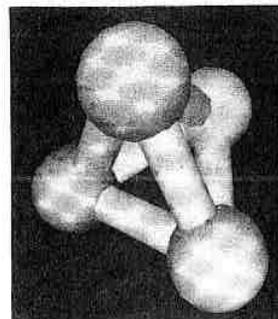


Figure I.4 : L'importance de l'intensité lumineuse dans la perception du relief [NEW - 79].

Le recouvrement partiel ou l'occultation est un autre élément important dans l'estimation de la distance. Dans la figure I.5, le cube cache une partie du cylindre ; il est donc plus proche de l'observateur que le cylindre.

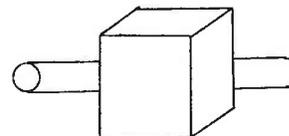


Figure I.5 : L'importance du recouvrement dans l'évaluation des distances.

Dans le cas de la vision rapprochée, la perception binoculaire permet la perception du relief indépendamment de tous les paramètres précédents. Chaque oeil perçoit une image de la scène. La mise en correspondance de ces images nous fournit une indication sur la distance : c'est la vision stéréoscopique [JUL - 71].

Tous ces éléments ajoutés à nos connaissances sur notre univers nous permettent de reconnaître les objets situés dans notre environnement, d'analyser et d'interpréter les scènes auxquelles nous sommes confrontés.

I .2 - PERCEPTION AUTOMATIQUE DE LA TROISIEME DIMENSION

a) Perspective et recouvrement

Les premiers travaux concernant la reconnaissance d'objets tridimensionnels sont ceux de ROBERTS [ROB - 65]. Son programme analyse l'image d'une scène et identifie tous les objets visibles. Il détermine leur orientation et position dans l'espace tridimensionnel. Les scènes qu'il interprète contiennent des objets polyédriques appelés "blocs". Le modèle utilisé ne contient que quelques objets tridimensionnels simples tels que des cubes, des parallélépipèdes rectangles et des prismes.

La première étape de son programme consiste à extraire des lignes (courbes ou droites) de l'image de départ. Dans la deuxième étape, un objet est mis en correspondance avec l'image. Le choix d'un tel objet est guidé par des caractéristiques simples telles que le nombre d'arêtes. Des rotations, changements d'échelle et projections permettent sa mise en correspondance avec l'image. Si cette correspondance est bonne, l'objet est reconnu ; sa position et sa taille sont alors connues.

De nombreux travaux sur le "monde des blocs" ont succédé à ceux de ROBERTS. Il s'agit essentiellement de ceux de GUZMANN (1968), FALK (1972), HUFFMAN et CLOWES (1971), WALTZ (1972), SHIRAI (1973), MACKWORTH (1973-74) et KANADE (1979-80). Une synthèse de ces travaux est proposée dans [COH - 82]. Ils utilisent tous la perspective et le recouvrement pour percevoir la troisième dimension. Leur principal inconvénient est de ne pas pouvoir se généraliser à des scènes naturelles (figure I.6) mais c'est encore le cas maintenant pour la plupart des systèmes de vision.

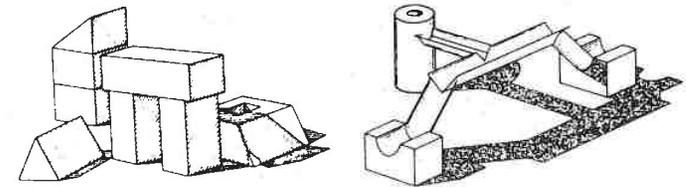


Figure I.6 : Exemples de scènes reconnues [BAL - 82].

b) Intensité lumineuse

L'intensité lumineuse fournit une indication sur la troisième dimension, principalement dans le cas de surfaces courbes et lisses. HORN [HOR - 75] a mené les premiers travaux à ce sujet. L'orientation d'une surface exprimée sous la forme $z = f(x,y)$ est donnée par

$$p = \frac{\partial f}{\partial x} \quad \text{et} \quad q = \frac{\partial f}{\partial y}$$

Sous certaines conditions, l'intensité I d'un point d'une image peut être calculée à l'aide de l'orientation (p,q) de la surface à laquelle appartient ce point. Nous pouvons donc définir une fonction $f:(p,q) \rightarrow I$. L'idée de HORN est d'inverser cette fonction. Mais, en général, la fonction inverse ne donne pas une orientation unique. Des contraintes supplémentaires doivent alors être introduites. On supposera par exemple que la surface admet un paramétrage continuellement différentiable. Une autre solution consiste à utiliser une nouvelle image de la scène prise sous un éclairage différent. On obtient ainsi une nouvelle

fonction $f_2: (p, q) \rightarrow I$. Les deux fonctions f_1^{-1} et f_2^{-1} fournissent une solution unique pour l'orientation. Il s'agit alors de "stéréovision photométrique".

c) La stéréovision

La stéréovision est une technique qui consiste à analyser plusieurs images bidimensionnelles (non identiques) d'une même scène. Comme nous l'avons vu, elle existe dans le système visuel humain. Il s'agit de mettre en correspondance tous les points images qui désignent le même point physique sur la scène. On en déduira une indication sur le relief.

Cette méthode pose deux problèmes importants : trouver des caractéristiques faciles à reconnaître dans une image, mettre en correspondance les caractéristiques de deux images.

Le premier problème est souvent résolu en choisissant les contours des objets présents dans l'image ; comme les contours correspondent à des changements brusques d'intensités, on peut les localiser rapidement. GRIMSON [GRI - 80] a utilisé de telles caractéristiques pour réaliser, sur les conclusions de MARR et POGGIO, un programme effectif de perception stéréoscopique. Il s'est inspiré de la vision humaine.

Une solution au deuxième problème consiste à formuler la mise en correspondance en terme de recherche de chemin optimal dans un graphe. Cette approche, qui ne prend en compte aucune référence à la vision humaine, est étudiée dans notre laboratoire [MOH - 84]. L'ensemble des

couples de points qui peuvent être mis en correspondance constitue les points d'un graphe ; les arcs sont valués par la qualité de chaque correspondance. Cette méthode n'a été testée actuellement que sur des stéréogrammes artificiels ; les résultats sont néanmoins prometteurs.

d) Les techniques de suréclairage

De nombreux systèmes utilisent des techniques de suréclairage pour capter la troisième dimension. Le plan de lumière en est un exemple. C'est une méthode simple relevant de la triangulation (figure I.7). Des plans de lumière sont projetés successivement sur l'ensemble d'une scène. Lors de chaque projection, une caméra capte une image de la scène. La zone éclairée, qui apparaît sur cette image, nous fournit des indications sur le relief. Cette méthode peut être utilisée directement pour rechercher des surfaces.

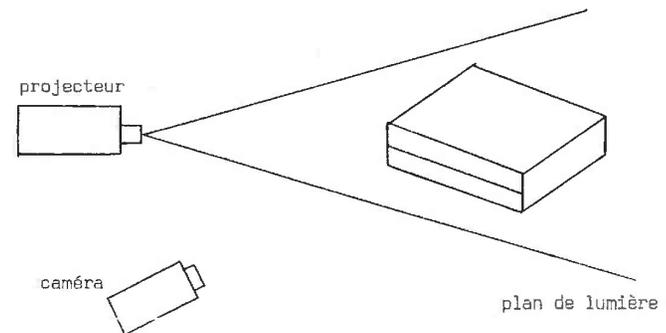


Figure I.7 : Perception de la troisième dimension par des plans de lumière.

Une autre méthode de suréclairage consiste à éclairer un ou plusieurs points d'une scène. Cette méthode a été utilisée par trois laboratoires français (INRIA, ENSMN, LETI) pour réaliser des maquettes de capteurs.

Le procédé de l'INRIA [AYA - 82] utilise un éclairage obtenu par un laser défléchi par un miroir. Le laser balaye la dimension x . L'objet à analyser est placé sur un support mobile, le long de l'axe y . Deux barres de détecteurs récupèrent le signal lumineux et permettent d'obtenir la valeur de z (figure I.8).

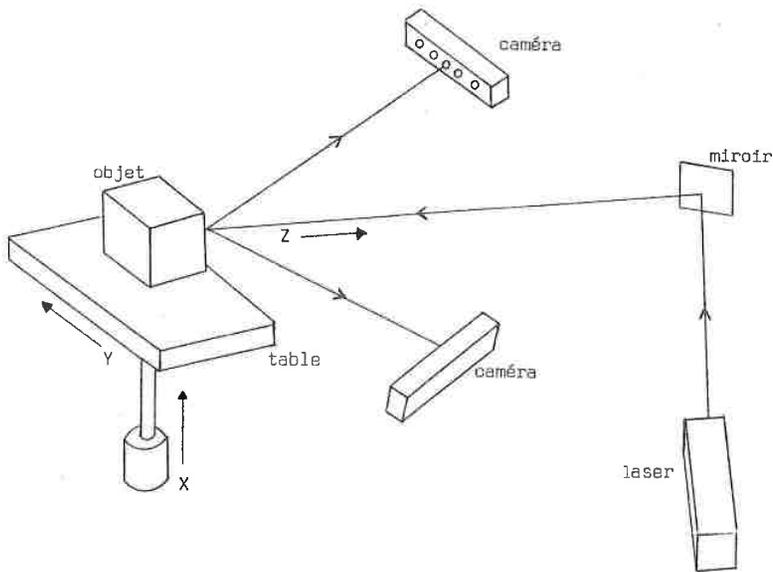


Figure I.8 : Le capteur de l'INRIA

Ce capteur privilégie la précision des résultats contrairement à ceux de l'ENSMN et du LETI qui privilégient la vitesse d'acquisition. C'est le capteur de l'INRIA qui nous fournit des données réelles pour tester notre système de détection de surfaces.

e) Conclusion

Les besoins de plus en plus importants d'informations tridimensionnelles ont rendu nécessaire le développement de systèmes de perception et d'acquisition du relief. Les méthodes utilisées sont variées et dépendantes du type d'application envisagée. Leurs avantages et inconvénients se mesurent principalement au niveau de la vitesse de saisie, de la précision des résultats et de la complexité du système.

I.3 - LA VISION PAR ORDINATEUR

Comprendre une scène à partir d'une image est une tâche élémentaire pour un être vivant, mais il est très difficile de construire un système de vision par ordinateur aussi performant que le système visuel humain. En effet, ce dernier est très complexe car il intègre une foule d'informations diverses. Pour certaines d'entre elles, nous connaissons un modèle mathématique mais la mise en oeuvre reste souvent complexe. C'est le cas pour les ombrages lorsque la source d'éclairage n'est pas ponctuelle ou que les surfaces ne sont pas lambertiennes. Pour d'autres informations, nous n'avons aucune connaissance ; c'est le cas pour le visage : comment modéliser la forme "visage" ? quels paramètres de ce modèle nous permettent de distinguer un visage d'un autre ?

On peut ordonner la complexité des problèmes de vision selon plusieurs axes ; en voici quelques exemples :

- la nature de l'image à traiter :
 - . image binaire
 - . image de niveaux de gris
 - . image multispectrale
 - . image des distances

- la visibilité des objets à reconnaître :
 - . objet complètement visible
 - . objet partiellement caché

- la forme des objets à reconnaître :
 - . objet de forme unique (exemple : pièces de fonderie)
 - . objet de forme variable mais modélisé (exemple : marteau)
 - . objet de forme non modélisé (exemple : visage)

- la position et la description exacte des objets :
 - . objet dont on connaît la position et la description exacte (exemple : mesure de l'épaisseur d'un objet)
 - . objet dont on connaît soit la position, soit la description exacte mais pas les deux à la fois (exemple : problème de localisation d'objets)
 - . objet dont on ne connaît ni la position ni la description exacte (exemple : reconnaissance de scènes naturelles).

I.4 - EXEMPLES DE SYSTEMES DE VISION TRIDIMENSIONNELLE

Un certain nombre de systèmes de vision tridimensionnelle se limitent à la reconnaissance d'objets dans une scène [OSH - 83], [FAN - 82], [FIS - 83], [BAL - 81b]. Le système développé par OSHIMA et SHIRAI en fait partie [OSH - 83]. Il s'intéresse à la reconnaissance d'objets formés de plans et de régions courbes ; il comporte deux phases : une phase d'apprentissage et une phase de reconnaissance. Dans la phase d'apprentissage, des objets connus sont présentés au système qui en fait une description en termes de surfaces et de relations entre ces surfaces. Un modèle est ainsi créé pour chaque objet. Si plusieurs vues sont nécessaires à la description d'un objet donné, plusieurs modèles seront créés. Dans la phase de reconnaissance, le système décrit une scène inconnue, puis il essaie de mettre cette description en correspondance avec des modèles existants.

D'autres systèmes permettent, en plus de la reconnaissance d'objets, d'estimer la position et l'orientation de chacun des objets trouvés. C'est le cas du système développé par FAUGERAS et HEBERT [FAU - 83].

Nous nous intéressons plus particulièrement aux systèmes de vision généraux, non limités à des domaines d'application particuliers. Le système d'interprétation de scènes naturelles de DHTA, KANADA et SAKAI [SAK - 76] en est un exemple. Une tendance actuelle consiste à définir des structures générales. Dans cet esprit, un système tel que ACRONYM [BRO - 78] présente une solution intéressante. Un objet est

décrit par un graphe relationnel à plusieurs niveaux . Au niveau supérieur, des relations "spatiales" précisent de façon sommaire les liens entre les parties de l'objet, puis à un niveau inférieur, des primitives de type "éléments de cône généralisé" décrivent la structure de ces objets. Un cône généralisé est décrit par un axe qui n'est pas nécessairement rectiligne et par une section qui n'est pas forcément constante. La figure I.9 montre quelques exemples d'objets engendrés par un cône généralisé.

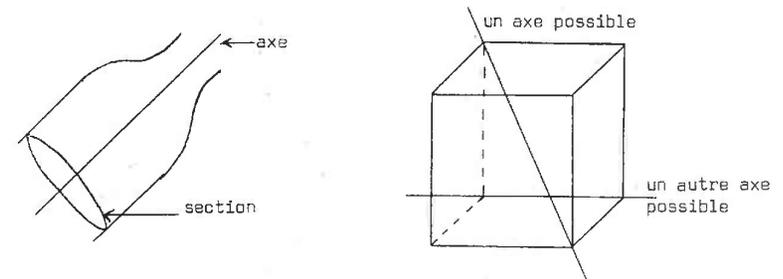


Figure I.9 : Exemples d'objets engendrés par des cônes généralisés.

ACRONYM est un système guidé par le modèle. Il cherche à prédire l'aspect de l'image à l'aide du modèle. Il en résulte un graphe d'observation. De l'image sont extraites un certain nombre d'informations permettant de créer le graphe de l'image. La mise en correspondance de ces deux graphes aboutit au graphe d'interprétation qui peut suggérer de nouvelles prédictions ou segmentations (figure I.10).

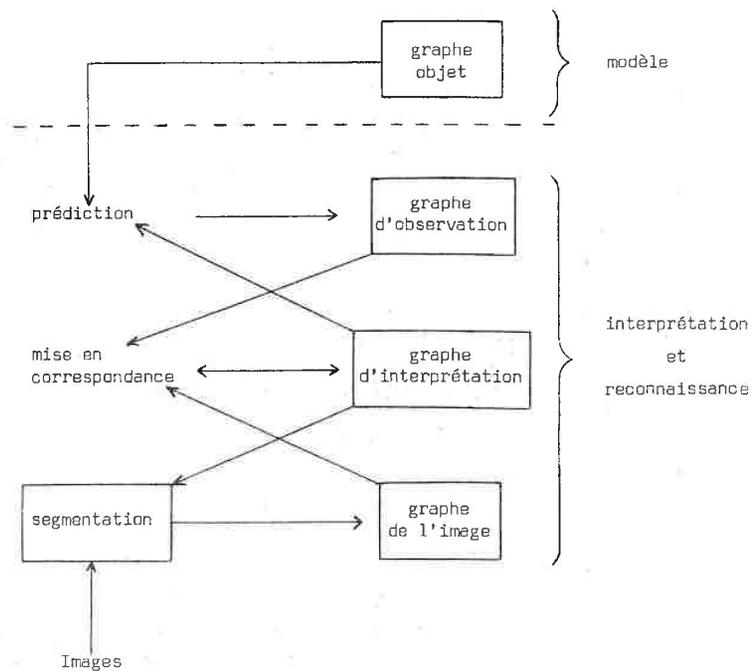


Figure I.10 : Fonctionnement général du système ACRONYM.

Une autre approche est utilisée par HANSON et RISEMAN dans le système VISIONS [HAN - 80]. Ils utilisent une représentation sous forme de niveaux (figure I.11).

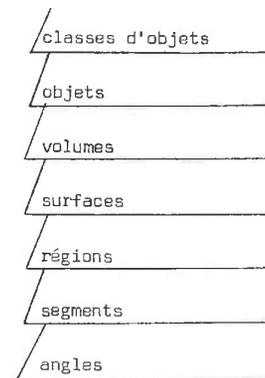


Figure I.11 : Les différents niveaux de représentation d'une scène dans VISIONS.

Ces niveaux se répartissent en deux catégories : les concepts 3D liés au monde physique (classe d'objets, objets, volumes, surfaces) et les concepts 2D liés aux projections des précédents sur un plan image (régions, segments, angles). Ces derniers réalisent l'interface entre le monde physique et une image de celui-ci.

Citons aussi le système SEER de FREUDER [FRE - 77] où les connaissances sont exprimées à travers un réseau exploré de façon ascendante et descendante. SEER illustre bien la complexité d'un système général intégrant tous les niveaux, y compris la perception.

I.5 - TRIDENT

TRIDENT est un projet à long terme mené dans notre équipe et tendant à la réalisation d'un système d'interprétation d'images pour une large classe d'application [MAS - 84]. La structure et le fonctionnement de ce système découlent de deux idées fondamentales : utiliser pleinement l'information contenue implicitement dans le modèle et extraire de la scène à analyser le maximum d'informations (couleur, texture, ...) avec la meilleure qualité possible. Avant d'expliquer le fonctionnement général de ce système, nous allons présenter le modèle utilisé.

a) Le modèle

Le modèle utilisé est inspiré d'ACRONYM sans en avoir la généralité. Il se présente sous la forme d'un ensemble de règles. Les objets sont décrits en termes de surfaces, à l'aide de trois types de primitives : des polygones plans, des secteurs circulaires et des cylindres de révolution (figure I.12).

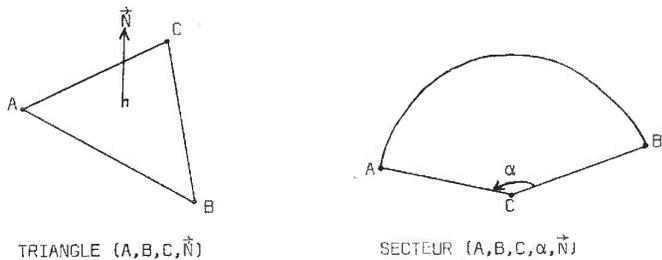


Figure I.12 : Exemples de primitives.

