HAL
open science

# Recherche par le contenu d'objets 3D

Mohamed Chaouch

## To cite this version:

Mohamed Chaouch. Recherche par le contenu d'objets 3D. Traitement du signal et de l'image [eess.SP]. Télécom ParisTech, 2009. Français. NNT : . pastel-00005168

**HAL Id: pastel-00005168**

**https://pastel.hal.science/pastel-00005168**

Submitted on 16 Nov 2010

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

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

# Thèse

présentée pour obtenir le grade de docteur
de l'École Nationale Supérieure des Télécommunications

Spécialité : **Signal et Images**


# Mohamed Chaouch


# Recherche par le contenu d'objets 3D

soutenue le 31 Mars 2009 devant le jury composé de

| | |
|---|---|
| Françoise Prêteux | Président |
| Michela Spagnuolo | Rapporteur |
| Atilla Baskurt | Rapporteur |
| Henri Maître | Examinateur |
| Nozha Boujemaa | Examinateur |
| Anne Verroust-Blondet | Directeur de thèse |

*À mes chers parents,*
*À ma chère femme,*
*À mes chères sœurs*

# Remerciements

Mes remerciements vont d'abord à Anne Verroust-Blondet pour m'avoir encadré tout au long de ma thèse. J'ai particulièrement été sensible à la cohérence et la pertinence dans le choix des directions scientifiques, tout en laissant une importante place à l'initiative personnelle. Elle a su aussi me faire partager ses reflexions sur les tendances de l'état de l'art et particulièrement sur les questions d'usage en recherche par le contenu d'objets 3D. Je lui suis sincèrement reconnaissant d'avoir fait que mes années de thèse soient passionnantes.

Je remercie vivement Nozha Boujemaa pour m'avoir accueilli avec bienveillance au sein de son équipe de recherche. J'ai eu ainsi le privilège d'apprécier ses qualités humaines et de bénéficier de ses précieux conseils.

C'est avec non moins de plaisir que je voudrais exprimer ma reconnaissance aux membres du jury pour s'être penché avec attention sur mon travail de thèse.

Je tiens à rendre hommage et dédier ce travail à mes précieux parents qui m'ont toujours soutenu, conseillé et encouragé dans mes choix. Leur présence et leur amour sont pour moi les piliers fondateurs de ce que je suis et de ce que je fais.

Aucun mot ne peut exprimer ma reconnaissance à ma chère épouse, Nouha, pour son soutien quotidien indéfectible et sa patience tout au long de mes recherches. Que cette thèse témoigne de mon amour et de ma tendresse pour elle.

Je tiens à saluer chaleureusement mes sœurs, Salha et Cherifa, pour leur inestimable soutien. Je profite de cette occasion pour faire un clin d'œil ému à mes deux petites nièces, Sirine et Nisriya.

Enfin, mes remerciements s'adressent à la famille, aux amis, aux membres du projet IMEDIA et à ceux et celles qui, de près ou de loin, ont contribué à l'aboutissement de ce travail. Qu'ils trouvent dans ce manuscrit l'expression de ma respectueuse reconnaissance et ma profonde estime.

# Résumé

Aujourd'hui, grâce aux technologies récentes de numérisation et de modélisation 3D, des bases d'objets 3D de taille de plus en plus grande deviennent disponibles. La recherche par le contenu apparaît être une solution nécessaire pour structurer, gérer ces données multimédia, et pour naviguer dans ces grandes bases. Dans ce contexte, nous nous sommes intéressés plus particulièrement aux trois phases qui constituent le processus de la recherche par le contenu : la normalisation, l'extraction de la signature, et la mesure de similarité entre objets 3D.

La première partie de la thèse porte sur la normalisation d'objets 3D, en particulier sur la recherche de la pose optimale. Nous proposons une nouvelle méthode d'alignement d'objets 3D basée sur la symétrie de réflexion et la symétrie de translation locale. L'approche repose sur les propriétés intéressantes de l'analyse en composantes principales vis-à-vis des symétries de réflexion pour retrouver d'éventuels axes d'alignement optimaux parmi les axes principaux.

