

THÈSE

présentée devant
l'UNIVERSITE CLAUDE BERNARD - LYON 1
pour l'obtention
du **DIPLÔME DE DOCTORAT**
(arrêté du 25 avril 2002)

préparée à l'école doctorale
Informatique et Information pour la Société
spécialité INFORMATIQUE

présentée et soutenue publiquement par

Julien RICARD

le 8 DÉCEMBRE 2005

Indexation et recherche d'objets 3D à partir de requêtes 2D et 3D

préparée au sein du laboratoire LIRIS
sous la direction de
David Coeurjolly et Atilla Baskurt

COMPOSITION DU JURY

Mme	Françoise Prêteux	Rapporteur	Professeur
M.	Francis Schmitt	Rapporteur	Professeur
M.	Marc Antonini	Examineur	Directeur de Recherche CNRS
Mme	Catherine Berrut	Examineur	Professeur
M.	Atilla Baskurt	Directeur de thèse	Professeur
M.	David Coeurjolly	Directeur de thèse	Chargé de Recherche CNRS
M.	Mohamed Daoudi	Membre invité	Maître de conférence, HDR



Résumé

Depuis l'émergence de la Conception Assistée par Ordinateurs (CAO), l'archivage, la structuration et la recherche dans des bases de modèles 3D constituent des domaines de recherche privilégiés de notre communauté. Cette thèse s'inscrit dans ce cadre, et plus généralement, dans la recherche d'information par le contenu. Dans ce contexte, nous nous intéressons plus particulièrement à l'indexation d'objets tridimensionnels, qui a pour objectif de caractériser la forme des objets, pour permettre de calculer des distances de similarité entre modèles. Pour cela, deux types d'approches ont été étudiées : les approches 3D/3D, comparant les objets entre eux et les approches 2D/3D, comparant les objets 3D à une image 2D.

La première partie de la thèse est consacrée à l'étude des descripteurs de forme 2D et 2D/3D. D'une part, nous avons généralisé le descripteur de forme 2D : Angular Radial Transform (ART), pour le rendre applicable sur les images couleurs et robuste aux déformations perspective. Dans un deuxième temps, nous avons étudié les approches 2D/3D, qui permettent de rechercher un objet 3D à partir d'une image 2D. La méthode d'indexation 2D/3D que nous proposons est une méthode dynamique qui compare directement l'image requête à l'objet 3D et s'adapte aux complexités de l'objet 3D et de la requête. Pour permettre une comparaison dynamique, notre système se base sur un système de représentation spectral des vues de l'objet, permettant d'extraire rapidement les vues et de mettre en place une recherche dynamique de la vue requête dans l'espace des vues.

La deuxième partie de la thèse porte sur l'indexation d'objets 3D. Pour indexer des grandes bases, nous avons réalisé une extension de la transformation ART en 3D, appelé : 3D Angular Radial Transform (3D ART). Ce descripteur possède un fort pouvoir de description, en gardant les propriétés de robustesse d'ART 2D. Pour permettre d'appréhender des grandes bases d'objets 3D, nous proposons également un système de navigation permettant de structurer les bases.

Mots clés : indexation, descripteur de forme, descripteur couleur, image, objet 3D.

Abstract

The increasing use of Computer-Aided Design (CAD) tools last decades has lead to huge 3D-model databases that still have to be correctly indexed. The context of this work is the content-based retrieval of images or models. We are interested in both 2D images and 3D-models indexing. We aim at extracting shape descriptors from images or models that allow to compute a similarity measure between them. We present two approaches. On the one hand retrieving a 3D-model from a 2D image-based query (called 2D/3D) and on the other hand, retrieving a 3D-model from another one (called 3D/3D).

The first part is dedicated to the 2D/3D approach. First, we generalize the 2D shape descriptor ART (Angular Radial Transform) by making it robust to rotation- and viewpoint-based deformations. In a second time, we propose a view-based approach that describes a 3D-model with a set of 2D views. More precisely, we compute for each model a set of views that are indexed by a 2D shape descriptor. This 2D/3D method is dynamic and adaptive to the complexity of both the 3D-model and the request. For a dynamic comparison we need a fast process for view extraction as it will be repeated a number of times. That is why we introduce a new spectral representation that characterizes all views from a given 3D-model.

In the second part of this work, we proposed a new fast 3D shape descriptor to index huge databases. The great number of 3D-models from the database imposes the use of a very compact shape descriptor. This one should contain maximum information in a minimum size to guarantee fast answer. We propose a 3D extension from ART called 3D Angular Radial Transform (3D ART). The latter has a strong descriptions capacity while keeping the 2D-ART robustness to deformations.

Keywords : indexation, shape descriptor, color descriptor, image, 3D object.

Table des matières

Introduction	3
I Etats de l'art	7
1 Etat de l'art en indexation de formes 2D	9
1.1 Introduction	11
1.2 Indexation de formes 2D	11
1.2.1 Descripteurs couleurs	11
1.2.2 Descripteurs formes	14
1.2.3 Transformation ART	16
1.2.4 Conclusion	23
1.3 Indexation d'objets 3D par requêtes images	23
1.3.1 Choix de la méthode de description	25
1.3.2 Choix du nombre de vues	26
1.4 Conclusion	29
2 Etat de l'art en indexation 3D	31
2.1 Introduction	33
2.2 Approches statistiques	33
2.2.1 Approches locales	33
2.2.2 Approches globales	34
2.3 Approches structurelles	35
2.3.1 Approches CAO	35
2.3.2 Approches basées sur les squelettes	37

TABLE DES MATIÈRES

2.3.3	Approche volumique	43
2.3.4	Représentation symbolique	45
2.4	Approches par transformées	46
2.4.1	Transformée de Hough	46
2.4.2	Approches basées sur les moments	47
2.5	Approches par vues	50
2.6	Pré-traitement	53
2.6.1	Centrage	54
2.6.2	Mise à l'échelle	54
2.6.3	Alignement	54
2.7	Critères d'évaluations	56
2.7.1	Base de données	56
2.7.2	Critères de comparaison	57
2.8	Conclusion	60
II Indexation par requêtes 2D		63
3 Indexation 2D		65
3.1	Introduction	66
3.2	Généralisation aux images couleurs	66
3.2.1	Analyse couleurs	66
3.2.2	Analyse de forme GART	71
3.2.3	Combinaison des recherches par la luminance et par la chrominance	78
3.3	Expérimentations	79
3.3.1	Recherche d'objets déformés par des transformations non planes	79
3.3.2	Recherche de logos dans une base d'images naturelles	81
3.4	Conclusion	85
4 Indexation 2D/3D		87
4.1	Introduction	88

4.2	Limitation de l'indexation 2D/3D	89
4.3	Recherche dynamique d'objets 3D	90
4.3.1	Système de représentation et de création des vues	90
4.3.2	Comparaison des vues	96
4.3.3	Processus d'optimisation	97
4.3.4	Processus de mise en correspondance	99
4.4	Expérimentations	101
4.4.1	Mise en correspondance d'un objet 3D à une vue 2D	101
4.4.2	Recherche sur une base d'objets	103
4.5	Conclusion	104
III Indexation par requêtes 3D		107
5 Indexation d'objets 3D		109
5.1	Introduction	110
5.2	3D Angular Radial Transform	110
5.2.1	Définition de ART 3D	110
5.2.2	Processus d'indexation	113
5.2.3	Expérimentations	113
5.3	Indexation 3D par vues	122
5.3.1	Indexation	122
5.3.2	Expérimentations	124
5.4	Conclusion	127
6 Expression analytique d'un maillage 3D		129
6.1	Introduction	130
6.2	Cadre de l'analyse	130
6.2.1	Principe général	130
6.2.2	Intégration sur un maillage 3D	132
6.3	Intégration d'un tétraèdre	133
6.3.1	Principe	133

TABLE DES MATIÈRES

6.3.2	Calcul des moments géométriques	134
6.3.3	Expression analytique de la transformée de Fourier	135
6.4	Application en analyse de formes 3D	139
6.4.1	Centrage et alignement d'un modèle 3D	139
6.4.2	Descripteur de forme 3D	144
6.4.3	Complexité	145
6.5	Conclusion	145
7	Navigation dans une base d'objets 3D	147
7.1	Introduction	148
7.2	Etat de l'art en navigation 3D	148
7.2.1	Indexation	149
7.2.2	Classification	149
7.2.3	Visualisation	150
7.3	Approche proposée	151
7.3.1	Indexation	151
7.3.2	Système de classification	153
7.3.3	Système de visualisation	156
7.4	Expérimentations	159
7.4.1	Critère de qualité de la partition	160
7.4.2	Performance du système de navigation	161
7.4.3	Validation du système de visualisation	162
7.5	Conclusion	163
8	Conclusions et perspectives	165
IV	Annexes	169
A	Développement d'une plate-forme logicielle pour Semantic-3D	173
A.1	Introduction	173
A.2	Usages	173
A.3	Architecture	174

A.4	Exemple de fonctionnement	175
A.4.1	Module de compression	176
A.4.2	Module d'indexation	176
A.4.3	Module de tatouage	176
A.4.4	Organisation des modules	176
A.4.5	Exemple d'exécution	177
A.5	Conclusion	177
B	Résolution de la transformée de Fourier d'un tétraèdre	179
C	Liste des publications de l'auteur	187
D	Index des auteurs cités	191

TABLE DES MATIÈRES

Table des figures

1	Organisation du projet RNRT Semantic-3D.	5
1.1	Exemple d'histogramme d'une image couleur.	12
1.2	Répartition des couleurs au sein des principaux espaces couleurs. . .	13
1.3	Exemple de similarités contour et région.	15
1.4	Classification des techniques de descriptions de formes 2D.	16
1.5	Parties réelles des fonctions de base d'ART.	17
1.6	Schéma de la procédure d'indexation.	18
1.7	Processus d'extraction de régions : extraction du carré englobant centré sur le centre de gravité de l'objet, mise à l'échelle avec les fonction de base sur l'espace $[-1; 1]^2$	19
1.8	Exemple d'images : à différentes échelles, à différentes orientations et à différents niveaux de bruits poivre et sel : 0, 35 et 50%.	21
1.9	Exemple de mise en correspondance 2D/3D et 3D/3D entre un ensemble de vues et une image requête et un autre ensemble de vues. .	24
1.10	Vues d'un objet 3D créées à partir des trois vues principales (a), sphère des vues et position des trois points de vues principaux (b).	25
1.11	Exemple de répartition de vues fixes sur la sphère des vues d'un objet : 4 vues, 7 vues, 8 vues et 9 vues.	27
1.12	Répartition aléatoire de points sur une sphère par la méthode de Sphere Picking (a) et répartition uniforme de points sur une demi-sphère par échantillonnage régulier des paramètres x_1 et x_2 sous la condition que $x_1^2 + x_2^2 \leq 0.5$ (b).	29
1.13	Exemple de répartition de vues par une méthode de sélection de vues variables. Un grand nombre de vues est créé uniformément répartis sur la Sphere des vues par une méthode de Sphere Picking, puis ne sont gardées que les plus caractéristiques.	30

TABLE DES FIGURES

2.1	Modèle CAD et son graphe AAG correspondant (LEE et SAITOU, 2001).	36
2.2	Décomposition par coque convexe (WACO et KIM, 1994).	37
2.3	Une feuille et son squelette avec diverses valeurs d'amincissement.	38
2.4	Un objet bidimensionnel, son squelette et la carte de distance de l'objet.	39
2.5	Exemple de squelette et d'exosquelette d'une image binaire et de leur fonction de répartition (LIPIKORN ET AL., 2002a).	41
2.6	Tore et graphe de Reeb correspondant pour une certaine fonction de poids. la fonction de point est sur cet exemple la hauteur (HILAGA ET AL., 2001).	42
2.7	Graphe de Reeb multi résolution (HILAGA ET AL., 2001).	42
2.8	Exemples d'appariements entre deux MRG (gauche) et deux aMRG (droite) : sans l'information géométrique, les jambes peuvent être appariées aux bras car ils sont topologiquement équivalents. En ajoutant l'information géométrique, tous les appendices sont bien appariés (TUNG et SCHMITT, 2005).	43
2.9	Les différents phases pour la reconnaissance par composantes proposée par Biederman.	44
2.10	Exemple de primitives d'objets définies par la théorie de Biederman : les géons et différentes connections créant des objets différents (IRANI et WARE, 2003).	45
2.11	Comparaison symbolique. (a) Caractéristiques de l'ontologie, (b) modèle caractérisé par la recherche de caractéristiques (NEAL et SHAPIRO, 2000).	46
2.12	Décomposition en harmoniques sphériques (KAZHDAN ET AL., 2003).	50
2.13	Comparaison de deux ensembles de vues représentant deux objets 3D (a) pour chaque objet 10 vues sont créées sur les sommets d'un dodécaèdre (b) les objets sont comparés en modifiant l'ordre de mise en correspondance des vues (c,d) (CHEN ET AL., 2003).	52
2.14	Exemple d'objets déformées : objet d'origine (a), objet translaté (b), objet ayant subit un changement d'échelle (c), objet ayant subit une rotation (d).	53
2.15	Exemple d'un objet mise à l'échelle, objet d'origine (a) et objet intégré à la sphère unité (b).	55
2.16	Objets de la base Princeton Shape Benchmark (SHILANE ET AL., 2004).	57
2.17	Objets de la base Renault Semantic-3D.	58

2.18	Exemple de classes de modèles 3D, première ligne : base Renault (classe : « attache de ceinture »), deuxième ligne : Princeton Shape Benchmark (classe : « escalier »).	59
3.1	Détection du seuil optimal par MEC.	68
3.2	Partitionnement du plan (x_1, x_2) par minimisation d'entropie croisée.	68
3.3	Segmentation d'une image couleur par partitionnement itératif de l'espace couleur CIE $L^*a^*b^*$.	69
3.4	Exemple d'histogramme 1D sur une des composantes CIE $L^*a^*b^*$ générée à partir de 3 couleurs dominantes.	71
3.5	Courbes Recall/Précision obtenue par ART binaire et ART niveaux de gris.	72
3.6	Exemple d'objets plans vus selon différents angles (a), exemple de fonctions de base d'ART projetées sur le plan support de l'objet à identifier (b).	73
3.7	Les coefficients ART obtenus en projetant l'image d'origine sur les FDB d'origine sont les mêmes qu'en projetant l'image projetée sur les FDB ayant subi la même projection.	74
3.8	Les fonctions de base sont projetées sur le plan image I selon les paramètres $\vec{n}(\theta, \phi)$ et p pour obtenir les fonctions de base projetées.	75
3.9	Exemple de fonctions de base projetées en respectant la perspective.	76
3.10	Schéma du processus d'indexation inverse.	77
3.11	Courbe de Recall pour 15 logos dans 50 positions aléatoirement tirées, pour 3 intervalles d'angle $\theta : [0^\circ, 25^\circ], [25^\circ, 45^\circ]$ et $[45^\circ, 65^\circ]$ et exemple de logos déformés (la première colonne contient les images d'origine).	80
3.12	Courbe de Recall sur une base de 15 logos dans 50 positions aléatoirement tirées avec et sans variations d'illuminations.	81
3.13	Schéma du processus d'indexation et de recherche. L'indexation calcule les coefficients ART généralisés en projetant les objets d'origine sur les fonctions de base projetées inverses et les stocke dans la base des coefficients. La recherche extrait l'objet à identifier d'une image naturelle, l'indexe en fonction des fonctions de base d'origine et compare les coefficients trouvés avec ceux de la base. Ce qui retourne un classement de cette dernière en fonction de l'objet clé.	82
3.14	Fenêtres de sélection des objets à extraire de l'image.	84

3.15	Courbes de Recall représentant la position moyenne de l'objet recherché suivant l'étude forme, l'étude couleur et un interclassement des distances forme et couleur.	85
4.1	Principe général de la mise en correspondance 2D/3D, l'espace des vues est échantillonné pour créer un ensemble de vues qui sera comparé à la vue requête	88
4.2	Principe du processus de recherche dynamique composé d'un système de représentation permettant l'extraction rapide de vues, d'une mesure de similarité entre vues et d'un processus d'optimisation.	90
4.3	Principe du théorème de Radon en deux dimensions. La coupe le long d'un axe passant par le centre d'une transformée de Fourier d'une forme discrète représente la transformée de Fourier de la projection de la forme.	91
4.4	Principe comparé de rendu par l'extraction de coupe spectral et par un processus basé sur le Z-buffer.	92
4.5	Exemple de rendu spectral : objet 3D maillé (a) est discrétisé pour obtenir un objet discret (b). Cet objet est transformé dans le domaine fréquentiel et stocké sous forme de spectre 3D discret (c). Le processus de rendu d'image, extrait un spectre 2D (d), le transforme dans l'espace réel, pour obtenir une image profondeur (e) et une image binaire (f).	93
4.6	Exemple de coupe dans le Fourier discret et image reconstruite : objet d'origine (a), sans interpolation (b) et avec interpolation (c).	95
4.7	Coupe de la transformée de Fourier, image profondeur et image binaire obtenues par le processus de récupération de vues.	95
4.8	Exemple de coupes extraites de la représentation de Fourier 3D et les images reconstruites à 4 résolutions différentes : 128, 64, 32 et 16.	96
4.9	Carte des distances calculées entre une vue requête (a) et un objet 3D (c). Pour chaque point de vue autour de l'objet, une vue est créée et comparée à la vue requête (c). L'image correspondant au minimum de la sphère des vues (b) est représentée par le point rouge sur la carte des distances (d).	97
4.10	Exemple de recherche du minimum dans un espace à deux dimensions. (a) Forme du simplex triangulaire pour un espace à deux dimensions. (b) Parcours vers le minimum pour un Simplex de taille fixe. (c) Parcours d'un processus du Simplex de taille variable.	99

4.11 Recherche de la vue minimum par l'algorithme du Simplex : la vue extraite, la vue trouvée et l'objet 3D (gauche), sphère des vues et parcours vers le minimum, en noir le point de départ de l'algorithme, en jaune le point d'arrivé (droite).	100
4.12 Processus dynamique de mise en correspondance de l'image requête avec un objet 3D.	101
4.13 Exemple de mise en correspondance entre un objet 3D et une image 2D : image requête (a), l'objet 3D (b), carte des distances sphériques et parcours vers le minimum (c), image la plus proche retrouvée (d) et objet aligné suivant le point de vue ayant fourni le minimum (e).	102
4.14 Courbe des Nearest Neighbor, First Tier et Second Tier, pour trois méthodes à sélection de vues fixes : 4, 7 et 94, et pour la méthode avec optimisation. L'abscisse représente le nombre d'itérations du processus d'optimisation.	104
5.1 Partie réelle des fonctions de base 3D.	112
5.2 Processus d'indexation.	114
5.3 Nuages de points représentant les critères : Nearest Neighbor, First Tier et Second Tier, pour plusieurs valeurs des paramètres n , m_θ et m_ϕ	116
5.4 Courbe de Recall/Precision pour certaines valeurs des paramètres n , m_θ et m_ϕ	117
5.5 Courbes de Recall/Precision pour plusieurs valeurs du paramètre S	118
5.6 Exemple de déformation due au bruit pour trois valeurs de distance : 0%, 5% et 10%.	119
5.7 Courbes de Recall/Precision des méthodes ART 3D et des harmoniques sphériques sur les bases d'objets 3D Renault et Princeton.	120
5.8 Position des vues pour les trois systèmes de vues caractéristiques fixe : 3, 7 et 25.	124
5.9 Processus de comparaison d'un objet 3D aux objets d'une base, en comparant les trois types de vues créées par le processus de rendu spectral : le spectre profondeur est extrait du spectre 3D de l'objet, l'image profondeur est créé par une transformée de Fourier inverse, qui peut être binarisée pour créer l'image binaire. Cette dernière peut également être créée directement à partir de l'objet 3D par un processus de rendu basé sur le Z-Buffer.	126

6.1	Principe de l'intégration, un cône Δv est formé par un élément de surface Δs est l'origine du repère.	132
6.2	Exemple de décomposition tétraédrique d'un objet 3D. L'objet Q peut être vu comme la somme CSG : $Q = Q_4 - Q_1 - Q_2 - Q_3$	133
6.3	Exemple de Fourier Volume Rendering de l'objet 3D Bunny. La première ligne représente, l'objet 3D maillé et les parties réelles et imaginaires du spectre extrait. La seconde ligne représente les parties réelles, imaginaires et la valeur absolue de l'image profondeur reconstruite.	138
6.4	Exemple de Fourier Volume Rendering de l'objet 3D maillé Fandisk, pour deux orientations quelconques et pour deux tailles du spectre extrait.	139
6.5	Exemple de rendu spectral par un processus de Fourier Volume Rendering de l'objet 3D maillé Fandisk par trois types de rendu : discret avec interpolation (a), discret sans interpolation (b), par l'expression analytique(c).	140
6.6	Exemple de centrage, d'alignements et de mise à l'échelle sur une classe d'objets de la base Semantic-3D, avec une normalisation sur les points (a) et une normalisation volumique(b).	142
6.7	Exemple d'alignement et de centrage d'un objet 3D remaillé, en calculant les paramètres sur le nuage de points(a), sur l'objet discrétisé (b) et sur la représentation tétraédrique de l'objet(c).	143
7.1	Exemple de carte 2D affichant des images (IDRISSI ET AL., 2004).	151
7.2	Exemple d'affichage d'une base d'objets en trois dimensions. (a) FSN de ROGERS ET AL. (1996), (b) Pathfinders Networks de CHEN et MORRIS (2003) et les arbres de cônes proposés par ROBERTSON ET AL. (1991).	152
7.3	Pourcentage d'information apportée par les dimensions après une analyse en composantes principales.	153
7.4	La première composante principale représentée par une flèche permet de diviser le nuage de points en 2 groupes (a) et de définir leurs centres de gravité comme centre des groupes (b).	154
7.5	Les croix sont les centres de classe, les ellipses noires les groupes et les cercles rouges la distance moyenne des points d'un groupe à son centre. (a) écart entre 2 groupes tenant compte de la distance moyenne des points de chaque groupe à leur centre respectif. (b) écart entre 2 groupes basés sur la distance entre les centres de groupe.	156

7.6	Exemple de création et d'affinage de l'arbre hiérarchique. (a) Création de l'arbre hiérarchique ascendant.(b) élimination de 2 niveaux sur 3.	157
7.7	Illustration du fonctionnement de l'interface. (a) Au cours de la navigation, l'utilisateur sélectionne l'objet en bas de l'écran. (b) Celui-ci devient rouge et l'affichage est modifié en conséquence.	158
7.8	Processus de Relevance Feedback. (A) Structure originelle (B) Carte 2D des similarités des objets du groupe sélectionné (C) Modification de l'agencement des objets (D) Nouvelle structure calculée en tenant compte des modifications.	160
A.1	Architecture globale de l'application Semantic-3D.	175
A.2	Architecture générale de l'application.	178

TABLE DES FIGURES

Liste des tableaux

1.1	Tableau des intervalles de quantification et des valeurs de la quantification inverse.	20
3.1	Nombre de descripteurs ART calculé, en ligne et hors ligne, durant le processus d'origine et le processus optimisé.	77
4.1	Comparaison en terme de Nearest Neighbor, First Tier et Second Tier, des 4 méthodes testées.	105
5.1	Nearest Neighbor, First Tier, Second Tier en fonction du nombre de descripteurs.	115
5.2	Nearest Neighbor, First Tier, Second Tier en fonction de la taille de la discrétisation S	115
5.3	Distances obtenues pour plusieurs déformations.	117
5.4	Comparaison des performances des méthodes ART 3D et des Harmoniques sphériques en terme de Nearest Neighbor, First Tier et Second Tier, sur les deux bases Renault et sur la base Princeton.	119
5.5	Comparaison de la taille (en nombre flottant) et du temps d'indexation par objet (en secondes) entre ART 3D et la représentation en harmoniques sphériques (HS).	119
5.6	Comparaison des descripteurs de forme (DF) et performances associées. N_{iD} est le nombre d'éléments d'un vecteur iD , $i \in \{1, 2, 3\}$, V_{type} désigne les valeurs classiquement associées aux paramètres N_{iD} , N le nombre des facettes du maillage, D la taille de la discrétisation et V le nombre de vues nécessaires pour les approches 2D/3D.	121
5.7	Performances des différents descripteurs de formes selon les critères : SBE , FT , SBE_{tot} et FT_{tot}	122
5.8	Statistique sur la base RNRT Semantic-3D Renault contenant 116 objets répartis en 25 classes.	125

5.9	Statistique sur la base RNRT Semantic-3D Renault contenant 758 objets répartis en 75 classes.	125
6.1	Comparaison des complexités des processus d'indexation par discrétisation de l'objet ou par intégration sur le maillage par décomposition tétraédrique. T est le nombre de triangles composant l'objet 3D et N est la taille de la grille de discrétisation.	146
7.1	Qualité de partitionnement et temps d'exécution comparés pour notre algorithme et les nuées dynamiques floues avec des jeux d'essais de 116 et de 758 points.	161
7.2	Performances du système de Relevance Feedback en terme de pourcentage moyen d'objets bien classés suite aux modifications de la base par l'utilisateur.	162

f

Introduction

Les rapides avancées des technologies de l'informatique ont révolutionné les moyens de communications permettant d'échanger de nombreux documents. Les documents multi-médias peuvent être constitués d'objets de multiples natures : textes, images, vidéos, modèles 3D... La rapide extension de ce domaine a demandé la mise en place de nouveaux outils permettant de manipuler efficacement ces données et en particulier, des outils d'indexation pour rechercher et classer les grandes bases de données disponibles. La facilité de création de ces documents, que ce soit par des appareils photos numériques personnels, des caméscopes, par des scanners 2D ou 3D ou par des logiciels de création, a entraîné une forte augmentation des bases d'objets multi-médias et, par la même, des besoins de recherches et de classifications. Si depuis une vingtaine d'années, des recherches ont été réalisées sur l'indexation d'images 2D, il a fallu attendre la fin des années 90 pour voir apparaître les premiers travaux en indexation d'objets 3D.

Dans le cas des objets 3D, les processus de conception assistée par ordinateurs (CAO) ont supplanté le dessin technique et se sont répandus rapidement dans les bureaux de designs et des cabinets d'ingénieries. Ces applications professionnelles créèrent les premiers modèles 3D et commencèrent à échanger ces derniers entre les concepteurs et les sites de fabrication. La révolution des communications de la fin des années 90, provoqua l'apparition de nouvelles applications grand public utilisant des objets 3D. Ces applications, comme les forums de discussions 3D, où les utilisateurs se déplacent dans un environnement 3D et sont identifiés par un avatar en trois dimensions, ou la diffusion aux grandes publiques d'objets 3D sous forme de notice de réparation, ont fortement augmenté le nombre d'utilisateurs de modèles 3D et les bases d'objets.

Pour répondre à ces problématiques, il faut développer de nouvelles fonctionnalités permettant de manipuler efficacement des objets 3D : format d'échange, programme de rendu, programme de conception, compression, sécurisation... Une fonctionnalité importante demandée par les utilisateurs de bases d'objets 3D, est la possibilité de structurer les bases et de permettre de rechercher des objets au sein d'une grande collection. La nature des objets 3D demande la mise en place de procédés d'indexations propre à ces données. De nombreuses méthodes ont été proposées par la communauté et conduisent à la mise en place de systèmes permettant de décrire les objets 3D en fonction de leur forme et de faire des requêtes par similarité.

Les recherches peuvent être de plusieurs natures. Il est possible de rechercher un objet 3D dans une base à partir d'un objet requête, mais il est également possible de rechercher des objets 3D à partir de requête 2D. Il est plus facile de prendre une photographie ou de dessiner un schéma d'une vue d'un objet 3D et de rechercher au départ de cette vue que de numériser ou de modéliser un objet. D'autre part, pour structurer les bases d'objets 3D, des systèmes de navigation ont vu le jour. Ces systèmes affichent à l'utilisateur la base structurée et permet de naviguer entre les différentes classes. Ceci permet d'appréhender la base dans sa globalité et de trouver un objet sans avoir besoin de disposer d'une requête.

Le projet RNRT Semantic-3D

Le projet RNRT Semantic-3D¹ a été proposé pour mettre en place des nouvelles fonctionnalités de manipulation d'objets 3D. Ce projet regroupe quatre laboratoires de recherche : ENIC TELECOM de Lille, EURECOM de Sophia-Antipolis, INT d'Evry et le LIRIS de Lyon et deux entreprises : le constructeur automobile RE-NAULT et l'entreprise d'informatique Mercury. Ce projet est financé par le ministère français de la recherche dans le cadre du réseau national de recherche en télécommunication² (RNRT). Son objectif principal est l'exploration de techniques et d'outils pour la réalisation de nouveaux services pour l'exploitation de contenus 3D à travers le Web ou différents objets communicants.

Le projet Semantic-3D a pour but en particulier la transmission de méthodes de réparation composées d'objets 3D animés et d'instructions. Typiquement, un réparateur, en concession ou en dépannage, interroge le central des méthodes de réparation par des mots clés, des images ou des objets 3D. La méthode de réparation lui est alors envoyée sous forme de scènes animées, permettant de faire la réparation de manière optimum. Il faut pour envoyer les scènes 3D dans les meilleures conditions, mettre en place des processus de compression qui limitent la taille des données à transmettre tout en garantissant un rendu suffisant, et de tatouage, qui cachent une marque dans un objet 3D et permettent de repérer la provenance des pièces 3D.

Pour répondre aux trois problématiques principales : la compression, le tatouage et l'indexation, et mettre en place l'architecture de l'application, le projet est organisé autour de quatre sous-projets, comme le montre la figure 1. Le sous-projet 1 concerne la compression de modèle 3D, le sous-projet 2 l'indexation, le sous-projet 3 le tatouage et le sous-projet 4 l'intégration des fonctionnalités dans une interface.

¹<http://www.semantic-3d.net>

²<http://www.telecom.gouv.fr/rnrt/>

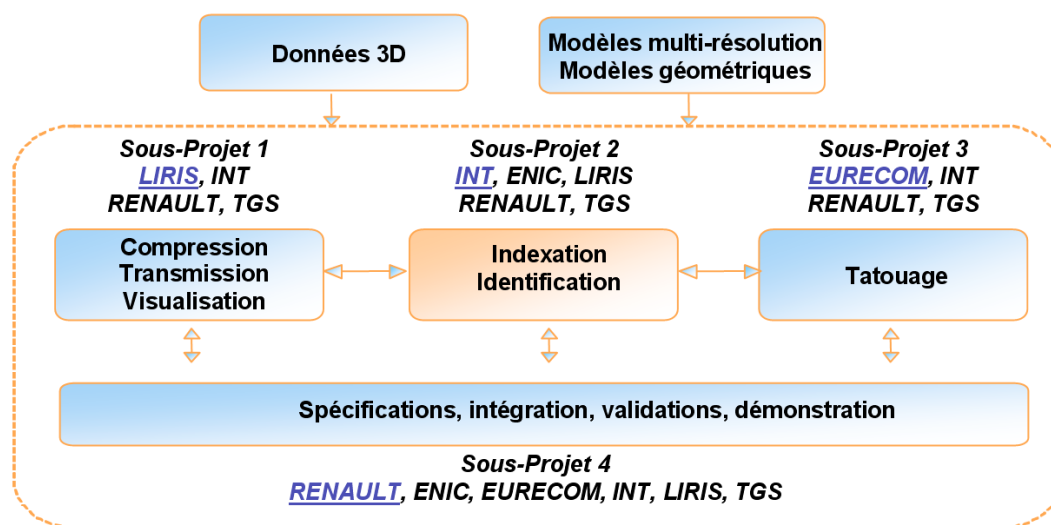


FIG. 1 – Organisation du projet RNRT Semantic-3D.

Objectifs de la thèse

Les objectifs de cette thèse sont de proposer des nouvelles méthodes d’indexation, à la fois sur des images 2D que sur des objets 3D par des requêtes 2D et 3D. Ils s’inscrivent dans le contexte du deuxième sous-projet du projet RNRT Semantic-3D. Les travaux que nous avons réalisés portent sur des images 2D et sur des objets 3D. Cette thèse est séparée en trois parties. Après un état de l’art (première partie), nous avons choisi de séparer nos travaux en fonction du type de données traitées. La deuxième partie contient l’étude des images 2D et l’indexation 2D/3D et la troisième partie, les travaux en indexation 3D.

Afin de correctement poser les problématiques traitées dans cette thèse, la première partie est consacrée à l’étude de l’état de l’art des deux domaines qui nous intéressent : l’indexation d’images 2D et l’indexation d’objets 3D. Cet état de l’art présente les principales méthodes d’indexation de forme, que ce soit pour des objets 2D ou des objets 3D et dévoile les différentes méthodes proposées au sein de la communauté en soulignant les limites et les avantages de chacune.

La deuxième partie de cette thèse est consacrée aux travaux que nous avons réalisés en indexation et recherche à partir de requêtes 2D. Ceci regroupe les méthodes que nous avons mises en place en indexation d’images couleurs et les méthodes d’indexation d’objets 3D par vues. Nous avons pris le parti de regrouper au sein de cette partie, les méthodes d’indexation 2D et les méthodes d’indexation 3D par vues, car elles se basent toutes les deux sur les mêmes méthodes d’indexation de formes 2D. Nous présentons en particulier la généralisation de la transformation ART pour la rendre applicable aux images couleurs et pour lui permettre de retrouver des objets

déformés par des transformations homographiques. Puis, nous expliquons la méthode de recherche d'objets 3D par des requêtes 2D.

La troisième partie de cette thèse porte sur les méthodes que nous avons proposées en indexation d'objets 3D par des requêtes 3D. Nous détaillons en particulier le nouveau descripteur de formes 3D : 3D Angular Radial Transform, descripteur de formes 3D par approche région basé sur les moments. Nous présentons également les études que nous avons menées sur les calculs d'expression analytique sur le maillage des objets 3D et l'application que nous en avons faite en analyse d'objets 3D. Le dernier chapitre de cette partie est consacré au système de navigation dans une base d'objets 3D.

Première partie

Etat de l'art

1

Etat de l'art en indexation de formes 2D

Sommaire

1.1	Introduction	11
1.2	Indexation de formes 2D	11
1.2.1	Descripteurs couleurs	11
1.2.2	Descripteurs formes	14
1.2.3	Transformation ART	16
1.2.3.1	Définition de la transformation ART	17
1.2.3.2	Indexation	18
	Pré-traitements	18
	Calcul des coefficients	19
	Normalisation des coefficients	19
	Quantification des coefficients	20
	Mise en correspondance	20
1.2.3.3	Propriétés de la transformation ART	21
	Robustesse aux changements d'échelle et aux transformations.	21
	Invariance aux rotations.	22
	Robustesse à la présence de bruit.	22
1.2.4	Conclusion	23
1.3	Indexation d'objets 3D par requêtes images	23
1.3.1	Choix de la méthode de description	25
1.3.2	Choix du nombre de vues	26
1.3.2.1	Méthodes de sélection de vues fixes	26
1.3.2.2	Méthodes de sélection de vues variables	28

1.4 Conclusion 29

1.1 Introduction

Ce chapitre présente l'état de l'art en indexation de formes 2D et en indexation d'objets 3D par des requêtes 2D. Ceci regroupe les principaux travaux sur lesquels nous nous sommes basés pour développer nos approches, que ce soit pour indexer des bases d'images en fonction de la forme, ou pour caractériser des objets 3D en fonction de leurs vues, pour permettre de faire de la recherche d'objets 3D à partir de requêtes image. Les travaux que nous présenterons dans la suite de cette thèse, et principalement les chapitres 3 et 4, sont inspirés de cet état de l'art.

1.2 Indexation de formes 2D

L'indexation d'images consiste à extraire l'information significative des images pour permettre de faciliter la comparaison ultérieure. La nature de ces documents multimédia exige la séparation des phases de description et de recherche pour permettre que cette dernière se déroule en temps utilisateur, temps relativement court durant lequel l'utilisateur attend la réponse du système. Elle varie principalement en fonction du choix des descripteurs employés et des techniques associées à leurs extractions. On peut différencier les descripteurs en deux grands niveaux :

- les descripteurs bas-niveau : décrivent le contenu bas niveau de l'image, principalement en terme de couleurs, de textures et de formes. Ce sont les descripteurs les plus utilisés dans les systèmes actuels, car les plus simples à mettre en place.
- Les descripteurs haut-niveau : décrivent le contenu « sémantique » de l'image et sont principalement des mots clés fournis par l'utilisateur, car l'extraction automatique est un problème non résolu actuellement.

Les travaux que nous avons réalisés portent principalement sur l'étude de descripteurs bas-niveau, par la mise en place de descripteurs couleur et forme. Ce chapitre reviendra sur ces deux classes de descripteurs et nous présenterons en particulier le descripteur de formes par approche région, « Angular Radial Transform »(ART) qui sera étudié en détail et généralisé à la couleur et au 3D dans la suite de cette thèse.

1.2.1 Descripteurs couleurs

La plus grande majorité des systèmes de recherche d'images par le contenu se base sur la description des couleurs composant les images. L'approche la plus courante dans la littérature est l'histogramme couleur. Pour les images couleurs, où chaque pixel est décrit par trois valeurs représentant l'intensité de rouge, de bleu et de vert, trois histogrammes sont construits. La figure 1.1 représente une image couleur et son histogramme couleur Rouge, Vert, Bleu (RVB ou RGB).

De nombreux travaux utilisent l'histogramme couleur comme descripteur et pro-

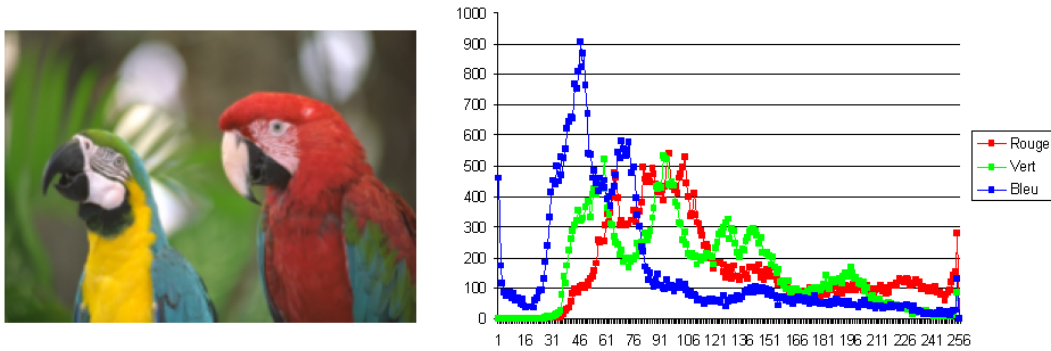


FIG. 1.1 – Exemple d’histogramme d’une image couleur.

posent diverses distances associées permettant de mesurer la similarité entre deux histogrammes. Ces approches se basent sur les travaux de SWAIN et BALLARD (1991), qui ont proposé de décrire la couleur d’une image à l’aide de son histogramme couleur et ont défini l’intersection d’histogrammes comme une mesure de similarité :

$$d(Q, I) = \sum_{j=1}^n \min(Q_j, I_j) \quad (1.1)$$

que l’on peut normaliser par :

$$H(Q, I) = \frac{\sum_{j=1}^n \min(Q_j, I_j)}{\sum_{j=1}^n I_j} \quad (1.2)$$

où n est le nombre de discrétisation des histogrammes utilisés dans le calcul et Q_i et I_j désignent respectivement deux valeurs des histogrammes des images Q et I à comparer. Cette mesure semble la plus référencée dans la littérature et fut à la base de nombreux travaux. Les principales variantes proposées se firent sur le choix de l’espace couleur de représentation des images, sur la distance de similarité entre histogrammes et sur le processus de quantification.

De nombreuses méthodes de descriptions d’images proposent de caractériser la couleur dans certains espaces couleurs pour profiter des propriétés de ces derniers. La figure 1.2 montre les principaux espaces couleurs utilisés en indexation d’images. On peut principalement citer :

- R, G, B (Rouge, Vert, Bleu) : est le plus utilisé car la plupart des images originelles sont codées dans cet espace couleur, ce qui ne nécessite pas de transformation inter espace couleur, donc facilement applicable.
- H, S, V : chaque composante représente respectivement la teinte, la saturation et la luminance.
- Y, C_b, C_r : est utilisé dans les normes MPEG 1, 2 et 4, ses composantes sont décorréelées et de faibles dynamiques, ce qui permet de bons taux de compression.

- *CIE L*a*b** ou *CIE Luv*, : sont des espaces couleurs perceptuellement uniformes. Ces espaces ont été créés dans le but de rendre plus homogène l'espace des couleurs et de permettre de mesurer uniformément les distances entre couleurs en tout point de l'espace. Deux couleurs proches dans ces espaces couleurs sont proches perceptuellement. Ces espaces sont grandement utilisés dans les systèmes de comparaison d'images.

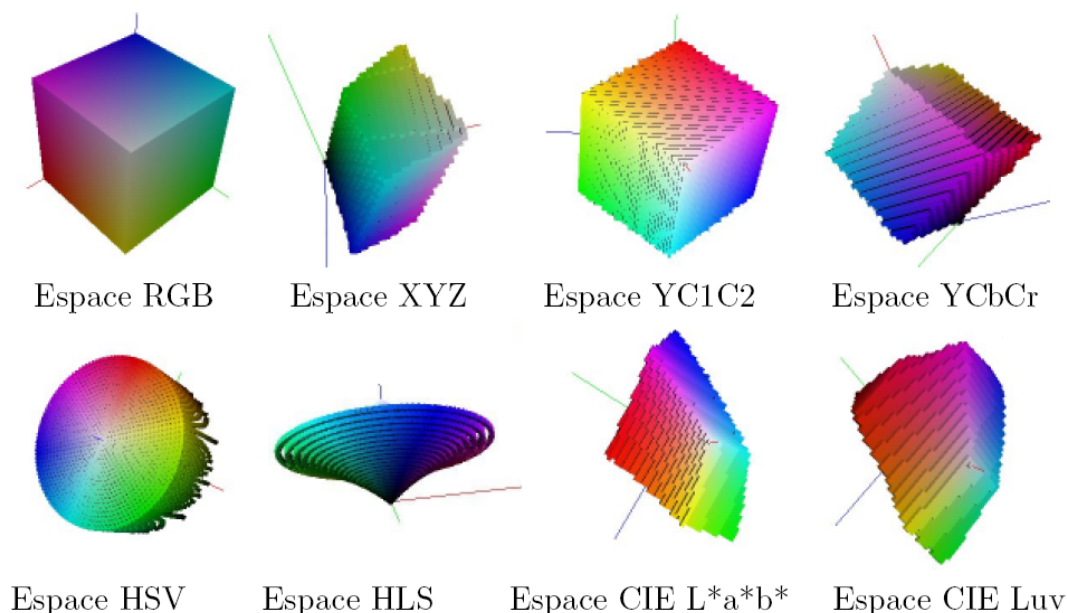


FIG. 1.2 – Répartition des couleurs au sein des principaux espaces couleurs.

Les méthodes de caractérisation de la couleur par des histogrammes varient en fonction de la distance inter histogrammes choisie. Ces distances se basent principalement sur les travaux de SWAIN et BALLARD (1991). On peut noter l'approche de HAFNER ET AL. (1995) qui propose une distance pondérée entre histogrammes afin de tenir compte de la similarité entre couleurs :

$$d(Q, I) = \sqrt{\sum_i \sum_j a_{ij} (Q_i - I_j)(Q_i - I_j)} \quad (1.3)$$

où Q_i et I_j désignent deux valeurs des histogrammes des images Q et I , et a_{ij} représente la similarité entre les deux couleurs i et j , basé sur la distance euclidienne entre deux couleurs selon la formule :

$$a_{ij} = 1 - \frac{d_{ij}}{\max_{ij}(d_{ij})} \quad (1.4)$$

où d_{ij} est la distance euclidienne entre deux couleurs dans l'espace choisi.

Une autre approche pour caractériser la couleur est l'utilisation des moments statistiques, l'histogramme étant une densité de probabilité. STRICKER et DIMAI

(1996) proposent de calculer pour chaque canal (axe de l'espace de représentation) la moyenne, la variance et le moment statistique d'ordre 3, afin de caractériser la distribution de la couleur. Ils définissent ensuite une distance entre deux images par :

$$d(Q, I) = \sum W_{i1} |\mu_i^Q - \mu_i^I| + W_{i2} |\sigma_i^Q - \sigma_i^I| + W_{i3} |s_i^Q - s_i^I| \quad (1.5)$$

où μ_i^I , σ_i^I et s_i^I représentent, respectivement, la moyenne, la variance et le moment statistique d'ordre 3 du canal i de l'image I , les W_{ij} étant des pondérations associées à chaque terme.

CINQUE ET AL. (2001) proposent l'histogramme chromatique spatial, qui intègre la position des pixels ayant la même couleur. Après quantification de l'image, les auteurs calculent, pour chaque couleur dominante k , le barycentre $b_I(k)$ des pixels de couleurs identiques, ainsi que la dispersion $\sigma_I(k)$ de ces pixels. La mesure de similarité est définie par :

$$f_s(Q, I) = \sum_{k=1}^c \min(h_Q(k), h_I(k)) \left(\frac{\sqrt{2} - d(b_Q(k), b_I(k))}{\sqrt{2}} + \frac{\min(\sigma_Q(k) - \sigma_I(k))}{\max(\sigma_Q(k) - \sigma_I(k))} \right) \quad (1.6)$$

où c est le nombre de couleur h de l'image.

Globalement, l'histogramme couleur reste le descripteur le plus utilisé et différentes distances lui ont été associées. Bien qu'il ne contienne qu'une information partielle en raison de l'absence d'indication sur les caractéristiques spatiales, l'histogramme garde un fort pouvoir de description, ce qui explique sa grande utilisation.

1.2.2 Descripteurs formes

En deux dimensions, on peut définir deux notions de similarité entre formes, la similarité de région et la similarité de contour. Ceci est montré dans la figure 1.3. Les objets de la première ligne ont des similarités spatiales dans la distribution des pixels et sont similaires selon des critères régions. Alors qu'ils sont clairement différents selon leur contour. En considérant la similarité de contour, les objets de chaque colonne sont similaires. En prenant pour requête l'objet situé dans la première ligne et la deuxième colonne, la similitude région fournira les objets de la première ligne, alors que la similarité contour fournira les résultats de la deuxième colonne.

Des techniques de représentation et de description de formes basées contours et les méthodes basées régions peuvent être divisées en deux sous parties : les approches structurelles et les approches globales. Ces sous-classes sont caractérisées si les descripteurs de forme sont calculés sur l'ensemble de la région ou du contour, ou sur un ensemble de primitives. La hiérarchie entière de la classification et les principales méthodes couramment citées dans la littérature sont montrées dans la figure 1.4.

On peut noter que la plupart des méthodes sont basées sur des approches globales. Les approches structurelles basées sur le contour ou basées sur la région,

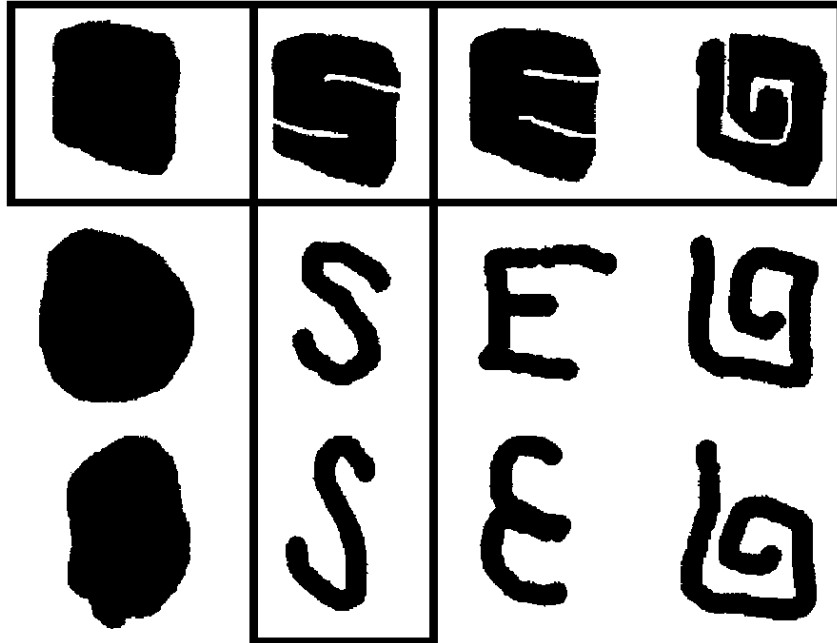


FIG. 1.3 – Exemple de similarités contour et région.

décomposent la forme en éléments de base et comparent les ensembles de primitives, ou les graphes les représentant, pour mesurer la similarité entre formes. En effet, de nombreuses méthodes structurales décrivent la forme comme un graphe reliant les éléments de base et comparent deux objets en faisant de la comparaison de graphe. Deux problèmes majeurs ont freiné ces méthodes : la segmentation et la comparaison. Il faut pour représenter convenablement une forme par un ensemble de primitives, avoir une phase de segmentation, qui extrait de la forme une représentation structurale. Il n'y a pas une définition formelle de la forme et l'on ne peut pas connaître le nombre de primitives requises pour correctement la décrire, ce qui entraîne un calcul complexe. D'autre part, le calcul de la mesure de similarité entre formes est non métrique, contrairement aux méthodes sur la forme globale, ce qui permet de faire de la comparaison de sous graphes, mais exige un coût de calcul élevé. De plus, ces méthodes ne réagissent pas bien à l'apparition de bruit sur les objets à décrire, ce qui limite leur utilisation dans des systèmes de reconnaissance d'images naturelles.

Parallèlement à cela, les méthodes globales permettent de calculer directement les descripteurs de forme sans une phase de sous-segmentation et avec une phase de comparaison numérique, ce qui permet de garantir un coût calculatoire d'indexation et de recherche faible. Les méthodes basées sur les régions traitent la forme dans son ensemble et utilisent l'information de tous les pixels de la région. Ces méthodes mesurent la distribution des pixels de la région et sont peu affectées par le bruit

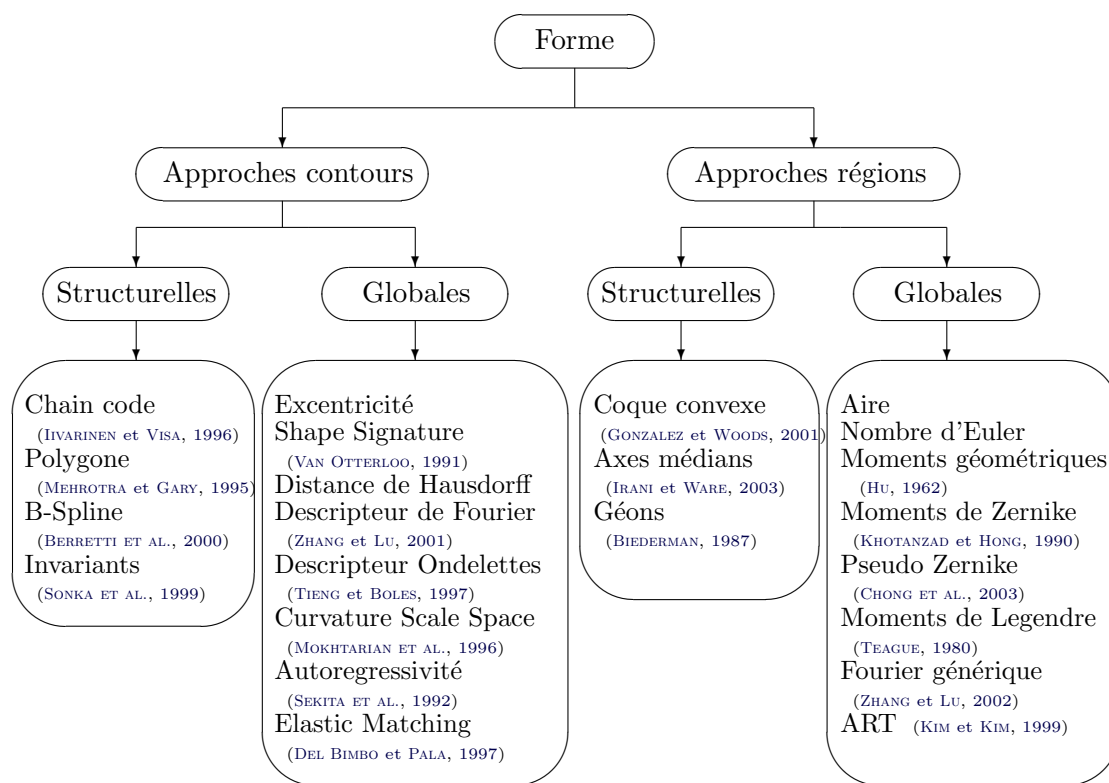


FIG. 1.4 – Classification des techniques de descriptions de formes 2D.

et par les variations de formes. En particulier, les méthodes populaires basées sur les régions sont des méthodes de moments. Ces méthodes extraient la distribution statistique des pixels de la région. Les moments portent des significations physiques liées à la distribution de pixels de la région.

Pour ces différentes raisons, nous nous sommes principalement intéressés aux méthodes d'analyses de formes basées régions, et plus précisément à la transformation ART, Angular Radial Transform, que nous avons généralisée pour la rendre applicable aux images naturelles et aux objets 3D.

1.2.3 Transformation ART

Un descripteur de formes par approche région caractérise la distribution des pixels dans un objet ou une région 2D. Il est basé sur l'étude des frontières et des pixels internes de la région à décrire. Il s'applique à un grand nombre d'objets, comme des objets complexes constitués de multiples régions discontinues ou des objets simples avec ou sans trou. Les techniques d'analyse de formes basées sur les moments appartiennent à la classe des descripteurs de formes par approche région

(TEH et CHIN, 1988), comme les moments de Zernike (KHOTANZAD et HONG, 1990), qui furent longtemps pressentis comme le descripteur de formes par approche région à intégrer à MPEG-7 (BOBER ET AL., 1999; JEANNIN, 2001). La principale différence entre les moments de Zernike et la transformation ART porte sur leurs fonctions de base qui sont utilisées pour extraire les coefficients de l'image qui décrivent la forme. Conceptuellement, le descripteur de formes par approche régions travaille en décomposant la forme sur plusieurs fonctions de base 2D orthogonales, définies par la transformation ART. Les coefficients ainsi obtenus sont utilisés après normalisation et quantification pour décrire la forme.

1.2.3.1 Définition de la transformation ART

La transformation ART, Angular Radial Transform, proposée par KIM et KIM (1999), est une transformation 2D complexe, définie sur un disque unité, en coordonnées polaires avec la formule suivante :

$$F_{nm} = \int_0^{2\pi} \int_0^1 V_{nm}(\rho, \theta) f(\rho, \theta) \rho d\rho d\theta \quad (1.7)$$

Dans cette formule, F_{nm} est un coefficient d'ART d'ordre n et m , $f : (\rho, \theta) \rightarrow \{0, 1\}$ est une fonction image binaire définie en coordonnées polaires, et $V_{nm}(\rho, \theta)$ est une fonction de base d'ART qui est séparable dans la direction angulaire et radiale :

$$V_{nm}(\rho, \theta) = A_m(\theta)R_n(\rho) \quad \text{où} \quad \begin{cases} A_m(\theta) = \frac{1}{2\pi} e^{jm\theta} \\ R_n(\rho) = \begin{cases} 1 & n = 0 \\ 2 \cos(\pi n \rho) & n \neq 0 \end{cases} \end{cases} \quad (1.8)$$

Les indices n et m sont des indices respectivement radiaux et angulaires qui définissent l'ordre du coefficient F_{nm} .

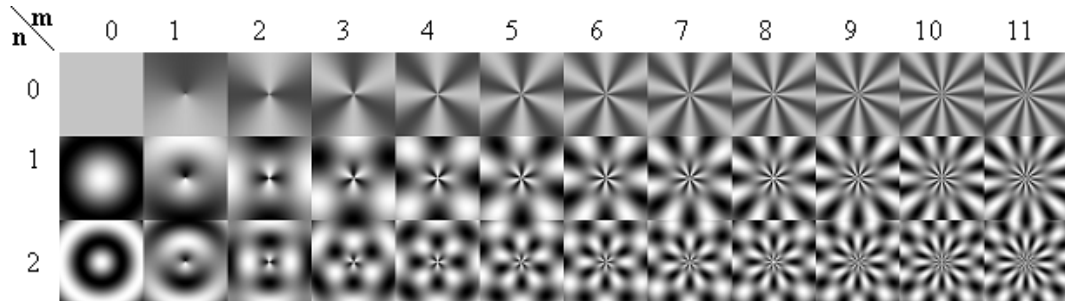


FIG. 1.5 – Parties réelles des fonctions de base d'ART.

La fonction de base d'ART V_{nm} est une fonction complexe. Sur la figure 1.5, la partie réelle de chaque fonction de base est présentée sous forme d'images. Les

parties imaginaires sont identiques aux parties réelles, avec une différence de phase. La transformation peut être considérée comme la projection d'une image sur chaque fonction de base. Ainsi, pour une image donnée, le coefficient F_{nm} est d'autant plus grand que l'image est semblable à la fonction de base d'ordre n et m .

Pour chaque image, il y aura $n * m$ coefficients ART. Le problème est de savoir combien sont nécessaires pour la description d'une image. Les coefficients sont ensuite utilisés pour mettre en correspondance deux images. Pour déterminer le nombre de coefficients, il faut analyser les covariances entre les fonctions de base. Pour des descripteurs orthogonaux, il n'y a aucune covariance entre les fonctions de base. Cependant, on montre que la transformation ART n'est pas orthogonale et il y a une forte redondance entre les fonctions de base de même valeur m . Ainsi, il est raisonnable de maintenir le nombre des valeurs de n plus petit que celle de m . Le comité MPEG-7 (JEANNIN, 2001) propose de fixer n à 3 et m à 12. Ce sont les valeurs que nous utilisons dans notre étude.

1.2.3.2 Indexation

La procédure d'indexation est composée de quatre étapes successives : le pré-traitement de l'image, le calcul des coefficients, leur normalisation et leur quantification, comme le montre la figure 1.6.

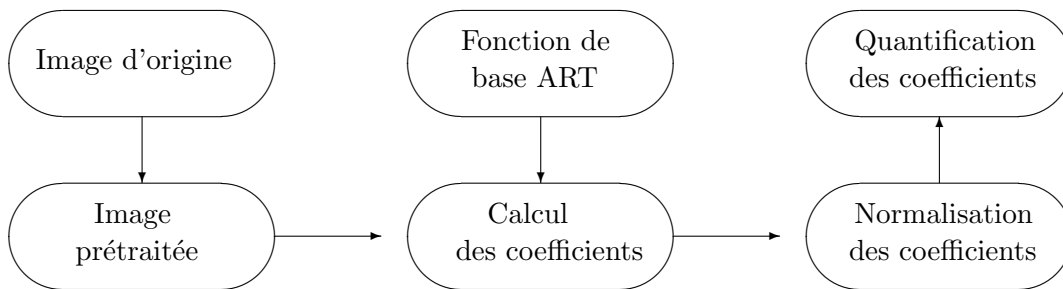


FIG. 1.6 – Schéma de la procédure d'indexation.

Pré-traitements Cette étape se compose de deux parties : l'extraction de la région d'intérêt et l'interpolation. Tandis que la première partie élimine le problème du décalage dans l'image, la deuxième partie standardise la taille de l'image à transformer, dans le but de simplifier la mise en correspondance avec les fonctions de base.

Pour extraire la région d'intérêt, le centre de gravité de l'objet et la dimension de l'image qui contient l'objet sont calculés. La dimension vaut deux fois la distance du centre de gravité au point de l'objet le plus éloigné. Au départ, l'image peut avoir n'importe quelle dimension, mais pour pouvoir comparer les images entre elles, elles

doivent être à la même échelle. Toutes les images sont centrées, recadrées et mises à l'échelle à la dimension choisie par le comité MPEG-7 : $101 * 101$. La figure 1.7 présente le processus d'extraction.

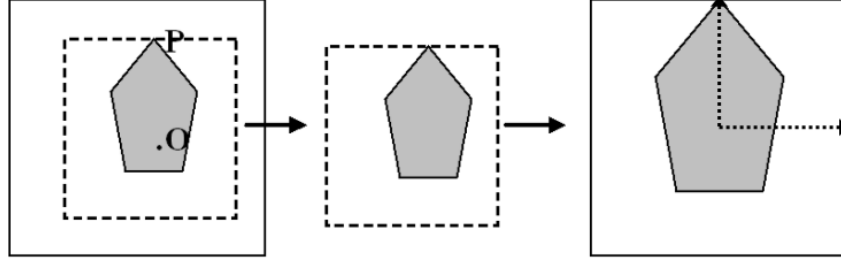


FIG. 1.7 – Processus d'extraction de régions : extraction du carré englobant centré sur le centre de gravité de l'objet, mise à l'échelle avec les fonction de base sur l'espace $[-1; 1]^2$.

Calcul des coefficients Dans cette étape, la transformation ART est appliquée sur image prétraitée. Pour chaque valeur de n et m , le coefficient correspondant d'ART F_{nm} est calculé. Comme cité précédemment, nous gardons $n = 3$ et $m = 12$. Ainsi, pour chaque image, il y aura $3 * 12 = 36$ coefficients de l'ART. Dans la définition formelle, F_{nm} est calculé par :

$$F_{nm} = \int_0^{2\pi} \int_0^1 V_{nm}(\rho, \theta) f(\rho, \theta) \rho d\rho d\theta \quad (1.9)$$

où F_{nm} est un coefficient d'ART de l'ordre n et m , $f(\rho, \theta)$ est une fonction image en coordonnées polaires et V_{nm} est la fonction de base d'ART. Pour la simplicité du calcul, l'équation est reformulée en coordonnées cartésiennes par :

$$F_{nm} = \int_x \int_y V_{nm}(x, y) f(x, y) dx dy \quad (1.10)$$

Dans cette équation, au lieu de convertir l'image f en coordonnées polaires, les fonctions de base d'ART V_{nm} sont converties en coordonnées cartésiennes. Il convient de noter que les coefficients d'ART calculés sont complexes.

Normalisation des coefficients Comme les coefficients de F_{nm} sont complexes, on se sert des modules des coefficients $\|F_{nm}\|$ qui sont normalisés par $\|F_{00}\|$. Ceci est appelé la normalisation de surface, puisque le premier coefficient F_{00} est égal à la somme des pixels non nuls dans l'image. Nous disposons ainsi de $n * m - 1$ coefficients, puisque F_{00} est utilisé pour la normalisation.

Intervalle de quantification	ArtDE	Valeur de quantification inverse
[0.000000000, 0.003585473[0000	0.001763817
[0.003585473, 0.007418411[0001	0.005468893
[0.007418411, 0.011535520[0010	0.009438835
[0.011535520, 0.015982337[0011	0.013714449
[0.015982337, 0.020816302[0100	0.018346760
[0.020816302, 0.026111312[0101	0.023400748
[0.026111312, 0.031964674[0110	0.028960940
[0.031964674, 0.038508176[0111	0.035140141
[0.038508176, 0.045926586[1000	0.042093649
[0.045926586, 0.054490513[1001	0.050043696
[0.054490513, 0.064619488[1010	0.059324478
[0.064619488, 0.077016351[1011	0.070472849
[0.077016351, 0.092998687[1100	0.084434761
[0.092998687, 0.115524524[1101	0.103127662
[0.115524524, 0.154032694[1110	0.131506859
[0.154032694, $+\infty$ [1111	0.192540857

TAB. 1.1 – Tableau des intervalles de quantification et des valeurs de la quantification inverse.

Quantification des coefficients Les coefficients normés d'ART sont quantifiés et codés sur 4 bits. La table de quantification peut être construite par une distribution statistique des coefficients. Le tableau 1.1 (JEANNIN, 2001) est proposé par le comité MPEG-7. Il est basé sur une distribution exponentielle avec le paramètre $\lambda = 18$. La table non uniforme de quantification est produite par l'équation suivante :

$$\int_{Q_i}^{Q_{i+1}} \lambda e^{-\lambda x} = \frac{1}{16} \quad (1.11)$$

où $k \in [0, 16[$ et $[Q_i, Q_{i+1}[$ est l'intervalle de quantification.

Mise en correspondance La similarité entre deux images est mesurée par la distance L_1 correspondant à la somme des différences entre deux ensembles des coefficients Art quantifiés inverses, en utilisant le tableau de quantification inverse (Tableau 1.1). La distance entre les images Q et I est mesurée par :

$$Distance(Q, I) = \frac{1}{35} \sum_{i=0}^{n.m-1} \|Quant^{-1}(ART_Q(i)) - Quant^{-1}(ART_I(i))\| \quad (1.12)$$

D'autres types de mesure de distance que L_1 peuvent également être envisagés.

1.2.3.3 Propriétés de la transformation ART

Les méthodes d'indexation et de recherche doivent posséder les propriétés d'invariance aux transformations subies par un objet à l'intérieur d'une image. Ces transformations peuvent être de plusieurs types : rotations planes, variations d'échelle, occultations, bruits, décalages, rotations non planes et phénomènes de perspective. Pour qu'une méthode de classification d'objets soit efficace, il faut qu'elle soit invariante à ces transformations. La figure 1.8 montre une image déformée suivant les principales transformations.

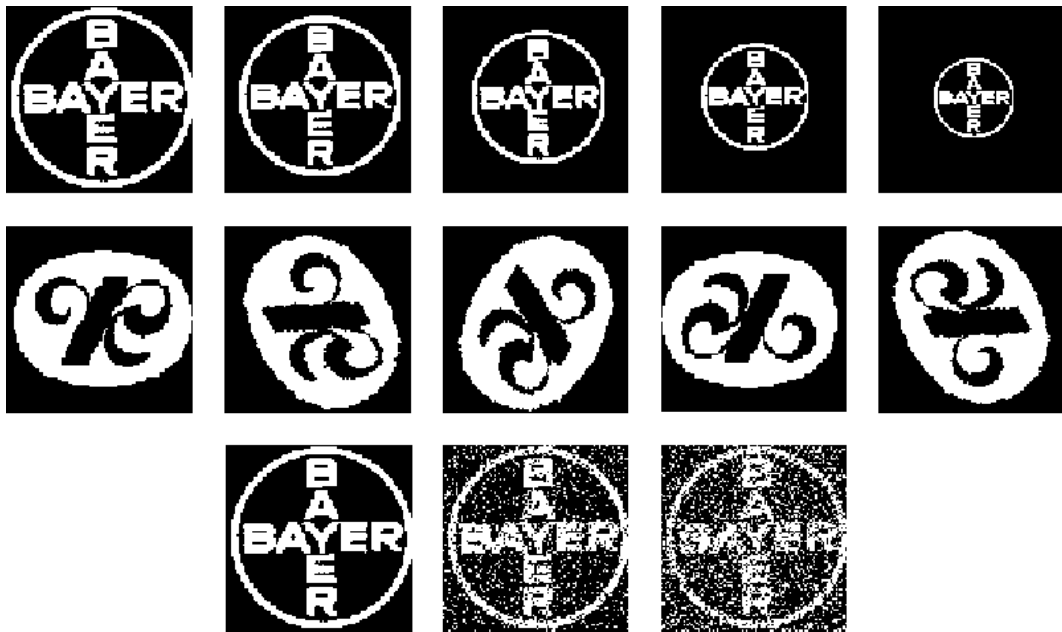


FIG. 1.8 – Exemple d'images : à différentes échelles, à différentes orientations et à différents niveaux de bruits poivre et sel : 0, 35 et 50%.

Robustesse aux changements d'échelle et aux translations. Les variations d'échelle sont prises en compte par le pré-traitement de la transformation ART, qui centre et met à l'échelle l'image à traiter avec les fonctions de base. Il en est de même pour les problèmes de décalage. Un objet, à des positions différentes dans le plan image, fournira les mêmes descripteurs. Si l'extraction de l'objet n'est pas nette et sélectionne les zones voisines, le centrage et la mise à l'échelle ne se feront pas de la même manière, ce qui produira des coefficients ART différents.

Des tests ont permis de quantifier la robustesse aux petits décalages. Un décalage de trois pixels vers le bas produit une distance entre descripteur de 0.16, ce qui correspond à l'ajout de 65% de bruit. Ceci montre la haute sensibilité de la transformation ART au décalage. Pour éliminer ce problème, tous les objets devront être

centrés, comme décrit dans l'étape de pré-traitement.

Invariance aux rotations. Dans une scène, un objet peut être présent sous différentes orientations. Les descripteurs doivent pouvoir classer comme similaire, un objet quelle que soit son orientation. La transformation ART est invariante aux rotations d'un objet dans le plan image. Les fonctions de base sont périodiques dans leur direction angulaire, ce qui rend leurs normes invariantes aux rotations planes. Deux images d'un même objet ayant subi une rotation plane, fourniront deux séries de coefficients ART de même norme, comme le montre les équations suivantes :

$$\begin{aligned}
 F_{nm}^{\alpha} &= \int_0^{2\pi} \int_0^1 V_{nm}(\rho, \theta_0) f^{\alpha}(\rho, \theta_0) \rho d\rho d\theta_0 \\
 &= \int_0^{2\pi} \int_0^1 \frac{1}{\pi} e^{jm\theta_0} \cos(\pi n\rho) f(\rho, \theta_0 + \alpha) \rho d\rho d\theta_0 \\
 &= \int_0^{2\pi} \int_0^1 \frac{1}{\pi} e^{jm(\theta-\alpha)} \cos(\pi n\rho) f(\rho, \theta) \rho d\rho d\theta \\
 &= \int_0^{2\pi} \int_0^1 \frac{1}{\pi} e^{-jm\alpha} e^{jm\theta} \cos(\pi n\rho) f(\rho, \theta) \rho d\rho d\theta \\
 &= e^{-jm\alpha} F_{nm} \Rightarrow \|F_{nm}^{\alpha}\| = \|F_{nm}\|
 \end{aligned}$$

où $f^{\alpha}(\rho, \theta_0)$ est l'image tournée par une rotation d'angle α . Nous obtenons l'égalité : $F_{nm}^{\alpha} = e^{-jm\alpha} F_{nm}$. Ainsi, pour une image modifiée par une rotation plane, les coefficients sont liés à ceux de l'image d'origine et leur norme est identique, ce qui rend les coefficients ART normalisés égaux.

Les rotations non planes, qui produisent un changement de profil de l'objet et une déformation due à la perspective, ne sont pas prises en compte. Ces transformations seront traitées dans la partie suivante par la généralisation que nous proposons pour la transformation ART.

Robustesse à la présence de bruit. L'extraction de l'objet image et la nature du document multimédia entraînent l'apparition de bruit et de petits décalages dans la représentation d'un même objet. La transformation ART est invariante à la présence de bruit. Des tests ont été faits en intégrant à une image binaire un bruit « poivre et sel ». Les descripteurs peuvent tolérer jusqu'à 35% de bruit sans perturber sensiblement les coefficients. La robustesse au bruit implique les faibles variations des coefficients lors de faibles occultations ou de petites perturbations. Si les perturbations sont de trop grande amplitude, le centre de gravité et le contour de la forme seront modifiés et le pré-traitement ne pourra se faire de manière identique. Ce qui se traduira par l'apparition de décalages entre les images à classer et de grandes variations de leurs coefficients.