Des volumes de base (cube, pyramide, ...) ont été introduits pour faciliter la description des objets du modèle. Ils sont composés d'un ensemble de primitives (figure I.13).

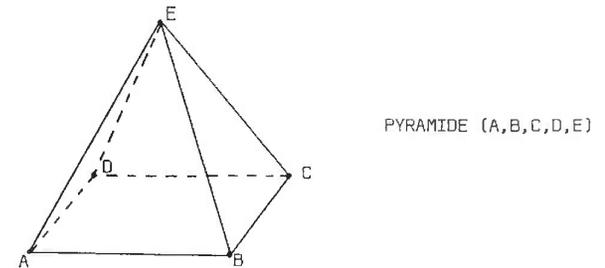
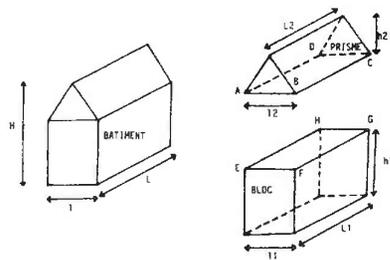


Figure I.13 : Exemple de volume de base.

A chaque objet, volume ou primitive est associée une liste d'attributs représentant ses paramètres (longueur, diamètre, ...) et ses caractéristiques (couleur, texture, ...). Des contraintes sur les attributs précisent le mode de construction de l'objet, son échelle ainsi que sa position dans un référentiel absolu. Un modèle d'univers est donc constitué d'un ensemble de règles dont certaines définissent des objets (figure I.14), alors que d'autres précisent les relations topographiques ou spatiales entre les différents objets présents (exemple : le téléphone est POSE SUR le bureau).



BATIMENT (H, I, L, couleur) --> BLOC (h1, I1, L1, coul), E, F, G, H)
 PRISME (h2, L2, A, B, C, D)

- | | |
|---|--|
| $H = h1 + h2$
$I = I1$
$L > L1$
couleur = coul | définition des attributs de
la forme décrite |
| $L2 < L1 < L2 + I$
$I2 < I1 < I2 + I$
$I1 < L1 / 2$
$5 < I2 < 7$ | contraintes sur les valeurs
des attributs |
| $A_y = E_y$ et $A_x = E_x$
$C_y = G_y$ et $C_x = G_x$ | règles d'assemblage des sous-formes dans
un repère absolu |

- le membre gauche indique la forme qui est décrite
- le membre droit donne les sous-formes utilisées pour construire le membre gauche
- les attributs associés aux formes apparaissent entre parenthèses

Figure I.14 : Un exemple de règle de description d'un objet [MAS - 84].

b) Le fonctionnement du système TRIDENT

Le système TRIDENT est constitué de différents niveaux : le prétraitement, l'extraction de renseignements, la segmentation, le compilateur de description et l'interpréteur (figure I.15).

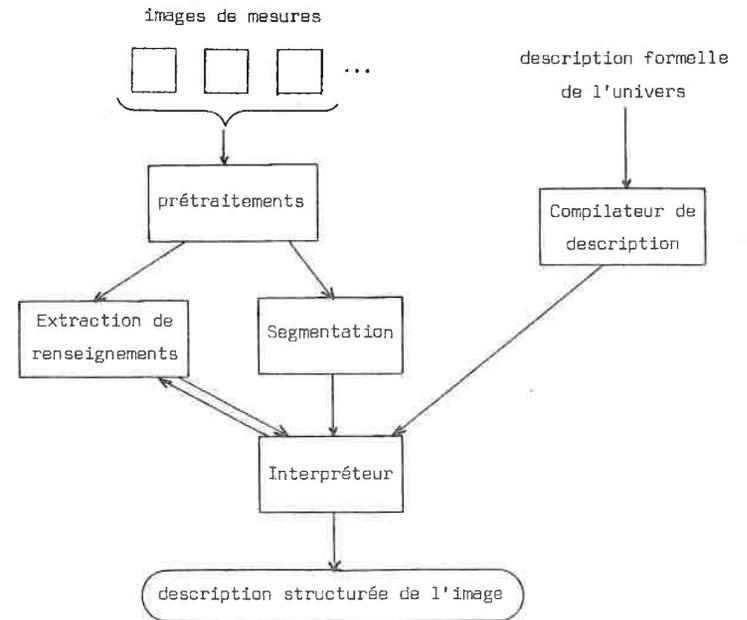


Figure I.15 : Schéma du système TRIDENT.

Les traitements sont facilités par la présence de plusieurs images de mesures sur la scène comme, par exemple, les niveaux de gris, les distances par rapport au capteur, les couleurs, les textures et les contours.

La phase de prétraitement a pour but d'améliorer la qualité des images numérisées. Les techniques utilisables sont nombreuses ; en voici quelques unes : le lissage, le seuillage, la modification d'histogramme.

La segmentation aboutit à un graphe des régions présentes dans l'image de la scène. Ce graphe est obtenu à l'aide d'un algorithme de regroupement de régions qui opère par fusions, en utilisant plusieurs images de mesures [ZAR - 83]. L'utilisation d'un critère de fusion élaboré et de paramètres de contrôle de la segmentation limite les cas de mauvaises segmentations.

L'extraction de renseignements se fera principalement en utilisant le système de détection de surfaces décrit dans le chapitre III. Il fournira les réponses aux requêtes formulées par l'interpréteur. Ces requêtes peuvent être du type : "*Est-ce que la région R est sphérique ? Y a-t-il un cylindre d'axe vertical dans cette image ?*". Il est donc nécessaire d'envisager un système de détection de surfaces qui peut prendre en compte des directives de recherche. C'est cet objectif qui nous a guidé vers l'utilisation de la transformée de Hough. Nous appliquerons cette technique à des images de distances.

Le compilateur de description a pour but de vérifier la cohérence du modèle proposé et d'en tirer le maximum d'informations pour nourrir l'interpréteur. Un algorithme de propagation de contraintes est développé actuellement pour vérifier la cohérence des données numériques. Il s'appuie sur les idées de BROOKS [BRO - 81] et SHOSTAK [SHO - 77].

L'interpréteur est formé de deux parties : l'analyseur et le mécanisme de déductions. L'analyseur est chargé d'étiqueter les régions de l'image avec les noms des formes qu'elles représentent, en se fondant sur les règles de description du modèle. La stratégie utilisée est du type ascendante/descendante. Elle permet, d'une part, de démarquer l'analyse par n'importe quelle forme caractéristique de l'image et, d'autre part, de prendre en compte le contexte d'une forme. Le mécanisme de déductions a été réalisé sous forme d'un système à règles de productions. Chaque déduction entraîne un certain nombre d'actions à réaliser par l'analyseur.

Une première maquette du système a été réalisée par F. ZAROLI [ZAR - 83] ; elle a permis de tester différentes stratégies d'analyse et d'observer le comportement de quelques outils de vision.

I.6 - CONCLUSION

La troisième dimension a ouvert de nouveaux axes de recherche en traitement d'images. Le premier concerne la perception et l'acquisition de cette nouvelle dimension. Différentes techniques ont été développées et des maquettes de capteurs tridimensionnels ont été réalisées. Ces capteurs nous fournissent des images de distances, nouvelles sources d'information pour un grand nombre de systèmes de vision. Des techniques classiques peuvent être appliquées à ces images de distances : c'est le cas de la transformée de Hough que nous allons utiliser pour détecter des surfaces.

CHAPITRE II
LA TRANSFORMÉE DE HOUGH

Nous voulons rechercher, dans une image, une forme F appartenant à une famille \mathcal{F} . Par exemple, on considère l'ensemble \mathcal{F} des sphères et on recherche une sphère F que l'on détermine par son centre et son rayon.

L'idée de la transformée de Hough est d'inverser ce problème. On recherche parmi l'ensemble des paramètres possibles ceux qui "*conviennent le mieux*". L'ensemble des paramètres possibles est appelé "espace des paramètres".

Cette méthode permet de rechercher toute surface analytique et, plus généralement, toute forme dans une famille de formes paramétrées. Pour nous, le terme "analytique" signifie ayant une équation cartésienne polynomiale. Les surfaces qui nous intéressent sont les plans et les surfaces du second degré (exemples : sphères, cylindres, cônes, ellipsoïdes).

Elle permet également de prendre en compte des informations sur la surface à trouver. Ces informations peuvent porter sur un ou plusieurs paramètres de la surface. Elles réduisent l'espace des paramètres et donc accélèrent le processus de recherche. Par exemple, si l'on recherche un cylindre quelconque, l'espace des paramètres est de dimension 5, mais si l'on recherche un cylindre de rayon R , cet espace ne sera plus que de dimension 4.

II.1 - PRINCIPE DE LA TRANSFORMEE DE HOUGH

La transformée de Hough est une méthode permettant de rechercher des formes géométriques dans un ensemble de points de \mathbb{R}^p (en général, $p = 2$ ou $p = 3$).

Pour nous, une forme géométrique F est un ensemble de points de \mathbb{R}^p défini par une équation $F(P, X) = 0$ où P est le paramètre de cette forme ($P \in \mathbb{R}^n$).

Une forme est dans une image si, et seulement si, pour une certaine valeur du paramètre P , il y a suffisamment de points X de l'image vérifiant $F(P, X) = 0$.

a) Exemple : (recherche de droites dans un espace à 2 dimensions)

Commençons par expliciter la méthode sur un exemple simple :

Soit un espace de départ à 2 dimensions contenant 6 points X_1, X_2, X_3, X_4, X_5 et X_6 (figure II.1).

Recherchons, dans cet espace, toutes les droites d'équation $y = ax + b$.

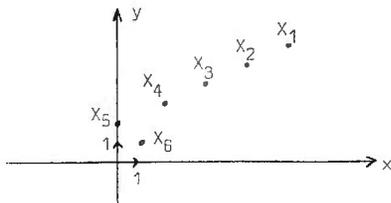


Figure II.1 : Espace de départ

L'espace des paramètres est l'espace à deux dimensions formé par les couples (a, b) . A chaque point X_i de coordonnées (x_i, y_i) de l'espace de départ est associé un ensemble E_i de points de l'espace des paramètres tel que :

$$y_i = ax_i + b \text{ ou encore } b = -x_i a + y_i.$$

Les ensembles E_i sont donc des droites (figure II.2).

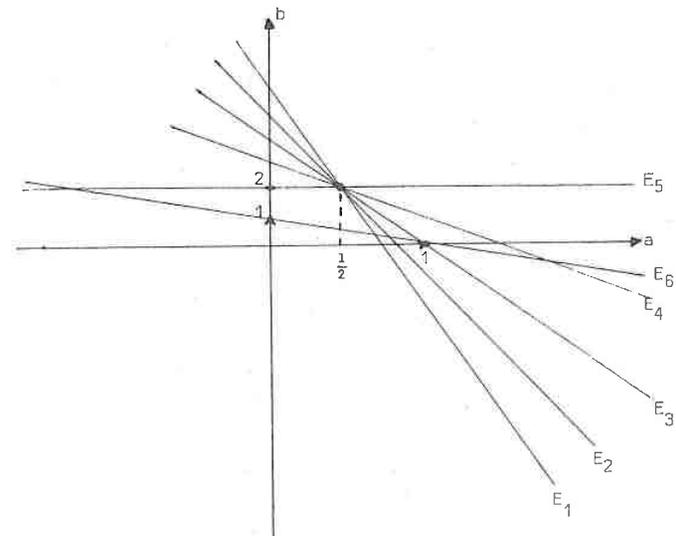


Figure II.2 : Espace des paramètres

Cinq de ces droites ont une intersection commune en $a = 1/2$ et $b = 2$. Les cinq points de l'espace de départ leur correspondant appartiennent donc à une même droite d'équation $y = \frac{1}{2}x + 2$.

Une droite est dans l'espace de départ si, et seulement si, elle est commune à un suffisamment grand nombre d'ensembles E_i . Cette notion de "suffisamment grand nombre" est évidemment impossible à formaliser ; tout ce que l'on peut dire pour l'instant c'est que ce nombre doit être supérieur à la dimension de l'espace des paramètres. Nous discuterons de ce problème au paragraphe II.3.

Le problème d'origine qui consistait à trouver des droites a été remplacé par un problème mathématiquement équivalent consistant à trouver des points d'intersections d'ensembles de paramètres.

b) Définition

La transformée de Hough consiste à passer de l'ensemble des valeurs X_i parmi lesquelles on recherche une forme $F(P, X_i)$ à l'ensemble des paramètres P de cette forme. Appelons n le nombre de paramètres de la forme F recherchée. A chaque point X_i de l'espace de départ est associé un ensemble E_i de points de \mathbb{R}^n tel que $F(P, X_i) = 0$ pour tout P de E_i .

Le principe de la transformée de Hough est d'étudier les ensembles E_i de l'espace des paramètres. S'il existe un nombre suffisant de ces ensembles E_i qui ont une intersection commune en un point P de \mathbb{R}^n alors la forme F de paramètre P est présente dans l'espace de départ ; les points qui la composent sont les antécédents X_i des ensembles E_i trouvés.

c) Applications

La transformée de Hough, qui a tout d'abord été appliquée à des droites et des cercles, a ensuite été généralisée à des courbes plus complexes, analytiques [SHA - 78], [SLO - 80] puis non analytiques [BAL - 81 a], [SLO - 80], [YAM - 81].

Pour représenter une courbe non analytique, il existe plusieurs méthodes. Celle de BALLARD [BAL - 81 a] consiste à représenter une courbe à l'aide d'une table. Cette table est construite en choisissant un point de référence ou centre et en calculant pour chaque point de la courbe la direction du gradient en ce point. La table associe alors à chaque direction du gradient la ou les positions possibles du point de référence. Si on permet à la courbe de changer d'orientation et d'échelle, il faut introduire deux paramètres supplémentaires : l'angle de rotation et le facteur de changement d'échelle. Des expériences ont été réalisées avec cette méthode sur des images aériennes [SLO - 80].

Une autre représentation possible consiste à choisir un point de la courbe et à mémoriser les coordonnées de tous les vecteurs reliant les différents points de la courbe à ce point particulier [YAM - 81].

L'espace de départ a été, dans un premier temps, une image binaire [DUD - 72]. Les points significatifs auxquels on applique la transformée de Hough sont localisés rapidement dans une telle image. Dans de nombreux cas, il s'agit des points de contours. Ces points contiennent, dans certains cas, assez d'informations pour reconnaître un objet.

La méthode a ensuite été étendue à des images de niveaux de gris. Les points les plus souvent utilisés sont, comme pour les images binaires, les points de contours.

Pour appliquer la transformée de Hough à des images de distances, des données intermédiaires peuvent être introduites : les normales représentées sur la sphère gaussienne [HEB - 82] (voir au chapitre III, paragraphe III.1) ou une image segmentée en morceaux de plans [BAL - 81 b].

Nous avons choisi d'utiliser l'image des distances comme espace de départ. L'espace des paramètres est celui de la surface recherchée. A chaque point de l'image de départ sont associés les paramètres d'une surface. La nature de cette surface et ses paramètres sont calculés pour chaque point en fonction des points voisins. La méthode utilisée sera développée dans le chapitre III. Les ensembles E_1 de l'espace des paramètres sont donc réduits à des points et la recherche d'intersections d'ensembles est transformée en une recherche de points d'accumulation. La notion de point d'accumulation est une approximation discrète de celle de l'analyse mathématique. Ainsi, même dans un espace à grandes dimensions, cette recherche est facilement réalisable.

II.2 - CHOIX DES PARAMETRES

Les paramètres sont fonction de la famille de formes recherchée. Pour une famille donnée, plusieurs espaces des paramètres sont possibles. Par exemple, dans le cas de droites, l'équation $y = ax + b$ suggère un espace des paramètres infini à deux dimensions a et b . Cependant, on remarque que si tous les angles de droites sont équiprobables, on aura autant de valeurs de a entre $[-1, 1]$ qu'à l'extérieur de ce segment. Pour cette raison, la paramétrisation $x \cos \theta + y \sin \theta + r = 0$ est préférée.

La taille de l'espace des paramètres croît très vite : 3 pour un plan, 4 pour une sphère, 5 pour un cylindre de révolution, 6 pour un cône de révolution et 7 pour un ellipsoïde de révolution. Nous avons choisi des paramètres dont les valeurs semblent, intuitivement, être équidistribuées dans leur domaine de définition ; ce sont les suivants :

- pour un plan :
 - . 2 angles φ et θ pour l'orientation de la normale au plan (figure II.3),
 - . la distance d du plan à l'origine.
- pour une sphère :
 - . les 3 coordonnées cartésiennes de son centre,
 - . son rayon.
- pour un cylindre :
 - . 3 angles (μ, θ, ω) et une distance e pour caractériser l'axe de révolution (figure II.4),
 - . son rayon.

- pour un cône :

- . les 3 coordonnées cartésiennes de son sommet (qui est aussi son centre),
- . 2 angles Ψ et Θ pour l'orientation de l'axe de révolution (figure II.3),
- . un angle α pour caractériser le demi-angle au sommet.

- pour un ellipsoïde :

- . les 3 coordonnées cartésiennes de son centre,
- . 2 angles Ψ et Θ pour l'orientation de l'axe de révolution (figure II.3),
- . 2 distances a et c pour caractériser l'excentricité.

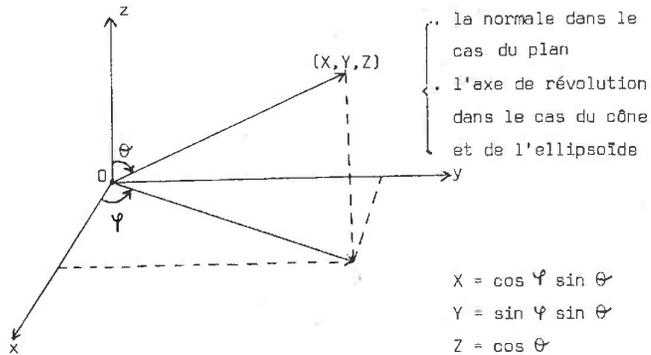
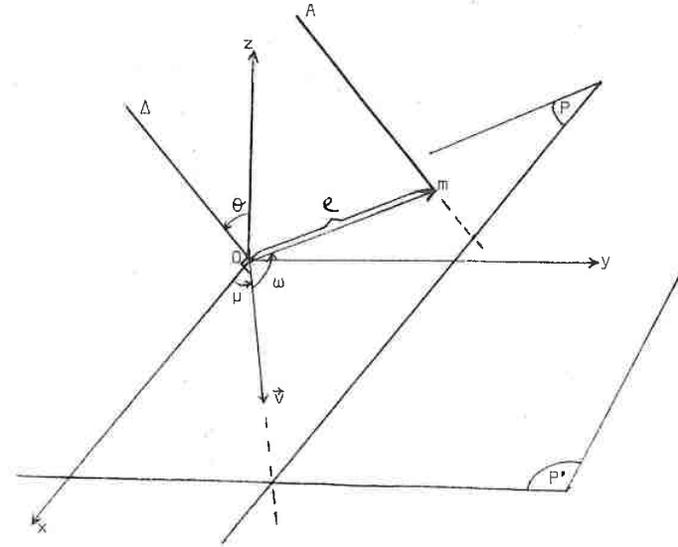


Figure II.3 : Les angles Ψ et Θ



- A : l'axe de révolution du cylindre
- Δ : droite parallèle à A et passant par O
- P : plan perpendiculaire à Δ
- P' : plan $O \times y$
- m : intersection de A et de P
- v : intersection de P et de P'

Figure II.4 : Les paramètres utilisés pour caractériser l'axe de révolution d'un cylindre

II.3 - LE PROBLEME DES GRANDS ESPACES

La transformée de Hough nous amène à rechercher des intersections d'ensembles (ou, dans notre cas particulier, des points d'accumulation) dans l'espace des paramètres de la forme recherchée. Dans la plupart des applications, cet espace est de dimension 2 et il est possible de le discrétiser. La méthode la plus simple consiste à le subdiviser de façon uniforme en petites portions d'espace, assimilables à des points.