La deuxième partie de la thèse est consacrée à l'étude des descripteurs de forme et des mesures de similarité associées. Dans un premier temps, nous proposons un nouveau descripteur 3D dérivé de la transformée de Gauss, appelé *3D Gaussian descriptor*. Calculé sur une partition de l'espace englobant l'objet, ce descripteur permet de caractériser localement la géométrie de la surface. Dans un deuxième temps, nous avons étudié les approches multi-vues caractérisant l'objet 3D à partir des images de projection. Nous introduisons une technique d'amélioration, appelée *Enhanced Multi-views Approach*, pouvant s'appliquer à la plupart des approches multi-vues. Des valeurs de pertinence sont définies et introduites dans la mesure de similarité afin de pondérer les contributions des projections dans la description de la forme 3D. La dernière méthode d'indexation que nous proposons, appelée *Depth Line Approach*, est une approche multi-vues fondée sur les images de profondeur. Pour les caractériser, nous introduisons une méthode de transcodage qui transforme les lignes de profondeur extraites en séquence d'états. Pour permettre une comparaison efficace, la mesure de similarité associée se base sur la programmation dynamique.

# Abstract

Recent technologies of digitizing and modelling have lead to huge available 3D model databases on the web, as well as in domain-specific applications. 3D shape-content based retrieval seems to be a necessary and promising solution to navigate in large databases. It also appears to be useful way to structure and manage this new multimedia data. This thesis deals with 3D shape similarity search. More precisely, we focus on the main steps of the 3D shape matching process : normalization of 3D models, signature extraction from models, and similarity measure between them.

The first part of the thesis concerns the normalization of 3D models, in particular the search for the optimal pose. We propose a new alignment method of 3D models based on the reflective symmetry and the local translational symmetry. We use the properties of the principal component analysis with respect to the planar reflective symmetry in order to select the eventual optimal alignment axes within the principal axes.

The second part of the thesis is dedicated to the shape descriptors and the associated similarity measures. Firstly, we propose a new 3D descriptor, called *3D Gaussian descriptor*, which is derived from the Gauss transform. Based on a partition of the enclosing 3D model space, this descriptor provides a local characterization of the boundary of the shape. Secondly, we study the multi-views based approaches that characterize the 3D model using their projection images. We introduce an augmented approach, named *Enhanced Multi-views Approach*, which can be applied in most of the multi-views descriptors. The relevance indices are defined and used in the similarity computation in order to normalize the contributions of the projections in the 3D-shape description. Finally, we propose a robust 3D shape indexing approach, called *Depth Line Approach*, which is based on the appearance of a set of depth-buffer images. To extract a compact 3D model signature, we introduce a sequencing method that transforms the depth lines into sequences. Retrieval is improved by using dynamic programming to compare sequences.

# Table des matières

**Introduction** **5**

## I Cadre général 9

**1 Recherche par le contenu d'objets 3D** **11**

1.1 Recherche par le contenu . . . . . 13
1.2 Recherche par le contenu d'objets 3D . . . . . 14
1.2.1 Similarité entre objets 3D . . . . . 14
1.2.2 Recherche par similarité . . . . . 15
1.2.3 Architecture d'un système de recherche par le contenu d'objets 3D . . . . . 16
1.2.4 Critères de la recherche d'objets 3D . . . . . 18
1.3 Objet 3D . . . . . 20
1.3.1 Maillage polyédrique . . . . . 21
1.3.2 Formats d'un maillage 3D . . . . . 23
1.4 Bases d'Objets 3D . . . . . 24
1.4.1 Base de MPEG7 . . . . . 27
1.4.2 Princeton Shape Benchmark . . . . . 28
1.4.3 Base de SHREC'09 . . . . . 29

## II Normalisation 31

**2 Normalisation : Problématique & Etat de l'art** **33**

2.1 Problématique . . . . . 34
2.2 Etat de l'art . . . . . 35
2.2.1 Centrage . . . . . 35

2.2.2 Alignement . . . . . 36
2.2.3 Mise à échelle . . . . . 41
2.2.4 Normalisation complète . . . . . 43
2.3 Conclusion . . . . . 44