1.2.4 Conclusion

Comme nous venons de le voir, la transformation ART est reconnue pour être une méthode efficace et robuste d'analyse de formes basée sur les moments. En particulier, de nombreux tests, menés au sein du processus de standardisation MPEG-7, ont montré la meilleure efficacité de la transformation ART par rapport aux autres méthodes d'indexation basées régions.

Dans la suite de ce travail, nous étendrons la transformation ART, pour la rendre applicable sur des images naturelles et pour dépasser les limitations dues au fait que la transformation ART ne prenne pas en compte les composantes couleur des images et ne soit pas robuste aux déformations homographiques (chapitre 3). D'autre part, nous étendrons ART pour la rendre applicable aux objets 3D, tout en gardant un faible coût d'indexation et de recherche et ses propriétés de robustesse aux principales déformations (chapitre 5). De plus, nous utiliserons la transformation ART pour caractériser les vues des objets 3D, pour faire de l'indexation d'objets 3D par vues ou de la recherche d'objets 3D à partir de vues 2D (chapitre 4 et 5.3).

1.3 Indexation d'objets 3D par requêtes images

Dans cette section, nous présenterons l'état de l'art en indexation d'objets 3D par requête 2D. Comme nous l'avons énoncée en introduction, une réelle demande émane des industriels qui manipulent des objets 3D, quant à la mise en place d'outils permettant de rechercher et de classer des bases d'objets 3D. Pour répondre à ces problématiques, plusieurs approches d'indexation 3D ont été proposées pour permettre de rechercher des objets 3D. Deux classes de méthodes se caractérisent par la nature des données qu'elles traitent. Une première classe de techniques considère que l'on peut rechercher des objets 3D par rapport à un autre objet 3D. Ce sont les méthodes d'indexation 3D/3D, que nous traiterons dans la chapitre 4.5. D'autre part, une autre série de méthodes caractérisent les objets 3D en fonction des vues pouvant être créées à partir de l'objet. Ceci permet de rechercher des objets 3D à partir de vues 2D. Les applications visées par les méthodes 2D/3D sont la recherche de modèles 3D à partir d'une photo ou du croquis d'une pièce. L'objectif est de rechercher le ou les modèles tridimensionnels les plus proches à partir d'une image 2D.

Le processus de comparaison d'objets 3D par vues peut se décomposer en deux phases : une phase d'indexation et une phase de recherche. Dans la phase d'indexation, pour tout modèle tridimensionnel de la base, on calcule les vues caractéristiques et leurs index associés. Lors de la phase de recherche, l'image requête subit un traitement similaire aux vues de la base, à l'issue duquel un descripteur (invariant à certaines déformations) est calculé et comparé avec les descripteurs de la base. La figure 1.9 montre un exemple d'une pièce caractérisée par une ensemble de vues 2D

et les deux types de recherches réalisables, 2D/3D en comparant l'ensemble des vues descripteurs à une image requête et 3D/3D en comparant les deux ensembles de vues.

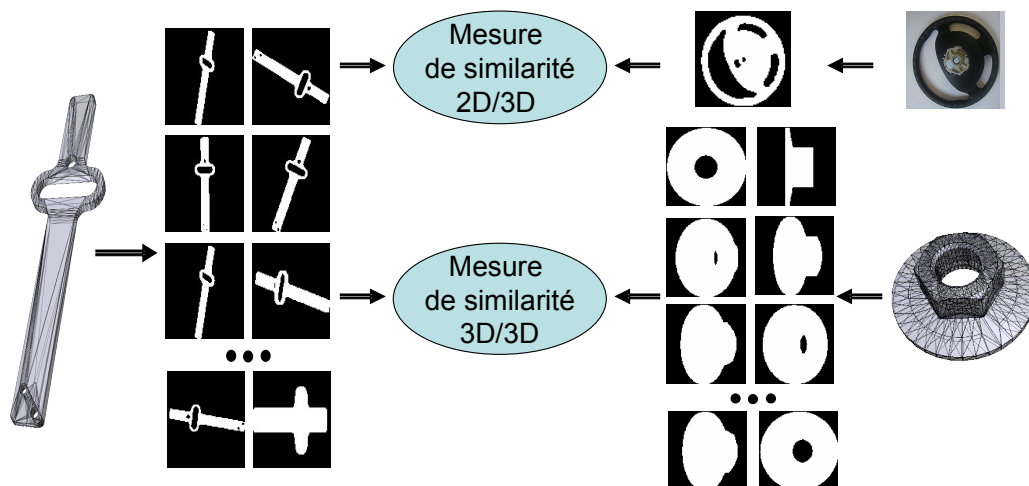


FIG. 1.9 – Exemple de mise en correspondance 2D/3D et 3D/3D entre un ensemble de vues et une image requête et un autre ensemble de vues.

L'idée générale de ce type de méthodes est de représenter un objet 3D par un ensemble de vues caractéristiques. Pour cela, l'espace des vues de l'objet est discrétisé par un ensemble de N points de vues répartis autour du modèle et pour chaque point, une image du profil de la pièce 3D est prise. La figure 1.10 montre les trois vues créées à partir d'un objet 3D en plaçant les points de vues sur les trois axes principaux.

Chaque image est indexée par une méthode d'analyse de formes 2D et lors de la recherche, la méthode indexe la vue 2D requête et la compare avec les N vues de l'objet et retourne la vue la plus proche, pour dire si la vue vient de ce modèle 3D. Les deux points longuement discutés dans la littérature sont le choix de la méthode de description des vues 2D et le choix des vues et leur nombres à prendre en compte. Le choix de la méthode de description des vues 2D ne modifiera pas le principe général de la méthode mais peut en modifier les performances et l'efficacité.

L'espace des vues est la sphère entourant l'objet sur lequel on définit les points de vues à partir desquelles sont créées les images de l'objet. Un autre paramètre modifiant les vues extraites est le coefficient perspective. Pour reconnaître des formes naturelles, il faut échantillonner l'espace des vues en fonction de la position de la caméra et de la perspective, pour prendre en compte ces déformations. Généralement, les méthodes d'indexation 2D/3D fixent l'indice de focale à une valeur moyenne, ce qui permet de négliger ce paramètre, ou utilisent un coefficient de perspective infini, ce qui revient à faire une projection orthographique et permet de réduire l'espace des vues en ne considérant que la demi-sphère des vues.

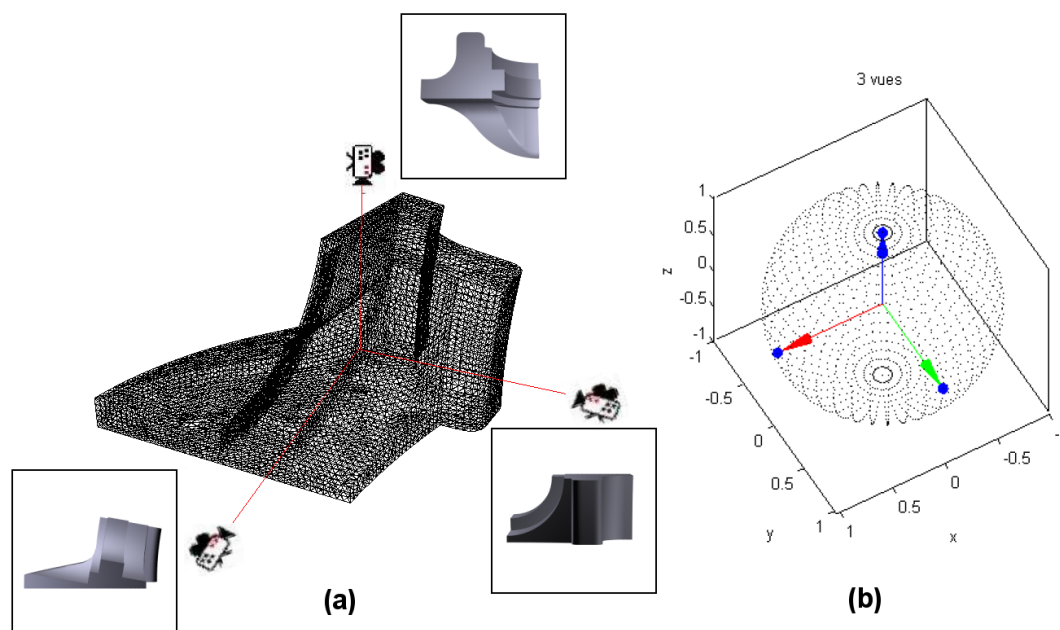


FIG. 1.10 – Vues d'un objet 3D créées à partir des trois vues principales (a), sphère des vues et position des trois points de vues principaux (b).

1.3.1 Choix de la méthode de description

Plusieurs méthodes d'indexation d'objets 3D par vues varient en fonction de la méthode de description des vues choisies. De façon globale, tous les descripteurs de forme 2D peuvent être utilisés pour caractériser les vues extraites des objets 3D. Il faut cependant que la méthode de description choisie possède des propriétés d'invariance aux principales transformations subies par les formes 2D lors de la création des images. Il faut en particulier, qu'elle soit invariante en rotation pour ne pas avoir à prendre en compte les rotations autour de l'axe de prise de vues. On peut en particulier citer, MOKHTARIAN et ABBASI (2000) qui a utilisé les Curvature Scale Space (CSS, MOKHTARIAN et MACKWORTH (1992)), CHEN et OUHYOUNG (2002) qui ont utilisé l'Angular Radial Transform (ART, KIM et KIM (1999)), (KHOTANZAD et HONG, 1990) qui ont utilisé les moments de Zernike, etc.

MOKHTARIAN et ABBASI (2000) propose l'utilisation du Curvature Scale Space (CSS) ou l'espace d'échelle de contour (EEC) pour l'indexation des vues d'objets tridimensionnels (MOKHTARIAN et MACKWORTH, 1992). L'EEC est un invariant par rapport au groupe de similitudes. DAUDI et MATUSIAK (2000) montrent qu'il est stable par rapport aux erreurs dues aux imprécisions et au bruit, et est complet. MAHMOUDI et DAUDI (2002) proposent d'indexer les modèles 3D par les vues. Les EEC sont organisés sous forme d'une structure d'arbre dite M-Tree. Cette étape est importante pour l'accélération de la recherche et la gestion de la mémoire considérant

le nombre important de vues associées aux modèles tridimensionnels. Les résultats obtenus sur une base de cent modèles montrent l'intérêt de cette approche.

CHEN et STOCKMAN (1998) et YI et CHELBERG (1998) proposent une méthode basée sur l'approche bayésienne. Formellement cela revient à calculer la probabilité a posteriori de reconnaître le modèle M_i quand les descripteurs M_k sont observés, c'est-à-dire $P(M_i/M_k)$. La méthode utilisée donne des résultats très satisfaisants. La modularité des LSG (Local Surface Group) utilisés comme descripteurs permet une intégration d'autres propriétés comme la couleur ou les textures. L'algorithme est facile à implémenter. Cependant, la base de données utilisée pour les tests est constituée uniquement de vingt modèles.

L'approche développée par CHEN et STOCKMAN (1998) se décompose en quatre étapes : la modélisation de l'objet par un ensemble de cinq vues, l'extraction des descripteurs, l'indexation et la phase d'alignement pour vérifier l'hypothèse du modèle trouvé lors de la phase d'indexation. L'étape d'indexation est basée une approche bayésienne. Cette méthode, testée sur une base de six cents modèles tridimensionnels, nécessite une phase de segmentation.

Dans le cas des objets polyédriques, COSTA et SHAPIRO (2000) proposent la détection des contours des vues par le filtre de Canny. Les segments de lignes et les segments d'arcs circulaires sont extraits à partir de l'image contour. Ces primitives sont utilisées pour générer des propriétés de haut niveau comme les paires de lignes parallèles, les jonctions, les triplets et les ellipses. Différentes relations sont ensuite calculées pour définir un graphe. Chaque vue est décrite par un graphe relationnel de toutes ses propriétés.

1.3.2 Choix du nombre de vues

Un point qui va modifier l'efficacité et la rapidité de ces méthodes est le nombre et la position de vues à prendre en compte. Ce nombre est directement lié aux performances du système. Il faut le diminuer au maximum pour avoir une méthode rapide, tout en conservant la capacité descriptive des vues. Deux approches ont été exploitées pour diminuer le nombre de vues, soit en ne prenant au départ qu'un nombre restreint de vues, ce sont les méthodes à sélection de vues fixes, soit en prenant un grand nombre de vues et en les regroupant pour ne garder que les plus caractéristiques, ce sont les méthodes à sélection de vues variables.

1.3.2.1 Méthodes de sélection de vues fixes

Pour garantir un faible nombre de vues et par la même une recherche rapide, une classe de techniques ne crée qu'un nombre restreint de vues à des positions fixes, indépendamment de l'objet à indexer. Ces méthodes fixent sur la sphère des vues entourant l'objet un certain nombre de points et créent des images à partir de ces

points de vues. Pour augmenter la similarité entre vues, les objets 3D sont alignés et mis à l'échelle par rapport à la sphère des vues. Pour cela, les axes principaux des objets sont calculés par une analyse en composante principale (ACP) ou par le calcul de la matrice des covariances. Ceci permet d'aligner les axes principaux des objets avec ceux du repère de l'espace des vues. La figure 1.11 montre quatre exemples de répartition de vues sur la sphère des vues.

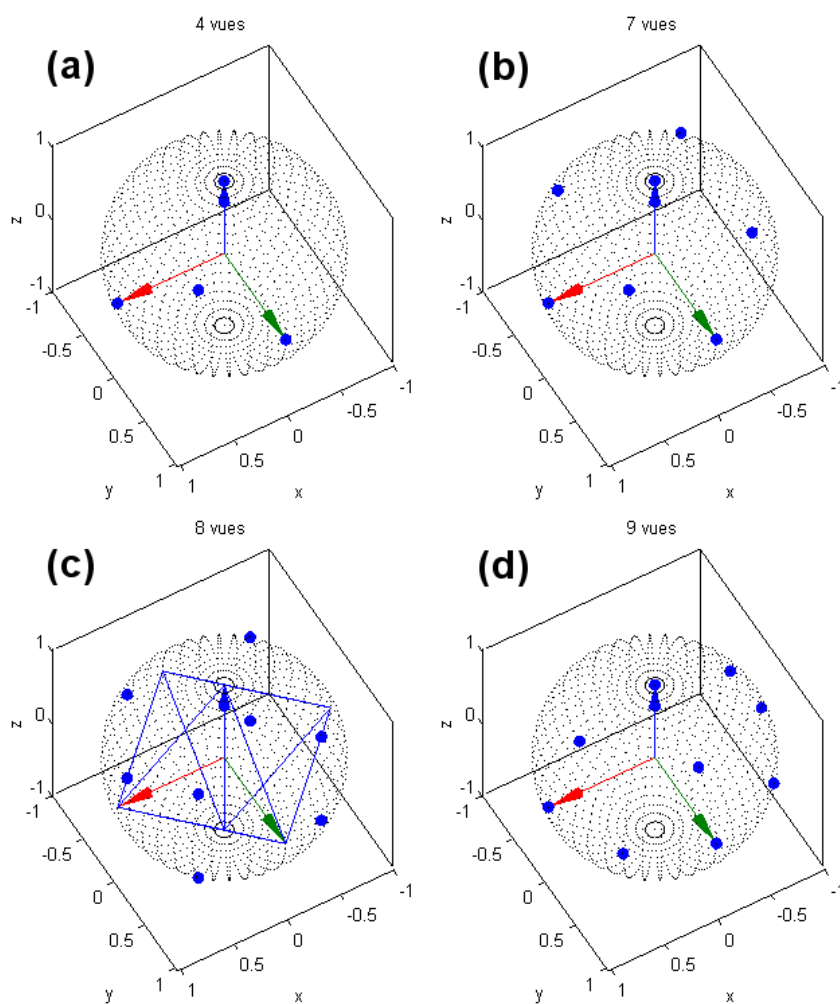


FIG. 1.11 – Exemple de répartition de vues fixes sur la sphère des vues d'un objet : 4 vues, 7 vues, 8 vues et 9 vues.

Les points de vues sont choisis de façon à être répartis sur la sphère des vues. On peut citer les travaux de [ABBASI et MOKHTARIAN \(2001\)](#), qui fixent le nombre de vues associées à chaque objet tridimensionnel à neuf vues : quatre vues, correspondant à la vue de haut, de côté, de face et une vue arrière, et cinq vues intermédiaires (figure 1.11.d).

CHEN et STOCKMAN (1998) proposent une facettisation de la sphère englobante pour chaque objet. Cette sphère est facétisée en huit faces triangulaires. La caméra est placée au barycentre de chaque face pour obtenir les vues (figure 1.11.c).

NAYAR ET AL. (1996), ainsi que MAHMOUDI et DAOUDI (2002), considèrent le modèle tridimensionnel comme un nuage de points. Les axes principaux du modèle sont calculés grâce aux valeurs propres de la matrice de co-variance de son nuage de points. Le nuage de points tridimensionnels est projeté en 2D suivant les trois directions principales ainsi calculées. Sept vues caractéristiques sont créées, les trois vues principales le long des axes principaux et quatre vues suivant les quatre directions correspondant à une vue de 45° entre la vue primaire et la vue secondaire, une vue de 45° entre la vue primaire et la vue tertiaire, une vue de 45° entre la vue secondaire et la vue tertiaire et une dernière vue de 45° entre les trois vues précédentes (figure 1.11.b).

1.3.2.2 Méthodes de sélection de vues variables

Pour réduire le nombre de vues descripteurs en décrivant convenablement l'espace des vues de l'objet, il faut mettre en place une méthode de sélection de vues variables. Ces méthodes discrétisent uniformément l'espace des vues et regroupent les points de vues en fonction de la similarité entre les vues créées, pour ne garder qu'un faible nombre de vues caractéristiques. Ceci permet d'avoir un nombre de vues qui s'adapte à la complexité de l'objet à décrire. Pour que les points de vues soient uniformément répartis sur la sphère des vues, il faut discrétiser l'espace par une méthode de Sphere Picking, proposée par MARSAGLIA (1972), selon les formules :

$$\begin{cases} x = 2x_1\sqrt{1-x_1^2-x_2^2} \\ y = 2x_2\sqrt{1-x_1^2-x_2^2} \\ z = 1-2(x_1^2+x_2^2) \end{cases} \quad (1.13)$$

où x_1 et x_2 sont deux distributions uniformes indépendantes sur $[-1, 1]$, tel que $x_1^2 + x_2^2 < 1$. Le nombre de points de vues est dépendant de la taille des distributions uniformes x_1 et x_2 . Ceci permet de positionner aléatoirement des points sur une sphère unité, où de répartir uniformément des points sur une sphère, comme le montre la figure ???. Pour obtenir une répartition uniforme des points sur une demi sphère, il suffit de prendre une répartition uniforme des paramètres x_1 et x_2 et de ne garder que les points tel que $x_1^2 + x_2^2 \leq 0.5$.

Le principe de réduction des vues caractéristiques a été souvent utilisé, et on peut citer les travaux de BOWYER et DYER (1990) et de CYR et KIMIA (2004). La figure 1.13 montre un exemple de groupement de vues semblables.

Plus récemment, FILALI ANSARY ET AL. (2004, 2005) ont proposé une méthode de réduction de vues basée sur une sélection des vues par une approches probabiliste. Un grand nombre de vues est créé à partir d'un objet 3D et les auteurs utilisent

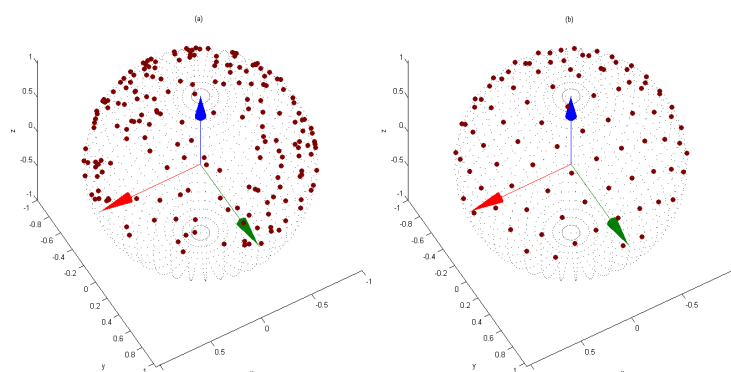


FIG. 1.12 – Répartition aléatoire de points sur une sphère par la méthode de Sphere Picking (a) et répartition uniforme de points sur une demi-sphère par échantillonnage régulier des paramètres x_1 et x_2 sous la condition que $x_1^2 + x_2^2 \leq 0.5$ (b).

un algorithme de classification adaptative inspiré de l'algorithme X-means (PELLEG et MOORE, 2000) pour extraire les représentants des classes. Le critère d'arrêt de la subdivision est basé sur un critère Bayésien, Bayesian Information Criteria (BIC). Ce critère permet de caractériser la subdivision et de définir si une itération supplémentaire fournit une meilleure représentation de l'espace des vues.

1.4 Conclusion

Nous venons de voir le principe général des méthodes d'indexation 2D et 2D/3D. Les méthodes 2D/3D varient en fonction de deux paramètres, le nombre de vues et le descripteur de forme 2D. Ces méthodes pré-calculent l'intégralité des descripteurs des vues caractéristiques sélectionnées pour décrire l'objet et comparent l'image requête à ces descripteurs. Ceci ne permet pas de s'adapter à la requête. Dans le cas où elle se situe loin d'une vue caractéristique, les résultats se dégradent. Nous présenterons dans le chapitre 4, les travaux que nous avons réalisés pour mettre en place une méthode de comparaison directe, qui compare la requête à une représentation compacte de l'ensemble des vues d'un objet 3D.

Une discussion générale sur les différentes méthodes d'indexation, tant 2D que 3D, est menée dans la conclusion de cette partie dans le paragraphe 2.8

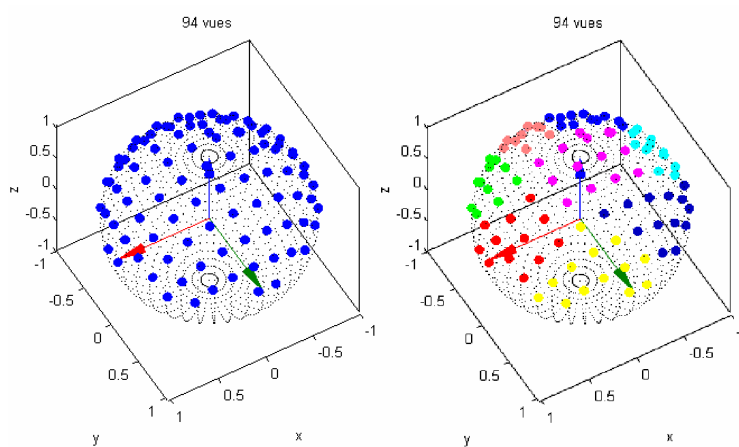


FIG. 1.13 – Exemple de répartition de vues par une méthode de sélection de vues variables. Un grand nombre de vues est créé uniformément répartis sur la Sphere des vues par une méthode de Sphere Picking, puis ne sont gardées que les plus caractéristiques.

2

Etat de l'art en indexation 3D

Sommaire

2.1	Introduction	33
2.2	Approches statistiques	33
2.2.1	Approches locales	33
2.2.2	Approches globales	34
2.3	Approches structurelles	35
2.3.1	Approches CAO	35
2.3.1.1	<i>Model signature graphs</i> MSG	36
2.3.1.2	Approche par décomposition volumique	37
2.3.2	Approches basées sur les squelettes	37
2.3.2.1	Algorithmes d'amincissement	38
2.3.2.2	Squelette à partir de la carte de distance	39
2.3.2.3	Simulation de la propagation de feu de prairie	39
2.3.2.4	Squelette à partir du diagramme de Voronoï	40
2.3.2.5	Exosquelette	40
2.3.2.6	Graphe de Reeb multi résolution	41
2.3.3	Approche volumique	43
2.3.4	Représentation symbolique	45
2.4	Approches par transformées	46
2.4.1	Transformée de Hough	46
2.4.2	Approches basées sur les moments	47
2.4.2.1	Les moments géométriques	48
2.4.2.2	Descripteur de Fourier 3D	48
2.4.2.3	Les moments de Zernike 3D	49

2.4.2.4	Les harmoniques sphériques	49
2.5	Approches par vues	50
2.6	Pré-traitement	53
2.6.1	Centrage	54
2.6.2	Mise à l'échelle	54
2.6.3	Alignement	54
	Remarque :	55
2.7	Critères d'évaluations	56
2.7.1	Base de données	56
	2.7.1.1 Base Princeton Shape Benchmark	56
	2.7.1.2 Base Renault Semantic-3D	56
	2.7.1.3 Taxinomie	56
2.7.2	Critères de comparaison	57
	2.7.2.1 Critères algorithmiques	57
	2.7.2.2 Critères évaluant la pertinence des requêtes par similarités	58
2.8	Conclusion	60

2.1 Introduction

Ce chapitre dresse un état de l'art en indexation d'objets tridimensionnels. Plusieurs grandes voies se distinguent dans les différentes approches proposées, et nous avons fait le choix de les séparer en quatre groupes :

- **les approches statistiques**, qui caractérisent les modèles tridimensionnels à indexer par une ou plusieurs distributions de descripteurs mathématiques de formes ;
- **les approches structurelles**, qui décrivent les objets 3D en s'appuyant sur des informations de haut niveau sur la structure de la scène ;
- **les approches par transformées** qui visent à déterminer des représentations de forme globales, définies en terme de transformation intégrale ;
- **les approches multi-vues** qui caractérisent les objets 3D par un ensemble de vues représentant des projections du modèle 3D.

Cet état de l'art en indexation d'objets 3D a été réalisé dans le cadre du projet RNRT Semantic-3D. Les principales méthodes en indexation 3D, utiles pour la compréhension des travaux de ce document, sont expliquées dans cette partie. Le lecteur pourra se reporter au livrable 2.1 du projet RNRT Semantic-3D (BASKURT ET AL., 2003) pour de plus amples explications sur d'autres méthodes d'indexation d'objets 3D.

2.2 Approches statistiques

Les approches statistiques consistent à caractériser les objets 3D par une ou plusieurs distributions de descripteurs mathématiques de formes. Ces distributions sont généralement représentées sous la forme d'histogrammes. La comparaison des objets se fait en calculant une distance entre histogrammes, ce qui fournit un classement des modèles dans l'ordre de ressemblance trouvé par le moteur de recherche. On peut donc dire que des modèles 3D sont considérés comme géométriquement (ou visuellement) proches si leurs distributions statistiques le sont aussi.

De nombreux descripteurs de formes se basant sur ces approches ont été proposés et il est possible de les classer selon deux grandes catégories : les descripteurs locaux, qui utilisent les caractéristiques locales de la forme, et les descripteurs globaux, qui caractérisent les modèles 3D dans leur globalité.

2.2.1 Approches locales

Les descripteurs locaux s'appuient sur les caractéristiques locales des modèles à indexer et utilisent principalement la courbure calculée en tout point de l'objet 3D. L'idée générale de ce type de méthode est de calculer la courbure en tous points d'un l'objet et de comparer ces mesures aux distributions des autres objets.

On peut citer les travaux de ZAHARIA et PRÊTEUX (2002a), qui ont défini un descripteur de formes qui s'appuie sur une distribution des courbures principales telles qu'elles ont été introduites par KOENDERINK et VAN DOORN (1992). L'étape d'estimation de la courbure utilise une approximation de la surface par une quadrique. Cette solution n'est cependant valable qu'à condition que les maillages respectent certaines hypothèses d'orientation et de régularité. La distribution de courbures ainsi formée est invariante aux transformations géométriques, mais n'est pas robuste face aux représentations topologiques multiples (niveau de facettisation par exemple).

ASSFALG ET AL. (2003a,b) développent l'idée que la forme d'un objet 3D peut être décrite par une carte des courbures de sa surface. Après une phase de pré-traitement pendant laquelle les propriétés différentielles de la surface de l'objet 3D sont calculées, la surface du modèle 3D est alors déformée pour épouser la surface d'une ellipsoïde, tout en gardant les informations lors de la déformation. La comparaison est effectuée en comparant les cartes de courbures comme des images 2D classiques. Les résultats obtenus par cette approche sont meilleurs qu'une méthode qui utilise les histogrammes de courbures calculés directement sur les objets 3D.

2.2.2 Approches globales

Les descripteurs globaux s'appuient sur une fonction de forme qui mesure les propriétés géométriques de l'objet dans son intégralité en calculant par exemple des distances entre des points sur la surface de l'objet, ou des moments 3D.

OSADA ET AL. (2001) présentent le problème de l'indexation de modèles 3D maillés quelconques (même des modèles présentant des problèmes de topologie) en trois étapes : la sélection d'une fonction de forme discriminante, la construction efficace de la distribution de cette fonction pour chaque modèle de la base et le calcul d'une distance de similarité entre paires de distributions. Les auteurs expérimentent pour cela cinq fonctions de forme choisies pour leur simplicité de calcul :

- la mesure des angles entre trois points pris au hasard sur la surface du modèle 3D ;
- la mesure de la distance euclidienne entre un point fixe du modèle (le centre de gravité par exemple) et un point pris aléatoirement sur sa surface ;
- la mesure de la distance euclidienne entre deux points pris au hasard sur la surface du modèle ;
- la mesure de l'aire des triangles de trois points pris au hasard sur la surface du modèle ;
- la mesure du volume tétraédrique entre quatre points pris au hasard sur la surface du modèle.

Ces descripteurs sont robustes aux rotations et aux translations. Pour avoir une robustesse à la mise à l'échelle, les auteurs choisissent de normaliser les distributions. Des tests plus poussés sont effectués avec la troisième fonction, et montrent les bons résultats de l'approche choisie dans un but de pré-classification avant utilisation

d'une méthode de recherche plus fine mais également plus coûteuse.

De la même manière, l'objectif d' ADAN ET AL. (2001) est d'utiliser des invariants globaux pour simplifier et accélérer le processus de calcul des descripteurs. Ils proposent de scinder le processus de reconnaissance d'objets en deux étapes, la discrimination et l'identification. La première étape consiste à utiliser deux caractéristiques globales invariantes aux rotations, translations et mises à l'échelle : une valeur scalaire simple, qui donne une mesure indirecte de la surface de l'objet, et un ensemble de vecteurs correspondant à l'orientation cohérente de la surface du modèle. Ces deux caractéristiques permettent de réduire considérablement le nombre de réponses potentiellement correctes dans la base de données avant de passer à la deuxième étape d'identification. Cette étape consiste à effectuer des tests d'hypothèses de la distribution de courbures sur la base précédemment réduite pour identifier l'objet correct.

2.3 Approches structurelles

L'objectif de ces méthodes est de caractériser des objets 3D en s'appuyant sur des informations de hauts niveaux sur leurs structures. Au cours des dernières années les graphes ont été utilisés pour représenter mathématiquement la connaissance structurelle d'un objet. Ce type de représentation permet de représenter les variations et les différences structurelles entre objets et de les combiner avec des algorithmes de mise en correspondance de graphes. Durant la phase d'indexation, ces méthodes extraient des objets 3D une représentation structurelle de l'objet suivant différentes méthodes que nous détaillerons par la suite. La comparaison des modèles se faisant alors en comparant leurs graphes.

Dans un premier temps, nous présenterons les méthodes qui se sont intéressées principalement aux objets issus de la CAO et qui par leur nature très spéciale permet de définir des méthodes qui leur sont propres. Puis, nous nous intéresserons aux deux principales méthodes permettant d'écrire la forme d'un objet sous forme de graphe, qui sont les approches squelettiques et les approches volumiques qui représentent l'objet par un ensemble de primitives volumiques simples, principalement les Géons. Enfin, nous présenterons les méthodes essayant de retrouver directement l'appartenance des objets à une classe en se basant sur une représentation symbolique.

2.3.1 Approches CAO

Les objets CAO par leur nature particulière et en particulier par leurs propriétés topologiques, permettent de mettre en place des procédés d'identification qui leur sont propres.

2.3.1.1 Model signature graphs MSG

MCWHERTER ET AL. (2001) proposent d'utiliser le *Model Signature Graphs* (MSG), construit par la représentation des surfaces d'un objet. La même approche fut proposée par SUN ET AL. (1995) et par JOSHI et CHANG (1988) dans le graphe d'attribut adjacent ou *Attributed Adjacency Graph* AAG (LEE et SAITOU, 2001). Dans un graphe AAG, un noeud représente les faces (par exemple, cylindrique ou plan) et les arêtes sans direction représentent l'adjacence entre les faces (par exemple, les bords circulaires et droits). Le graphe contient également des arêtes dirigées qui représentent les dépendances dans les tolérances et les dimensions entre deux faces. La figure 2.1 montre un objet simple et sa représentation en graphe AAG. Par exemple, les noeuds, *PF* et *CF*, représentent les faces planaires et cylindriques, et les arêtes sans direction, *s+* et *c*, représentent les bords droits et les bords circulaires.

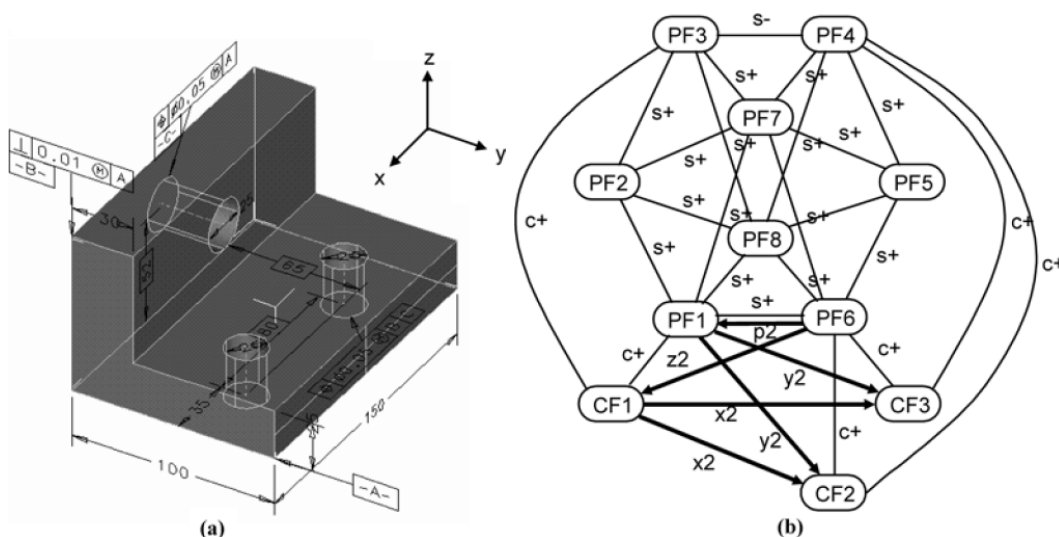


FIG. 2.1 – Modèle CAD et son graphe AAG correspondant (LEE et SAITOU, 2001).

Les MSG proposent de rajouter des informations sur les noeuds représentant les arêtes et les faces, comme : l'aire de la surface, les normales, la longueur de la courbe ou la concavité. La comparaison se fait à deux niveaux distincts. Dans un premier temps, les auteurs proposent de comparer un vecteur topologiquement invariant, Invariant Topology Vector (ITV), composé de 29 caractéristiques. Puis de comparer les matrices d'adjacences, par le calcul d'une distance de similarité entre leurs vecteurs propres.

Le point principal ne permettant pas d'appliquer cette méthode sur un objet quelconque est que les faces des objets doivent être des faces facilement identifiables (plan, cylindre, sphère,...). Ce qui ne permet pas d'appliquer ce type de méthode sur un objet naturel.

2.3.1.2 Approche par décomposition volumique

KYPRIANOU (1980) et WOO (1982) proposent les techniques de décomposition par coques convexes, qui décomposent récursivement l'objet 3D en fonction de la convexité de la pièce. Un objet est représenté par un arbre de construction composé d'objets convexes simples. Une évolution proposée par WACO et KIM (1994) est d'ajouter un opérateur d'addition permettant de fusionner des pièces convexes. La décomposition par tuile convexe est très intéressante d'un point de vue de la modélisation géométrique, mais l'algorithme de décomposition ne garantit pas de fournir un modèle correct. La figure 2.2 montre un exemple de décomposition d'un objet 3D en coque convexe.

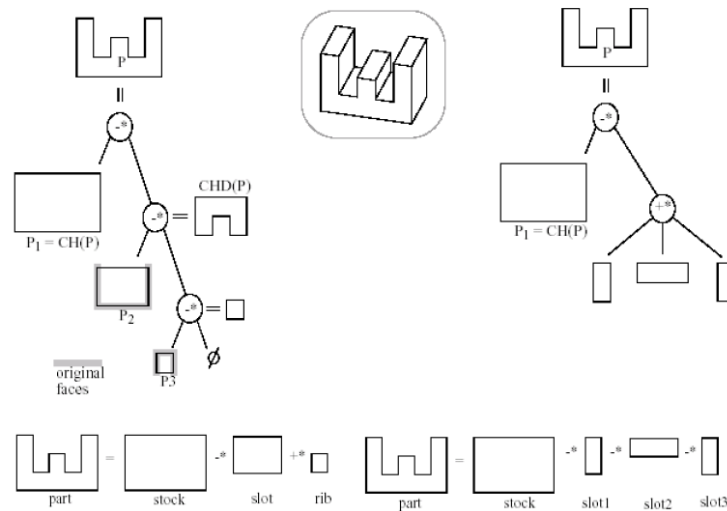


FIG. 2.2 – Décomposition par coque convexe (WACO et KIM, 1994).

2.3.2 Approches basées sur les squelettes

Le squelette et la transformée en axe médian sont des outils nés du besoin de décrire de manière compacte les propriétés globales d'un objet et en particulier la forme. Ces approches caractérisent la forme d'un objet 3D comme une suite de lignes courbes, représentant un graphe caractérisant le squelette. La structure de ce graphe nous donne l'information sur la forme de l'objet. Si l'on attache à chaque point de ce graphe la valeur de la distance au point le plus proche du contour (le rayon de la boule maximale centré en ce point) nous obtenons ce qui, dans la littérature, est nommé axe médian. Grâce à celui-ci, nous aurons aussi accès à l'information concernant le contour (frontière) de l'objet. Puisque le squelette est destiné à représenter l'objet, le transformer en axe médian ou, plus généralement, le processus de squelettisation doit satisfaire les conditions suivantes :

- être homotopique (BEUCHER, 1990) : cela signifie qu'il faut que la topologie de l'objet soit préservée. Par exemple, si un objet est connexe, il faut que son squelette le soit aussi ;
- assurer la bonne localisation du squelette afin d'obtenir, entre autres, l'invariance du squelette par rapport aux transformations isométriques. Il faut donc que la méthode s'ajuste aux définitions données par BLUM (1967) et MONTANARI (1968) ;
- être réversible. L'objet doit pouvoir être reconstruit à partir de son axe médian. On peut remarquer que la plupart des méthodes ne vérifient pas ces trois propriétés.

Le principal problème de toutes ces méthodes est que le squelette est sensible au moindre changement dans la frontière de l'objet. Une petite saillie ou bruit du contour ajoute une nouvelle branche, ou barbule (IRANI et WARE, 2003), au squelette. La plupart des algorithmes de squelettisation font l'usage de techniques de simplification (SHAKED et BRUCKSTEIN, 1998), afin d'éliminer ce type de branches.

La suite de cette partie est consacrée aux principales méthodes d'extractions de squelettes utilisées en indexation d'objets 3D.

2.3.2.1 Algorithmes d'amincissement

GONG et BERTRAND (1990) utilisent des approches qui cherchent à amincir l'objet tout en gardant sa topologie initiale. Ce sont des algorithmes itératifs, qui partent du contour initial de l'objet, étudient la connexité de chaque point du contour dans son voisinage et enlèvent ceux dont la suppression n'affecte pas la topologie de l'objet. Ainsi, l'amincissement de l'objet consiste à enlever successivement des couches. Il y a aussi, dans le cadre de la morphologie mathématique, des algorithmes d'amincissement basés sur l'érosion successive du contour de l'objet. En général, ces types d'algorithmes sont non invariants par rotation. Par contre, ils sont applicables aux objets tridimensionnels et sont largement employés (GAGVANI et SILVER, 1999). La figure 2.3 montre un exemple de squelette 2D obtenue par amincissement d'une forme 2D.



FIG. 2.3 – Une feuille et son squelette avec diverses valeurs d'amincissement.

2.3.2.2 Squelette à partir de la carte de distance

L'utilisation de la carte de distance pour le calcul du squelette est une idée très répandue. Dans la figure 2.4, qui représente la carte de distance d'un objet bidimensionnel, comme une surface d'élévation, la hauteur pour chaque point de l'objet est égale à la distance de ce point au contour de l'objet.

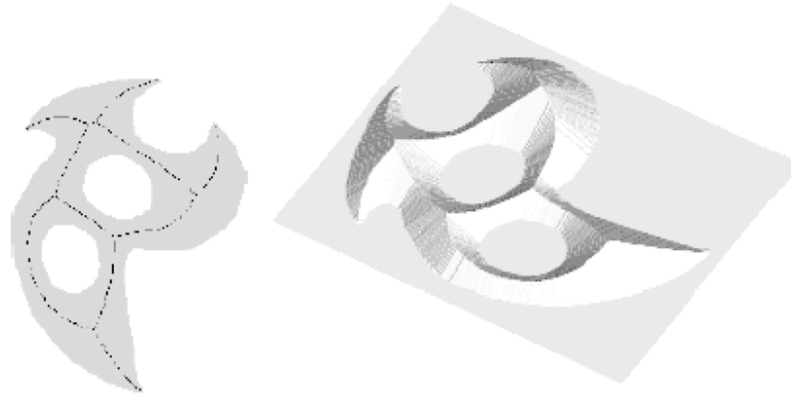


FIG. 2.4 – Un objet bidimensionnel, son squelette et la carte de distance de l'objet.

Visuellement, la relation entre cette surface et le squelette de l'objet est très claire : la projection des arêtes de cette surface dans le plan, donne une structure très similaire au squelette de l'objet. Dans le cas continu, il suffit de calculer la dérivée de la carte des distances et le squelette est défini comme le lieu des discontinuités. Dans la littérature, plusieurs solutions existent pour le calcul de la carte de distance discrète ou transformée en distance de l'objet (HIRATA, 1996; MAURER JR. ET AL., 2001). Nous retiendrons de ces méthodes qu'elles permettent un calcul en temps optimal de la carte en utilisant la métrique euclidienne exacte (donc sans erreurs) quel que soit la dimension.

Une approche classique calcule le squelette en deux étapes, dont la première consiste à trouver un sous-ensemble initial du squelette, par exemple les maxima locaux de la carte de distance (COEURJOLLY, 2003). Dans la deuxième étape, on essaie de le rendre connexe, on cherche des configurations de voisinage dans la carte de l'ensemble des maxima locaux. Par la suite, des processus de filtrage peuvent être mis en place pour simplifier le squelette obtenu (ATTALI ET AL., 1997).

2.3.2.3 Simulation de la propagation de feu de prairie

Un troisième type d'approche est la simulation de la propagation de feu de prairie à partir des contours vers l'intérieur de l'objet. On détecte les intersections ou point d'extinction des fronts de ces feux au fur et à mesure qu'ils se propagent dans le domaine discret ou analytique. Le principe consiste à déclarer tous les pixels du

contour comme émetteur d'une onde et de la propager au voisinage. Un pixel appartenant à deux ondes au même instant appartiendra au squelette. Dans le cas discret, les fronts peuvent s'intercepter sur des pixels qui ne sont pas points d'extinction, il faut donc mettre en place un processus pour traiter ces anomalies.

2.3.2.4 Squelette à partir du diagramme de Voronoï

Le point de départ de cette approche est donné par [KIRKPATRICK \(1979\)](#), où il postule que le squelette d'une forme polygonale est un sous-ensemble du diagramme de Voronoï de cette forme (diagramme de Voronoï d'un ensemble de segments, pas d'un ensemble de points). Cette idée est mise en oeuvre par [LEE \(1982\)](#). Il calcule d'abord le diagramme de Voronoï du polygone, et ensuite il supprime les arêtes partant des sommets des parties concaves de ce polygone. Mais l'utilisation de cette idée pour créer des squelettes d'objets ayant une forme arbitraire est complexe à mettre en place. D'une part, l'approximation polygonale d'une forme quelconque est un problème assez difficile, et d'autre part le calcul du diagramme de Voronoï est long et produit un très grand nombre de barbules, surtout lorsqu'il y a des trous dans l'objet.

Pour ces raisons, [ATTALI ET AL. \(1994\)](#); [OGNIEWICZ et ILG \(1992\)](#) proposent des approches qui calculent le squelette à partir du diagramme de Voronoï (DV) d'un ensemble de points, cet ensemble étant un échantillonnage discret du contour (continu) de l'objet. Les diagrammes de Voronoï, et les squelettes trouvés via ces approches convergent vers le cas idéal au fur et à mesure qu'on augmente la densité de l'échantillonnage. Donc, dans ces approches, le problème principal est de trouver un échantillonnage discret qui approxime au mieux la forme de l'objet, afin que cette idée puisse être mise en pratique. Ces méthodes sont aussi désignées sous le terme de méthodes continues car elles utilisent des techniques propres à la géométrie algorithmique, ce qui permet de s'abstraire de la grille discrète. De nombreuses variantes de cette méthode, se différencient par la sélection des points et par la méthode d'échantillonnage. Cette méthode est dans la plupart des cas associée à une méthode de simplification qui permet de réduire le nombre de branches du squelette.

2.3.2.5 Exosquelette

[SERRA \(1982\)](#) fut le premier à proposer de caractériser la forme d'un objet par son exosquelette, qui possédait des propriétés d'invariance aux rotations et à la présence de bruit. L'exosquelette contrairement au squelette classique ne représente pas l'objet par son squelette fil de fer, qui est calculé sur l'objet lui même, mais calcule le squelette du complémentaire de l'objet.

[LIPIKORN ET AL. \(2002b\)](#) ont mis en place les exosquelettes modifiés (mES), qui adoptent le concept des exosquelettes classiques mais différents par le fait de ne

plus utiliser le complémentaire mais le cercle circonscrit pour réduire l'ensemble des points complémentaires de l'objet X . La figure 2.5 montre un exemple de squelette et d'exosquelette sur une forme 2D.

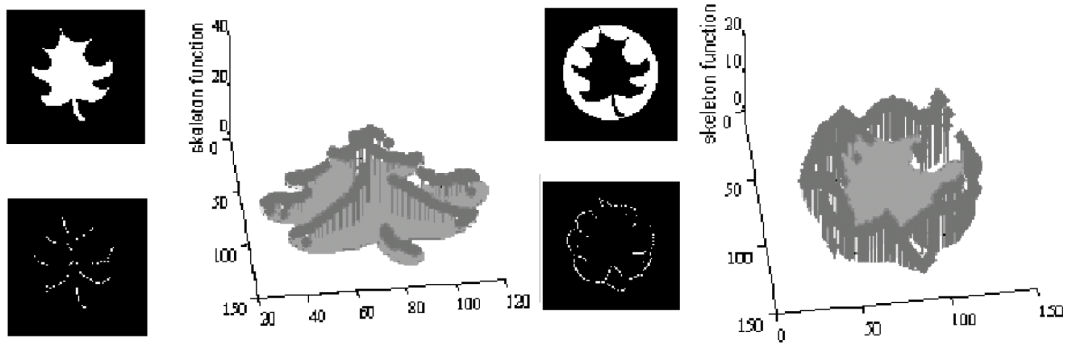


FIG. 2.5 – Exemple de squelette et d'exosquelette d'une image binaire et de leur fonction de répartition (LIPIKORN ET AL., 2002a).

LIPIKORN ET AL. (2002a) ont étendu le concept d'exosquelette modifié aux objets 3D. Ils reprennent le même algorithme en remplaçant le cercle circonscrit par une sphère et en faisant les calculs dans un espace à trois dimensions. La mesure de similarité entre mES recherche les similarités locales entre les deux mES en appariant chaque voxel du mES objet avec celui le plus proche du mES modèle.

2.3.2.6 Graphe de Reeb multi résolution

Un graphe de Reeb (REEB, 1946) est une structure topologique et squelettique pour un objet de dimension arbitraire. La définition d'un graphe de Reeb est comme suit : soit $\mu : C \rightarrow \mathbb{R}$ est une fonction continue défini sur l'objet C . Le Graphe de Reeb est un sous espace des graphes μ dans $C \times \mathbb{R}$ par la relation d'équivalence $(X_1, \mu(X_1)) \sim (X_2, \mu(X_2))$ qui tient si et seulement si $\mu(X_1) = \mu(X_2)$ et X_1 et X_2 sont dans la même composante connexe de $\mu^{-1}(\mu(X_1))$. On peut remarquer que le choix de la fonction de poids μ va modifier le graphe de Reeb. La figure 2.6 montre la distribution de la fonction de poids sur la surface d'un tore et le graphe de Reeb correspondant.

Un graphe de Reeb multi résolution (MRG) consiste à développer une série de graphes de Reeb à différents niveaux de détails, les résolutions étant définies en fonction de la fonction de poids. Les noeuds du graphe de Reeb représentent une région de l'objet et des noeuds adjacents sont reliés par des arrêtes si les régions de l'objet ainsi représentées sont mitoyennes, comme le montre la figure 2.7.

Un graphe de Reeb est toujours produit par une fonction continue μ . Si une fonction différente est employée le graphe de Reeb changera. Il est important que la fonction μ soit soigneusement définie pour l'application. Par exemple, une fonction de

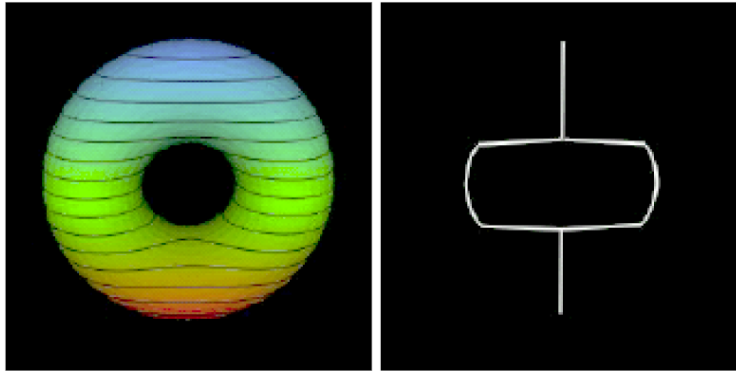


FIG. 2.6 – Tore et graphe de Reeb correspondant pour une certaine fonction de poids. la fonction de point est sur cet exemple la hauteur (HILAGA ET AL., 2001).

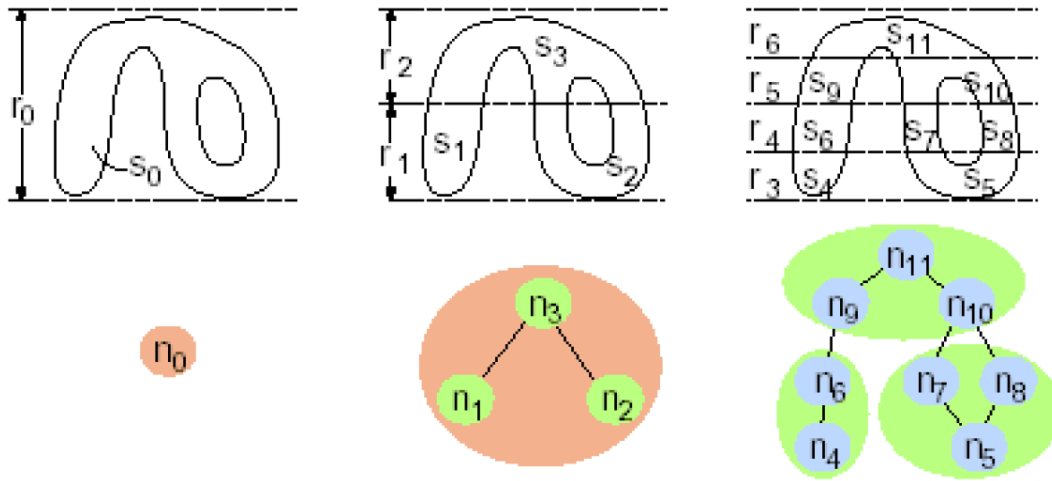


FIG. 2.7 – Graphe de Reeb multi résolution (HILAGA ET AL., 2001).

poids (figure 2.6) n'est pas appropriée pour l'identification d'objets 3D car elle n'est pas invariante aux transformations géométriques comme la rotation. Pour palier à ces problèmes, HILAGA ET AL. (2001) propose d'utiliser une distance géodésique, c'est à dire, la distance entre deux points sur une surface. Employer la distance géodésique fournit l'invariance aux rotations, aux bruits et petites ondulations ou déformations. La fonctions μ à un point v sur une surface S est construite par :

$$\mu(v) = \int_{p \in S} g(v, p) ds \quad (2.1)$$

où la fonction $g(v, p)$ retourne la distance géodésique entre deux points v et p sur la surface S . Cette fonction est définie comme la somme de la distance géodésique de v à tous les points de S , une petite valeur signifie que la distance de v aux points de la surface est relativement petite, donc le point v est proche du centre de l'objet.

Par contre, la fonction n'est pas invariante à la graduation de l'objet, pour le rendre indépendant aux variations d'échelles, une version normalisée de $\mu(v)$ est employée :

$$\mu_n(v) = \frac{\mu(v) \min_{p \in S} \mu(p)}{\max_{p \in S} \mu(p)} \quad (2.2)$$

La stratégie de comparaison de deux MRG consiste à appairer tous les noeuds ayant une configuration topologique similaire en terme de connexions avec les voisins et de relations parent/fils entre les niveaux de résolutions successifs. Dans un premier temps, on détecte les paires de noeuds respectant un ensemble de critères de cohérence topologique prédéfinis. Puis, on retient parmi ces paires de noeuds ceux qui serviront à estimer la similarité des deux MRG.

Pour parvenir à des appariements plus optimaux, TUNG et SCHMITT (2004, 2005) ont enrichi les MRG en rajoutant des critères géométriques et colorimétriques et en étendant les critères de cohérence topologique. Les graphes ainsi créés sont appelés les graphes de Reeb multi résolution augmentés (aMRG). Ces graphes contiennent, en particulier, sur chacun des noeuds des informations géométriques sur la position du noeud par rapport à l'objet, ce qui permet de diminuer les mauvais appariements, comme le montre la figure 2.8.

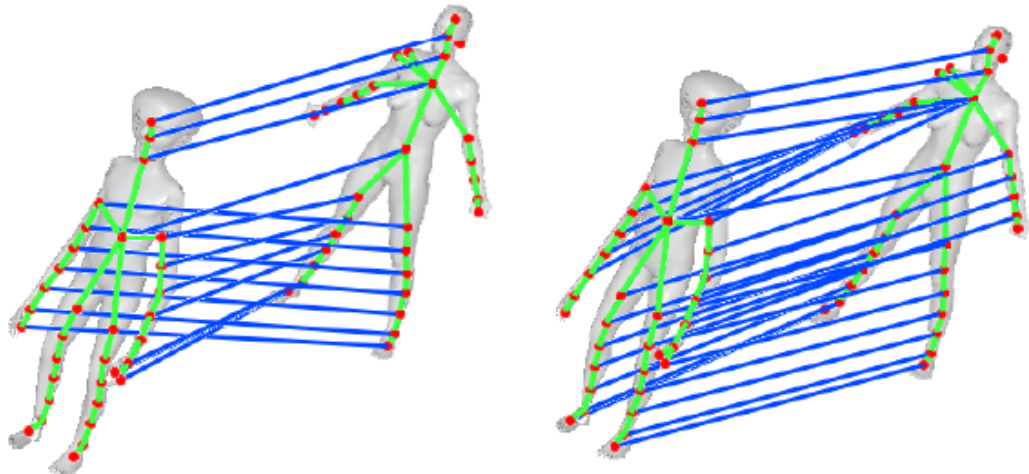


FIG. 2.8 – Exemples d'appariements entre deux MRG (gauche) et deux aMRG (droite) : sans l'information géométrique, les jambes peuvent être appariées aux bras car ils sont topologiquement équivalents. En ajoutant l'information géométrique, tous les appendices sont bien appariés (TUNG et SCHMITT, 2005).

2.3.3 Approche volumique

Les méthodes volumiques représentent les objets tridimensionnels comme une suite d'objets 3D élémentaires. Plusieurs types de primitives volumiques ont été

utilisés pour représenter des objets tridimensionnels. On peut citer les méthodes utilisant des cylindres, des modèles ellipsoïdes, des Géons, des super-quadriques (CHEVALIER ET AL., 2003) et des hyper-quadriques.

En 1985, BIEDERMAN (1986) propose sa théorie concernant la compréhension humaine des images, les méthodes de reconnaissance par composants : *recognition by component* (RBC). Ces recherches dans ce domaine suggèrent que les humains comprennent les objets en les décomposant en modules élémentaires (composants). Partant de ces observations, BERGEVIN et LEVINE (1993) définissent des formes de base ou « géons » (géométriques). Un géon est un objet géométrique élémentaire, ils sont au nombre de 36. Chaque objet peut être approximé par une structure formée de ces géons géométriques. Ces techniques peuvent s'appliquer aussi bien à la classification d'objets 2D, à la recherche d'un objet tridimensionnel au sein d'une scène 2D et au classement d'objets 3D entre eux.

La structure hiérarchique menant à l'identification des objets 3D se décompose alors selon le schéma de la figure 2.9, dont les cinq niveaux sont :

- extraction des contours ;
- détection des configurations non accidentelles et extraction des zones concaves, dans lesquelles on regroupe les contours pour mettre en évidence des propriétés comme l'alignement, le parallélisme etc. ;
- activation des géons et les relations entre eux, qui se basent sur les informations précédemment obtenues pour formuler des hypothèses sur l'existence des géons dans des zones de l'image ainsi que les relations de position entre ces derniers ;
- activation des modèles d'objets, qui émet des hypothèses quant à la présence ou non de modèles, basés sur la présence des géons le composant ;
- identification des objets, qui sélectionne les modèles effectivement présents dans la scène.

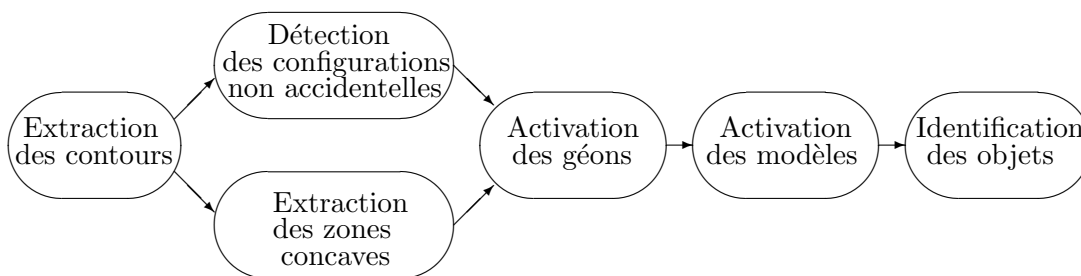


FIG. 2.9 – Les différentes phases pour la reconnaissance par composants proposée par Biederman.

Différents accordements de mêmes géons peuvent mener à différents objets, comme le montre la figure 2.10. Pour cela, un ensemble de relations entre les géons a été défini pour décrire des objets complexes (la verticalité, la taille relative, l'adjacence...).

D'autres travaux plus récents s'intéressent à interclasser des objets entre eux en

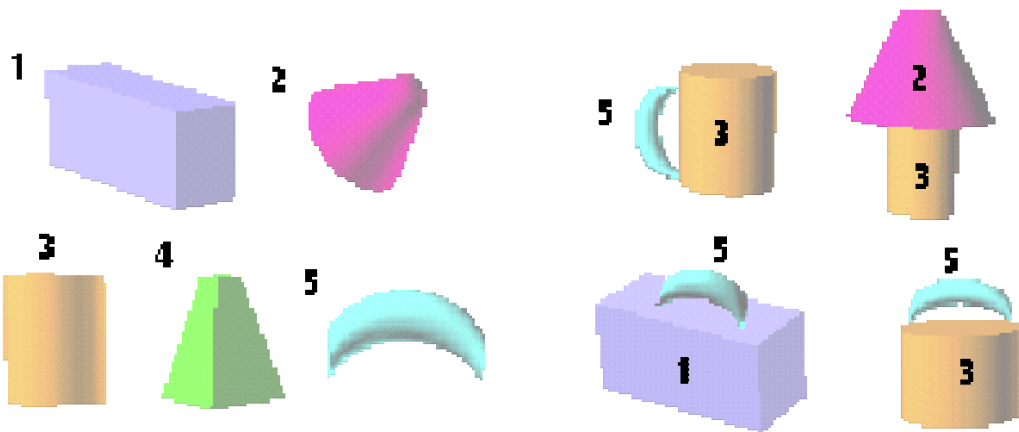


FIG. 2.10 – Exemple de primitives d’objets définies par la théorie de Biederman : les géons et différentes connections créant des objets différents (IRANI et WARE, 2003).

les indexant grâce aux géons. Nous pouvons citer les travaux de IRANI et WARE (2003) qui transforment le graphe de géons 3D en un diagramme UML 2D et compare par la suite les diagrammes UML entre eux.

2.3.4 Représentation symbolique

NEAL et SHAPIRO (2000) proposent de suivre le point de vue de HERNANDEZ (1994), qui est d’utiliser une représentation symbolique du modèle tridimensionnel pour définir son appartenance à une classe d’objets. L’algorithme définit des modèles abstraits qui caractérisent les classes d’objets qui seront par la suite retrouvées. Le modèle abstrait se compose d’une ontologie spatiale d’objets et de trois réseaux : le réseau de topologie, le réseau de partie et le réseau de rapports spatiaux. Une ontologie est un ensemble d’informations dans lequel sont définis les concepts utilisés dans un langage donné et qui décrit les relations logiques qu’ils entretiennent entre eux.

La mise en correspondance se fait en plusieurs étapes. Dans un premier temps, le programme cherche pour chaque caractéristique du modèle courant des ensembles de points du maillage où la caractéristique peut se trouver. La deuxième étape consiste à affiner ces ensembles de points en fonction des propriétés de forme et des valeurs des paramètres de forme pour garder uniquement les caractéristiques vraiment présentes dans l’objet. La dernière étape consiste à regrouper les caractéristiques présentes sur le modèle et en fonction de leur label, statuer sur l’appartenance ou non du modèle à la classe du modèle abstrait. La figure 2.11 montre un exemple de comparaison symbolique.

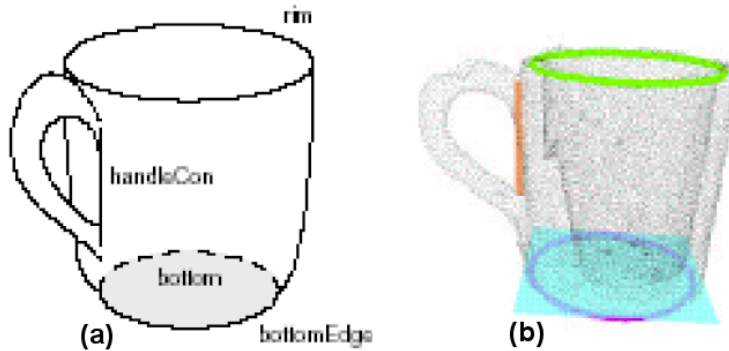


FIG. 2.11 – Comparaison symbolique. (a) Caractéristiques de l'ontologie, (b) modèle caractérisé par la recherche de caractéristiques (NEAL et SHAPIRO, 2000).

2.4 Approches par transformées

Caractéristiques du domaine de la reconnaissance, les représentations d'objets 3D fondées sur des transformées visent à déterminer des représentations de forme globales, définies en terme de transformation intégrale. De nombreuses techniques existent, et un certain nombre d'entre elles sont expliquées dans le livrable 2.1 du projet RNRT Semantic-3D. Nous nous sommes principalement intéressé dans cette partie, aux méthodes basées sur la transformée de Hough 3D et aux approches basées sur les moments.

2.4.1 Transformée de Hough

Une littérature très riche souligne tout l'intérêt porté aux approches à base de transformée de Hough. La transformée de Hough (HOUGH, 1962; ILLINGWORTH et KITTLER, 1988) s'appuie sur un principe d'accumulation de différentes primitives élémentaires paramétrées selon des variables spécifiques. Les maxima locaux de l'accumulateur indiquent alors la présence de ces différents éléments au niveau des données et permettent leur détection. Initialement appliquée à des droites du plan, paramétrées en coordonnées polaires, la transformée de Hough a été ensuite étendue à d'autres types d'éléments, tels que droites et plans de \mathbb{R}^3 , coniques, sphères et ellipsoïdes, (DUDA et HART, 1972; TSUJI et MATSUMOTO, 1978; BALLARD, 1981; PAO ET AL., 1992; HU, 1995; BECKER, 1998; DAVIES et NIXON, 1998; BENNETT ET AL., 1999; GUIL ET AL., 1999; BHATTACHARYA ET AL., 2000).

D'autres axes de recherche considèrent des extensions probabilistes de la transformée de Hough (XU ET AL., 1990; McLAUGHLIN, 1998; KIRYATI ET AL., 2000), qui visent à réduire la complexité de calcul en échantillonnant aléatoirement les données, sans pour autant diminuer les performances de la représentation en terme de pouvoir de discrimination. Enfin, mentionnons également les versions floues de

la transformée de Hough (BHANDARKAR, 1994; PHILIP ET AL., 1994; SOODAMANI et LIU, 1998), se proposant de combiner les principes de la transformée de Hough classique avec ceux de la théorie des ensembles flous.

Ces différentes méthodes relèvent d'une approche purement surfacique, consistant à détecter différentes variétés de dimension $(n - 1)$ plongées dans l'espace \mathbb{R}^n . Cependant, remarquons qu'un autre point de vue, cette fois de nature volumique, peut être associé à la transformée de Hough. Cela établit en effet le lien avec la transformée de Radon (DEANS, 1983) qui définit les fondements théoriques de la reconstruction d'objets à partir de ses projections.

Cette logique permet d'élargir le cadre applicatif de la transformée de Hough au-delà des applications élémentaires de détection de primitives et notamment dans le cadre de l'indexation et des requêtes par similarité. ZAHARIA et PRÊTEUX (2002a,b) proposent une nouvelle représentation, fondée sur un descripteur de Hough 3D optimal (DH3DO) qui permet de s'affranchir des problèmes d'alignement spatial par un échantillonnage de la sphère unité invariant aux repères issus d'une analyse en composantes principales de l'objet 3D. Cela permet de définir des mesures de similarité intrinsèquement symétriques, qui conduisent à une réduction importante du temps de calcul des réponses aux requêtes.

2.4.2 Approches basées sur les moments

Les approches basées sur les moments peuvent être définies comme la projection de la fonction définissant l'objet sur un ensemble de fonctions caractéristiques. Ces approches ont été utilisées dans la reconnaissance de forme 2D avec des moments 2D, et on peut citer : les moments géométriques, Legendre, Fourier-Mellin, Zernike, pseudo-Zernike (TEH et CHIN, 1988), ART (voir chapitre 1.2.3). Plusieurs de ces moments ont été étendus en 3D : Fourier 3D (ELAD ET AL., 2002), Ondelette 3D (PAQUET et RIOUX, 1999), Zernike 3D (CANTERAKIS, 1999) et la décomposition en harmoniques sphériques FUNKHOUSER ET AL. (2003); KAZHDAN ET AL. (2003).

Le calcul des moments 3D consiste à projeter une fonction f , représentant un objet 3D, sur une famille de fonctions $\Phi = \{\phi_i\}_{i \in N}$, dans un domaine de définition Ω . Ceci peut s'écrire :

$$\mu_i = \langle f, \phi_i \rangle = \int_{\Omega} f(x) \cdot \phi_i(x) dx \quad (2.3)$$

où, $f(x)$ est la fonction représentant l'objet 3D, ϕ_i la i_{me} fonction de base et Ω le domaine de définition où est intégrable la fonction f . Les coefficients μ définissent l'objet 3D et sont utilisés comme descripteurs. Le niveau de précision est défini en fonction du nombre de fonctions de base ϕ et des descripteurs choisis. Différentes familles de fonctions de base peuvent être utilisées pour caractériser la forme d'un objet 3D, et on peut citer :

- Les moments statistiques centrés 3D, calculés sur la surface de l'objet, propo-

sées par PAQUET et RIOUX (1999), qui pour chaque triangle calcule :

$$\mu_{pqr} = \sum_{i=0}^n m_i (x_i - x_0)^p (y_i - y_0)^q (z_i - z_0)^r \quad (2.4)$$

où n est le nombre de triangles du maillage, (x_i, y_i, z_i) le barycentre du triangle, m_i le poids associé au triangle i et (x_0, y_0, z_0) le centre de l'objet.

- Les moments géométriques d'ordre (p, q, r) , utilisés par SAUPE et VRANIC (2001),
- La transformée de Fourier 3D utilisé par VRANIC et SAUPE (2001),
- Les moments de Zernike 3D proposé par CANTERAKIS (1999) et utilisé par NOVOTNI et KLEIN (2004), et
- Les harmoniques sphériques de FUNKHOUSER ET AL. (2003).

2.4.2.1 Les moments géométriques

Les moments géométriques sont définis sur les ensembles de base $\{x^p y^q z^r\}$ sur la représentation discrète des objets 3D. Les moments géométriques tridimensionnels d'ordre $(p + q + r)$ sont notés par m_{pqr} et peuvent être calculés par :

$$m_{pqr} = \int \int \int_{\Omega} x^p y^q z^r f(x, y, z) dx dy dz \quad p, q, r = 0, 1, 2, 3, 4 \dots \quad (2.5)$$

où Ω est la région de l'espace dans laquelle est définie la fonction image $f(x, y, z)$.

Les moments géométriques ont été utilisés comme descripteurs de forme 2D pour caractériser la forme de la région extrait d'une image 2D. SAUPE et VRANIC (2001) ont appliqué ceci sur les objets 3D discrets en appliquant la formule 2.5. Ceci peut être mis en parallèle des approches par moments statistiques centrés 3D de PAQUET et RIOUX (1999), qui calculent les moments géométriques directement sur les points du maillage de l'objet. Cette approche est moins robuste aux remaillages comme nous le verrons dans le chapitre 6, ou nous comparerons les deux types d'approche et nous proposerons une méthode pour calculer les moments géométriques sur le volume d'un maillage 3D.

2.4.2.2 Descripteur de Fourier 3D

VRANIC et SAUPE (2001) proposent d'utiliser les coefficients de la transformée de Fourier comme descripteur de forme. Sur l'objet discrétisé préalablement centré, mis à l'échelle et aligné par une analyse en composante principale, les premiers coefficients de Fourier sont calculés selon la formule :

$$g_{uvw} = \frac{1}{N^3} \sum_{i=-\frac{N}{2}}^{\frac{N}{2}-1} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \sum_{l=-\frac{N}{2}}^{\frac{N}{2}-1} q_{ikl} \exp \left(-j \frac{2\pi}{N} (iu + kv + lw) \right) \quad (2.6)$$

ou q_{ikl} est l'ensemble des voxels composant l'objet. La valeur absolue des coefficients g_{uvw} avec des indices $-K \leq u, v, w \leq K$ sont sélectionnées comme descripteurs. Le vecteur descripteurs est composé de $((2K - 1)^3 + 1)/2$ valeurs réelles. Les meilleurs résultats sont obtenus pour $K = 3$.