Cette méthode conduit à de sérieux problèmes d'encombrement mémoire dans le cas d'espace de dimension supérieure ou égale à 3. C'est pourquoi un certain nombre de méthodes de subdivision dynamique ont été proposées [O'R - 81], [SLO - 81].

Appelons "cellule" le pavé d'espace délimité par une valeur minimum et une valeur maximum sur chacune des dimensions de l'espace des paramètres. O'ROURKE [O'R - 81] a proposé des espaces quantifiés dynamiquement. Les cellules sont divisées en deux (pas forcément en leur milieu) ou fusionnées selon le nombre et la répartition des éléments leur appartenant. Une structure hiérarchique telle un arbre binaire s'impose pour connaître rapidement les voisins d'une cellule. Les divisions sont simples à réaliser mais les fusions sont difficiles à mettre en oeuvre.

C'est pourquoi SLOAN [SLO - 81] a proposé une autre solution plus rigide mais plus facile à utiliser : les pyramides quantifiées dynamiquement. Une cellule est divisée en n cellules de même taille si l'espace des paramètres est de dimension n . Un grand nombre de cellules inintéressantes peuvent donc être créées. Cette perte de place est compensée par une mise en oeuvre plus simple.

Nous avons préféré une nouvelle solution fondée sur le principe "diviser pour régner" : elle est rapide, simple à mettre en oeuvre et écarte immédiatement les cellules inintéressantes. Une cellule est divisée en deux cellules de même taille, tour à tour, suivant chacune de ses dimensions, tant qu'elle n'est pas assimilable à un point et qu'elle est intéressante : elle est assimilable à un point lorsque sa taille est inférieure à un seuil donné T ; elle est intéressante tant qu'elle contient un nombre suffisant de points. Soit NP le nombre minimum de points suffisants. Ce seuil est choisi comme suit : dans l'espace de départ, il désigne le nombre minimum de points qui doivent appartenir à une surface pour que l'existence de celle-ci soit reconnue. Des deux seuils T et NP dépend la précision du résultat. Le choix de ces seuils est fonction des surfaces recherchées et de la qualité des images à traiter. Nous proposerons quelques exemples lors des résultats expérimentaux. L'algorithme correspondant à cette méthode est le suivant, pour une cellule donnée :

DIVISE (cellule)

début

- choix d'une dimension i
- division en 2 de la cellule suivant la dimension i
=> 2 nouvelles cellules $S1$ et $S2$

- Si nombre d'éléments de $S1 > NP$ alors
 Si taille de $S1 > T$ alors DIVISE ($S1$)
 sinon SAUVEGARDE ($S1$)

fsi

fsi

- Si nombre d'éléments de $S2 > NP$ alors
 Si taille de $S2 > T$ alors DIVISE ($S2$)
 sinon SAUVEGARDE ($S2$)

fsi

fsi

fin

La dimension i suivant laquelle s'effectue la division correspond, tour à tour, à chacune des dimensions de l'espace des paramètres. Une dimension n'est prise en compte que si sa taille minimum n'est pas encore atteinte.

Appliqué à l'espace des paramètres tout entier, cet algorithme trouve tous les points d'accumulation. Nous connaissons alors les surfaces qui existent dans l'image de départ et les points leur appartenant.

Cette méthode a l'avantage de ne subdiviser finement que les cellules intéressantes. Elle est très rapide : si nous voulons diviser un espace des paramètres de dimension 7 en 100 parties égales sur chacune de ses dimensions, nous avons à considérer 100^7 cellules. Si un seul point d'accumulation est présent dans cet espace, il est localisé après $\log_2 (10^{14}) \approx 47$ divisions.

La division dichotomique ne se justifie que si la probabilité d'avoir une surface est la même dans chacune des deux nouvelles cellules ainsi obtenues. Comme nous l'avons dit au paragraphe II.2, les paramètres ont été choisis tels que cette probabilité semble être équidistribuée.

Cette méthode pose néanmoins un problème lorsque le découpage nous amène à diviser un point d'accumulation en son milieu. Lorsque l'espace des paramètres est de dimension n , un point d'accumulation peut, dans le pire des cas, être divisé en 2^n (figure II.5).

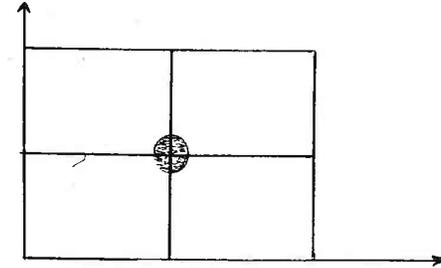


Figure II.5 : Dans un espace à 2 dimensions, un point d'accumulation peut être divisé en 4

Si la cellule correspondant à ce point d'accumulation contient tout juste un nombre suffisant de points pour être conservée (soit NP points) alors les 2^n cellules résultant des n divisions successives ne contiennent, dans le pire des cas, plus que $NP/2^n$ points. Elles sont alors ignorées et le point d'accumulation est perdu.

Pour remédier à ce problème, nous sauvegardons non seulement les cellules intéressantes mais toutes les cellules contenant au moins $NP/2^n$ points. C'est le rôle de la procédure SAUVEGARDE. Puis nous effectuons un traitement spécial pour fusionner plusieurs cellules.

Dans un premier temps, cette méthode consistait à fusionner une cellule sauvegardée contenant moins de NP points avec $(2^n - 1)$ de ses voisins de même taille. Si le nouvel espace ainsi obtenu contenait

au moins NP points alors la fusion avait effectivement lieu et le point d'accumulation était conservé sinon il était oublié. Lors des expériences, nous avons constaté qu'un point d'accumulation peut recouvrir plus que 2 cellules élémentaires sur une même dimension. Dans ce cas, la méthode précédente ne nous permet pas de localiser l'ensemble du point d'accumulation. C'est pourquoi nous avons adopté une nouvelle solution ; elle consiste à fusionner toutes les cellules élémentaires qui sont adjacentes (figure II.6).

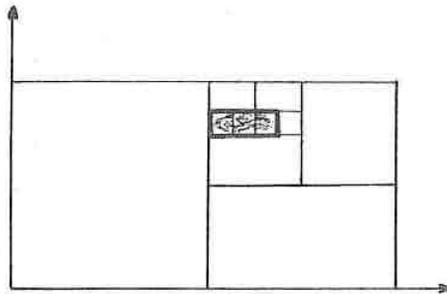


Figure II.6 : Fusion de 3 cellules élémentaires

Si la cellule résultant de ces fusions contient plus de NP points, elle est conservée. Elle correspond à une surface dans l'espace de départ, son centre de gravité nous donne les paramètres de cette surface.

Cette méthode de subdivision nous permet de détecter tous les points d'accumulation de l'espace des paramètres. Nous en déduisons la liste des surfaces leur correspondant ainsi que les points de l'espace de départ leur appartenant.

La figure II.7 illustre le découpage d'un espace à 2 dimensions. Cet espace est une projection de l'espace des paramètres d'un ellipsoïde (de dimension 7). Les deux dimensions visualisées correspondent aux deux premières coordonnées du centre de l'ellipsoïde.

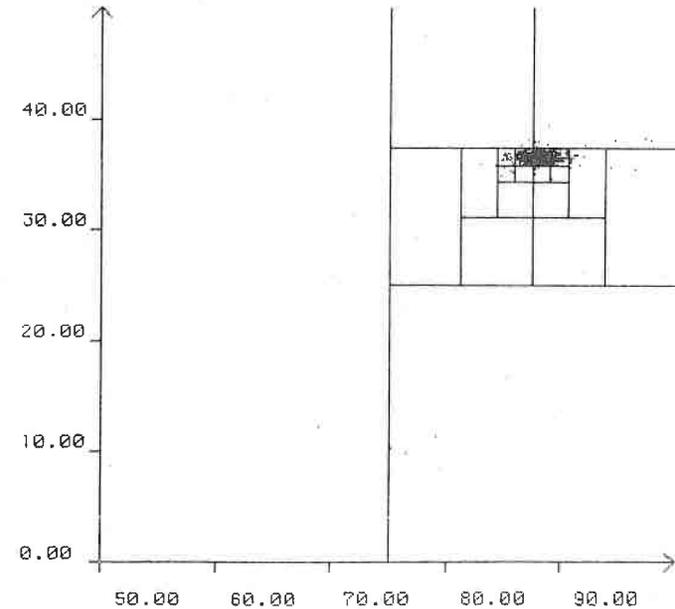


Figure II.7 : Découpage dans l'espace des 2 premières coordonnées du centre d'un ellipsoïde

Les divisions successives et la fusion des cellules élémentaires adjacentes nous ont permis de localiser le point d'accumulation. Il contient plus de 4 000 points. Dans notre image de départ, il existe donc probablement un ellipsoïde dont les deux premières coordonnées du centre sont 88 et 37. Cette illustration pourrait nous inciter à ne travailler que dans des espaces de petite dimension par projections successives. Malheureusement ces projections sont souvent accompagnées de bruit. L'origine de ce bruit peut être soit la mauvaise qualité de l'image étudiée, soit le manque de précision au niveau des calculs. Dans l'espace des paramètres tout entier, ce bruit est très dispersé et donc ignoré mais sa projection dans un espace à 2 dimensions peut créer de petites accumulations entraînant des découpages superflus. La figure II.8 en est un exemple : elle illustre le découpage d'un espace à 2 dimensions dans lequel ont été projetées les deux premières coordonnées du centre d'un ellipsoïde et d'un cône. En plus des deux points d'accumulation cherchés, il apparaît une petite zone d'accumulation qui a entraîné quelques divisions inutiles. Il ne semble donc pas intéressant de vouloir réduire la taille des espaces par projections successives.

Une suite de divisions dichotomiques et des fusions de cellules élémentaires adjacentes nous permettent de localiser rapidement des points d'accumulation dans des espaces de toutes tailles.

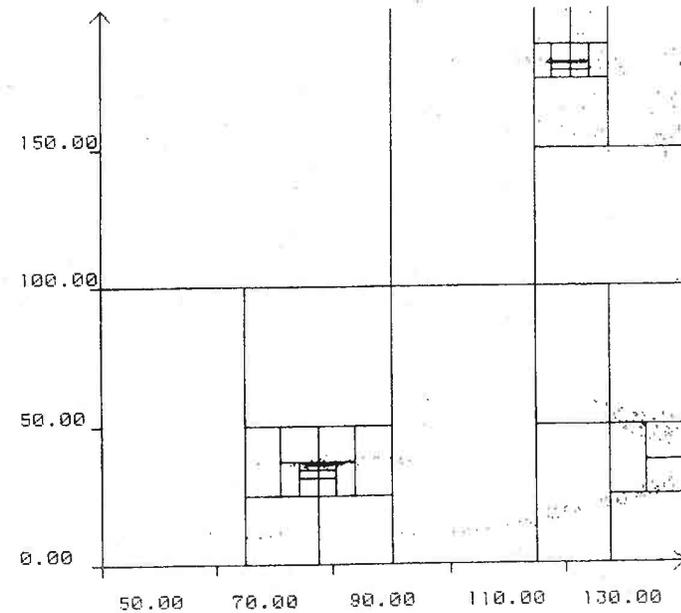


Figure II.8 : Découpage dans l'espace des 2 premières coordonnées du centre d'un ellipsoïde et d'un cône

II.4 - CONCLUSION

La transformée de Hough nous permet de transformer notre problème de recherche de surfaces en un problème plus simple de recherche de points d'accumulation dans l'espace des paramètres des surfaces. Ce nouveau problème est résolu par une suite de divisions dichotomiques suivie d'éventuelles fusions. Le choix des paramètres justifie cette subdivision dynamique par dichotomie. Un point d'accumulation dans l'espace des paramètres indique la présence d'une surface dans l'image, son centre de gravité fournit les paramètres correspondants.

CHAPITRE III
RECHERCHE DE SURFACES

Contours et régions sont les éléments primitifs de systèmes de vision [PEL - 82], [FEL - 74]. Dans une image tridimensionnelle, les régions correspondent à des surfaces. Nous présenterons quelques exemples de systèmes de recherche de surfaces, puis nous proposerons une nouvelle solution qui pourra prendre en compte des directives de recherches fournies par un utilisateur.

III.1 - EXEMPLES DE SYSTEMES EXISTANTS

Nous ne présenterons, dans ce paragraphe, que quelques exemples caractéristiques de systèmes de détection de surfaces. Ils ont tous pour but de segmenter une image en plans et en quadriques.

La méthode utilisée par M. OSHIMA et Y. SHIRAI [OSH - 79] consiste à diviser une image en petits éléments de surfaces assimilables à des plans. Ces éléments de surface sont ensuite fusionnés en régions élémentaires à peu près planes. Ces régions sont divisées en trois catégories : des plans, des régions courbes (c'est-à-dire non planes) et des régions non définies (il s'agit de régions pour lesquelles il n'a pas été possible de décider si elles sont planes ou non). Les régions courbes et les régions non définies qui sont adjacentes sont fusionnées en de grandes régions courbes dont la nature est déterminée par une recherche de surfaces, au sens des moindres carrés. Une scène peut ainsi être décrite en terme de plans et de quadriques. Cette méthode a été testée sur des scènes contenant des polyèdres et des objets courbes simples (sphère, cylindre, ...). Les résultats obtenus sont satisfaisants.

C. DANE et R. BAJCSY [DAN - 81] ont choisi une approche totalement différente. Ils ignorent les informations spatiales dans un premier temps. Ils associent à chaque point de l'image tridimensionnelle la normale à la surface en ce point. Ces normales sont mémorisées dans une structure de données appelée "approximation de l'image gaussienne étendue" [SMI - 79]. Dans cette structure de données, toutes les normales ayant la même orientation sont confondues en une seule et même normale. La longueur de cette normale est proportionnelle à l'aire des surfaces qu'elle représente. Cette information permet d'identifier un certain nombre de classes de surfaces. Par exemple, un plan correspond à une normale de

très grande taille, un cylindre à un arc de cercle formé de normales de taille moyenne et une sphère à une région de normales de petite taille. Un regroupement de régions est effectué dans l'approximation de l'image gaussienne étendue. Chaque région trouvée correspond à un type de surface. Pour séparer des surfaces de même type, non adjacentes, les informations spatiales sont nécessaires. Un histogramme spatial est alors créé pour chaque région trouvée. De nouvelles régions sont formées à partir de ces histogrammes ; chacune de ces régions correspond à une surface. L'équation de chacune des surfaces ainsi trouvées est obtenue par un algorithme au sens des moindres carrés.

M. HEBERT et J. PONCE [HEB - 82] ont décrit une méthode qui se rapproche de celle qui vient d'être décrite. La première étape de cette méthode consiste à utiliser la sphère gaussienne comme structure de données intermédiaire. Sur cette sphère de rayon unité sont mémorisées les orientations des normales associées à chaque point de l'image. Elles ont toute la même longueur car leur fréquence d'apparition ne nous intéresse pas. Toutes les normales ayant la même orientation correspondent à un même point sur la sphère. Les primitives choisies peuvent être représentées de façon simple sur cette sphère gaussienne. Un point d'accumulation correspond à un plan, un cercle de rayon 1 à un cylindre et un cercle de rayon inférieur à 1 à un cône. La transformée de Hough est utilisée pour rechercher ces trois types de primitives. Les régions en résultant sont séparées lors de la deuxième étape en régions connexes. Les petites régions sont éliminées. Les informations manipulées durant cette étape sont les coordonnées des points de l'image. Un algorithme de recherche de sous-graphe connexe dans un graphe est utilisé pour la séparation en régions connexes. La détermination des paramètres restant constitue la dernière étape.

Ces différentes méthodes segmentent une image en plans et en quadriques mais elles ne permettent pas de prendre en compte des directives de recherche (exemple : rechercher une sphère de rayon R). Un utilisateur (qui pourrait être un système d'interprétation d'images) est susceptible de formuler des demandes plus ou moins précises (exemples : rechercher toutes les surfaces, rechercher tous les plans horizontaux,...) ou de fournir des renseignements supplémentaires pour guider le système vers une solution particulière (exemple : recherche de cylindres ayant un rayon compris entre R_1 et R_2 , ...). C'est typiquement le processus qui intervient dans un système de vision automatique. Nous avons donc envisagé un nouveau système de détection de surfaces répondant à ces besoins. Il utilise la transformée de Hough dans ce but.

III.2 - LA RECHERCHE DE PLANS

Nous avons choisi d'utiliser la transformée de Hough pour pouvoir introduire des contraintes sur les paramètres du ou des plans recherchés. Nous souhaitons que les ensembles de l'espace des paramètres soient réduits à des points pour faciliter la recherche d'intersections. Nous devons donc associer à chaque point de l'image les paramètres (φ, θ, d) d'au maximum un plan. Nous avons envisagé deux méthodes différentes. La première consiste à associer un plan à chaque point de l'image. Par contre, dans la deuxième méthode, seuls les points qui appartiennent "probablement" à un plan ont une image dans l'espace des paramètres.

a) Première méthode

A chaque point d'une image, nous pouvons associer sa normale calculée sur un petit voisinage. Nous en déduisons les deux paramètres φ et θ . L'équation du plan appliquée au point considéré nous fournit le troisième paramètre qui est d , la distance du plan à l'origine. Ainsi, nous pouvons associer à chaque point de l'image les paramètres du plan tangent à la surface en ce point. Le nombre de points de l'espace des paramètres est alors identique au nombre de points de l'image. Si un seul plan est présent dans l'image, les points de ce plan créent, dans l'espace des paramètres, une zone d'accumulation. Tous les autres points de l'image engendrent des points isolés et dispersés qui peuvent entraîner des divisions inutiles lors de la subdivision dynamique. C'est pourquoi nous n'avons pas retenu cette méthode.

b) Deuxième méthode

L'équation générale d'un plan peut être exprimée à l'aide d'une fonction $f(x,y,z) = 0$ avec $f(x,y,z) = Ax + By + Cz + D$ où A, B, C et D sont des coefficients à valeurs réelles et (x,y,z) les coordonnées d'un point de l'image à traiter.