**3 Alignement** **47**
3.1 Objectif . . . . . 49
3.2 Réflexions & Objets 3D . . . . . 50
3.2.1 Classes de symétrie miroir . . . . . 51
3.2.2 Réflexions & Composantes principales . . . . . 52
3.3 Alignement . . . . . 57
3.3.1 Estimation de la symétrie . . . . . 58
3.3.2 Invariance locale de translation . . . . . 60
3.4 Ordre et sens des axes d'alignements . . . . . 68
3.4.1 Méthode basée sur les moments . . . . . 68
3.4.2 Méthode alternative . . . . . 70
3.5 Résultats expérimentaux . . . . . 75
3.6 Conclusion . . . . . 80

**III Recherche par le contenu d'objets 3D** **83**

**4 Etat de l'art** **85**
4.1 Approches 3D . . . . . 87
4.1.1 Approches statistiques . . . . . 87
4.1.2 Approches basées sur une partition de l'espace . . . . . 90
4.1.3 Approches par transformées . . . . . 95
4.1.4 Approches structurelles . . . . . 101
4.2 Approches 2D/3D . . . . . 105
4.2.1 Approches multi-vues . . . . . 105
4.2.2 Autres approches . . . . . 111
4.3 Conclusion . . . . . 113

**5 Indexation d'objets 3D** **117**
5.1 Approche 3D basée sur le descripteur gaussien . . . . . 119
5.1.1 *3D gaussian descriptor* . . . . . 119
5.1.2 Mesure de similarité adoptée dans 3DGA . . . . . 125

5.2 Amélioration d'approches 2D/3D . . . . . 127
5.2.1 Principe . . . . . 127
5.2.2 Indices de pertinence . . . . . 129
5.2.3 Approches 2D/3D améliorées . . . . . 133
5.2.4 Mesure de similarité adoptée dans ESA et EDBA . . . . . 138
5.3 Approches 2D/3D basées sur les lignes de profondeur . . . . . 139
5.3.1 Principe . . . . . 139
5.3.2 Descripteurs à base de lignes de profondeur . . . . . 140
5.3.3 Mesure de similarité adoptée dans DLA, MDLA et CDLA 145
5.4 Conclusion . . . . . 153

**6 Résultats expérimentaux 155**
6.1 Techniques d'évaluation . . . . . 156
6.1.1 Mesures de performance . . . . . 156
6.1.2 Courbe précision-rappel . . . . . 158
6.2 Évaluation des approches proposées . . . . . 159
6.2.1 *3D Gaussian Descriptor based Approach* . . . . . 159
6.2.2 *Enhanced Silhouette based Approach* . . . . . 167
6.2.3 *Enhanced Depth-Buffer based Approach* . . . . . 170
6.2.4 *Depth Line Approach* . . . . . 172
6.2.5 *Multi-view Depth Line Approach* . . . . . 177
6.2.6 *Cylindrical Depth Line Approach* . . . . . 181
6.3 Étude comparative . . . . . 184
6.4 Conclusion . . . . . 192

**Conclusion générale 193**

**IV Annexes 199**

**A Outils pour la description d'un objet 3D 201**
A.1 Volumes englobants . . . . . 201
A.1.1 Boîtes englobantes . . . . . 201
A.1.2 Sphère englobante . . . . . 202
A.1.3 Cylindre englobant . . . . . 203
A.2 Projections d'objet 3D . . . . . 203
A.2.1 Image Silhouette . . . . . 203

A.2.2 Image de profondeur . . . . . 204

**B 3D SHape REtrieval Contests - SHREC 207**

B.1 Résultats de SHREC'06 . . . . . 207

B.2 Résultats de SHREC'07 . . . . . 209

B.3 Résultats de SHREC'09 . . . . . 211

**C Moteur de recherche 213**

**Bibliographie 229**

**Liste des publications 245**

# Introduction

Les objets 3D sont très sollicités dans de nombreux domaines tels que la conception assistée par ordinateur, la synthèse d'image 3D, la télésurveillance, l'imagerie bio-médicale, les effets spéciaux 3D, les jeux vidéo, les univers virtuels 3D, etc. Ainsi, parmi les volumes considérables de documents multimédia disponibles, de grandes quantités de modèles 3D sont générées et utilisées, grâce aux technologies récentes de numérisation et de modélisation 3D et grâce aux avancées des moyens de communication. La rapide extension de ce domaine a demandé la mise en place de nouveaux outils permettant de structurer et gérer ces données multimédia, et en particulier, d'outils de recherche et de navigation dans de grandes bases de modèles 3D.