2.4.2.3 Les moments de Zernike 3D

La description d'une forme par le descripteur de Zernike 3D a été définie par CANTERAKIS (1999) et utilisée par NOVOTNI et KLEIN (2003, 2004). Le calcul du descripteur de Zernike 3D, tel qu'il est proposé par les auteurs, se fait de façon discrète sur un objet discrétisé et normalisé, en combinant linéairement les moments géométriques m_{pqr} , pour tous $p, q, r \geq 0$ et $p + q + r \leq n$, avec $n \in [0, N]$ et $l \in [0, n]$, tel que $(n - l)$ est paire, et $m \in [-l, l]$. Le descripteur de forme de Zernike 3D est la norme du vecteur Ω_{nl} :

$$F_{nl} = \|\Omega_{nl}\| \quad \text{ou} : \quad \Omega_{nl}^m = \frac{3}{4} \sum_{p+q+r \leq n} \overline{\chi_{nlm}^{pqr}} m_{pqr} \quad (2.7)$$

Les fonctions Ω_{nl}^m sont les moments de Zernike 3D d'un objet, et sont une combinaison linéaire des moments géométriques d'ordre n . La combinaison linéaire se fait grâce à la fonction χ_{nlm}^{rst} selon la formule :

$$\begin{aligned} \chi_{nlm}^{pqr} = & c_l^m 2^{-lm} \sum_{v=0}^k q_{kl}^v \sum_{\alpha=0}^v \binom{v}{\alpha} \sum_{\beta=0}^{v-\alpha} \binom{v-\alpha}{\beta} \sum_{u=0}^m (-1)^{m-u} \binom{m}{u} i^u \\ & \cdot \sum_{\mu=0}^{\lfloor \frac{l-m}{2} \rfloor} (-1)^\mu 2^{-2\mu} \binom{l}{\mu} \binom{l-\mu}{m+\mu} \sum_{v=0}^{\mu} \binom{\mu}{v} \end{aligned} \quad (2.8)$$

où $2k = n - 1$, c_l^m est le facteur de normalisation et q_{kl}^v garantit l'orthogonalité des fonctions avec la sphère unité. Ces deux termes se calculent avec les formules suivantes :

$$c_l^m = c_l^{-m} = \frac{\sqrt{(2l+1)(l+m)!(l-m)!}}{l!} \quad (2.9)$$

$$q_{kl}^v = \frac{(-1)^k}{2^{2k}} \sqrt{\frac{2l+4k+3}{3}} \binom{2k}{k} (-1)^v \frac{\binom{k}{k} \binom{2(k+l+v)+1}{2k}}{\binom{k+l+v}{k}} \quad (2.10)$$

2.4.2.4 Les harmoniques sphériques

La description en harmoniques sphériques (HS) est une des méthodes les plus efficaces de la littérature et sera utilisée dans les expérimentations du paragraphe

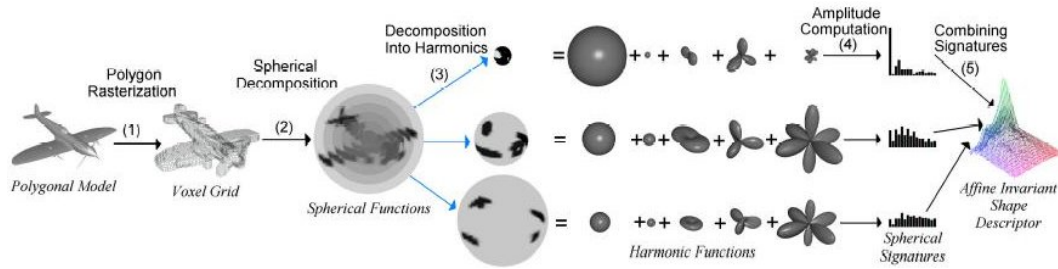


FIG. 2.12 – Décomposition en harmoniques sphériques (KAZHDAN ET AL., 2003).

5.2.3, pour estimer les performances de la méthode proposée. La description HS a été introduite par FUNKHOUSER ET AL. (2003). Cette description décompose les formes 3D centrées en un ensemble de composants indépendants à la rotation par l'échantillonnage de l'espace tridimensionnel en coquilles concentriques. Après une étape de centrage, la description en harmoniques sphériques décompose les formes 3D en un ensemble irréductible de composants indépendants à la rotation par l'échantillonnage de l'espace tridimensionnel en coquilles concentriques, où les coquilles sont définies par des intervalles radiaux égaux. Les fonctions sphériques sont décomposées comme la somme des 16 premières composantes harmoniques (KAZHDAN ET AL., 2003), de façon analogue à une décomposition de Fourier dans différentes fréquences. Utilisant le fait que la rotation ne change pas la norme des composantes harmoniques, la signature de chaque fonction sphérique est définie comme la liste des 16 normes. Ainsi calculées, les différentes signatures sont combinées pour obtenir un descripteur de taille $32 * 16$ pour chaque objet. Durant les phases de recherche, la similarité entre objets est calculée par une distance Euclidienne entre les vecteurs. La figure 2.12 montre le principe de la décomposition en harmoniques sphériques.

2.5 Approches par vues

Les méthodes utilisant des vues des modèles 3D sont motivées par les résultats psychophysiques, comme énoncé par RIESENHUBER et POGGIO (2000), qui montrent que dans le système visuel humain, un objet en trois dimensions est représenté par un ensemble de vues 2D plutôt que par un modèle tridimensionnel. L'idée générale de ce type de méthode est que « si deux objets sont similaires, leurs vues seront aussi similaires de tous les points de vues ». Comme nous l'avons vu dans le chapitre 1.3 de telles méthodes peuvent être utilisées pour classer des objets 3D au départ d'une requête 2D. Ces approches peuvent aussi être utilisées pour classer des objets 3D. Deux objets similaires en terme de forme auront leurs ensembles de vues extraites similaires.

Les problématiques de l'indexation d'objets 3D par vues sont similaires à celle de la recherche d'objets 3D à partir de requêtes images. Il s'agit de caractériser un

objet 3D par un ensemble de vues 2D. Comme pour les approches 2D/3D, deux points sont discutés pour convenablement caractériser les objets 3D :

- le nombre de vues à prendre en compte ;
- la position des vues dans l'espace des vues.

Le point sur lequel l'indexation d'objets 3D par vues diffère de la recherche d'objets 3D par requêtes 2D, est la mesure de similarité à pratiquer. En effet, il ne s'agit plus de comparer une image 2D à un ensemble d'images caractérisant l'espace des vues d'un objet, mais de comparer deux ensembles d'images caractérisant deux espaces des vues. La mesure de similarité en 2D/3D revient à rechercher l'image de l'ensemble des vues la plus proche de l'image requête et correspond à la recherche d'un minimum sur l'ensemble. En 3D/3D par vues, il s'agit d'apparier les deux ensembles de vues pour trouver le meilleur agencement permettant de minimiser la somme des distances entre deux images des ensembles. Il faut trouver un appariement où chaque vue du première ensemble corresponde à une vue du second en respectant les positions des vues. Pour cela, différentes distances entre ensembles de vues ont été proposées.

Les approches les plus répandues consistent à comparer deux ensembles de vues en recherchant l'arrangement minimum, correspondant à ce que chaque vue caractéristique de l'objet requête soit appariée avec une vue de l'objet testé. Ceci peut être vu comme la somme des minimums et peut s'écrire :

$$D(R, O) = \sum_{i=1}^N \min_{j \in \{1..M\}} d(R_i, O_j) \quad (2.11)$$

où N est le nombre de vues caractéristiques de l'objet R , M le nombre de vues caractéristiques de l'objet O et $d(R_i, O_j)$ la mesure de similarité entre la vue i de l'objet R et la vue j de l'objet O .

CHEN ET AL. (2003) proposent de caractériser un objet 3D par dix vues orthographiques pris selon les dix premières faces d'un dodécaèdre centré autour de l'objet. Les vues sont des vues binaires représentant la silhouette de l'objet. Les auteurs ne considèrent que les dix premiers points de vues du dodécaèdre, car les images créées sur les faces inverses du dodécaèdre sont identiques par la projection orthographique. La mesure de similarité entre les deux ensembles de vues est faite en recherchant le meilleur appariement entre les vingt points de vues. Considérant les vingt points de vues formés par les centres des faces d'un dodécaèdre, il existe soixante rotations possibles entre les deux ensembles de points pour vérifier toute les combinaisons, comme le montre la figure 2.13.

La mesure de similarité doit être invariante en rotation et doit prendre en compte les soixante mises en correspondance. La formule, pour comparer deux ensembles de vues ainsi créés, est :

$$D_A = \min_i \left(\sum_{k=1}^{10} d(I_{1k}, I_{2k}) \right), \quad i = 1 \dots 60 \quad (2.12)$$

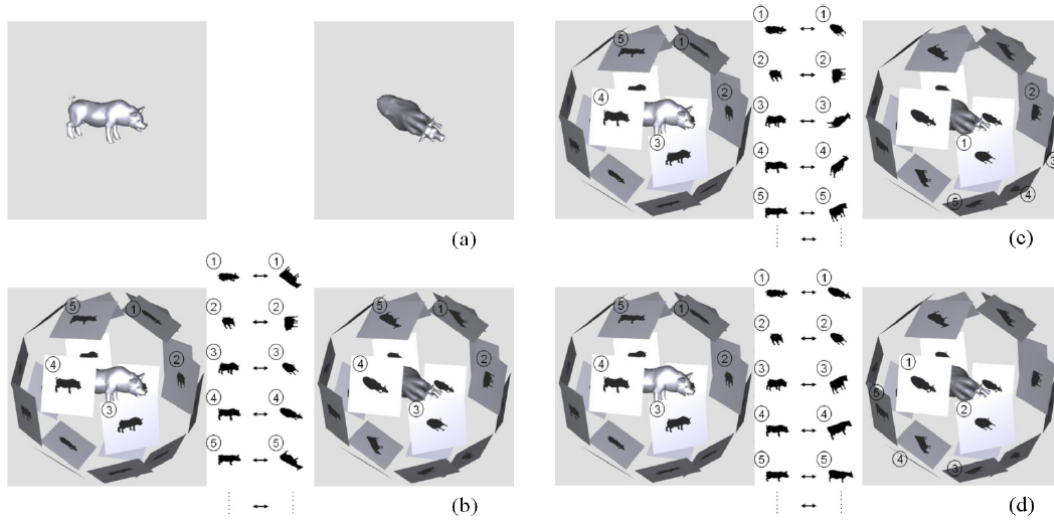


FIG. 2.13 – Comparaison de deux ensembles de vues représentant deux objets 3D (a) pour chaque objet 10 vues sont créées sur les sommets d'un dodécaèdre (b) les objets sont comparés en modifiant l'ordre de mise en correspondance des vues (c,d) (CHEN ET AL., 2003).

où, d est la mesure de similarité entre deux images et i représente les rotations entre les positions de caméra. Pour augmenter la précision du partitionnement et la robustesse, les auteurs décrivent un objet 3D par plusieurs ensembles de vues, en modifiant l'orientation de la caméra. Ils proposent de créer dix ensembles de vues, ce qui revient à caractériser les objets 3D selon cent vues caractéristiques. La mesure de similarité entre objets devient :

$$D_B = \min(D_A(L_j, L_k)) \quad j, k = 1 \dots 10 \quad (2.13)$$

où D_A est définie par l'équation 2.12, et L_j et L_k sont les ensembles de vues. Cette mesure de similarité est la plus complète, car elle prend à la fois en compte la similarité entre les ensembles de vues et la position des vues les unes par rapport aux autres.

Les approches comparant les vues des objets en fonction de leurs positions sur la sphère des vues demandent que les vues soient équi-reparties ce qui ne permet pas de mettre en place de telles méthodes avec un processus de sélection de vues variables. Le point limitant ce type d'approche est le coût de la comparaison. Pour palier à ces problèmes, CHEN ET AL. (2003) ont proposé des mesures de similarité plus légères pour permettre de faire des recherches dans des grandes bases d'objets 3D. Ils proposent de mettre en place plusieurs descripteurs de formes et plusieurs versions de leur distance entre deux ensemble de vues. Ceci leur permet de faire plusieurs études de la base d'objets 3D. Une première étude, avec un descripteur de forme et une distance entre objets de faible coût, permet de définir un premier sous

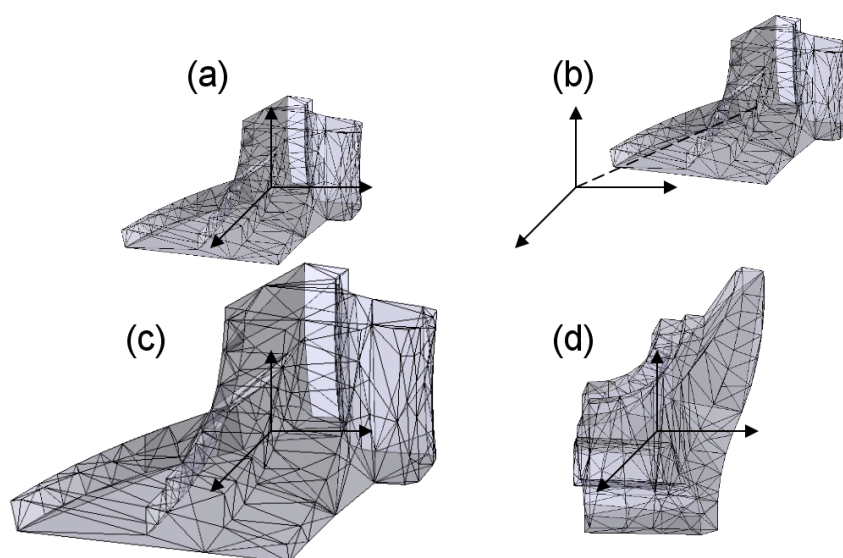


FIG. 2.14 – Exemple d’objets déformées : objet d’origine (a), objet translaté (b), objet ayant subi un changement d’échelle (c), objet ayant subi une rotation (d).

échantillon sur lequel seront exécutées les études suivantes.

Pour combiner les approches à sélections de vues variables et une mesure de similarité plus précise, [FILALI ANSARY ET AL. \(2005\)](#) ont proposé une mesure de similarité basée sur une mesure bayésienne, nommée AVC (Adaptive Views Clustering). Cette approche calcule la probabilité qu’une vue de l’objet requête appartienne aux objets de la base et somme les contributions de toutes les vues de l’objet pour calculer une mesure de similarité.

2.6 Pré-traitement

Une grande partie des méthodes d’indexation d’objets 3D que nous venons dénoncer sont robustes aux principales déformations usuelles par le fait que le pré-traitement normalise les objets. Les déformations qui peuvent déformer un objet 3D sont de trois types :

- la translation, qui déplace un objet par rapport au centre du repère ;
- le changement d’échelle, qui modifie la taille de l’objet ;
- la rotation : qui le fait tourner.

La figure 2.14 montre un exemple d’objet ayant subi ces trois déformations.

Certaines méthodes sont naturellement invariantes à certaines de ces déformations (le descripteur ne change pas). On parle de méthode invariante aux rotations, aux changements d’échelle et aux translations. A l’opposé, on qualifie des méthodes de robustes à ces déformations, si elles possèdent un pré-traitement, qui normalise

les objets avant de les indexer.

La normalisation porte sur trois points, qui sont :

- le centrage, qui place le centre de gravité des objets au centre du repère pour rendre les méthodes robustes aux translations,
- la mise à l'échelle, qui modifie la taille de l'objet pour rendre les méthodes robustes aux changements d'échelles,
- l'alignement, qui aligne les objets par rapport aux axes pour rendre les méthodes robustes aux rotations.

2.6.1 Centrage

Le centrage consiste à aligner le centre de gravité de l'objet avec le centre du repère. Le calcul du centre de gravité de l'objet, $G = (g_x, g_y, g_z)$, peut se faire en sommant les composantes de chaque point du maillage de l'objet, selon la formule :

$$G = \left(\frac{\sum_{i=0}^N x_i}{N}, \frac{\sum_{i=0}^N y_i}{N}, \frac{\sum_{i=0}^N z_i}{N} \right) \quad (2.14)$$

où N est le nombre de points constituant le maillage de l'objet.

2.6.2 Mise à l'échelle

La mise à l'échelle a pour but de rendre les objets robustes aux changements d'échelle. Cette normalisation consiste à mettre tous les objets à la même taille. Pour cela, les objets sont intégrés dans la sphère unité, de centre, le centre du repère et de rayon 1. Il faut rechercher le point le plus loin du centre du repère pour tous les points de l'objet centré :

$$D = \max_{i=1..N} d(O, (X_i, Y_i, Z_i)) \quad (2.15)$$

où d est la fonction qui calcule la distance entre les points i et le centre du repère, et pratiquer une normalisation par D , pour ramener tous les points à l'intérieur de la sphère unité selon la formule :

$$(X_i, Y_i, Z_i) = \left(\frac{x_i}{D}, \frac{y_i}{D}, \frac{z_i}{D} \right) \quad (2.16)$$

La figure 2.15 montre un exemple de ce processus de mise à l'échelle.

2.6.3 Alignement

L'alignement permet de rendre les méthodes d'indexation robustes aux rotations 3D. Le principe de l'alignement est de calculer les axes principaux des objets et

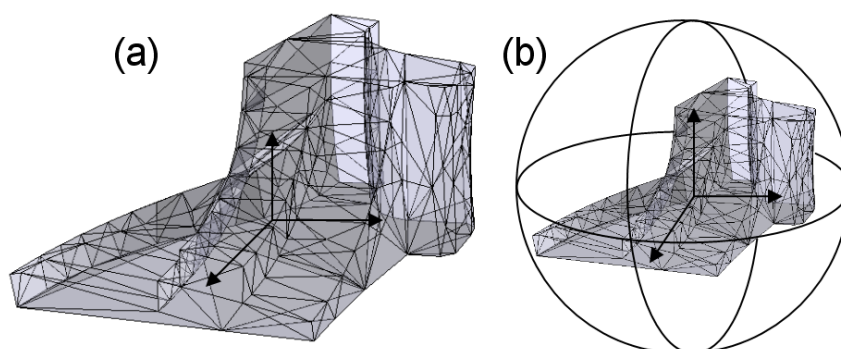


FIG. 2.15 – Exemple d'un objet mise à l'échelle, objet d'origine (a) et objet intégré à la sphère unité (b).

d'aligner ces axes avec ceux du repère. Le calcul des axes principaux d'un objet peut se faire en extrayant les vecteurs propres de la matrice contenant les moments géométriques d'ordres deux :

$$C = \begin{pmatrix} m_{200} & m_{110} & m_{101} \\ m_{110} & m_{020} & m_{011} \\ m_{101} & m_{011} & m_{002} \end{pmatrix}. \quad (2.17)$$

Les axes principaux sont obtenus en calculant les vecteurs propres de la matrice C . Ceci est connu comme l'Analyse en Composante Principale (ACP). Le calcul des valeurs propres et des vecteurs propres de la matrice peut se fait par la méthode de Jacobi (GOLDSTINE ET AL., 1959). Il suffit alors d'aligner les vecteurs propres aux axes du repères en procédant à la rotation de tous les points de l'objet par :

$$\begin{pmatrix} X_i \\ Y_i \\ Z_i \end{pmatrix} = \begin{pmatrix} V_{11} & V_{12} & V_{13} \\ V_{21} & V_{22} & V_{23} \\ V_{31} & V_{32} & V_{33} \end{pmatrix} \times \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} \quad (2.18)$$

où V_i sont les vecteurs propres. L'ordre des vecteurs propres est défini en fonction des valeurs propres. Le vecteur propre de plus grande valeur propre doit être aligné avec le première axe du repère x , le deuxième avec l'axe y et le troisième avec l'axe z .

Remarque : il est également possible de calculer les paramètres de normalisation, non plus sur le maillage de l'objet, mais sur le volume de l'objet. Ceci créer une normalisation plus robuste, qui est moins dépendante aux remaillages. La différence entre ces deux type de normalisation sera étudiée plus précisément dans le chapitre 6.

2.7 Critères d'évaluations

2.7.1 Base de données

Les différentes expérimentations réalisées au cours de cette thèse ont été principalement faites sur deux bases de données d'objets 3D :

- la base « Princeton Shape Benchmark » ;
- la base « Renault Semantic-3D ».

2.7.1.1 Base Princeton Shape Benchmark

La base Princeton Shape Benchmark (SHILANE ET AL., 2004) est une base de modèles 3D mise à disposition sur le Web par l'équipe du « Princeton Shape Retrieval and Analysis Group », pour que la communauté évalue leurs algorithmes d'indexation 3D sur une même base d'objets 3D. La base de données de Princeton contient 1814 modèles groupés dans les classes sémantiques de haut-niveaux où les objets d'une même classe sont hétérogènes. Par exemple, une classe des escaliers contient les modèles 3D, qui représentent des escaliers de formes très différentes mais avec la même sémantique (figure 2.18). La figure 2.16 montre des objets de cette base.

2.7.1.2 Base Renault Semantic-3D

Dans le cadre du projet Semantic-3D, les expérimentations seront conduites sur le corpus "Laguna", comportant 4933 objets, provenant des bases de données du constructeur industriel Renault. Une évaluation objective des performances des requêtes par similarité nécessite de disposer d'une catégorisation (vérité terrain) des objets 3D en classes. Une première vérité terrain, de 758 objets répartis en 75 classes, a été définie. La figure 2.17 montre des objets de cette base.

2.7.1.3 Taxinomie

La base de modèles Renault est une base de modèles techniques qui contient des pièces mécaniques. Les formes des modèles d'une même classe sont similaires. La base Princeton contient des classes hautement sémantiques où les objets d'une classe sont plus hétérogènes. La figure 2.18 montre deux exemples de classes de la base Renault et de la base Princeton. On peut voir que les objets d'une même classe Renault sont proches et correspondent à une même pièce selon différentes versions, alors que les classes de la base Princeton sont hautement sémantiques et les objets peuvent être de forme fortement dissemblable.

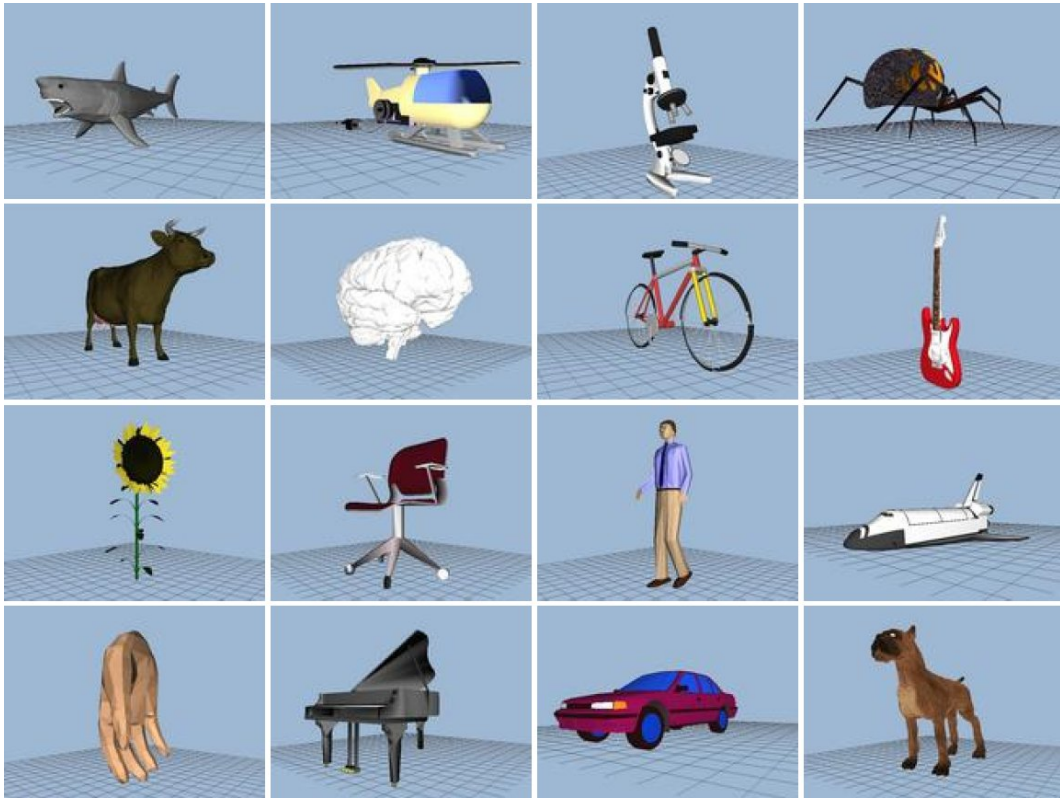


FIG. 2.16 – Objets de la base Princeton Shape Benchmark (SHILANE ET AL., 2004).

2.7.2 Critères de comparaison

La comparaison des différents descripteurs de forme implique de définir un ensemble de critères prenant en compte :

- les aspects algorithmiques associés,
- les performances au regard d'applications de requête par similarité dans des bases 3D génériques.

2.7.2.1 Critères algorithmiques

Les critères suivants ont été définis par le projet Semantic-3D comme critère commun de comparaison :

- coût de stockage, noté T , représentant la taille du descripteur pour chaque objet exprimée en nombre de composantes du vecteur de représentation ;
- complexité de l'algorithme d'extraction du descripteur à partir des données 3D, notée CE ,
- complexité de calcul de la mesure de similarité associée, notée CMS ;

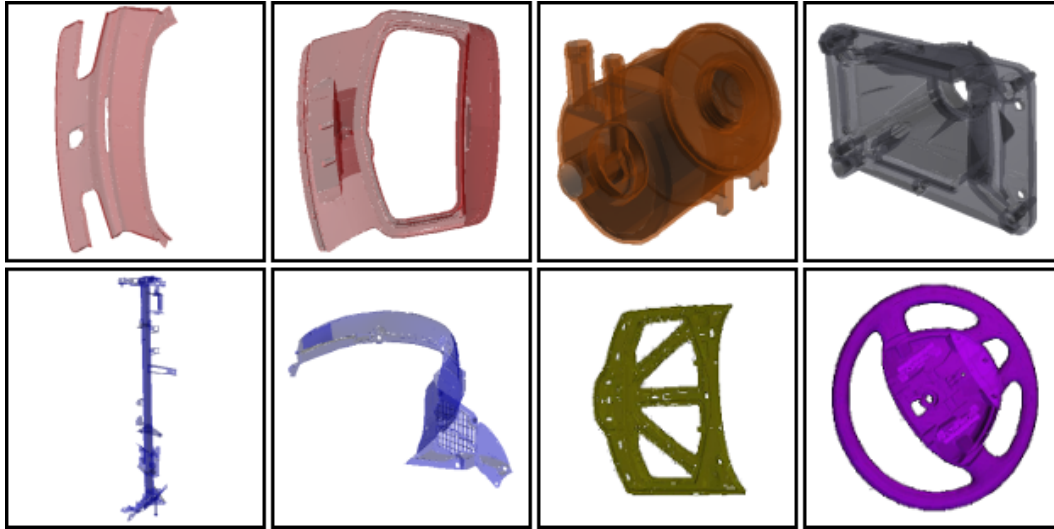


FIG. 2.17 – Objets de la base Renault Semantic-3D.

- généralité (si le descripteur s'applique ou non aux maillages de géométrie et topologie arbitraires), notée G ;
- invariance géométrique (IG) aux transformées de similarité et topologique (IT) aux représentations maillées multiples.

Un descripteur sera considéré comme invariant si et seulement si les conditions d'invariance sont respectées de façon intrinsèque sans recourir aux procédures de normalisation, pré-traitement, filtrage... Chaque descripteur de forme est représenté comme un vecteur $1D, 2D, \dots, nD$ composé respectivement d'un nombre $N_{1D}, N_{2D}, \dots, N_{nD}$ d'éléments (quantifiés ou non). Les valeurs de $N_{1D}, N_{2D}, \dots, N_{nD}$ classiquement associées (notées V_{type}) à chaque descripteur seront également indiquées. Un deuxième volet de cette comparaison concerne l'évaluation des performances des descripteurs dans le cadre des applications de requête par similarité.

2.7.2.2 Critères évaluant la pertinence des requêtes par similarités

Pour évaluer quantitativement et objectivement les performances des descripteurs de forme proposés, les deux paramètres suivants seront calculés :

- le First Tier (FT), proposé par (OSADA ET AL., 2001), ainsi que ;
- le Second Tiers (ST), ou le score Bull-Eye (SBE), retenu comme critère d'évaluation de tous les descripteurs de forme considérés dans MPEG-7.

Pour chaque maillage catégorisé, ces deux mesures sont définies par :

$$SBE = \frac{N_{cor}(2Q)}{Q} \quad FT = \frac{N_{cor}(Q)}{Q} \quad (2.19)$$

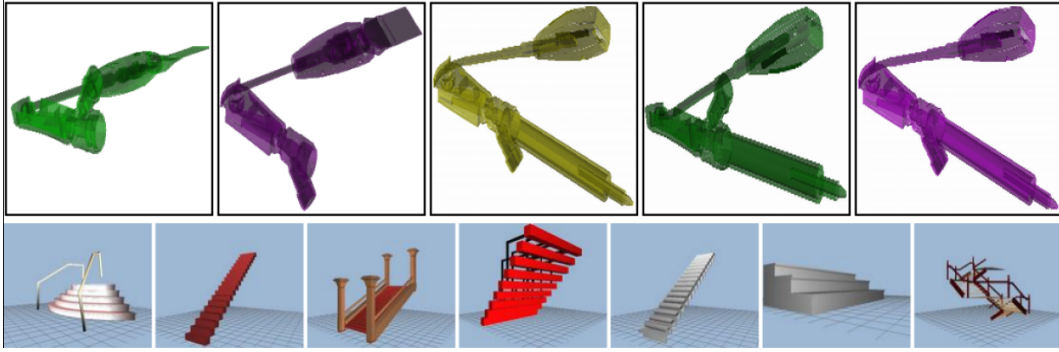


FIG. 2.18 – Exemple de classes de modèles 3D, première ligne : base Renault (classe : « attache de ceinture »), deuxième ligne : Princeton Shape Benchmark (classe : « escalier »).

où $N_{cor}(x)$ est le nombre de résultats corrects (i.e. maillage appartenant à la même catégorie que celle de la requête) parmi les x premiers résultats retrouvés, et Q est le nombre d'éléments de la catégorie correspondant à la requête. Si le SBE est relativement permissif, admettant une marge d'erreur de 50 % pour retrouver l'ensemble des modèles recherchés, le FT est beaucoup plus drastique, un FT de 100% n'autorisant aucune erreur.

Ces deux scores seront calculés pour tous les objets appartenant à une classe de la base. La moyenne des scores de chaque objet sera faite ce qui permet de définir le SBE_{tot} et FT_{tot} . Ces critères permettront d'analyser et de comparer rapidement les performances des différents descripteurs.

Pour une analyse plus fine des performances, une étude en terme de courbe rappel/précision, sera présentée. Deux critères sont utilisés pour calculer cette courbe : le *Recall* qui représente le nombre de réponses positives obtenues au rang k sur le nombre d'éléments de la classe de l'objet requête, et la *Precision* qui représente le nombre de réponses positives au rang k sur le nombre total de réponses obtenues. Ceci peut se noter de la façon suivante :

$$Recall(k) = \frac{N_{cor}(k)}{Q} \quad Precision(K) = \frac{N_{cor}(k)}{k} \quad (2.20)$$

Le Recall mesure la capacité du système à retrouver tous les documents qui sont pertinents, alors que la précision mesure la capacité du système à retrouver seulement les documents qui sont pertinents. Ces deux critères sont calculés pour tous les rangs de la base et sont affichés sous forme de courbes : la courbe de Recall, la courbe de Precision et la courbe de Recall/Precision. Cette dernière est couramment utilisée car elle représente de façon visuelle les informations données par le Recall et la Precision. Les FT et les SBE sont liés aux courbes de *Recall*. La valeur du FT correspond à la valeur de la courbe de *Recall* à l'abscisse du cardinal de la classe

et respectivement à deux fois le cardinal de la classe pour le SBE.

Pour l'indexation des modèles 3D à partir de vues, le protocole d'expérimentation sera le suivant. Pour chaque objet 3D, on génère 10 vues aléatoires que l'on définit comme vue requête. Ces vues seront prises aléatoirement et uniformément réparties autour du modèle 3D. Chacune de ces vues sera une requête au système. L'indexation 2D/3D utilisera les mêmes critères d'évaluation que ceux de l'indexation 3D/3D (FT, SBE et courbes Rappel-Précision). Le score par modèle sera la moyenne des scores pour les 10 vues requêtes. Les scores SBE et FT seront calculés pour chaque élément catégorisé. Des scores moyens pour chaque catégorie seront ensuite estimés. Enfin, le score global moyen sur l'ensemble des catégories sera calculé. Afin d'étudier la scalabilité, le même protocole suivi pour l'indexation 3D/3D sera utilisé dans le cas d'indexation 3D à partir de vues.

2.8 Conclusion

Dans ce chapitre, nous avons présenté les principales méthodes d'indexation d'objets 3D. Une étude plus précise, de laquelle est inspirée cette partie, a été réalisée dans le cadre du projet RNRT Semantic-3D (BASKURT ET AL., 2003). Cet état de l'art nous a permis d'avoir une vision claire du domaine de recherche et des méthodes proposées et est à la base du travail que nous avons réalisé.

Cette étude approfondie nous conduit à formuler quelques remarques sur les différentes approches 2D et 3D en vue d'en extraire des pistes originales de recherche que nous souhaitons approfondir.

Premièrement, un grand nombre de méthodes proposées sont dépendants du type d'objets à traiter et aux caractéristiques de l'application visée. Comme nous l'avons montré dans le paragraphe 2.3, les méthodes structurelles permettent de comparer les structures des objets et de classifier comme similaire des objets de formes différentes mais de structures proches. Ces méthodes ont été parmi les premières approches évaluées car elles sont intuitives et créent des graphes de description compréhensibles. Par contre, elles ne peuvent pas être appliquées sur des grandes bases d'objets généralistes. En effet, des problèmes existent lors de la phase de segmentation pour extraire les représentations. Ceci est particulièrement vrai pour les méthodes représentant les objets 3D par une famille de formes, comme les Géons, qui sont peu robustes à l'apparition de bruit sur le maillage des objets à segmenter. Ce problème est atténué par les méthodes extrayant la structure par des fonctions mathématiques, comme les axes médians ou les graphes de Reeb, mais n'est toujours pas applicable sur tous les objets, en particulier les objets surfaciques. D'autre part, ces méthodes ne permettent pas d'être exécutées sur des grandes bases d'objets (plusieurs milliers d'objets), car la mesure de similarité entre graphes « coûte » en temps de calcul.

Pour ces raisons, les études se sont portées sur des méthodes moins haut-niveau,

comme les méthodes par transformée ou les méthodes statistiques. Ces méthodes sont dites bas-niveau, car elles extraient des caractéristiques des objets par des fonctions mathématiques, et ne permettent pas de faire un lien intuitif entre les descripteurs et la forme des objets. Par contre, elles ont un coût de calcul de descripteur et de mesure de similarités faibles, ce qui permet de les employer sur des grandes bases et de garantir des temps de réponses courts. Cependant, elles ne comparent pas la structure des objets et se limitent à une comparaison de similarité de forme proche. La forme de l'objet est caractérisée par un vecteur de descripteurs, difficilement interprétable mais rapide à calculer et à comparer.

Les méthodes à base de vues sont à mi-chemin entre les deux approches précédentes. Elles extraient l'information bas-niveau des images et comparent leurs ensembles de vues lors du calcul de la mesure de similarité, ce qui peut être fait en temps linéaire. Cette description est bas-niveau lors de l'extraction de l'information des images mais le nombre et la position des images fournissent une représentation de plus haut-niveau, ce qui permet de comprendre les similarités au niveau des images. Nous pouvons donc extraire plus d'information que les méthodes par transformées, en ne s'intéressant pas uniquement à la forme globale des objets et aux ressemblances des points de vues.

Le choix de telle ou telle méthode doit être dicté par le type d'applications envisagées, et par les performances et les coûts escomptés. Dans le cadre de ces travaux et du projet RNRT Semantic-3D, nous devons indexer des grandes bases d'objets mécaniques de mauvaise qualité (trous, surfacique, plusieurs parties, ...), ce qui ne nous a pas permis de mettre en place des approches structurales. D'autre part, il fallait pour les scénarios d'usages, définis par le projet et par notre partenaire Renault, mettre en place des processus de recherche d'objets par des requêtes 2D et 3D.

Pour ces différentes raisons, nous avons opté pour l'étude et le développement d'une méthode d'indexation basée sur les moments et d'une méthode d'indexation 2D/3D. Le fait de devoir indexer des grandes bases d'objets 3D et les résultats de la méthode ART sur les images 2D, nous ont poussé à mettre en place un nouveau descripteur de forme 3D par transformée en généralisant la transformation ART en 3D. D'autre part, aux vues des limitations des méthodes d'indexation d'objets 3D par vues et des problèmes dans le choix du types et du nombres de vues, nous avons mis en place une méthode dynamique de comparaison d'un objet 3D à une image 2D. La méthode se base sur un système de rendu spectral permettant la création rapide de vues de l'objet, et nous l'avons appliquée, à la fois à la recherche d'objets 3D à partir d'une requête 2D mais aussi à la recherche à partir d'un objet 3D.

Deuxième partie

Indexation par requêtes 2D

3

Indexation 2D

Sommaire

3.1	Introduction	66
3.2	Généralisation aux images couleurs	66
3.2.1	Analyse couleurs	66
3.2.1.1	Segmentation des images couleurs	67
3.2.1.2	Indexation des objets par des paramètres couleur	70
3.2.1.3	Mesure de similarité entre objets couleur	70
3.2.2	Analyse de forme GART	71
3.2.2.1	Application de la transformation ART aux images niveaux de gris	71
3.2.2.2	Généralisation de l'ART aux déformations homo- graphiques	73
	Création des fonctions de bases déformées.	75
	Inversion du processus.	76
3.2.3	Combinaison des recherches par la luminance et par la chro- minance	78
3.3	Expérimentations	79
3.3.1	Recherche d'objets déformés par des transformations non planes	79
3.3.2	Recherche de logos dans une base d'images naturelles	81
3.3.2.1	Extraction d'un logo d'une image naturelle	83
3.3.2.2	Résultats	84
3.4	Conclusion	85

3.1 Introduction

La deuxième partie de cette thèse est consacrée aux études que nous avons menées en indexation d'images 2D, et particulièrement aux généralisations que nous avons apportées aux descripteurs de formes par approche région de la norme MPEG-7 : Angular Radial Transform (ART).

Depuis une vingtaine d'années, l'indexation d'images couleurs est un thème de recherche important. De nombreux travaux ont été proposés sur la recherche d'images similaires et se basant sur la description couleur. Comme nous l'avons vu dans le chapitre 1, la transformation ART possède des propriétés d'invariance et de robustesse et une complexité réduite, qui en font un descripteur de formes très efficace. Par contre, on peut noter qu'elle n'est pas définie pour s'appliquer sur les images couleur et qu'elle n'est pas robuste aux déformations homographiques. Pour répondre à ces deux attentes, nous nous proposons de mettre en place deux évolutions pour permettre à la transformation ART d'être utilisée sur un grand nombre de documents multimédia. Ces évolutions sont :

- l'extension du domaine de définition de la transformation aux images niveaux de gris et couleurs ;
- la généralisation de la transformation ART aux rotations non planes, qui modifient la forme de l'objet lors de la projection de la scène sur le plan image.

Ce chapitre présente ces deux évolutions et le processus que nous avons mis en place pour indexer les régions issues d'images naturelles. Nous présentons également l'application qui a guidé ces travaux, qui consiste à indexer des logos extraits d'images (cf. paragraphe 3.3.2).

3.2 Généralisation aux images couleurs

Le processus que nous avons mis en place caractérise les formes issues d'images naturelles. Pour extraire le maximum d'information des régions, nous avons opté pour la mise en place de deux analyses simultanées : l'une caractérise les régions en fonctions de leurs couleurs, l'autre s'attache à la forme. Ceci fournit deux classements différents qui seront combinés en pondérant l'influence des deux caractéristiques (cf. paragraphe 3.2.3).

3.2.1 Analyse couleurs

L'étude de la couleur est basée sur les travaux de FLICKNER ET AL. (1995), PENTLAND ET AL. (1996), SALEMBIER et MANJUNATH (2000) et plus précisément sur les travaux de IDRISSE ET AL. (2002, 2004). La recherche de similarité se fait en deux étapes. Dans un premier temps, la base d'images est indexée, en extrayant les N premières couleurs dominantes (voir ci-dessous). Pour chaque classe couleur,

nous stockons plusieurs paramètres comme la couleur dominante, le pourcentage et la variance dans l'espace couleurs CIE $L^*a^*b^*$. Ce sont ces paramètres qui permettront de mesurer la similarité entre un objet clé et les objets de la base. L'étude se divise en trois étapes :

1. La segmentation des images couleurs ;
2. L'indexation des objets segmentés avec les couleurs dominantes ;
3. La recherche de similarité entre objets par une mesure adaptée.

3.2.1.1 Segmentation des images couleurs

L'approche proposée pour détecter les couleurs dominantes est basée sur un seuillage itératif de l'histogramme couleur obtenu par minimisation de l'entropie croisée (MEC). Comme énoncé par IDRISSI ET AL. (2004), l'Entropie Croisée est une mesure que l'on effectue entre deux distributions de probabilité, permettant de quantifier la distance et le degré de ressemblance (ou de dissemblance), qui existe entre ces deux distributions. Elle est donnée pour deux distributions P et Q par :

$$E(P, Q) = \sum_{k=1}^N q_k \log_2 \left(\frac{q_k}{p_k} \right) \quad (3.1)$$

où $\{p_k\}$ et $\{q_k\}$ sont les N valeurs de l'histogramme.

En traitement d'images, cette technique peut être utilisée pour seuiller une image à niveaux de gris, l'image étant modélisée par sa distribution de probabilité p (histogramme normalisé). On démontre que l'entropie croisée est proportionnelle à :

$$\eta(t) = \sum_{j=1}^{t-1} j h_j \ln \left(\frac{j}{\mu_1(t)} \right) + \sum_{j=t}^L j h_j \ln \left(\frac{j}{\mu_2(t)} \right) \quad (3.2)$$

où t représente le niveau de seuil choisi, L le nombre de niveaux de gris présents dans l'image, h_j le nombre de pixels ayant le niveau de gris j , μ_1 et μ_2 , le niveau de gris attribué respectivement à la classe 1 et à la classe 2 de l'image seuillée, μ_1 et μ_2 représentent les valeurs moyennes des niveaux de gris des pixels appartenant respectivement aux deux classes. La minimisation de $\eta(t)$ permet de trouver le niveau optimal L_{opt} séparant l'image en deux classes, comme il est montré sur la figure 3.1.

Une extension de cette méthode à l'utilisation de 2 (ou N) variables est possible, comme énoncé par LI et LEE (1993). En effet, si nous considérons une image représentant une nouvelle variable M , dont les pixels sont calculés à partir d'une combinaison de 2 (ou N) variables, alors l'application de la MEC à cette nouvelle image permettra d'obtenir un seuillage où il a été tenu compte de 2 (ou N) variables, en fonction de la combinaison choisie (linéaire, pondérée, quadratique, etc.). La Figure 3.2 donne le partitionnement du plan (x_1, x_2) pour $M = x_1 + x_2$ (Figure 3.2.a) et pour $M = \sqrt{x_1^2 + x_2^2}$ (Fig. 3.2.b).

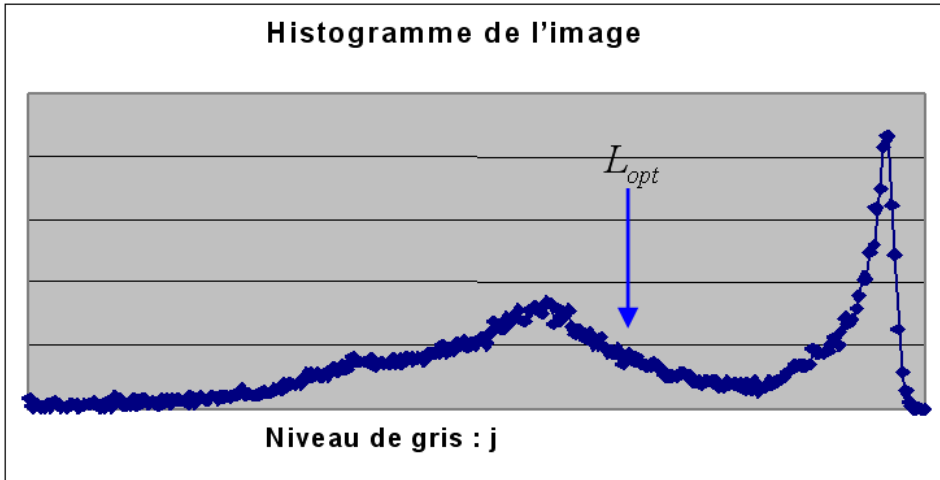


FIG. 3.1 – Détection du seuil optimal par MEC.

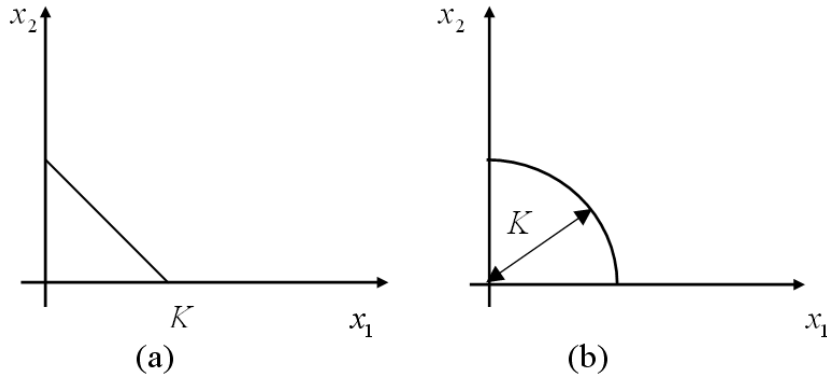


FIG. 3.2 – Partitionnement du plan (x_1, x_2) par minimisation d'entropie croisée.

Dans les deux cas, la MEC détecte le seuil optimal K partitionnant le plan en deux sous plans vérifiant respectivement $M < K$ et $M > K$. Pour appliquer cette méthode à la quantification couleur, nous calculons, à chaque itération, l'histogramme CIE $L^*a^*b^*$ 3D de la classe à décomposer et en déduisons le maximum appelé centroïde, désigné par $(O_{k,1}, O_{k,2}, O_{k,3})$ et représentant en fait la couleur dominante de l'image à ce stade de décomposition (l'indice k est relatif à la k^{me} itération). Ceci conduit à une image "distance" définie par :

$$I_k(m, n) = \sqrt{\sum_{i=1}^3 (C_i - O_{k,i})^2} \quad (3.3)$$

où $I_k(m, n)$ représente la distance couleur entre la couleur du pixel (m, n) et la couleur du maximum d'histogramme détecté $(O_{k,1}, O_{k,2}, O_{k,3})$ et où C_i est la i^{me}

composante couleur du pixel (m, n) . Le choix de l'espace de travail CIE $L^*a^*b^*$ autorise l'utilisation de la distance Euclidienne.

Le seuillage de $I_k(m, n)$ par entropie croisée fournit le seuil optimal qui permet de séparer l'image distance en deux classes : celle des pixels dont la couleur est proche du centroïde, et donc de la couleur dominante, et celle des pixels dont la couleur en est éloignée. La méthode peut alors être exécutée itérativement sur chacune des classes trouvées, décomposant chaque classe en deux nouvelles classes. Il est ainsi possible de respecter le nombre de classes imposé initialement.

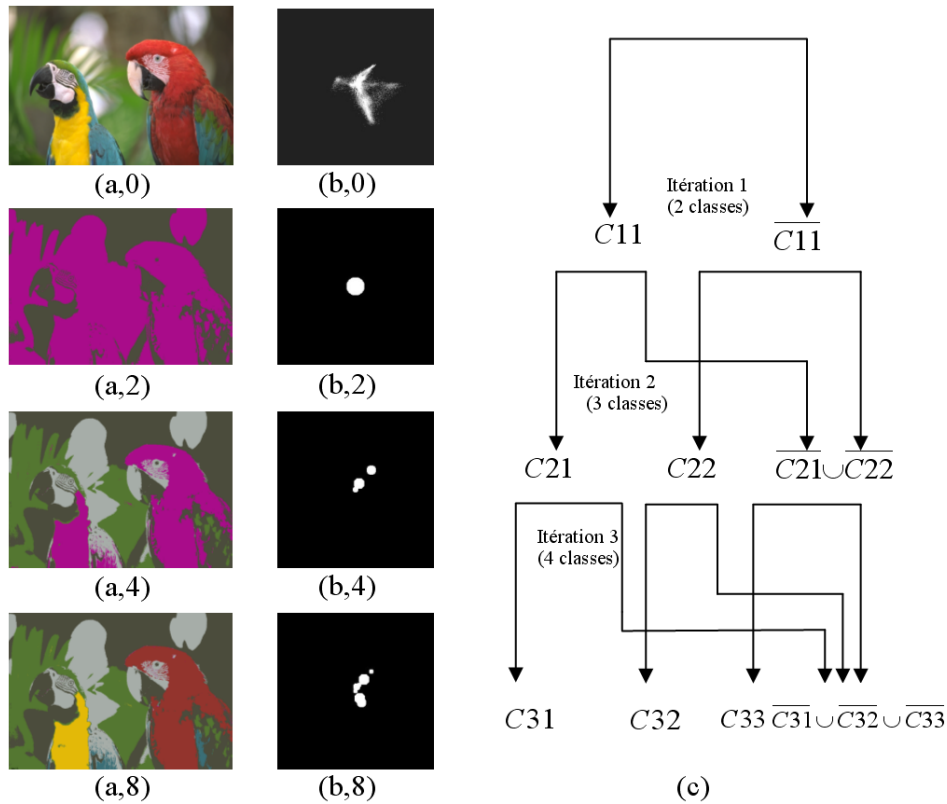


FIG. 3.3 – Segmentation d’une image couleur par partitionnement itératif de l’espace couleur CIE $L^*a^*b^*$.

La figure 3.3 illustre la méthode de segmentation : à partir de l’image originale $(a, 0)$, l’histogramme couleur 3D est calculé et son maximum déterminé. Pour la lisibilité de la figure, seule une projection de l’histogramme est affichée $(b, 0)$. Ensuite l’image distance est calculée, puis seuillée par MEC, donnant ainsi naissance à deux classes C_{11} et $\overline{C_{11}}$. Ce même processus est appliqué itérativement à chacune des classes trouvées, et à chaque étape, les classes correspondant aux couleurs non similaires sont fusionnées dans une même classe. Les figures 3.3.(a.4) et (a.8) montrent les images segmentées respectivement au bout de 4 et 8 itérations. La figure 3.3.c

montre le découpage successif des classes pour 3 itérations.

3.2.1.2 Indexation des objets par des paramètres couleur

L'indexation couleur est directement issue de la segmentation couleur réalisée précédemment. Chaque classe couleur i est caractérisée par 3 paramètres : la couleur dominante CD_i que nous avons appelée Centroïde lors de la segmentation, le pourcentage p_i que représente la classe par rapport à l'image et la variance V_i^2 de la couleur dans la classe. Chaque image est représentée par son descripteur qui est constitué de N triplets (CD_i, p_i, V_i^2) , N étant le nombre de couleurs dominantes détectées dans l'image.

Dans ce travail, nous nous intéressons aux objets constituant l'image. La segmentation et l'indexation vont donc travailler sur l'image de l'objet où les pixels qui n'appartiennent pas à la zone de l'image le représentant sont ignorés.

3.2.1.3 Mesure de similarité entre objets couleur

Notre point de départ étant les descripteurs couleur des images de la base, nous proposons d'utiliser l'histogramme pour évaluer la similarité couleur entre images. En supposant que l'histogramme couleur peut être modélisé par un mélange de distributions gaussiennes, un nouvel histogramme est créé à partir des N couleurs dominantes fournies par le descripteur de l'image et intégrant les 3 paramètres qui constituent ce descripteur (couleur dominante, pourcentage et variance), et en combinant N distributions gaussiennes. La contribution de la couleur dominante est donnée par :

$$H_i(x) = \frac{p_i}{V_i \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x-CD_i)^2}{2 \cdot V_i^2}} \quad (3.4)$$

où CD_i est la i^{me} couleur dominante, p_i son pourcentage et V_i^2 sa variance. La figure 3.4 montre un exemple d'histogramme 1D créé.

Le calcul de la distance inter descripteurs couleur consiste à recréer les histogrammes 1D en fonction des paramètres des couleurs dominantes et à calculer la distance inter histogramme. La distance couleur que nous proposons est issue de la distance de KULLBACK (1959) dans sa forme symétrique, donnée par :

$$d_c(R, I) = \sum_{n=1}^N \sum_{m=1}^3 (r_{n,m} - i_{n,m}) \log_2 \left(\frac{r_{n,m} + 1}{i_{n,m} + 1} \right) \quad (3.5)$$

où N est le nombre d'éléments de l'histogramme de chaque composante couleur (classiquement $N = 256$), et $r_{n,m}$ et $i_{n,m}$ sont les pourcentages de la composante m de la couleur n , respectivement de l'image clef et d'une image de la base.

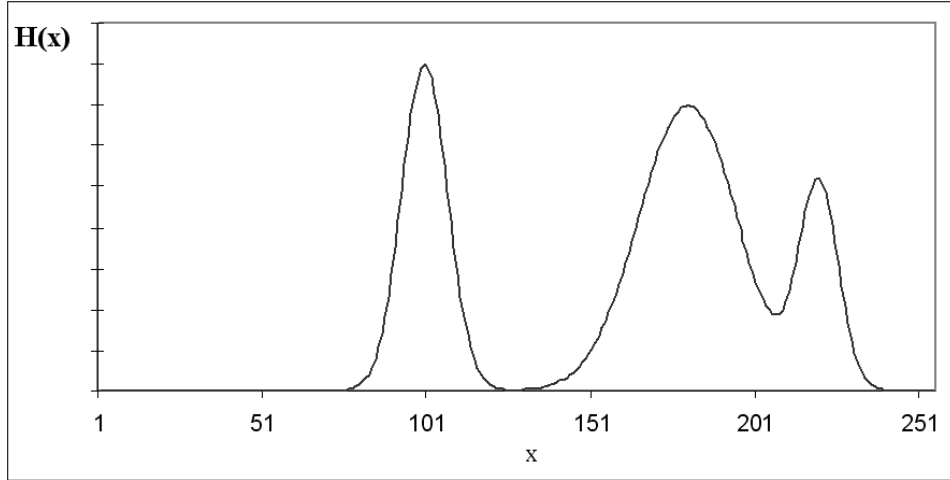


FIG. 3.4 – Exemple d’histogramme 1D sur une des composantes CIE L*a*b* générée à partir de 3 couleurs dominantes.

3.2.2 Analyse de forme GART

La seconde analyse porte sur la forme des régions. Comme nous l’avons déjà énoncé dans le paragraphe 1.2.3.3, la transformation ART est appliquée sur les régions binaires et est peu robuste aux rotations non planes. Pour augmenter les performances d’ART nous avons mis en place une généralisation de ART (GART), portant sur deux points : l’application d’ART sur le canal luminance de la régions à décrire et le fait de rendre ART robuste aux rotations non planes.

3.2.2.1 Application de la transformation ART aux images niveaux de gris

Rappelons que la transformation ART, *Angular Radial Transform*, est une transformation 2D complexe, définie sur un disque unité, en coordonnées polaires avec la formule :

$$F_{nm} = \int_0^{2\pi} \int_0^1 V_{nm}(\rho, \theta) f(\rho, \theta) \rho d\rho d\theta \quad (3.6)$$

où F_{nm} est un coefficient d’ART d’ordre n et m , $f(\rho, \theta)$ est une fonction image définie en coordonnées polaires, et $V_{nm}(\rho, \theta)$ est une fonction de base d’ART qui est séparable dans la direction angulaire et radiale :

$$V_{nm}(\rho, \theta) = A_m(\theta) R_n(\rho) \quad \text{où} \quad \begin{cases} A_m(\theta) = \frac{1}{2\pi} e^{jm\theta} \\ R_n(\rho) = \begin{cases} 1 & n = 0 \\ 2 \cos(\pi n \rho) & n \neq 0 \end{cases} \end{cases} \quad (3.7)$$

La transformation ART, telle qu’elle a été proposée par le comité MPEG-7, est

une transformation dont le domaine d'application est limité aux images binaires, bien qu'il n'y ait aucune limitation dans la définition de la transformation. L'image binaire permet de prendre en compte le contour de l'objet. Pour décrire des objets complexes, il faut représenter les variations internes de l'objet, qui peuvent être dues à la présence d'arrêtes ou de contours internes. Pour décrire la forme d'un objet complexe, il ne faut pas s'intéresser à l'image binaire de ces contours mais travailler sur l'image niveaux de gris représentant les variations internes, comme il est suggéré par FANG et QIU (2003) et AKCAY ET AL. (2002). Dans ce cas, on considère $f : (\rho, \theta) \rightarrow [0, 1]$.

Le canal luminance représente l'aspect global de la forme de l'objet. Nous avons comparé les deux approches qui sont : décrire l'objet par une image binaire intérieur/extérieur (ART classique), ou décrire l'objet par son image luminance. Pour créer l'image luminance de l'objet, nous prenons en compte les pixels intérieurs de l'objet et remplaçons les triplés RGB par la valeur de luminance correspondante, dans l'espace couleur CIE $L^*a^*b^*$. Pour comparer ces deux approches, nous avons défini P classes d'objets similaires, en créant une base d'images composées de P objets. Pour comparer les deux méthodes, nous cherchons à identifier celle qui permettra de trouver le plus rapidement possible les autres membres de la classe à laquelle l'image clé appartient. La figure 3.5 montre ces résultats en terme de courbe Recall/Precision.

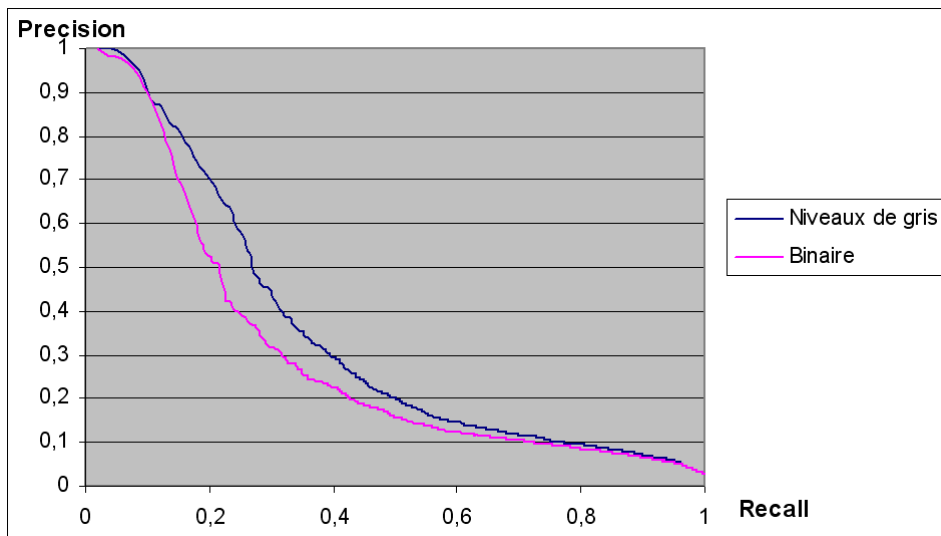


FIG. 3.5 – Courbes Recall/Precision obtenue par ART binaire et ART niveaux de gris.

Ceci montre que la description d'une forme par son image luminance fournit de meilleurs résultats que la description par son image binaire intérieur/extérieur. Dans la suite de ce travail, sur les images 2D, l'ART sera appliquée sur la composante

luminance des images.

3.2.2.2 Généralisation de l'ART aux déformations homographiques

Comme nous l'avons déjà énoncé dans la partie 1.2.3.3, la transformation ART possède des propriétés d'invariance en rotation plane et aux variations d'échelles, ce qui garantit une forte similarité entre les coefficients de deux images d'un même objet ainsi déformé. Une autre transformation courante subie par les objets, à l'intérieur d'une image, est la rotation non plane. Un objet plan vu selon deux angles différents n'aura pas la même forme et ne sera pas détecté comme similaire par la méthode ART.

Un objet plan, ayant subi une rotation 3D, peut être perçu comme une image projetée sur un plan qui a subi une série de rotations. La rotation du plan support de l'objet à détecter entraîne des déformations de la forme de l'objet, lors de sa projection sur le plan image. Par l'application des phénomènes de perspective, la partie proche de l'objet semble plus grande que la partie la plus éloignée et les bords ne sont plus parallèles. L'objet subit de fortes déformations et n'est plus classé comme similaire par la méthode ART classique, comme le montre la figure 3.6.a. Pour permettre à la transformation ART de détecter comme invariant un objet vu selon deux angles, il faut faire subir, aux fonctions de base, la même série de rotations que l'objet à détecter, ce qui revient à projeter les fonctions de base sur le plan support de l'objet déformé. Les applications visées par cette généralisation sont la détection de logos plans dans des images naturelles comme nous l'expliquerons plus précisément dans la partie 3.3.

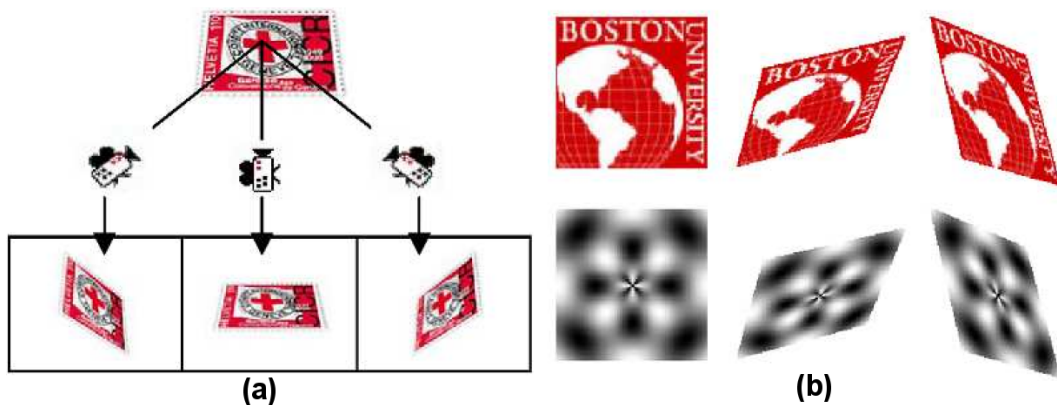


FIG. 3.6 – Exemple d'objets plans vus selon différents angles (a), exemple de fonctions de base d'ART projetées sur le plan support de l'objet à identifier (b).

Si l'on fait subir aux fonctions de base (FDB) les mêmes perturbations que celles subies par l'image déformée, les coefficients ART calculés, entre l'image d'origine et les fonctions de base d'origine et entre l'image déformée et les fonctions de base

projetées sur le plan support de l'objet, seront identiques. Les expressions suivantes montrent cette identité :

$$\forall \begin{pmatrix} i' \\ j' \end{pmatrix} = M \begin{pmatrix} i \\ j \end{pmatrix}, \begin{cases} V_1(i', j') = V_1 \left(M \begin{pmatrix} i \\ j \end{pmatrix} (0), M \begin{pmatrix} i \\ j \end{pmatrix} (1) \right) = V_0(i, j) \\ F_1(i', j') = F_1 \left(M \begin{pmatrix} i \\ j \end{pmatrix} (0), M \begin{pmatrix} i \\ j \end{pmatrix} (1) \right) = F_0(i, j) \end{cases} \quad (3.8)$$

Ce qui permet d'obtenir : $F_0(i, j).V_0(i, j) = F_1(i', j').V_1(i', j')$, où $F_0(i, j)$ représente le pixel (i, j) de l'image d'origine, $V_0(i, j)$ représente la fonction de base d'origine au point (i, j) , $F_1(i', j')$ l'image de l'objet projeté suivant la transformation M et $V_1(i', j')$ la fonction de base projetée suivant la transformation M . On montre que les coefficients ART trouvés, lors du calcul, entre l'image d'origine et les fonctions de base d'origine et entre l'image projetée et les fonctions de base projetées, seront les mêmes, comme le montre la figure 3.7.

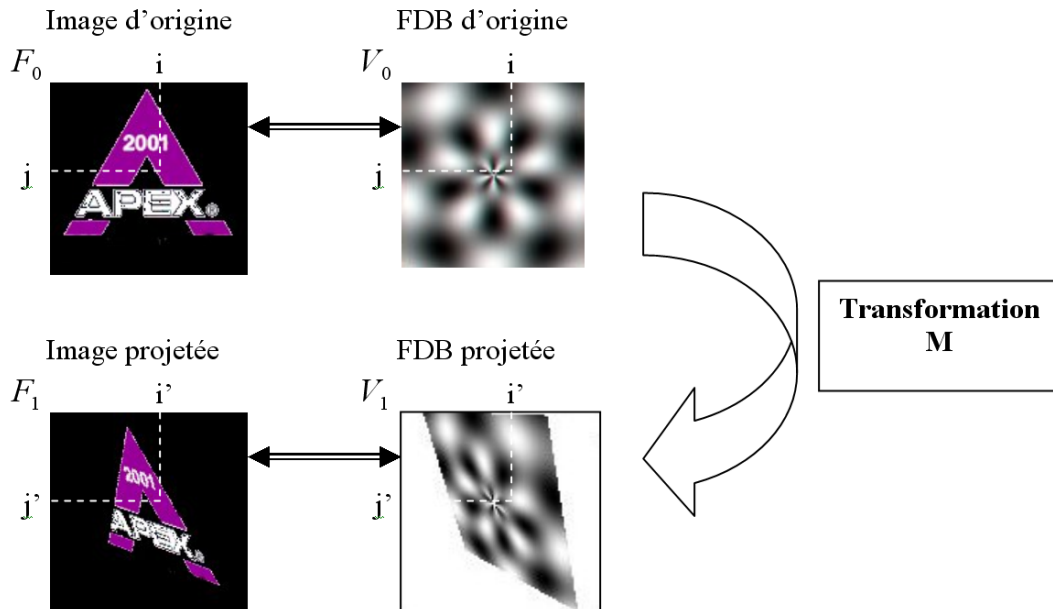


FIG. 3.7 – Les coefficients ART obtenus en projetant l'image d'origine sur les FDB d'origine sont les mêmes qu'en projetant l'image projetée sur les FDB ayant subi la même projection.

Il n'est pas possible de détecter les coordonnées du plan support d'un objet quelconque, afin de déformer, au mieux, les fonctions de base. Pour détecter le plus grand nombre d'objets, nous avons décidé de partitionner l'espace des déformations en plans de projection et de projeter les fonctions de base sur chacun de ces plans.

Cette généralisation de la transformation ART (GCART) va permettre de détecter un même objet vu suivant différents profils. ART devient robuste aux rotations

non planes et aux effets de perspective. Par contre, lors de la rotation d'un objet 3D non plan, l'apparition et l'occultation des bords de l'objet provoquent une forte variation du profil, ce qui ne sera pas détecté par ART généralisé. GCART « General Color Angular Radial Transform » permet de détecter un objet tourné dans la limite d'un faible changement de profil. Cette méthode s'applique particulièrement aux objets plans.

Création des fonctions de bases déformées. Pour définir les transformations subies par l'objet durant les rotations et la projection perspective sur le plan image, nous considérons l'espace des transformations défini par le coefficient de perspective p et la vecteur normal au plan image, noté \vec{n} . Le coefficient de perspective p est défini comme la distance entre la fonction de base d'origine et le plan image. Le vecteur normal \vec{n} est constitué par les angles d'Euler θ (direction radial) et ϕ (angle de rotation) définis par entre le vecteur \vec{n} et l'axe du repère \vec{x} , comme le montre la figure 3.8. Les deux premiers paramètres définissent l'orientation de la normale du plan image et le troisième, le coefficient de perspective, définit les déformations homographiques subies par l'objet lors de sa projection sur le plan image. La figure 3.8) montre ces trois paramètres.

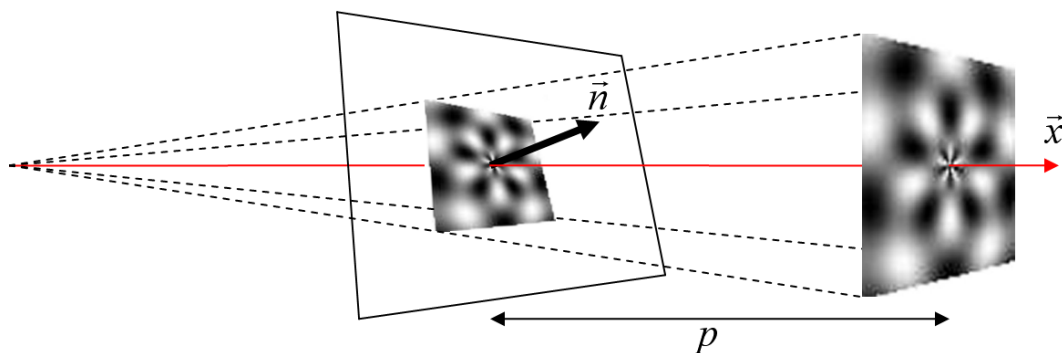


FIG. 3.8 – Les fonctions de base sont projetées sur le plan image I selon les paramètres $\vec{n}(\theta, \phi)$ et p pour obtenir les fonctions de base projetées.

Cet espace de transformation est échantillonné suivant chacun des paramètres selon les valeurs k_θ , k_ϕ et k_p . Nous obtenons ainsi : $K = k_\theta * k_\phi * k_p$ transformations. Les fonctions de base sont déformées selon ces K transformations, pour obtenir K séries de fonctions de base projetées. Chaque objet est indexé avec ces K séries de fonctions de base. Le nombre de projections K est limité pour garder un coût d'exécution relativement faible et les valeurs $k_\theta = 12$, $k_\phi = 3$ et $k_p = 3$ ont été choisies par notre protocole d'expérimentation, présenté dans le paragraphe 3.3, car elles fournissent le meilleur ratio coût/efficacité. L'échantillonnage de l'espace des transformations a été réalisé en pratiquant une discretisation uniforme sur les paramètres θ , ϕ et ρ . Une étude plus approfondie de l'espace des transformations permettrait de réaliser une discretisation uniforme au sein de cette espace, ce qui

permettrait sans doute d'améliorer les résultats. Cette étude n'a pas été réalisée durant les recherches qui ont été mise en place, mais cela fait partie des perspectives qu'ils seraient intéressant d'étudier.