Nous voulons déterminer la valeur des coefficients pour chaque point (x,y,z) . En appliquant l'équation $f(x,y,z) = 0$ à un point et à au moins 4 de ses voisins (car il y a 4 coefficients à trouver), nous obtenons un système d'équations qui peut être résolu au sens des moindres carrés. Nous en déduisons A, B, C et D . Pour ne pas trouver la solution évidente $(0, 0, 0, 0)$, il faut fixer l'un des coefficients; nous avons fixé C à 1. En effet, dans une image, les plans parallèles à l'axe de vision sont réduits à une droite. Ils ne peuvent donc jamais être pris en compte. Or ces plans sont tels que $C = 0$. Donc en fixant C à 1, nous éliminons seulement des plans qui, de toute façon, ne sont pas visibles.

En résolvant notre système au sens des moindres carrés, nous trouvons toujours un plan qui passe par un point, même si ce point appartient en réalité à une autre surface (par exemple, une sphère). Nous devons donc distinguer, parmi les paramètres trouvés, ceux qui correspondent effectivement à un plan présent dans l'image de ceux qui représentent des plans fictifs. Cette sélection se fait en testant l'erreur au sens des moindres carrés. Si elle est faible (de l'ordre de 10^{-6}), nous supposons qu'il existe effectivement un plan passant par le point considéré. Nous calculons alors ses paramètres (φ, θ, d) à partir des coefficients trouvés, puis nous les mémorisons dans l'espace des paramètres.

Nous n'avons ainsi, dans l'espace des paramètres, que les points intéressants, soit ceux qui correspondent probablement à des plans. Leur nombre est en général beaucoup moins important que le nombre de points de l'image. De plus, ils sont presque toujours regroupés en quelques zones d'accumulation très nettes. Ces zones sont très facilement localisées car elles sont entourées de très peu de bruit. Nous avons donc retenu cette solution pour rechercher des plans.

III.3 - LA RECHERCHE DE QUADRIQUES

Dans l'esprit de la transformée de Hough, nous cherchons à associer à chaque point d'une image les paramètres de la quadrique à laquelle appartient ce point. Ces paramètres sont mémorisés dans l'espace des paramètres de la quadrique. Une recherche de points d'accumulation dans cet espace nous indique s'il existe une quadrique de ce type dans l'image (voir chapitre II).

a) Principe

Les surfaces du second degré peuvent être décrites à l'aide d'une équation

$$f(x,y,z) = 0 \quad (1)$$

$$\text{avec } f(x,y,z) = Ax^2 + A'y^2 + A''z^2 + 2Byz + 2B'zx + 2B''xy + 2Cx + 2C'y + 2C''z + D$$

où A, A', A'', B, B', B'', C, C', C'' et D sont des coefficients à valeurs réelles et (x,y,z) les coordonnées d'un point de l'image à traiter.

Si, pour un point donné (x,y,z), nous arrivons à déterminer la valeur des différents coefficients, nous pourrions en déduire la nature de la quadrique qui passe par ce point, ainsi que ses paramètres.

L'équation (1) est linéaire par rapport aux coefficients qui, dans notre cas, sont les inconnues. Ainsi, en appliquant cette équation à un point et à au moins 10 de ses voisins (car il y a 10 coefficients à trouver), nous aurons un système d'équations qui peut être résolu au sens des moindres carrés (pour prévenir d'éventuelles erreurs sur les coordonnées des points). Nous connaîtrons alors l'équation de la quadrique qui passe par ce point et ses voisins. Pour éviter de trouver la

solution évidente où tous les coefficients sont nuls, nous avons fixé l'un d'entre eux, ce qui a réduit le nombre d'inconnues à 9. Nous avons fixé D à 1. Nous éliminons ainsi les quadriques dont la distance à l'origine est nulle. Il suffit de translater l'origine du repère associé à l'image pour éviter de rencontrer de telles quadriques.

Pour chaque point de l'image, nous connaissons donc l'équation de la quadrique qui passe par ce point et un certain nombre de ses voisins. Si cette quadrique est de la nature de celle que nous recherchons, nous déterminons ses paramètres et nous les marquons dans l'espace des paramètres, sinon ce point n'a pas d'image. La nature et les paramètres de la quadrique sont déduits de l'équation (1) [CAS - 65].

Si $\{x_0, y_0, z_0\}$ sont les coordonnées cartésiennes d'un centre O' d'une surface (S) du second degré et si $f(x,y,z) = 0$ est l'équation de (S), son équation dans un système d'axes équipollents et d'origine O' sera

$$\begin{aligned} f(x'+x_0, y'+y_0, z'+z_0) &= f(x_0, y_0, z_0) + x'f'_x(x_0, y_0, z_0) \\ &+ y'f'_y(x_0, y_0, z_0) + z'f'_z(x_0, y_0, z_0) + x'^2f''_{x^2}(x_0, y_0, z_0) \\ &+ y'^2f''_{y^2}(x_0, y_0, z_0) + z'^2f''_{z^2}(x_0, y_0, z_0) \\ &+ 2x'y'f''_{xy}(x_0, y_0, z_0) + 2x'z'f''_{xz}(x_0, y_0, z_0) \\ &+ 2y'z'f''_{yz}(x_0, y_0, z_0) = 0 \end{aligned}$$

d'après l'identité de TAYLOR.

O' sera un centre si, et seulement si, $f(x'+x_0, y'+y_0, z'+z_0)$ ne change pas quand on change simultanément x' en -x', y' en -y' et z' en -z'.

$$\text{D'où les conditions : } f'_x(x_0, y_0, z_0) = 0, f'_y(x_0, y_0, z_0) = 0 \text{ et } f'_z(x_0, y_0, z_0) = 0.$$

Pour trouver un centre, il suffit donc de résoudre ce système de 3 équations à 3 inconnues. Ainsi, pour les quadriques à centre (sphère, cône, ellipsoïde), nous connaissons dès maintenant trois de leurs paramètres. Pour le cylindre, nous connaissons un des points de la droite centre. Nous remarquons qu'en traduisant l'origine du repère pour la faire coïncider avec le centre ou un des centres de la quadrique cherchée, l'équation (1) devient :

$$Ax^2 + A'y^2 + A''z^2 + 2B'yz + 2B''xz + 2B'''xy + D' = 0 \quad (2)$$

L'équation (2) peut aussi s'écrire sous forme matricielle :

$$\begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} A & B'' & B' \\ B'' & A' & B \\ B' & B & A'' \end{bmatrix}^t \begin{bmatrix} x & y & z \end{bmatrix} + D' = 0$$

Nous remarquons, grâce à cette écriture, qu'en diagonalisant la matrice $\Delta = \begin{bmatrix} A & B'' & B' \\ B'' & A' & B \\ B' & B & A'' \end{bmatrix}$

nous obtenons l'équation réduite de la quadrique

$$Sx^2 + S'y^2 + S''z^2 + D' = 0 \quad (3)$$

La matrice Δ étant réelle symétrique, elle est toujours diagonalisable. Pour trouver ses valeurs propres S, S' et S'' et ses vecteurs propres V, V' et V'' , nous utilisons la méthode de JACOBI. Le passage de l'équation (2) à l'équation (3) se traduit par un changement de repère par rotation. L'orientation du nouveau repère est donnée par les vecteurs propres V, V' et V'' , son origine reste (x_0, y_0, z_0) .

De l'équation (3), nous pouvons déduire la nature de la surface car nous connaissons la forme des équations réduites des quadriques que nous recherchons. Des calculs simples nous donneront les paramètres.

Le principal avantage de cette méthode consiste à ne mémoriser dans l'espace des paramètres que les points intéressants. Par exemple, si nous recherchons des cônes, seuls les points pour lesquels l'équation (1) correspond à celle d'un cône ont une image dans l'espace des paramètres associé au cône. Les points d'accumulation de l'espace des paramètres sont, en général, peu bruités. La recherche est donc facile et rapide.

b) Mise en oeuvre

Nous nous sommes intéressés plus particulièrement à la recherche de sphères, de cylindres de révolution, de cônes de révolution et d'ellipsoïdes de révolution. Leurs paramètres ont été définis dans le chapitre II, au paragraphe 2. Nous supposons ici, dans un but de simplification, que l'axe z est l'axe de révolution. Les équations réduites de ces quadriques de révolution sont alors les suivantes :

$$x^2 + y^2 + z^2 = R^2 \quad \text{pour une sphère,}$$

$$x^2 + y^2 = R^2 \quad \text{pour un cylindre de révolution,}$$

$$x^2 + y^2 - tg^2 \alpha z^2 = 0 \quad \text{pour un cône de révolution,}$$

$$\frac{x^2}{a^2} + \frac{y^2}{a^2} + \frac{z^2}{c^2} - 1 = 0 \quad \text{pour un ellipsoïde de révolution.}$$

Nous cherchons la nature et les paramètres d'une quadrique dont nous connaissons l'équation réduite $Sx^2 + S'y^2 + S''z^2 + D' = 0$, un des centres (x_0, y_0, z_0) et l'orientation de l'axe de révolution qui est donnée par V'' de coordonnées (α, β, γ) .

Cette quadrique est une sphère si les 3 valeurs propres sont identiques et non nulles. Ses paramètres sont le centre (x_0, y_0, z_0) et le rayon $\sqrt{\frac{-D'}{S}}$.

Nous avons un cylindre de révolution si les 2 valeurs propres S et S' sont identiques et non nulles et si la valeur propre S'' correspondant à l'axe de révolution est nulle. Le rayon de ce cylindre est donné par $\sqrt{\frac{-D'}{S}}$. Les autres paramètres μ, θ, ω et ρ sont calculés de la façon suivante :

$$\theta : \gamma > 0 \quad \theta = \text{Arctg} \frac{\sqrt{\alpha^2 + \beta^2}}{\gamma^2}$$

$$\gamma < 0 \quad \theta = \pi - \text{Arctg} \frac{\sqrt{\alpha^2 + \beta^2}}{\gamma^2}$$

$$\gamma = 0 \quad \theta = \frac{\pi}{2}$$

$$\mu : \beta \neq 0 \quad \beta > 0 \quad \mu = \text{Arctg} \left(\frac{-\alpha}{\beta} \right) + \pi$$

$$\beta < 0 \quad \mu = \text{Arctg} \left(\frac{-\alpha}{\beta} \right)$$

$$\beta = 0 \quad \alpha < 0 \quad \mu = \frac{3\pi}{2}$$

$$\alpha > 0 \quad \mu = \frac{\pi}{2}$$

$$\alpha = 0 \quad \mu = 0$$

$$\rho : \rho = \sqrt{m_1^2 + m_2^2 + m_3^2} \quad \text{avec} \quad m_1 = x_0 + \lambda\alpha$$

$$m_2 = y_0 + \lambda\beta$$

$$m_3 = z_0 + \lambda\gamma$$

$$\lambda = -(x_0\alpha + y_0\beta + z_0\gamma)$$

$$\omega : \alpha m_2 - \beta m_1 \neq 0 : \omega' = \text{Arctg} \frac{\sqrt{m_3^2 (\alpha^2 + \beta^2) + (m_1\alpha + \beta m_2)^2}}{\alpha m_2 - \beta m_1}$$

$$\omega' < 0 : \alpha m_1 + \beta m_2 < 0 : \omega = \pi - \omega'$$

$$\alpha m_1 + \beta m_2 > 0 : \omega = \omega' + \pi$$

$$\omega' \geq 0 : \alpha m_1 + \beta m_2 < 0 : \omega = 2\pi - \omega'$$

$$\alpha m_1 + \beta m_2 \geq 0 : \omega = \omega'$$

$$\alpha m_1 - \beta m_2 = 0 : \alpha = \beta = 0 : m_1 = 0 : m_2 < 0 : \omega = \frac{3\pi}{2}$$

$$m_2 > 0 : \omega = \frac{\pi}{2}$$

$$m_2 = 0 : \omega = 0$$

$$m_1 < 0 : \omega = \pi + \text{Arctg} \left(\frac{m_2}{m_1} \right)$$

$$m_1 > 0 : \omega = \text{Arctg} \left(\frac{m_2}{m_1} \right)$$

$$\alpha \neq 0 \vee \beta \neq 0 : \rho = 0 : \omega = 0$$

$$\rho \neq 0 : m_1\alpha + m_2\beta < 0 : \omega = \frac{3\pi}{2}$$

$$m_1\alpha + m_2\beta > 0 : \omega = \frac{\pi}{2}$$

La quadrique recherchée est un cône de révolution si les deux valeurs propres S et S' sont identiques et non nulles, si la troisième valeur propre S'' est non nulle et si le terme constant est nul. Son sommet est le centre (x_0, y_0, z_0) ; son demi-angle au sommet est donné par $\text{Arctg} \left(\sqrt{\frac{-S''}{S}} \right)$. Les deux angles ψ et θ sont calculés de la façon suivante :

$$\begin{aligned} \Psi: \alpha > 0 \quad \beta > 0 & \quad \Psi = \frac{\pi}{2} \\ \alpha > 0 \quad \beta < 0 & \quad \Psi = \frac{3\pi}{2} \\ \alpha < 0 & \quad \Psi = \text{Arctg}\left(\frac{\beta}{\alpha}\right) + \pi \\ \alpha > 0 & \quad \Psi = \text{Arctg}\left(\frac{\beta}{\alpha}\right) \end{aligned}$$
$$\begin{aligned} \Theta: \gamma = 0 & \quad \Theta = \frac{\pi}{2} \\ \gamma > 0 & \quad \Theta = \text{Arctg}\left(\frac{\sqrt{\alpha^2 + \beta^2}}{\gamma}\right) \\ \gamma < 0 & \quad \Theta = \text{Arctg}\left(\frac{\sqrt{\alpha^2 + \beta^2}}{\gamma}\right) + \pi \end{aligned}$$

La quadrique est un ellipsoïde de révolution si les deux valeurs propres S et S' sont identiques, si les trois valeurs propres sont de même signe et non nulles et si le terme constant n'est pas nul. Le centre est connu : c'est le point de coordonnées (x_0, y_0, z_0) . Les deux paramètres a et c sont donnés respectivement par $\sqrt{\frac{1}{S}}$ et $\sqrt{\frac{1}{S''}}$. Le calcul des paramètres Ψ et Θ donnant l'orientation de l'axe de révolution se fait de la même façon que pour le cône.

III.4 - CONCLUSION

Nous avons présenté, dans ce chapitre, une méthode permettant de passer d'une image des distances à l'espace des paramètres d'une surface. Seuls les points appartenant à une surface du type recherché ont une image dans l'espace des paramètres. Le nombre de points de cet espace est donc peu important et, en général, les points sont peu dispersés.

Le traitement qui consiste à associer une surface à chaque point de l'image à analyser peut se faire en parallèle, s'il nécessite beaucoup de calcul.

CHAPITRE IV
REALISATION PRATIQUE
ET RESULTATS

Cette détection de surfaces est expérimentée en Pascal sur un Exormacs (Motorola 68000). Pour la tester, nous avons conçu un système de création et d'affichage d'images tridimensionnelles. Ce système fournit des images de distances à partir d'une scène et d'une position de l'observateur dans cette scène.

Dans un premier temps, nous détaillerons le fonctionnement de ce système, puis nous présenterons les résultats obtenus sur des images synthétiques créées à l'aide de ce système. Nous montrerons, ensuite, un exemple de recherche de surfaces sur une image réelle obtenue à l'aide du capteur de l'INRIA (voir chapitre I, paragraphe I.2.d).

IV.1 - SYSTEME DE CREATION ET D'AFFICHAGE D'IMAGES TRIDIMENSIONNELLES

Ce système, réalisé en collaboration avec Martine CAMONIN, fournit, à partir d'un modèle de scène tridimensionnelle, l'image des distances et l'image des intensités lumineuses perçues par un observateur placé dans la scène. La position de l'observateur, la lumière ambiante et la source de lumière sont des paramètres du système.

a) Le modèle

Une scène est décrite dans un repère absolu à l'aide de primitives. Nous appelons "facette" une partie plane délimitée par un polygone et pouvant contenir des trous délimités eux aussi par des polygones. Les primitives utilisées classiquement sont des facettes [NEW - 79]. Nous avons choisi d'ajouter à ces facettes quelques quadriques de révolution simples : sphères, cylindres de révolution, cônes de révolution et ellipsoïdes de révolution.

Une facette est déterminée par la liste ordonnée de ses sommets et par la liste de ses trous, un trou étant lui-même une liste ordonnée de sommets. Une sphère est caractérisée par son centre et son rayon. Un cylindre de révolution est représenté par une extrémité P de son axe, par sa hauteur H, son rayon R et par la direction de son axe de révolution donnée par les angles ψ et θ (figure IV.1)

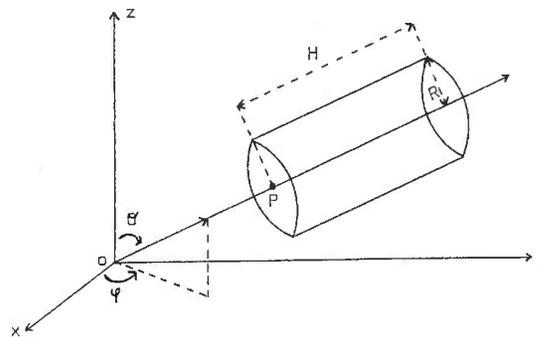


Figure IV.1 : Description d'un cylindre de révolution

Un cône de révolution est décrit par son sommet S, sa hauteur H, son demi-angle au sommet α et par la direction de son axe de révolution donnée par ψ et θ (figure IV.2).

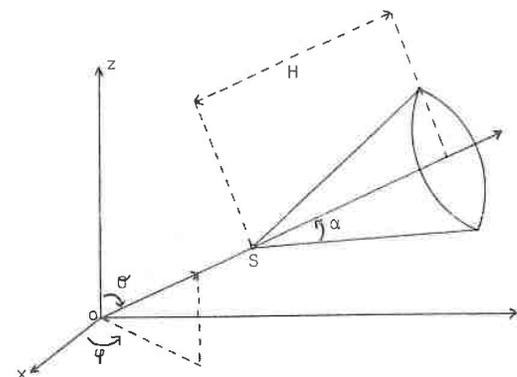


Figure IV.2 : Description d'un cône de révolution

Un ellipsoïde de révolution est caractérisé par son centre C , par 2 distances a et c définissant l'excentricité et par 2 angles ψ et θ indiquant la direction de l'axe de révolution (figure IV.3).