La complexité de ce nouveau type de données requiert de nouvelles techniques d'indexation et d'interprétation. La description textuelle d'un objet 3D n'étant pas suffisamment riche pour caractériser le contenu visuel de sa forme, des algorithmes de description automatique sont alors appliqués sur le contenu de l'objet 3D pour en extraire des signatures numériques. La recherche par le contenu d'objets 3D paraît être une solution prometteuse pour consulter et parcourir des bases d'objets.

La recherche par le contenu d'objets 3D consiste à sélectionner dans une base de données les objets visuellement proches d'un objet requête. L'indexation automatique d'une base d'objets calcule pour chaque objet un ensemble d'attributs descriptifs qui définit sa signature. La description des modèles 3D peut être envisagée de plusieurs manières en fonction de ce que l'on souhaite décrire dans la forme 3D et suivant l'utilisation qu'on en fait. Une mesure de similarité utilisant ces descripteurs permet de comparer deux objets entre eux et de répondre aux requêtes sur l'ensemble de la base. Une requête peut être exprimée sous la forme d'un objet exemple choisi à l'intérieur ou à l'extérieur de la base, ou esquissé

manuellement par l'utilisateur. Généralement, le résultat de la requête consiste en une liste ordonnée d'objets classés par ordre de similarité.

## Objectifs de la thèse

L'objectif principal de cette thèse est de proposer et développer de nouvelles méthodes efficaces pour l'indexation et la recherche par le contenu d'objets 3D dans des bases de données généralistes. Pour ce faire, la signature de l'objet 3D doit être compacte, pertinente et robuste aux transformations géométriques et topologiques. Avec une mesure de similarité adaptée, il faut que la méthode parvienne à discerner au mieux les objets de forme différente et à regrouper de manière intuitive les objets de forme similaire. Le souci d'interactivité implique de disposer d'une méthode de recherche rapide.

Les objets 3D sont donnés dans une position, orientation et échelle arbitraires. Les méthodes d'indexation 3D ne sont pas, en général, invariantes aux transformations affines. Une étape de normalisation est alors nécessaire. Elle se décompose en trois phases : le centrage, l'alignement et la mise à échelle. Alors que des solutions simples et robustes ont été proposées pour le centrage et la mise à échelle, l'alignement reste un problème complexe. Pour améliorer la précision des systèmes de recherche par le contenu, un autre objectif de cette thèse a été de proposer une méthode d'alignement satisfaisant à la fois aux contraintes d'invariance aux rotations, d'efficacité d'alignement et de faible complexité.

## Contributions de la thèse

Nous avons tout d'abord introduit une nouvelle méthode d'alignement d'objets 3D. Elle utilise les bonnes propriétés des symétries vis-à-vis de la perception humaine. Deux types de symétrie ont été considérés, la symétrie de réflexion et la symétrie locale de translation. Cette méthode calcule des alignements naturels et pertinents, et permet d'assurer l'invariance aux rotations dans un processus de recherche par le contenu.

Nous nous sommes ensuite intéressés à l'indexation d'objets 3D, et avons proposé et développé un ensemble d'approches :
- La *3D Gaussian descriptor based approach* permet de caractériser localement la

surface de l'objet 3D. Elle est basée sur des fonctions gaussiennes qui mesurent l'influence des points de la surface sur des points régulièrement répartis dans l'espace englobant l'objet 3D. Cette approche offre une description compacte et robuste de la forme 3D.
- Une technique d'amélioration, appelée *Enhanced multi-views approach*, pouvant s'insérer dans la plupart des approches multi-vues, tient compte de la disparité d'information contenue dans les différentes projections. Des valeurs de pertinence ont été définies et introduites dans la mesure de similarité afin de pondérer les contributions des projections dans la description de la forme 3D.
- Une famille d'approches multi-vues construites sur les lignes de profondeur extraites des images de profondeur. La description est fondée sur une méthode de transcodage qui transforme les lignes extraites en séquence d'états. Combinées à une mesure de similarité fondée sur la programmation dynamique, ces méthodes donnent d'excellents résultats sur les bases de données généralistes.