En d'autres termes, nous décrivons les formes des objets par $K = 108$ séries de coefficients, et nous devons calculer 108 mesures de similarités entre l'objet requête et les objets de la base. La complexité de ART classique est en $O(n * m * N^2)$ car on calcule $n * m$ fonctions de base pour $N \times N$ pixels de l'image. La généralisation de ART aux déformations homographiques, qui crée K séries de coefficients a une complexité en $O(K * n * m * N^2)$. La figure 3.9 montre les 108 séries de fonctions de base créées.

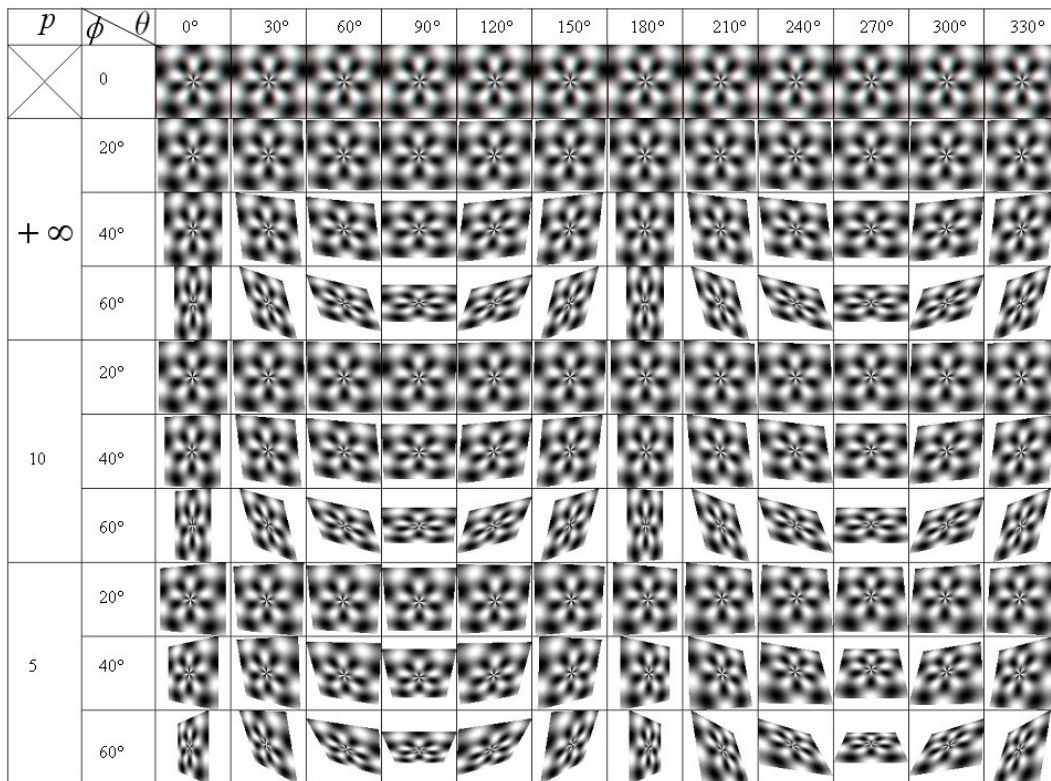


FIG. 3.9 – Exemple de fonctions de base projetées en respectant la perspective.

Inversion du processus. Pour rendre le processus de recherche plus rapide, nous avons choisi d'inverser les processus d'indexation et de recherche. Sans optimisation, le processus d'indexation calcule un descripteur ART entre l'objet d'origine et les fonctions de base d'origines, alors que durant la recherche, 108 descripteurs ART sont calculés entre l'objet extrait d'une image naturelle et toutes les séries de fonctions de base projetées. Le processus d'indexation qui est hors ligne, a un coût K fois moindre que le processus de recherche qui est en ligne.

Heureusement, il est possible d'inverser ces deux processus et d'indexer l'objet d'origine sur les K fonctions de base projetées inverses alors que l'objet extrait lors de la recherche, sera indexé seulement sur les fonctions de base d'origine. Ceci augmente le coût du processus d'indexation hors ligne mais diminue le processus de recherche en ligne sans modifier les propriétés de la descriptions (Table 3.1).

	Processus d'origine	Processus optimisé
En ligne	K	1
hors ligne	1	K

TAB. 3.1 – Nombre de descripteurs ART calculé, en ligne et hors ligne, durant le processus d'origine et le processus optimisé.

Comme il est montré sur la figure 3.10, l'image d'origine F_0 est transformée en l'image projetée F_1 par la transformation T et les fonctions de base d'origine V_0 sont transformées en fonctions de bases projetées inverses V_{-1} par la transformation T^{-1} . On peut voir que l'on obtient les mêmes valeurs de descripteurs, si l'on indexe l'image déformée F_1 sur les fonctions de base d'origine V_0 , ou si l'on indexe l'image d'origine sur F_0 sur les fonctions de base inverse V_{-1} .

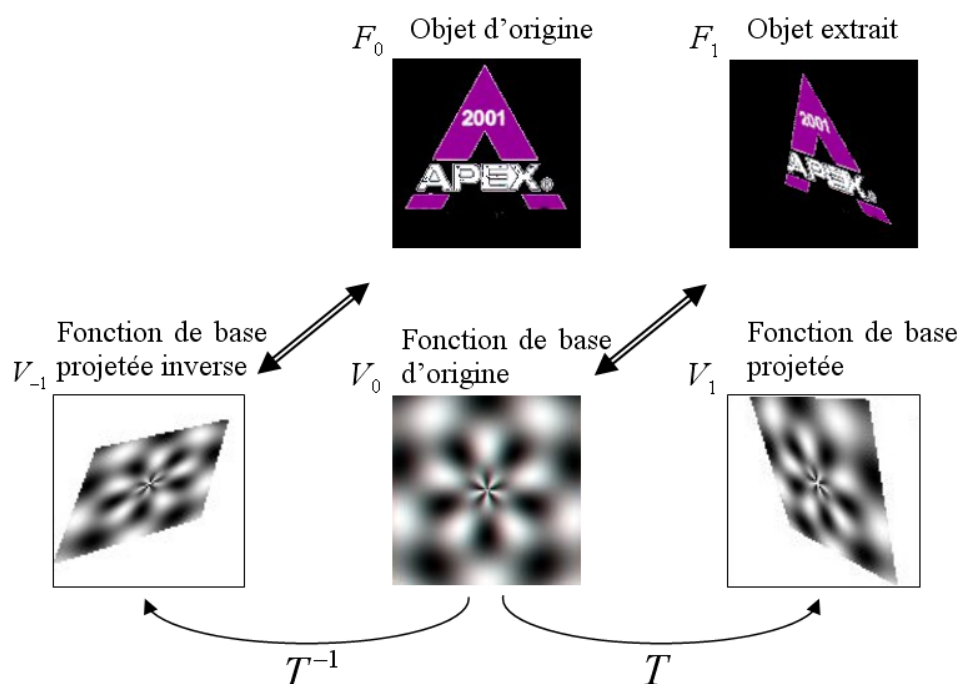


FIG. 3.10 – Schéma du processus d'indexation inverse.

Chaque objet est décrit par $K = k_\theta * k_\phi * k_p$ séries de coefficients ART, créées à partir des fonctions de base projetées sur les K plans de projection. La distance de

similarité entre formes est faite en calculant l'ensemble des distances $d_{ART}(Q, I_j)$, distance ART entre la requête et le descripteurs ART obtenue sur la série de fonctions de base j , $j \in K$. Pour toutes les valeurs de j , les coefficients ART de Q , image requête calculée sur les fonctions de base d'origine, et ceux de I , image d'origine calculée sur les fonctions de base projetées inverses, sont comparés en utilisant l'équation (1.12).

Alors, la distance de forme entre Q et I est donnée par :

$$d_{shape}(Q, I) = \min_{j \in K} \sum_{i=0}^{n \cdot m} \left\| ART_Q[i] - ART_I^j[i] \right\| \quad (3.9)$$

où ART_Q sont les coefficients ART de l'objet requête Q et ART_I^j sont les coefficients de l'objet I , calculé sur l'ensemble de fonctions de base projetées inverses j .

3.2.3 Combinaison des recherches par la luminance et par la chrominance

Les recherches, couleur (chrominance) et forme (luminance), se font en parallèle ce qui fournit deux classements. L'inter-classement permet de prendre en compte les deux critères et, en particulier, de garder le meilleur des deux distances. En préalable à la classification, il faut normer les deux répartitions. La normalisation peut être effectuée de différentes façons. Nous utilisons l'approche énoncée par KULLBACK (1959) qui se sert de la normalisation de la Gaussienne comme suit :

$$\tilde{D}_k(X_i, X_j) = \frac{D_k(X_i, X_j) - \mu_k}{K \cdot \sigma_k} \quad (3.10)$$

où X_i et X_j sont deux images, D_k est la distance à intégrer dans la combinaison et μ_k et σ_k sont la moyenne et la variance des distances D_k . Supposant que les distances ont une distribution gaussienne, on peut montrer qu'il y a une probabilité de 68% pour que les valeurs du dispositif se situent dans la gamme de $[-1, 1]$, si $K = 1$ et une probabilité de 99% si $K = 3$ (ORTEGA ET AL., 1997). Dans la pratique, la distribution réelle ne sera pas parfaitement gaussienne. Cependant, l'expression suivante peut s'appliquer :

$$p \left(-1 \leq \frac{D_k - \mu_k}{K \cdot \sigma_k} \leq 1 \right) = 1 - \frac{1}{K^2} \quad (3.11)$$

Selon cette expression, la probabilité que la distance soit dans l'intervalle $[-1, 1]$ est de 89%, si nous plaçons K à 3 et à 94% pour un réglage de K à 4. Un décalage additionnel simple permet de recentrer les distances dans l'intervalle $[0, 1]$:

$$d_k = \frac{\tilde{D} + 1}{2} \quad (3.12)$$

L'inter-classement des répartitions normalisées, suivant la forme et suivant la couleur, se fait en pratiquant une somme pondérée, en affectant un poids w_k à chaque composante de la distance k :

$$dist(X_i, X_j) = \sum_{k=1}^K w_k d_k(X_i, X_j) \quad (3.13)$$

Dans ce cas précis, nous ne possédons que deux distances à interclasser, et définissons un paramètre α qui sera le poids à donner aux composantes formes et $(1 - \alpha)$ aux composantes couleurs. La formule globale de normalisation et d'interclassement devient :

$$d(X_i, X_j) = \alpha \frac{\frac{\tilde{D}_{Forme}(X_i, X_j) - \mu_{Forme}}{K \cdot \sigma_{Forme}} + 1}{2} + (1 - \alpha) \frac{\frac{\tilde{D}_{Couleur}(X_i, X_j) - \mu_{Couleur}}{K \cdot \sigma_{Couleur}} + 1}{2} \quad (3.14)$$

3.3 Expérimentations

Afin de valider les approches que nous proposons, deux procédés de recherche ont été mis en place. Dans un premier temps, nous perturbons un objet d'origine en simulant sa rotation 3D et les effets de perspective. Ceci permet de valider la méthode GCART et de connaître ses limites. Dans un deuxième temps, nous utilisons la méthode GCART pour reconnaître des objets extraits d'images naturelles. Le point important de cette partie est de mettre en place un procédé d'extraction des objets qui permette de récupérer les pixels appartenant à l'objet. Il faut qu'il soit particulièrement rigoureux sur les frontières de l'objet, car nous savons que la transformation ART est peu robuste aux faibles décalages.

3.3.1 Recherche d'objets déformés par des transformations non planes

Pour valider la méthode GCART de recherche et connaître sa fiabilité, nous mettons en place un test qui consiste à déformer artificiellement un logo d'origine, avec une déformation réaliste, afin de voir à quel rang se situe le logo d'origine après le classement. Ce test a été fait pour 50 logos et chacun a été déformé par plusieurs déformations aléatoirement choisies. Nous caractérisons la méthode de recherche en utilisant le critère Recall et Precision, défini dans le paragraphe 3.2.2.1, qui représente la somme des probabilités d'avoir trouvé l'objet d'origine au rang k . En faisant varier les logos d'origine et les intervalles des déformations aléatoires, on peut définir le type d'objets pour lesquels la méthode GCART est efficace et les perturbations limites jusqu'où la méthode fonctionne avec un taux de détection satisfaisant.

Dans un premier temps, nous définissons les perturbations sans variations d'illuminants. Pour vérifier, la robustesse de GART à retrouver un objet ayant subi une

rotation non plane, nous avons plongé des logos plans dans un environnement de rendu 3D, en appliquant une rotation du plan supportant l'objet et en éclairant les objets par une lumière de face, ce qui ne modifie pas les couleurs du logo. Une deuxième série de tests a été réalisée en modifiant les positions de caméra, ce qui entraîne des modifications d'illuminant, et donc de la luminance des logos. Dans ces conditions, sans variations d'illuminants, le taux de retour, exprimé en terme de courbe de Recall, est bon : dans 86% des tests, la méthode GART retrouve le logo d'origine en première position et 97% dans les dix premières positions. Ceci montre que la méthode fonctionne correctement. Un partitionnement de l'espace des déformations subies par le logo d'origine, nous permet de connaître la fiabilité de la méthode GCART pour chaque type de déformation. Nous ne détectons pas de variations significatives du taux moyen de retour en fonction des variations de l'angle azimuthal ϕ . Il en est de même lors de variations du coefficient de perspective p , définissant la transformation homographique à faire subir à l'objet. Des variations importantes des taux de retour satisfaisants sont notées en fonction des valeurs prises par l'angle θ , définissant l'angle de la rotation non plane subie par l'objet. La figure 3.11 montre des courbes de Recall pour des images perturbées par des déformations appartenant à trois intervalles d'angle θ .

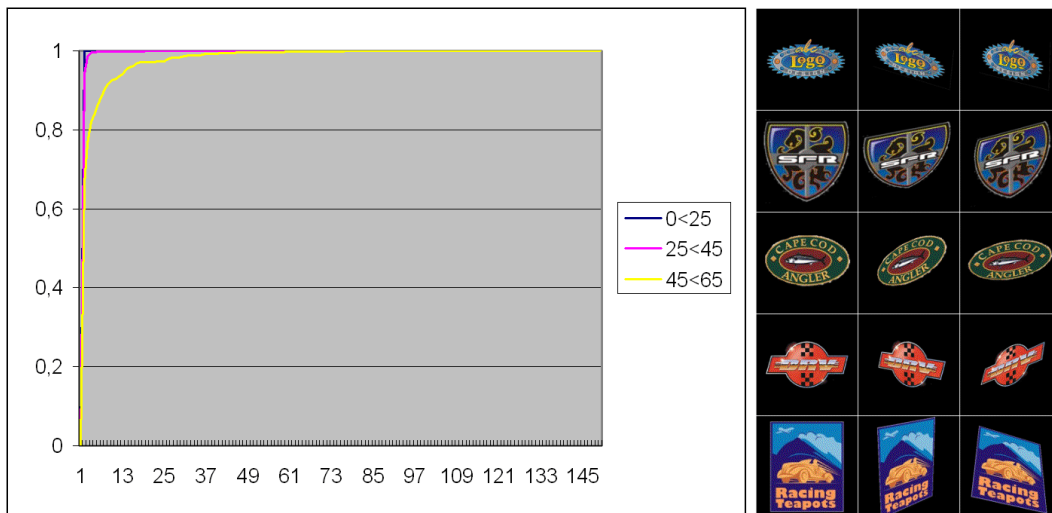


FIG. 3.11 – Courbe de Recall pour 15 logos dans 50 positions aléatoirement tirées, pour 3 intervalles d'angle θ : $[0^\circ, 25^\circ]$, $[25^\circ, 45^\circ]$ et $[45^\circ, 65^\circ]$ et exemple de logos déformés (la première colonne contient les images d'origine).

Nous pouvons constater que le taux moyen de retour satisfaisant diminue lors de l'augmentation de la valeur moyenne de l'angle θ . Ceci est principalement dû au fait que pour une forte rotation, un objet voit sa surface exposée lors de la projection sur le plan image diminuer de manière significative. Les coefficients ART étant normés par le nombre de pixels rentrant dans le calcul (F_{00}), si l'image de l'objet diminue, les coefficients normalisés seront modifiés ce qui perturbera la recherche.

Les mêmes tests ont été réalisés en créant des images en intégrant des variations d'illumination. Les courbes Recall, de la figure 3.12, montrent la probabilité de sortie de l'image recherche en fonction du rang, sur une base d'images aléatoirement déformées avec et sans variation d'illumination. L'avantage de mettre en place des tests automatiques, par un programme de visualisation 3D, est de ne pas avoir à lancer la recherche sur de nombreuses images naturelles, ce qui oblige à indexer manuellement un grand nombre d'images et interdit de faire des tests de grandes ampleurs. On peut remarquer que les résultats, avec variation d'illuminations, sont moins bons que ceux sans variation.

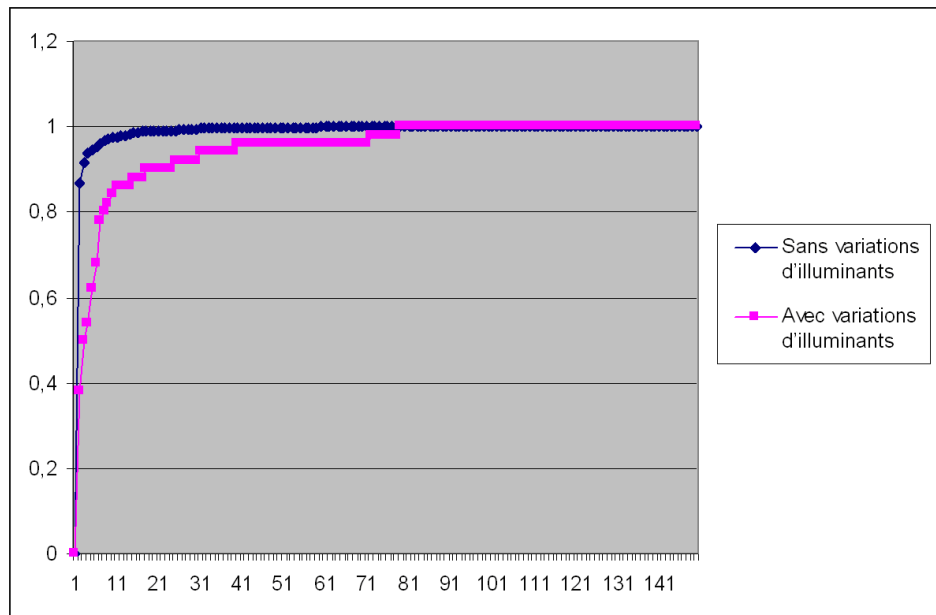


FIG. 3.12 – Courbe de Recall sur une base de 15 logos dans 50 positions aléatoirement tirées avec et sans variations d'illuminations.

3.3.2 Recherche de logos dans une base d'images naturelles

L'étude a pour but d'identifier un objet extrait d'une image naturelle. De nombreuses applications sont possibles, comme la détection et l'identification de panneaux de signalisation routière ou le temps d'exposition d'un logo publicitaire dans un document multimédia.

En préalable à la recherche, la phase d'indexation calcule les coefficients GCART des objets de la base qui seront retrouvés. Pour chaque objet, nous calculons une série de coefficients GCART en la projetant sur les fonctions de base projetées inverses. Ceci nous fournit un descripteur de l'objet dans toutes les positions définies par la discrétisation de l'espace des déformations.

La recherche se compose de trois étapes. Dans un premier temps, il s'agit d'extraire l'objet de l'image naturelle. Ceci peut se faire de différentes façons : par l'extraction manuelle de la ROI (Regions Of Interest) ou par la mise en place d'un système automatique ou supervisé recherchant les zones de l'image répondant à des critères simples comme la similarité d'histogramme. La ROI sélectionnée est indexée par rapport aux fonctions de base d'origine, ce qui fournit une série de coefficients ART. La dernière étape consiste à comparer les coefficients GCART des objets d'origine avec ceux de la ROI. Nous obtenons ainsi un classement par ordre décroissant de similarité des objets de la base les plus ressemblants à la ROI. La figure 3.13 montre le schéma général de l'application, avec en particulier, la phase de recherche en ligne et la phase d'indexation hors ligne.

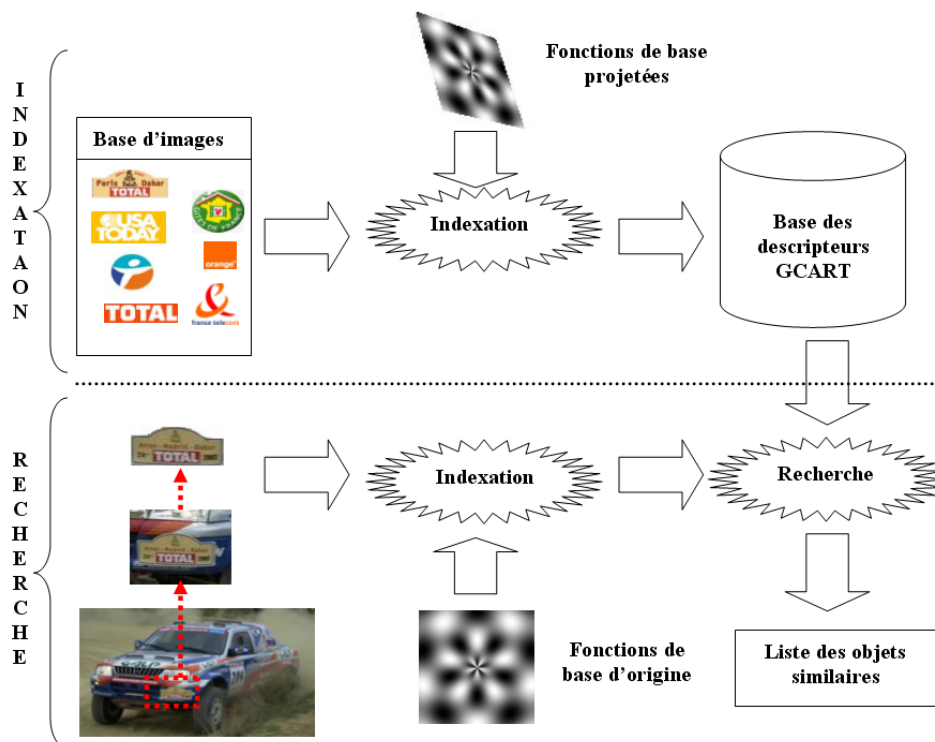


FIG. 3.13 – Schéma du processus d'indexation et de recherche. L'indexation calcule les coefficients ART généralisés en projetant les objets d'origine sur les fonctions de base projetées inverses et les stocke dans la base des coefficients. La recherche extrait l'objet à identifier d'une image naturelle, l'indexe en fonction des fonctions de base d'origine et compare les coefficients trouvés avec ceux de la base. Ce qui retourne un classement de cette dernière en fonction de l'objet clé.

3.3.2.1 Extraction d'un logo d'une image naturelle

La faible robustesse de la méthode ART aux décalages implique que l'extraction des zones de l'image représentant l'objet soit précise. Le rajout de quelques pixels sur le bord de l'objet va entraîner un déplacement du barycentre et une variation du recadrage ce qui ne fournira pas les mêmes coefficients ART. Différents procédés d'extractions peuvent être envisagés : les méthodes manuelles, les méthodes automatiques et les méthodes supervisées.

Les méthodes manuelles fonctionnent en sélectionnant les zones ou les pixels qui composent l'objet. Ces méthodes permettent d'extraire très précisément l'objet, mais demandent une grande interaction avec l'utilisateur ce qui diminue la rapidité et la facilité d'utilisation. Les méthodes automatiques recherchent un objet précis en définissant des critères colorimétriques ou spatiaux colorimétriques pour sélectionner les zones à extraire. De nombreux travaux sur la détection de visages (SENIOR ET AL., 2002) ou la détection d'objets précis, comme des plaques d'immatriculation ou des panneaux de signalisation (WEI ET AL., 2001; GAO ET AL., 2002), utilisent des méthodes automatiques. Elles se basent sur des propriétés propres de l'objet à rechercher, comme la répartition couleur ou la distance entre les différents composants pour sélectionner les zones potentiellement ROI. Ces méthodes sont efficaces et permettent de décharger l'utilisateur de la phase d'extraction manuelle, mais elles restent liées à une application précise et ne permettent pas de détecter des objets variés. Les méthodes supervisées font cohabiter les deux principes, l'utilisateur sélectionne les pixels de l'objet, mais sa sélection est assistée par un programme qui fusionne les pixels voisins selon certains critères.

Nous avons décidé de mettre en place un procédé supervisé. L'utilisateur sélectionne les zones intéressantes. Le pixel sélectionné est considéré comme un germe pour la fusion des pixels adjacents, par une méthode de croissance de région (PLATANIOTIS et VENETSANOPOULOS, 2000). La distance inter régions est calculée en utilisant la distance euclidienne sur l'espace couleur CIE $L^*a^*b^*$, espace qui garantit les meilleures performances de l'algorithme (GAUCH et HSIA, 1992). Deux régions adjacentes sont fusionnées si leur distance est inférieure au seuil de fusion. Le seuil de fusion est défini itérativement par l'utilisateur. Afin de rendre le procédé plus aisé, une première phase de sélection d'un carré englobant est rajoutée, ce qui permet de diminuer la taille de l'image à afficher et, ainsi, de pouvoir zoomer sur l'objet ce qui facilite d'autant la sélection. La figure 3.14 montre un exemple de sélection de la zone extraite par notre procédé pseudo supervisé.

Ce procédé de sélection d'objets supervisés permet d'extraire facilement les pixels de l'image représentant l'objet recherché. L'avantage de cette méthode est sa grande adaptabilité aux différentes images. L'utilisateur fixe lui même le seuil de fusion de chaque région de manière intuitive. La précision du procédé est libre, ce qui permet de faire une sélection rapide, si l'image est non ambiguë, et de rentrer dans le détail de la forme si l'image est complexe.

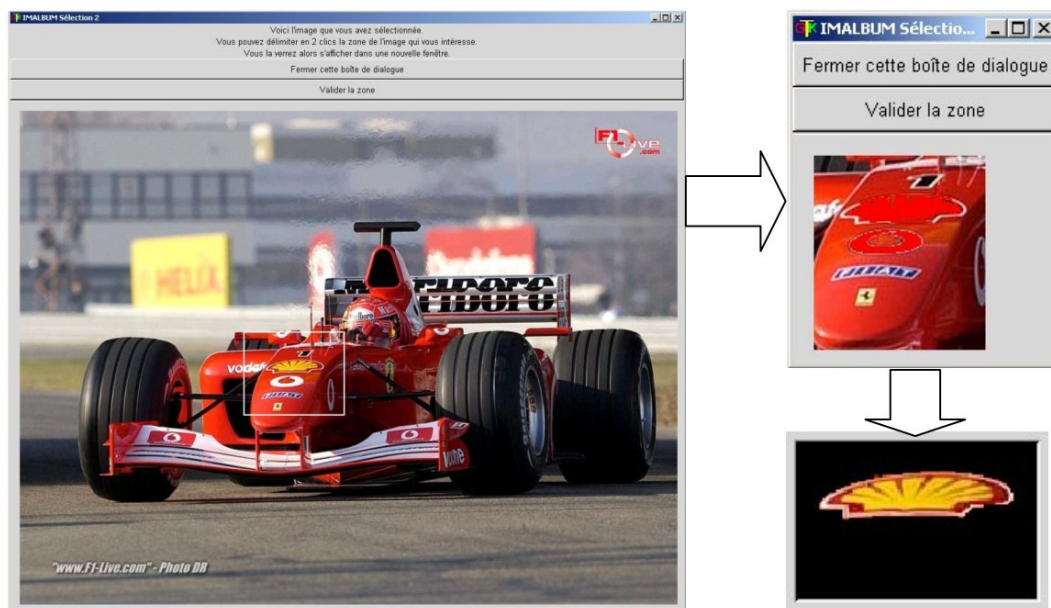


FIG. 3.14 – Fenêtres de sélection des objets à extraire de l'image.

3.3.2.2 Résultats

Afin de montrer le bien fondé de cette méthode de classification GCART, basée sur la forme des objets et sur leurs couleurs, nous indexons une base de 150 logos couleurs selon la forme par la méthode ART et selon la couleur par le procédé énoncé dans les sections précédentes. Nous extrayons 50 objets d'images naturelles et lançons le procédé de recherche.

La figure 3.15 montre les courbes de probabilités de sortie des trois études : forme, couleur et forme et couleur. On peut remarquer que la courbe qui représente l'étude suivant la forme et la couleur est supérieure aux courbes représentant l'étude suivant la forme et l'étude suivant la couleur. Ceci montre l'intérêt de la méthode. La méthode de classification permet de retrouver l'image d'origine en première position dans 51% des études et dans les 6 premiers rangs dans 94% des cas avec une position moyenne de l'objet recherché au rang $k = 2$. On peut, d'autre part, remarquer que la courbe de l'étude forme est inférieure à celle que nous obtenions lors de la recherche sur les objets créés par des perturbations aléatoires de leur position et de l'illuminant ambiant par des programmes de visualisation d'objets 3D. Ceci est dû au fait que l'extraction des objets perturbés ne fournit pas une image aussi précise. L'image de l'objet extrait est de taille nettement inférieure et la précision sur les bords n'est pas aussi nette, ce qui explique ces résultats moins bons. Par contre, la courbe qui représente l'étude suivant la forme et la couleur est supérieure aux deux autres, ce qui montre l'intérêt de la méthode.

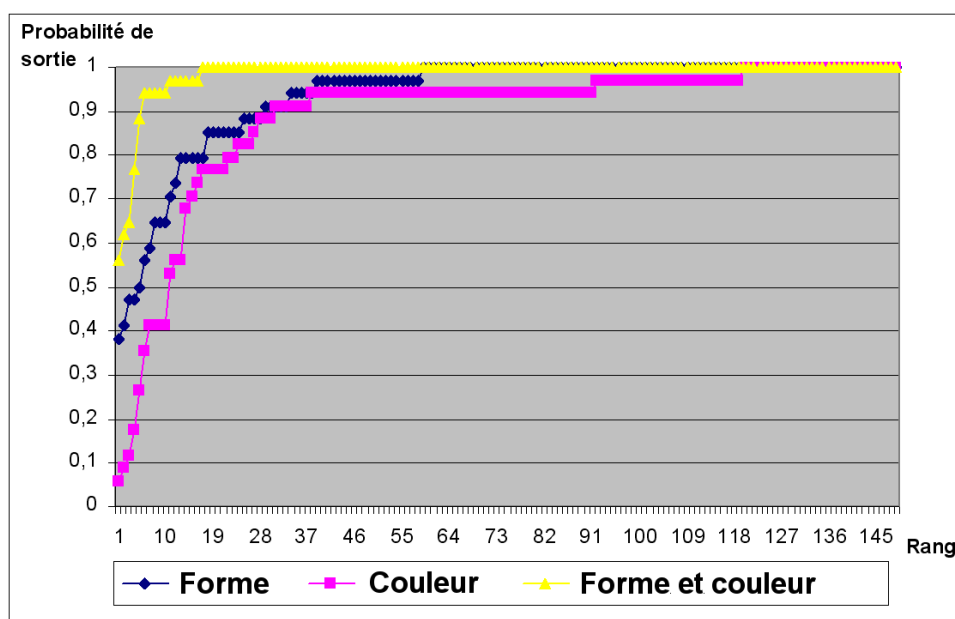


FIG. 3.15 – Courbes de Rappel représentant la position moyenne de l'objet recherché suivant l'étude forme, l'étude couleur et un interclassement des distances forme et couleur.

3.4 Conclusion

La description de la forme des objets composant une image naturelle est un grand axe de recherche visant à mettre en place des méthodes de description et de recherche d'images. La méthode ART, proposée par Mpeg-7, remplit cette tâche, mais ne permet pas de traiter efficacement les images couleurs et n'est pas invariante à toutes les perturbations couramment subies par les objets.

Nous avons proposé deux évolutions à la transformation ART. Dans un premier temps, nous l'avons étendue aux images niveaux de gris et aux images couleurs par l'application de l'étude ART sur l'image luminance et d'une étude des couleurs composant les objets. Nous avons proposé une méthode permettant de combiner l'étude couleur et l'étude forme (CART). Dans un deuxième temps, nous avons rendu ART invariant aux transformations non planes que peut subir un objet au sein d'une image naturelle. Ces travaux ont fait l'objet de plusieurs publications, en particulier, dans la conférence, *ICIP 2004*, (RICARD ET AL., 2004) et dans les Revues, *Pattern Recognition Letters* (RICARD ET AL., 2005) et *Computer Vision On Image Understanding* (IDRISSI ET AL., 2004).

4

Indexation 2D/3D

Sommaire

4.1	Introduction	88
4.2	Limitation de l'indexation 2D/3D	89
4.3	Recherche dynamique d'objets 3D	90
4.3.1	Système de représentation et de création des vues	90
4.3.1.1	Représentation par le Fourier discret d'un objet	92
4.3.1.2	Extraction des coupes	94
4.3.2	Comparaison des vues	96
4.3.3	Processus d'optimisation	97
4.3.3.1	Méthode du Simplex	98
4.3.3.2	Exemple de parcours	99
4.3.4	Processus de mise en correspondance	99
4.4	Expérimentations	101
4.4.1	Mise en correspondance d'un objet 3D à une vue 2D	101
4.4.2	Recherche sur une base d'objets	103
4.5	Conclusion	104

4.1 Introduction

Cette partie présente les approches que nous proposons pour répondre au problème de recherche d'objets 3D à partir de requêtes 2D. La complexité de création d'objets 3D rend la recherche d'objets par requête 3D difficile à utiliser, si on ne dispose pas d'un objet requête qui nous convienne. Pour palier à ce problème, la recherche 2D/3D permet de retrouver des objets 3D à partir de requêtes images. La facilité de création d'une image 2D, par un processus de dessins est facilement utilisable et permet à l'utilisateur de dessiner une vue de la pièce 3D qu'il souhaite retrouver. Une telle recherche est plus complexe, car elle compare des objets de nature différente, mais permet à l'utilisateur de faire lui même sa requête.

Comme énoncé dans le paragraphe 1.3, les approches d'indexation d'objets 3D par vues caractérisent un objet 3D par une ensemble de vues 2D et classe les objets 3D en comparant leurs ensembles de vues à la vue requête. Pour chaque objet 3D, un certain nombre de vues est créé et indexé par un descripteur de forme 2D. La figure 4.1 montre le principe général de la comparaison 2D/3D. L'espace des vues de l'objet est échantillonné et de chaque point de vue est créé une vue, qui est comparée à la vue requête.

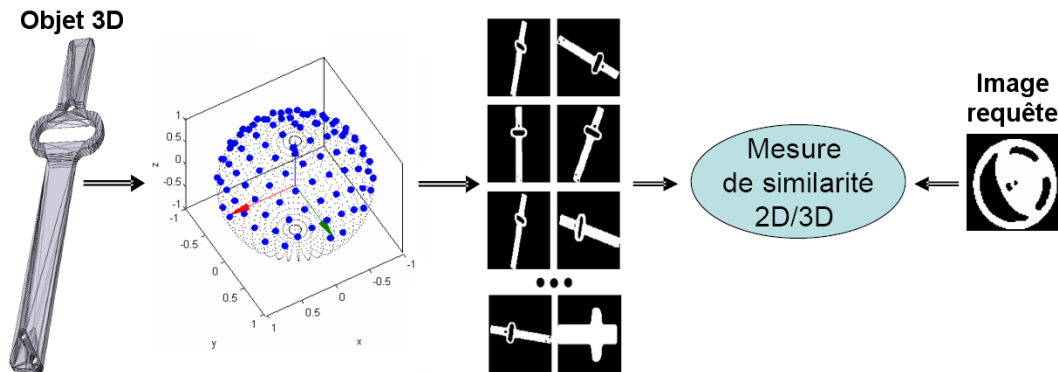


FIG. 4.1 – Principe général de la mise en correspondance 2D/3D, l'espace des vues est échantillonné pour créer un ensemble de vues qui sera comparé à la vue requête

Les méthodes caractérisant les objets 3D par des ensembles de vues, varient en fonction du nombre et de la position des vues et de la méthode de description pour les décrire. Ces méthodes permettent de s'adapter à la complexité de l'objet à décrire en fonction du nombre de vues choisies, mais ne sont pas adaptatives à la complexité de la requête. Pour permettre une double adaptation à l'objet et à la requête, il faut mettre en place un processus dynamique qui, en fonction de la requête, va parcourir l'espace des vues de l'objet et rechercher la vue la plus similaire. Un tel processus demande la mise en place :

- d'un système de représentation, contenant les vues de l'objet et permettant de les extraire rapidement ;

- d'une métrique entre les vues pour les comparer ;
- d'un processus d'optimisation parcourant l'espace des vues pour rechercher la vue la plus similaire.

Cette comparaison dynamique entre l'objet 3D et la vue requête permettra de fournir des résultats plus précis, car l'espace des vues de l'objet sera convenablement représenté. D'autre part, cette représentation fournira des informations supplémentaires en indiquant la position précise de la vue extraite sur la sphère des vues.

Cette partie sera organisée comme suit : dans un premier temps, un état de l'art reviendra sur les méthodes d'indexation d'objets 3D par vues et leurs limites, puis nous présenterons la méthode que nous proposons, enfin, les expériences que nous avons menées en indexation d'objets 3D par requêtes 2D seront détaillées.

4.2 Limitation de l'indexation 2D/3D

Comme nous l'avons vu dans le paragraphe 1.3, les méthodes utilisant des vues des modèles 3D sont motivées par les résultats psychophysiques (RIESENHUBER et POGGIO, 2000) qui montrent que dans le système visuel humain, un objet en trois dimensions est représenté par un ensemble de vues 2D plutôt que par un modèle tridimensionnel. L'idée de ce type de méthode est de créer un certain nombre de vues au départ de l'objet à indexer et de comparer ces vues à la vue requête.

Les méthodes d'indexation d'objets 3D par vues sont souvent adaptatives à l'objet à décrire mais ne le sont pas à la requête. Quelque soit la requête, la comparaison entre l'ensemble des vues de l'objet et la vue requête, se fait de façon pré-définie et ne dépend pas de l'image requête. La recherche va comparer la vue requête à toutes les vues contenues dans le descripteur, quelque soit la vue requête. Si cette vue est très proche de la première vue descripteur, la recherche va quand même la comparer à toutes les vues. Dans le cas contraire, si la vue requête est loin des vues descripteurs, ou n'est pas décrite par l'une d'elle, les résultats ne seront pas exactes malgré une recherche exhaustive sur les vues descripteurs. Le seul moyen pour avoir un processus permettant à la fois de s'adapter à la complexité de l'espace des vues de l'objet à décrire et à celui de l'image requête, est de mettre en place un processus dynamique pour extraire la meilleure vue dans l'espace des vues en fonction de l'image requête. Pour mettre en place un processus dynamique de comparaison de vues, il faut :

- **un système de représentation et de calcul des vues**, qui permet d'extraire rapidement des vues de l'objet 3D.
- **une mesure de similarité** permettant de comparer les vues extraites à la vue requête.
- **un processus d'optimisation** qui, en fonction de la distance entre les vues extraites, recherche la vue optimale.

Un tel processus sera doublement adaptatif à la requête et à l'objet 3D et permettra

de retrouver la vue de l'espace des vues de l'objet la plus proche de l'image requête. La figure 4.2 montre ce processus composé de ces trois points. De plus, nous pouvons récupérer l'information sur la position de la caméra ayant créé la vue, ce qui permettra de faire de la mise en correspondance entre l'objet 3D et la vue requête.

L'approche que nous proposons représentera l'intégralité des vues de l'objet par une transformée de Fourier 3D (paragraphe 4.3.1). La précision d'un tel système sera nettement supérieure à une méthode de sélection de vues, ce qui permettra de fournir la position de la vue requête sur la sphère de vues pour faire de la mise en correspondance entre l'objet 3D et l'image requête.

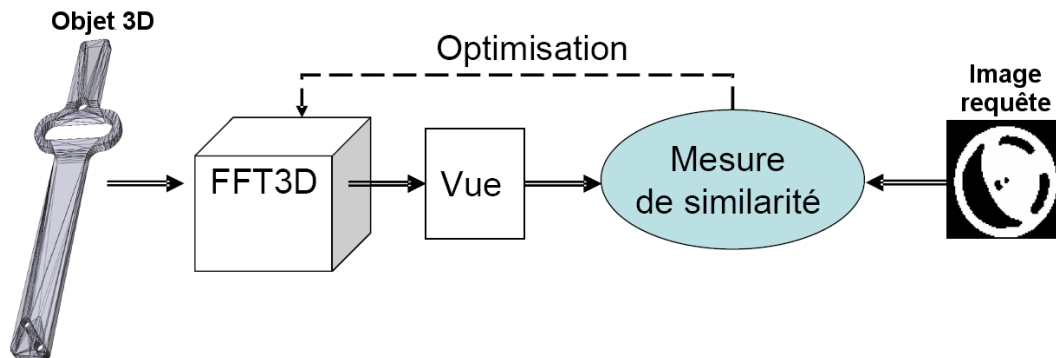


FIG. 4.2 – Principe du processus de recherche dynamique composé d'un système de représentation permettant l'extraction rapide de vues, d'une mesure de similarité entre vues et d'un processus d'optimisation.

4.3 Recherche dynamique d'objets 3D

Comme nous venons de le voir, le processus dynamique de recherche d'objets 3D sera basé sur trois éléments qui vont être explicités dans cette partie : un système de représentation, une mesure de similarité et un processus d'optimisation.

4.3.1 Système de représentation et de création des vues

Les méthodes d'indexation d'objets 3D par vues se basent sur une représentation triangulaire de l'objet et obtiennent les vues par un processus de rendu type Z-buffer. Pour garantir une création de vues rapide, nous avons opté pour la mise en place d'un processus de rendu spectral basé sur le Fourier discret d'un objet 3D et sur le théorème de Radon, identique aux méthodes de Fourier Volume Rendering (MALZBENDER, 1993). Le théorème de Radon stipule, en deux dimensions, que la valeur de la transformée bidimensionnelle d'un signal le long d'un axe de pente $1/m$ passant par le centre d'un spectre 2D est égale à la transformée de Fourier

de la projection orthographique de l'objet parallèlement à un axe de pente $-m$. La fonction g , projection orthographique de la fonction f , et sa transformée de Fourier G peuvent s'écrire :

$$g(t) = m \int_{-\infty}^{\infty} f[x, m(t-x)] dx \quad (4.1)$$

$$G(w) = m \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f[x, m(t-x)] e^{-j\omega t} dx dt \quad (4.2)$$

En faisant un changement de variable $y = m(t-x)$ et donc $t = x + \frac{y}{m}$, on peut écrire :

$$G(w) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f[x, y] e^{-j\omega(x+\frac{y}{m})} dx dy \quad (4.3)$$

En posant $u = w$ et $v = w/m$, on obtient :

$$G(w) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f[x, y] e^{-j(u x + v y)} dx dy = F(u, v) \quad (4.4)$$

on reconnaît dans $G(w)$ la valeur de la transformée de Fourier bidimensionnelle $F(u, v)$ de $f(x, y)$ le long d'une droite passant par l'origine et de pente $1/m$. La figure 4.3 montre un exemple sur une forme 2D.

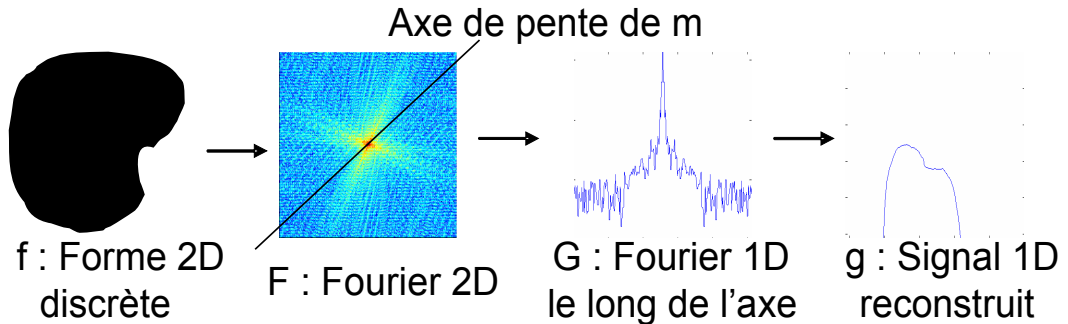


FIG. 4.3 – Principe du théorème de Radon en deux dimensions. La coupe le long d'un axe passant par le centre d'une transformée de Fourier d'une forme discrète représente la transformée de Fourier de la projection de la forme.

Le théorème de Radon est transposable en 3D, où la coupe 2D passant par le centre du spectre 3D d'un objet discret représente la transformée de la projection orthographique volumique de l'objet. Dans une première étape de pré-traitement, l'objet 3D est discrétisé sur une grille de N^3 voxels. La transformation de Fourier 3D (FFT 3D) de l'objet est gardée comme espace de représentation des vues. En effet, une telle représentation contient l'ensemble des vues de l'objet 3D. Pour créer une vue, il suffit d'extraire une coupe passant par le centre de la FFT 3D, qui représente la transformée de Fourier de la projection orthographique de l'objet vu suivant cette

direction. Ceci permet d'extraire une vue plus rapidement que par une projection des triangles, par une méthode basée sur le Z-buffer, car la complexité passe de $O(n^2 * \log T)$ à $O(n^2 \log n)$, mais il faut pour cela manipuler la FFT 3D de l'objet de taille N^3 . La figure 4.4 montre le principe de rendu par l'extraction de coupes d'une FFT 3D.

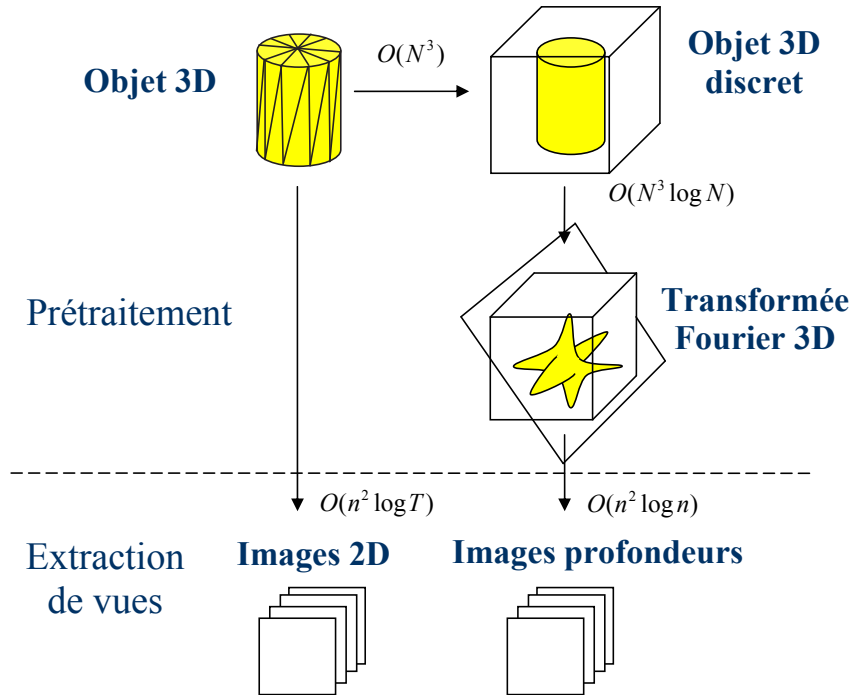


FIG. 4.4 – Principe comparé de rendu par l'extraction de coupe spectral et par un processus basé sur le Z-buffer.

4.3.1.1 Représentation par le Fourier discret d'un objet

La représentation de l'objet par le Fourier discret est naturelle et facile d'utilisation car la transformation de Fourier sur une grille discrète, en deux ou en trois dimensions, est quelque chose de largement utilisé et optimisé.

Cette représentation consiste à discrétiser le modèle 3D sur une grille carrée de taille N^3 et à faire la transformée de Fourier discrète de cette grille. On obtient une grille de même taille contenant les coefficients de Fourier. Pour faire cette transformation, nous utilisons l'algorithme de la transformée de Fourier rapide, Fast Fourier Transform (FFT), de complexité $O(N^3 \log N)$. La transformation discrète correspond à une triple sommation que l'on peut appliquer en trois temps ; d'abord sur x , puis sur y , puis sur z . L'algorithme de calcul rapide de la (FFT) peut s'appliquer de

la manière suivante :

$$F(u, v, w) = \sum_{z=0}^{N-1} \left[\sum_{y=0}^{N-1} \left[\sum_{x=0}^{N-1} f(x, y, z) e^{-2\pi j \frac{ux}{N}} \right] e^{-2\pi j \frac{vy}{N}} \right] e^{-2\pi j \frac{wz}{N}} \quad (4.5)$$

On peut reconnaître, à valeur de y et z fixée, une transformée de Fourier monodimensionnelle sur chacune des lignes de la grille N^3 . Cette transformée monodimensionnelle peut se calculer grâce à l'algorithme de la transformée de Fourier rapide et avoir un coût global en $O(N^2 \log N)$. La figure 4.5 montre le système de rendu spectral que nous avons mis en place.

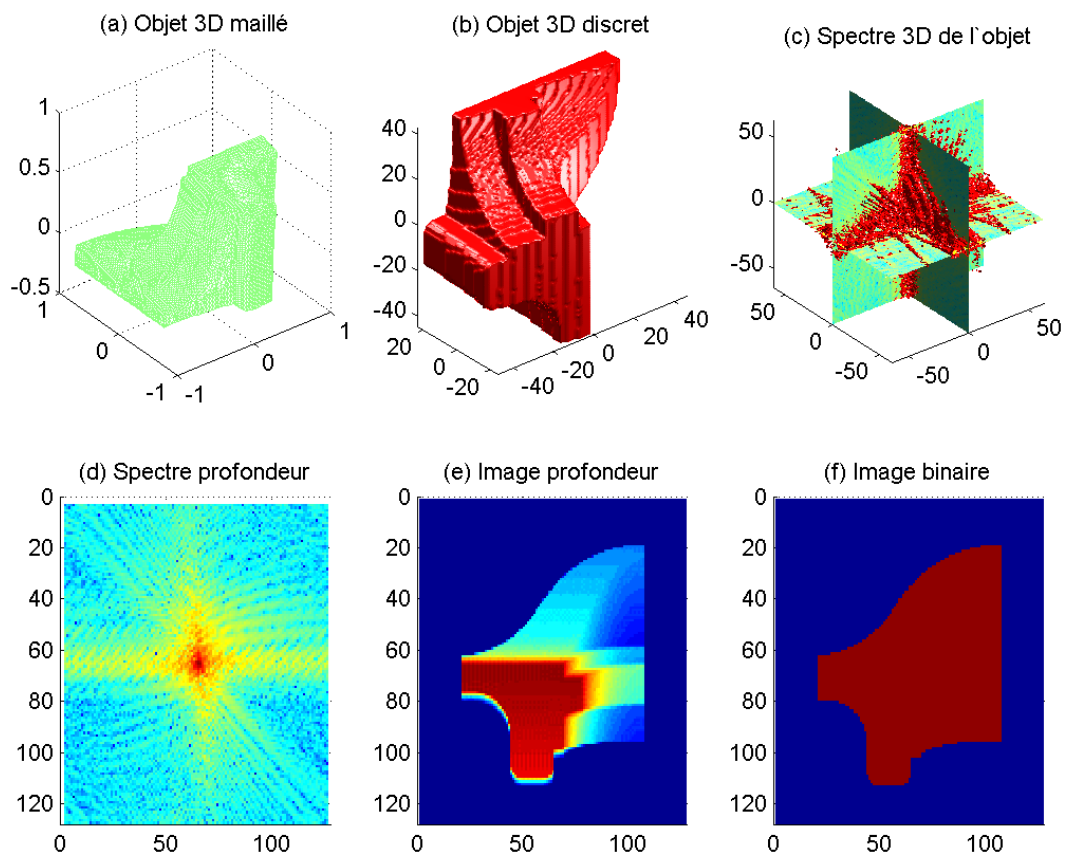


FIG. 4.5 – Exemple de rendu spectral : objet 3D maillé (a) est discrétisé pour obtenir un objet discret (b). Cet objet est transformé dans le domaine fréquentiel et stocké sous forme de spectre 3D discret (c). Le processus de rendu d'image, extrait un spectre 2D (d), le transforme dans l'espace réel, pour obtenir une image profondeur (e) et une image binaire (f).

4.3.1.2 Extraction des coupes

Les objets sont représentés sous la forme de FFT discrètes et nous allons les utiliser pour extraire des coupes 2D de la transformée de Fourier, représentant la transformée de Fourier de l'image profondeur de la vue de l'objet. L'extraction de plan discret d'un volume discret N^3 se fait avec un coût de $O(n^2)$, n étant la taille de la vue extraite. Ceci revient à placer les n^2 points de l'image au centre du volume 3D suivant la bonne orientation. Ce placement peut se faire de manière continue en transformant le plan image par une transformation étant la composition de rotation. Le volume étant discret, on peut également mettre en place une méthode discrète qui extrait le plan. La méthode discrète a l'avantage de mettre en place des calculs plus simples que le calcul matriciel en ne demandant que des additions.

Pour faire une interpolation trilineaire, nous devons interpoler les huit valeurs entourant le points (x, y, z) , dont nous voulons obtenir la valeur. Soit x_f, y_f, z_f , la partie décimal de (x, y, z) donnée par :

$$x_f = x - \lfloor x \rfloor \quad (4.6)$$

$$y_f = y - \lfloor y \rfloor \quad (4.7)$$

$$z_f = z - \lfloor z \rfloor \quad (4.8)$$

Premièrement, on interpole selon z :

$$i_1 = FFT[\lfloor x \rfloor, \lfloor y \rfloor, \lfloor z \rfloor](1 - z_f) + FFT[\lfloor x \rfloor, \lfloor y \rfloor, \lceil z \rceil]z_f \quad (4.9)$$

$$i_2 = FFT[\lceil x \rceil, \lceil y \rceil, \lfloor z \rfloor](1 - z_f) + FFT[\lceil x \rceil, \lceil y \rceil, \lceil z \rceil]z_f \quad (4.10)$$

$$j_1 = FFT[\lfloor x \rfloor, \lfloor y \rfloor, \lceil z \rceil](1 - z_f) + FFT[\lfloor x \rfloor, \lceil y \rceil, \lceil z \rceil]z_f \quad (4.11)$$

$$j_2 = FFT[\lceil x \rceil, \lceil y \rceil, \lceil z \rceil](1 - z_f) + FFT[\lceil x \rceil, \lfloor y \rfloor, \lceil z \rceil]z_f. \quad (4.12)$$

Ensuite, on interpole les valeurs suivant y :

$$w_1 = i_1(1 - y_f) + i_2y_f \quad (4.13)$$

$$w_2 = j_1(1 - y_f) + j_2y_f \quad (4.14)$$

Puis on interpole suivant x :

$$= w_1(1 - x_f) + w_2x_f. \quad (4.15)$$

pour obtenir la valeur interpolée du points. On peut noter que le résultat de l'interpolation trilineaire est indépendant de l'ordre de l'interpolation. Cette interpolation est nécessaire, car le fait de ne pas interpoler, et de prendre les valeurs discrètes les plus proches, engendre des artefacts de discrétisation et l'apparition de répétitions de forme dans l'image reconstruite. La figure 4.6 montre un exemple de coupe avec et sans interpolation.

Pour récupérer la vue de l'objet, nous devons procéder à une transformée de Fourier inverse. Ceci nous fournit l'image profondeur de la vue. Pour la comparer avec

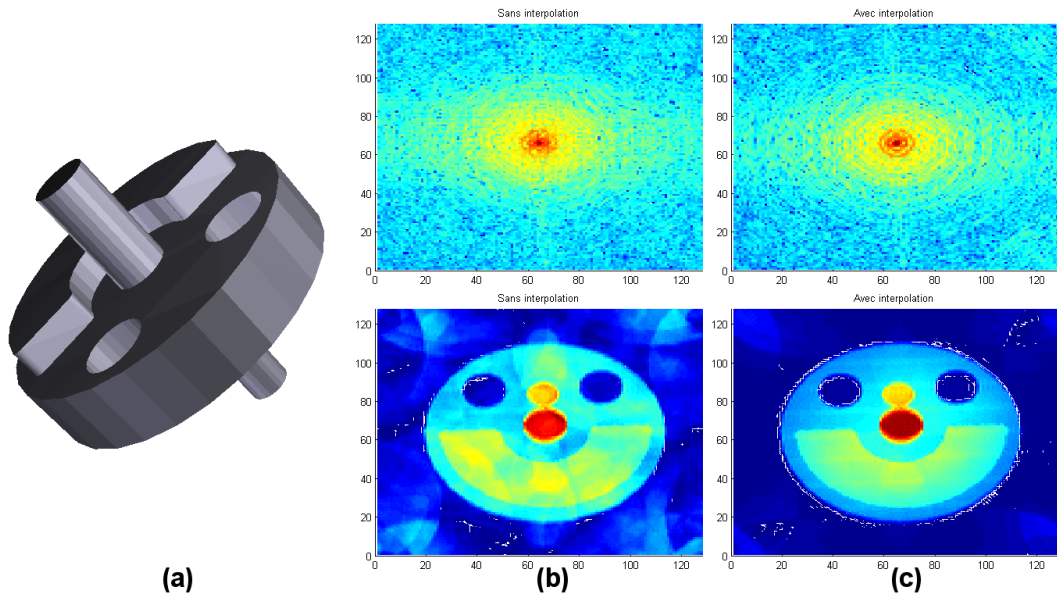


FIG. 4.6 – Exemple de coupe dans le Fourier discret et image reconstruite : objet d'origine (a), sans interpolation (b) et avec interpolation (c).

l'image requête, nous procédons à une binarisation de cette image par l'utilisation d'un algorithme de seuillage par hystérésis. Ce seuillage permet une plus grande précision, car il se base sur deux seuils, un seuil haut et un seuil bas et sur la continuité des régions, comme expliqué par COCQUEREZ et PHILIPP (1995). La figure 4.7 montre la coupe extraite du spectre 3D discret, l'image profondeur et l'image binaire récupérées par ce processus.

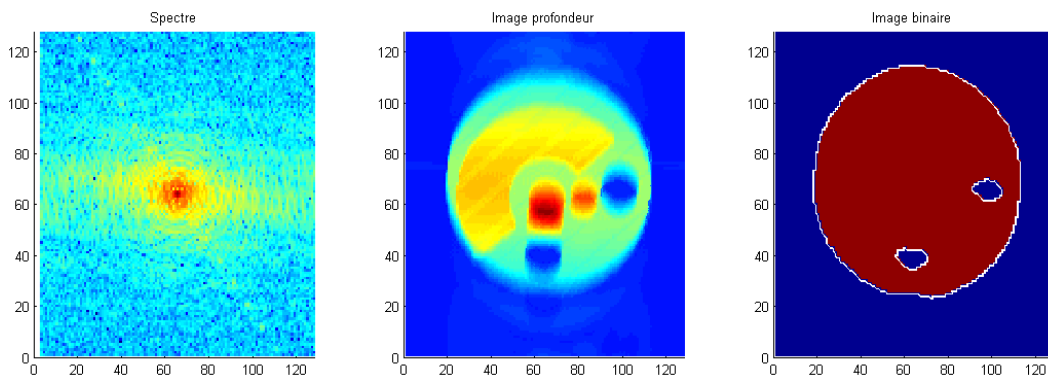


FIG. 4.7 – Coupe de la transformée de Fourier, image profondeur et image binaire obtenues par le processus de récupération de vues.

Cette représentation permet de créer les vues de l'objet de façon rapide et pré-

cise. D'autre part, étant basée sur la transformée de Fourier, elle possède les mêmes propriétés que cette dernière. En particulier, on peut n'extraire que les basses fréquences du Fourier 3D, ce qui fournira une vue basse résolution de l'objet. Ceci permettra de réduire fortement la complexité du processus d'extraction de vues. La figure 4.8 montre un exemple de coupes extraites à quatre plages de fréquences et les vues reconstruites correspondantes.

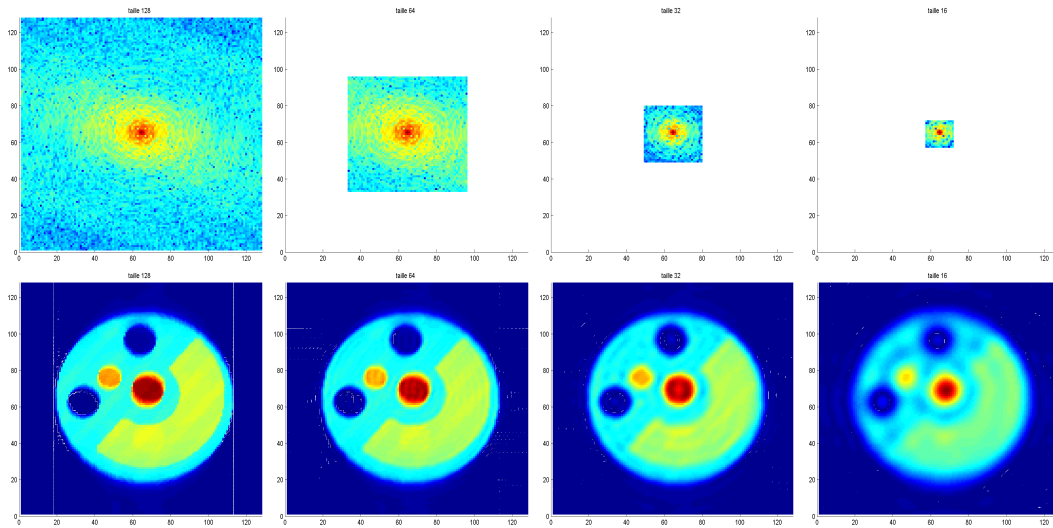


FIG. 4.8 – Exemple de coupes extraites de la représentation de Fourier 3D et les images reconstruites à 4 résolutions différentes : 128, 64, 32 et 16.

4.3.2 Comparaison des vues

La comparaison entre la vue requête et les vues extraites se fait par le calcul du descripteur de forme Angular Radial Transform (ART) présenté dans le chapitre 3. Cette transformation est en particulier invariante en rotation, ce qui permet de ne pas prendre en compte l'orientation de la prise de vue et de ne s'intéresser qu'à sa position. Ce descripteur fournit une distance entre les deux images et permet de savoir si la vue extraite est proche de la vue requête. L'axe de prise de vues est défini sur la demi-sphère des vues par un angle d'Euler (θ, ϕ) . Notons que les vues créées sont des projections orthographiques de l'objet 3D, ce qui revient à ne pas définir de coefficients de perspective lors de la prise de vue, et à limiter la recherche à une demi-sphère des vues, l'autre étant identique. On peut calculer cette distance pour tous les points de l'espace des vues en discrétisant ce dernier. Pour chaque point, on crée une vue extraite et on la compare avec la vue requête. Ceci fournit une carte des distances représentant les distances entre les vues de l'objet et la vue requête. La figure 4.9 montre une carte de distance entre une vue requête et un objet 3D, où pour chaque point de l'espace des vues, $(\theta, \phi) \in [0, 2\pi] * [0, \pi/2]$, une distance est

calculée entre la vue extraite et la vue requête. On peut noter que l'espace des vues est plutôt uniforme et possède peu de maxima locaux, ce qui permet de mettre en place un processus d'optimisation.

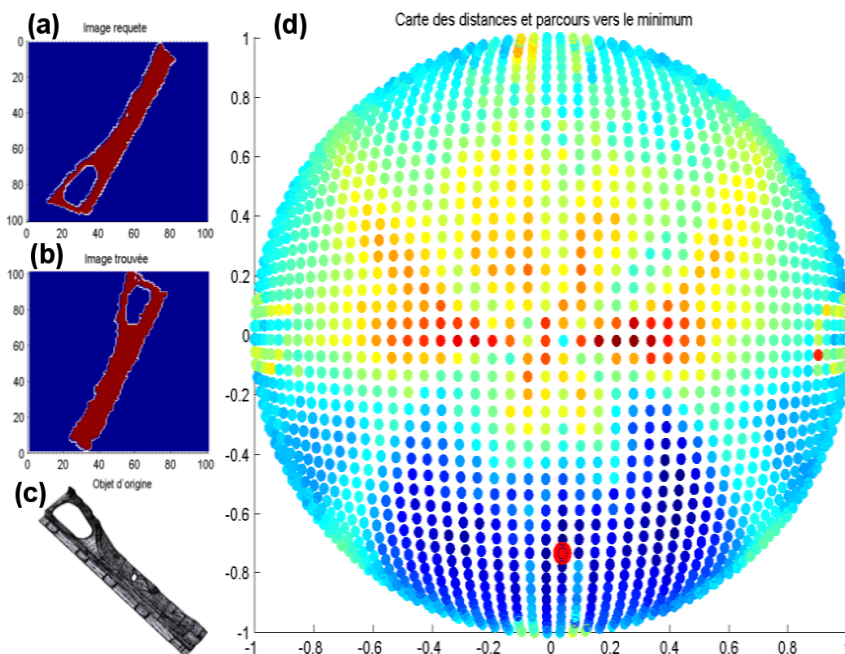


FIG. 4.9 – Carte des distances calculées entre une vue requête (a) et un objet 3D (c). Pour chaque point de vue autour de l'objet, une vue est créée et comparée à la vue requête (c). L'image correspondant au minimum de la sphère des vues (b) est représentée par le point rouge sur la carte des distances (d).

4.3.3 Processus d'optimisation

Le calcul de la carte des distances exige la création prohibitive d'un grand nombre de vues. Pour garantir une comparaison rapide, un processus d'optimisation a été mis en place pour trouver le minimum de la carte des distances en un nombre restreint d'extractions. Un processus d'optimisation dépend de deux paramètres :

- **un espace de recherche** ici θ et ϕ ;
- **une fonction d'évaluation**, qui va être la mesure de similarité entre vues basée sur ART.

L'espace de recherche est la demi-sphère des vues définies en coordonnée sphérique avec $\theta \in [0, 2\pi]$ et $\phi \in [0, \pi/2]$. Il peut être affiché sur la sphère des vues en trois dimensions comme sur la figure 4.9. On peut voir sur ces figures que la fonction d'évaluation est cohérente sur l'espace de recherche et ne possède pas de discontinuités ou de nombreux minima locaux, ce qui garantit des résultats cohérents.

4.3.3.1 Méthode du Simplex

Le coût du test d'un point de l'espace de recherche demande l'extraction d'une coupe de l'espace de représentation des vues, l'indexation de la vue extraite et la comparaison avec la vue requête. Cette opération doit être exécutée le moins possible de fois, tout en parcourant tout l'espace des vues pour retrouver la meilleure vue extraite. Pour diminuer au maximum le nombre de vues testées, tout en gardant des résultats cohérents. Nous avons mis en place un processus de recherche du minimum basé sur la méthode du Simplex de taille variable.

La méthode du Simplex appartient à la classe des méthodes d'optimisation permettant de rechercher le minimum dans un espace à N dimensions, par un nombre restreint de pas, et fut proposée par SPENDLEY ET AL. (1962). Ceci permet en particulier de ne pas faire une recherche exhaustive sur l'intégralité de l'espace de recherche et de s'adapter à la complexité de l'espace. Le Simplex est une figure géométrique ayant un nombre de sommets égal à un de plus que la dimension de l'espace. Comme montré sur la figure 4.10.a, le Simplex d'un espace à deux dimensions a trois sommets et peut être représenté par un triangle. Cette figure montre comment la forme du Simplex est déplacée en rejetant un point et en créant un nouveau simplex formé des deux points du précédent Simplex et du nouveau point. Le choix de la position du nouveau point est défini en fonction des trois points précédents. Les deux points de plus faibles valeurs sont utilisés pour créer le nouveau triangle. Ce processus crée un nouveau Simplex adjacent, en effectuant seulement une nouvelle expérience. La forme du simplex reste la même, un triangle équilatéral de même taille, ce qui ne permet pas de modifier la précision, qui dépend de la taille du Simplex fixée au démarrage. La figure 4.10.b montre comment le Simplex progresse vers l'optimum en répétant ce principe.

Pour palier à ces problèmes, NELDER et MEAD (1965) modifient l'algorithme du Simplex de taille fixe, en autorisant le fait de ne plus utiliser uniquement des triangles équilatéraux. Le nouveau point est placé sur l'espace de recherche en fonction des trois points précédents et il ne faut plus seulement choisir le sens par rapport au triplé précédent, comme dans la méthode du simplex de taille fixe, mais définir les coordonnées du nouveau point, en fonction de la valeur des points déjà calculés. Pour cela, la surface est approximée par un parabolöide passant par les trois points et le minimum du parabolöide est défini comme nouveau point. La figure 4.10.c montre le parcours réalisé par un simplex de taille variable. On peut voir que ce processus converge plus rapidement vers la bonne solution et conduit à une précision supérieure à la méthode de taille fixe.

Les paramètres de la méthode du Simplex de taille variable sont : la taille du pas de recherche, qui définit la taille des triangles de recherche et le critère d'arrêt, qui en fonction du gain de chaque étape de l'algorithme décide de stopper la recherche. Un troisième paramètre a été utilisé, qui est le nombre de processus d'optimisation à réaliser. Nous souhaitons une recherche relativement rapide, et en fonction des

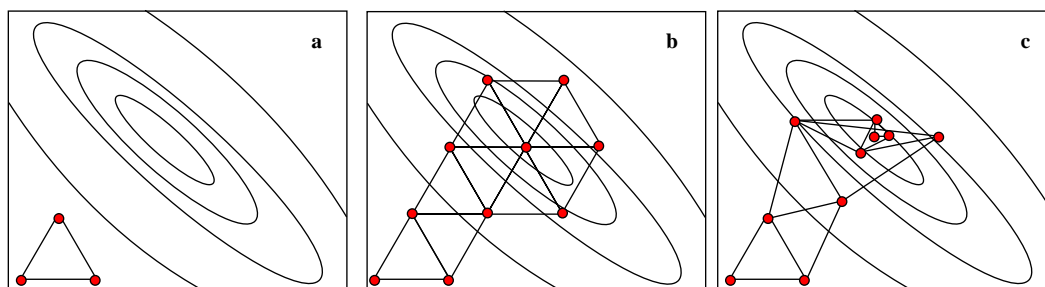


FIG. 4.10 – Exemple de recherche du minimum dans un espace à deux dimensions. (a) Forme du simplex triangulaire pour un espace à deux dimensions. (b) Parcours vers le minimum pour un Simplex de taille fixe. (c) Parcours d'un processus du Simplex de taille variable.

expérimentations que nous avons faites, nous avons limité la longueur du processus, comme nous le verrons plus précisément dans le paragraphe 4.4 consacrée aux expérimentations.