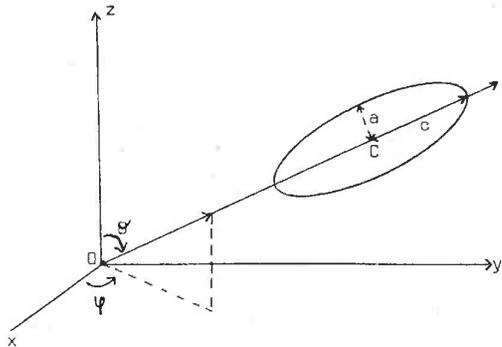


Figure IV.3 : Description d'un ellipsoïde de révolution

A ces différentes caractéristiques permettant de placer les objets dans l'espace, sont associées deux propriétés concernant les surfaces : la couleur et le coefficient de réflexion.

b) La perspective

L'image perçue par l'observateur est la projection de la scène sur son plan de vision (figure IV.4). Les coordonnées en x et en y d'un point P' projection d'un point P de la scène sur le plan de vision se calculent simplement à l'aide des propriétés géométriques sur les triangles [NEW - 79].

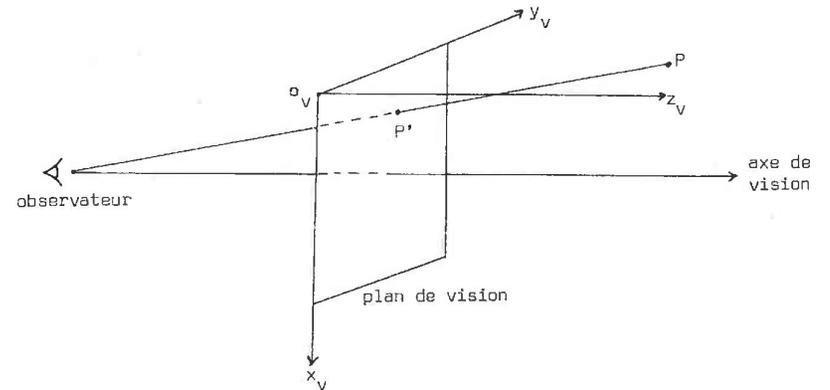


Figure IV.4 : Projection d'un point P de la scène sur le plan de vision de l'observateur

La notion de perspective peut donc être introduite facilement dans un système de création d'images.

Mais, dans un but de simplification, nous avons supposé, dans la première version de notre système, que l'observateur se trouve à l'infini. L'image qu'il perçoit est alors la projection orthogonale de la scène sur le plan de vision. Les coordonnées en x et en y d'un point P' du plan de vision sont donc les mêmes que celles en x et en y du point P dont P' est la projection orthogonale. Indiquer la position de l'observateur consiste alors à fixer le repère (o_v, x_v, y_v, z_v) en donnant les coordonnées cartésiennes de l'origine o_v et de chacun des trois vecteurs x_v, y_v et z_v .

c) Elimination des parties cachées

Dans une scène, certains objets peuvent être cachés partiellement ou entièrement par d'autres objets. De nombreux algorithmes ont été développés pour éliminer les parties cachées d'une scène [NEW - 79]. Le choix d'un tel algorithme est fonction du modèle utilisé et de la complexité de la scène. Le plus simple et le plus connu est sans doute l'algorithme de "Depth-Buffer" ; il utilise deux tableaux : un tableau DISTANCE pour l'image des distances et un tableau INTENSITE pour l'image des intensités lumineuses. Le principe de cet algorithme est le suivant :

- Pour tous les points (x,y) du plan de vision, initialisation de $DISTANCE(x,y)$ et de $INTENSITE(x,y)$
- Pour chaque primitive de la scène, recherche des coordonnées (x,y) de tous les points P' du plan de vision qui correspondent à la projection d'un point P de cette primitive. Pour chacun de ces points P' :
 - . calcul de la distance z de P à P'
 - . si $z < DISTANCE(x,y)$, mémorisation de z dans $DISTANCE(x,y)$ et de l'intensité lumineuse en ce point dans $INTENSITE(x,y)$.

Le principal inconvénient de cet algorithme est l'importance de la place occupée par les deux tableaux représentant l'image des distances et l'image des intensités lumineuses. C'est pour cette raison que nous avons choisi d'utiliser un autre algorithme. Il se rapproche des algorithmes de type "Scan-line" ; son principe est le suivant :

Pour chaque point P' du plan de vision :

- Initialisation de la distance D ,
- Pour chaque primitive pour laquelle il existe un point P dont la projection orthogonale sur le plan de vision coïncide avec P' :
 - . calcul de la distance D' de P à P'
 - . si $D' < D$ alors mémorisation de D' dans D et de l'intensité lumineuse en ce point dans I
- Sauvegarde de D et I pour P' .

L'intérêt de cet algorithme réside dans le fait qu'il n'utilise que très peu de place mémoire. Par contre, son temps d'exécution est plus important que celui de l'algorithme précédent car chaque primitive est prise en compte n fois si n est le nombre de points du plan de vision.

d) L'intensité lumineuse

L'intensité lumineuse en un point est fonction de l'éclairement et des propriétés de la surface à laquelle il appartient. Les deux propriétés principales d'une surface sont la transparence et le coefficient de réflexion. C'est ce dernier qui permet de calculer la quantité de lumière réfléchiée par une surface. Une scène peut être éclairée par une lumière réfléchiée par une surface. Une scène peut être éclairée par une lumière diffuse (ou ambiante) uniforme dans toutes les directions et par une ou plusieurs sources ponctuelles. L'intensité lumineuse en un point est composée de trois parties :

- la contribution de la lumière ambiante,
- la contribution des sources ponctuelles,
- l'effet de transparence.

La première version de notre système ne prend en compte qu'une source ponctuelle. De plus, les surfaces composant une scène sont supposées opaques.

Pour calculer l'intensité lumineuse en un point, nous supposons que les surfaces suivent la loi de Lambert. Elle indique que la quantité de lumière réfléchie par une surface parfaitement mate est la même dans toutes les directions. Dans le cas de l'éclairage par une lumière diffuse et une source ponctuelle, l'intensité réfléchie en un point P est alors donnée par la formule suivante :

$$I_R = R * I_A + R * I_S * \cos i$$

où R est le coefficient de réflexion, I_A l'intensité de la lumière ambiante, I_S l'intensité de la source ponctuelle et i l'angle d'incidence (figure IV.5).

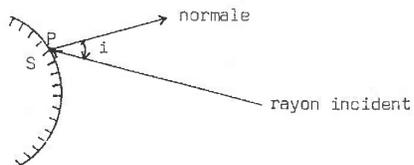


Figure IV.5 : L'angle d'incidence i en un point P d'une surface S

Si l'angle d'incidence est compris entre $\frac{\pi}{2}$ et $\frac{3\pi}{2}$, le point P n'est pas éclairé par la source ponctuelle. Seule la lumière ambiante intervient alors dans le calcul de l'intensité réfléchie.

D'autres méthodes de calcul de l'intensité lumineuse ont été proposées par BUI-TUONG PHONG [PHO - 75] et par BLINN [BLI - 77]. Elles intègrent la transparence et la brillance.

e) Mise en oeuvre

L'acquisition des données, qui est un problème difficile, a été résolu temporairement de manière très simple mais peu agréable à utiliser : une scène est décrite par l'ensemble de ses primitives (facettes, sphères, cylindres de révolution, cônes de révolution, ellipsoïdes de révolution) ; chaque primitive est déterminée comme indiqué dans le modèle (voir paragraphe IV.1.a). Ainsi, pour acquérir, par exemple, une facette, il faut donner les coordonnées cartésiennes de chacun de ses sommets en respectant leur ordre.

L'élaboration de l'image des distances et de l'image des intensités lumineuses se déroule en plusieurs étapes. La première étape consiste à associer à chaque facette l'équation du plan auquel elle appartient. Cette équation est donnée dans le repère image. A chaque cylindre, cône et ellipsoïde est attaché le repère principal d'inertie lui correspondant. De plus, à chaque primitive sont jointes ses coordonnées minimum et maximum exprimées dans le repère image dans le but d'améliorer les performances de l'algorithme utilisé. Nous évitons ainsi, par un simple test, de calculer la distance d'un point à une primitive si les coordonnées x et y de ce point n'ont aucune chance de coïncider avec les coordonnées x et y d'un point de la primitive.

La deuxième étape consiste à associer à chaque point (x,y) de l'image une distance, une couleur, un coefficient de réflexion et une normale. Les trois dernières valeurs permettent de calculer l'intensité lumineuse. Pour tout point (x,y) du plan de vision et pour toute primitive telle que (x,y) se trouve dans ses limites, nous calculons la distance z du point à la primitive. Nous mémorisons la plus petite distance trouvée. Le calcul de la distance dépend de la primitive.

Pour une facette, la distance z est calculée à l'aide de l'équation du plan de la facette. Nous vérifions ensuite que le point ainsi déterminé se trouve à l'intérieur de la facette et qu'il n'appartient à aucun trou. Il se trouve dans la facette si la somme des angles entre les différents vecteurs reliant le point à chacun des sommets de la facette vaut $\pm 2\pi$, sinon il se trouve à l'extérieur de celle-ci. Le même calcul pour chaque trou nous permet de savoir si le point appartient à un trou ou non. Si le point vérifie toutes les conditions précédentes, la valeur z est prise en compte. Elle est mémorisée si elle est inférieure à toutes les distances trouvées en ce point.

Pour les quadriques, la distance z est calculée à l'aide des équations réduites. Un point appartient à une quadrique si ses coordonnées dans le repère lié à la quadrique vérifient l'équation réduite de celle-ci. La valeur de z s'obtient donc en résolvant une équation du second degré à laquelle s'ajoute une contrainte sur la hauteur pour les cylindres et les cônes. S'il existe une telle valeur de z , elle est mémorisée si elle est inférieure à toutes les distances trouvées en ce point.

Pour chaque point de l'image, nous mémorisons en même temps que la distance de ce point à la primitive la plus proche, la normale

à cette primitive au point considéré, la couleur de cette primitive et son coefficient de réflexion. A la fin de cette étape, nous disposons de l'image des distances et de tous les éléments nécessaires au calcul de l'image des intensités lumineuses. Ce calcul constitue la dernière étape.

La visualisation de l'image perçue par l'observateur consiste à afficher les couleurs et les intensités lumineuses de chaque point de son plan de vision.

f) Conclusion

Ce système de création et d'affichage d'images permet donc, à partir d'un modèle de scènes tridimensionnelles, de visualiser l'image perçue par un observateur placé dans la scène et de créer l'image des distances correspondante. Les possibilités actuelles du système sont limitées. L'étude d'une version optimisée et améliorée est en cours. Les principales améliorations faciles à ajouter sont la transparence [PHO - 75] et les ombres [ATH - 76], [CRO - 77]. Une méthode possible pour introduire les ombres consiste à considérer la source ponctuelle comme un observateur pour lui appliquer un algorithme d'élimination des parties cachées. La transparence ne pose pas de problème particulier ; il suffit de multiplier l'intensité de tout point caché par la transparence du point qui le cache. Pour plus de réalisme, la notion de texture devrait également être introduite. Nous ne l'envisageons pas dans l'immédiat car les techniques existantes sont difficiles à mettre en oeuvre [BLI - 76], [BLI - 78]. L'utilisation d'un modèle tridimensionnel plus complet pourrait aussi être envisagée mais sort du cadre du projet TRIDENT.

IV.2 - RESULTATS

Nous allons présenter quelques exemples de recherche de surfaces. Pour chacun de ces exemples, nous indiquons quelles valeurs nous avons choisies pour les différents paramètres qui peuvent être fixés par l'utilisateur. Ces paramètres sont fonction de la qualité de l'image à traiter et de la précision souhaitée sur les résultats. Il s'agit des cinq variables suivantes :

- NP , le nombre minimum de points dont doit être formée une surface (voir chapitre II, paragraphe II.3),
- T , la taille à partir de laquelle une cellule de l'espace des paramètres est considérée comme élémentaire (c'est-à-dire assimilée à un point) (voir chapitre II, paragraphe II.3),
- ϵ , le seuil à partir duquel l'erreur au sens des moindres carrés (dans le cas de la recherche d'un plan) est considérée comme nulle (voir chapitre III, paragraphe III.2 b),
- E , l'erreur relative utilisée pour tester si deux valeurs peuvent être considérées comme égales ou si une valeur peut être considérée comme nulle comparées à un ensemble d'autres valeurs (utile dans le cas de la recherche de quadriques),
- V , la taille du voisinage utilisé pour associer à un point les paramètres d'une surface (voir chapitre III, paragraphes III.2 b et III.3 a) ; le nombre de points de ce voisinage est constant, il est de 25 dans nos exemples.

a) Sur des images synthétiques

Les premiers tests ont été effectués sur des images de synthèse de résolution 256 X 256.

Première image test -

Dans l'image de la figure IV.6, nous avons recherché tous les plans qui composent la maison. La zone de recherche correspond à une fenêtre contenant la maison. Les paramètres choisis sont les suivants :

$$NP = 100$$

$$T = \begin{cases} 0,1 & \text{sur la dimension } \theta' \\ 0,1 & \text{sur la dimension } \psi \\ 1 & \text{sur la dimension } d \end{cases}$$

$$\epsilon = 0,01$$

$$V = 5 \times 5$$

Les résultats qui apparaissent figures IV.7a et IV.7b montrent que tous les plans d'au moins 100 points ont été trouvés. Les points leur appartenant sont visualisés en gris-bleu dans la figure IV.7a. Les paramètres de chacun des cinq plans trouvés sont donnés dans la figure IV.7b. Le nombre de points trouvés pour chaque plan y figure également. Les points de contour des surfaces ne sont pas trouvés car nous utilisons un voisinage pour calculer l'équation de la surface passant par un point. Ce problème peut être résolu par un traitement ultérieur, rapide et simple (voir chapitre V, paragraphe V.1 g).

Nous avons, dans un deuxième temps, effectué une recherche de quadriques dans l'ensemble de l'image de la figure IV.6. Les paramètres

utilisés sont liés à chaque type de quadrique. Ils sont les suivants pour celles qui nous intéressent :

pour les cylindres : NP = 50

$$E = 0,0015$$

$$V = 5 \times 5$$

$$T = \begin{cases} 0,2 & \text{pour les dimensions} \\ & \text{correspondant à des angles} \\ 2 & \text{pour les autres} \end{cases}$$

pour les cônes : NP = 50

$$E = 0,0005$$

$$V = 9 \times 9$$

$$T = \begin{cases} 0,2 & \text{pour les dimensions} \\ & \text{correspondant à des angles} \\ 3 & \text{pour les autres} \end{cases}$$

pour les ellipsoïdes : NP = 50

$$E = 0,005$$

$$V = 13 \times 13$$

$$T = \begin{cases} 0,3 & \text{pour les dimensions} \\ & \text{correspondant à des angles} \\ 3 & \text{pour les autres} \end{cases}$$

Les quatre quadriques présentes dans l'image ont été détectées. La liste de leurs paramètres est donnée dans la figure IV.8b. Les points qui ont été trouvés sont affichés en gris dans la figure IV.8a. On remarque que, en plus des points des bords des surfaces, un certain nombre d'autres points n'ont pas été trouvés. Les irrégularités locales et le

manque de précision dans certains calculs (par exemple, lorsque l'on recherche les coefficients d'une surface au sens des moindres carrés ou lorsque l'on recherche les valeurs et les vecteurs propres) sont les deux principales causes de ces oublis. Ce problème peut être résolu simplement à l'aide d'un lissage fondé sur le principe suivant : tout point isolé, entouré de points appartenant à la surface considérée est assimilé à cette surface.

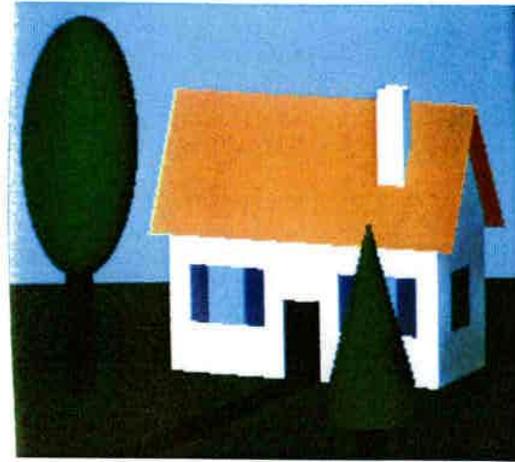


Figure IV.8 :
Scène de maison

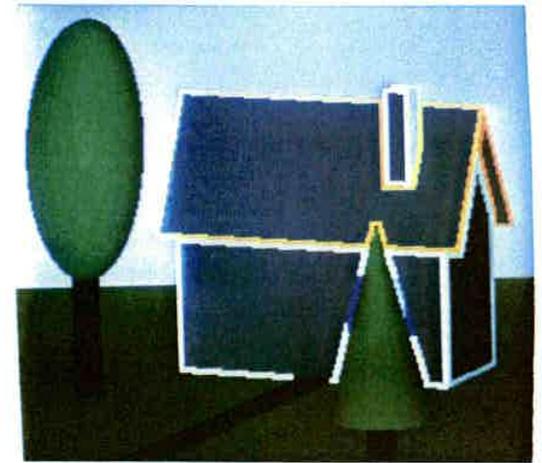


Figure IV.7a :
Détection de plans

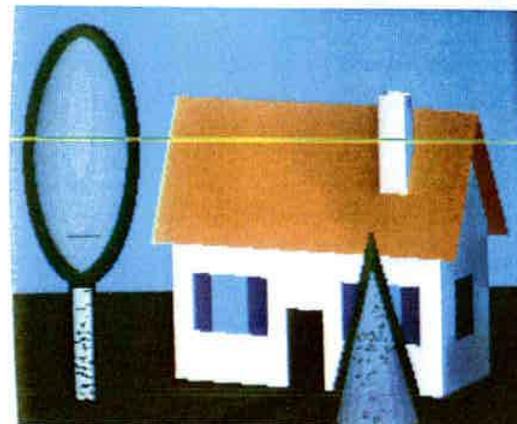


Figure IV.8a :
Détection de quadrangles

Liste des plans trouvés :

```
-----
plan no 1 : teta = 1.186   phi = 4.632   d = 92.952
            nb. d'el. = 1373
plan no 2 : teta = 0.760   phi = 2.626   d = -263.051
            nb. d'el. = 186
plan no 3 : teta = 0.436   phi = 2.012   d = -329.630
            nb. d'el. = 6108
plan no 4 : teta = 0.436   phi = 2.012   d = -339.624
            nb. d'el. = 358
plan no 5 : teta = 0.457   phi = 0.889   d = -380.501
            nb. d'el. = 9323
```

Figure IV.7b : Liste des paramètres des plans trouvés

Liste des cylindres trouvés :

```
-----
cylindre no 1 : nu = 1.57
                teta = 1.57
                ro = 54.89
                omega = 0.84
                rayon = 5.56
                nb. d'el. = 479
cylindre no 2 : nu = 1.57
                teta = 1.57
                ro = 277.11
                omega = 0.86
                rayon = 7.05
                nb. d'el. = 50
```

Liste des cônes trouvés :

```
-----
cône no 1 : centre = ( 130.43 , 180.56 , -185.53 )
            alpha = 0.16
            teta = 1.57
            phi = 0.00
            nb. d'el. = 1559
```

Liste des ellipsoïdes trouvés :

```
-----
ellipsoïde no 1 : centre = ( 87.83 , 36.65 , -22.93 )
                  a = 29.95
                  c = 64.60
                  teta = 1.57
                  phi = 0.00
                  nb. d'el. = 4274
```

Figure IV.8b : Liste des paramètres des quadriques trouvées

Deuxième image test -

La deuxième image test utilisée (figure IV.9) représente un ensemble de figures géométriques disposées en vrac. Nous avons recherché dans cette image toutes les sphères. Les paramètres utilisés sont les suivants :

NP = 50
E = 0,0005
V = 13 X 13
T = 1 sur toutes les dimensions

Les quatre sphères présentes dans l'image ont été détectées. Les points leur appartenant sont visualisés en gris dans la figure IV.10a. Les paramètres et le nombre de points de chacune de ces sphères sont donnés dans la figure IV.10b.