Dans toute cette thèse, le critère de rapidité de calcul a été pris en compte dans l'élaboration des algorithmes. Ceux-ci ont été évalués et comparés sur quatre bases généralistes, ainsi qu'à des méthodes existantes. Il ressort qu'une recherche utilisant notre descripteur fondé sur les lignes de profondeur et notre méthode d'alignement possède un fort pouvoir discriminant et permettent de retrouver efficacement les classes d'objets 3D similaires.

## Description des chapitres

Ce mémoire décrit l'ensemble des travaux menés dans le cadre de cette thèse. Il comporte trois parties :

- La **partie** I est consacrée au cadre général de la thèse.
Le **chapitre** 1 introduit le problème spécifique de la recherche par le contenu d'objets 3D. Nous y présentons les principaux critères que doit vérifier un système de recherche par le contenu 3D et les principales bases de données utilisées dans la littérature.
- La **partie** II est dédiée au problème d'alignement d'objets 3D.
Ce problème est resitué dans le cadre général de la normalisation dans le **chapitre** 2. Puis, notre nouvelle méthode d'alignement est présentée, évaluée et comparée aux méthodes existantes dans le **chapitre** 3. Nous proposons également une

perspective permettant de retrouver d'une façon semi-automatique la pose naturelle, résolvant par là même les problèmes liés à l'étiquetage et aux choix de sens des trois axes d'alignement.
- La **partie** III porte sur la recherche par le contenu d'objets 3D.
Une synthèse bibliographique avec principes, avantages et limitations des méthodes les plus représentatives est présentée dans le **chapitre** 4. Le **chapitre** 5 présente nos méthodes d'indexation d'objets 3D. Nous détaillons pour chacune son principe, ses étapes d'extraction de signature, et de mesure de similarité. Afin d'optimiser les performances de chaque méthode proposée, une évaluation expérimentale des différents paramètres est réalisée sur des bases de données généralistes. Dans le **chapitre** 6, les résultats des différentes approches introduites sont détaillés, discutés et comparés.

La conclusion générale présente une synthèse des travaux effectués dans cette thèse. Elle décrit les possibles évolutions futures de nos approches ainsi que les perspectives de cette thèse.

# Première partie

# Cadre général

# Chapitre 1

# Recherche par le contenu d'objets 3D

L'objectif de la recherche par le contenu d'objets 3D consiste à retrouver dans une base de données, les modèles visuellement similaires à un objet-requête, le but étant principalement la reconnaissance de forme. La recherche des $k$-plus proches voisins est la recherche par similarité la plus usité dans ce type de systèmes.

Alors que la couleur et la texture sont souvent utilisées dans la recherche d'images similaires, la forme est la base des descriptions visuelles des objets 3D. Ainsi, l'indexation d'objets 3D consiste à caractériser la forme des objets pour en extraire des signatures numériques. Celles-ci sont ensuite utilisées dans le calcul de la similarité entre objets pour répondre aux requêtes des utilisateurs. Généralement sensibles aux transformations géométriques et topologiques, ces méthodes de recherche d'objets 3D font appel à une étape de prétraitement pour résoudre ces problèmes. Nous présentons dans la section 1.2.1 les différentes étapes pour mesurer la similarité entre objets 3D. L'architecture d'un système général de recherche d'objets 3D par le contenu est présentée dans la section 1.2.3.

Afin de poser correctement les problématiques traitées dans cette thèse, une étude est consacrée à l'efficacité d'un système de recherche par le contenu d'objets 3D. Celle-ci reste très dépendante des besoins et des préférences de l'utilisateur. La section 1.2.4 présente les principaux critères que doit vérifier une recherche par le contenu, que ce soit de précision ou de rapidité.

Nous nous limitons ici aux objets polyédriques, c'est-à-dire aux modèles dont

la surface est décrite par un maillage polygonal qui représente un échantillonnage discret d'une surface continue. Nous présentons dans la section 1.4 les bases de données d'objets 3D utilisées par la communauté scientifique, et plus en détails, les bases de Princeton, de MPEG7 et de SHREC'09, sur lesquelles ont été effectuées les différentes expérimentations réalisées au cours de cette thèse.