4.3.3.2 Exemple de parcours

Comme énoncé précédemment, nous avons mis en place un parcours de l'espace des vues par un processus de recherche du minimum basé sur un Simplex de taille variable. La figure 4.11 montre le parcours réalisé pour obtenir la vue minimum. Le paramètre principal de cette méthode est la longueur du processus et donc le nombre de vues testées. Plus ce nombre est grand plus la comparaison sera longue mais plus le minimum global aura une chance d'être correctement détecté.

4.3.4 Processus de mise en correspondance

La comparaison d'un objet 3D à une vue requête va se faire suivant le protocole expliqué dans les paragraphes précédents. Le schéma global est expliqué dans la figure 4.12. La vue requête est indexée suivant le descripteur de forme 2D choisi (ART). Les objets 3D sont discrétisés sur une grille de N^3 voxels et transformés dans l'espace spectral par une FFT 3D. Cette représentation est utilisée pour créer les vues par l'extraction de plans de coupe selon le théorème de Radon. Ces vues sont ramenées dans le domaine spatial par une FFT inverse, puis binarisées et indexées pour être comparées avec le descripteur de la vue requête. Le processus d'optimisation évalue la vue extraite pour continuer la recherche.

Deux applications peuvent être visées par ce type de processus : la mise en correspondance entre une image et un objet 3D et la recherche d'un objet 3D dans une base d'objets à partir d'une image requête. Le première processus doit fournir

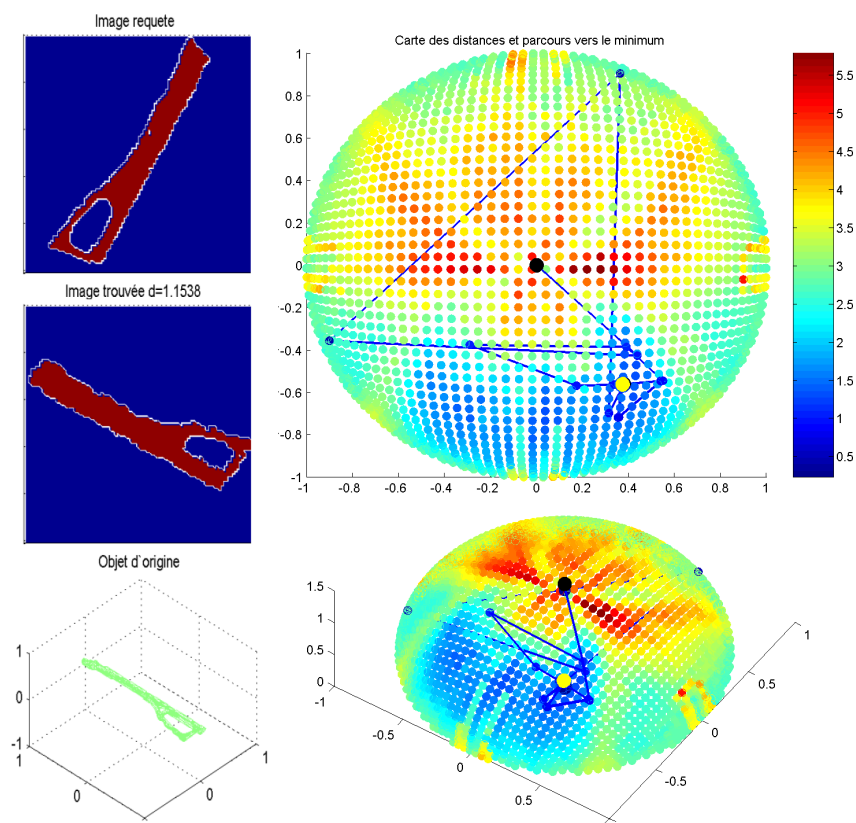


FIG. 4.11 – Recherche de la vue minimum par l'algorithme du Simplex : la vue extraite, la vue trouvée et l'objet 3D (gauche), sphère des vues et parcours vers le minimum, en noir le point de départ de l'algorithme, en jaune le point d'arrivée (droite).

la position de la vue la plus proche de la vue requête. Pour augmenter la précision, deux recherches de minimum par l'algorithme du Simplex ont été mises en place.

Pour faire de la mise en correspondance, nous avons mis en place un parcours en trois temps. Premièrement, l'image requête est comparée à sept vues créées à partir des trois vues principales et des quatre vues sur les quatre grandes diagonales. Le point de vue ayant créé la vue la plus proche est défini comme point de départ de la première recherche du minimum. Deuxièmement, une première recherche du minimum est mise en place. Ce parcours est limité à 10 itérations et les paramètres de l'algorithme sont ajustés pour que la recherche se fasse sur une large zone. Cela permet de rechercher le minimum sur une grande région et de ne pas s'arrêter sur le premier minimum local trouvé. Troisièmement, un deuxième processus de recherche du minimum est exécuté pour converger vers le minimum en parcourant localement la région trouvée. Ce processus teste 27 points de vues pour trouver le minimum de la sphère des vues.

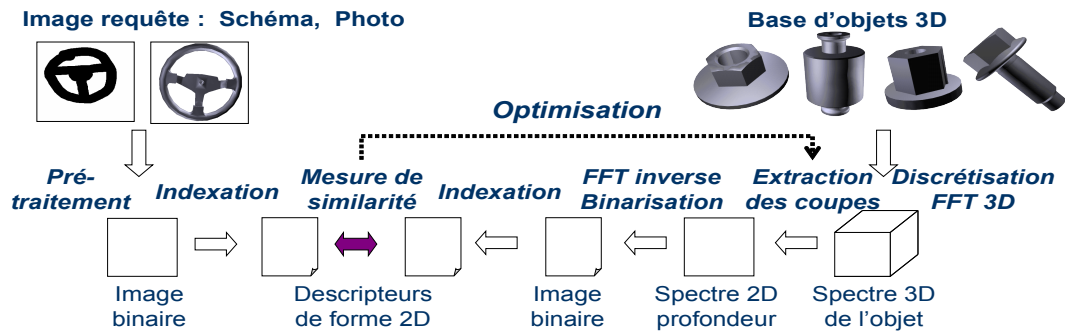


FIG. 4.12 – Processus dynamique de mise en correspondance de l'image requête avec un objet 3D.

Pour faire de la recherche d'un objet 3D dans une base d'images, nous avons opté pour un processus de parcours moins coûteux. Ce parcours se compose en deux phases successives. Premièrement, l'image requête est comparée aux quatre vues, créées sur les trois axes principaux et sur la première grande diagonale. Deuxièmement, une recherche du minimum par l'algorithme du Simplex est exécutée avec un nombre de tests limité à 5. Ceci permet de faire une comparaison rapidement sans trop pénaliser les performances.

4.4 Expérimentations

Le processus que nous avons mis en place permet de mettre en correspondance un objet 3D avec une image 2D et de rechercher le point de vue ayant créé la vue requête. Ceci nous permet de faire de la mise en correspondance entre un objet 3D et une vue 2D. Ce processus permet aussi de comparer une image requête à une base d'objets 3D et faire une recherche d'objets 3D au départ d'une image 2D. Comme nous l'avons énoncé dans la partie 1, les tests ont été réalisés sur une base d'objets 3D issues des bureaux de style de Renault. Cette base contient 758 objets 3D répartis en 75 classes correspondant aux différentes versions d'une même pièce.

4.4.1 Mise en correspondance d'un objet 3D à une vue 2D

La mise en correspondance d'un objet 3D et d'une image 2D permet de mesurer une distance entre deux objets de nature différente, et de rechercher le point de vue ayant créé la vue requête, en utilisant le processus que nous venons d'expliquer. La figure 4.13 montre un exemple de mise en correspondance entre un objet 3D et une image 2D. Cette figure représente la comparaison entre l'image requête (a) et l'objet 3D que l'on souhaite comparer (b). La partie centrale représente la carte des distances sphériques (c). Pour chaque point de la demi-sphère des vues, une image

est créée et comparée avec l'image requête. Ceci permet de vérifier la continuité des distances sur l'espace des vues et de valider la mise en place d'un processus d'optimisation. Cette recherche exhaustive n'est pas réalisée lors de la comparaison, mais permet de voir si la recherche du minimum est correcte.

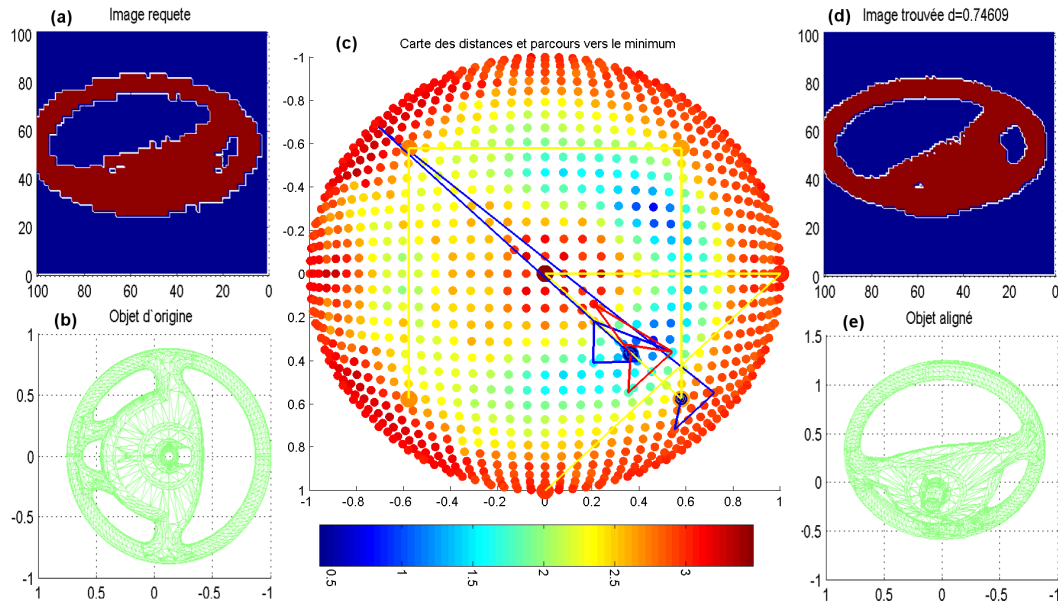


FIG. 4.13 – Exemple de mise en correspondance entre un objet 3D et une image 2D : image requête (a), l'objet 3D (b), carte des distances sphériques et parcours vers le minimum (c), image la plus proche retrouvée (d) et objet aligné suivant le point de vue ayant fourni le minimum (e).

Il est possible d'exécuter plusieurs études successives pour obtenir de meilleurs résultats. Pour faire de la mise en correspondance entre un objet 3D et une image 2D, nous avons opté pour la mise en place d'un processus d'optimisation précis, et nous avons défini un processus en trois temps, comme le montre la figure 4.13. Premièrement, l'image requête est comparée aux trois images principales, créées à partir des trois axes principaux de l'objet, et aux quatre images secondaires, créées à partir des quatre grandes diagonales (figure 4.13.c : parcours jaune). Ceci permet de sélectionner une première zone de la carte de distance. Deuxièmement, un premier parcours par l'algorithme du Simplex est réalisé. Cette première optimisation recherche le minimum sur une large zone, ce qui permet de rechercher le minimum sur une grande région et de ne pas s'arrêter sur le premier minimum local trouvé (figure 4.13.c : parcours bleu). Troisièmement, un deuxième parcours du simplex est exécuté pour converger vers le minimum en parcourant localement la région trouvée (figure 4.13.c : parcours rouge). La partie base de la figure 4.13 représente l'image minimum trouvé (d) et l'objet 3D aligné suivant ce point de vue (e). On peut remarquer, qu'une rotation est apparue entre l'image requête et l'image retrouver. Cela

est dû au descripteur de forme choisi, qui est invariant en rotation. Cela permet de ne pas considérer l'orientation de la caméra lors de la comparaison des vues, ce qui diminue fortement l'espace des vues à parcourir, mais ne permet pas de retrouver la même orientation.

4.4.2 Recherche sur une base d'objets

Le processus de comparaison d'objets permet de comparer un objet à une vue et fournit une mesure de similarité. Ceci peut être appliqué sur une base d'objets 3D, ce qui permet de classer la base en fonction d'une image. Un tel processus aura un coût supérieur à un processus d'indexation 2D/3D classique, qui pré-calculé l'intégralité des vues et compare les descripteurs lors de la phase de recherche.

Pour évaluer les performances de notre méthode, 3 méthodes ayant un nombre de vues fixes de 4 vues, 7 vues et 94 vues ont été mises en place. Cette dernière méthode possède un nombre de vues très élevé et représentera l'information maximum portée par les vues. Les performances des méthodes sont exprimées en terme de Nearest Neighbor, First Tier et Second Tier, critères définis dans le chapitre 2.7. Le protocole expérimental dans laquelle à été réalisé ces testes est le suivant. Pour chaque objets de la base, une image requête est créée aléatoirement. Le point de vue est positionner sur la sphère des vues par une méthode de Sphere Picking, vu dans le paragraphe 1.3.2.2. Pour chaque image requête une recherche est lancé sur la base et les rangs de sortie des objets de la classe de l'objets ayant créé la vue requête sont notés pour calculer les scores. Les courbes que nous affichons sont les courbes moyennes obtenues pour les 758 études réalisées.

La figure 4.14 montre les courbes de NN, FT, et ST des quatre méthodes testées, sur la base Renault de 758 objets. L'abscisse est la longueur du processus d'optimisation de la méthode dynamique, soit le nombre d'itérations faites par le système pour comparer une vue à un objet 3D. Pour les trois méthodes ayant un nombre de vues fixes l'ordonnée ne change pas. Elle est affichée sur ce graphique comme élément de comparaison.

Ces courbes montrent que les résultats augmentent à mesure que la longueur du processus d'optimisation s'accroît. Pour une longueur de processus entre 3 et 10, les courbes NN et FT augmentent respectivement de 10 et de 7%. Si le processus continu, avec une longueur supérieure à 10, le gain s'amenuise, voir s'annule. Ceci est dû aux paramètres de la méthode du Simplex qui ne permettent pas de ressortir des minima locaux trouvés pour parcourir le reste de l'espace. Le tableau 4.1 montre les valeurs des trois indicateurs pour les 4 méthodes avec une longueur du processus d'optimisation de 5 pour la méthode de comparaison dynamique.

Ces résultats montrent que la méthode que nous proposons fournit des résultats supérieurs aux méthodes de sélection de vues fixes, pour un nombre de vues testées moindre. En effet, pour un processus d'optimisation de longueur 5, qui comparera

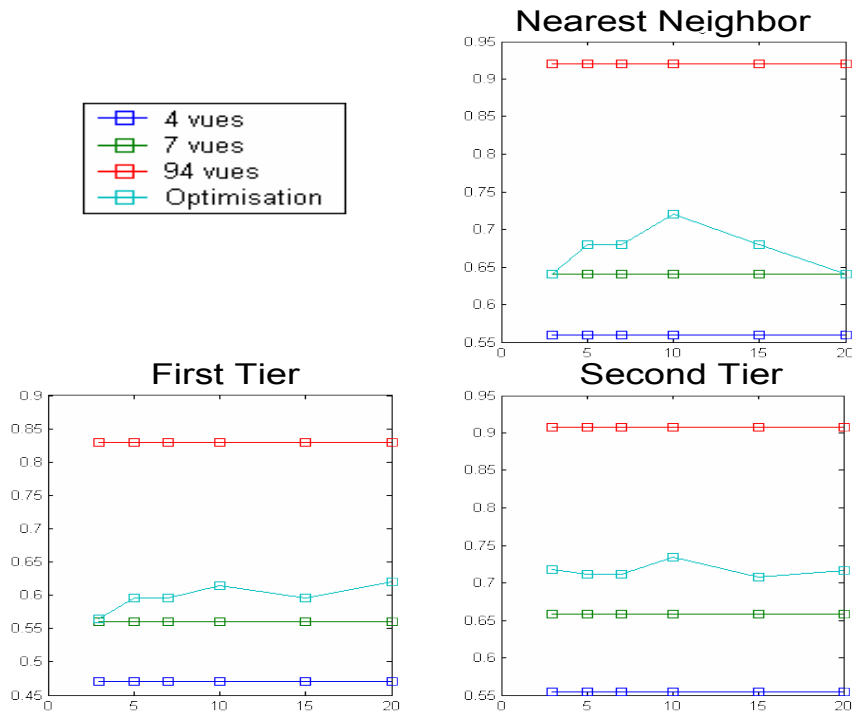


FIG. 4.14 – Courbe des Nearest Neighbor, First Tier et Second Tier, pour trois méthodes à sélection de vues fixes : 4, 7 et 94, et pour la méthode avec optimisation. L’abscisse représente le nombre d’itérations du processus d’optimisation.

la vue requête à 5 vues extraites, les résultats sont meilleurs que les méthodes de sélection de vues fixes à 4 et 7 vues, pour un nombre de vues testées moindre.

4.5 Conclusion

Dans cette partie, nous avons présenté une nouvelle méthode de recherche dynamique d’objets 3D à partir de requêtes 2D. Celle-ci est basée sur un système de représentation et de création de vues, couplé à un processus d’optimisation. Le système de représentation des vues choisies utilise la transformée de Fourier 3D et permet d’extraire rapidement les vues 2D de l’objet. Le processus d’optimisation parcourt l’espace des vues de l’objet en fonction de la requête par une méthode du Simplex, ce qui s’adapte à la fois à l’objet 3D et à la requête.

Cette nouvelle méthode a été proposée pour résoudre le problème d’indexation 2D/3D et permet de rechercher des objets dans une base à partir d’une requête 2D. Nous avons montré qu’avec un nombre de vues générées restreint, comparable aux méthodes d’indexation à vues fixes, nous obtenons de très bons résultats tout en gardant un processus hautement adaptatif. La recherche de la *meilleure vue* mise

Nombre de vues	NN	FT	ST
4 vues	0,56	0,47	0,554
7 vues	0,64	0,56	0,658
94 vues	0,92	0,83	0,908
Optimisation 5	0,69	0,60	0,716

TAB. 4.1 – Comparaison en terme de Nearest Neighbor, First Tier et Second Tier, des 4 méthodes testées.

en place dans l'optimisation proposée, nous permet aussi d'obtenir une information précise sur l'orientation de la vue requête.

Le choix du modèle de représentation que nous avons choisi, basé sur la transformée de Fourier, constitue un élément qui nous paraît très prometteur pour de futures recherches. Les perspectives liées à cela seront présentées plus en détail dans la conclusion générale (chapitre 8).

Troisième partie

Indexation par requêtes 3D

5

Indexation d'objets 3D

Sommaire

5.1	Introduction	110
5.2	3D Angular Radial Transform	110
5.2.1	Définition de ART 3D	110
5.2.2	Processus d'indexation	113
5.2.3	Expérimentations	113
5.2.3.1	Paramètres de ART 3D	114
5.2.3.2	Robustesse aux déformations	116
5.2.3.3	Comparaison aux harmoniques sphériques	118
5.2.3.4	Comparaison au sein de Semantic-3D	120
	Comparaison par critères algorithmique.	121
	Performances des requêtes par similarité.	121
5.3	Indexation 3D par vues	122
5.3.1	Indexation	122
5.3.1.1	Création des vues caractéristiques	123
5.3.1.2	Position des vues caractéristiques	123
5.3.1.3	Mesure de similarité entre deux ensembles de vues	123
5.3.1.4	Processus de recherche	124
5.3.2	Expérimentations	124
5.4	Conclusion	127

5.1 Introduction

Dans le cadre du projet RNRT Semantic-3D¹, de très grandes bases d'objets 3D sont à indexer. Pour exemple, les tests réalisés durant ce projet ont été faits sur une base de 5000 objets, ne représentant que le quart avant droit d'une voiture Laguna II. Un rapide calcul nous permet de voir qu'une base d'objets 3D contenant les modèles 3D de l'ensemble de la gamme Renault contiendrait plusieurs centaines de milliers de pièces (voir paragraphe 2.7.1). De telles bases requièrent la mise en place de descripteurs de formes de petite taille, ayant un coût d'indexation et de recherche faible pour garantir des réponses rapides. Pour indexer des objets 3D, des approches basées sur les moments 3D existent et l'on peut en particulier citer la représentation en harmoniques sphériques (FUNKHOUSER ET AL., 2003; KAZHDAN ET AL., 2003; VRANIC et SAUPE, 2001, 2002), les moments géométriques (SAUPE et VRANIC, 2001) et les moments de Zernike (CANTERAKIS, 1999), comme nous les avons présentées dans le chapitre 2.

Sur les images 2D, le comité MPEG-7 a proposé d'indexer les images en utilisant un descripteur de forme par approche région, Angular Radial Transform (ART) (cf. chapitre 1.2.3). Ce descripteur de formes possède de nombreuses propriétés : petite taille, robuste aux bruits et aux variations d'échelle, invariance en rotation, possibilité de décrire des objets complexes. Cette partie présente les travaux que nous avons réalisés sur les descripteurs de forme 3D basés sur les moments et en particulier, un nouveau descripteur de forme tridimensionnel, nommé 3D Angular Radial Transform, qui correspond à une généralisation d'ART en trois dimensions pour l'indexation d'objets 3D.

L'indexation d'objets 3D par vues est également une voie de recherche prometteuse, comme nous l'avons vu dans le paragraphe 2.5). Nous étudions dans ce chapitre les méthodes d'indexations d'objets 3D par vues. Nous allons en particulier utiliser le système de rendu spectral, présenté dans le chapitre 4, pour évaluer différentes méthodes en fonction du types de vues décrites (image binaire, image profondeur, spectre de forme...).

5.2 3D Angular Radial Transform

5.2.1 Définition de ART 3D

Pour appliquer la transformation ART 3D, les objets sont représentés en coordonnées sphériques où ϕ est l'angle azimuthal dans le plan xy par rapport à l'axe x , θ est l'angle polaire par rapport à l'axe z et ρ est la distance du point à l'origine. L'ART 3D est une transformation unitaire orthogonale complexe définie sur la boule

¹<http://www.semantic-3d.net>

unité. Nous définissons les coefficients ART 3D par :

$$F_{nm_\theta m_\phi} = \int_0^{2\pi} \int_0^\pi \int_0^1 V_{nm_\theta m_\phi}^*(\rho, \theta, \phi) f(\rho, \theta, \phi) \rho d\rho d\theta d\phi \quad (5.1)$$

où $F_{nm_\theta m_\phi}$ est le coefficient ART d'ordre n , m_θ et m_ϕ , $f(\rho, \theta, \phi)$ est l'objet 3D exprimé en coordonnées sphériques et $V_{nm_\theta m_\phi}(\rho, \theta, \phi)$ est la fonction de base d'ART 3D (FDB). Ces fonctions sont séparables suivant la direction radiale et les deux directions angulaires :

$$V_{nm_\theta m_\phi}(\rho, \theta, \phi) = A_{m_\theta}(\theta) A_{m_\phi}(\phi) R_n(\rho) \quad (5.2)$$

Comme en 2D, la fonction radiale est définie par une fonction cosinus :

$$R_n(\rho) = \begin{cases} 1 & n = 0 \\ 2 \cos(\pi n \rho) & n \neq 0 \end{cases} \quad (5.3)$$

Les fonctions angulaires sont définies par des fonctions exponentielles pour avoir une stabilité en rotation :

$$\begin{aligned} A_{m_\theta}(\theta) &= \frac{1}{2\pi} e^{2jm_\theta\theta} \\ A_{m_\phi}(\phi) &= \frac{1}{\pi} e^{jm_\phi\phi} \end{aligned} \quad (5.4)$$

Deux fonctions angulaires sont définies pour garantir la continuité des fonctions de base en fonction des valeurs de θ et ϕ . Les valeurs des paramètres n , m_θ et m_ϕ sont fixées au regard du ratio coût/efficacité, en calculant les courbes de *Recall* pour différentes valeurs de ces paramètres. Sur la base de modèles 3D techniques présentés dans le paragraphe 2.7.1, nous fixons les valeurs : $n = 3$, $m_\theta = 5$ et $m_\phi = 5$, comme expliqué dans le paragraphe 5.2.3.1. Les parties réelles des fonctions de base sont visibles dans la figure 5.1.

La mesure de similarité est calculée en utilisant une distance L_1 entre les descripteurs ART 3D :

$$d(R, I) = \sum_{i=1}^{n \cdot m_\theta \cdot m_\phi} \|ART3D_R[i] - ART3D_I[i]\| \quad (5.5)$$

où R et I représentent l'objet requête et un objet de la base et $ART3D$ est le tableau contenant les valeurs des descripteurs ART3D.

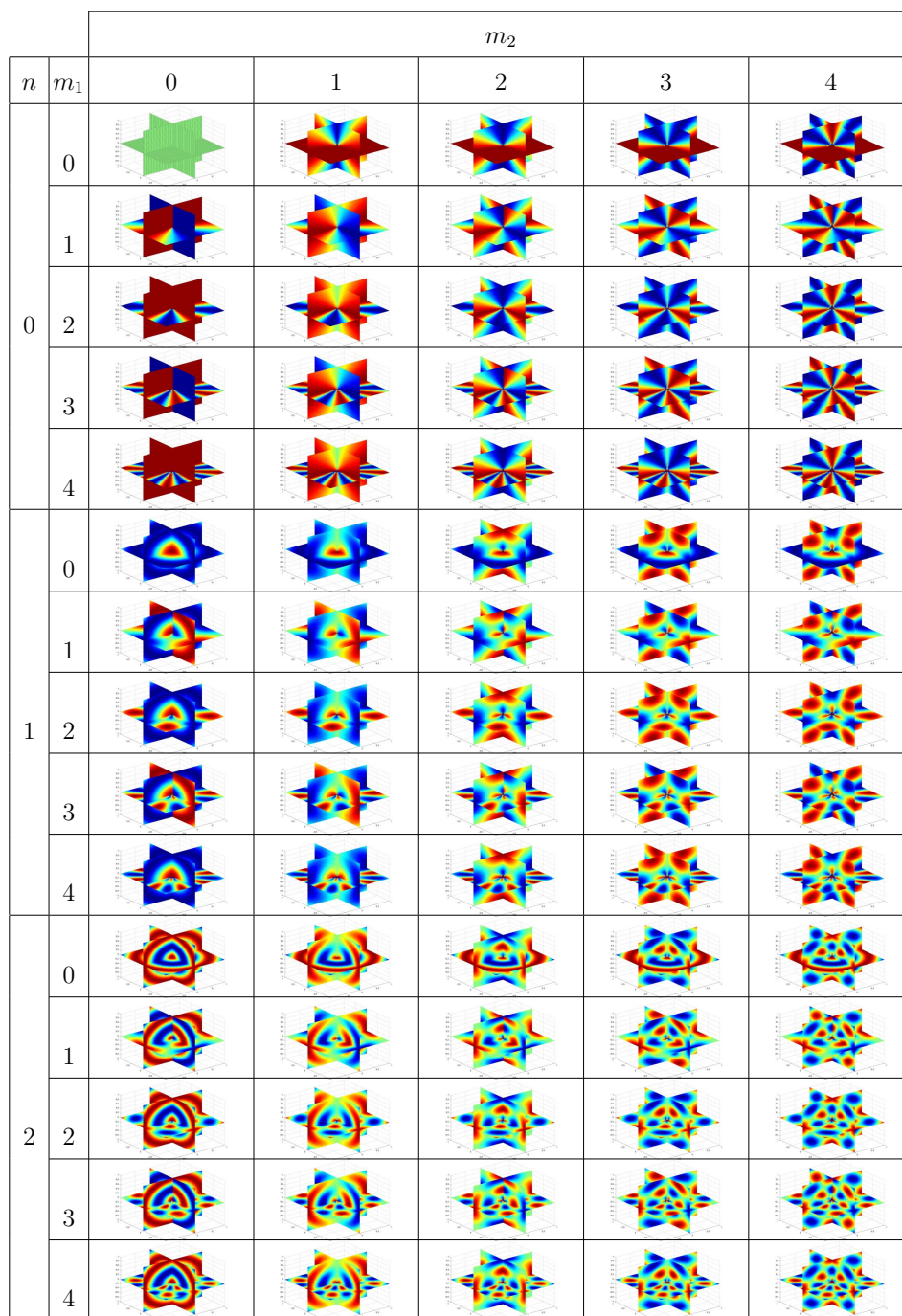


FIG. 5.1 – Partie réelle des fonctions de base 3D.

5.2.2 Processus d'indexation

Une propriété importante de l'ART 2D est son invariance à la rotation. En effet, une rotation en coordonnée polaire peut s'exprimer comme la somme d'un terme constant sur les composantes angulaires.

$$(\rho, \theta) \xrightarrow{Rot_\alpha} (\rho, \theta + \alpha) \quad (5.6)$$

Ceci ne modifie pas la norme de la fonction $A_n(\theta)$ et le descripteur ART, comme défini dans le paragraphe 1.2.3. En 3D, une rotation quelconque ne peut être exprimée comme la somme de valeurs constantes sur les composantes angulaires. Ce qui modifie les fonctions exponentielles A_{m_i} et le descripteur. Les seules rotations qui sont invariantes par cette description sont les rotations autour de l'axe z . Ces rotations ne modifient pas les composantes θ des points de l'objet. Par conséquent, les rotations peuvent être exprimées comme une somme linéaire sur la composante ϕ qui est invariante. Pour résoudre ce problème, nous alignons les objets selon leurs directions principales. Une analyse en composante principale (ACP) est appliquée pour obtenir les directions principales des objets et nous alignons les objets seulement selon l'axe de la première direction principale (direction confondue avec l'axe de z). L'obtention de la robustesse aux rotations par l'alignement sur un seul axe est plus stable qu'un alignement sur les trois axes. La figure 5.2 montre le processus d'indexation.

Avant de projeter l'objet 3D sur les fonctions de base, l'objet subit un pré-traitement. Premièrement, l'objet est discrétisé sur une grille régulière et les voxels sont étiquetés intérieur/extérieur. La discrétisation est utilisée pour calculer les paramètres de centrage, de mise à l'échelle et d'alignement suivant l'axe z . Ensuite, l'objet discrétisé est aligné selon ces paramètres : l'objet 3D est centré sur le centre de gravité et mis à l'échelle de façon à ce que les bords de l'objet coïncident avec ceux de la grille. Le pré-traitement rend la description robuste aux translations, aux variations d'échelle et aux rotations. Enfin, la discrétisation est projetée sur les fonctions de base d'ART 3D pour obtenir les coefficients ART 3D. Ces coefficients sont normalisés et stockés dans le descripteur de forme 3D.

5.2.3 Expérimentations

Trois séries d'expériences ont été réalisées. Dans un premier temps, des tests ont été mis en place pour fixer le nombre de fonctions de base et la taille de la discrétisation. Puis nous avons évalué la robustesse de la méthode et enfin, nous avons comparé ART 3D à la description par harmoniques sphériques (HS) (KAZHDAN ET AL., 2003). Ces tests ont été faits sur deux bases de modèles : Princeton Shape Benchmark (SHILANE ET AL., 2004) et une base de modèles techniques 3D Renault, présentée dans le paragraphe 2.7.1.

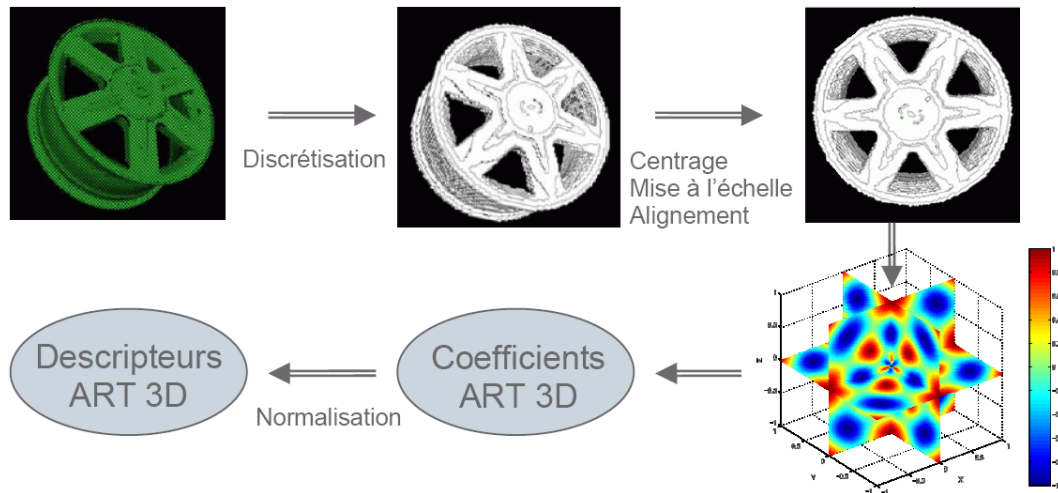


FIG. 5.2 – Processus d'indexation.

5.2.3.1 Paramètres de ART 3D

Pour fixer les valeurs des paramètres, les courbes de Recall calculée sur la base Semantic-3D contenant des objets 3D mécaniques sont comparées. Deux paramètres sont à fixer :

- le nombre des fonctions de base utiles pour indexer les objets 3D. Ce nombre correspond à la taille du descripteur qui sera créé et utilisé pour comparer les objets entre eux ;
- la taille de la discrétisation, qui correspond à la précision de la discrétisation.

Pour évaluer le nombre de fonctions de base et, par la même, la précision de la description en fonction de la taille du descripteur, 25 valeurs des paramètres n , m_θ et m_ϕ ont été évaluées. La figure 5.3 et le tableau 5.1 montrent les résultats obtenus en terme de Nearest Neighbor, First Tier et Second Tier. La taille des descripteurs est directement reliée à la complexité des processus d'indexation et de recherche. Pour ces raisons, nous avons choisi de conserver 75 descripteurs, ce qui correspond à $n = 3$ et $m_\theta = m_\phi = 5$. Ces valeurs fournissent un bon rapport coût/efficacité, comme on peut le voir sur les expériences réalisées. La figure 5.4 montre ces résultats en terme de courbe Recall/Precision.

L'évaluation du paramètre S , qui correspond à la taille de la discrétisation, a été faite, en comparant sept valeurs de S , $S \in \{25, 51, 64, 75, 101, 128, 256\}$. La figures 5.5 et le tableau 5.2 montrent les résultats, en terme de courbe Recall/Precision et selon les critères de performance : Nearest Neighbor, First Tier, Second Tier, pour plusieurs tailles de discrétisation S . Les meilleurs résultats sont obtenus sur la base de modèles techniques pour $S = 101$. Ces valeurs sont utilisées dans la suite de ce travail.

(n, m_θ, m_ϕ)	Taille du descripteur	Nearest Neighbor	First Tier	Second Tier
(2, 1, 1)	2	0.477	0.262	0.360
(3, 1, 1)	3	0.813	0.541	0.642
(4, 1, 1)	4	0.885	0.599	0.688
(5, 1, 1)	5	0.903	0.650	0.734
(1, 3, 3)	9	0.965	0.792	0.849
(2, 3, 3)	18	0.977	0.847	0.896
(1, 5, 5)	25	0.975	0.842	0.895
(3, 3, 3)	27	0.990	0.879	0.924
(4, 3, 3)	36	0.989	0.896	0.926
(5, 3, 3)	45	0.989	0.894	0.928
(1, 7, 7)	49	0.982	0.852	0.896
(2, 5, 5)	50	0.989	0.879	0.921
(3,5,5)	75	0.993	0.893	0.928
(1, 9, 9)	81	0.979	0.847	0.887
(2, 7, 7)	98	0.991	0.892	0.927
(4, 5, 5)	100	0.993	0.907	0.942
(5, 5, 5)	125	0.993	0.905	0.942
(3, 7, 7)	147	0.990	0.894	0.930
(2, 9, 9)	162	0.985	0.878	0.923
(4, 7, 7)	196	0.991	0.898	0.935
(3, 9, 9)	243	0.988	0.889	0.924
(5, 7, 7)	245	0.994	0.905	0.942
(4, 9, 9)	324	0.993	0.898	0.931
(5, 9, 9)	405	0.997	0.901	0.935

TAB. 5.1 – Nearest Neighbor, First Tier, Second Tier en fonction du nombre de descripteurs.

Taille de la discrétisation	Nearest Neighbor	First Tier	Second Tier
25	0.956	0.745	0.810
50	0.969	0.830	0.865
64	0.976	0.838	0.870
75	0.985	0.869	0.913
101	0.993	0.893	0.928
128	0.986	0.882	0.907
256	0.987	0.883	0.908

TAB. 5.2 – Nearest Neighbor, First Tier, Second Tier en fonction de la taille de la discrétisation S .

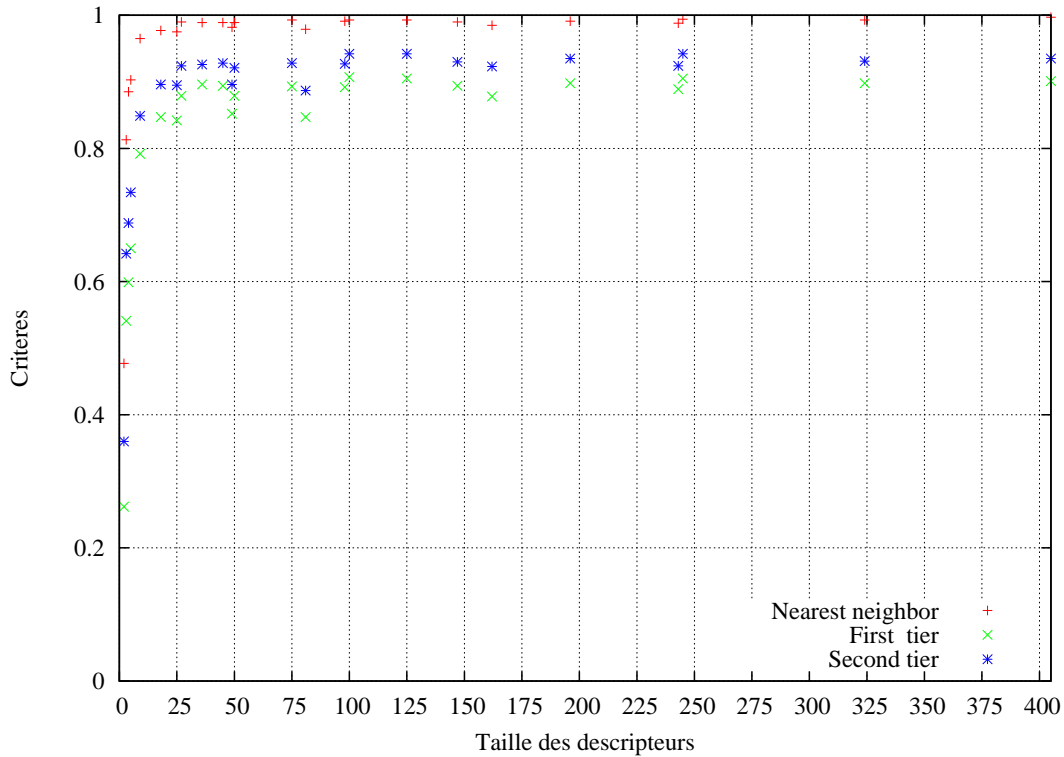


FIG. 5.3 – Nuages de points représentant les critères : Nearest Neighbor, First Tier et Second Tier, pour plusieurs valeurs des paramètres n , m_θ et m_ϕ .

5.2.3.2 Robustesse aux déformations

Pour évaluer la robustesse du processus, nous déformons des objets 3D selon certaines transformations et mesurons les variations de distance observées entre l'objet d'origine et les objets déformés. Les transformations couramment subies par un objet 3D sont :

- la translation ;
- le changement d'échelle ;
- la rotation ;
- le bruit.

La table 5.3 montre les distances maximum et moyenne obtenues pour ces quatre déformations. Pour chaque transformation, nous créons des ensembles d'objets 3D déformés et pour chaque objet, nous calculons la distance à l'objet d'origine.

La translation n'a pas effet sur la distance car le pré-traitement centre et met à l'échelle les objets. Pour les mêmes raisons, le changement d'échelle n'a qu'un très faible effet dû aux artefacts de discrétisation. La distance maximum obtenue est de 0.16 alors que la distance moyenne entre deux objets de classes différentes est autour

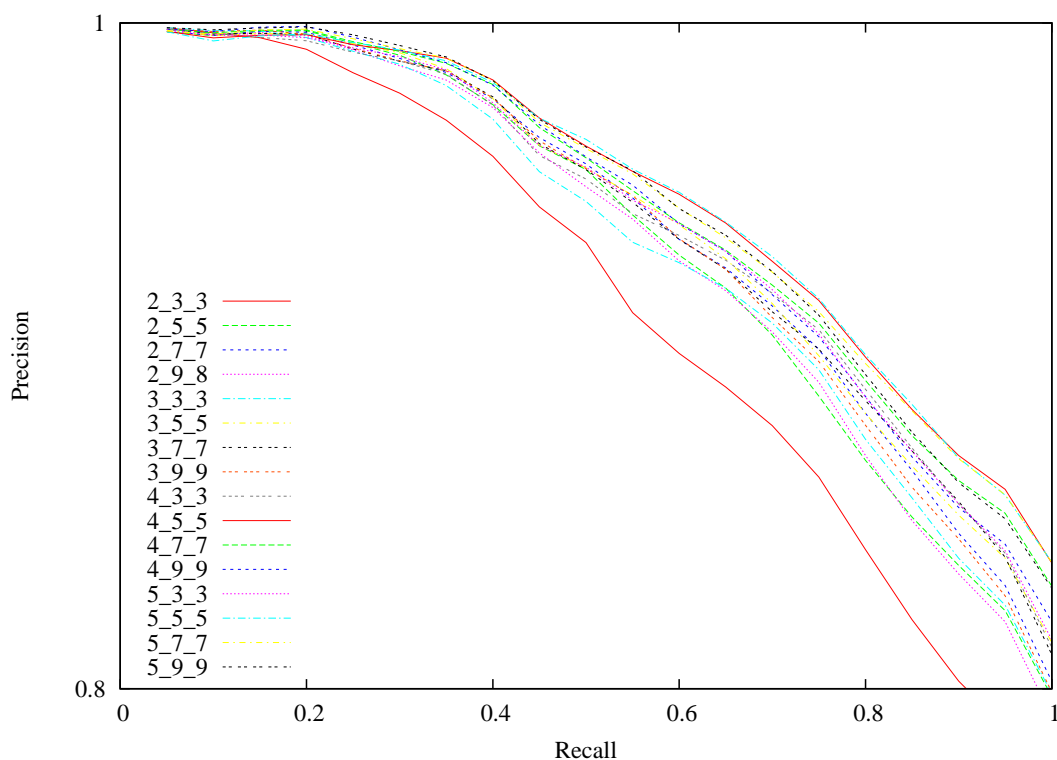
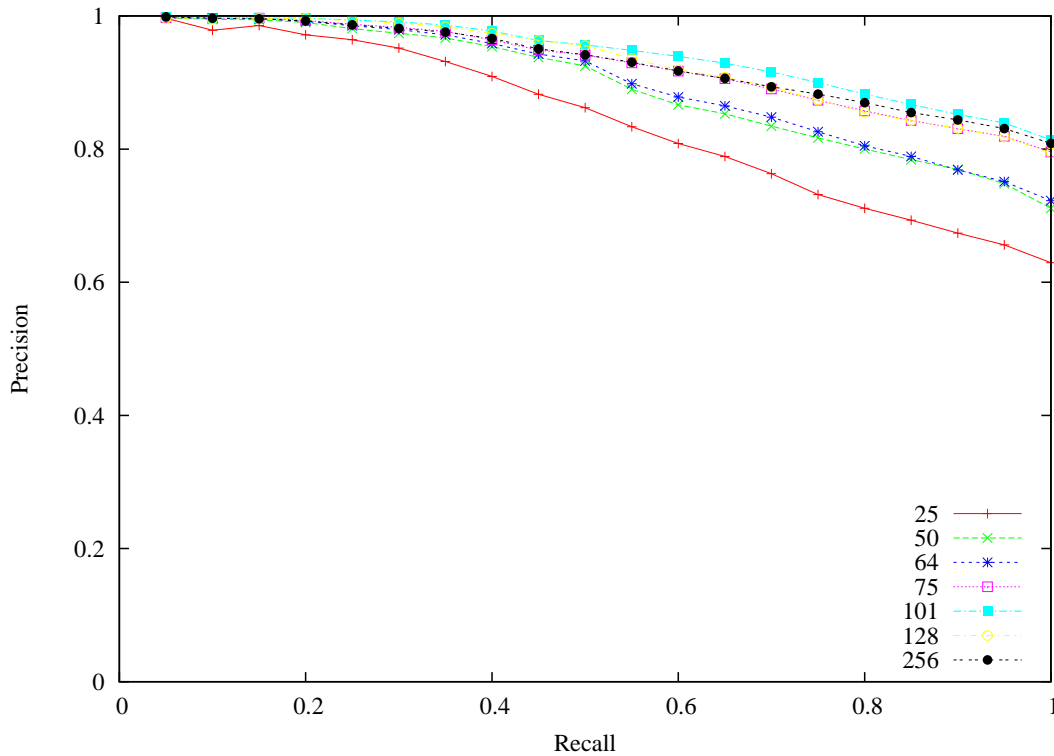


FIG. 5.4 – Courbe de Recall/Precision pour certaines valeurs des paramètres n , m_θ et m_ϕ .

de 3. La distance obtenue est inférieure à la distance intra classe et la classification sera bonne. Le test de déformation aux rotations crée un ensemble d'objets perturbés suivant trois rotations aléatoires selon les axes et donne une distance maximum de 1.272 et une distance moyenne de 0.75. La déformation due au bruit est un déplacement aléatoire des sommets de l'objet ; chaque sommet est déplacé suivant un vecteur gaussien aléatoire. La longueur moyenne des vecteurs de déplacement est un pourcentage de la taille de l'objet. Si la distance est de 5%, respectivement 10%, la surface de l'objet est très déformée, mais la mesure de similarité moyenne est de 1.012, respectivement 1.6, et l'objet est encore bien classé. La figure 5.6 montre un objet déformé par ce procédé.

	Translation	Variation d'échelle	Rotation	Bruit
Nombre d'objets	100	30	109	100
Distance maximum	0	0.016	1.272	2.217
Distance moyenne	0	0.003	0.750	1.012

TAB. 5.3 – Distances obtenues pour plusieurs déformations.

FIG. 5.5 – Courbes de Recall/Precision pour plusieurs valeurs du paramètre S .

5.2.3.3 Comparaison aux harmoniques sphériques

Un autre ensemble d'expérimentations a été fait pour comparer la description ART 3D à celle par des harmoniques sphériques (HS). Cette expérience a été réalisée sur les deux bases d'objets. La figure 5.7 et le tableau 5.4 montrent les courbes de Recall/Precision et les critères Nearest Neighbor, First Tier et Second Tier, pour les descriptions HS et ART 3D sur les bases d'objets 3D Renault et Princeton Shape Benchmark. L'implémentation utilisée pour expérimenter les descripteurs en Harmoniques Sphériques est celle disponible sur la page Web des auteurs². Ceci nous permet de considérer que les temps d'exécution des deux méthodes peuvent être comparées.

Sur la base Princeton, les HS donnent une meilleure description que la méthode ART. Sur la base Renault les résultats sont sensiblement équivalents suivant les deux méthodes de description. Ceci montre que la description ART fonctionne efficacement lorsque la distance intra classe est faible. Le coût de calcul et la taille des descripteurs sont aussi un point significatif pour comparer des méthodes (Table 5.5). Le coût de l'indexation avec la méthode ART 3D est 2.5 fois plus rapide que l'indexa-

²<http://www.cs.jhu.edu/~misha/HarmonicSignatures/>

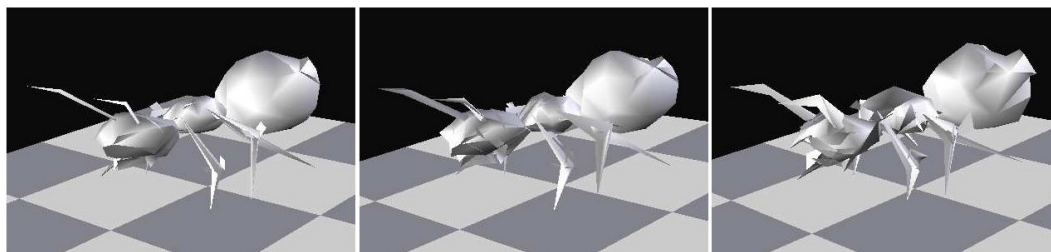


FIG. 5.6 – Exemple de déformation due au bruit pour trois valeurs de distance : 0%, 5% et 10%.

Base	Méthode	Nearest Neighbor	First Tier	Second Tier
Renault 758	ART 3D	0.993	0.893	0.928
	HS 3D	0.993	0.921	0.942
Renault 5000	ART 3D	0,964	0,770	0,844
	HS 3D	0,946	0,804	0,910
Princeton	ART 3D	0.462	0.238	0.321
	HS 3D	0.598	0.324	0.437

TAB. 5.4 – Comparaison des performances des méthodes ART 3D et des Harmoniques sphériques en terme de Nearest Neighbor, First Tier et Second Tier, sur les deux bases Renault et sur la base Princeton.

tion des HS et le coût du calcul de la mesure de similarité est 7.8 fois moindre pour ART. Ces différences sont dues au fait que la description ART ne fait pas de transformation fréquentielle et reste dans l'espace réel contrairement aux harmoniques sphériques qui procèdent à une transformation fréquentielle.

La transformation ART 3D, que nous proposons et le processus d'indexation fournissent un descripteur de formes robustes aux translations, rotations, variations d'échelle, multi-représentations (remaillage, faible déformation), bruit et rotation 3D. Le descripteur de formes remplit les conditions induites par l'analyse de base de modèles techniques : processus d'indexation et de recherche robuste, précis et calcul

	Temps d'indexation	Taille du descripteur
HS	10	544
3D ART	4	74

TAB. 5.5 – Comparaison de la taille (en nombre flottant) et du temps d'indexation par objet (en secondes) entre ART 3D et la représentation en harmoniques sphériques (HS).

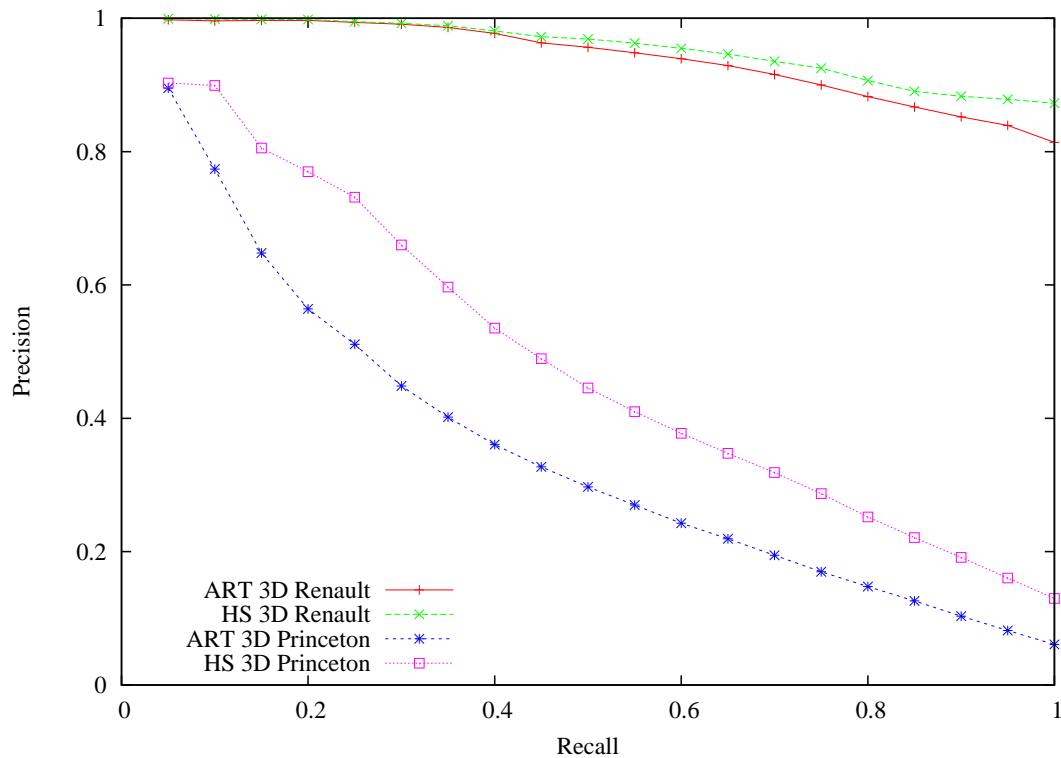


FIG. 5.7 – Courbes de Recall/Precision des méthodes ART 3D et des harmoniques sphériques sur les bases d'objets 3D Renault et Princeton.

de similarités rapide.

5.2.3.4 Comparaison au sein de Semantic-3D

Dans le cadre du projet RNRT Semantic-3D, une étude comparative des méthodes proposées par les trois partenaires impliqués en indexation a été réalisée. Quatre méthodes d'indexation 3D/3D ont été évaluées :

- le descripteur de Hough 3D optimisé, DH3DO, (ZAHARIA et PRÊTEUX, 2004) ;
- le descripteur de Hough 3D optimisé avec quantification vectorielle (DH3DO-QV), toutes deux proposées par le partenaire INT, (ZAHARIA et PRÊTEUX, 2004) ;
- le descripteur à base de cluster de vues adaptatives où « Adaptive Views Clustering »(AVC), qui caractérise un objet 3D par un ensemble de vues 2D, proposé par le partenaire ENIC (FILALI ANSARY ET AL., 2005) ;
- le descripteurs ART 3D.

Les quatre méthodes retenues par les trois partenaires ont été évaluées en terme algorithmique et en terme de pertinence aux requêtes par similarité selon les critères

DF	T	V_{type}	CE	CMS	IG	IT	G
DH3DO	N_{3D}	1280	$O(N)$	$O(N3D)$	Oui	Oui	Oui
DH3DO-QV	N_{2D}	128	$O(N)$	$O(N_{2D})$	Oui	Oui	Oui
ART 3D	N_{1D}	74	$O(N.D^3)$	$O(N_{1D})$	Non	Oui	Oui
AVC	$V.N_{1D}$	49	$O(320.D^3)$	$O(V^2N_{1D})$	Non	—	—

TAB. 5.6 – Comparaison des descripteurs de forme (DF) et performances associées. N_{iD} est le nombre d'éléments d'un vecteur iD , $i \in \{1, 2, 3\}$, V_{type} désigne les valeurs classiquement associées aux paramètres N_{iD} , N le nombre des facettes du maillage, D la taille de la discrétisation et V le nombre de vues nécessaires pour les approches 2D/3D.

d'évaluation définis au sein du projet et énoncés dans la partie 2.7.

Comparaison par critères algorithmique. Le tableau 5.6 synthétise la façon dont les différents descripteurs de forme 3D considérés répondent aux divers critères algorithmique retenus (paragraphe 2.7.2).

En terme de taille, le descripteur ART 3D est le plus compact avec une représentation sur seulement 74 flottants. Une compacité comparable (128 indices entiers) est acquise par le DH3DO-QV, avec l'avantage d'une complexité de calcul de la mesure de similarité réduite, grâce aux distances pré-calculées entre vecteurs prototypes, mais ce pré-calcul des mesures de similarité peut être réalisé pour n'importe quelle méthode d'indexation et il n'est donc pas à prendre en compte dans la comparaison. Les descripteurs par transformée de Hough bénéficient d'un algorithme d'extraction rapide, de complexité linéaire avec le nombre des facettes du maillage. Quant au descripteur AVC 2D/3D, il est pénalisé par le nombre important de vues (320) nécessaires pour les requêtes 3D/3D, aussi bien en taille qu'en temps de calcul de la mesure de similarité associée. Dans leur ensemble, les descripteurs proposés sont généraux et peuvent s'appliquer à des formes de géométrie et de topologie arbitraires.

Performances des requêtes par similarité. En ce qui concerne l'évaluation des performances des requêtes par similarité, les valeurs globales des scores retenus, sur l'ensemble des catégories sont synthétisées dans le tableau 5.7.

Ces scores globaux montrent des performances très élevées pour l'ensemble des descripteurs retenus. En ce qui concerne les approches 3D/3D, les performances du DH3DO sont supérieures à celles correspondant au descripteur ART 3D, avec un gain de 4% (resp. 7%) en terme de score SBE (resp. FT). La DH3DO-QV présente une dégradation mineure des performances (1-2%), avec l'avantage d'une taille et d'un temps de calcul de la mesure de similarité considérablement réduits. Enfin, notons les excellentes performances du descripteur 2D/3D par moments de Zernike

Approche	DH3DO	DH3DO-QV	ART 3D	AVC
SBE	0,97	0,96	0,92	0,95
FT	0,95	0,93	0,89	0,95
SBE_{tot}	0,94	0,92	0,84	0,94
FT_{tot}	0,86	0,83	0,78	0,92

TAB. 5.7 – Performances des différents descripteurs de formes selon les critères : SBE , FT , SBE_{tot} et FT_{tot} .

dans le cadre de requêtes 3D/3D.

5.3 Indexation 3D par vues

Comme nous avons vu dans le chapitre 2, lors de l'état de l'art sur l'indexation d'objets 3D par requêtes 3D, une classe de méthodes caractérise les objets 3D par un ensemble de vues 2D et compare les ensembles de vues pour classer les objets 3D. Ces méthodes s'appellent les approches multi-vues, ou les méthodes d'indexation 3D par vues. Dans cette partie, nous avons étudié les possibilités d'utiliser le système de rendu spectral, que nous avons mis en place dans le chapitre 4, pour faire de l'indexation d'objets 3D par vues.

Le système de rendu spectral, basé sur l'extraction de coupes dans le Fourier discret des objets 3D, peut fournir des vues des objets possédant plus d'information que des vues créées par un processus basé sur le lancé de rayon. Dans le chapitre 4, nous comparons les vues extraites à une image requête binaire et il faut que les vues extraites soient binaires pour être correctement comparées. Pour la comparaison d'objets 3D par une processus 3D/3D par vues, il n'est pas obligatoire de comparer des vues binaires et nous allons évaluée l'hypothèse selon laquelle les images profondeurs portent plus d'information dans un processus d'indexation 3D/3D par vues.

5.3.1 Indexation

Les méthodes d'indexation d'objets 3D par vues (paragraphe 2.5) se différencient en fonction du choix du nombre et de la position des vues caractéristiques, des descripteurs de forme 2D et de la mesure de similarité mise en place entre les ensembles de vues caractéristiques. Ces méthodes créent les vues d'un objet 3D par un processus de Z-buffer et, en fonction de la méthode de description, caractérisent des vues binaires ou des images luminances entre elles. Ces vues possèdent une certaine quantité d'information utile pour les comparer, mais sont moins caractéristiques et moins porteuses d'information que les vues profondeurs.

On est en droit de penser, qu'utiliser les vues profondeurs de l'objet permettra de caractériser plus précisément la forme des objets et par la même de mettre en place un processus de comparaison plus rapide en limitant le nombre de vues caractéristiques.

5.3.1.1 Création des vues caractéristiques

Le système de rendu spectral extrait les coupes représentant la transformée de Fourier des vues de l'objet, comme nous l'avons vu dans la paragraphe 4.3.1. La description des vues caractéristiques peut se faire, soit en décrivant le spectre de l'image profondeur, soit l'image profondeur reconstruite ou l'image binaire seuillée. La figure 4.5 du chapitre 4 montre les différentes vues qu'il est possible d'extraire du système de rendu spectral que nous avons mis en place.

Les vues créées par le processus de rendu spectral sont des vues orthographiques. Il est plus performant de caractériser les objets 3D par des vues orthographiques, car cela permet de diminuer le nombre de vues à caractériser en se limitant à une demi-sphère des vues. En effet, les deux images symétriques par rapport au centre de la sphère sont identiques, ce qui n'est pas le cas avec un processus de rendu perspectif.

5.3.1.2 Position des vues caractéristiques

Pour évaluer les performances des descriptions en fonction du type de vues sélectionné, nous avons mis en place un système de vues caractéristiques fixes, basé sur trois ensembles de vues :

- 3 vues, placées sur les 3 axes principaux de l'objet,
- 7 vues placées sur les trois axes principaux et sur les quatre diagonales,
- 25 vues « uniformément » réparties sur la demi-sphère des vues par une méthode Sphère Picking, présenté dans le paragraphe 1.3.2.2, avec $x_1^2 + x_2^2 \leq 0.5$ et x_1 et $x_2 \in [-1, -0.75, -0.5, \dots, 1]$.

La figure 5.8 montre ces trois ensembles de vues.

5.3.1.3 Mesure de similarité entre deux ensembles de vues

La mesure de similarité que nous avons mise en place entre deux ensembles de vues, recherche le meilleur appariement possible. Si deux objets sont similaires, il est possible de trouver, pour chaque vue du première objet, une vue appartenant à l'ensemble des vues du second lui ressemblant. La mesure de similarité doit permettre de détecter le meilleur appariement possible. Pour chaque vue de l'ensemble E_1 , elle doit trouver la vue de l'ensemble E_2 fournissant la distance la plus petite. Il faut noter qu'il ne faut pas apparier une vue de E_2 à plusieurs vues de E_1 . Comme énoncé dans le chapitre 2.5, ceci revient à rechercher les N appariement minimums, où N

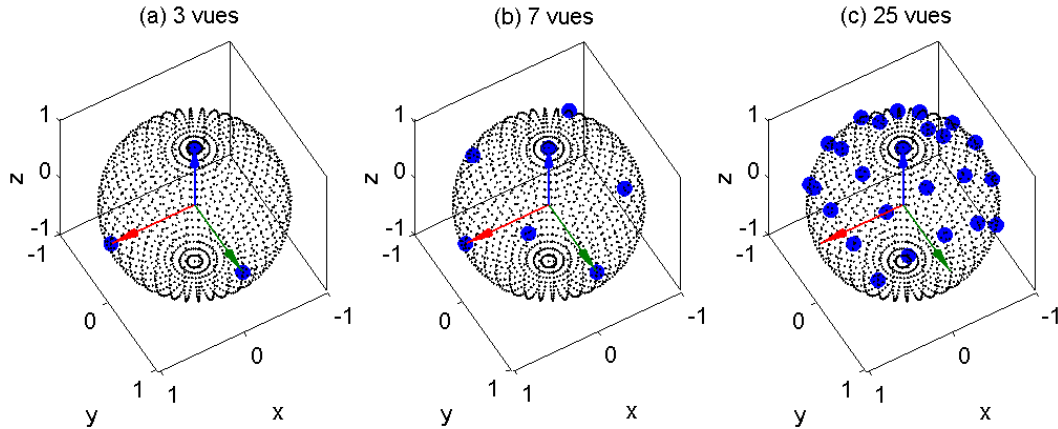


FIG. 5.8 – Position des vues pour les trois systèmes de vues caractéristiques fixe : 3, 7 et 25.

est le nombre de vues caractéristiques :

$$D(E_1, E_2) = \sum_{i=0}^N \min_{j \in \overline{E_2}} d(E_{1i}, E_{Ej}) \quad (5.7)$$

où $\overline{E_2}$ est l'ensemble E_2 privé des vues déjà utilisées par une vue précédente. Le coût de la mesure de similarité est en $O(N^2)$.

5.3.1.4 Processus de recherche

Les méthodes d'indexation d'objets 3D par vues caractérisent les objets par un ensemble de vues 2D. La méthode de rendu spectral présentée dans le chapitre 4, permet d'en extraire plusieurs types (spectral, reconstruite, binaire). La figure 5.9 représente le schéma global de la comparaison d'un objet 3D requête à une base d'objets.

5.3.2 Expérimentations

L'expérimentation que nous avons menée, compare les trois types de vues : binaire, profondeur et spectral, pour les trois ensembles de vues caractéristiques basés sur 3, 7 et 25 vues. Les tests ont été réalisés sur les bases d'objets 3D RNRT Semantic-3D Renault de 116 et de 758 objets (cf. paragraphe 2.7.1). Les tableaux 5.8 et 5.9 montrent les résultats en termes de Nearest Neighbor, First Tier et Second Tier (cf. paragraphe 2.7.2). Les tableaux contiennent également les tailles des descripteurs et les résultats des méthodes d'indexation d'objets 3D : ART 3D (cf. chapitre 5.2) et les harmoniques sphériques (cf. paragraphe 2.4.2.4) à titre de comparaison.

Nombre de vues	Type de vues	Nearest Neighbor	First Tier	Second Tier	Taille des descripteurs
3	Binaire	1	0.979	0.979	108
	Profondeur	1	0.960	0.976	108
	Spectral	1	0.903	0.937	108
7	Binaire	1	0.960	0.968	252
	Profondeur	1	0.951	0.971	252
	Spectral	1	0.950	0.965	252
25	Binaire	1	0.956	0.958	900
	Profondeur	1	0.944	0.969	900
	Spectral	1	0.936	0.964	900
ART 3D		0.982	0.923	0.945	74
HS 3D		0.984	0.944	0.964	544

TAB. 5.8 – Statistique sur la base RNRT Semantic-3D Renault contenant 116 objets répartis en 25 classes.

Nombre de vues	Type de vues	Nearest Neighbor	First Tier	Second Tier	Taille des descripteurs
3	Binaire	1	0.900	0.923	108
	Profondeur	1	0.876	0.911	108
	Spectrale	1	0.876	0.917	108
7	Binaire	1	0.896	0.933	252
	Profondeur	1	0.881	0.915	252
	Spectrale	1	0.873	0.913	252
25	Binaire	1	0.876	0.915	900
	Profondeur	1	0.871	0.909	900
	Spectrale	1	0.848	0.883	900
ART 3D		0.993	0.893	0.928	74
HS 3D		0.993	0.921	0.942	544

TAB. 5.9 – Statistique sur la base RNRT Semantic-3D Renault contenant 758 objets répartis en 75 classes.

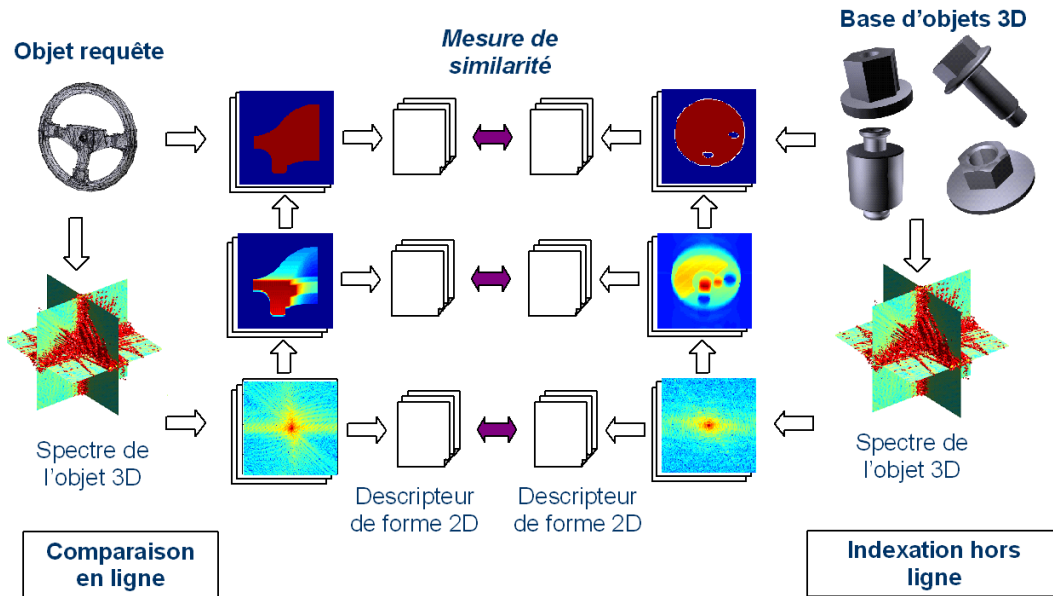


FIG. 5.9 – Processus de comparaison d'un objet 3D aux objets d'une base, en comparant les trois types de vues créées par le processus de rendu spectral : le spectre profondeur est extrait du spectre 3D de l'objet, l'image profondeur est créé par une transformée de Fourier inverse, qui peut être binarisée pour créer l'image binaire. Cette dernière peut également être créée directement à partir de l'objet 3D par un processus de rendu basé sur le Z-Buffer.

Ces résultats montrent la grande performance des méthodes d'indexation d'objets 3D basées sur les vues. Ces méthodes, bien que relativement simples fournissent des résultats meilleurs que les méthodes basées sur les moments : ART 3D et HS 3D. Par contre, il faut garder un nombre de vues caractéristiques relativement petit pour rester compétitif, en terme de taille de descripteurs et par la même en terme de coût de la mesure de similarité par rapport aux descripteurs de formes basés sur les moments.

Par contre, contrairement à nos attentes, l'information supplémentaire fournie par les vues profondeurs, ou par les spectres des vues profondeurs, n'est pas correctement utilisée et ne permet pas de fournir des résultats supérieurs que ceux obtenus avec les vues binaires. On peut comprendre que bien que contenant plus d'information, le fait d'indexer les vues profondeurs ne fournit pas de meilleurs résultats car la mesure de similarité et le descripteur de formes utilisé ne semble pas prendre en compte l'information supplémentaire. Il en est de même pour les spectres des vues profondeurs, qui fournissent des résultats inférieurs que ceux obtenus avec les vues binaires. Ceci peut s'expliquer par le fait que le descripteur de formes utilisé, ART, n'est pas un descripteur spectral. Il serait intéressant de mettre en place un descripteur spécialement pour ce type de vues.

Les résultats montrent également, que les performances des systèmes de vues caractéristiques décroissent à mesure que le nombre de vues augmente. On peut voir que entre 3 et 25 vues, le First Tier passe de 97% à 95%, pour la base de 116 objets. Ceci peut être dû à la mesure de similarité et aux descripteurs de formes choisis qui ne permettent pas de sommer correctement les contributions de chaque vue et rajoutent seulement du bruit sans tenir compte de l'information significative.

5.4 Conclusion

Pour répondre aux besoins définis dans le cadre du projet RNRT Semantic-3D, nous avons proposé un nouveau descripteur de formes 3D basé sur les moments pour indexer des grandes bases de données d'objets, ART 3D. La rapidité d'indexation et de recherche, requise par la problématique d'indexation de grandes bases d'objets, a été le critère principal qui a guidé la mise en place de ce descripteur. Le nouveau descripteur de formes 3D a montré au cours des différents tests réalisés, sa grande rapidité et sa bonne précision par rapport aux principaux descripteurs d'objets 3D existant.

D'autre part, nous avons utilisé le système de rendu spectral, que nous avons développé dans le chapitre 4, pour faire de l'indexation d'objets 3D par vues. Les expériences que nous avons réalisées ont montré le fort potentiel de ces méthodes et leur facilité d'utilisation et de mise en oeuvre.

Nous avons également montré que l'utilisation des images profondeurs, ou que l'accroissement du nombre de vues caractéristiques, ne permet pas d'avoir des résultats meilleurs. Bien qu'augmentant la quantité d'information, la mesure de similarité ne parvient pas à correctement sommer le peu d'information supplémentaire fournie par la description. Dans ces conditions, l'ajout d'information non utilisée peut être vue comme une augmentation du bruit.