Les figure IV.11a et IV.11b montrent les résultats d'une recherche de cône sur cette même image avec les paramètres suivants :

NP = 50
E = 0,0005
V = 13 X 13
T = { 0,2 pour toutes les dimensions
 correspondant à des angles
 1 pour les autres

Le dernier exemple correspond à une recherche de cylindres d'axe de révolution à peu près horizontal. Cette contrainte se traduit par une réduction de deux dimensions de l'espace des paramètres :

$$1,54 \leq \theta \leq 1,61$$

$$-0,3 \leq \mu \leq 0,3$$

Les paramètres utilisés pour cette recherche sont les suivants :

NP = 50
E = 0,001
V = 13 X 13
T = { 0,2 pour les dimensions
 correspondant à des angles
 2 pour les autres

Les résultats apparaissent dans les figures IV.12a et IV.12b.

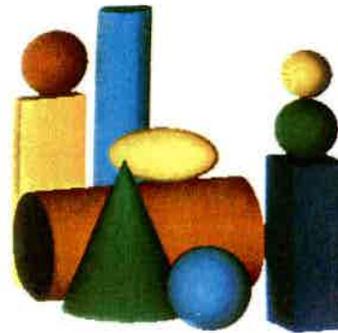


Figure IV.9 :
Figures géométriques

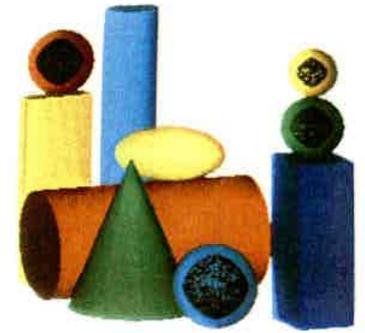


Figure IV.10a :
Détection de sphères

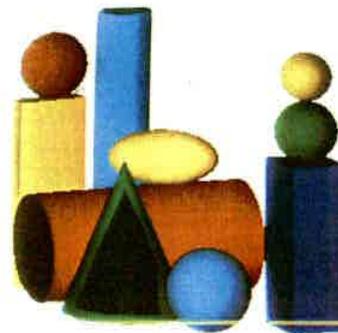


Figure IV.11a :
Détection de cônes

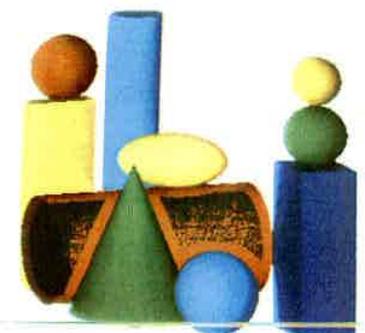


Figure IV.12a :
Détection de cylindres
d'axe de révolution à
peu près horizontal

Liste des spheres trouvees :

sphere no 1 : centre = (64.33 , 216.01 , 186.39)
rayon = 17.97
nb. d'el. = 285
sphere no 2 : centre = (103.48 , 215.98 , 194.37)
rayon = 21.87
nb. d'el. = 585
sphere no 3 : centre = (54.88 , 38.09 , 323.74)
rayon = 24.69
nb. d'el. = 916
sphere no 4 : centre = (217.02 , 149.37 , 194.42)
rayon = 29.69
nb. d'el. = 1373

Figure IV.10b : Liste des paramètres des sphères trouvées

Liste des cones trouvees :

cone no 1 : centre = (125.84 , 91.04 , 239.36)
alpha = 0.36
teta = 1.57
phi = 0.00
nb. d'el. = 3198

Figure IV.11b : Liste des paramètres des cônes trouvés

Liste des cylindres trouvees :

cylindre no 1 : nu = 0.00
teta = 1.57
ro = 329.37
omega = -0.95
rayon = 40.29
nb. d'el. = 4071

Figure IV.12b : Liste des paramètres des cylindres trouvés

Conclusion -

Un certain nombre de remarques ou conclusions découlent de l'ensemble de ces résultats :

- si aucune prévision ne peut être effectuée sur la taille des surfaces recherchées, il paraît raisonnable de choisir NP égal à 50 pour des images 256 X 256. Dans le cas d'un voisinage 5 X 5, la plus petite surface qui peut être trouvée contient alors 126 points dont 50 seront pris en compte, les autres étant des points du bord des surfaces,

- le choix de la taille maximale d'une cellule élémentaire dépend de la précision souhaitée. Raisonnablement, cette taille peut être fixée à 2 pour les distances et 0,2 pour les angles. Si cette taille est inférieure à celle d'un point d'accumulation, des fusions seront effectuées,

- pour des images de bonne qualité (c'est le cas des images de synthèse utilisées), le seuil relatif à l'erreur au sens des moindres carrés, dans le cas d'une recherche de plans, peut être fixé à 10^{-2} . Cette valeur peut aussi convenir pour l'erreur relative utilisée dans le cas des recherches de quadriques. Ainsi, si on a un ensemble de valeurs dont le plus grande est V_1 , on considère que deux de ces valeurs sont égales si leur différence est inférieure à $V_1 \cdot 10^{-2}$. De même, une valeur est considérée comme nulle par rapport à d'autres si elle est inférieure à 10^{-2} fois la plus grande. Le choix de cette erreur influence beaucoup les résultats. Il serait intéressant d'effectuer un grand nombre d'essais pour en déduire une méthode qui fixerait cette erreur en fonction du type de surfaces recherchées, de la précision souhaitée et des résultats obtenus pour les points déjà traités à un instant donné de

la recherche. Ce travail nécessitant un grand nombre de résultats expérimentaux, il ne sera réalisé que lorsque le système de détection de surfaces sera implanté sur une machine plus performante (nous envisageons de l'implanter sur un VAX 750),

- le problème du choix de la taille d'un voisinage est facile à résoudre ; pour des plans, un voisinage 5 X 5 semble convenir parfaitement. Pour des quadriques dont le rayon de courbure est important, un grand voisinage est nécessaire pour effacer les irrégularités locales et pour que la courbure soit sensible sur le voisinage choisi. Un voisinage 13 X 13 est un bon choix dans ce cas.

b) Sur des images réelles

Les résultats sur des images de synthèse étant satisfaisants, nous avons ensuite testé notre système de détection de surfaces sur une image réelle, obtenue à l'aide du capteur de l'INRIA (voir chapitre I, paragraphe I.2 d). La scène numérisée comprend une maquette de maison (figure IV.14) et un arbre. La figure IV.15 représente l'image des distances correspondant à cette scène. Nous avons recherché dans cette image tous les plans composant la maison. Les paramètres utilisés sont les suivants :

$$NP = 200$$

$$\epsilon = 5$$

$$V = 9 \times 9$$

$$T = \begin{cases} 2 & \text{pour la dimension correspondant à la distance} \\ 0,2 & \text{pour les deux autres.} \end{cases}$$

La liste des plans trouvés, les paramètres et le nombre de points appartenant à chacun de ces plans apparaissent figure IV.16b. Le premier plan est celui du toit de la maison ; les points lui appartenant sont représentés par le caractère "1" dans la figure IV.16a. Presque tous les points du toit (sauf ceux du bord et ceux de la cheminée) ont été trouvés. Le troisième plan correspond au côté gauche de la maison ; les points lui appartenant sont visualisés à l'aide du caractère "3" dans la figure IV.16a. Les résultats sont satisfaisants sauf en ce qui concerne une zone située à mi-hauteur de la maison. Cette zone correspond effectivement à une discontinuité. Le deuxième plan est celui de la façade avant de la maison ; les points lui appartenant sont affichés à l'aide du caractère "2" dans la figure IV.16a. Ces points sont peu nombreux pour plusieurs raisons :

- la façade avant comprend une porte et deux fenêtres en retrait par rapport au plan de la façade,
- chaque fenêtre est entourée de deux volets qui ont une épaisseur non négligeable,
- de plus, la façade est rugueuse.

Cette façade n'est donc effectivement pas plane ; si on veut tout de même l'assimiler à un plan, il faut augmenter ϵ .

Nous avons volontairement choisi ϵ grand car aucune partie de la maison n'est vraiment plane. Le toit est composé de tuiles, avec une valeur de ϵ faible, il n'aurait pas été assimilé à un plan. La partie supérieure du côté gauche de la maison est constituée d'un alignement de poutres ; elle est donc également rugueuse. Seule la partie inférieure est presque plane (elle correspond à du crépi). Elle peut donc être détectée avec une valeur de ϵ un peu plus faible et une taille de voisinage également plus petite : $\epsilon = 3$ et $V = 5 \times 5$ (figures IV.17a et IV.17b).

Ces tests montrent que la méthode utilisée pour rechercher des surfaces est encore satisfaisante pour des images réelles. Elle peut même être utilisée pour rechercher des surfaces à peu près planes. Il suffit de fixer les paramètres en fonction de la précision souhaitée.

Des tests semblables n'ont pas pu être réalisés sur des quadriques car nous ne possédons pas d'images réelles comportant de telles surfaces. Nous pensons que les résultats devraient être comparables.

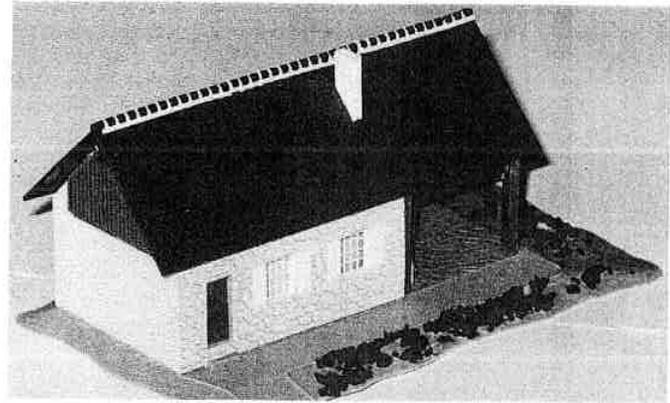


Figure IV.14 : Maquette de maison



Figure IV.15 : Image des distances de cette maquette de maison et d'un arbre



Figure IV.16a : Détection de plans avec $\epsilon = 5$ et $V = 9 \times 9$

Liste des plans trouvés :

plan no 1 :	teta = 0.835	phi = 5.715	d = -111.365
	nb. d'el. = 3752		
plan no 2 :	teta = 0.527	phi = 4.777	d = -82.692
	nb. d'el. = 523		
plan no 3 :	teta = 1.021	phi = 1.577	d = -189.786
	nb. d'el. = 851		

Figure IV.16b : Liste des paramètres des plans trouvés
avec $\epsilon = 5$ et $V = 9 \times 9$



Figure IV.17a : Détection d'un plan -codé par des "1"-
avec $\epsilon = 3$ et $V = 5 \times 5$

Liste des plans trouvés :

plan no 1 : teta = 1.022 phi = 1.579 d = -189.454
nb. d'al. = 1112

Figure IV.17b : Liste des paramètres du plan trouvé
avec $\epsilon = 3$ et $V = 5 \times 5$

IV.3 - CONCLUSION

Nous avons testé notre système de détection de surfaces sur des images de synthèse et sur une image réelle. Dans les deux cas, les résultats sont satisfaisants. Des traitements ultérieurs seront cependant nécessaires pour traiter les points des bords des surfaces et, éventuellement, quelques points oubliés à l'intérieur des surfaces. La précision des résultats est variable ; c'est un paramètre du système. De nombreuses utilisations de ce système de détection de surfaces peuvent être envisagées ; nous étudierons ces différentes possibilités dans le chapitre V.

CHAPITRE V
INTEGRATION DANS UN SYSTEME
D'INTERPRETATION D'IMAGES

Les chapitres II et III ont décrit un système de détection de surfaces pour des images de distances. Au chapitre I, nous avons vu quelques exemples de systèmes de vision dont TRIDENT. Ce chapitre montre comment intégrer notre système de détection de surfaces dans un système de vision. Nous envisageons les différentes utilisations possibles :

- recherche de toutes les surfaces d'une certaine nature (exemple : recherche de toutes les sphères),
- recherche avec contraintes sur les paramètres des surfaces (exemple : recherche d'une sphère de rayon R),
- recherche dans une fenêtre de l'image,
- recherche dans une région,
- recherche à grands pas,
- recherche de surfaces évidentes.

Nous estimons les gains correspondants et nous proposons des solutions à deux problèmes qui se posent dans certains cas d'utilisation :

- comment intégrer les points des bords des surfaces à la liste des points de cette surface ?
- comment décider facilement si un point appartient à une surface de paramètres connus ?

Nous suggérons ensuite une stratégie d'utilisation du système de détection de surfaces dans un système de vision.

V.1 - LES UTILISATIONS POSSIBLES ET LES GAINS CORRESPONDANTS

Le temps T nécessaire à la recherche de surfaces est essentiellement la somme de deux temps T₁ et T₂ : T₁ est celui nécessaire à la recherche de points d'accumulation et T₂ celui nécessaire au passage de l'espace de départ à l'espace des paramètres.

T₁ est fonction de la taille et du nombre de points de l'espace des paramètres :

- la taille de l'espace des paramètres intervient dans le calcul du nombre de divisions de l'espace à effectuer : par exemple, pour diviser un espace de dimension p selon k parties égales sur chacune de ses dimensions, nous avons à considérer k^p cellules ; si un seul point d'accumulation est présent, il est localisé après log₂(k^p) = p log₂ k divisions,

- le nombre de points de l'espace des paramètres intervient dans le nombre de tests à effectuer pour répartir les points dans les cellules : pour un espace des paramètres de dimension p contenant n points accumulés et m points répartis aléatoirement, le nombre TE de tests est donné par la formule suivante :

$$TE = nd + m + \frac{m}{2} + \dots + \frac{m}{2^d} \leq nd + 2m$$

où d est le nombre de divisions.

Avec les hypothèses précédentes, T₁ est de l'ordre de np log₂ k + 2m . Les expériences montrent que m est souvent petit par rapport à n . Dans ce cas, T₁ est proportionnel à np log₂ k .

T₂ est fonction du nombre de points à traiter dans l'espace initial. Pour chaque point, on effectue un calcul pour obtenir les paramètres de la surface à associer à ce point. Si ce calcul nécessite un temps T₃ et s'il y a n points à traiter, on a T₂ = n T₃ . Dans T₃ sont inclus, d'une part, la résolution d'un système d'équations au sens des moindres carrés et, d'autre part, dans certains cas, la recherche des valeurs propres et des vecteurs propres d'une matrice 3 X 3 .

T₂ est souvent beaucoup plus coûteux que T₁ . Par exemple, sur une machine 16 bits où les opérations en virgule flottante ne sont pas câblées, T₂ peut atteindre 1 heure 20 dans le cas où l'on traite 64 X 64 points ; par contre T₁ ne dépasse jamais la minute.

a) Recherche de toutes les surfaces d'une certaine nature

Si on connaît la nature des surfaces à rechercher, un seul espace des paramètres est à prendre en compte. La recherche de points d'accumulation dans ce seul espace est évidemment plus rapide que la recherche de points d'accumulation dans l'ensemble des espaces des paramètres possibles. Un gain intervient donc au niveau du temps T₁ en comparaison avec le cas où on recherche toutes les familles de surfaces.

Un gain plus ou moins important selon la nature des surfaces recherchées intervient également au niveau du temps T₂ . Si on recherche une quadrique de nature N , il est inutile, pour certains points, de calculer l'équation réduite de la quadrique à associer à ce point. En effet, si ce point appartient à une autre famille de surfaces, il est souvent possible de s'en rendre compte avant d'obtenir l'équation réduite. Prenons l'exemple d'une recherche de sphères : pour chaque point de la zone de recherche, on obtient une équation de quadrique. Si dans la première équation obtenue (voir chapitre III, paragraphe III.3 a, équation (1)), les coefficients des termes en x², y² et z² sont significativement

différents, on peut en déduire que cette quadrique ne sera pas une sphère. Des optimisations de ce type ont lieu à tous les niveaux qui se situent entre l'obtention de la première équation et l'obtention de l'équation réduite.

b) Recherche avec contraintes sur les paramètres des surfaces

L'utilisation de la transformée de Hough nous permet de prendre en compte des contraintes sur les paramètres des surfaces recherchées (par exemple, pour une sphère, on peut imposer un rayon de 5). Ces contraintes se traduisent par une réduction de la taille des espaces des paramètres et donc accélèrent le processus de recherche de points d'accumulation. Les requêtes imposant des contraintes sur les paramètres sont, par exemple, du type :

- rechercher une sphère de rayon R ,
- rechercher un cylindre d'axe de révolution vertical,
- rechercher un cône dont le demi-angle au sommet est compris entre α_1 et α_2 ,
- ...

Prenons l'exemple d'une recherche de cylindre d'axe de révolution vertical. La contrainte "axe de révolution vertical" fixe les deux paramètres θ et μ du cylindre. La recherche se fait alors dans un espace des paramètres de dimension 3 au lieu de 5 ; elle est donc plus rapide que dans le cas d'une recherche de cylindre quelconque. Par exemple, si nous voulons diviser l'espace des paramètres en 100 parties égales sur chacune de ses dimensions, nous avons à considérer 100^5 cellules si la dimension est 5 et 100^3 cellules si la dimension est 3. Si un seul point d'accumulation est présent dans cet espace, il est localisé après $\log_2(10^{10}) \sim 33$ divisions si la dimension est 5 et après $\log_2(10^6) \sim 20$ divisions si la dimension est 3. La contrainte sur l'axe de révolution réduit donc le nombre de divisions de 33 à 20.