Les perspectives envisageables de ces travaux sont de deux ordres. Le travail sur ART 3D a montré les performances de ce descripteur et il serait intéressant de l'utiliser sur d'autres bases d'objets et dans d'autres secteurs applicatifs. D'autre part, l'étude que nous avons réalisée sur l'indexation d'objets 3D par vues ouvre de nombreuses perspectives à moyen et long terme. D'une part, il serait intéressant d'approfondir l'étude en modifiant la métrique et les descripteurs choisis pour caractériser les images extraites. Ceci permettra d'augmenter les performances de la recherche. De plus, il paraît prometteur de définir un descripteur de formes applicable sur les spectres de Fourier extraits et non sur les images reconstruites. Cela permettrait de ne pas avoir à reconstruire les images et de mesurer la similarité directement dans le domaine spectral.

6

Expression analytique d'un maillage 3D

Sommaire

6.1	Introduction	130
6.2	Cadre de l'analyse	130
6.2.1	Principe général	130
6.2.2	Intégration sur un maillage 3D	132
6.3	Intégration d'un tétraèdre	133
6.3.1	Principe	133
6.3.2	Calcul des moments géométriques	134
6.3.3	Expression analytique de la transformée de Fourier	135
6.4	Application en analyse de formes 3D	139
6.4.1	Centrage et alignement d'un modèle 3D	139
6.4.2	Descripteur de forme 3D	144
	Descripteur de Fourier 3D	144
	Descripteur de Zernike 3D	144
6.4.3	Complexité	145
6.5	Conclusion	145

6.1 Introduction

Une grande partie des méthodes d'indexation de modèles 3D calcule les descripteurs de formes en intégrant une fonction sur le volume de l'objet. Pour cela, l'objet est dans un premier temps discrétisé sur une grille de voxels et l'intégration est faite dans le domaine discret, comme dans les chapitres 2, 4, 5 et 5.3. C'est le cas par exemple de l'analyse en composantes principales, des moments géométriques, des moments de Zernike, ou de la transformée de Fourier.

Cette partie présente une méthode permettant d'intégrer un objet 3D sans passer par une représentation discrète mais directement à partir de son maillage. Ces méthodes, qui se basent sur l'intégration du contour de l'objet, peuvent, comme nous le montrerons dans la dernière partie de cette section, servir au calcul des paramètres d'alignement, à l'extraction de descripteurs de forme ou à la création de vues par une méthode de Fourier Volume Rendering.

6.2 Cadre de l'analyse

6.2.1 Principe général

Le volume, le centre de gravité, les moments d'inertie et d'autres propriétés d'un solide sont définis par une intégrale triple dans un espace à trois dimensions définissant le volume de l'objet. L'intégrale d'une fonction f sur le domaine Q définit par le volume s'écrit :

$$I = \int_Q f dv \quad (6.1)$$

Cette fonction est difficilement intégrable, car même si la fonction à intégrer f est simple, le domaine d'intégration Q représentant le volume de l'objet est complexe. Pour intégrer sur une volume quelconque, si f est intégrable, le théorème de la divergence produit une méthode alternative pour évaluer les propriétés de l'intégrale d'un solide en intégrant sur les frontières du solide :

$$\int_Q f dv = \int_Q \text{div}(g) dv = \int_{\delta Q} g \cdot n ds \quad (6.2)$$

où g est une fonction vecteur, telle que la divergence de g est égale à la fonction f , δQ est la frontière de Q , n est le vecteur normal de la frontière et ds est la dérivée de la surface.

L'intégrale sur un volume 3D peut être vue comme la somme de l'intégrale de toutes les composantes frontières, comme le montre LIEN (1984) et ZHANG et CHEN (2001). Dans le cas d'objets 3D maillés, le maillage représente la frontière du solide et il est dans ce cas facile de le parcourir. Ceci est exact sous l'hypothèse que les

frontières de l'objet soient correctement définies et que le maillage soit une variété topologique (sans trous, orientable...). Cette méthode fournit l'expression analytique exacte de l'intégrale sur le volume de l'objet, pour une complexité linéaire au nombre de sommets du polyèdre.

Les éléments de surface de l'objet 3D peuvent être séparés en fonction du signe de leurs normales par rapport à un point fixe, généralement par rapport à l'origine du repère $(0, 0, 0)$, et l'on peut écrire :

$$\int_Q f(x, y, z)dv = \int_{\delta Q} g(x, y, z).nds = \int_{\delta Q_1} g(x, y, z).nds + \int_{\delta Q_2} g(x, y, z).nds \quad (6.3)$$

où δQ_1 est le sous ensemble de δQ tel que \vec{n} est dans le sens du vecteur (x, \vec{y}, z) , formé entre l'origine et le centre de l'élément de surface, tel que $(x, \vec{y}, z).\vec{n} > 0$ et δQ_2 est le sous ensemble de δQ tel que $(x, \vec{y}, z).\vec{n} \leq 0$ (voir figure 6.1).

On peut retransformer les deux termes pour retourner en volumique :

$$\begin{aligned} \int_{\delta Q_1} g(x, y, z).nds + \int_{\delta Q_2} g(x, y, z).nds &= \int_{Q_1} f(x, y, z)dv + \int_{Q_2} f(x, y, z)dv \\ &= \sum_{\Delta v_i} S(\Delta v_i) \int_{\Delta v_i} f(x, y, z)dv \quad (6.4) \end{aligned}$$

où Δv_i le cône formé entre l'élément de surface Δs_i et l'origine du repère, comme le montre la figure 6.1. $S(\Delta v_i)$ est une fonction de signe calculé en fonction du sens des normales \vec{n}_i :

$$\begin{aligned} S(\Delta v_i) &= +1 \text{ si } (x, \vec{y}, z).\vec{n}_i > 0, \\ &= -1 \text{ sinon.} \end{aligned} \quad (6.5)$$

La fonction de signe $S(\Delta v_i)$ détermine si l'intégrale Δv_i a une contribution positive ou négative à l'intégrale de Q . Si \vec{n}_i est dans le même sens que (x, \vec{y}, z) , l'intégrale ΔV_1 a une contribution positive à Q , autrement elle a une contribution négative.

Ceci peut être illustré par un exemple, comme le montre la figure 6.2. Le tétraèdre $Q = (v_1, v_2, v_3, v_4)$ à quatre faces, $F_1 = (v_1, v_4, v_2)$, $F_2 = (v_1, v_3, v_4)$, $F_3 = (v_4, v_3, v_2)$ et $F_4 = (v_1, v_3, v_2)$. L'ordre des sommets de chaque face est défini par la règle de la main droite, afin que le sens des normales de tous les triangles soit cohérent. La face F_i de Q forme avec l'origine O un tétraèdre $Q_i = (O, v_{i_1}, v_{i_2}, v_{i_3})$ où $(v_{i_1}, v_{i_2}, v_{i_3})$ sont les sommets de F_i . Quatre nouveaux tétraèdres, $Q_1 = (O, v_1, v_4, v_2)$, $Q_2 = (O, v_1, v_3, v_4)$, $Q_3 = (O, v_4, v_3, v_2)$ et $Q_4 = (O, v_1, v_3, v_2)$, sont ainsi formés. On peut noter que le tétraèdre Q_4 est positif selon S et que les tétraèdres Q_1 , Q_2 et Q_3 sont négatif. Le tétraèdre Q est la somme signée des quatre tétraèdres Q_1 , Q_2 , Q_3 et Q_4 :

$$Q = Q_4 - Q_1 - Q_2 - Q_3. \quad (6.6)$$

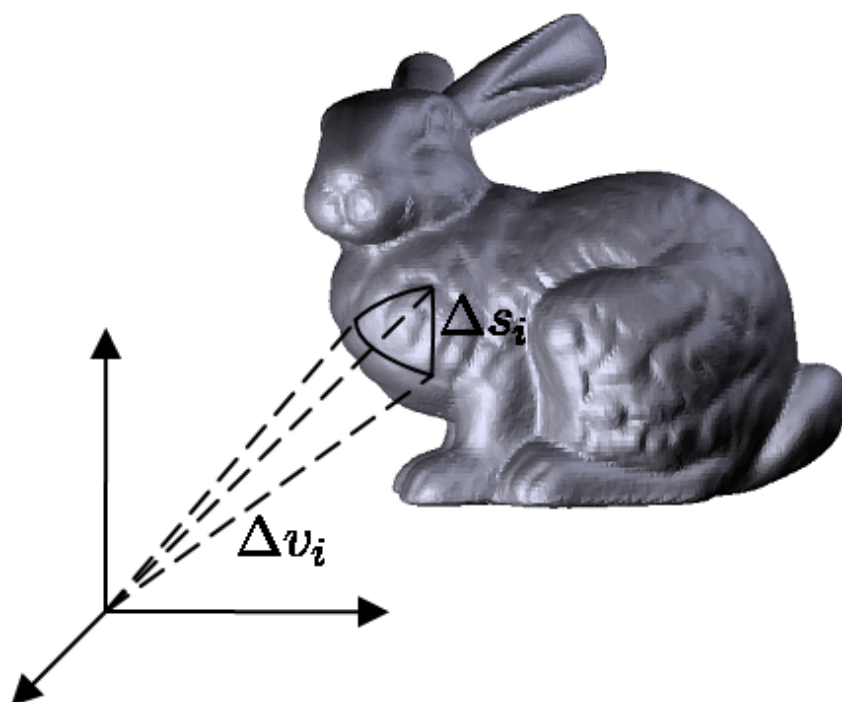


FIG. 6.1 – Principe de l'intégration, un cône Δv est formé par un élément de surface Δs est l'origine du repère.

6.2.2 Intégration sur un maillage 3D

Comme nous venons de le voir, le volume d'un objet 3D est égal à la somme signée des volumes des tétraèdres le composant et il en est de même pour les autres fonctions intégrales que l'on peut calculer sur un objet 3D. Pour calculer la valeur de l'intégrale d'une fonction $f(x, y, z)$ sur le volume d'un objet maillé, on peut procéder de la manière suivante :

1. orientation des faces : vérifier que le maillage est une variété combinatoire sans bord, et orienter les faces pour garantir que les normales des triangles adjacents sont cohérentes du point de vue de leur sens ;
2. pour chaque triangle de l'objet :
 - intégration de la fonction $f(x, y, z)$ sur le tétraèdre construit à partir du triangle et de l'origine du repère ;
 - sommation de la valeur de l'intégrale en fonction du signe de la fonction de signe S .

Pour orienter les normales et calculer les fonctions de signe S , le sens de la normale de chaque triangle adjacent est vérifié. Ceci revient à parcourir tous les triangles de l'objet et à vérifier que l'ordre des points fournit une normale dans le bon sens en fonction des triangles adjacents déjà parcourus.

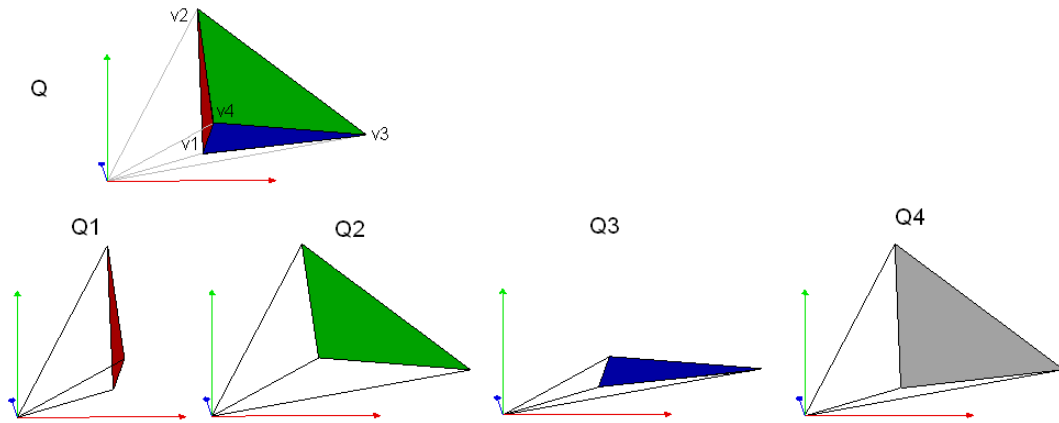


FIG. 6.2 – Exemple de décomposition tétraédrique d'un objet 3D. L'objet Q peut être vu comme la somme CSG : $Q = Q_4 - Q_1 - Q_2 - Q_3$.

6.3 Intégration d'un tétraèdre

6.3.1 Principe

Comme énoncé précédemment, le calcul d'une intégrale sur un objet tridimensionnel peut se faire en additionnant les contributions signées de l'intégrale de chaque tétraèdre. Ce qui revient à sommer les volumes des tétraèdres le composant :

$$I = \sum_{Q_i} S(Q_i) \int_{Q_i} f(x, y, z) dv. \quad (6.7)$$

Soit F_i un triangle composé des points (x_{i1}, x_{i2}, x_{i3}) , (y_{i1}, y_{i2}, y_{i3}) et (z_{i1}, z_{i2}, z_{i3}) . On peut définir une transformation linéaire T_i telle que :

$$T_i = \begin{pmatrix} x_{i1} & x_{i2} & x_{i3} \\ y_{i1} & y_{i2} & y_{i3} \\ z_{i1} & z_{i2} & z_{i3} \end{pmatrix} \quad (6.8)$$

qui transforme le tétraèdre unité orthogonale (tétraèdre formé de l'origine et des points $(1, 0, 0)$, $(0, 1, 0)$ et $(0, 0, 1)$), en le tétraèdre défini par $(0, 0, 0)$ et les faces de F_i . On peut montrer que le déterminant de la matrice T_i est de même signe que la fonction de signe S . Cela conduit à :

$$\int_{Q_i} f(x, y, z) dx dy dz = \int_{W_i} \|T_i\| f(X, Y, Z) dX dY dZ \quad (6.9)$$

où $\|T_i\|$ est la norme de la matrice de transformation T_i et W_i est le tétraèdre unité orthogonal selon la transformation :

$$\begin{cases} x &= x_{i1}X + x_{i2}Y + x_{i3}Z \\ y &= y_{i1}X + y_{i2}Y + y_{i3}Z \\ z &= z_{i1}X + z_{i2}Y + z_{i3}Z \end{cases} \quad (6.10)$$

On peut noter que l'intégration d'une fonction f sur le tétraèdre Q_i est égale au produit de l'intégrale de la fonction f sur le tétraèdre unité multiplié par la norme de la matrice de transformation $\|T_i\|$. D'autre part, le déterminant $|T_i|$ représente implicitement les effets de la fonction de signe $S(Q_i)$ et de la matrice $\|T_i\|$, ce qui peut s'écrire :

$$S(Q_i) \cdot \|T_i\| = \|T_i\| \quad (6.11)$$

Sous cette condition, l'intégrale d'une fonction f dévient :

$$I = \sum_{Q_i} |T_i| \int_{W_i} f(X, Y, Z) dX dY dZ \quad (6.12)$$

L'intégrale sur le volume d'un tétraèdre quelconque peut être calculée en se ramenant au tétraèdre unité orthogonale formé des points $(0, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$ et $(0, 0, 1)$. Une intégrale sur ce tétraèdre permet d'avoir un domaine intégration Q simple et facilement intégrable. Il suffit d'intégrer z sur $[0, 1]$, y sur $[0, 1-z]$ et x sur $[0, 1-z-y]$ et on peut ainsi écrire :

$$I = \int_0^1 \int_0^{1-z} \int_0^{1-z-y} f(X, Y, Z) dX dY dZ \quad (6.13)$$

6.3.2 Calcul des moments géométriques

Comme nous l'avons présenté dans le chapitre 2, les moments géométriques tri-dimensionnels sont largement utilisés en indexation d'objets 3D, soit pour calculer des paramètres de normalisation (centre de gravité, axes principaux), soit comme descripteurs comme l'ont montré PAQUET et RIOUX (1999).

Les moments géométriques d'ordre $(p + q + r)$, notés par m_{pqr} , peuvent être calculés par :

$$m_{pqr} = \int_Q x^p y^q z^r dx dy dz \quad (6.14)$$

où Q est l'objet. La façon la plus répandue de calculer une telle intégrale est de discrétiser l'objet et de procéder dans l'espace discret par la formule :

$$m_{pqr} = \sum_{i=-N/2}^{N/2} \sum_{j=-N/2}^{N/2} \sum_{k=-N/2}^{N/2} i^p j^q k^r h(i, j, k) \quad p, q, r = 0, 1, 2, 3, 4 \dots \quad (6.15)$$

où N est la taille de la discrétisation et $h(i, j, k) \rightarrow \{0, 1\}$ est une fonction définissant l'objet (intérieur/extérieur).

Une autre manière de procéder est d'intégrer la surface. Soit un tétraèdre Q composé des points $(0, 0, 0)$, (x_1, y_1, z_1) , (x_2, y_2, z_2) et (x_3, y_3, z_3) , le calcul de l'expression

analytique des moments géométriques peuvent se calculer en intégrant :

$$\begin{aligned} m_{pqr} &= \int_Q f(x, y, z) dx dy dz \\ &= |T| \int_0^1 \int_0^{1-Z} \int_0^{1-Z-Y} f(X, Y, Z) dX dY dZ \end{aligned} \quad (6.16)$$

avec :

$$f(X, Y, Z) = (x_1 X + x_2 Y + x_3 Z)^p (y_1 X + y_2 Y + y_3 Z)^q (z_1 X + z_2 Y + z_3 Z)^r \quad (6.17)$$

Soit l'expression du moment géométrique d'ordre $(0, 0, 0)$, la fonction à intégrer est égale à : $f(x, y, z) = x^0 y^0 z^0 = 1$. L'expression analytique d'un tel moment le long d'un tétraèdre unité peut se calculer :

$$\begin{aligned} m_{000} &= \int_0^1 \int_0^{1-z} \int_0^{1-z-y} dx dy dz \\ &= \int_0^1 \int_0^{1-z} (1-z-y) dy dz \\ &= \int_0^1 \frac{1}{2} z^2 - z + \frac{1}{2} dz = \frac{1}{6} \end{aligned} \quad (6.18)$$

Ceci est applicable quel que soit l'ordre des moments géométriques. Par exemple, les moments d'ordres inférieurs à 2 ont pour valeurs :

$$\begin{aligned} m_{000} &= \frac{1}{6} (x_1 y_2 z_3 - x_1 y_3 z_2 - y_1 x_2 z_3 + y_1 x_3 z_2 + z_1 x_2 y_3 - z_1 x_3 y_2) \\ m_{100} &= \frac{m_{000}}{4} (x_1 + x_2 + x_3) \\ m_{110} &= \frac{m_{000}}{20} (x_1(2y_1 + y_2 + y_3) + x_2(y_1 + 2y_2 + y_3) + x_3(y_1 + y_2 + 2y_3)) \\ m_{200} &= \frac{m_{000}}{10} (x_1^2 + x_2^2 + x_3^2 + x_1 x_2 + x_1 x_3 + x_2 x_3) \end{aligned} \quad (6.19)$$

Nous verrons dans la dernière partie, les applications que peuvent avoir les moments géométriques en analyse de formes tridimensionnelles et les avantages que nous apporte le fait de les calculer de la sorte.

Le calcul des moments géométriques a été réalisé grâce à un logiciel de calcul formel Maple, pour tous les moments géométriques jusqu'à l'ordre 20. Le code source des formules analytiques des moments géométriques est disponible à l'adresse : <http://liris.univ-lyon1.fr/julien.ricard/publications.html>

6.3.3 Expression analytique de la transformée de Fourier

La transformée de Fourier est également un outil largement utilisée en indexation d'objets 3D, comme nous l'avons présenté dans le chapitre 2. Elle est basée sur une

intégrale du volume de l'objet

$$F_{uvw} = \int_Q e^{-i(xu+yv+zw)} dx dy dz. \quad (6.20)$$

La transformée de Fourier est aussi une intégration sur le volume interne d'un objet et peut être calculée en décomposant l'objet en tétraèdre et en sommant les contributions. Ceci peut se calculer selon la formule :

$$F(u, v, w) = \sum_{Q_i} |T_i| \int_{W_i} f(X, Y, Z) dX dY dZ \quad (6.21)$$

L'expression analytique de la transformée de Fourier d'un tétraèdre peut être calculée en intégrant sur le tétraèdre unité orthogonale l'expression :

$$f(x, y, z) = e^{-i(u(x_1 X + x_2 Y + x_3 Z) + v(y_1 X + y_2 Y + y_3 Z) + w(z_1 X + z_2 Y + z_3 Z))} \quad (6.22)$$

L'expression générale à intégrer et la solution peuvent s'écrire :

$$\begin{aligned} F(u, v, w) &= |T| \int_W e^{-i(u(x_1 X + x_2 Y + x_3 Z) + v(y_1 X + y_2 Y + y_3 Z) + w(z_1 X + z_2 Y + z_3 Z))} dV \\ &= |T| \left[\frac{ie^{-iU_1}}{U_1(U_1-U_2)(U_1-U_3)} + \frac{ie^{-iU_2}}{U_2(U_2-U_1)(U_2-U_3)} \right. \\ &\quad \left. + \frac{ie^{-iU_3}}{U_3(U_3-U_2)(U_3-U_1)} - \frac{i}{U_1 U_2 U_3} \right] \end{aligned} \quad (6.23)$$

avec :

$$\begin{cases} U_1 = ux_1 + vy_1 + wz_1 \\ U_2 = ux_2 + vy_2 + wz_2 \\ U_3 = ux_3 + vy_3 + wz_3 \end{cases} \quad (6.24)$$

La formule 6.23 représente l'expression analytique de la transformée de Fourier d'un tétraèdre et permet de calculer la transformée de Fourier du maillage d'un objet 3D. Ceci est utilisé dans de nombreuses applications d'analyse de formes tridimensionnelles, comme nous l'avons vu dans les chapitres 2 et 4.

Cependant, cette expression n'est pas bien définie et ne peut pas être calculée pour tous les tétraèdres en tous les points de l'espace fréquentiel. En effet, cette expression est une somme de quotients qui ne sont pas définis lorsque les dénominateurs s'annulent. Il faut, dans ces cas, modifier la fonction à intégrer, $f(x, y, z)$ en fonction de la valeur de la solution et intégrer la nouvelle fonction. On peut expliciter 19 solutions S annulant la fonction dénominateur de $F(u, v, w)$:

$$F_{denum}(u, v, w) = U_1 U_2 U_3 (U_1 - U_2)(U_1 - U_3)(U_2 - U_3), \quad (6.25)$$

ce qui fournit les solutions suivantes :

$$\begin{aligned}
 S = & \{ \{u = 0, v = 0, w = 0\}, \\
 & \{u = 0, v = 0, z_1 = 0\}, \{u = 0, v = 0, z_2 = 0\}, \{u = 0, v = 0, z_3 = 0\}, \\
 & \{u = 0, v = 0, z_1 = z_2\}, \{u = 0, v = 0, z_2 = z_3\}, \{u = 0, v = 0, z_1 = z_3\}, \\
 & \left\{ u = 0, y_3 = -\frac{wz_3}{v} \right\}, \left\{ u = 0, y_2 = -\frac{wz_2}{v} \right\}, \left\{ u = 0, y_1 = -\frac{wz_1}{v} \right\}, \\
 & \left\{ u = 0, y_2 = \frac{vy_3 + wz_3 - wz_2}{v} \right\}, \\
 & \left\{ u = 0, y_1 = \frac{vy_3 + wz_3 - wz_1}{v} \right\}, \\
 & \left\{ u = 0, y_1 = \frac{vy_2 + wz_2 - wz_1}{v} \right\}, \\
 & \left\{ x_3 = -\frac{vy_3 + wz_3}{u} \right\}, \left\{ x_2 = -\frac{vy_2 + wz_2}{u} \right\}, \left\{ x_1 = -\frac{vy_1 + wz_1}{u} \right\}, \\
 & \left\{ x_2 = -\frac{-ux_3 - vy_3 - wz_3 + vy_2 + wz_2}{u} \right\}, \\
 & \left\{ x_1 = -\frac{-ux_3 - vy_3 - wz_3 + vy_1 + wz_1}{u} \right\}, \\
 & \left\{ x_1 = -\frac{-ux_2 - vy_2 - wz_2 + vy_1 + wz_1}{u} \right\} \} \tag{6.26}
 \end{aligned}$$

Pour chaque solutions S_i annulant le dénominateur de F , il faut modifier la fonction à intégrer f pour intégrer selon la solution S_i . Par exemple, pour les solutions $S_1 = \{u = 0, v = 0, w = 0\}$ et $S_2 = \{u = 0, v = 0, z_1 = 0\}$, les fonctions à intégrer deviennent :

$$\begin{aligned}
 f_{S_1} &= 1 \Rightarrow F_{S_1} = \frac{1}{6} \\
 f_{S_2} &= e^{-iw(z_2Y+z_3Z)} \\
 &\Rightarrow F_{S_2} = -\frac{wz_2^2z_3 - wz_2z_3^2 + z_3^2ie^{-iwz_2} + iz_2^2 - z_3^2i - z_2^2ie^{-iwz_3}}{w^3z_2^2z_3^2(z_2 - z_3)} \tag{6.27}
 \end{aligned}$$

Il faut expliciter les écritures analytiques de la transformation de Fourier pour toutes les valeurs de $\{u, v, w, x_1, y_1, z_1, x_2, y_2, z_2, x_3, y_3, z_3\}$ et en particulier pour toutes les solutions des équations S . Ceci fait, le calcul de l'expression analytique de la transformée de Fourier peut se faire pour tous les triangles et en tous points de l'espace spectral. Ces calculs ont été réalisés au sein d'un logiciel de calcul formel Maple ce qui a permis d'explicitier l'expression analytique de la transformée de Fourier d'un objet 3D en tous points de l'espace. Les résultats et le code permettant de procéder à la transformée de Fourier d'un objet 3D maillé sont visibles dans l'annexe B et sur la page Web : <http://liris.univ-lyon1.fr/julien.ricard/publications.html>.

Pour illustrer le calcul analytique de la transformée de Fourier d'un objet maillé, nous avons mis en place un procédé de Fourier Volume Rendering (cf. chapitre 4).

Cette méthode de rendu d'objets permet de créer une image profondeur de l'objet sans projeter les voxels, mais en pratiquant une coupe dans l'espace spectral et en reconstruisant l'image réelle par un FFT inverse. L'image obtenue est une projection orthogonale de l'objet discret sur le plan image. Comme nous l'avons fait dans le chapitre 4, l'extraction de la coupe de Fourier est faite sur la transformée de Fourier de l'objet discret.

Pour évaluer le procédé de calcul analytique de la transformée de Fourier, nous avons, au départ de l'objet maillé, calculé la transformée de Fourier pour tous les points d'une image spectrale et reconstruit l'image réel. La construction de la coupe est faite en définissant le plan de coupe comme N^2 points de l'espace spectral. L'expression analytique de la transformée de Fourier du maillage est calculée pour chaque point pour remplir le spectre 2D extrait. La figure 6.3 montre ce procédé sur l'objet 3D Bunny.

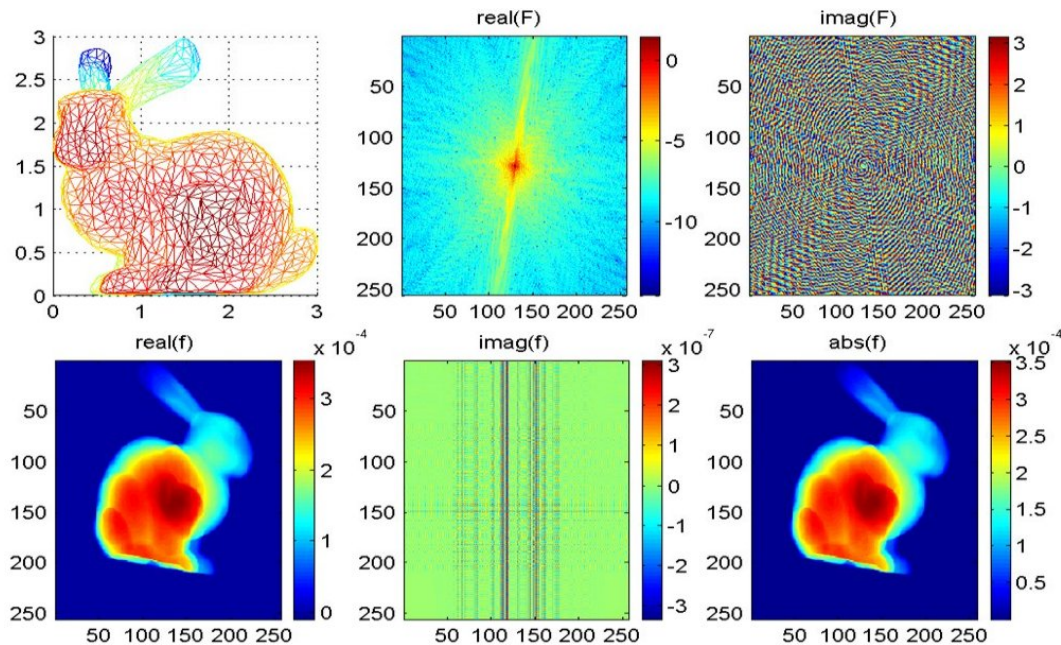


FIG. 6.3 – Exemple de Fourier Volume Rendering de l'objet 3D Bunny. La première ligne représente, l'objet 3D maillé et les parties réelles et imaginaires du spectre extrait. La seconde ligne représente les parties réelles, imaginaires et la valeur absolue de l'image profondeur reconstruite.

La figure 6.4 présente deux exemples de rendu spectral sur l'objet Fandisk pour deux orientations et pour deux tailles de spectre extrait. Comme énoncé dans le chapitre 4, le système de rendu spectral extrait des vues par interpolation dans le domaine Fourier discret des objets 3D. Le fait de pratiquer le rendu par le processus que nous venons de présenter, en utilisant l'expression analytique de la transformée

de Fourier, permet de ne pas faire d'interpolation sur le spectre discret et d'obtenir des vues exactes sans artefacts dus à la discrétisation. La figure 6.5 montre un exemple pour les trois types de rendu spectral : discret avec (a) et sans (b) interpolation et par l'expression analytique(c).

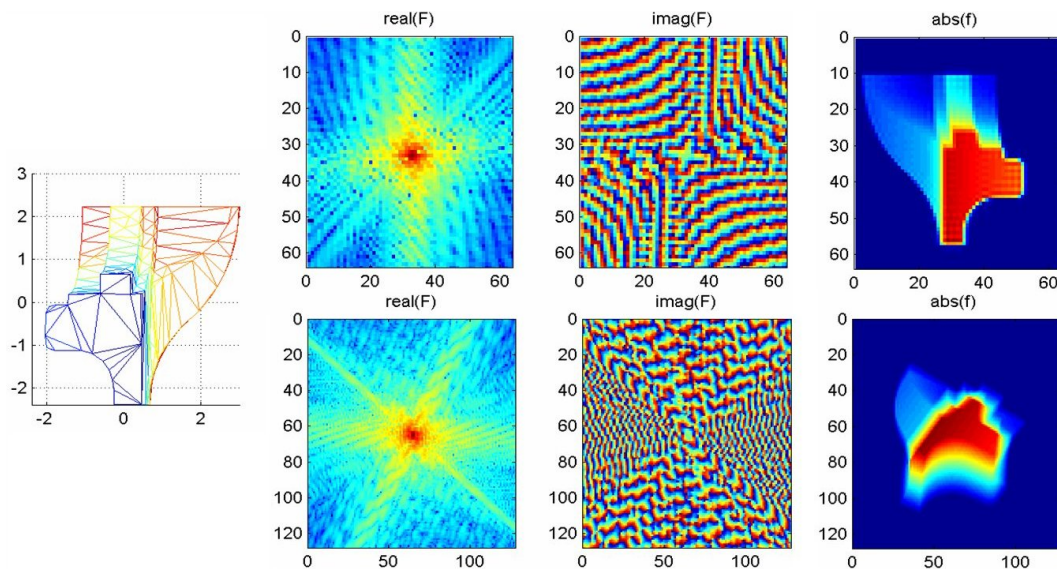


FIG. 6.4 – Exemple de Fourier Volume Rendering de l'objet 3D maillé Fandisk, pour deux orientations quelconques et pour deux tailles du spectre extrait.

6.4 Application en analyse de formes 3D

6.4.1 Centrage et alignement d'un modèle 3D

De nombreuses méthodes d'indexation d'objets 3D ne sont pas intrinsèquement robustes aux translations, aux variations d'échelles et aux rotations (chapitre 2). Pour palier à ces problèmes, la plupart de ces méthodes proposent de mettre en place un pré-traitement qui normalise l'objet 3D pour rendre les méthodes robustes à ces transformations. Pour rendre la méthode robuste aux translations, l'objet est centré par rapport au centre du repère par le calcul de son centre de gravité. Pour les variations d'échelles, l'objet est intégré dans une boîte englobante et pour les rotations, les axes principaux de l'objet sont calculés afin de les aligner avec les axes du repère.

Le pré-traitement peut, soit se faire sur le maillage de l'objet 3D, et les paramètres sont calculés sur le nuage de points, soit se faire sur une représentation discrète de l'objet. Le calcul des paramètres sur le nuage de points constituant le maillage ne permet pas d'avoir une normalisation robuste car elle est dépendante de

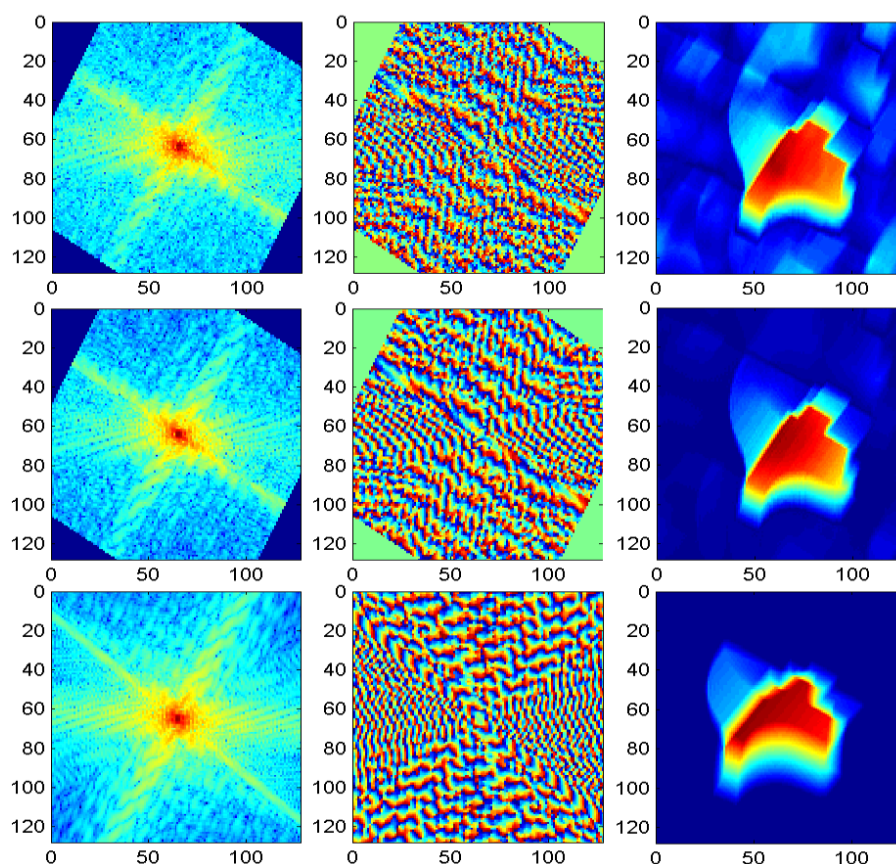


FIG. 6.5 – Exemple de rendu spectral par un processus de Fourier Volume Rendering de l'objet 3D maillé Fandisk par trois types de rendu : discret avec interpolation (a), discret sans interpolation (b), par l'expression analytique(c).

la répartition des points sur le maillage. Un même objet remaillé ne fournira pas les mêmes paramètres et ne sera pas normalisé de façon similaire.

Les paramètres de normalisation de l'objet peuvent aussi se calculer directement à partir du maillage en calculant les moments géométriques à partir de la surface des objets 3D, comme énoncé précédemment. La normalisation d'un objet 3D peut se faire en calculant les moment géométriques. Le centrage se fait grâce au moment géométrique : m_{100} , m_{010} et m_{001} en les additionnant aux coordonnées des points composant l'objet. Ces moments géométriques d'ordre 1 correspondent aux coordonnées du centre de gravité de l'objet. L'alignement se fait en calculant les vecteurs propres de la matrice I ce qui est aussi appelé analyse en composante principale

(ACP). S est la matrice des moments d'ordre 2 :

$$I = \begin{bmatrix} m_{200} & m_{110} & m_{101} \\ m_{110} & m_{020} & m_{011} \\ m_{101} & m_{011} & m_{002} \end{bmatrix} \quad (6.28)$$

Le calcul des paramètres de normalisation sur une représentation directe de l'objet permet d'obtenir une normalisation indépendante du maillage et qui dépend du volume de l'objet. Une telle normalisation exige de transformer l'objet dans le domaine discret pour extraire ces paramètres. Pour ce faire, nous allons discrétiser l'objet et labelliser les voxels internes. Ces opérations ont un coût en $O(N^3 * T)$, où N est la taille de la discrétisation et T est le nombre de triangles composant l'objet. La figure 6.6 montre cinq objets d'une même classe de la base Semantic-3D normalisés en fonction de leurs nuages de points (a) ou de leur volume (b). La normalisation correspond à un centrage sur le centre de gravité de l'objet, une mise à l'échelle sur la boîte englobante $[-1, 1]^3$ et un alignement sur les axes principaux. On peut voir que la normalisation suivant le nuage de points n'est pas robuste, car les objets d'une même classe n'ont pas la même répartition de leurs points. L'alignement volumique fournit des résultats plus cohérents, les centres de gravités et les paramètres de mise à l'échelle sont plus proches, ce qui fournit des alignements plus précis.

La figure 6.7 montre les différentes normalisations calculées en fonction du nuage de points de l'objet (a), de la représentation discrète de l'objet (b) et des moments géométriques calculés à partir de la surface de l'objet (c), pour un même objet 3D ayant subi différents remaillages. On peut tout de suite noter que la normalisation par rapport au nuage de points de l'objet n'est pas robuste au remaillage et que les objets normalisés n'ont pas le même alignement. Les deux autres normalisations, en fonction de la représentation discrète de l'objet et de la représentation tétraédrique, fournissent des bons résultats, car il n'y a pas de différences d'alignement entre les différents objets re-maillés.

L'extraction des paramètres de normalisation sur une représentation discrète des objets 3D doit se faire à une certaine résolution. Le choix de la résolution va modifier les paramètres de normalisation et par la même les objets normalisés. L'extraction des paramètres de normalisation sur la surface de l'objet est indépendante du choix de la résolution de discrétisation. Ce qui fournit des objets normalisés indépendants de tout paramètres.

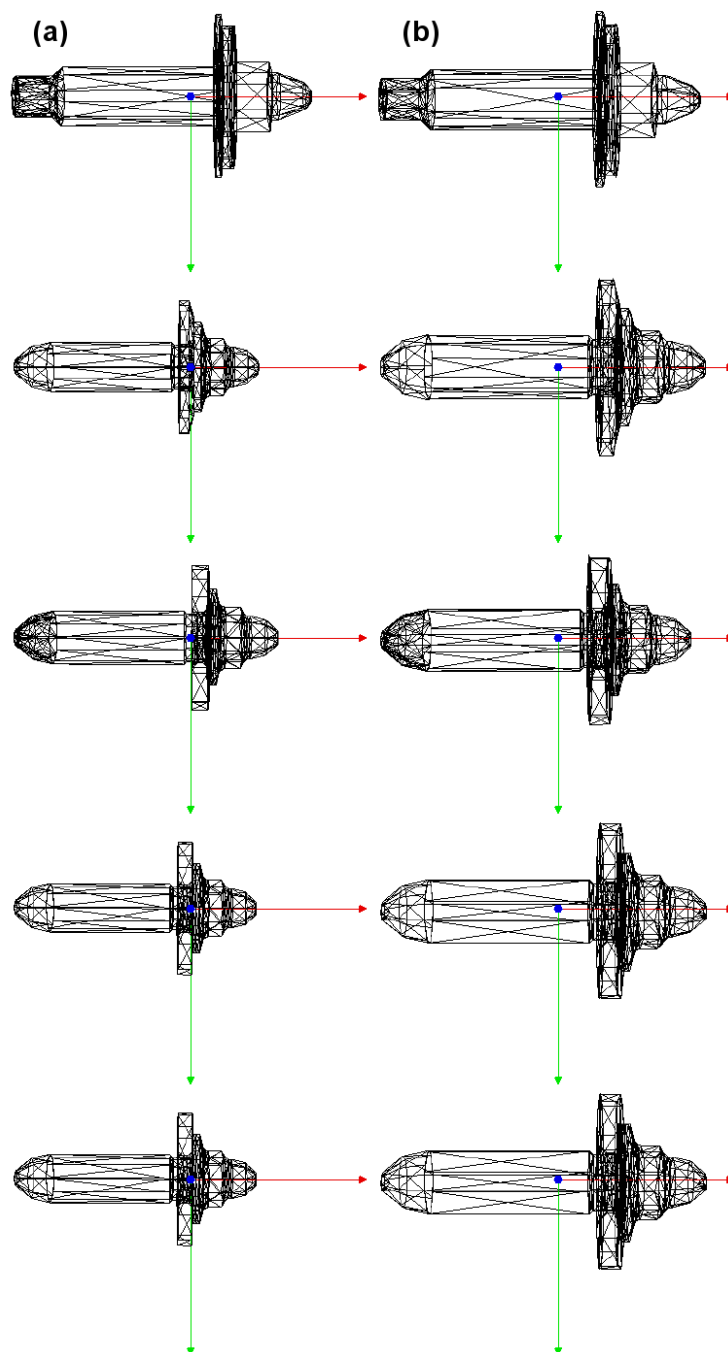


FIG. 6.6 – Exemple de centrage, d'alignements et de mise à l'échelle sur une classe d'objets de la base Semantic-3D, avec une normalisation sur les points (a) et une normalisation volumique(b).

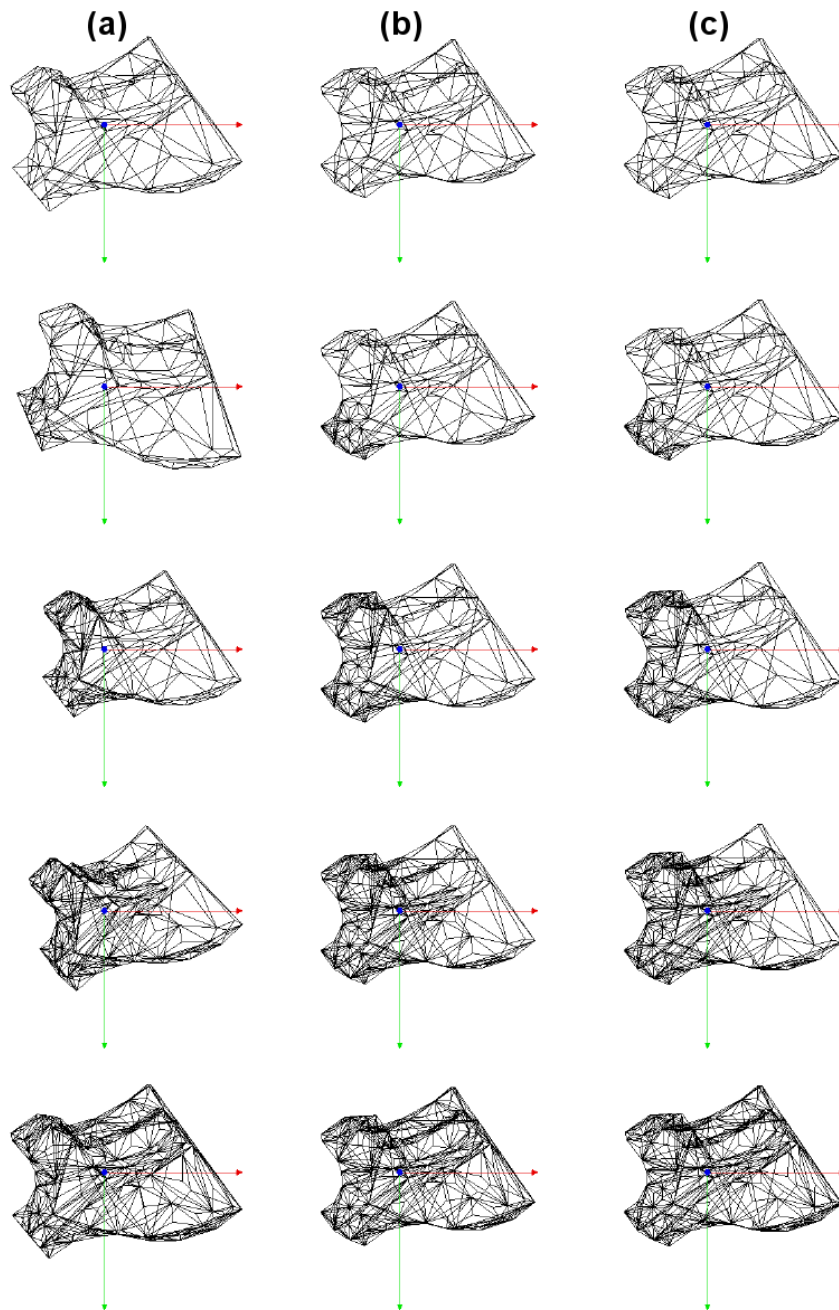


FIG. 6.7 – Exemple d’alignement et de centrage d’un objet 3D remaillé, en calculant les paramètres sur le nuage de points(a), sur l’objet discrétisé (b) et sur la représentation tétraédrique de l’objet(c).

6.4.2 Descripteur de forme 3D

De nombreux descripteurs de forme calculent une intégrale sur le volume des objets 3D pour extraire une signature caractérisant la forme. On peut en particulier citer les méthodes régions basées sur les moments : Fourier 3D (VRANIC et SAUPE, 2001), Zernike 3D (CANTERAKIS, 1999; NOVOTNI et KLEIN, 2004) et ART 3D (RICARD ET AL., 2005). L'intégration sur le volume d'un objet se fait en discrétisant l'objet sur une grille de voxels de taille N^3 et en calculant une somme discrète sur cette représentation.

Descripteur de Fourier 3D. VRANIC et SAUPE (2001) proposent d'utiliser les coefficients de la transformée de Fourier comme descripteur de forme. Les objets sont centrés, mis à l'échelle et alignés par une analyse en composante principale et discrétisés. Le calcul des premiers coefficients de Fourier se fait selon la formule :

$$g_{uvw} = \frac{1}{N^3} \sum_{i=-\frac{N}{2}}^{\frac{N}{2}-1} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \sum_{l=-\frac{N}{2}}^{\frac{N}{2}-1} q_{ikl} e^{-j\frac{2\pi}{N}(iu+kv+lw)} \quad (6.29)$$

où q_{ikl} est l'ensemble des voxels. Les valeurs absolues des coefficients g_{uvw} d'indices $-K \leq u, v, w \leq K$ sont sélectionnées comme descripteurs. Excepté le coefficient g_{000} , tous les coefficients complexes sont conjugués deux à deux (g_{uvw} est le conjugué de g_{-u-v-w}). Le vecteur descripteur consiste à $((2K+1)^3 + 1)/2$ valeurs réelles. Les meilleurs résultats sont obtenus pour $K = 3$.

Comme expliqué dans le paragraphe 6.3.3, la transformation de Fourier d'un objet 3D maillé peut être calculée en décomposant l'intégrale sur la forme sur chaque élément de forme élémentaire par la formule 6.23.

Descripteur de Zernike 3D. La description d'une forme par le descripteur de Zernike 3D a été définie par CANTERAKIS (1999) et utilisé par NOVOTNI et KLEIN (2003, 2004) et est largement reconnue au sein de la communauté. Le calcul du descripteur de Zernike 3D, tel qu'il est proposé par les auteurs, se fait de façon discrète sur un objet discrétisé et normalisé, en combinant linéairement les moments géométriques m_{pqr} , pour tous $p, q, r > 0$ et $p+q+r \leq n$, avec $n \in [0, N]$ et $l \in [0, n]$, tel que $(n-l)$ est paire, et $m \in [-l, l]$. Le descripteur de forme de Zernike 3D est la norme du vecteur Ω_{nl} selon la formule :

$$F_{nl} = \|\Omega_{nl}\| \quad \text{ou} : \quad \Omega_{nl}^m = \frac{3}{4} \sum_{p+q+r \leq n} \overline{\chi_{nlm}^{pqr}} m_{pqr} \quad (6.30)$$

Les fonctions Ω_{nl}^m sont les moments de Zernike 3D d'un objet, et sont une combinaison linéaire des moments géométriques d'ordre n . La combinaison linéaire se fait grâce

à la fonction χ_{nlm}^{pqr} selon la formule :

$$\begin{aligned} \chi_{nlm}^{pqr} &= c_l^m 2^{-lm} \sum_{v=0}^k q_{kl}^v \sum_{\alpha=0}^v \binom{v}{\alpha} \sum_{\beta=0}^{v-\alpha} \binom{v-\alpha}{\beta} \sum_{u=0}^m (-1)^{m-u} \binom{m}{u} i^u \\ &\quad \cdot \sum_{\mu=0}^{\lfloor \frac{l-m}{2} \rfloor} (-1)^\mu 2^{-2\mu} \binom{l}{\mu} \binom{l-\mu}{m+\mu} \sum_{v=0}^{\mu} \binom{\mu}{v} \end{aligned} \quad (6.31)$$

ou $2k = n - 1$, c_l^m est le facteur de normalisation et q_{kl}^v garantit l'orthogonalité des fonctions avec la sphère unité. Ces deux termes se calculent avec les formules suivantes :

$$c_l^m = c_l^{-m} = \frac{\sqrt{(2l+1)(l+m)!(l-m)!}}{l!} \quad (6.32)$$

$$q_{kl}^v = \frac{(-1)^k}{2^{2k}} \sqrt{\frac{2l+4k+3}{3}} \binom{2k}{k} (-1)^v \frac{\binom{k}{k} \binom{2(k+l+v)+1}{2k}}{\binom{k+l+v}{k}} \quad (6.33)$$

Le calcul des descripteurs de Zernike 3D se base sur les moments géométriques calculés de façon discrète. Comme énoncé précédemment, il est plus efficace de calculer les moments géométriques en intégrant la surface de l'objet et en sommant la contribution signée des tétraèdres la composant que de les calculer à partir d'une discrétisation.

6.4.3 Complexité

Le principal avantage d'intégrer le volume d'un objet 3D sur son maillage et non sur une représentation volumique discrète est le gain de complexité dû principalement au fait de ne pas avoir à discrétiser l'objet 3D. Le coût d'une discrétisation d'un objet 3D est en $O(N^3 * T)$ où T est le nombre de triangles composant l'objet et N est la taille de la discrétisation. L'objet discrétisé doit ensuite être labellisé pour marquer les voxels internes, ce qui peut se faire en $O(N^3)$. L'intégration a alors un coût en $O(M)$ où M est le nombre de voxels marqués. On peut noter que M est généralement supérieur à T le nombre de triangles composant l'objet. Alors que le coût de l'intégration sur le maillage de l'objet a un coût en $O(T)$.

6.5 Conclusion

Ces approches permettent de calculer directement l'intégrale d'un volume sur le maillage d'un objet 3D, sans passer par une phase de discrétisation, qui rallonge

	Discrétisation	Tétraédrique
Fourier 3D descriptor	$O(N^3 * T)$	$O(T)$
Zernike 3D descriptor	$O(N^3 * T)$	$O(T)$

TAB. 6.1 – Comparaison des complexités des processus d'indexation par discrétisation de l'objet ou par intégration sur le maillage par décomposition tétraédrique. T est le nombre de triangles composant l'objet 3D et N est la taille de la grille de discrétisation.

les calculs et est moins précis qu'un calcul analytique. Nous avons montré que l'expression analytique d'un maillage 3D peut être utilisée pour exprimer les moments géométriques ou la transformée de Fourier. Nous l'avons utilisé pour calculer les paramètres de normalisation et pour calculer des descripteurs : moments géométriques, descripteurs de Fourier et moments de Zernike.

L'intégration tétraédrique fournit l'expression analytique de l'intégrale de la fonction à intégrer sur le tétraèdre unité. Ce qui demande d'exprimer cette fonction sur tout le domaine à intégrer et surtout où la fonction n'est pas correctement définie, comme nous l'avons explicité pour la transformation de Fourier. Cette étape limite la généralisation pour n'importe quelle intégrale, car il est difficile d'intégrer une fonction quelconque comme les fonctions de base ART 3D où la présence de cosinus complique le calcul.

Notre travail a principalement porté sur la mise en place des calculs et sur leur implantation au sein d'un logiciel de calcul formel pour expliciter les formules et l'utilisation des expressions analytiques dans différents programmes d'indexation d'objets 3D. Ces travaux se basent sur l'étude de l'intégration surfacique faite par LIEN (1984) et reprise par ZHANG et CHEN (2001), mais sont originaux, par la résolution de l'expression de Fourier analytique en tout point de l'espace et par les applications que nous avons présentées en analyse de formes.

Comme nous l'avons vu, le calcul de l'expression analytique d'un maillage 3D n'est possible que sous les conditions que le maillage soit une variété topologique (sans trous, orientable...). Dans le cadre du projet RNRT Semantic-3D, une grande partie des objets de la base ne possède pas ces propriétés et nous n'avons pas pu utiliser ces méthodes, par exemple lors des phases de pré-traitements (alignement, normalisation,...).

7

Navigation dans une base d'objets 3D

Sommaire

7.1	Introduction	148
7.2	Etat de l'art en navigation 3D	148
7.2.1	Indexation	149
7.2.2	Classification	149
7.2.3	Visualisation	150
7.3	Approche proposée	151
7.3.1	Indexation	151
7.3.2	Système de classification	153
7.3.2.1	Critère de qualité d'un groupe	154
	L'énergie intra-groupe	155
	L'énergie inter-groupe	155
7.3.2.2	Arbre hiérarchique	155
7.3.3	Système de visualisation	156
7.3.3.1	Description de l'interface	157
7.3.3.2	Relevance feedback	158
7.4	Expérimentations	159
7.4.1	Critère de qualité de la partition	160
7.4.2	Performance du système de navigation	161
7.4.3	Validation du système de visualisation	162
7.5	Conclusion	163

7.1 Introduction

Dans le cadre du projet RNRT Semantic-3D et plus généralement dans cette thèse, nous nous sommes employé à mettre en place des descripteurs de forme pour mesurer la similarité entre objets et à faire des requêtes par similarité. Une autre approche est de mettre en place un processus de navigation permettant de parcourir les différentes classes de la base pour rechercher un objet. Ce travail a fait l'objet d'un stage de Master Recherche réalisé par M. Rémi Trichet et co-encadré avec M. Bruno Tellez, et a fait l'objet d'une publication nationale (TRICHET ET AL., 2005).

Les deux problèmes majeurs de l'exploration d'une base inconnue sont : la difficulté à appréhender la base dans sa globalité et l'absence de représentation des relations entre les objets. Les méthodes de classification automatique apportent une aide précieuse dans la résolution de ce type de problème, en découpant la base en groupes similaires. Une fois les objets de la base regroupés, plutôt que de représenter la totalité de la base, les systèmes de visualisation choisissent de définir un objet résumé pour chaque groupe. Ainsi l'utilisateur peut naviguer rapidement et efficacement dans la base en ne conservant que l'essentiel de l'information. Outre la structuration de la base, le choix de l'espace de visualisation est également important.

Cette partie présente le système de navigation dans une base d'objets 3D que nous avons développé. Nous montrerons que la navigation dans une collection d'objets à l'aide d'une interface en trois dimensions peut apporter à l'utilisateur une aide significative dans la visualisation de la collection dans son ensemble. En effet, en plus de la représentation d'objets 3D, l'ajout d'une dimension supplémentaire permet d'afficher un plus grand nombre d'informations. Après un état de l'art sur la navigation, nous présenterons notre approche qui propose un nouvel algorithme de classification par division, ainsi qu'un système de navigation 3D. Ce dernier est basé sur une interface utilisant des arbres de cônes. Enfin, nous concluons en présentant les résultats obtenus et le prototype que nous avons implanté.

7.2 Etat de l'art en navigation 3D

Un processus de navigation peut se décomposer selon trois phases :

1. **l'indexation**, qui permet de représenter les objets par un ensemble de descripteurs qui facilitent leur comparaison. Ces descripteurs permettent de représenter l'objet comme un point dans l'espace de représentation.
2. **la classification**, qui structure l'espace en découpant le nuage de points créé en groupes cohérents.
3. **la visualisation**, qui permet de consulter la disposition des objets et les informations qui leurs sont associées.

7.2.1 Indexation

La première phase d'un processus de navigation consiste à extraire des descripteurs des objets pour représenter la base d'objets 3D comme un nuage de points dans un espace de grande dimension, qui sera utilisé par la suite pour structurer la base. Le choix des descripteurs à prendre en compte est un paramètre important car il va directement influencer sur les performances du système, tant d'un point de vue du coût que de l'efficacité. Les descripteurs doivent être pertinents pour les types d'objets considérés dans la collection et doivent éviter la redondance d'information (non corrélation des descripteurs) qui faussent leur importance relative. Enfin, l'utilisation d'un trop grand nombre de descripteurs fournira, certes, une information plus précise mais plus fragmentaire et difficile à interpréter et augmentera le coût d'une analyse de la base qui, comme nous le montrerons par la suite, doit être faible pour permettre de recalculer la structure en fonction du parcours réalisé.

7.2.2 Classification

La classification a pour but de découper automatiquement le nuage de points en groupes de telle sorte que les éléments d'un même groupe soient les plus similaires possibles et que les éléments de groupes différents soient les plus dissemblables possibles. On distingue deux catégories de méthodes : la classification hiérarchique et le partitionnement.

Les algorithmes de classification hiérarchique divisent le nuage de points en groupes pouvant être représentés par un arbre hiérarchique ou un dendrogramme (les feuilles représentant les points). Les méthodes de classification hiérarchique ont l'avantage de ne pas souffrir de problèmes d'initialisation et de créer des partitions aisément structurables mais présentent de nombreux inconvénients, car elles ne tiennent pas compte du nuage de points dans sa globalité mais seulement des proximités immédiates de chaque point. Ce sont des méthodes statiques, autrement dit une fois un point attribué à un groupe, il est impossible de le ré-affecter à un autre groupe. Plusieurs méthodes de ce type sont recensées. CRAVER ET AL. (1999) proposa d'utiliser des arbres de division, CHEN ET AL. (2000) présentèrent une méthode fortement inspirée du quad-tree et KOHONEN (1997) développa les cartes auto organisées ou cartes de Kohonen utilisées dans de nombreux systèmes de navigation (HONKELA ET AL., 2000).

Les algorithmes de partitionnement, par opposition aux algorithmes de classification hiérarchique, divisent un nuage de points en groupes sans aucune relation entre eux. L'algorithme le plus connu, les nuées dynamiques assument l'hypothèse que le nombre de groupes est connu et affecte à chacun un centre de classe. La partition retenue est celle qui minimise la distance des points de chaque classe à leur centre. Tout d'abord, l'algorithme est initialisé en plaçant aléatoirement les centres de classe. Puis, tant que le système n'est pas stabilisé, chaque point est alloué à la

classe dont le centre est le plus proche et le centre de chaque groupe est placé à son centre de gravité. Cet algorithme nécessite de connaître le nombre de groupes : chaque point ne peut être affecté qu'à une classe et une seule. Ceci conduisit BEZEDEK (1981) à développer le partitionnement flou. Dans ce type d'algorithme, à tout point est attribué un pourcentage d'appartenance à chaque groupe. Si celui-ci dépasse un certain seuil, préalablement fixé, le point est attribué à ce groupe. Les algorithmes de partitionnement évoluèrent pour pouvoir fonctionner avec un nombre de groupes à priori inconnu. On en distingue deux catégories : le partitionnement par agglomération comme l'agglomération compétitive (FRIGUI et KRISHNAPURAM, 1997) ou l'ARC (Adaptative Robust Competition) proposé par LE SAUX et BOUJEMAA (2004), qui considère les points comme autant de groupes isolés et élimine, au fur et à mesure des itérations, les groupes non pertinents en minimisant une fonction objectif jusqu'à l'obtention du nombre correct de groupes. Le partitionnement par division, au contraire, ajoute un nouveau groupe à la partition et réajuste celle-ci tant qu'elle ne satisfait pas à un critère d'optimalité. Le Principal Direction Divisive Partitionning (PDDP) proposé par BOLEY (1998) utilise les propriétés géométriques de l'analyse en composantes principales pour diviser itérativement le nuage de points. Les nuées dynamiques bissectrices, par contre, sont une variante des nuées dynamiques. Débutant avec deux groupes, il augmente itérativement le nombre de groupes et relance l'algorithme des nuées dynamiques tant qu'un critère d'optimalité n'est pas satisfait.

7.2.3 Visualisation

Une fois la base catégorisée et structurée, il est nécessaire de représenter les informations obtenues de façon à les rendre facilement accessibles par l'utilisateur. L'interface est la finalisation de la conception d'un système de navigation. Elle doit permettre à l'utilisateur de visualiser le maximum d'informations apportées par les étapes précédentes, d'intuitivement comprendre les mécanismes mis en oeuvre par le système et être ergonomique.

Les collections étant trop grandes pour être affichées dans leur intégralité, il faut choisir un objet résumé pour représenter chaque groupe. De plus il faut représenter en faible dimension (1,2 ou 3 dimensions) les données tout en limitant la perte d'informations. Le système de visualisation le plus largement répandu est la carte 2D, où les images sont représentées par des vignettes réparties sur un plan, la distance entre celles-ci représentant la dissemblance entre les images. Les groupes étant structurés selon un arbre hiérarchique, chaque vignette est l'image représentative d'un sous-groupe, cliquer sur une vignette permet d'accéder au sous-groupe correspondant. La figure 7.1 montre le système de navigation basé sur une visualisation par des cartes 2D développé au sein du LIRIS par IDRISI ET AL. (2004). Ce type de visualisation apporte une meilleure compréhension des relations entre les images, comme la montré MOGHADDAM ET AL. (2001). Par contre, l'absence

de représentation de la structure de navigation empêche l'utilisateur d'appréhender celle-ci dans sa globalité. En effet, si celui-ci effectue un mauvais choix au cours du processus de navigation, dans la majorité des cas, il recommencera le processus du début plutôt que de revenir en arrière. Les travaux de [SANTINI ET AL. \(2001\)](#) ou de [MOGHADDAM ET AL. \(2001\)](#) visent à permettre l'inter-activité entre l'utilisateur et le système

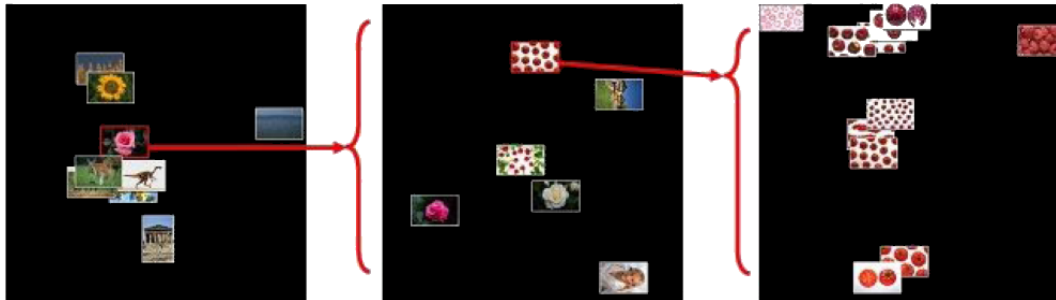


FIG. 7.1 – Exemple de carte 2D affichant des images ([IDRISSI ET AL., 2004](#)).

L'ajout d'une troisième dimension comporte de nombreux avantages : cela rend l'affichage d'objets 3D possible, permet d'afficher un plus grand nombre d'objets sans surcharger l'écran et détruit moins l'information. Toutefois, une compréhension difficile des relations entre les objets et une interface souvent moins facile à prendre en main rebutent généralement les concepteurs. Les interfaces 3D les plus utilisées jusqu'à présent sont les arbres de cônes développés par [ROBERTSON ET AL. \(1991\)](#), FSN de [ROGERS ET AL. \(1996\)](#) dont l'organisation hiérarchique est directement inspirée de celle de villes composées de bâtiments et reliées par des routes et Pathfinders Networks de [CHEN ET MORRIS \(2003\)](#) inspirés de représentations 3D utilisées en psychologie. La figure 7.2 montre des exemples de ces trois approches.

7.3 Approche proposée

Pour répondre à la problématique de la navigation dans une base d'objets 3D, nous avons proposé un système complet de navigation. Notre système se base sur le schéma classique en trois parties que nous venons d'énoncer.

7.3.1 Indexation

Les descripteurs que nous avons choisis d'utiliser sont les 74 descripteurs de forme ART 3D présentés dans le chapitre 5. Le choix de ce descripteur de forme est motivé par le fait de vérifier les performances d'ART 3D, mais il faut noter que l'approche que nous proposons est complètement indépendante de la base de représentation choisie.

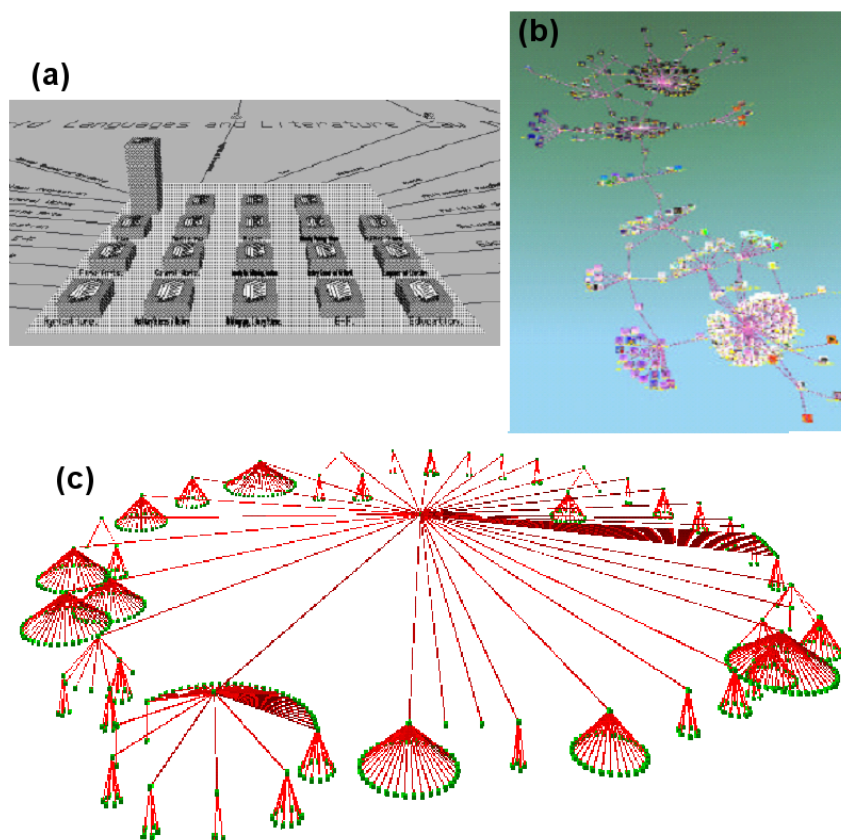


FIG. 7.2 – Exemple d’affichage d’une base d’objets en trois dimensions. (a) FSN de ROGERS ET AL. (1996), (b) Pathfinders Networks de CHEN ET MORRIS (2003) et les arbres de cônes proposés par ROBERTSON ET AL. (1991).

Pour augmenter la rapidité du système, nous avons opté pour la réduction de l’espace de représentation. La méthode de réduction de dimensionnalité que nous avons choisie est l’analyse en composantes principales. Une étude de la variance expliquée des composantes principales (Figure 7.3) montre que 99% de l’information est contenue dans les 42 dimensions à plus forte variance, que l’apport des 32 dimensions suivantes est quasi nul et que 6 dimensions suffisent pour apporter 75% de l’information. Aussi, on peut effectuer, lors d’une étape préalable à la classification non supervisée, une importante réduction de l’espace avec une perte d’information minimale. Cette étape permet d’éliminer l’information redondante et de limiter le temps de calcul de la classification non supervisée qui va suivre. Dans notre système de navigation, nous avons choisi de ne prendre en compte que les 25 premières dimensions qui constituent un bon compromis entre le nombre de descripteurs et le pourcentage d’information apportée (95%).

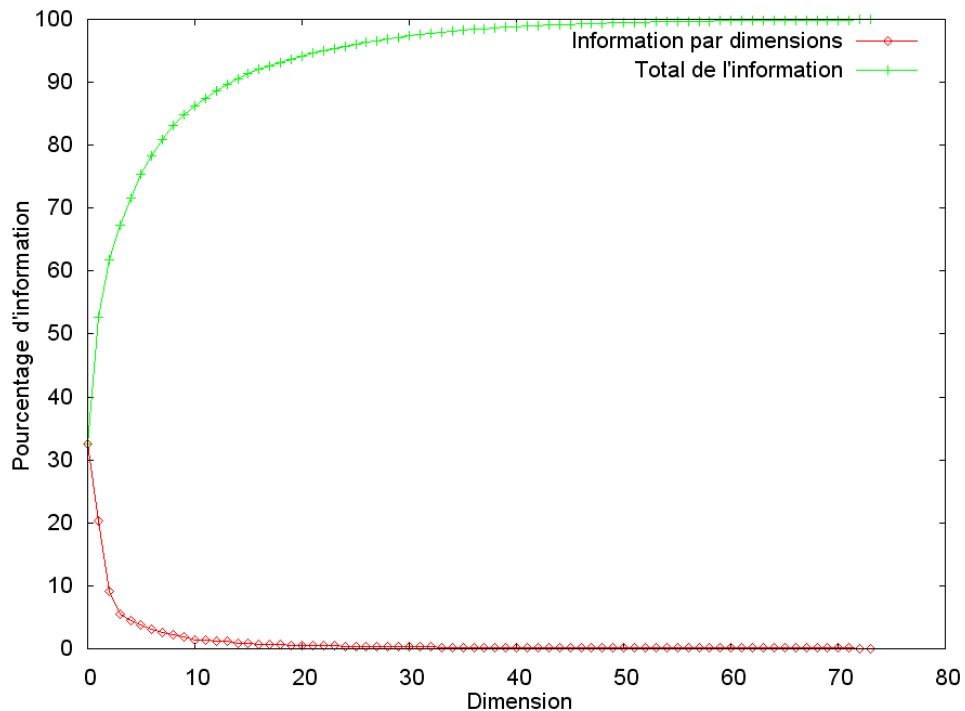


FIG. 7.3 – Pourcentage d’information apportée par les dimensions après une analyse en composantes principales.

7.3.2 Système de classification

Notre système de classification doit obéir à certaines contraintes, à savoir que le nombre de groupes effectifs est inconnu et que la forme des groupes ainsi que celle du nuage de points initial sont quelconques. L’algorithme des nuées dynamiques bissectives répond à ces contraintes, mais il présente également certains désavantages comme le fait que la qualité des groupes créés et la rapidité de convergence de l’algorithme dépendent du placement initial des centres des groupes.

Nous avons choisi d’adapter cet algorithme en utilisant les propriétés géométriques de l’analyse en composantes principales pour guider le placement des centres des groupes. La première composante principale d’un nuage de points suit l’allongement maximal du nuage et passe par son centre de gravité (Figure 7.4). Cette heuristique fonctionne dans la grande majorité des cas et permet de gagner un temps de calcul important, en positionnant les nouveaux centres des classes.

Le critère d’arrêt usuellement défini consiste à évaluer la qualité de la partition. Si celle-ci est inférieure à un seuil donné, les groupes sont considérés satisfaisants et l’algorithme s’arrête. Dans notre algorithme, au lieu d’établir un critère de qualité pour la partition entière, un critère de qualité est défini pour chaque groupe. L’algorithme s’achève lorsque le critère de chaque groupe est inférieur à un seuil de qualité

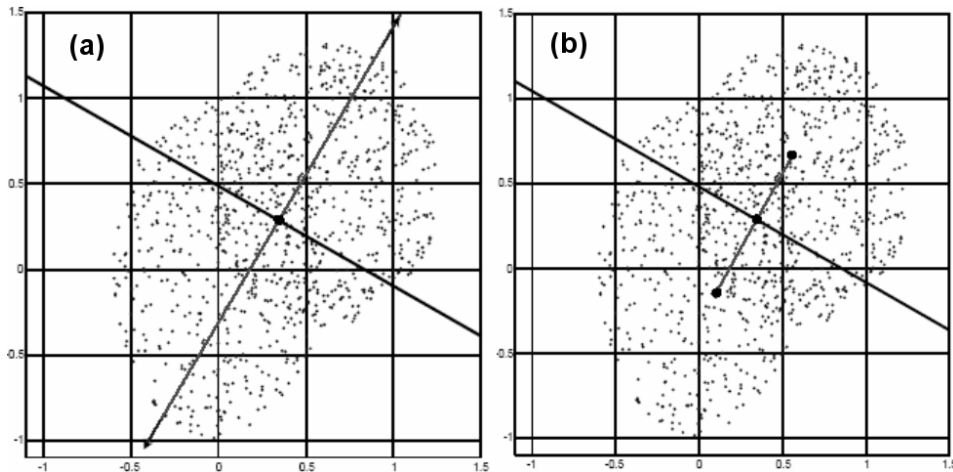


FIG. 7.4 – La première composante principale représentée par une flèche permet de diviser le nuage de points en 2 groupes (a) et de définir leurs centres de gravité comme centre des groupes (b).

défini par l'utilisateur et, pour l'itération suivante, l'algorithme n'est relancé que sur les groupes dont le critère de qualité est supérieur au seuil.

7.3.2.1 Critère de qualité d'un groupe

Le critère de qualité de la partition est le critère d'arrêt des algorithmes de partitionnement. Celui-ci se compose d'une énergie inter-groupe représentant l'écartement moyen entre les différents groupes et d'une énergie intra-groupe représentant la compacité moyenne des groupes. Notre apport consiste à définir un critère pour chaque groupe, ce qui comporte plusieurs avantages :

- la qualité du partitionnement est évaluée de façon plus précise puisque chaque groupe est évalué indépendamment ;
- lorsqu'un groupe a une qualité satisfaisante, il n'est plus nécessaire de le traiter à nouveau ce qui permet de diminuer le temps de calcul pour des résultats identiques ;
- on peut ainsi obtenir des groupes de tailles différentes.

Nous avons choisi de représenter la qualité de notre partitionnement par l'énergie intra-groupe E_{intra} , qui caractérise la compacité du groupe et par l'énergie inter-groupe E_{inter} , qui décrit l'isolement du groupe par rapport aux autres groupes. La qualité d'un groupe se définit comme la somme pondérée des deux énergies selon la formule :

$$S(c) = \lambda_1 \cdot E_{intra} + \lambda_2 E_{inter} \quad (7.1)$$

où λ_1 et λ_2 sont les poids accordés respectivement à l'énergie intra et inter classes et représentant l'importance relative de chacune des deux énergies.

L'énergie intra-groupe d'un groupe C est défini par :

$$E_{intra} = f(N_C) \sqrt{\frac{\sum_{i=0}^{N_C} \|x_i - \beta_c\|^2}{\frac{N_C}{2} \max_i \|x_i - \beta_c\|^2}} \quad (7.2)$$

où N_C est la cardinalité du groupe, et β_C le centre du groupe. Ce critère est évalué en comparant la distance moyenne des points au centre à celle, divisée par 2, du point le plus éloigné du centre de classe. Il sera fort si la majorité des points sont répartis entre le centre et cet axe médian, autrement dit pour les groupes compacts. Mais ce critère n'est plus représentatif pour les groupes à faible cardinalité, c'est pourquoi il est pondéré par la fonction f qui limite son importance lorsque la cardinalité du groupe est faible. Pour notre système de navigation, nous avons défini empiriquement la fonction f suivante pour qu'elle respecte le profil de courbe attendu :

$$\begin{cases} f(x) = x/10 & \text{si } x \leq 1 \\ f(x) = 1 + \log(x - 9)/30 & \text{sinon} \end{cases} \quad (7.3)$$

L'énergie inter-groupe d'un groupe C est défini par :

$$E_{inter} = \frac{1}{N} \sum_{i=0}^N \sqrt{\frac{\frac{\sum_{j=0}^{N_C} \|x_j - \beta_C\|^2}{N_C} + \frac{\sum_{j=0}^{N_A} \|x_j - \beta_i\|^2}{N_i}}{\|\beta_C - \beta_i\|^2}} \quad (7.4)$$

où N est le nombre de groupes avec lesquels le groupe C est comparé, β_C et β_i les centres de classe respectifs des groupes C et i , N_C et N_i leurs cardinalités respectives.

La distance entre deux groupes est évaluée en comparant l'écart entre les distances moyennes des points de ces deux groupes à leur centre et l'écart entre les centres de classe de ces deux groupes (Figure 7.5). Ainsi, deux groupes volumineux juxtaposés seraient considérés comme éloignés si les distances entre les centres de classe seules étaient considérées, alors que l'utilisation de notre critère de proximité permet de considérer ces deux groupes comme plus ou moins proches en fonction de leur compacité.

Le critère inter-groupes est alors calculé en effectuant la moyenne de la distance entre le groupe considéré et les N groupes qui lui sont les plus proches. Dans notre système de navigation, nous avons estimé que cinq groupes voisins suffisent pour déterminer efficacement ce critère.

7.3.2.2 Arbre hiérarchique

La création de l'arbre hiérarchique est cruciale pour le système de navigation. Lors du processus de navigation, l'utilisateur va se déplacer suivant cette arborescence. Chaque feuille représente un objet et chaque noeud représente un groupe, les

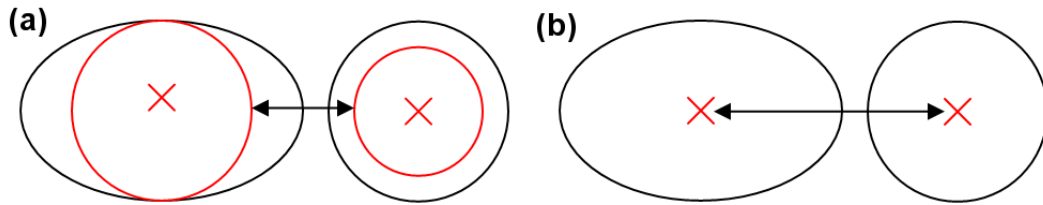


FIG. 7.5 – Les croix sont les centres de classe, les ellipses noires les groupes et les cercles rouges la distance moyenne des points d'un groupe à son centre. (a) écart entre 2 groupes tenant compte de la distance moyenne des points de chaque groupe à leur centre respectif. (b) écart entre 2 groupes basés sur la distance entre les centres de groupe.

fil de ce groupe étant les objets et/ou les sous-groupes le composant. Notre arbre hiérarchique devra avoir certaines caractéristiques :

- les fils de chaque sommet devront présenter des similarités selon les critères établis ;
- plus la longueur du chemin entre une feuille et le sommet est importante, plus la recherche sera longue et plus les risques de mauvaise classification seront élevés. En conséquence, l'arbre devra être de hauteur réduite ;
- en contrepartie, chaque groupe ne devra pas avoir trop de sous-groupes afin de ne pas surcharger l'affichage. L'affichage de trop d'objets simultanément rendrait les relations entre les objets illisibles. Dans notre interface, nous avons fixé ce seuil à dix sous-groupes.

Nous avons comparé pour 3 types d'arbres hiérarchiques différents leur efficacité à représenter une base d'objets. Il est à noter que la bonne répartition ou non des objets selon l'arbre hiérarchique répondant à une notion sémantique, il ne nous a pas été possible d'établir un critère d'évaluation de cette répartition.

La figure 7.6 montre un exemple de construction d'arbres hiérarchiques. L'arbre hiérarchique ascendant initialisé avec les groupes issus de l'algorithme de classification est affiné pour obtenir l'arbre qui sera affiché par l'interface de visualisation. A chaque itération, les 2 groupes les plus similaires sont fusionnés. Afin d'obtenir un nombre suffisant de sous-groupes, on ne conserve qu'un niveau sur 3, chaque groupe a donc entre 2 et 8 fils.

7.3.3 Système de visualisation

Le choix de notre système de visualisation s'est naturellement orienté vers la 3D, d'une part car il s'agissait du moyen le plus sûr de représenter des objets 3D aux formes peu communes sans nuire à leur lisibilité, et d'autre part car la 3D offre des possibilités de navigation qui sont encore peu exploitées. Afin d'en tester ses possibilités, nous avons codé un prototype en OpenGL.

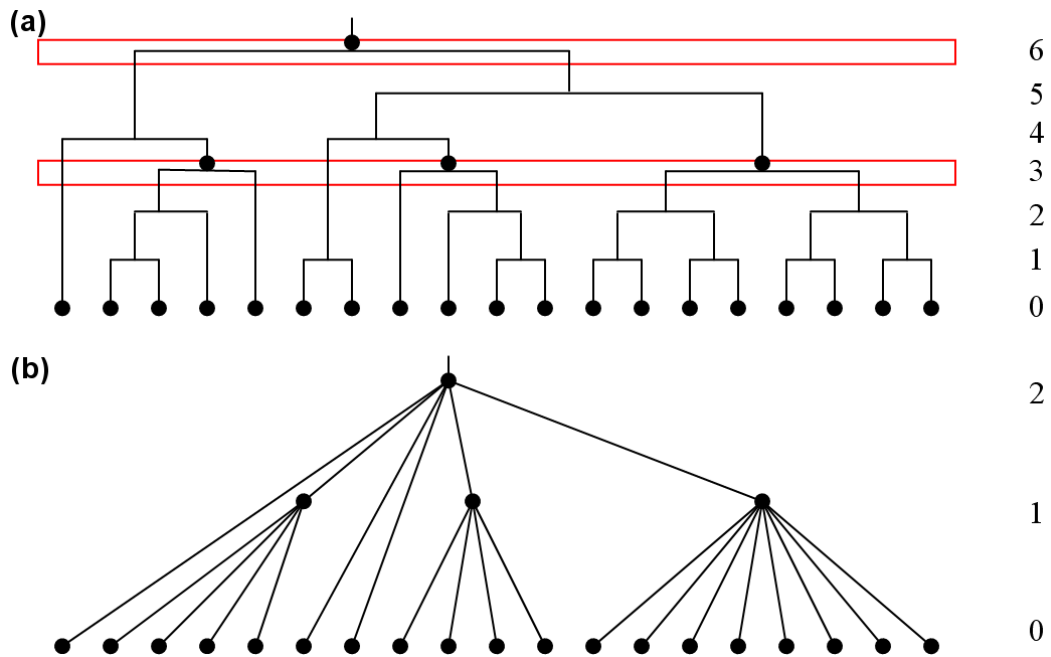


FIG. 7.6 – Exemple de création et d’affinage de l’arbre hiérarchique. (a) Création de l’arbre hiérarchique ascendant.(b) élimination de 2 niveaux sur 3.

7.3.3.1 Description de l’interface

Notre interface est une interface 3D inspirée des arbres de cônes développés par ROBERTSON ET AL. (1991). Chaque groupe est modélisé par un cône, son représentant est placé sur le sommet du cône et les objets ou les représentants des sous-groupes sont modélisés sur le pourtour du cône. La similarité entre les objets est représentée par la proximité entre les objets sur la surface de ce cône. Afin d’éviter la superposition d’objets très similaires voir identiques, un écart minimum entre les objets est imposé. Enfin, pour plus de clarté, aucun objet n’est modélisé dans le tiers supérieur du cône, leur sélection conduisant à représenter les cônes très proches les uns des autres et nuisant à leur visualisation. Les objets sont représentés par des couleurs différentes : les représentants de groupes en gris, les feuilles de l’arborescence en vert et l’objet actuellement sélectionné en rouge.

Le fonctionnement de l’interface est le suivant. A tout instant est représenté l’objet sélectionné, ses objets fils (i.e les objets contenus dans le groupe dont il est le représentant), les objets constituant le chemin depuis le sommet de l’arborescence jusqu’à lui, et ses objets frères (i.e les autres objets constituant le groupe dont son père est le représentant). L’interface est initialisée par l’affichage de l’objet au sommet de l’arborescence ainsi que de ses fils. Par la suite, l’utilisateur navigue dans la base en sélectionnant, à chaque étape, l’objet affiché qui lui semble le plus proche

de celui qu'il cherche. Il dispose pour cela de 2 moyens, la souris qui permet de cliquer sur les objets pour afficher leurs fils et le pavé numérique, qui permet de monter et de descendre ou de faire défiler les frères dans l'arbre. Lorsqu'un objet est sélectionné (Figure 7.7), la fenêtre est centrée sur celui-ci, une rotation est effectuée autour de son cône afin de le placer à l'avant plan et les objets adéquats sont ajoutés ou ôtés afin de respecter la règle d'affichage énoncée plus haut.

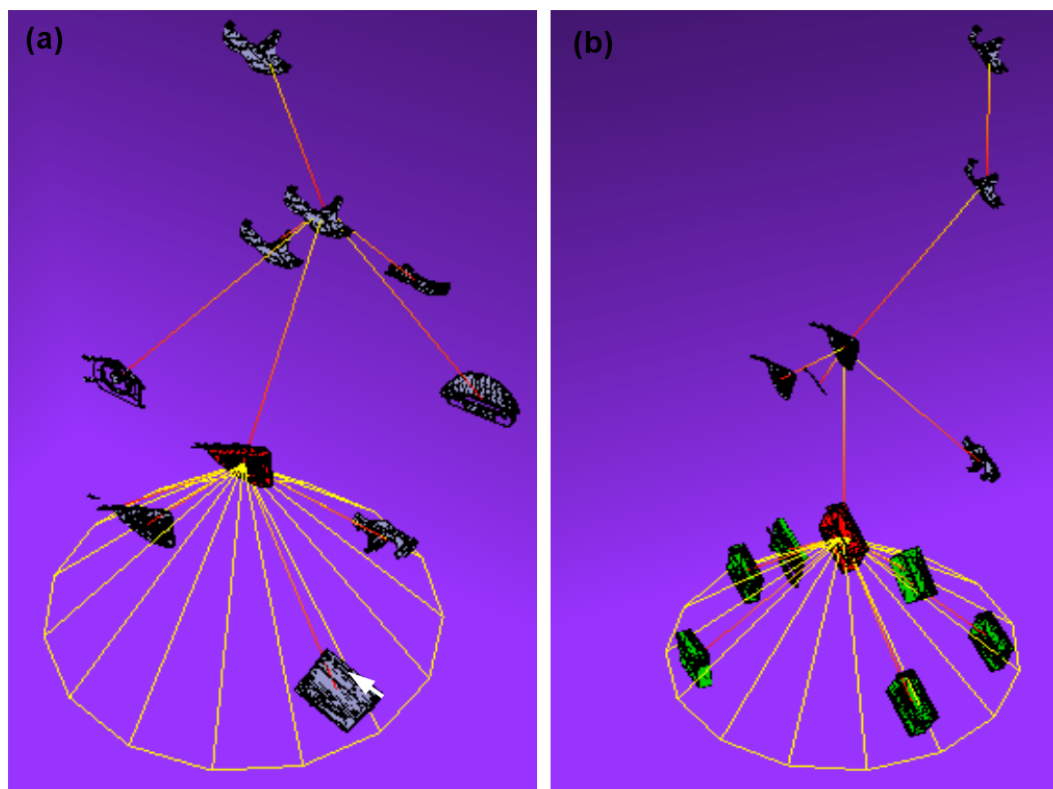


FIG. 7.7 – Illustration du fonctionnement de l'interface. (a) Au cours de la navigation, l'utilisateur sélectionne l'objet en bas de l'écran. (b) Celui-ci devient rouge et l'affichage est modifié en conséquence.

Ce type d'interface permet à la fois de pouvoir afficher des objets en 3D et de visualiser le chemin suivi dans l'arborescence tout en restant clair et facile d'utilisation. Son seul désavantage est que la clarté souhaitée n'est réalisable qu'au prix d'une perte d'information.

7.3.3.2 Relevance feedback

Pour être réellement efficace, la visualisation d'information doit être interactive. L'être humain est particulièrement habile à extraire des informations d'un environnement qu'il contrôle directement par rapport à un environnement qu'il ne peut

qu'observer de manière passive. Selon le psychologue GIBSON (1979), la perception est indissociable de l'action. Nous avons modifié notre système de navigation pour permettre à l'utilisateur de modifier lui-même l'agencement de la base en rectifiant, entre des objets, certaines similarités qui lui paraissent erronées. Des poids appliqués aux descripteurs sont alors ajustés afin que les similarités entre les objets correspondent au mieux aux modifications de l'utilisateur.

Le fonctionnement, inspiré des travaux de SANTINI ET AL. (2001) est montré dans la figure 7.8. A tout instant, l'utilisateur peut accéder, pour le groupe actuellement sélectionné, à une carte 2D. Les similarités entre les objets, représentées par leurs distances, sont ici plus précises que dans l'interface d'arbres de cônes. L'utilisateur peut ré-agencer les objets sur le plan, pour représenter des similarités qu'il considère comme plus pertinentes. Ensuite, il retourne au système de navigation. S'il n'a pas effectué de modifications, il retourne à la position où il se trouvait à l'origine. Par contre, si un nouvel agencement a été proposé, le système re-calcule un nouvel arbre plus en accord avec les nouvelles similarités tout en tenant compte des précédentes modifications.

La modification de la structure se fait en modifiant les poids attribués à chaque descripteur. Ceci revient à mettre en place une pondération globale de la base, qui ré-agencera toutes les classes de la base d'objets 3D. Le calcul des nouveaux poids s'effectue par minimisation de la fonction de coût suivante :

$$J = \sum_{i=0}^N \sum_{j=0}^i |J_{ij}| \quad \text{avec} \quad J_{ij} = d_{ij} - \sum_{k=0}^{dim} \sqrt{(W_k(X_{ik} - X_{jk}))^2} \quad (7.5)$$

où N le nombre d'objets considérés, dim le nombre total de descripteurs, X_{ik} la valeur du k^{ime} descripteur de l'objet i , W_k le poids associé au k^{ime} descripteur, et d_{ij} la distance entre les objets i et j estimée par l'utilisateur. Ceci permet de modifier les poids en fonction de la modification des distances entre objets fournie par l'utilisateur.

Le fait de modifier, durant le parcours, l'arbre en fonction des informations récupérées auprès de l'utilisateur n'est possible que grâce au système de classification que nous avons mis en place. Son coût faible permet de recalculer les classes en temps utilisateur et d'intégrer directement les informations collectées.

7.4 Expérimentations

Afin d'évaluer la pertinence de notre système de navigation, nous avons mis en place des tests quantitatifs visant à comparer notre système de classification aux nuées dynamiques floues et, d'autre part, des tests qualitatifs pour juger de l'ergonomie du système.

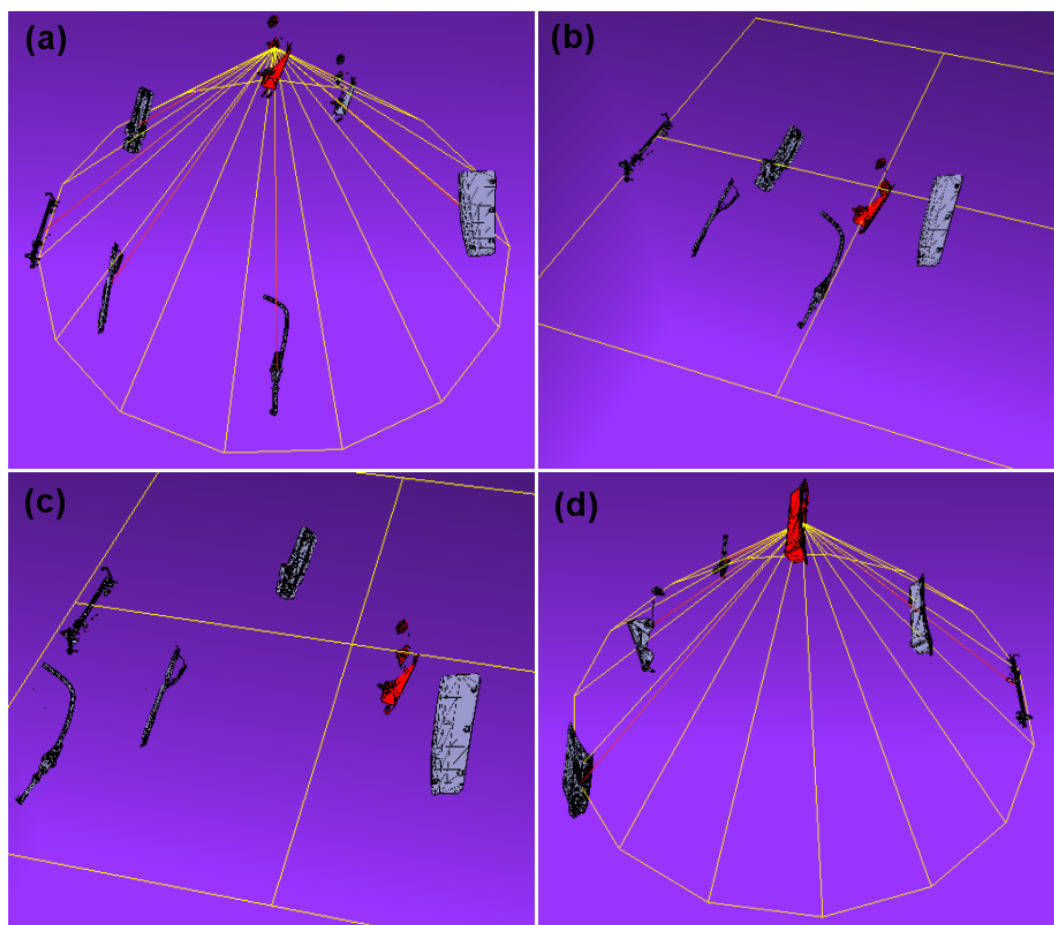


FIG. 7.8 – Processus de Relevance Feedback. (A) Structure originelle (B) Carte 2D des similarités des objets du groupe sélectionné (C) Modification de l'agencement des objets (D) Nouvelle structure calculée en tenant compte des modifications.

7.4.1 Critère de qualité de la partition

Les objets 3D sur lesquels nous travaillons dans le cadre du projet Semantic-3D sont les pièces mécaniques de la Lagune II préalablement classées. Le fait de connaître les groupes effectifs nous permet de pouvoir vérifier la qualité de notre partitionnement. Nous avons défini le pourcentage d'objets bien classés comme suit :

$$P(E) = \frac{\sum_{i=0}^G \max_{j \in [0..k]} C(P_i(E_j))}{N} \quad (7.6)$$

avec E la k -partition trouvée par l'algorithme testé, $C(P_i(E_j))$ le nombre de points du groupe réel i présents dans le groupe E_j , G le nombre de groupes réels et N le nombre d'objets du jeu d'essai considéré.

Jeu d'essai de 116 points	Qualité	Nombre de groupes trouvés	Groupes entiers	Temps d'exécution
Notre algorithme	91,37%	24	18/25	0,19 sec
Notre algorithme + Nuées dynamiques floues	92,24%	24	19/25	1,59sec
Nuées dynamiques floues	89,36%	25	17/25	1,07 sec

Jeu d'essai de 758 points	Qualité	Nombre de groupes trouvés	Groupes entiers	Temps d'exécution
Notre algorithme	77,17%	81	43/75	3,55 sec
Notre algorithme + Nuées dynamiques floues	80,47%	81	49/75	78,41 sec
Nuées dynamiques floues	74,23%	75	45/75	156,50 sec

TAB. 7.1 – Qualité de partitionnement et temps d'exécution comparés pour notre algorithme et les nuées dynamiques floues avec des jeux d'essais de 116 et de 758 points.

Ce critère est un critère qui favorise les partitions conservant le maximum de groupes entiers. Par exemple, si l'algorithme trouve un seul groupe contenant tous les objets, le pourcentage d'objets bien classés sera de 100. Pour contrebalancer ceci, nous affichons également le nombre de groupes trouvés par l'algorithme et le nombre de groupes réels qui ne sont pas fragmentés (i.e les groupes entiers).

7.4.2 Performance du système de navigation

Les expériences ont été faites sur deux bases d'objets 3D de la base Semantic-3D : une base de 116 objets répartis en 25 classes et une base de 758 objets regroupés en 75 classes. Le tableau 7.1 montre les performances des trois approches que nous avons testées sur les deux bases d'objets 3D.

Les tests ont été effectués sur un Pentium IV de 2,66GHz avec 512 Mo de Ram. Les résultats pour les nuées dynamiques floues sont une moyenne sur trois exécutions. On constate que notre algorithme donne des résultats sensiblement meilleurs pour des temps de calculs nettement inférieurs. Le fait que le temps d'exécution de notre algorithme suivi des nuées dynamiques floues soit parfois inférieur au temps d'exécution des nuées dynamiques floues seules peut être expliqué par le fait que les centres de groupe soient déjà placés de façon quasi optimale.

Pourcentage moyen d'objets bien classés	Notre algorithme	Notre algorithme + nuées dynamiques floues
Sans modification	91.38%	92.24%
1 ^{re} modification	90.18%	92.93%
2 ^{me} modification	89.91%	93.01%

TAB. 7.2 – Performances du système de Relevance Feedback en terme de pourcentage moyen d'objets bien classés suite aux modifications de la base par l'utilisateur.

7.4.3 Validation du système de visualisation

Afin de tester l'ergonomie de notre système de visualisation, nous avons fait tester celui-ci à un groupe de dix personnes ayant des compétences en informatique, mais aucune connaissance sur les systèmes de navigation. La perception de la qualité étant très subjective, ce test n'avait pas l'ambition de valider celle-ci, mais de faire apparaître les traits significatifs de notre système de visualisation. La façon d'utiliser l'interface a été montrée aux utilisateurs sans en expliquer les mécanismes sous-jacents. Ensuite, trois exercices leur ont été proposés. Le premier d'entre eux consistait à retrouver un objet bien classé dans la base, le second, un objet mal classé (i.e un objet que nous ne classerions pas intuitivement dans le groupe où le système le place). Enfin, le troisième consistait à effectuer deux modifications de la base en utilisant le système de Relevance Feedback. Pour les deux premiers tests, le nombre de clics souris leur permettant de parvenir à l'objet recherché a été comptabilisé. Pour le dernier test, le pourcentage d'objets bien classés nous a permis d'estimer l'amélioration moyenne apportée par une modification de l'utilisateur à la base.

Les résultats obtenus ont montré que la prise en main de l'interface n'est pas immédiate et nécessite quelques minutes d'utilisation. En moyenne, les utilisateurs trouvent un objet bien classé en 6.4 clics de souris pour une recherche optimale de 3 clics de souris. L'objet est donc trouvé presque immédiatement, ce qui signifie que le système de navigation offre, dans ce cas, une aide substantielle. L'objet mal classé est, par contre, trouvé en moyenne en 11 clics de souris pour une recherche optimale de 3 clics de souris. Bien que l'objet soit retrouvé dans tous les cas, l'aide apportée par le système de navigation dans ce cas est minime. Le pourcentage moyen d'objets bien classés suite aux modifications de la base par l'utilisateur est donné par le tableau 7.2.

On constate que l'aide apportée par le Relevance Feedback est faible et ce, quelle que soit la taille de la base. Ces résultats peuvent, peut être, s'expliquer par le fait des très bons résultats obtenus par notre processus de navigation sur la base d'objets Semantic-3D. D'autre part, on peut penser que le fait de modifier la totalité de la base d'après des changements sur un groupe correspond à modifier la base dans sa globalité d'après des informations locales. Ce qui améliore localement le classement,

mais le détériore dans une autre partie de la base.

7.5 Conclusion

Nous avons présenté ici un système de navigation offrant un algorithme de partitionnement efficace et très rapide. De plus, la possibilité de le coupler avec les nuées dynamiques floues permet d'obtenir une structure souple et plusieurs chemins existants pour accéder à un objet. Enfin, le système de visualisation 3D proposé présente plusieurs avantages. Outre sa clarté et sa simplicité d'utilisation, la représentation du chemin suivi dans l'arborescence permet à l'utilisateur de mieux appréhender l'organisation de la base et de la parcourir plus rapidement. La possibilité donnée à l'utilisateur d'interagir avec le système en modifiant l'agencement des objets en cours de navigation lui permet, si elle n'améliore que très peu la qualité de la structure, d'avoir une base plus en accord avec sa vision personnelle et donc, qui peut être parcourue plus facilement. Mais le caractère le plus novateur de ce système de navigation reste son application originale à des objets 3D ainsi que le développement inédit d'un système de visualisation 3D.

8

Conclusions et perspectives

Trois domaines de l'indexation ont été étudiés durant cette thèse : l'indexation d'images 2D, la recherche d'objets 3D par des requêtes 2D et l'indexation d'objets 3D. Pour ces trois points, nous présentons un résumé de notre contribution et quelques perspectives qui en découlent.

Indexation d'images 2D

Les travaux que nous avons réalisés en indexation d'images 2D portent à la fois sur l'étude de la couleur et de la forme. Nous avons proposé une généralisation de la transformation ART pour permettre l'indexation des formes issues d'images naturelles. Cette étude caractérise la couleur en extrayant les couleurs dominantes et en calculant une mesure de similarité associée. D'autre part, nous avons proposé une généralisation de la transformation ART pour la rendre robuste aux déformations perspectives. Ces deux points nous ont permis de développer une méthode de détection pseudo-supervisée de logos dans des images naturelles fonctionnant en fusionnant les études couleur et forme.

Les perspectives de ce travail sont de deux ordres. D'une part, l'approche permettant de faire un classement multi-critère, sur la forme et sur la couleur, donne des résultats supérieurs aux deux approches séparées et permet d'affiner les résultats. Il serait intéressant d'étudier plus précisément les mécanismes de fusion et tester plusieurs descripteurs pour évaluer les performances de ces procédés. D'autre part, l'application que nous avons faite étant pseudo-supervisée, elle requiert l'aide de l'utilisateur pour la sélection des zones d'intérêts des images et la segmentation des

objets. Notre système pourrait être couplé à une méthode de segmentation automatique, permettant d'extraire toutes les régions des images et de tester leur similarité à un logo de la base.

Indexation 2D/3D

La recherche d'objets 3D par des requêtes 2D a été abordée d'un point de vue novateur. Notant les limites de ces approches quant aux choix et au nombre des vues caractéristiques, nous avons proposé une méthode de comparaison dynamique se basant sur un système de rendu spectral. L'intégralité des vues de l'objet 3D est stockée sous forme d'un spectre 3D, ce qui permet d'extraire rapidement les vues et de parcourir dynamiquement l'espace des vues pour rechercher la plus proche. Ce procédé a été utilisé pour faire de la mise en correspondance entre une vue requête et un objet 3D et pour faire de la recherche dans une base d'objets 3D à partir d'une requête 2D.

Cette approche a montré son fort potentiel, mais peut être approfondie de différentes façons. A court terme, une perspective serait de faire la comparaison entre la vue extraite et la vue requête dans l'espace réel. Ceci oblige de transformer les coupes spectrales extraites dans ce domaine. Il serait moins coûteux de transformer la vue requête dans l'espace spectral et de mesurer la similarité entre les spectres de Fourier. Ceci serait possible, à condition de « binariser » les spectres extraits, car ils représentent les transformées des images profondeurs alors que le spectre requête représente une vue binaire. Il serait également intéressant de définir une mesure de similarité entre spectres pour garantir des résultats efficaces dans cette espace. D'autre part, une perspective à plus long terme serait d'améliorer le processus d'optimisation en utilisant les propriétés de la transformée de Fourier. Une recherche à plusieurs niveaux de précisions permettrait de faire un premier parcours rapide de l'espace des vues pour sélectionner des zones pouvant potentiellement contenir des minima locaux.

Indexation 3D

L'indexation d'objets 3D a été abordée suivant deux approches durant cette thèse. Dans un premier temps, nous avons étudié les approches par transformées et proposé une nouvelle méthode d'indexation basée sur les moments. Notre méthode extrait de façon rapide et précise une information bas-niveau caractérisant la forme des objets et fournit une approche intéressante pour indexer rapidement de grandes bases d'objets 3D. D'autre part, nous avons étudié les approches d'indexations d'objets 3D par vues et comparé les performances obtenues en fonction du type d'images caractéristiques choisies et du nombre de vues choisies.

Cette étude ouvre différentes perspectives. D'une part, il serait intéressant de modifier la métrique utilisée lors de l'indexation d'objets 3D par vues, ainsi que le descripteur de forme 2D employé. Nous avons constaté une réduction des performances à mesure qu'augmente le nombre de vues. Ceci est dû au fait que la mesure de similarité et que le descripteur de formes n'arrivent pas à sommer correctement les informations extraites des vues. Une piste de recherche intéressante, serait de tester d'autres descripteurs de formes et d'autres métriques associées, pour augmenter les résultats. D'autre part, comme en indexation 2D/3D, des futurs travaux pourraient consister à créer un descripteur de forme applicable sur les spectres extraits, ce qui permettrait de ne pas avoir à reconstruire les images. Au delà de ces perspectives à court terme sur la méthode que nous avons proposée, la nature même de l'indexation d'objets 3D par vues est une voie de recherche à approfondir. Le fait de caractériser les vues d'un objet 3D par des descripteurs de formes 2D bas-niveau et de posséder une représentation de plus haut-niveau contenue dans l'ensemble des vues caractéristiques, permet de tirer profit de ces deux descriptions. Ceci permet d'avoir une caractérisation rapide et efficace des images tout en décrivant la structure de l'objet 3D.

En parallèle des deux méthodes d'indexation d'objets 3D étudiées, nous avons travaillé sur le calcul d'expression analytique sur un maillage 3D et sur un système de navigation dans une base d'objets 3D. De nombreuses méthodes d'analyse d'objets 3D calculent des fonctions intégrales sur le volume d'un objet 3D et pour cela les discrétisent. Nous avons montré qu'il est également possible de calculer ces intégrales sur les frontières des objets et d'obtenir l'expression analytique exacte. Nous avons en particulier calculé l'expression analytique de la transformée de Fourier d'un objet 3D et l'avons utilisé en analyse de formes. Rendering.

Les perspectives de ces travaux seraient d'étendre le calcul d'expression analytique à d'autres fonctions intégrales, comme nous l'avons fait avec la transformée de Fourier. De nombreux descripteurs de formes utilisent une intégrale sur le volume des objets pour extraire une description bas-niveau. On peut en particulier citer le descripteur ART 3D, qu'il serait intéressant d'étudier. L'utilisation de ces approches permet d'obtenir l'expression exacte sur le volume d'un objet 3D et ouvre une nouvelle voie de recherche qu'il faudra dans le futur continuer à étudier.

Enfin, nous avons proposé un système de navigation dans des bases d'objets 3D. Les limitations de la recherche d'objets 3D, par le fait qu'il faille posséder une requête et que les résultats affichés soient sous forme linéaire, ne permettent pas d'appréhender la base dans sa globalité. Notre système structure la base d'objets 3D et fournit un affichage en trois dimensions des classes d'objets similaires. La structuration de la base et la visualisation se font de façon indépendante par rapport aux descripteurs choisis pour caractériser la base.

Les perspectives issues de l'étude du système de navigation sont à rapprocher de l'analyse sur la fusion de similarité faite en indexation de formes 2D. Ils semblent en effet intéressant d'étudier les possibilités de naviguer sur plusieurs descriptions en

parallèle et par la même de fusionner les navigations. Comme nous l'avons vu, les descripteurs de natures différentes (structurel, par vue et par transformée) fournissent plusieurs niveaux de description et apportent des informations différentes quant aux similarités entre objets. Un système de navigation basé sur plusieurs descripteurs permettra de combiner les avantages de chacun des descripteurs.

Quatrième partie

Annexes

Contenu des annexes

A	Développement d'une plate-forme logicielle pour Semantic-3D	173
A.1	Introduction	173
A.2	Usages	173
A.3	Architecture	174
A.4	Exemple de fonctionnement	175
A.4.1	Module de compression	176
A.4.2	Module d'indexation	176
A.4.3	Module de tatouage	176
A.4.4	Organisation des modules	176
A.4.5	Exemple d'exécution	177
A.5	Conclusion	177
B	Résolution de la transformée de Fourier d'un tétraèdre	179
C	Liste des publications de l'auteur	187
D	Index des auteurs cités	191



Développement d'une plate-forme logicielle pour Semantic-3D

A.1 Introduction

Cette thèse s'est déroulée dans le cadre du projet RNRT Semantic-3D. Le but de ce projet était de développer des nouvelles fonctionnalités en indexation, tatouage et compression d'objets 3D et plus spécialement d'utiliser ces nouvelles fonctionnalités dans une application d'échange de modèles 3D mécaniques.

Pour cela, nous avons défini, avec notre partenaire industriel Renault, les scénarii d'usages et les fonctionnalités de l'application en fonction des besoins de notre partenaire. L'application mise en place a pour but la création et la consultation de méthodes de réparations 3D, consultable sur de nombreux terminaux. Elle contient en particulier, des fonctionnalités de tatouage, de compression et d'indexation pour permettre l'échange facile et sécurisé de modèles 3D.

A.2 Usages

Les objectifs attendus par notre partenaire Renault sont d'une part de contribuer à l'amélioration des services liés à la maintenance des véhicules, et d'autre part de participer à la réduction des coûts de la production de la documentation technique. Les usages envisagés dans le cadre du projet se sont ainsi concentrés sur les activités de réalisation de la documentation technique du véhicule et de sa consultation. Après

analyse des besoins, ils ont été caractérisés en cinq services principaux :

- **rédaction de la documentation technique** : le futur processus de description de la documentation technique se substituera au processus actuel de rédaction textuelle de documents. Le coût de « rédaction » de cette documentation devra rester dans des proportions comparables à celles de l'existant.
- **Continuité numérique des informations** : il s'agit de mettre au point un processus de diffusion de la documentation technique utilisant les données amonts, créées et utilisées par les bureaux d'études. Ces données se distinguent entre les données IAO (Ingénierie Assistée par Ordinateur) et celles qui permettent de décrire la structure et la composition du véhicule.
- **Consultation de la documentation technique** : la consultation de la documentation technique concerne l'accès aux informations nécessaires à la réalisation d'une intervention technique. On retrouve notamment les méthodes de réparation, les pièces de rechange (avec des moyens de recherche par photographie numérique ou par numérisation de pièce), les notices d'utilisation, ... Cette consultation se fera via différents supports de visualisation (PC, PDA, Terminal véhicule) et différents canaux de communication (Intranet/Internet, Wifi, GPRS,...).
- **Protection des données transmises** : l'échange de nombreuses informations liées à la modélisation, à la structuration d'un véhicule et à la description de sa documentation technique, nécessite une protection forte, et un contrôle de sa dissémination possible.
- **Réalisation d'un diagnostic** : l'une des principales tâches d'un technicien lors de son intervention technique est l'identification de l'origine d'un problème sur le véhicule.

Pour répondre à ce besoin, et dans les limites du périmètre technique du projet, le système pourra assister l'intervenant en exploitant la description de la documentation technique. Différents scénarii d'usage ont été mis en oeuvre dans le cadre du projet de manière à illustrer tout ou partiellement ces services.

A.3 Architecture

L'architecture du système mis en place permet une utilisation optimale des réseaux (charge, gestion de cache,...) et une adaptation au média de consultation. Le mode d'implantation permet aux modules composant le système de dialoguer sans se soucier des aspects de communication (réseau, local). Six entités métiers ou fonctionnelles ont été identifiées (voir figure A.1) :

- serveur Métier RENAULT : en charge des données CAO (conversion CATIA, fourniture des méta-données) ;
- serveur Documentaire : en charge du stockage des méthodes de réparation ainsi que des données associées (géométries, méta-données,...) ;
- poste de Rédaction : en charge de la gestion des manuels de réparations (créa-

- tion, mise à jour, .. .);
- serveur de Consultation : en charge de la consultation des méthodes de réparation avec la gestion des identifications, des caches réseaux ;
 - poste Concession/Assistance : en charge de la consultation pour les demandes d'assistances ou au sein de la concession ;
 - médias de consultation : ceci regroupe toutes les entités susceptibles de se connecter et d'interagir avec le système SEMANTIC (PDA, tablette PC, ordinateur de bureau,...).

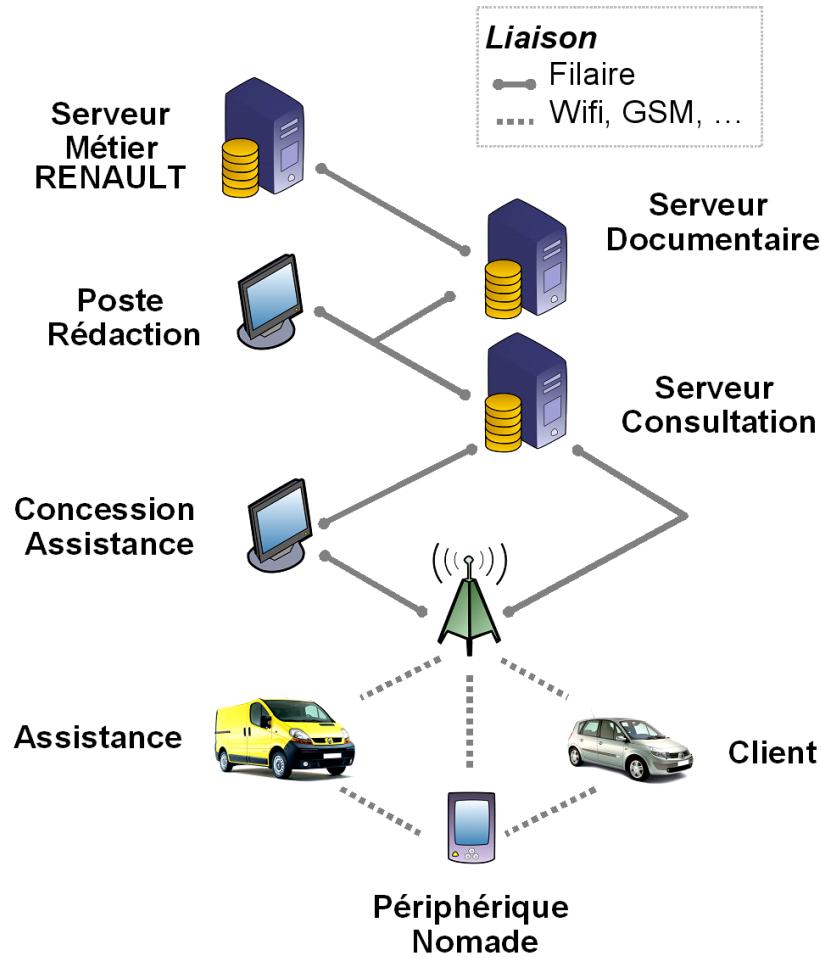


FIG. A.1 – Architecture globale de l'application Semantic-3D.

A.4 Exemple de fonctionnement

Dans le cadre du projet RNRT Semantic-3D, il a été réalisé une plate-forme de démonstration permettant de valider la faisabilité du projet et des méthodes

choisies. L'application est organisée autour de l'architecture définie par le sous-projet 4 et des trois modules de compression, d'indexation et de tatouage, regroupant les fonctionnalités de l'application.

A.4.1 Module de compression

Le module de compression est le fruit du travail réalisé au sein du sous-projet 1 par les partenaires INT et LIRIS. Le but de ce module est de mettre en place des nouvelles méthodes de compression permettant de diminuer la taille des données à transmettre sur les réseaux utilisés par l'application (Internet, Wifi, GPRS). Ces méthodes s'adaptent aux qualités des réseaux et permettent de faire de la transmission progressive.

A.4.2 Module d'indexation

Le module d'indexation contient les méthodes d'indexation d'objets 3D par des requêtes 2D et par des requêtes 3D, réalisées par les trois partenaires du sous-projet 2 : le LIRIS, l'INT et l'ENIC. Plusieurs méthodes ont été proposées et nous avons opté pour l'intégration dans l'application de plusieurs méthodes pour permettre de les comparer. Les méthodes intégrées dans le module d'indexation sont :

- le descripteur de formes 3D, Hough 3D dans sa version quantifiée et non quantifiée (ZAHARIA et PRÊTEUX, 2004) ;
- le descripteur de formes 3D, ART 3D (5.2) ;
- le descripteur de formes 3D par vues, AVC 3D/3D (FILALI ANSARY ET AL., 2005) ;
- le descripteur de formes 2D/3D, AVC 2D/3D (FILALI ANSARY ET AL., 2005).

A.4.3 Module de tatouage

Le module de tatouage a été réalisé par le sous-projet 3 par les laboratoires INT et EURECOM. Il regroupe les méthodes de tatouages faites par les deux partenaires. Les méthodes développées au sein du projet permettent d'insérer une marque dans un objet 3D sans modifier l'aspect du maillage. Cette marque permet d'identifier l'objet de façon certaine et doit être robuste aux déformations pouvant être subies par l'objet.

A.4.4 Organisation des modules

L'application générale et la structure des protocoles de communication ont été organisées par le sous-projet 4 par les partenaires industriels Renault et Mercury. Comme nous l'avons vu précédemment, l'architecture générale de l'application est

structurée autour de trois serveurs : le serveur métier, le serveur documentaire et serveur de consultation. La figure A.2 représente l'architecture de l'application avec les différents objets stockés sur les serveurs.

Le serveur de contenu est extérieur à l'application et est géré par notre partenaire Renault. Il est le point d'entrée de l'application pour les modèles CAO. Les modèles sont convertis par ce serveur pour répondre au besoin d'affichage et envoyés au serveur documentaire qui les stocke dans leurs versions allégées. Le but de ce serveur est de garantir la sécurité des données Renault en filtrant les données sorties. Le serveur documentaire décrit les modèles qu'il stocke, soit par des méthodes d'indexations, soit par des méthodes de réparation rédigées sur le poste de rédaction. Les données sont ensuite envoyées sur le serveur de consultation pour être transmises sur les postes de réparation. Les données véhiculant sur les réseaux sont compressées et tatouées par le serveur de documentaire et stockées sur le serveur de consultation pour être directement envoyées.

A.4.5 Exemple d'exécution

L'application finale est en cours de finalisation pour être présentée lors de la réunion finale du projet. Il n'est pas possible à l'heure actuelle de montrer les différentes fonctionnalités. Celles-ci seront rajoutées dans la version finale de la thèse et présentée durant la soutenance.

A.5 Conclusion

Ce projet m'a donné un sacré bonus pour cette thèse. Le fait de réaliser une thèse dans le cadre d'un projet de recherche exploratoire était réellement motivant. D'autant plus que ce projet RNRT avait des objectifs d'étude de faisabilité dans le monde industriel. Le fait de travailler avec des industriels sur un projet de grandes ampleurs et avec de nombreux partenaires académiques, a permis de nombreux échanges toujours bénéfiques. Les partenaires industriels ont amené le dynamisme et un cadre « professionnel », alors que les laboratoires académiques fournissaient un environnement propice aux confrontations scientifiques.

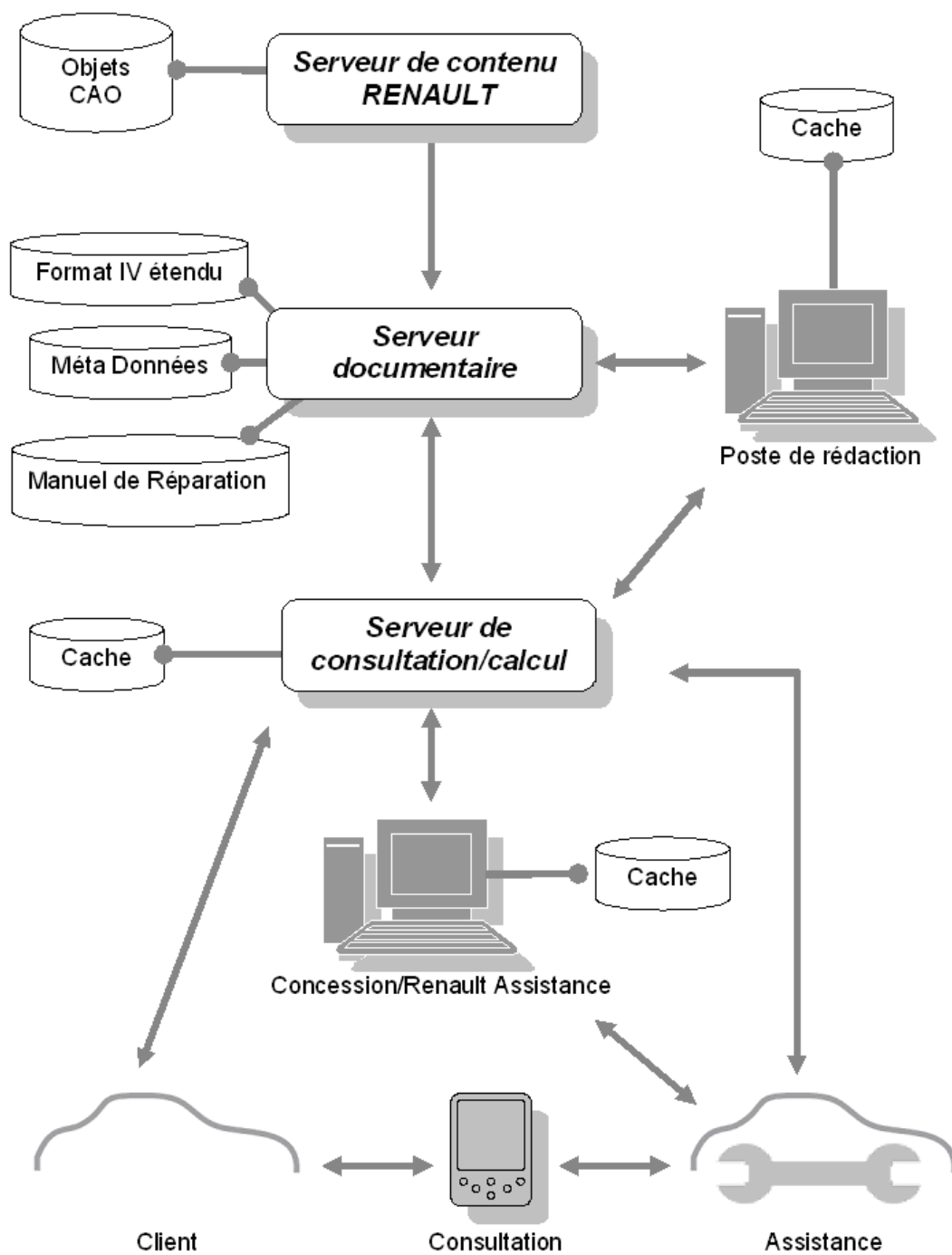


FIG. A.2 – Architecture générale de l'application.

B

Résolution de la transformée de Fourier d'un tétraèdre

Comme nous l'avons vue dans le chapitre 6, la transformée de Fourier d'un objet 3D maillé peut se calculer directement sur le maillage, ceci est faisable si le maillage est une variété topologique (sans trous, orientable...). Sous cette condition, il faut intégrer sur le volume de l'objet la fonction :

$$f(x, y, z) = e^{-i(u(x_1X+x_2Y+x_3Z)+v(x_2X+y_2Y+z_2Z)+w(z_1X+z_2Y+z_3Z))}. \quad (\text{B.1})$$

Cela revient à intégrer sur le tétraèdre unité selon l'expression :

$$F(u, v, w) = \int_0^1 \int_0^{1-Z} \int_0^{1-Y-Z} e^{-i \begin{pmatrix} u(x_1X + x_2Y + x_3Z) \\ +v(y_1X + y_2Y + y_3Z) \\ +w(z_1X + z_2Y + z_3Z) \end{pmatrix}} dXdYdZ \quad (\text{B.2})$$

L'expression générale à intégrer et la solution générale peuvent s'écrire :

$$\begin{aligned} F(u, v, w) &= |T| \int_0^1 \int_0^{1-Z} \int_0^{1-Y-Z} e^{-i(u(x_1X+x_2Y+x_3Z)+v(y_1X+y_2Y+y_3Z)+w(z_1X+z_2Y+z_3Z))} dXdYdZ \\ &= |T| \left[\frac{ie^{-iU_1}}{U_1(U_1-U_2)(U_1-U_3)} + \frac{ie^{-iU_2}}{U_2(U_2-U_1)(U_2-U_3)} \right. \\ &\quad \left. + \frac{ie^{-iU_3}}{U_3(U_3-U_2)(U_3-U_1)} - \frac{i}{U_1U_2U_3} \right] \end{aligned} \quad (\text{B.3})$$

avec :

$$\begin{cases} U_1 &= ux_1 + vy_1 + wz_1 \\ U_2 &= ux_2 + vy_2 + wz_2 \\ U_3 &= ux_3 + vy_3 + wz_3 \end{cases} \quad (\text{B.4})$$

Cette solution générale n'est pas applicable en tout point de l'espace (x, y, z) pour toutes les valeurs de (u, v, w) . Il faut rechercher les solutions où cette expression n'est pas définie et modifier la fonction à intégrer en fonction de ces solutions pour obtenir une définition analytique exacte en tout point.

Les solutions annulant le dénominateur de l'expression B.3 sont :

$$\begin{aligned}
 S = & \{ \{u = 0, v = 0, w = 0\}, \\
 & \{u = 0, v = 0, z_1 = 0\}, \{u = 0, v = 0, z_2 = 0\}, \{u = 0, v = 0, z_3 = 0\}, \\
 & \{u = 0, v = 0, z_1 = z_3\}, \{u = 0, v = 0, z_2 = z_3\}, \{u = 0, v = 0, z_1 = z_2\}, \\
 & \left\{ u = 0, y_1 = -\frac{wz_1}{v} \right\}, \left\{ u = 0, y_2 = -\frac{wz_2}{v} \right\}, \left\{ u = 0, y_3 = -\frac{wz_3}{v} \right\}, \\
 & \left\{ u = 0, y_2 = -\frac{-y_3v - wz_3 + wz_2}{v} \right\}, \\
 & \left\{ u = 0, y_1 = -\frac{-y_3v - wz_3 + wz_1}{v} \right\}, \\
 & \left\{ u = 0, y_1 = -\frac{-wz_2 + wz_1 - y_2v}{v} \right\}, \\
 & \left\{ x_3 = -\left(\frac{y_3v + wz_3}{u}\right) \right\}, \left\{ x_2 = -\frac{y_2v + wz_2}{u} \right\}, \left\{ x_1 = -\frac{y_1v + wz_1}{u} \right\}, \\
 & \left\{ x_1 = -\frac{-wz_2 + y_1v - x_2u + wz_1 - y_2v}{u} \right\}, \\
 & \left\{ x_1 = -\frac{-y_3v - wz_3 - x_3u + y_1v + wz_1}{u} \right\}, \\
 & \left. \left\{ x_2 = -\frac{-y_3v - wz_3 - x_3u + y_2v + wz_2}{u} \right\} \right\} \tag{B.5}
 \end{aligned}$$

Il faut pour chaque solution, notée S_i , modifier la fonction f en f_i , en fonction de la solution et exprimer l'intégrale de f_i , notée F_i . Dans certain cas, l'intégrale F_i n'est pas définie en tout point, et il faut réitérer le processus en fonction des valeurs pour lesquelles elle n'est pas définie :

- F s'annule si : $S_1 = \{u = 0, v = 0, w = 0\}$, sous cette condition :
 $f_1 = 1 \Rightarrow F_{S_1} = \frac{1}{6}$
- F s'annule si : $S_2 = \{u = 0, v = 0, z_2 = 0\}$, sous cette condition :
 $f_2 = e^{-Iw(z_1X+z_2Y+z_3Z)}$
 $\Rightarrow F_2 = \frac{-i \left(\begin{aligned} & e^{-iwz_3} z_1^2 z_2 - z_1^2 z_2 - z_1^2 e^{-iwz_2} z_3 + z_1^2 z_3 - e^{-iwz_3} z_2^2 z_1 + z_2^2 z_1 \\ & + z_1 e^{-iwz_2} z_3^2 - z_1 z_3^2 + z_2^2 e^{-iwz_1} z_3 - z_2^2 z_3 - z_2 e^{-iwz_1} z_3^2 + z_2 z_3^2 \end{aligned} \right)}{z_2 z_3 w^3 (-z_3 + z_1)(-z_3 + z_2)(-z_2 + z_1) z_1}$
- F_2 s'annule si : $S_{2,1} = \{z_3 = 0\}$
 $\Rightarrow F_{2,1} = -I \frac{2e^{(-Iwz_1)-2+2Iz_1w+z_1^2w^2}}{2w^3 z_1^3}$

-
- $F_{2,1}$ s'annule si : $\{z_1 = 0\} \Rightarrow F_{2,1,1} = \frac{1}{6}$
 - F_2 s'annule si : $S_{2,2} = \{z_1 = 0\}$
 $\Rightarrow F_{22} = -I \frac{w^2 z_3^2 - 2 + 2Iwz_3 + 2e^{-Iwz_3}}{2z_3^3 w^3}$
 - F_2 s'annule si : $S_{2,3} = \{z_1 = z_3\}$
 $\Rightarrow F_{23} = -I \frac{w^2 z_3^2 - 2 + 2Iwz_3 + 2e^{-Iwz_3}}{2z_3^3 w^3}$
 - F s'annule si : $S_2 = \{u = 0, v = 0, z_3 = 0\}$, sous cette condition :
 $f_3 = e^{-Iw(z_1 X + z_2 Y)} \Rightarrow F_3 = -\frac{-Ie^{-Iwz_2} z_1^2 + Ie^{-Iwz_1} z_2^2 - Iz_2^2 + z_1^2 w z_2 + Iz_1^2 - z_2^2 z_1 w}{w^3 z_2^2 (z_1 - z_2) z_1^2}$
 - F_3 s'annule si : $S_{3,1} = \{z_1 = 0\}$
 $\Rightarrow F_{3,1} = -\frac{I(2e^{-Iwz_2} + z_2^2 w^2 + 2Iz_2 w - 2)}{2w^3 z_2^3}$
 - F_3 s'annule si : $S_{3,2} = \{z_1 = z_2\}$
 $\Rightarrow F_{3,2} = -\frac{-Ie^{-Iwz_2} + e^{-Iwz_2} z_2 w + z_2 w + 2I}{2z_2^3 w^3}$
 - F s'annule si : $S_4 = \{u = 0, v = 0, z_1 = 0\}$, sous cette condition :
 $f_4 = e^{-Iw(z_2 Y + z_3 Z)} \Rightarrow F_4 = -\frac{z_2^2 z_3 w - z_2 z_3^2 w + Ie^{-Iwz_2} z_3^2 + Iz_2^2 - Iz_3^2 - Ie^{-Iwz_3} z_2^2}{z_3^2 w^3 (z_2 - z_3) z_2^2}$
 - F_4 s'annule si : $S_{4,1} = \{z_2 = z_3\}$
 $\Rightarrow F_{4,1} = -\frac{wz_3 + 2I + e^{-Iwz_3} w z_3 - 2Ie^{-Iwz_3}}{w^3 z_3^3}$
 - F s'annule si : $S_5 = \{u = 0, v = 0, z_1 = z_3\}$, sous cette condition :
 $f_5 = e^{-Iw(z_3 X + z_2 Y + z_3 Z)}$
 $\Rightarrow F_5 = \frac{Iz_2^2 - 2Iz_2 z_3 + Iz_3^2 - Ie^{-Iwz_2} z_3^2 - Ie^{-Iwz_3} z_2^2 - e^{-Iwz_3} z_2 z_3^2 w + 2Ie^{-Iwz_3} z_2 z_3 + e^{-Iwz_3} z_2^2 z_3 w}{z_2 (z_2 - z_3)^2 z_3^2 w^3}$
 - F_5 s'annule si : $S_{5,1} = \{z_2 = z_3\}$
 $\Rightarrow F_{5,1} = \frac{I(2 - 2e^{-Iwz_3} - 2Ie^{-Iwz_3} w z_3 + e^{-Iwz_3} w^2 z_3^2)}{w^3 z_3^3}$
 - F s'annule si : $S_6 = \{u = 0, v = 0, z_2 = z_3\}$, sous cette condition :
 $f_6 = e^{-Iw(z_1 X + z_3 Y + z_3 Z)}$
 $\Rightarrow F_6 = \frac{-Iz_1 + 2Iz_3 - z_3^2 w + z_1 w z_3}{w^3 z_3^2 (z_1 - z_3)^2} e^{-Iwz_3} + \frac{I(z_1^2 - 2z_1 z_3 + z_3^2 - z_3^2 e^{-Iwz_1})}{w^3 z_3^2 (z_1 - z_3)^2 z_1}$
 - F s'annule si : $S_7 = \{u = 0, v = 0, z_1 = z_2\}$, sous cette condition :
 $f_7 = e^{-Iw(z_2 X + z_2 Y + z_3 Z)}$
 $\Rightarrow F_7 = -\frac{(-2Iz_2 + Iz_3 + z_2^2 w - z_2 z_3 w) e^{-Iwz_2}}{w^3 (z_2 - z_3)^2 z_2^2} - \frac{Ie^{-Iwz_3}}{w^3 z_3 (z_2 - z_3)^2} + \frac{I}{z_2^2 w^3 z_3}$
 - F s'annule si : $S_8 = \{u = 0, y_1 = \frac{-wz_1}{v}\}$, sous cette condition :
 $f_8 = e^{-I(y_2 Y v + y_3 Z v + w z_2 Y + w z_3 Z)}$
 $\Rightarrow F_8 = -\frac{Ie^{-I(y_2 v + z_2 w)}}{(y_2 v + z_2 w)^2 (y_2 v + z_2 w - v y_3 - w z_3)} + \frac{Ie^{-I(v y_3 + w z_3)}}{(v y_3 + w z_3)^2 (y_2 v + z_2 w - v y_3 - w z_3)}$

- $$- \frac{y_2 y_3 v^2 + I v y_2 + y_2 v w z_3 + v z_2 w y_3 + I v y_3 + z_2 w^2 z_3 + I w z_2 + I w z_3}{(v y_3 + w z_3)^2 (y_2 v + z_2 w)^2}$$

– F_8 s'annule si : $S_{8,1} = \{y_3 = \frac{-w z_3}{v}\}$,
 sous cette condition : $f_{8,1} = e^{-IY(y_2 v + w z_2)}$
 $\Rightarrow F_{8,1} = \frac{-I e^{-I(y_2 v + z_2 w)}}{(y_2 v + z_2 w)^3} - \frac{I(2I v y_2 - 2 + z_2^2 w^2 + 2y_2 v z_2 w + 2I z_2 w + y_2^2 v^2)}{2(y_2 v + z_2 w)^3}$
- $F_{8,1}$ s'annule si : $S_{8,1.1} = \{y_2 = -z_2 w / v\} \Rightarrow F_{8,1.1} = \frac{1}{6}$
- F_8 s'annule si : $S_{8,2} = \{y_2 = \frac{-z_2 w}{v}\}$,
 sous cette condition : $f_{8,2} = e^{-IZ(y_3 v + w z_3)}$
 $\Rightarrow F_{8,2} = \frac{-I e^{-I(v y_3 + w z_3)}}{(v y_3 + w z_3)^3} - \frac{I(v^2 y_3^2 + 2v y_3 w z_3 + w^2 z_3^2 + 2I w z_3 - 2 + 2I v y_3)}{2(v y_3 + w z_3)^3}$
- F_8 s'annule si : $S_{8,3} = \{y_2 = \frac{v y_3 + w z_3 - z_2 w}{v}\}$,
 sous cette condition : $f_{8,3} = e^{-I(y_3 v + w z_3)(Y+Z)}$
 $\Rightarrow F_{8,3} = \frac{-(w z_3 + v y_3 - 2I) e^{-I(v y_3 + w z_3)}}{(v y_3 + w z_3)^3} - \frac{v y_3 + 2I + w z_3}{(v y_3 + w z_3)^3}$
- F s'annule si : $S_9 = \{u = 0, y_2 = \frac{-w z_2}{v}\}$,
 sous cette condition : $f_9 = e^{-I(y_1 X v + y_3 Z v + w z_1 X + w z_3 Z)}$
 $\Rightarrow F_9 = \frac{-I e^{-I(y_1 v + z_1 w)}}{(y_1 v + z_1 w)^2 (y_1 v + z_1 w - v y_3 - w z_3)} + \frac{I e^{-I(v y_3 + w z_3)}}{(y_1 v + z_1 w - v y_3 - w z_3)(v y_3 + w z_3)^2}$

$$- \frac{y_1 y_3 v^2 + y_1 v w z_3 + I y_1 v + I v y_3 + v y_3 z_1 w + I w z_3 + w^2 z_3 z_1 + I z_1 w}{(v y_3 + w z_3)^2 (y_1 v + z_1 w)^2}$$
- F_9 s'annule si : $S_{9,1} = \{y_3 = \frac{-w z_3}{v}\}$,
 sous cette condition : $f_{9,1} = e^{-IX(y_1 v + z_1 w)}$
 $\Rightarrow F_{9,1} = \frac{-I e^{-I(y_1 v + z_1 w)}}{(y_1 v + z_1 w)^3} - \frac{I(2I y_1 v + 2I z_1 w - 2 + w^2 z_1^2 + v^2 y_1^2 + 2y_1 v z_1 w)}{2(y_1 v + z_1 w)^3}$
- F_9 s'annule si : $S_{9,1} = \{y_1 = \frac{v y_3 + w z_3 - z_1 w}{v}\}$,
 sous cette condition : $f_{9,2} = e^{-I(y_3 v + w z_3)(X+Z)}$
 $\Rightarrow F_{9,2} = \frac{-(w z_3 + v y_3 - 2I) e^{-I(v y_3 + w z_3)}}{(v y_3 + w z_3)^3} - \frac{v y_3 + 2I + w z_3}{(v y_3 + w z_3)^3}$
- F s'annule si : $S_{10} = \{u = 0, y_3 = \frac{-w z_3}{v}\}$,
 sous cette condition : $f_{10} = e^{-I(y_1 X v + y_2 Y v + w z_1 X + w z_2 Y)}$
 $\Rightarrow F_{10} = \frac{I e^{-I(y_2 v + z_2 w)}}{(y_2 v + z_2 w)^2 (y_1 v - y_2 v + z_1 w - z_2 w)} - \frac{I e^{-I(y_1 v + z_1 w)}}{(y_1 v + z_1 w)^2 (y_1 v - y_2 v + z_1 w - z_2 w)}$

$$- \frac{y_2 v^2 y_1 + I y_1 v + z_2 w y_1 v + I v y_2 + y_2 v z_1 w + z_2 w^2 z_1 + I w z_2 + I z_1 w}{(y_2 v + z_2 w)^2 (y_1 v + z_1 w)^2}$$
- F_{10} s'annule si : $S_{10,1} = \{y_1 = \frac{y_2 v - z_1 w + z_2 w}{v}\}$,
 sous cette condition : $f_{10,1} = e^{-I(v y_2 + w z_2)(X+Y)}$
 $\Rightarrow F_{10,1} = \frac{-(y_2 v + z_2 w - 2I) e^{-I(y_2 v + z_2 w)}}{(y_2 v + z_2 w)^3} - \frac{y_2 v + 2I + z_2 w}{y_2 v + z_2 w)^3}$
- F s'annule si : $S_{11} = \{u = 0, y_2 = -\frac{-v y_3 - w z_3 + w z_2}{v}\}$,
 sous cette condition : $f_{11} = e^{I(y_1 X v + Y y_3 v + Y w z_3 + y_3 Z v + w z_1 X + w z_3 Z)}$
 $\Rightarrow F_{11} = \frac{-I e^{-I(y_1 v + z_1 w)}}{(y_1 v + z_1 w - v y_3 - w z_3)^2 (y_1 v + z_1 w)} + \frac{I}{(v y_3 + w z_3)^2 (y_1 v + z_1 w)}$

$$+ \frac{(y_1 y_3 v^2 - I v y_1 + y_1 v w z_3 - v^2 y_3^2 + 2 I v y_3 - 2 v y_3 w z_3 + v y_3 z_1 w - w^2 z_3^2 - I w z_1 + w^2 z_3 z_1 + 2 I w z_3) e^{-I(v y_3 + w z_3)}}{(v y_3 + w z_3)^2 (y_1 v + z_1 w - v y_3 - w z_3)^2}$$

- F_{11} s'annule si : $S_{11.1} = \{y_1 = \frac{v y_3 + w z_3 - z_1 w}{v}\}$,
sous cette condition : $f_{11.1} = e^{-I(y_3 v + w z_3)}(X + Y + Z)$
 $\Rightarrow F_{11.1} = \frac{I(-2 I v y_3 - 2 I w z_3 + w^2 z_3^2 - 2 + 2 v y_3 w z_3 + v^2 y_3^2) e^{-I(v y_3 + w z_3)}}{2(v y_3 + w z_3)^3} + \frac{I}{(v y_3 + w z_3)^3}$

- F s'annule si : $S_{12} = \{u = 0, y_1 = -\frac{-v y_3 - w z_3 + w z_1}{v}\}$,
sous cette condition : $f_{12} = e^{-I(X y_3 v + X w z_3 + y_2 Y v + y_3 Z v + w z_2 Y + w z_3 Z)}$
 $\Rightarrow F_{12} = \frac{-I e^{-I(y_2 v + z_2 w)}}{(y_2 v + z_2 w - v y_3 - w z_3)^2 (y_2 v + z_2 w)} + \frac{I}{(v y_3 + w z_3)^2 (y_2 v + z_2 w)}$
 $+ \frac{(y_2 y_3 v^2 + y_2 v w z_3 - I v y_2 - v^2 y_3^2 + v z_2 w y_3 - 2 v y_3 w z_3 + 2 I v y_3 - I w z_2 - w^2 z_3^2 + z_2 w^2 z_3 + 2 I w z_3) e^{-I(v y_3 + w z_3)}}{(y_2 v + z_2 w - v y_3 - w z_3)^2 (v y_3 + w z_3)^2}$