Les contraintes sur les paramètres des surfaces recherchées réduisent aussi le nombre de points de l'espace des paramètres. Il en résulte un gain au niveau des tests effectués pour répartir les points d'une cellule dans deux nouvelles cellules obtenues après division. Pour un espace des paramètres de dimension p contenant n points accumulés et m points répartis aléatoirement, nous avons vu que le nombre TE de tests est de l'ordre de $n.d + 2.m$ où d est le nombre de divisions. Si $p = 5$, $d = 33$, $n = 10\ 000$ et $m = 1\ 000$, alors $TE = 332\ 000$. Si n est réduit à 1 000 et m à 100, alors $TE = 33\ 200$. Le gain au niveau des tests est donc, dans ce cas, de 298 800.

Toutes sortes de demandes entraînant des contraintes sur les paramètres des surfaces recherchées engendrent un gain sur $T1$. Ce gain, bien que minime vis-à-vis du calcul préliminaire, accélère la recherche de points d'accumulation.

c) Recherche dans une fenêtre de l'image

Nous pouvons envisager d'effectuer une recherche dans une partie seulement d'une image. Dans la pratique, ce type de restriction se traduit le plus souvent par une recherche dans une fenêtre.

Pour nous, une image est représentée sous la forme matricielle $I(l,c)$ où l est le nombre de lignes et c le nombre de colonnes. Une fenêtre délimite une zone rectangulaire de l'image comprise entre deux lignes l_1 et l_2 et entre deux colonnes c_1 et c_2 . Une recherche sur une fenêtre réduit la dimension de l'espace de départ. Si le nombre de points de la fenêtre est n fois plus petit que le nombre de points de l'image entière, le temps $T2$ est divisé par n . Le gain peut donc dans ce cas être très important ; par exemple, pour une fenêtre de taille $\frac{1}{8}$ de celle de l'image initiale, le nombre de points à traiter est divisé par 64 (= 8×8).

d) Segmentation en régions et recherche dans une région

Un système de vision peut bénéficier d'un certain nombre de renseignements provenant de différentes images de mesures et de différents outils de segmentation. Il peut, par exemple, disposer de l'ensemble des régions d'une image. On appelle région d'une image I tout sous-ensemble R de points de I vérifiant les propriétés suivantes [CAS - 84] :

$\forall p_1 \in R, \forall p_2 \in R, p_1$ et p_2 sont connectés dans R par une propriété P .

(un exemple de propriété P peut être : "être de même couleur", d'autres exemples seront cités plus loin dans le texte).

Deux points p_1 et p_2 d'une image sont dits connectés par la propriété P , définie sur tout point de l'image :

- s'ils vérifient cette propriété : $P(p_1)$ et $P(p_2)$
- et s'il existe une séquence de points q_1, q_2, \dots, q_n telle que

$$\forall i \in [1, \dots, n], P(q_i)$$

$$q_1 = p_1 \text{ et } q_n = p_2$$

$$\forall i \in [1, \dots, n], q_i \text{ voisin de } q_{i+1}.$$

Si on traduit ceci en théorie des graphes, on considère le graphe $G = (E, \Gamma)$ où E est l'ensemble des points et Γ la relation $x \Gamma y \iff x$ est voisin de y et $P(x)$ et $P(y)$. Dans cette terminologie, une région est une composante connexe de G .

Deux points $p_1(i,j)$ et $p_2(k,l)$ sont voisins si la distance de p_1 à p_2 vaut 1 vu les métriques utilisées. Citons pour mémoire la distance de la valeur absolue ($d(p_1, p_2) = |i-k| + |j-l|$) et la distance du maximum ($d(p_1, p_2) = \text{Max}(|i-k|, |j-l|)$) .

Segmentation en régions -

La segmentation d'une image en régions (appelée "regroupement de régions") peut se faire selon plusieurs critères (continuité dans la variation des distances, dans la couleur, ...) et sur plusieurs images de mesures (image des distances, image des niveaux de gris, image des couleurs, ...). Nous pouvons envisager de faire un regroupement de régions sur les critères suivants : continuité dans la variation des distances, des normales et des couleurs et même type de convexité.

Recherche globale -

Après une telle segmentation, les points d'une même région appartiennent généralement à une même surface. Pour en connaître la nature et les paramètres, il suffit alors de traiter un seul point de la région. Pour déterminer l'équation de la surface à associer à ce point, il faut prendre en compte un certain nombre de ses voisins (voir chapitre III, paragraphes III.2 b et III.3 a). On peut considérer comme voisins, soit tous les autres points de la région, soit un échantillon de ces points (on choisit par exemple un point sur p). Dans les deux cas, le temps T2 est très faible ($T2 = T3$) . Il en est de même pour le temps T1 car un seul point est présent dans l'espace des paramètres. La recherche de points d'accumulation est donc immédiate. En plus du gain très important sur T1 et T2 , cette méthode de recherche globale dans une région supprime le problème des points des bords des surfaces (voir chapitre IV, paragraphe IV.2).

Recherche par échantillonnage -

Des cas de mauvaises segmentations peuvent intervenir dans certains regroupements de régions. La méthode précédente ne permet pas de les détecter ; elle n'est donc applicable que lorsqu'on est sûr que la région considérée correspond à une seule surface. Dans les autres cas, une méthode plus "prudente" s'impose.

Au lieu de traiter seulement un point de la région, on traite un échantillon de ces points (par exemple, un point sur p). Si la détection de surfaces appliquée à cet échantillon de points confirme la présence d'une seule surface, on en déduit que l'ensemble des points de la région appartient à cette même surface dont on connaît la nature et les paramètres. Dans ce cas, la recherche par échantillonnage, bien que plus lente que la recherche globale dans une région, est beaucoup plus rapide qu'une recherche point par point. En effet, T_2 est de l'ordre de $\frac{n}{p} T_3$ au lieu de nT_3 où n est le nombre de points de la zone de recherche. Un gain non négligeable intervient également sur T_1 car le nombre de points de l'espace des paramètres est très faible : de l'ordre de $\frac{n}{p}$. Enfin, cette méthode a aussi l'avantage de supprimer le problème des points des bords des surfaces.

Si, par contre, plusieurs surfaces sont présentes dans la région considérée, on connaît la nature et les paramètres de ces surfaces mais pas la liste des points leur appartenant. Dans ces cas exceptionnels, on peut soit refaire un traitement point par point, soit envisager des traitements plus rapides qui, partant d'équations de surfaces, donnent la liste des points leur appartenant (voir paragraphe V.1 h).

e) Recherche à grands pas

Pour associer à un point P les paramètres d'une surface, il faut que ce point et n de ses voisins vérifient l'équation de cette surface (voir chapitre III, paragraphe III.3 a). Il est donc inutile de traiter un à un les n voisins de ce point ; on peut les assimiler directement à la surface associée à P . Le temps T_2 est ainsi divisé par un facteur n . Un petit gain intervient également au niveau du temps T_1 car le nombre de points de l'espace des paramètres est divisé par n (donc T_1 est approximativement divisé par n). Malheureusement, la recherche est parfois moins bonne que dans le cas où on traite tous les points. En effet, les expériences ont montré que, pour des raisons de précision, certains points d'une surface ne sont pas trouvés (voir chapitre IV, paragraphe IV.2). En ne traitant qu'un point sur n , le nombre de points ainsi oubliés est dans le pire des cas multiplié par n . Mais, statistiquement, la proportion devrait être semblable.

f) Recherche de surfaces évidentes

Dans une image, il peut y avoir un certain nombre de surfaces évidentes. Il s'agit de surfaces très larges. Pour les détecter, on peut se contenter de parcourir l'image à très grands pas, en ne traitant qu'un point sur m (m pouvant être de l'ordre de 200 voire 400 pour une image 256 X 256). Le temps T_2 est ainsi divisé par m . Le temps T_1 est également diminué car le nombre de points des espaces des paramètres est divisé par m (donc T_1 est approximativement divisé par m). Cette méthode fournit les paramètres des grandes surfaces présentes dans l'image mais elle ne donne pas l'ensemble des points appartenant à chacune de ces surfaces. Il faut donc envisager un traitement ultérieur pour décider rapidement si un point appartient à une surface donnée (voir paragraphe V.1 h).

g) Problème des bords des surfaces

Dans la recherche de surfaces, les points des bords des surfaces ne sont pas pris en compte (sauf dans le cas des recherches dans des régions). Ce problème est dû à l'utilisation d'un voisinage pour les calculs relatifs à un point. Pour nous, le voisinage d'un point P est composé de l'ensemble ou d'une partie des points qui appartiennent à une fenêtre centrée sur P. La taille de cette fenêtre peut être variable. Les expériences montrent que les tailles les plus intéressantes sont 5 X 5, 9 X 9 ou 13 X 13 (voir chapitre IV, paragraphe IV.2). Un point du bord d'une surface a un voisinage dont tous les points n'appartiennent pas à la même surface. Nous ne trouverons donc pas de surface passant par l'ensemble de ces points. Aucun traitement spécial ne peut être prévu pour ces points car ils ne sont pas connus à l'avance. Nous pouvons cependant envisager des traitements ultérieurs destinés à résoudre ce problème.

Nous pouvons, par exemple, ajouter à la surface considérée tous les points qui sont voisins de celle-ci et qui vérifient son équation avec une certaine tolérance ($f(x,y,z) = 0 \pm \epsilon$). Cette méthode nécessite une grande précision sur les paramètres. De plus, il est difficile d'obtenir, par exemple par un calcul d'erreur, la tolérance à appliquer.

Une autre méthode consiste à utiliser des critères comme la couleur, la continuité dans les distances et dans les tangentes pour ajouter à la surface les points voisins de son contour.

Mais la méthode la plus simple consiste à parcourir la surface considérée ligne par ligne. Si une ligne donnée débute à la colonne c_1 et se termine à la colonne c_2 , on l'étend (si les limites de l'image le

permettent) de la colonne c_1-x à la colonne c_2+x où x vaut 2,4 ou 6 selon que le voisinage est du type 5 X 5, 9 X 9 ou 13 X 13. On effectue ensuite un traitement similaire pour chaque colonne. Tous les points des bords (sauf quelques exceptions) sont ainsi systématiquement intégrés à l'ensemble des points de la surface considérée. Les exceptions sont les points pour lesquels il n'est pas possible de trouver un voisinage dont tous les points appartiennent à la surface (figure V.1).

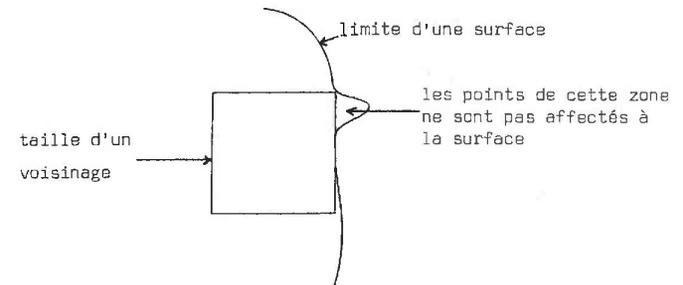


Figure V.1 : Exemple de points non intégrés

h) Recherche des points appartenant à une surface connue par ses paramètres

Si on connaît les paramètres, et donc l'équation, d'une surface présente dans une zone de recherche, on peut envisager plusieurs méthodes pour décider quels sont les points qui lui appartiennent.

La méthode qui consiste à appliquer à chaque point l'équation de la surface et à voir si elle est vérifiée à une tolérance près, ne permet pas de déduire, dans le cas général, une indication sur la distance de ce point à la surface. Nous avons donc envisagé d'autres méthodes : elles sont dépendantes du type de surface considérée.

α) s'il s'agit d'un plan, d'une sphère, d'un cylindre de révolution ou d'un cône de révolution, on peut calculer pour chaque point la distance de ce point à la surface. On en déduit la liste des points appartenant à cette surface.

β) s'il s'agit d'un ellipsoïde de révolution de paramètres $(x_c, y_c, z_c, \varphi, \theta, a, c)$, on peut décider qu'un point appartient à cet ellipsoïde E_T s'il est à l'extérieur de l'ellipsoïde E_P de paramètres $(x_c, y_c, z_c, \varphi, \theta, a-c, c-\epsilon)$ et à l'intérieur de l'ellipsoïde E_G de paramètres $(x_c, y_c, z_c, \varphi, \theta, a+\epsilon, c+\epsilon)$ où ϵ est un nombre positif (ϵ est choisi en fonction de la qualité de l'image traitée et de la précision souhaitée). L'ellipsoïde E_T se trouve entre les deux ellipsoïdes E_P et E_G . Si $f_1(x, y, z) = 0$ est l'équation de E_P et $f_2(x, y, z) = 0$ celle de E_G , alors le point (x, y, z) appartient à l'ellipsoïde E_T s'il vérifie les deux inéquations suivantes :

$$f_1(x, y, z) > 0 \quad \text{et} \quad f_2(x, y, z) < 0 \quad (\text{figure V.2}).$$

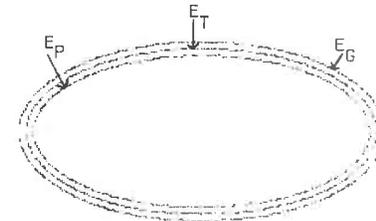


Figure V.2 : Coupe des trois ellipsoïdes E_T , E_P et E_G

La distance qui sépare E_T de E_G et E_T de E_P est irrégulière. Pour le montrer, nous considérons une coupe des deux ellipsoïdes E_T et E_G (figure V.3).

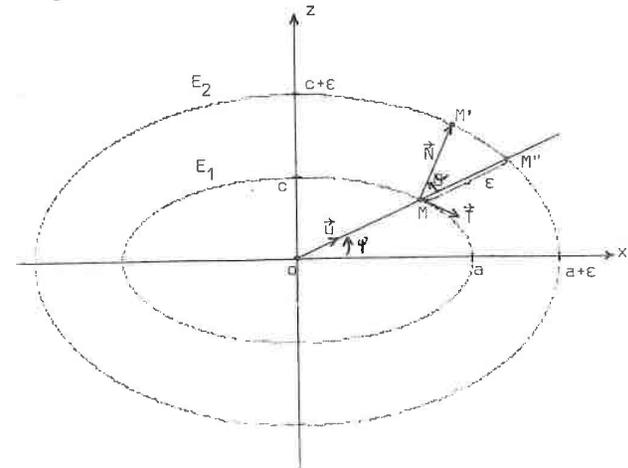


Figure V.3 : Coupe de deux ellipsoïdes

Nous cherchons alors à déterminer la distance minimale D_{\min} et la distance maximale D_{\max} séparant les deux ellipses E_1 et E_2 d'équations respectives

$$\frac{x^2}{a^2} + \frac{z^2}{c^2} = 1$$

et

$$\frac{x^2}{(a+\epsilon)^2} + \frac{z^2}{(c+\epsilon)^2} = 1$$

Calcul de la distance maximale -

La distance de M à M' est maximale lorsque θ est minimum. La valeur minimale de θ étant 0, on a

$$D_{\max} = \epsilon$$

Calcul de la distance minimale -

ϵ étant petit, nous supposons en première approximation que le triangle $MM'M''$ est rectangle en M' .

D'où

$$D_{\min} = \epsilon \cos \theta_{\max}$$

Or

$$\vec{OM} = \begin{pmatrix} a \cos \psi \\ c \sin \psi \end{pmatrix} \text{ car } M \text{ appartient à } E_1$$

On a alors :

$$\vec{t} = \frac{d\vec{OM}}{d\psi} = \begin{pmatrix} -a \sin \psi \\ c \cos \psi \end{pmatrix} \text{ et } \vec{N} = \begin{pmatrix} c \cos \psi \\ a \sin \psi \end{pmatrix}$$

$$\text{avec } D = \sqrt{a^2 \sin^2 \psi + c^2 \cos^2 \psi}$$

$$\text{On sait que } \vec{t} \cdot \vec{N} = \cos \theta, \text{ or } \vec{t} = \begin{pmatrix} a \cos \psi \\ c \sin \psi \\ D' \end{pmatrix}$$

$$\text{avec } D' = \sqrt{a^2 \cos^2 \psi + c^2 \sin^2 \psi}$$

$$\text{d'où } \cos \theta = \frac{ac}{DD'}. \text{ Posons } M = DD'.$$

Si θ est maximum, $\cos \theta$ est minimum et M^2 est maximum. Or on sait que, si la somme de deux nombres est constante, leur produit est maximum lorsque ces deux nombres sont égaux.

$$\text{D'où } M^2 \text{ maximum} \Rightarrow D^2 = D'^2$$

$$\text{On a donc } (a^2 - c^2) \cos 2\psi = 0. \text{ Or } a \neq c, \text{ d'où } \cos 2\psi = 0.$$

$$\text{Donc } \psi = \frac{\pi}{4} + k \frac{\pi}{2} \text{ avec } k \in \{0, 1, 2, 3\}$$

$$\text{On en déduit que } \cos \theta_{\max} = \frac{2ac}{a^2 + c^2}$$

Par suite,

$$D_{\min} = \frac{2ac}{a^2 + c^2} \epsilon \text{ en première approximation.}$$

Exemple : Soit $a = 50$ et $c = 30$, on a alors $D_{\max} = \epsilon$ et $D_{\min} = 0,88 \epsilon$.

La variation de la distance séparant E_G (ou E_P) de la surface exacte E_T est toujours faible. La méthode proposée peut donc être utilisée sans problème.

γ) des traitements de ce type peuvent être envisagés pour toutes sortes d'autres quadriques. Le coût de tels calculs est évidemment fonction du nombre de points de la zone de recherche. Mais il est beaucoup plus rapide que le calcul qui consiste à associer à un point les paramètres d'une surface.

Par exemple, pour une sphère, les opérations suivantes :

- résoudre au sens des moindres carrés un système de 25 équations à 9 inconnues,
- rechercher le centre d'une quadrique dont on connaît l'équation générale,
- rechercher les valeurs propres et les vecteurs propres d'une matrice 3 X 3 ,
- déduire de l'équation réduite la nature et les paramètres de la quadrique,

sont remplacées par le calcul de la distance d'un point à la sphère et un test permettant de décider si le point appartient ou non à la sphère.

i) Conclusion

La figure V.4 nous rappelle à quels niveaux se situent les gains relatifs aux différentes utilisations.