- F s'annule si : $S_{13} = \{u = 0, y_1 = -\frac{-v y_2 - w z_2 + w z_1}{v}\}$,
sous cette condition : $f_{13} = e^{-I(X v y_2 + X w z_2 + y_2 Y v + y_3 Z v + w z_2 Y + w z_3 Z)}$
 $\Rightarrow F_{13} = \frac{-(-y_2 y_3 v^2 + I v y_3 - v z_2 w y_3 + y_2^2 v^2 - 2 I y_2 v - y_2 v w z_3 + 2 y_2 v z_2 w - 2 I w z_2 + I w z_3 + z_2^2 w^2 - z_2 w^2 z_3) e^{-I(y_2 v + z_2 w)}}{(y_2 v + z_2 w)^2 (y_2 v + z_2 w - v y_3 - w z_3)^2}$
 $- \frac{I e^{-I(v y_3 + w z_3)}}{(v y_3 + w z_3)(y_2 v + z_2 w - v y_3 - w z_3)^2} + \frac{I}{(y_2 v + z_2 w)^2 (v y_3 + w z_3)}$

- F s'annule si : $S_{14} = \{x_1 = -\frac{y_1 v + w z_1}{u}\}$,
sous cette condition : $f_{14} = e^{-I(x_2 Y u + x_3 Z u + y_2 Y v + y_3 Z v + w z_2 Y + w z_3 Z)}$
 $\Rightarrow F_{14} = \frac{-I e^{-I(u x_2 + y_2 v + z_2 w)}}{(u x_2 + y_2 v + z_2 w)^2 (u x_2 - u x_3 + y_2 v - v y_3 + z_2 w - w z_3)} +$
 $\frac{I e^{-I(u x_3 + v y_3 + w z_3)}}{(u x_2 - u x_3 + y_2 v - v y_3 + z_2 w - w z_3)(u x_3 + v y_3 + w z_3)^2}$
 $- \frac{\left(x_2 x_3 u^2 + x_2 u v y_3 + x_2 u w z_3 + I u x_2 + u x_3 y_2 v + u x_3 z_2 w + I u x_3 + w^2 z_3 z_2 \right.}{(u x_3 + v y_3 + w z_3)^2 (u x_2 + y_2 v + z_2 w)^2}$
 $\left. + I v y_2 + I w z_2 + v^2 y_3 y_2 + w z_3 y_2 v + v y_3 z_2 w + I z_3 w + I y_3 v \right)$

- F_{14} s'annule si : $S_{14.1} = \{x_3 = -\frac{v y_3 + w z_3}{u}\}$,
sous cette condition : $f_{14.1} = e^{-I Y (u x_2 + v y_2 + w z_2)}$
 $\Rightarrow F_{14.1} = \frac{-I e^{-I(y_2 v + u x_2 + z_2 w)}}{(y_2 v + u x_2 + z_2 w)^3}$
 $- \frac{I(2 I v y_2 + 2 I z_2 w + u^2 x_2^2 + 2 y_2 v u x_2 + 2 y_2 v z_2 w + 2 u x_2 z_2 w + y_2^2 v^2 + 2 I u x_2 + z_2^2 w^2 - 2)}{2(y_2 v + u x_2 + z_2 w)^3}$

- $F_{14.1}$ s'annule si : $S_{14.1.1} = \{x_2 = -\frac{y_2 v + z_2 w}{u}\} \Rightarrow F_{14.1.1} = \frac{1}{6}$

- F_{14} s'annule si : $S_{14.2} = \{x_2 = -\frac{y_2 v + z_2 w}{u}\}$,
sous cette condition : $f_{14.2} = e^{-I Z (u x_3 + y_3 v + w z_3)}$
 $\Rightarrow F_{14.2} = \frac{-I e^{-I(u x_3 + v y_3 + w z_3)}}{(u x_3 + v y_3 + w z_3)^3}$
 $- \frac{I(-2 + 2 I w z_3 + 2 I v y_3 + v^2 y_3^2 + w^2 z_3^2 + 2 u x_3 v y_3 + 2 u x_3 w z_3 + 2 v y_3 w z_3 + u^2 x_3^2 + 2 I u x_3)}{(u x_3 + v y_3 + w z_3)^3}$

- F_{14} s'annule si : $S_{14.3} = \{x_2 = \frac{v y_3 + u x_3 + w z_3 - y_2 v - z_2 w}{u}\}$,
sous cette condition : $f_{14.3} = e^{-I(u x_3 + y_3 v + w z_3)}(Y + Z)$
 $\Rightarrow F_{14.3} = \frac{-(u x_3 + w z_3 + v y_3 - 2 I) e^{-I(u x_3 + v y_3 + w z_3)}}{(u x_3 + v y_3 + w z_3)^3} - \frac{u x_3 + v y_3 + 2 I + w z_3}{(u x_3 + v y_3 + w z_3)^3}$

- F s'annule si : $S_{15} = \{x_2 = -\frac{vy_2+wz_2}{u}\}$,
 sous cette condition : $f_{15} = e^{-I(x_1Xu+x_3Zu+y_1Xv+y_3Zv+wz_1X+wz_3Z)}$
 $\Rightarrow F_{15} = \frac{-Ie^{-I(u x_3+vy_3+wz_3)}}{(vy_3+ux_3+wz_3-y_1v-z_1w-x_1u)(ux_3+vy_3+wz_3)^2} +$
 $\frac{Ie^{-I(x_1u+y_1v+z_1w)}}{(x_1u+y_1v+z_1w)^2(vy_3+ux_3+wz_3-y_1v-z_1w-x_1u)}$
 $\left(\begin{array}{l} u^2x_3x_1 + Iux_3 + ux_3z_1w + x_3uy_1v + Ix_1u + ux_1wz_3 + ux_1vy_3 + vy_3z_1w \\ +y_1vwz_3 + Iy_1v + y_1y_3v^2 + Iwz_3 + Iz_1w + w^2z_3z_1 + Ivy_3 \end{array} \right)$
 $-\frac{\hspace{10em}}{(x_1u+y_1v+z_1w)^2(ux_3+vy_3+wz_3)^2}$

- F_{15} s'annule si : $S_{15.1} = \{x_3 = -\frac{vy_3+wz_3}{u}\}$,
 sous cette condition : $f_{15.1} = e^{-IX(xu+y_1v+z_1w)}$
 $\Rightarrow F_{15.1} = \frac{-Ie^{-I(x_1u+y_1v+z_1w)}}{(x_1u+y_1v+z_1w)^3}$
 $-\frac{I(w^2z_1^2+2y_1vz_1w+2x_1uy_1v+2x_1uz_1w+x_1^2u^2+2Iz_1w+2Iux_1+v^2y_1^2-2+2Iy_1v)}{2(x_1u+y_1v+z_1w)^3}$

- F_{15} s'annule si : $S_{15.2} = \{x_1 = \frac{vy_3+ux_3+wz_3-y_1v-z_1w}{u}\}$,
 sous cette condition : $f_{15.2} = e^{-I(ux_3+y_3v+wz_3)(X+Z)}$
 $\Rightarrow F_{15.2} = \frac{-(ux_3+wz_3+vy_3-2I)e^{-I(ux_3+vy_3+wz_3)}}{(ux_3+vy_3+wz_3)^3} - \frac{ux_3+vy_3+2I+wz_3}{(ux_3+vy_3+wz_3)^3}$

- F s'annule si : $S_{16} = \{x_3 = -\frac{vy_3+wz_3}{u}\}$,
 sous cette condition : $f_{16} = e^{-I(x_1Xu+x_2Yu+y_1Xv+y_2Yv+wz_1X+wz_2Y)}$
 $\Rightarrow F_{16} = \frac{Ie^{-I(y_2v+ux_2+z_2w)}}{(y_2v+ux_2+z_2w)^2(x_1u-ux_2+y_1v-y_2v+z_1w-z_2w)}$
 $-\frac{Ie^{-I(x_1u+y_1v+z_1w)}}{(x_1u+y_1v+z_1w)^2(x_1u-ux_2+y_1v-y_2v+z_1w-z_2w)}$
 $-\frac{y_2v^2y_1+Iy_1v+z_2wy_1v+ux_2y_1v+y_2vz_1w+Ivy_2+z_2w^2z_1+Iz_1w+Ix_1u+u^2x_2x_1+z_2wx_1u+ux_2z_1w+Iwz_2+Iux_2}{(y_2v+ux_2+z_2w)^2(x_1u+y_1v+z_1w)^2}$

- F_{16} s'annule si : $S_{16.1} = \{x_1 = -\frac{-ux_2+y_1v-y_2v+z_1w-z_2w}{u}\}$,
 sous cette condition : $f_{16.1} = e^{-I(wz_2+ux_2+vy_2)(Y+X)}$
 $\Rightarrow F_{16.1} = \frac{-(y_2v+ux_2-2I+z_2w)e^{-I(y_2v+ux_2+z_2w)}}{(y_2v+ux_2+z_2w)^3} - \frac{y_2v+ux_2+z_2w+2I}{(y_2v+ux_2+z_2w)^3}$

- F s'annule si : $S_{17} = \{x_1 = -\frac{-wz_2+vy_1-ux_2+wz_1-vy_2}{u}\}$,
 sous cette condition : $f_{17} = e^{-I(Xux_2+Xvy_2+Xwz_2+x_2Yu+x_3Zu+y_2Yv+y_3Zv+wz_2Y+wz_3Z)}$
 $\Rightarrow F_{17} = \frac{-Ie^{-I(ux_3+vy_3+wz_3)}}{(ux_3+vy_3+wz_3)(vy_3+ux_3+wz_3-ux_2-y_2v-z_2w)^2}$
 $+\frac{I}{(y_2v+ux_2+z_2w)^2(ux_3+vy_3+wz_3)}$
 $\left(\begin{array}{l} (u^2x_2x_3 + 2Iz_2w + x_3uy_2v + z_2wux_3 - u^2x_2^2 + ux_2vy_3 - 2y_2vux_2 - 2ux_2z_2w \\ -Iwz_3 + ux_2wz_3 - y_2^2v^2 - z_2^2w^2 + z_2w^2z_3 + y_2vwz_3 - Ivy_3 + y_2y_3v^2 \\ -Iux_3 + 2Ivy_2 + vz_2wy_3 + 2Iux_2 - 2y_2vz_2w) \end{array} \right)$
 $+\frac{\hspace{10em}}{(y_2v+ux_2+z_2w)^2(vy_3+ux_3+wz_3-ux_2-y_2v-z_2w)^2}$

- F_{17} s'annule si : $S_{17.1} = \{x_2 = \frac{vy_3+ux_3+wz_3-y_2v-z_2w}{u}\}$,
 sous cette condition : $f_{17.1} = e^{-I(ux_3+y_3v+wz_3)(X+Y+Z)}$
 $\Rightarrow F_{17.1} = \frac{I}{(ux_3+vy_3+wz_3)^3}$
 $+\frac{I(v^2y_3^2-2Iwz_3-2Ivy_3+u^2x_3^2+2ux_3vy_3+w^2z_3^2-2-2Iux_3+2ux_3wz_3+2vy_3wz_3)e^{-I(ux_3+vy_3+wz_3)}}{2(ux_3+vy_3+wz_3)^3}$

- F s'annule si : $S_{18} = \left\{ x_1 = -\frac{-vy_3-wz_3-ux_3+vy_1+wz_1}{u} \right\}$,
sous cette condition : $f_{18} = e^{-I(Xy_3v+Xu_3+Xwz_3+x_2Yu+x_3Zu+y_2Yv+y_3Zv+wz_2Y+wz_3Z)}$
 $\Rightarrow F_{18} = \frac{Ie^{-I(y_2v+ux_2+z_2w)}}{(vy_3+ux_3+wz_3-ux_2-y_2v-z_2w)^2(y_2v+ux_2+z_2w)} +$
 $\frac{I}{(ux_3+vy_3+wz_3)^2(y_2v+ux_2+z_2w)}$
 $\left(\begin{array}{l} (-y_2y_3v^2 - x_3uy_2v + Iwz_2 - y_2vwz_3 + v^2y_3^2 - ux_2vy_3 - vz_2wy_3 - 2Iwz_3) \\ + 2vy_3wz_3 + 2ux_3vy_3 - 2Ivy_3 + w^2z_3^2 + 2ux_3wz_3 - 2Iux_3 - z_2wux_3 \\ - z_2w^2z_3 + u^2x_3^2 - u^2x_2x_3 - ux_2wz_3 + Ivy_2 + Iux_2 \end{array} \right) e^{-I(ux_3+vy_3+wz_3)}$
 $\frac{I}{(ux_3+vy_3+wz_3)^2(vy_3+ux_3+wz_3-ux_2-y_2v-z_2w)^2}$

- F_{18} s'annule si : $S_{18.1} = \left\{ x_2 = \frac{vy_3+ux_3+wz_3-y_2v-z_2w}{u} \right\}$,
sous cette condition : $f_{18.1} = e^{-I(ux_3+y_3v+wz_3)(X+Y+Z)}$
 $\Rightarrow F_{18.1} = \frac{I}{(ux_3+vy_3+wz_3)^3}$
 $+ \frac{I(v^2y_3^2-2Iwz_3-2Ivy_3+u^2x_3^2+2ux_3vy_3+w^2z_3^2-2-2Iux_3+2ux_3wz_3+2vy_3wz_3)e^{-I(ux_3+vy_3+wz_3)}}{(ux_3+vy_3+wz_3)^3}$

- F s'annule si : $S_{19} = \left\{ x_2 = -\frac{-vy_3-wz_3-ux_3+vy_2+wz_2}{u} \right\}$,
sous cette condition : $f_{19} = e^{-I(x_1Xu+Yy_3v+Yux_3+Ywz_3+x_3Zu+y_1Xv+y_3Zv+wz_1X+wz_3Z)}$
 $\Rightarrow F_{19} = \frac{I}{(ux_3+vy_3+wz_3)^2(x_1u+y_1v+z_1w)}$
 $\frac{Ie^{-I(x_1u+y_1v+z_1w)}}{(vy_3+ux_3+wz_3-y_1v-z_1w-x_1u)^2(x_1u+y_1v+z_1w)}$
 $\left(\begin{array}{l} (-y_1y_3v^2 - 2Iwz_3 - x_3uy_1v - y_1vwz_3 + v^2y_3^2 - vy_3z_1w + 2ux_3vy_3) \\ + 2vy_3wz_3 - ux_1vy_3 + Ix_1u - ux_1wz_3 + u^2x_3^2 - w^2z_3z_1 - 2Ivy_3 \\ + Iz_1w - 2Iux_3 + Iy_1v - u^2x_3x_1 + 2ux_3wz_3 - ux_3z_1w + w^2z_3^2 \end{array} \right) e^{-I(ux_3+vy_3+wz_3)}$
 $\frac{I}{(ux_3+vy_3+wz_3)^2(vy_3+ux_3+wz_3-y_1v-z_1w-x_1u)^2}$

Le code Maple de ces résolutions, ainsi que l'implémentation du calcul du Fourier d'un objet 3D dont le maillage est une variété topologique, sont disponibles, en langage C et en Matlab, sur la page Web <http://liris.univ-lyon1.fr/julien.ricard/publications.html>.

C

Liste des publications de l'auteur

Synthèse des publications réalisées :

- 2 publications dans des revues internationales avec comité de lecture.
- 6 articles dans des actes de conférences internationales avec comité de lecture.
- 4 articles dans des actes de conférences nationales avec comité de lecture.
- 4 rapports de contrat RNRT Semantic-3D.
- 2 rapports de recherche.

Publications internationales avec comité de lecture

1. RICARD, J., COEURJOLLY, D. et BASKURT, A. « Generalizations of Angular Radial Transform for 2D and 3D Shape Retrieval ». *Pattern Recognition Letters*, 26(14) :2174–2186. **2005**.
2. IDRISSE, K., LAVOUÉ, G., RICARD, J. et BASKURT, A. « Object of Interest based visual navigation, retrieval and semantic content identification system ». *Computer Vision On Image Understanding Special Issue On Colour For Image Indexing and Retrieval*, 94(1-3) :271–294. **2004**.

Conférences internationales avec comité de lecture

1. RICARD, J., COEURJOLLY, D. et BASKURT, A. « Generalization of Angular Radial Transform ». dans *IEEE International Conference on Image Processing, ICIP 2004*, volume 4, pages 2211-2214. Singapore. **2004**.
2. RICARD, J., COEURJOLLY, D. et BASKURT, A. « ART Extension for Description, Indexing and Retrieval of 3D Objects ». dans *17th International*

- Conference on Pattern Recognition, ICPR 2004*, volume 3, pages 79-82. Cambridge, United Kingdom. **2004**.
3. IDRISSE, K., LAVOUÉ, G., RICARD, J. et BASKURT, A. « Système de Recherche d'Images et de Navigation Visuelle Basé sur une Approche Locale ». dans *International Conference on Image and Signal Processing, ICISP 2003*, pages 119-128. Agadir, Maroc. **2003**.
 4. IDRISSE, K., RICARD, J. et BASKURT, A. « An Objective Performance Evaluation Tool For Color Based Image Retrieval Systems ». dans *IEEE International Conference on Image Processing, ICIP 2002*, volume 2, pages 389-392. Rochester, New York. **2002**.
 5. IDRISSE, K., RICARD, J., ANWANDER, A. et BASKURT, A. « An Image Retrieval System Based on Local and Global Color Descriptors ». dans *2nd IEEE Pacific-Rim Conference on Multimedia, PCM 2001*. pages 55-62, Beijing, China. **2001a**.
 6. IDRISSE, K., RICARD, J. et BASKURT, A. « Multi-component Cross Entropy segmentation for Color Image Retrieval ». dans *2nd International Symposium on Image and Signal Processing and Analysis, ISPA 2001*. pages 132-137, Pula, Croatia. **2001**.

Conférences nationales avec comité de lecture

1. RICARD, J., COEURJOLLY, D. et BASKURT, A. « Indexation et recherche dynamique d'objets 3D par vues par des requêtes 2D ». dans *COmpression et REprésentation des Signaux Audiovisuels, CORESA 2005*, accepté pour publication, Rennes, France. **2005**.
2. TRICHET, R., RICARD, J., TELLEZ, B. et BASKURT, A. « Navigation dans une base d'objets 3D ». dans *COmpression et REprésentation des Signaux Audiovisuels, CORESA 2005*, accepté pour publication, Rennes, France. **2005**.
3. RICARD, J., COEURJOLLY, D. et BASKURT, A. « Extension de la transformation ART pour la description, l'indexation et la recherche d'objets 3D ». dans *COmpression et REprésentation des Signaux Audiovisuels, CORESA 2004*. Lille, France. **2004**.
4. BASKURT, A., DUPONT, F., LAVOUÉ, G., RICARD, J., DAOUDI, M., VANDEBORRE, J., ANSARY, T. F., DUGELAY, J., GARCIA, E., PRÊTEUX, F., ZAHARIA, T., BLUM, F., SAVAGE, B., LICHAU, D., FRATANI, F. et DUTARTRE, A. « Semantic-3D : compression, indexation et tatouage de données 3D ». dans *COmpression et REprésentation des Signaux Audiovisuels, CORESA 2004*. Lille, France. **2004**.

Rapports de contrat

1. ZAHARIA, T., PRÊTEUX, F., RICARD, J., DUPONT, F., BASKURT, A., T.FILALI, DAOUDI, M. et VANDEBORRE, J.-P. *Comparaison des méthodes d'indexation*. Rapport technique, Contrat National RNRT. **2005**.
2. ZAHARIA, T., PRÊTEUX, F., RICARD, J., DUPONT, F., BASKURT, A., T.FILALI, DAOUDI, M. et VANDEBORRE, J.-P. *Critères pour la comparaison des méthodes d'indexation*. Rapport technique, Contrat National RNRT. **2004**.
3. BASKURT, A., DAOUDI, M., ANSARY, T. F., RICARD, J. et VANDEBORRE, J.-P. *Evaluation et comparaison des méthodes de recherche d'objets 3D à partir de données 2D*. Rapport technique, Contrat National RNRT. **2004**.
4. BASKURT, A., DAOUDI, M., ANSARY, T. F., MAHMOUDI, S., F.PRÊTEUX, RICARD, J., VANDEBORRE, J.-P. et ZAHARIA, T. *Etat de l'art sur l'indexation 3D*. Rapport technique, Contrat National RNRT. **2003**.

Rapports de recherche

1. RICARD, J., COEURJOLLY, D. et BASKURT, A. *ART Extension for Description, Indexing and Retrieval of 3D Objects*. Rapport technique, Rapport de Recherche RR-2004-001, LIRIS FRE 2672 CNRS, Claude Bernard University, Lyon. **2003a**.
2. RICARD, J., COEURJOLLY, D. et BASKURT, A. *Generalizations of Angular Radial Transform for 2D and 3D Shape Retrieval*. Rapport technique, Rapport de Recherche RR-2004-023, LIRIS FRE 2672 CNRS, Claude Bernard University, Lyon. **2003b**.

D

Index des auteurs cités

- Abbasi, S. 16, 25, 27
Abdel-Mottaleb, M. 83
Adan, A. 35
AKCAY, M. 72
Alaoutinen, S. 46
Ar, S. 47
Ashley, J. 66
Assfalg, J. 34
Attali, D. 39, 40

Ballard, D. H. 12, 13, 46
Baskurt, A. xvi, 33, 44, 60, 66, 67, 85, 144, 148, 150, 151
Batty, S. 83
Becker, J.-M. 46
Bennett, N. 46
Bergevin, R. 44
Berretti, S. 16

Bertolino, P. 40
Bertrand, G. 38
Beucher, S. 38
Bezedek, J. 150
Bhandarkar, S. M. 47
Bhattacharya, P. 46
Biederman, I. 16, 44
Blum, H. 38
Bober, M. 17
Boles, W. W. 16
Boley, D. 150
Boujemaa, N. 150
Bouman, C. A. 149
Bowyer, K. 28
Boyle, R. 16
Bruckstein, A. M. 38
Burridge, R. 46

- Canterakis, N. 47–49, 110, 144
Card, S. K. xvi, 151, 152, 157
Cerrada, C. 35
Chakrabarti, K. 78
Chandran, K. B. 47
Chang, T. C. 36
Chazelle, B. 34, 58
Chelberg, D. M. 26
Chen, C. xvi, 151, 152
Chen, D.-Y. xii, 25, 51, 52
Chen, J. 47, 48, 50, 110
Chen, J.-L. 26, 27
Chen, J.-Y. 149
Chen, T. 130, 146
Chevalier, L. 44
Chin, R. T. 16, 47
Chong, C.-W. 16
Cinque, L. 14
Ciocca, G. 14
Cocquerez, J.-P. 95
Coeurjolly, D. 39, 85, 144
Costa, M. S. 26
Cozar, J. R. 46
Craver, S. 149
Cunningham, S. J. xvi, 151, 152
Cyr, C. M. 28

Dalton, J. C. 149
Daoudi, M. 25, 28, 33, 53, 60, 120, 176
Davies, C. J. 46
Deans, S. R. 47
Del Bimbo, A. 16, 34
Dimai, A. 13
Dobkin, D. 34, 47, 48, 50, 58, 110
Dom, B. 66
Duda, R. O. 46
Dyer, C. 28

Elad, M. 47
Equitz, W. 13

Fang, J. 72
Feliu, V. 35
Filali Ansary, T. 28, 33, 53, 60, 120, 176
Flickner, M. 13, 66
Frigui, H. 150
Funkhouser, T. xii, 34, 47, 48, 50, 56–58, 110, 113

Gagvani, N. 38
Gao, X. 83
Gary, J. E. 16
Gauch, J. 83
Gibson, J. J. 159
Goldstine, H. H. 55
Golovan, A. 83
Gong, W. X. 38
Gonzalez, R. C. 16
Gorkani, M. 66
Gotteiner, N. L. 47
Guil, N. 46
Gupta, A. 151, 159
Gusakova, V. 83

Hafner, J. 13, 66
Hagihara, Y. xii, 40, 41
Halderman, A. 47, 48, 50, 110
Hart, P. E. 46
Hernandez, D. 45
Hext, G. 98
Hilaga, M. xii, 42
Himsworth, F. 98
Hirata, T. 39
Hlavac, V. 16
Holmes, G. xvi, 151, 152
Hong, K. 83
Hong, Y. H. 16, 17, 25
Honkela, T. 149
Hough, P. V. C. 46
Hsia, C. W. 83
Hsu, R.-L. 83
Hu, G. 46
Hu, M. 16
Huang, Q. 66
Huang, T. S. 78, 150, 151
Huang, Z. 83

Idrissi, K. xvi, 66, 67, 85, 150, 151
Iivarinen, J. 16
Ilg, M. 40
Illingworth, J. 46
Irani, P. xii, 16, 38, 45

Jacobs, D. 47, 48, 50, 110
Jaillet, F. 44
Jain, A. K. 83
Jain, R. 151, 159
Jayakumar, R. 46
Jeannin, S. 17, 18, 20
Joshi, S. 36

Kälviäinen, H. 46
Karvonen, S. 149
Kazhdan, M. xii, 47, 48, 50, 56, 57, 110, 113
Khotanzad, A. 16, 17, 25
Kim, J.-D. 17
Kim, W.-Y. 16, 17, 25
Kim, Y.-S. 16, 17, 25
Kimia, B. B. 28
Kirkpatrick, D. G. 40
Kiryati, N. 46
Kittler, J. 16, 46
Klein, R. 48, 49, 144
Kobatake, H. xii, 40, 41
Koenderink, J. 34
Kohmura, T. xii, 42
Kohonen, T. 149
Koskenniemi, T. 149
Koskinen, I. 149
Krishnapuram, R. 150
Kullback, S. 70, 78
Kultanen, P. 46
Kunii, T. L. xii, 42
Kurita, T. 16
Kyprianou, L. K. 37

Lavoué, G. xvi, 66, 67, 85, 150, 151

- Le Saux, B. 150
Lee, B. xii, 36
Lee, C. 67
Lee, D. 40, 66
Levialdi, S. 14
Levine, M. D. 44
Li, C. 67
Li, H. F. 46
Lien, S.-L.-C. 130, 146
Lipikorn, R. xii, 40, 41
Liu, H. 46
Liu, Z. 47
Love, E. L. 47
Lu, G. 16

Mackinlay, J. D. xvi, 151, 152, 157
Mackworth, A. K. 25
Mahmoudi, S. 25, 28, 33, 60
Malzbender, T. 90
Manjunath, B. S. 66
Marsaglia, G. 28
Matsumoto, F. 46
Matusiak, S. 25
Maurer Jr., C. 39
Mayer, R. J. 36
McLaughlin, R. A. 46
McPherson, D. D. 47
McWherter, D. 36
Mead, R. 98
Mehrotra, R. 16
Mehrotra, S. 78
Min, P. xii, 47, 48, 50, 56, 57, 110, 113
Moghaddam, B. 150, 151
Mokhtarian, F. 16, 25, 27
Montanari, U. 38
Montanvert, A. 40
Moore, A. W. 29
Morris, S. xvi, 151, 152
Mukundan, R. 16
Muller, K. 17
Murase, H. 28
Murray, F. J. 55

Nayar, S. 28
Neal, P. J. xii, 45, 46
Nelder, J. A. 98
Nene, S. A. 28
Niblack, W. 13, 66
Nixon, M. S. 46
Novotni, M. 48, 49, 144

Ogniewicz, R. 40
Oja, E. 46
Ortega, M. 78
Osada, R. 34, 58
Otsu, N. 16
Ouhyoung, M. xii, 25, 51, 52

Pala, P. 16, 34
Pao, D. C. W. 46
Paquet, E. 47, 48, 134
Peabody, M. 36
Pelleg, D. 29

Pellicanò, A. 14
Pentland, A. 66
Petkovic, D. 66
Philip, K. P. 47
Philipp, S. 95
Picard, R. 66
Plataniotis, K. N. 83
Podladchikova, L. 83
Poggio, T. 50, 89
Prêteux, F. 120, 176

Qi, R. 39
Qiu, G. 72

Raghavan, V. 39
Raveendran, P. 16
Reeb, G. 41
Regli, W. C. 36
Ricard, J. xvi, 33, 60, 66, 67, 85, 144, 148, 150, 151
Riesenhuber, M. 50, 89
Rioux, M. 47, 48, 134
Robertson, G. G. xvi, 151, 152, 157
Rogers, W. J. xvi, 151, 152
Rosenfeld, A. 46
Rui, Y. 78
Rusinkiewicz, S. xii, 47, 50, 110, 113

Saito, N. 46
Saitou, K. xii, 36
Salembier, P. 66
SANKUR, B. 72
Sanniti di Baja, G. 39
Santini, S. 151, 159
Saupe, D. 48, 110, 144
Sawhney, H. S. 13, 66
Schettini, R. 14
Schmitt, F. xii, 43
Sclaroff, S. 66
Sekita, I. 16
Senior, A. 83
Serra, J. 40
Shaked, D. 38
Shapiro, L. G. xii, 26, 45, 46
Shaposhnikov, D. 83
Shen, Y.-T. xii, 51, 52
Shevtsova, N. 83
Shilane, P. xii, 56, 57, 113
Shimizu, A. xii, 40, 41
Shinagawa, Y. xii, 42
Shokoufandeh, A. 36
Silver, D. 38
Sonka, M. 16
Soodamani, R. 47
Spendley, W. 98
Stanford, W. 47
Steele, D. 66
Stockman, G. 26, 27
Stricker, M. 13
Su, C.-J. 36
Sun, T.-L. 36
Swain, M. J. 12, 13

Tal, A. 47

- Teague, M. 16
- Teh, C.-H. 16, 47
- Tellez, B. 148
- Thiel, E. 39
- Thompson, S. 46
- Tian, Q. 150, 151
- Tian, X.-P. xii, 51, 52
- Tieng, Q. M. 16
- Trichet, R. 148
- Tsuji, S. 46
- Tung, T. xii, 43
- Van Doorn, A. 34
- Van Otterloo, P. J. 16
- Vandeborre, J.-P. 28, 33, 53, 60, 120, 176
- Venetsanopoulos, A. N. 83
- Visa, A. 16
- Von Neumann, J. 55
- Vranic, D. V. 48, 110, 144
- Waco, D. xii, 37
- Wang, M. 83
- Ware, C. xii, 16, 38, 45
- Wei, W. 83
- Woo, T. 37
- Woods, R. E. 16
- Wysk, R. A. 36
- Xu, L. 46
- Yanker, P. 66
- Yeo, B. 149
- Yeung, M. 149
- Yi, J. H. 26
- Zaharia, T. 33, 34, 47, 60, 120, 176
- Zapata, E. L. 46
- Zhang, D. 16, 130, 146

Bibliographie

- ABBASI, S. et MOKHTARIAN, F. « Affine-Similar Shape Retrieval : Application to Multi-View 3D Object Recognition ». *IEEE Transactions on Image Processing*, 10(1) :131–139. **2001**. pages 27
- ADAN, A., CERRADA, C. et FELIU, V. « Global shape invariants : a solution for 3D free-form object discrimination/identification problem ». *Pattern Recognition Journal*, 34(7) :1331–1348. **2001**. pages 35
- AKCAY, M., BASKURT, A. et SANKUR, B. « Object detection in images based on color features ». Dans *Signal and Image processing, Telecommunication and applications (SIU '02)*. **2002**. pages 72
- ASSFALG, J., DEL BIMBO, A. et PALA, P. « Curvature Maps for 3D CBR ». Dans *International Conference on Multimedia and Expo (ICME '03)*, volume 1, pages 301–304. Baltimore, MD, Etat-Unis. **2003a**. pages 34
- ASSFALG, J., DEL BIMBO, A. et PALA, P. « Retrieval of 3D Objects using Curvature Maps and Weighted Walkthroughs ». Dans *Proceedings of International Conference on Image Analysis and Processing (ICIAP '03)*, pages 348–353. Mantova, Italie. **2003b**. pages 34
- ATTALI, D., BERTOLINO, P. et MONTANVERT, A. « Using polyballs to approximate shapes and skeletons ». Dans *International Conference on Pattern Recognition (ICPR '94)*, pages 626–628. Jerusalem, Israël. **1994**. pages 40
- ATTALI, D., SANNITI DI BAJA, G. et THIEL, E. « Skeleton simplification through non significant branch removal ». *Image Processing and Communications*, 3(3-4) :63–72. **1997**. pages 39
- BALLARD, D. H. « Generalizing the Hough transform to detect arbitrary shapes ». *Pattern Recognition*, 13(2) :111–122. **1981**. pages 46
- BASKURT, A., DAOUDI, M., FILALI ANSARY, T., MAHMOUDI, S., PRÊTEUX, F., RICARD, J., VANDEBORRE, J.-P. et ZAHARIA, T. *Etat de l'art sur l'indexation 3D*. Rapport technique, Rapport de Contrat National RNRT-SEMANTIC 3D. **2003**. pages 33, 60

- BECKER, J.-M. *Méthodes géométriques pour l'imagerie*. Thèse de doctorat, Thèse de HDR, Université Jean Monnet, St.-Etienne, France. **1998**. pages 46
- BENNETT, N., BURRIDGE, R. et SAITO, N. « A method to detect and characterize ellipses using the Hough transform ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(7) :652–657. **1999**. pages 46
- BERGEVIN, R. et LEVINE, M. D. « Generic Object Recognition : Building and Matching Coarse Descriptions from Line Drawings ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(1) :19–36. **1993**. pages 44
- BERRETTI, S., DEL BIMBO, A. et PALA, P. « Retrieval by Shape Similarity with Perceptual Distance and Effective Indexing ». *IEEE Transactions on Multimedia*, 2(4) :225–239. **2000**. pages 16
- BEUCHER, S. *Segmentation d'Images et Morphologie mathématique*. Thèse de doctorat, École des Mines, Paris, France. **1990**. pages 38
- BEZEDEK, J. *Pattern Recognition with Fuzzy Objective Functions Algorithms*. Plenum, New York, NY, Etat-Unis. **1981**. pages 150
- BHANDARKAR, S. M. « A fuzzy probabilistic model for the generalized Hough transform ». *IEEE Transactions On Systems, Man and Cybernetics*, 24(5) :745–759. **1994**. pages 47
- BHATTACHARYA, P., LIU, H., ROSENFELD, A. et THOMPSON, S. « Hough-transform detection of lines in 3D space ». *Pattern Recognition Letters*, 21(9) :843–849. **2000**. pages 46
- BIEDERMAN, I. « Human image understanding : recent research and a theory ». Dans *Papers from the second workshop Vol. 13 on Human and Machine Vision II*, pages 13–57. Academic Press Professional, Inc., San Diego, CA, Etat-Unis. **1986**. pages 44
- BIEDERMAN, I. « Recognition by Components : A Theory of Human Image Understanding ». *Psychological Review*, 94 :115–147. **1987**. pages 16
- BLUM, H. « A transformation for extracting new descriptors of shape ». Dans W. Wathen-Dunn, éditeur, *Models for the Perception of Speech and Visual Forms*, pages 362–380. MIT Press, Amsterdam, Pays-Bas. **1967**. pages 38
- BOBER, M., KIM, J.-D., KIM, Y.-S., KIM, W.-Y. et MULLER, K. « Summary of the results in shape descriptor core experiment ». Dans *ISO/IEC JTC1/SC29/WG11/M4869, 48th Mpeg Meeting*. Vancouver, Canada. **1999**. pages 17
- BOLEY, D. « Principal Direction Divisive Partitioning ». *Data Mining and Knowledge Discovery*, 2(4) :325–344. ISSN 1384-5810. **1998**. pages 150

- BOWYER, K. et DYER, C. « Aspect graphs : an introduction and survey of recent results ». *International Journal of Imaging Systems and Technology*, 28(2) :315–328. **1990**. pages 28
- CANTERAKIS, N. « 3D Zernike Moments and Zernike Affine Invariants for 3D Image Analysis and Recognition ». Dans *11th International Conference on Image Analysis (ICIA '99)*. Kangerlussuaq, Groenland. **1999**. pages 47, 48, 49, 110, 144
- CHEN, C. et MORRIS, S. « Visualizing evolving networks : Minimum spanning trees versus Pathfinder networks ». Dans *IEEE Symposium on Information Visualization (InfoVis '03)*, pages 67–74. Seattle, Washington, Etat-Unis. **2003**. pages xvi, 151, 152
- CHEN, D.-Y. et OUHYOUNG, M. « A 3D Model Alignment and Retrieval System ». Dans *International Computer Symposium (ICS 2002), Workshop on Multimedia Technologies*, volume 2, pages 1436–1443. Hualien, Taiwan. **2002**. pages 25
- CHEN, D.-Y., TIAN, X.-P., SHEN, Y.-T. et OUHYOUNG, M. « On Visual Similarity Based 3D Model Retrieval ». *Computer Graphics Forum (EUROGRAPHICS'03)*, 22(3) :223–232. **2003**. pages xii, 51, 52
- CHEN, J.-L. et STOCKMAN, G. « 3D Free-Form Object Recognition Using Indexing by Contour Feature ». *Computer Vision and Image Understanding*, 71(3) :334–355. **1998**. pages 26, 27
- CHEN, J.-Y., BOUMAN, C. A. et DALTON, J. C. « Hierarchical Browsing and Search of Large Image Databases ». *IEEE TIP : IEEE Transactions on Image Processing*, 9(3) :442–455. **2000**. pages 149
- CHEVALIER, L., JAILLET, F. et BASKURT, A. « Segmentation and Superquadratic Modeling of 3D Objects ». *Journal of Winter School of Computer Graphics (WSCG '03)*, 11(1). ISSN 1213-6972. **2003**. pages 44
- CHONG, C.-W., RAVEENDRAN, P. et MUKUNDAN, R. « An Efficient Algorithm for Fast Computation of Pseudo-Zernike Moments ». *International Journal of Pattern Recognition and Artificial Intelligence*, 17(6) :1011–1023. **2003**. pages 16
- CINQUE, L., CIOCCA, G., LEVIALDI, S., PELLICANÒ, A. et SCHETTINI, R. « Color-based image retrieval using spatial-chromatic histograms ». *Image Vision Computer*, 19(13) :979–986. **2001**. pages 14
- COCQUEREZ, J.-P. et PHILIPP, S. *Analyse d'image : filtrage et segmentation*. édition Masson. **1995**. pages 95
- COEURJOLLY, D. « d-Dimensional Reverse Euclidean Distance Transformation and Euclidean Medial Axis Extraction in Optimal Time ». Dans G. S. I. Nystrom et

- S. Svensson, éditeurs, *Discrete Geometry for Computer Imagery, 11th International Conference, (DGCI '03)*, numéro 2886 dans LNCS, pages 327–337. Springer, Naples, Italie. **2003**. pages **39**
- COSTA, M. S. et SHAPIRO, L. G. « 3D Object Recognition and Pose with relational indexing ». *Computer Vision and Image Understanding*, 79 :364–407. **2000**. pages **26**
- CRAVER, S., YEO, B. et YEUNG, M. « Multi-Linearization Data Structure for Image Browsing ». Dans *SPIE Proceedings on Storage and Retrieval of Image and Video Databases (SPIE '99)*, pages 155–166. San Jose, CA, Etat-Unis. **1999**. pages **149**
- CYR, C. M. et KIMIA, B. B. « A Similarity-Based Aspect-Graph Approach to 3D Object Recognition ». *International Journal of Computer Vision*, 57(1) :5–22. **2004**. pages **28**
- DAOUDI, M. et MATUSIAK, S. « Visual Image Retrieval by Multiscale Description of User Sketches ». *Journal of Visual Languages and Computing*, 11(3) :287–301. **2000**. pages **25**
- DAVIES, C. J. et NIXON, M. S. « A Hough transform for detecting the location and orientation of three-dimensional surfaces via color encoded spots ». *IEEE Transactions on Systems, Man and Cybernetics, Part B*, 28(1) :90–95. **1998**. pages **46**
- DEANS, S. R. *The Radon transform and some of its applications*. Wiley Interscience, New York, NY, Etat-Unis. **1983**. pages **47**
- DEL BIMBO, A. et PALA, P. « Visual Image Retrieval by Elastic Matching of User Sketches ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(2) :121–132. **1997**. pages **16**
- DUDA, R. O. et HART, P. E. « Use of the Hough transformation to detect lines and curves in pictures ». *Communication ACM*, 15(1) :11–15. **1972**. pages **46**
- ELAD, M., TAL, A. et AR, S. « Content based retrieval of VRML objects : an iterative and interactive approach ». Dans *Proceedings of the sixth Eurographics workshop on Multimedia 2001*, pages 107–118. Springer-Verlag New York, Inc., New York, NY, Etat-Unis. **2002**. pages **47**
- FANG, J. et QIU, G. « Human face detection using angular radial transform and support vector machines. ». Dans *IEEE International Conference on Image Processing (ICIP '03)*, volume 1, pages 669–672. Barcelone, Espagne. **2003**. pages **72**

- FILALI ANSARY, T., DAOUDI, M. et VANDEBORRE, J.-P. « A Bayesian Approach for 3D Models Retrieval Based on Characteristic Views ». Dans *International Conference on Pattern Recognition (ICPR '04)*, volume 11(1), pages 898–901. Cambridge, Royaume-Unis. **2004**. pages **28**
- FILALI ANSARY, T., DAOUDI, M. et VANDEBORRE, J.-P. « 3D Model Retrieval Based on Adaptive Views Clustering ». Dans *Third international conference on Advances in Pattern Recognition (ICAPR '05)*. Springer Verlag, Bath, Royaume-Unis. **2005**. pages **28, 53, 120, 176**
- FLICKNER, M., SAWHNEY, H. S., NIBLACK, W., ASHLEY, J., HUANG, Q., DOM, B., GORKANI, M., HAFNER, J., LEE, D., PETKOVIC, D., STEELE, D. et YANKER, P. « Query by Image and Video Content : The QBIC System ». *IEEE Computer*, 28(9) :23–32. **1995**. pages **66**
- FRIGUI, H. et KRISHNAPURAM, R. « Clustering by competitive agglomeration ». *Pattern Recognition Journal*, 30(7) :1109–1119. **1997**. pages **150**
- FUNKHOUSER, T., MIN, P., KAZHDAN, M., CHEN, J., HALDERMAN, A., DOBKIN, D. et JACOBS, D. « A search engine for 3D models ». *ACM Transactions on Graphics*, 22(1) :83–105. **2003**. pages **47, 48, 50, 110**
- GAGVANI, N. et SILVER, D. « Parameter-controlled VOLUME thinning ». *CVGIP : Graphical Model and Image Processing*, 61(3) :149–164. **1999**. pages **38**
- GAO, X., SHEVTSOVA, N., HONG, K., BATTY, S., PODLADCHIKOVA, L., GOLOVAN, A., SHAPOSHNIKOV, D. et GUSAKOVA, V. « Vision Models Based Identification of Traffic Signs ». Dans *The First European Conference on Colour in Graphics, Image and Vision (CGIV '02)*, pages 47–51. Poitiers, France. **2002**. pages **83**
- GAUCH, J. et HSIA, C. W. « A comparison of three color image segmentation algorithms in four color spaces ». Dans *SPIE Proceedings on Visual Communications and Image Processing (SPIE '92)*, volume 1818, pages 1168–1181. San-Diego, CA, Etat-Unis. **1992**. pages **83**
- GIBSON, J. J. *The Ecological Approach to Visual Perception*. Houghton Mifflin, Boston, MA, Etat-Unis. **1979**. pages **159**
- GOLDSTINE, H. H., MURRAY, F. J. et VON NEUMANN, J. « The Jacobi method for real symmetric matrices ». *Journal of the ACM*, 6(1) :59–96. **1959**. pages **55**
- GONG, W. X. et BERTRAND, G. « A simple parallel 3D thinning algorithm ». Dans *10th International Conference on Pattern Recognition (ICPR '90)*, pages 188–190. Atlantic City, NJ, Etat-Unis. **1990**. pages **38**
- GONZALEZ, R. C. et WOODS, R. E. *Digital Image Processing*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, Etat-Unis. **2001**. pages **16**

- GUIL, N., COZAR, J. R. et ZAPATA, E. L. « Planar 3D object detection by using the generalized Hough transform ». Dans *10th International Conference on Image Analysis and Processing (ICIAP '99)*, pages 358–363. Venise, Italie. **1999**. pages [46](#)
- HAFNER, J., SAWHNEY, H. S., EQUITZ, W., FLICKNER, M. et NIBLACK, W. « Efficient Color Histogram Indexing for Quadratic Form Distance Functions ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(7) :729–736. **1995**. pages [13](#)
- HERNANDEZ, D. *Qualitative Representation of Spatial Knowledge*. Springer-Verlag New York, Inc., Secaucus, NJ, Etat-Unis. **1994**. pages [45](#)
- HILAGA, M., SHINAGAWA, Y., KOHMURA, T. et KUNII, T. L. « Topology matching for fully automatic similarity estimation of 3D shapes ». Dans *28th annual conference on Computer graphics and interactive techniques (SIGGRAPH '01)*, pages 203–212. ACM Press, New York, NY, Etat-Unis. **2001**. pages [xii](#), [42](#)
- HIRATA, T. « A unified linear-time algorithm for computing distance maps ». *Information Processing Letters*, 58(3) :129–133. ISSN 0020-0190. **1996**. pages [39](#)
- HONKELA, T., KOSKINEN, I., KOSKENNIEMI, T. et KARVONEN, S. « Kohonen's self-organizing maps in contextual analysis of data ». *Information organization and databases : foundations of data organization*, pages 135–148. **2000**. pages [149](#)
- HOUGH, P. V. C. « Method and means for recognizing complex patterns ». *U.S. Patent 3 069 654*. **1962**. pages [46](#)
- HU, G. « 3D object matching in the Hough space ». Dans *IEEE International Conference on Systems, Man and Cybernetics (SMC '95)*, volume 3, pages 2718–2723. Vancouver, Canada. **1995**. pages [46](#)
- HU, M. « Visual pattern recognition by moment invariants ». *IRE Transactions on Information Theory*, IT(8) :179–187. **1962**. pages [16](#)
- IDRISSI, K., LAVOUÉ, G., RICARD, J. et BASKURT, A. « Object of interest-based visual navigation, retrieval, and semantic content identification system ». *Computer Vision On Image Understanding Special Issue On Colour For Image Indexing and Retrieval*, 94(1-3) :271–294. ISSN 1077-3142. **2004**. pages [xvi](#), [66](#), [67](#), [85](#), [150](#), [151](#)
- IDRISSI, K., RICARD, J. et BASKURT, A. « An Objectuve Performance Evaluation Tool For Color Based Image Retrieval Systems ». Dans *IEEE 2002 International Conference on Image Processing (ICIP '02)*, volume 2, pages 389–392. Rochester, New York, Etat-Unis. **2002**. pages [66](#)

- IIVARINEN, J. et VISA, A. « Shape recognition of irregular objects ». Dans D. Casasent, éditeur, *Intelligent Robots and Computer Vision XV : Algorithms, Techniques, Active Vision, and Materials Handling (SPIE '96)*, pages 25–32. Boston, MA, Etat-Unis. **1996**. pages 16
- ILLINGWORTH, J. et KITTLER, J. « A survey of the Hough transform ». *Computer Vision, Graphics and Image Processing*, 44 :87–116. **1988**. pages 46
- IRANI, P. et WARE, C. « Diagramming information structures using 3D perceptual primitives ». *ACM Transactions on Computer-Human Interaction (TOCHI)*, 10 :1–19. **2003**. pages xii, 16, 38, 45
- JEANNIN, S. « MPEG-7 Visual part of experimentation model, Version 9.0 ». Dans *ISO/IEC JTC1/SC29/WG11/N3914, 55th Mpeg Meeting*. Pise, Italie. **2001**. pages 17, 18, 20
- JOSHI, S. et CHANG, T. C. « Graph-based heuristics for recognition of machined features from a 3D solid model ». *Computer-Aided Design*, 20 :58–66. **1988**. pages 36
- KAZHDAN, M., FUNKHOUSER, T. et RUSINKIEWICZ, S. « Rotation invariant spherical harmonic representation of 3D shape descriptors ». *Symposium on Geometry Processing*. **2003**. pages xii, 47, 50, 110, 113
- KHOTANZAD, A. et HONG, Y. H. « Invariant Image Recognition by Zernike Moments ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(5) :489–497. **1990**. pages 16, 17, 25
- KIM, W.-Y. et KIM, Y.-S. « A New region-Based Shape Descriptor ». Dans *Mpeg Meeting, TR 15-01*. Pise, Italie. **1999**. pages 16, 17, 25
- KIRKPATRICK, D. G. « Efficient computation of continuous skeletons ». Dans *IEEE 20th Annual Symposium on Foundations of Computer Science (FOCS '79)*, pages 18–27. **1979**. pages 40
- KIRYATI, N., KÄLVIÄINEN, H. et ALAOUTINEN, S. « Randomized or probabilistic Hough transform : unified performance evaluation ». *Pattern Recognition Letters*, 21(13–14) :1157–1164. **2000**. pages 46
- KOENDERINK, J. et VAN DOORN, A. « Surface shape and curvature scales ». *Image and Vision Computing*, 10(8) :557–565. **1992**. pages 34
- KOHONEN, T. *Self-organizing maps*. Springer-Verlag New York, Inc., Secaucus, NJ, Etat-Unis. **1997**. pages 149
- KULLBACK, S. *Information theory and statistics*. John Wiley and Sons, New York, NY, Etat-Unis. **1959**. pages 70, 78

- KYPRIANOU, L. K. *Shape Classification in Computer Aided Design*. Thèse de doctorat, Christ College. Cambridge, Royaume-Unis : Université de Cambridge. **1980**. pages **37**
- LE SAUX, B. et BOUJEMAA, N. « Image database clustering with SVM-based class personalization ». Dans *SPIE Proceedings on Conference on Storage and Retrieval Methods and Applications for Multimedia / Electronic Imaging symposium (SPIE '04)*, pages 9–19. San Jose, CA, Etat-Unis. **2004**. pages **150**
- LEE, B. et SAITOU, K. « Design of part families for reconfigurable machining systems based on manufacturability feedback ». Dans *International Conference On Engineering Design (ICED '01)*. Glasgow, Royaume-Unis. **2001**. pages **xii, 36**
- LEE, D. « Medial axis transformation of a planar shape ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 4(4) :363–369. **1982**. pages **40**
- LI, C. et LEE, C. « Minimum cross entropy thresholding ». *Pattern Recognition*, 26(4) :617–625. **1993**. pages **67**
- LIEN, S.-L.-C. *Combining Computation with Geometry*. Thèse de doctorat, Technical Report. California Institute of Technology, Pasadena, Californie. **1984**. pages **130, 146**
- LIPIKORN, R., SHIMIZU, A., HAGIHARA, Y. et KOBATAKE, H. « 3D Modified Exoskeleton and its Application for Shape Recognition ». Dans *International Conference on Acoustics Speech and Signal Processing (ICASSP '02)*, volume 4, pages 3568–3571. Orlando, FL, Etat-Unis. **2002a**. pages **xii, 41**
- LIPIKORN, R., SHIMIZU, A., HAGIHARA, Y. et KOBATAKE, H. « A modified exoskeleton and its application to object representation and recognition, ». *IEICE Transactions on Information and Systems*, E85-D(5) :884–896. **2002b**. pages **40**
- MAHMOUDI, S. et DAOUDI, M. « Une nouvelle méthode d'indexation 3D ». Dans *13ème Congrès de Reconnaissance des Formes et Intelligence Artificielle (RFIA '02)*, volume 1, pages 19–27. Angers, France. **2002**. pages **25, 28**
- MALZBENDER, T. « Fourier Volume rendering ». *ACM Transactions on Graphics*, 12(3) :233–250. **1993**. pages **90**
- MARSAGLIA, G. « Choosing a point from the surface of a sphere ». *Annals of mathematical statistics*, 43 :645–646. ISSN 0003-4851. **1972**. pages **28**
- MAURER JR., C., RAGHAVAN, V. et QI, R. « A Linear Time Algorithm for Computing the Euclidean Distance Transform in Arbitrary Dimensions ». Dans *Information Processing in Medical Imaging (IPMI '01)*, pages 358–364. Davis, CA, Etat-Unis. **2001**. pages **39**

- MCCLAUGHLIN, R. A. « Randomized Hough Transform : Improved ellipse detection with comparison ». *Pattern Recognition Letters*, 19(3-4) :299–305. **1998**. pages [46](#)
- MCWHERTER, D., PEABODY, M., REGLI, W. C. et SHOKOUFANDEH, A. « Transformation Invariant Shape Similarity Comparison of Solid Models. ». Dans *ASME Design Engineering Technical Confs., 6th Design for Manufacturing Conf. (DETC 2001/DFM-21191)*. Pittsburgh, PA, Etat-Unis. **2001**. pages [36](#)
- MEHROTRA, R. et GARY, J. E. « Similar-Shape Retrieval In Shape Data Management ». *Computer*, 28(9) :57–62. **1995**. pages [16](#)
- MOGHADDAM, B., TIAN, Q. et HUANG, T. S. « Spatial Visualization for Content-Based Image Retrieval ». Dans *IEEE International Conference on Multimedia and Expo (ICME '01)*, pages 42–47. Tokyo, Japon. **2001**. pages [150](#), [151](#)
- MOKHTARIAN, F. et ABBASI, S. « Automatic View Selection in Multi-view Object Recognition ». Dans *International Conference on Pattern Recognition (ICPR '00)*, volume 1, pages 13–16. Barcelone, Espagne. **2000**. pages [25](#)
- MOKHTARIAN, F., ABBASI, S. et KITTLER, J. « Robust and Efficient Shape Indexing through Curvature Scale Space ». Dans *Proceedings of the sixth British Machine Vision Conference (BMVC '96)*, pages 53–62. Edinburgh, Royaume-Unis. **1996**. pages [16](#)
- MOKHTARIAN, F. et MACKWORTH, A. K. « A Theory of Multiscale, Curvature-Based Shape Representation for Planar Curves ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(8) :789–805. **1992**. pages [25](#)
- MONTANARI, U. « A Method for Obtaining Skeletons Using a Quasi-Euclidean Distance ». *Journal of the ACM (JACM)*, 15(4) :600–624. ISSN 0004-5411. **1968**. pages [38](#)
- NAYAR, S., NENE, S. A. et MURASE, H. « Real-Time Object Recognition System ». Dans *IEEE International Conference on Robotics and Automation (ICRA '96)*. Minneapolis, MN, Etat-Unis. **1996**. pages [28](#)
- NEAL, P. J. et SHAPIRO, L. G. « A Symbolic Representation for 3D Object Feature Detection ». Dans *International Conference on Pattern Recognition (ICPR '00)*, pages 221–224. Barcelone, Espagne. **2000**. pages [xii](#), [45](#), [46](#)
- NELDER, J. A. et MEAD, R. « A simplex method for function minimization ». *Computer Journal*, 7 :308–313. **1965**. pages [98](#)
- NOVOTNI, M. et KLEIN, R. « 3D Zernike descriptors for content based shape retrieval ». Dans *the eighth ACM symposium on Solid modeling and applications (SM '03)*, pages 216–225. ACM Press. **2003**. pages [49](#), [144](#)

- NOVOTNI, M. et KLEIN, R. « Shape retrieval using 3D Zernike descriptors ». *Computer Aided Design*, 36(11) :1047–1062. **2004**. pages 48, 49, 144
- OGNIEWICZ, R. et ILG, M. « Voronoi skeletons : Theory and applications ». *Computer Vision and Pattern Recognition*, pages 63–69. **1992**. pages 40
- ORTEGA, M., RUI, Y., CHAKRABARTI, K., MEHROTRA, S. et HUANG, T. S. « Supporting similarity queries in MARS ». Dans *the fifth ACM international conference on Multimedia (MULTIMEDIA '97)*, pages 403–413. ACM Press, New York, NY, Etat-Unis. **1997**. pages 78
- OSADA, R., FUNKHOUSER, T., CHAZELLE, B. et DOBKIN, D. « Matching 3D Models with Shape Distributions ». Dans *Proceedings of the International Conference on Shape Modeling & Applications (SMI '01)*, pages 154–168. IEEE Computer Society, Washington, DC, Etat-Unis. **2001**. pages 34, 58
- PAO, D. C. W., LI, H. F. et JAYAKUMAR, R. « Shapes Recognition Using the Straight Line Hough Transform : Theory and Generalization ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(11) :1076–1089. ISSN 0162-8828. **1992**. pages 46
- PAQUET, E. et RIOUX, M. « The MPEG-7 Standard and the Content-Based Management of Three-Dimensional Data : A Case Study ». Dans *IEEE International Conference on Multimedia Computing Systems (ICMCS '99)*, volume 1, pages 375–380. Florence, Italie. **1999**. pages 47, 48, 134
- PELLEG, D. et MOORE, A. W. « X-means : Extending K-means with Efficient Estimation of the Number of Clusters ». Dans *the Seventeenth International Conference on Machine Learning (ICML '00)*, pages 727–734. Morgan Kaufmann Publishers Inc., San Francisco, CA, Etat-Unis. **2000**. pages 29
- PENTLAND, A., PICARD, R. et SCLAROFF, S. « PhotoBOOK : Tools for Content based manipulation of image databases ». *International Journal of Computer Vision*, 18(3) :233–254. **1996**. pages 66
- PHILIP, K. P., LOVE, E. L., MCPHERSON, D. D., GOTTEINER, N. L., STANFORD, W. et CHANDRAN, K. B. « The fuzzy Hough transform-feature extraction in medical images ». *IEEE Transactions on Medical Image Processing*, 13(2). **1994**. pages 47
- PLATANIOTIS, K. N. et VENETSANOPOULOS, A. N. *Color image processing and applications*. Springer-Verlag New York, Inc., New York, NY, Etat-Unis. **2000**. pages 83
- REEB, G. « Sur les points singuliers d'une forme de Pfaff completement integrable ou d'une fonction numerique [On the Singular Points of a Completely Integrable Pfaff Form or of a Numerical Function] ». *Comptes Rendus de Académie des Sciences de Paris*, 222 :847–849. **1946**. pages 41

- RICARD, J., COEURJOLLY, D. et BASKURT, A. « Generalization of angular radial transform ». Dans *IEEE International Conference on Image Processing (ICIP '04)*, volume IV, pages 2211–2214. Singapour. **2004**. pages 85
- RICARD, J., COEURJOLLY, D. et BASKURT, A. « Generalizations of Angular Radial Transform for 2D and 3D Shape Retrieval ». *Pattern Recognition Letters*, 26(14) :2174–2186. **2005**. pages 85, 144
- RIESENHUBER, M. et POGGIO, T. « Models of object recognition ». *Nature neuroscience*, 3 NUMBER Supp. **2000**. pages 50, 89
- ROBERTSON, G. G., MACKINLAY, J. D. et CARD, S. K. « Cone Trees : animated 3D visualizations of hierarchical information ». Dans *Conference on Human factors in computing systems (SIGCHI '91)*, pages 189–194. ACM Press, New York, NY, Etat-Unis. **1991**. pages xvi, 151, 152, 157
- ROGERS, W. J., CUNNINGHAM, S. J. et HOLMES, G. « Navigating the virtual library : a 3D browsing interface for information retrieval ». *Australian Journal of Intelligent Information Processing Systems*, 3(4) :43–49. **1996**. pages xvi, 151, 152
- SALEMBIER, P. et MANJUNATH, B. S. « AudioVisual Content Description and Retrieval : Tools and MPEG-7 Standardization activities ». Dans *IEEE International Conference on Image Processing (ICIP '00)*, pages 29–33. Vancouver, Canada. **2000**. pages 66
- SANTINI, S., GUPTA, A. et JAIN, R. « Emergent Semantics through Interaction in Image Databases ». *IEEE Transactions on Knowledge and Data Engineering*, 13(3) :337–351. ISSN 1041-4347. **2001**. pages 151, 159
- SAUPE, D. et VRANIC, D. V. « 3D Model Retrieval with Spherical Harmonics and Moments ». Dans *the 23rd Symposium on Pattern Recognition (DAGM '01)*, pages 392–397. Springer-Verlag, Londres, Royaume-Unis. **2001**. pages 48, 110
- SEKITA, I., KURITA, T. et OTSU, N. « Complex Autoregressive Model for Shape Recognition ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(4) :489–496. **1992**. pages 16
- SENIOR, A., HSU, R.-L., ABDEL-MOTTALEB, M. et JAIN, A. K. « Face Detection in Color Images ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5) :696–706. **2002**. pages 83
- SERRA, J. *Image Analysis and Mathematical Morphology*. Academic Press, New York, NY, Etat-Unis. **1982**. pages 40
- SHAKED, D. et BRUCKSTEIN, A. M. « Pruning medial axes ». *Computer Vision Image Understanding*, 69(2) :156–169. ISSN 1077-3142. **1998**. pages 38

- SHILANE, P., MIN, P., KAZHDAN, M. et FUNKHOUSER, T. « The Princeton Shape Benchmark ». Dans *Shape Modeling International (SMI '05)*, pages 167–178. Genève, Italie. **2004**. pages [xii](#), [56](#), [57](#), [113](#)
- SONKA, M., HLAVAC, V. et BOYLE, R. *Image Processing : Analysis and Machine Vision*. Thomson-Engineering ; 2 édition. **1999**. pages [16](#)
- SOODAMANI, R. et LIU, Z. « A novel fuzzy Hough transform for shape representation ». Dans *IEEE International Conference on Fuzzy Systems (FUZZ '98)*, volume 2, pages 1605–1608. Anchorage, Alaska, Etat-Unis. **1998**. pages [47](#)
- SPENDLEY, W., HEXT, G. et HIMSWORTH, F. « Sequential Application of Simplex Designs in Optimization and Evolutionary Operation ». *Technometrics*, 4 :441–61. **1962**. pages [98](#)
- STRICKER, M. et DIMAI, A. « Color Indexing with Weak Spatial Constraints ». Dans *Storage and Retrieval for Image and Video Databases (SPIE '96)*, volume IV 2670, pages 29–40. San Diego/La Jolla, CA, Etat-Unis. **1996**. pages [13](#)
- SUN, T.-L., SU, C.-J., MAYER, R. J. et WYSK, R. A. « Shape similarity assessment of mechanical parts based on solid models ». Dans *ASME Design for Manufacturing Conference, Symposium on Computer Integrated Concurrent Design*, volume 83(2), pages 953–962. Boston, MA, Etat-Unis. **1995**. pages [36](#)
- SWAIN, M. J. et BALLARD, D. H. « Color indexing ». *International Journal of Computer Vision*, 7(1) :11–32. **1991**. pages [12](#), [13](#)
- TEAGUE, M. « Image analysis via the general theory of moments ». *Journal of the Optical Society of America*, 70(8) :920–930. **1980**. pages [16](#)
- TEH, C.-H. et CHIN, R. T. « On image analysis by the methods of moments ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(4) :496–513. **1988**. pages [17](#), [47](#)
- TIENG, Q. M. et BOLES, W. W. « Recognition of 2D Object Contours Using the Wavelet Transform Zero-Crossing Representation ». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(8) :910–916. **1997**. pages [16](#)
- TRICHET, R., RICARD, J., TELLEZ, B. et BASKURT, A. « Navigation dans une base d'objets 3D ». Dans *COmpression et REprésentation des Signaux Audiovisuels (CORESA '05)*, accepté pour publication. Rennes, France. **2005**. pages [148](#)
- TSUJI, S. et MATSUMOTO, F. « Detection of ellipses by a modified Hough transform ». *IEEE Transactions on Computers*, C-27(8) :777–781. **1978**. pages [46](#)
- TUNG, T. et SCHMITT, F. « Augmented Reeb Graphs for Content-Based Retrieval of 3D Mesh Models ». Dans *International Conference on Shape Modeling and Applications (SMI '04)*, pages 157–166. IEEE Computer Society Press, Genève, Italie. **2004**. pages [43](#)

- TUNG, T. et SCHMITT, F. « The augmented multiresolution Reeb graph approach for content-based retrieval of 3D shapes ». *International Journal of Shape Modeling (IJSM), SMI04 Special issue, accepté pour publication*, 11(1) :91–120. **2005**. pages [xii](#), [43](#)
- VAN OTTERLOO, P. J. *A contour-oriented approach to shape analysis*. Prentice Hall International (UK) Ltd., Hertfordshire, Royaume-Unis. **1991**. pages [16](#)
- VRANIC, D. V. et SAUPE, D. « 3D Shape Descriptor Based on 3D Fourier Transform ». Dans *EURASIP Conference on Digital Signal Processing for Multimedia Communications and Services (ECMCS '01)*, pages 271–274. Budapest, Hongrie. **2001**. pages [48](#), [110](#), [144](#)
- VRANIC, D. V. et SAUPE, D. « Description of 3D-shape using a complex function on the sphere ». Dans *IEEE International Conference on Multimedia and Expo (ICME '02)*, pages 177–180. Lausanne, Suisse. **2002**. pages [110](#)
- WACO, D. et KIM, Y. S. « Geometric Reasoning for Machining Features using Convex Decomposition ». *Computer Aided Design*, 26 :477–489. **1994**. pages [xii](#), [37](#)
- WEI, W., WANG, M. et HUANG, Z. « An automatic method of localisation for NUMBER-plate using color features ». Dans *IEEE International Conference on Image Processing (ICIP '01)*, volume 1, pages 782–785. Thessaloniki, Grèce. **2001**. pages [83](#)
- WOO, T. « Feature Extraction by VOLUME Decomposition ». Dans *CAD/CAM Technology in Mechanical Engineering*, pages 76–94. Cambridge, MA, Etat-Unis. **1982**. pages [37](#)
- XU, L., OJA, E. et KULTANEN, P. « A new curve detection method : randomized Hough transform (RHT) ». *Pattern recognition Letters*, 11(5) :331–338. **1990**. pages [46](#)
- YI, J. H. et CHELBERG, D. M. « Model-Based 3D Object Recognition Using Bayesian Indexing ». *Computer Vision and Image Understanding*, 69(1) :87–105. **1998**. pages [26](#)
- ZAHARIA, T. et PRÊTEUX, F. « Indexation de maillages 3D par descripteurs de forme ». Dans *13ème Congrès Francophone de Reconnaissance de Forme et Intelligence Artificielle (RFIA '02)*, volume 1, pages 48–57. Angers, France. **2002a**. pages [34](#), [47](#)
- ZAHARIA, T. et PRÊTEUX, F. « Shape-based retrieval of 3D mesh models ». Dans *IEEE International Conference on Multimedia and Expo (ICME '02)*, volume 1, pages 437–440. Lausanne, Suisse. **2002b**. pages [47](#)

- ZAHARIA, T. et PRÊTEUX, F. « Descripteurs de forme : Etude comparée des approches 3D et 2D/3D ». Dans *14ème Congrès Francophone AFRIF-AFIA Reconnaissance des Formes et Intelligence Artificielle (RFIA '04)*. Toulouse, France. **2004**. pages 120, 176
- ZHANG, D. et CHEN, T. « Efficient Feature Extraction for 2D/3D Objects in Mesh Representation ». Dans *IEEE International Conference on Image Processing (ICIP '01)*, volume 1, pages 935–938. Thessaloniki, Grèce. **2001**. pages 130, 146
- ZHANG, D. et LU, G. « A Comparison of Shape Retrieval Using Fourier Descriptors and Short-Time Fourier Descriptors ». Dans *IEEE Pacific Rim Conference on Multimedia (PCM '01)*, pages 855–860. Pékin, chine. **2001**. pages 16
- ZHANG, D. et LU, G. « Generic Fourier descriptor for shape-based image retrieval ». Dans *IEEE International Conference on Multimedia and Expo (ICME '02)*, volume 1, pages 425–428. Lausanne, Suisse. **2002**. pages 16

Indexation et recherche d'objets 3D à partir de requêtes 2D et 3D

Depuis l'émergence de la Conception Assisté par Ordinateurs (CAO), l'archivage, la structuration et la recherche dans des bases de modèles 3D constituent des domaines de recherche privilégiés. Cette thèse s'inscrit dans ce cadre, et plus généralement, dans la recherche d'information par le contenu. Dans ce contexte, nous nous intéressons plus particulièrement à l'indexation d'objets tridimensionnels, qui a pour objectif de caractériser la forme des objets, pour permettre de calculer des distances de similarité entre modèles. Pour cela, deux types d'approches ont été étudiées : les approches 3D/3D, comparant les objets entre eux et les approches 2D/3D, comparant les objets 3D à une image.

La première partie de la thèse est consacrée à l'étude des descripteurs de forme 2D et 2D/3D. D'une part, nous avons généralisé le descripteur de forme 2D : Angular Radial Transform (ART), pour le rendre applicable sur les images couleurs et robuste aux déformation perspective. Dans un deuxième temps, nous avons étudié les approches 2D/3D, qui permettent de rechercher un objet 3D à partir d'une image 2D. La méthode d'indexation 2D/3D que nous proposons est une méthode dynamique qui compare directement l'image requête à l'objet 3D et s'adapte aux complexités de l'objet 3D et de la requête. Notre système se base sur un système de représentation spectral des vues de l'objet, permettant d'extraire rapidement les vues et de mettre en place une recherche dynamique de la vue requête.

La deuxième partie de la thèse porte sur l'indexation d'objets 3D. Pour indexer des grandes bases, nous avons réalisé une extension de la transformation ART en 3D, appelé : 3D Angular Radial Transform (3D ART). Ce descripteur possède un fort pouvoir de description, en gardant les propriétés de robustesse d'ART 2D. Pour permettre d'appréhender des grandes bases d'objets 3D, nous proposons également un système de navigation permettant de structurer les bases.

Mots clés : indexation, descripteur de forme, descripteur couleur, image, objet 3D.

3D-objects indexing and retrieval, from 2D and 3D queries

The increasing use of Computer-Aided (CAD) tools last decades has lead to huge 3D-model databases that still have to be correctly indexed. The context of this work is content-based retrieval of images or models. We are interested in both 2D images and 3D-models indexing. We aim at extracting shape descriptors from images or models that allow to compute a similarity measure between them. We present two approaches. On the one hand retrieving a 3D-model from a 2D image-based query and on the other hand, retrieving a 3D-model from another one.

The first part is dedicated to the 2D/3D approach. First, we generalize the 2D shape descriptor ART (Angular Radial Transform) by making it robust to rotation- and viewpoint-based deformations. In a second time, we propose a view-based approach that describes a 3D-model with a set of 2D views. More precisely, we compute for each model a set of views that are indexed by a 2D shape descriptor. This 2D/3D method is dynamic and adaptive to the complexity of both the 3D-model and the request. For a dynamic comparison we need a fast process for view extraction as it will be repeated a number of times. That is why we introduce a new spectral representation that characterizes all views from a given 3D-model.

In the second part of this work, we proposed a new fast 3D shape descriptor to index huge databases. The great number of 3D-models from the database imposes the use of a very compact shape descriptor. This one should contain maximum information in a minimum size to guarantee fast answer. We propose a 3D extension from ART called 3D Angular Radial Transform (3D ART). The latter has a strong descriptions capacity while keeping the 2D-ART robustness to deformations.

Keywords : indexation, shape descriptor, color descriptor, image, 3D object.