Recherche \	nombre de E_{par}	nombre de points de E_{par}	taille de E_{dep}	nombre de points de E_{dep}	taille de E_{dep}
a) de surfaces d'une certaine nature	X				
b) avec contraintes sur les paramètres		X	X		
c) dans une fenêtre					X
d) dans une région (*)				X	X
e) à grands pas				X	
f) de surfaces évidentes (*)				X	

E_{par} = espace des paramètres

E_{dep} = espace de départ

Une croix dans une case indique un gain

(*) des traitements ultérieurs sont nécessaires dans certains cas pour d) et dans tous les cas pour f)

Figure V.4 : Les gains correspondant aux différentes utilisations

V.2 UN EXEMPLE DE STRATEGIE D'UTILISATION

Prenons l'exemple du système d'interprétation de scènes TRIDENT (voir chapitre I, paragraphe I.5). Il dispose d'un outil de regroupement de régions et du système de détection de surfaces. Il peut combiner l'utilisation de ces deux outils de plusieurs façons ; en voici une :

L'analyse de scènes débute par la recherche de surfaces évidentes (voir paragraphe V.1 f). La détection des limites exactes de ces surfaces se fait par une méthode du type de celles indiquées au paragraphe V.1 h. Ces surfaces sont le point de départ d'un certain nombre d'analyses qui peuvent être menées en parallèle et qui peuvent communiquer entre elles [MAS - 84]. Pour pouvoir progresser, ces différentes analyses ont besoin de renseignements sur l'image. Les renseignements concernant les surfaces s'appliqueront à des régions dans un but de rapidité. On effectue donc un regroupement de régions sur l'ensemble des points qui n'appartiennent pas à une surface évidente. Puis, au fur et à mesure de ses besoins, chaque analyse interroge le système de détection de surfaces. Si elle dispose, par l'intermédiaire du modèle, d'informations sur la nature ou sur les paramètres de la surface à laquelle correspond la région considérée, elle les fournit au détecteur de surfaces. Pour prévenir des cas de mauvaises segmentations en régions, il nous paraît prudent d'adopter la méthode de recherche par échantillonnage proposée au paragraphe V.1 d β . Elle vérifie que la région considérée appartient effectivement à une seule surface.

Cette stratégie nous paraît intéressante car, tout en étant prudente, elle reste d'un coût modéré (voir paragraphe V.1). En effet, pour une image 256 X 256, elle implique une recherche de surfaces évidentes

sur environ 256 points. Si p_1 est la proportion de points appartenant à des surfaces évidentes, un regroupement de régions est effectué sur $(1 - p_1) 256^2$ points. Si t est la taille moyenne des régions et p_2 la proportion de région correspondant à une seule surface, le nombre total de points traité par la détection de surface est donné par la formule suivante :

$$NP = 256 + \frac{(1-p_1) 256^2}{t} e$$

où e est le nombre de points échantillonnés dans une région. En prenant $p_1 = \frac{1}{2}$, $p_2 = \frac{1}{3}$, $e = 10$ et $t = 200$, on a $NP = 420$. Un traitement ultérieur (du type de celui décrit au paragraphe V.1 h) pour déterminer si un point appartient à une surface (dont on connaît les paramètres) sera alors effectué sur le nombre de points suivants :

$$(1-p_1)(1-p_2) 256^2 + 256^2 p_1$$

soit 43 691 points avec les données précédentes.

Exemple pratique simple -

Analysons la scène de la figure V.5 dont nous disposons de l'image des distances et de l'image des couleurs.

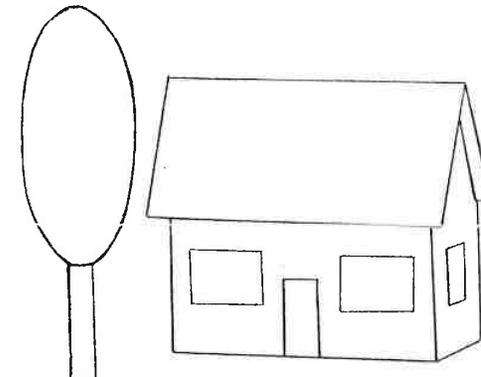


Figure V.5 : Scène à analyser

La recherche de surfaces évidentes permet d'y détecter un plan incliné rouge et un ellipsoïde vert. Ces deux surfaces sont hachurées et étiquetées S1 et S2 dans la figure V.6 .

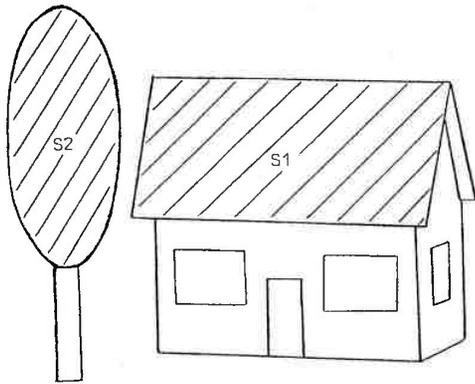


Figure V.6 : Détection de surfaces évidentes

Les points n'appartenant ni à S1 ni à S2 sont ensuite regroupés en 7 régions numérotées de R1 à R7 (figure V.7).

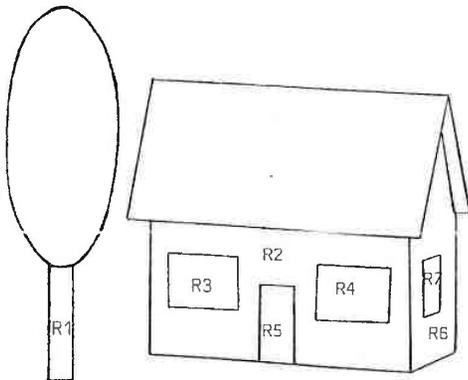


Figure V.7 : Regroupement de régions

Deux analyses peuvent alors débuter en parallèle : l'une à partir de S1 et l'autre de S2 .

Analyse partant de S1

Le modèle utilisé nous indique qu'un plan incliné rouge peut être un toit. On recherche donc un plan vertical en dessous du toit, soit dans la région R2 . Une détection de surfaces rapide (recherche par échantillonnage) confirme la présence de ce plan. Il reste alors à vérifier que la région voisine R6 est aussi un plan vertical et que R3, R4, R5 et R7 correspondent à des fenêtres ou à des portes. On en déduit qu'une maison est présente dans la scène.

Analyse partant de S2

Un ellipsoïde vert peut correspondre au feuillage d'un arbre. Pour confirmer cette hypothèse, on recherche un cylindre d'axe de révolution vertical en dessous du feuillage, soit dans la région R1 . Un tel cylindre est détecté par une recherche par échantillonnage sur R1 . Un arbre se trouve donc dans la scène.

Toutes les régions étant étiquetées, l'analyse est terminée.

V.3 CONCLUSION

L'intérêt principal de notre système de détection de surfaces est sa souplesse. Notre objectif a été de concevoir un outil utilisable de façon "intelligente" par un système de vision. Le choix d'une bonne utilisation dépend de ce dernier. Ses outils et sa stratégie guideront ce choix. Dans tous les cas, il est intéressant de fournir le maximum de renseignements pour limiter le nombre d'espaces des paramètres (en précisant la nature des surfaces recherchées) et la taille de ces espaces (en imposant des contraintes sur les paramètres des surfaces recherchées).

CONCLUSION

Nous avons réalisé un système de détection de plans et de quadriques pour des images de distances. Notre objectif principal a été de concevoir un outil utilisable de façon "intelligente" par un système de vision tel que TRIDENT. Nous nous sommes pour cela orientés vers l'utilisation de la transformée de Hough. Elle permet d'imposer des contraintes sur les paramètres des surfaces recherchées (exemple : recherche d'une sphère de rayon R). Ainsi, seuls les points de l'espace de départ appartenant à une surface du type recherché ont une image dans l'espace des paramètres. Le nombre de points de cet espace est donc peu important et, en général, les points sont peu dispersés. La stratégie "diviser pour régner" permet alors de converger rapidement vers les points d'accumulation.

Le traitement qui consiste à associer à un point de l'image à analyser une surface nécessite beaucoup de calculs : il faut, en particulier résoudre au sens des moindres carrés un système d'équations et, dans certains cas, rechercher les valeurs propres et les vecteurs propres d'une matrice 3×3 . Ce traitement peut se faire en parallèle s'il est répété un grand nombre de fois.

Cette détection de surfaces a été testée sur des images de synthèse et sur une image réelle. Les résultats obtenus sont satisfaisants et nous comptons étendre les tests dès que possible à d'autres images.

Grâce à sa souplesse, cet outil est bien adapté à un système d'interprétation de scènes. Nous avons étudié les différentes utilisations possibles et les gains correspondants. Nous attendons pour les tester que les autres parties de TRIDENT soient réalisées.

Une généralisation à d'autres familles de surfaces peut être envisagée. Il suffit de résoudre le problème géométrique qui consiste à représenter la famille de surfaces recherchée à l'aide d'une fonction $f(x,y,z)$ à partir de laquelle on pourra déduire la nature des différentes surfaces de cette famille et leurs paramètres.

BIBLIOGRAPHIE

- [ATH - 78] P. ATHERTON, K. WEILER, D. GREENBERG :
"Polygon Shadow Generation",
Computer Graphics, Volume 12, Number 3, August 1978,
pp. 275-281.
- [AYA - 82] M. AYACHE, J.D. BOISSONNAT, D. FAUGERAS, F. GERMAIN,
M. HEBERT, E. PAUCHON, J. PONCE :
"Vers un système de vision flexible",
Premières Journées Annuelles du Programme ARA, Septembre
1982, pp. 153-169.
- [BAL - 81a] D.H. BALLARD :
"Generalizing the Hough Transform to Detect Arbitrary
Shapes",
Pattern Recognition, Volume 13, Number 2, 1981, pp. 111-122.
- [BAL - 82] D.H. BALLARD, C.M. BROWN :
"Computer Vision",
Prentice Hall, 1982, p 293.
- [BAL - 81b] D.H. BALLARD, D. SABBAH :
"On Shapes",
Proceedings of the 7th IJCAI, Vancouver, B.C., Canada,
August 1981, pp. 607-612.
- [BLI - 77] J.F. BLINN :
"Models of Light Reflection for Computer Synthesized Pictures",
Computer Graphics, Volume 11, Number 2, Summer 1977, pp. 192-198.

- [BLI - 78] J.F. BLINN :
"Simulation of Wrinkled Surfaces",
Computer Graphics, Volume 12, Number 3, August 1978,
pp. 286-292.
- [BLI - 76] J.F. BLINN, M.E. NEWELL :
"Texture and Reflection in Computer Generated Images",
Communications of the ACM, Volume 19, Number 10, October
1976, pp. 542-546.
- [BRD - 81] R.A. BROOKS :
"Symbolic Reasoning among 3-D Models and 2-D Images",
Ph. D. Thesis, Stanford University, 1981.
- [BRO - 78] R.A. BROOKS, G. RUSSEL, T.O. BINFORD :
"The Acronym Model-Based Vision System",
Proc. ARPA Image Understanding Workshop Pittsburg, Nov. 1976,
pp. 145-151.
- [CAS - 65] G. CASANOVA :
"Mathématiques Spéciales",
Volume III, Librairie Classique Eugène Belin, Paris, 1965,
pp. 317-342.
- [CAS - 84] S. CASTAN :
"Extraction d'informations élémentaires dans les images",
journées de synthèse sur la vision, Rocquencourt,
23-24 janvier 1984.

- [COH - 82] P.R. COHEN, E.A. FEIGENBAUM :
"The Handbook of Artificial Intelligence",
Volume III, chapitre XIII, 1982, pp. 139-194.
- [CRO - 77] F.C. CROW :
"Shadow Algorithms for Computer Graphics",
Computer Graphics, Volume 11, Number 2, Summer 1977,
pp. 242-246.
- [DAN - 81] C. DANE, R. BAJCSY :
"Three-Dimensional Segmentation using the Gaussian Image and
Spatial Information",
Proceedings of the Conference on Pattern Recognition and
Image Processing, Dallas TX, 2-4 August 1981, pp. 54-56.
- [DUD - 72] R.O. DUDA, P.E. HART :
"Use the Hough Transform to Detect Lines and Curves in
Pictures",
Communications of the ACM, Volume 15, Number 1, January 1972,
pp. 11-15.
- [DUD - 76] R.O. DUDA, D. NITZAN :
"Low-Level Processing of Registered Intensity and Range Data",
Proceedings of the 3rd IJCP, Coronado, November 1976,
pp. 598-601.
- [FAN - 82] T.J. FANG, Z.H. HUANG, L.N. KANAL, B. LAMBIRO, D. LAVINE,
G. STOCKMAN, F.L. XIONG :
"Three Dimensional Object Recognition using a Transformation
Clustering Technique",
Proc. of the 6th International Conference on Pattern Recognition,
Munich, Germany, October 19-22, 1982, Volume 2, pp. 678-681.

- [FAU - 83] O.D. FAUGERAS, M. HEBERT :
"A 3-D Recognition and Positioning Algorithm using Geometrical
Matching between Primitive Surfaces",
Proceedings of the 8th IJCAI, Karlsruhe, West-Germany, 8-12
August 1983, Volume 2, pp. 996-1002.
- [FEL - 74] J.A. FELDMAN, Y. YAKIMOVSKY :
"Decision Theory and Artificial Intelligence : I.A. Semantics-
Based Region Analyzer",
Artificial Intelligence, Volume 5, 1974, pp. 349-371.
- [FIS - 83] R.B. FISHER :
"Using Surfaces and Object Models to Recognize Partially
Obscured Objects",
Proceedings of the 8th IJCAI, Karlsruhe, West-Germany, 8-12
August 1983, Volume 2, pp. 989-995.
- [FRE - 77] E.C. FREUDER :
"A Computer System for Visual Recognition using Active
Knowledge",
Proceedings of the 5th IJCAI, Cambridge, August 1977,
pp. 671-677.
- [GRI - 80] W.E.L. GRIMSON :
"Aspects of a Computational Theory of Human Stereo Vision",
Proceedings of the Image Understanding Workshop, April 1980,
pp. 128-149.

- [HAN - 80] A.R. HANSON, C.C. PARMA, E.M. RISEMAN :
"Experiments in Schema-Driven Interpretation of a Natural Scene",
Dig. Image Proc. and Anal., Simon and Haralick, editors,
Reidel, 1981.
- [HEB - 82] M. HEBERT, J. PONCE :
"A New Method for Segmenting 3-D Scenes into Primitives",
Proceedings of the 6th International Conference on Pattern
Recognition, Munich, Germany, October 19-22, 1982, Volume 2,
pp. 836-838.
- [HOR - 75] B.K.P. HORN :
"Obtaining Shape from Shading Information",
in the Psychology of Computer Vision, P.H. Winston, Ed.
New York : Mc Graw-Hill, 1975, pp. 115-155.
- [JUL - 71] B. JULESZ :
"Foundations of Cyclopean Perception", University of Chicago
Press, Chicago, 1971.
- [MAS - 84] G. MASINI, F. ZAROLI :
"Présentation de TRIDENT : un système d'interprétation d'images
tri-dimensionnelles",
Actes 4ème congrès AFCET, RF-IA, Paris, janvier 1984, tome I,
pp. 179-197.
- [MOH - 84] R. MOHR, B. WROBEL :
"La correspondance en stéréovision vue comme une recherche
d'un chemin optimal",
Actes 4ème congrès AFCET, RF-IA, Paris, janvier 1984, tome I,
pp. 71-79.

- [NEW - 79] W.M. NEWMAN, R.F. SPROULL :
"Principles of Interactive Computer Graphics",
Computer Science Series, Mc Graw-Hill Book Company, Second
Edition, part five, 1979, pp. 291-410.
- [O'R - 81] J. O'ROURKE :
"Dynamically Quantized Spaces for Focusing the Hough Transform",
Proceedings of the 7th IJCAI, Vancouver, B.C., Canada,
August 1981, pp. 737-739.
- [OSH - 79] M. OSHIMA, Y. SHIRAI :
"A Scene Description Method using Three-Dimensional Information",
Pattern Recognition, Volume 11, Number 1, 1979, pp. 9-17.
- [OSH - 83] M. OSHIMA, Y. SHIRAI :
"Object Recognition using Three-Dimensional Information",
IEEE Transactions on Pattern Analysis and Machine Intelligence,
Volume Pami-5, Number 4, July 1983, pp. 353-361.
- [PEL - 82] T. PELI, D. MALAH :
"A Study of Edge Detection Algorithms",
Computer Graphics and Image Processing, Volume 20, Number 1,
September 1982, pp. 1 -21.
- [PHO - 75] B.T. PHONG :
"Illumination for Computer Generated Pictures",
Communications of the ACM, Volume 18, Number 6, June 1975,
pp. 311-317.

- [ROB - 65] L.G. ROBERTS :
"Machine Perception of Three Dimensional Solids",
in Optical and Electro-Optical Information Processing,
J.T. Tippett et al., Eds., Cambridge, MA: MIT Press, 1965,
pp. 159-197.
- [SAK - 76] T. SAKAI, T. KANADE, Y. OHTA :
"Model-Based Interpretation of Outdoor Scene",
Proceedings of the 3rd IJCP, Coronado, November 1976,
pp. 581-585.
- [SHA - 78] S.D. SHAPIRO :
"Feature Space Transforms for Curve Detection",
Pattern Recognition, Volume 10, Number 3, 1978, pp. 129-143.
- [SHD- 77] R.E. SHOSTAK :
"On the sup-inf Method for Proving Presburger Formulas",
Journal of the ACM, Volume 24, Number 4, 1977, pp. 529-543.
- [SLO - 81] K.R. Jr SLOAN :
"Dynamically Quantized Pyramids",
Proceedings of the 7th IJCAI, Vancouver, B.C., Canada,
August 1981, pp. 734-736.
- [SLO - 80] K.R. Jr SLOAN, D.H. BALLARD :
"Experience with the Generalized Hough Transform",
Proceedings of the 5th International Conference on Pattern
Recognition, Miami Beach, Florida, December 1980, pp. 174-179.
- [SMI - 79] D.A. SMITH :
"Using Enhanced Spherical Images for Object Representation",
AI Memo, Number 530, AI Labo, MIT, 1979.

NOM DE L'ETUDIANT : MULLER Yolande

NATURE DE LA THESE : Doctorat 3ème cycle en Informatique

[YAM - 81] S. YAM, L.S. DAVIS :
"Image Registration using Generalized Hough Transform",
Proceedings of the Conference on Pattern Recognition and
Image Processing",
Dallas TX, 2-4 August 1981, pp. 526-533.

[ZAR - 83] F. ZAROLI :
"Réalisation d'un système d'analyse et d'interprétation
d'images tridimensionnelles",
Thèse de troisième cycle, Université de Nancy I, CRIN, 1983.

VU, APPROUVE ET PERMIS D'IMPRIMER

NANCY, le -4 MAI 1984

LE PRESIDENT DE L'UNIVERSITE DE NANCY I



RESUME

Nous avons réalisé un système de détection de plans et de quadriques dans des images tridimensionnelles. L'utilisation de la transformée de Hough permet de diriger la recherche vers un type particulier de surfaces, par exemple des cylindres d'axe de révolution vertical. Le problème de recherche de points d'accumulation dans l'espace des paramètres associé à une surface est résolu par une méthode fondée sur le principe "*diviser pour régner*". Des tests ont été réalisés sur des images artificielles et réelles. Grâce à sa souplesse, cet outil est bien adapté à un système d'interprétation de scènes en vision par ordinateur. Nous avons étudié les différentes utilisations possibles et les gains correspondants.

MOTS-CLES

Traitement d'images - Vision par ordinateur - Images de distances - Transformée de Hough - Détection de surfaces.