## 1.1 Recherche par le contenu

L'émergence des technologies d'acquisition numérique et les échanges de l'information électronique suscitent un besoin de plus en plus croissant d'organisation et d'exploitation de grandes quantités de données. Le scénario classique d'exploitation d'un système de recherche d'information est le suivant : un utilisateur soumet une requête pour un besoin d'information et le système identifie les documents pertinents pour la requête soumise puis les retourne à l'utilisateur. Le but d'un système de recherche d'information est de retrouver les documents pertinents par rapport à la requête donnée. Toutefois, l'évaluation de la pertinence d'un document par rapport à une requête donnée n'est pas toujours aisée, la notion de pertinence étant très dépendante des préférences de l'utilisateur.

La recherche traditionnelle des documents multimédias (images, vidéos, objets 3D, etc.) par mots-clés est l'approche la plus ancienne et la plus utilisée (Google, Yahoo, YouTube). Cependant, elle reste limitée par le faible pouvoir expressif des mots, par les contraintes linguistiques (le passage d'une langue à une autre, l'ambiguïté sémantique) et par le caractère subjectif des annotations (deux personnes annoteront-elles un document donné avec les mêmes mots-clés ?). Elle nécessite l'intervention humaine et est donc contraignante pour les collections de taille importante de données si les mots clés-sont générés manuellement. De plus, notons que l'annotation ne pourra jamais décrire le contenu d'un document de façon exhaustive.

Une alternative à l'annotation manuelle est apparue il y a une quinzaine d'années : la recherche par le contenu. Elle consiste à rechercher des fichiers multimédias en n'utilisant que le document lui-même, c'est-à-dire son contenu sans aucune autre information. Dans le cas des images, des vidéo ou des objets 3D, l'idée est de caractériser le contenu visuel des documents par des descripteurs visuels et d'effectuer des recherches par similarité visuelle à partir de ces descripteurs. Alors que l'annotation de documents conduit à une indexation de nature sémantique, les descripteurs visuels sont appropriés au medium qu'ils décrivent, c'est-à-dire de nature visuelle. Il devient alors possible d'indexer automatiquement les documents et d'interroger une base de données directement à partir de leur contenu visuel.

Cette nouvelle approche permet de répondre à de nouveaux besoins dans le domaine de la recherche d'information dans les bases de données.

## 1.2 Recherche par le contenu d'objets 3D

### 1.2.1 Similarité entre objets 3D

Une recherche par le contenu d'objets 3D nécessite de disposer d'une méthode automatique pour mesurer la similarité entre deux objets. Le principe général d'une telle méthode repose sur l'hypothèse que la mesure de similarité entre deux objets 3D peut se ramener au calcul de la distance entre deux descriptions de ces objets. Un processus comparant deux objets comporte généralement trois étapes principales qui sont le prétraitement, l'extraction de la signature, et la mesure de leur similarité (cf. Figure 1.1).

- **Prétraitement de l'objet 3D :** est souvent nécessaire avant l'extraction de la signature. Le but du prétraitement est de transformer l'objet 3D pour en obtenir une description plus propre et un repère intrinsèque. On distingue deux types de prétraitements résolvant :
  *les problèmes de définition de la surface :* l'objet 3D peut être incorrectement défini topologiquement et géométriquement. Par exemple, le maillage associé peut présenter des problèmes de bruit, de déformation, des facettes dégénérées... Des solutions comme le débruitage, le filtrage et le rééchantillonnage sont alors nécessaires pour les résoudre.
  *les invariances aux transformations géométriques (translation, rotation, réflexion, changement d'échelle) :* l'objet 3D est généralement donné sous une position, orientation et échelle arbitraires. Une normalisation (centrage, alignement et mise à l'échelle) suffit pour les fixer.
- **Extraction de la signature :** décrivant l'objet 3D (sous forme d'un vecteur, d'un graphe, d'une séquence...) obtenue au moyen d'un ou plusieurs descripteurs de forme. Dans un processus général de recherche d'objets 3D, la signature de l'objet requête est la clé de recherche avec laquelle les éléments de la base vont pouvoir être comparés.
- **Mesure de similarité entre deux objets 3D :** qui consiste à comparer les deux signatures extraites en utilisant une distance. Pour effectuer

une requête de similarité par l'exemple, cette mesure sert à évaluer la similarité entre l'objet exemple et les objets de la base. Avec ces distances, les objets vont pouvoir être classés selon leur ordre de ressemblance.

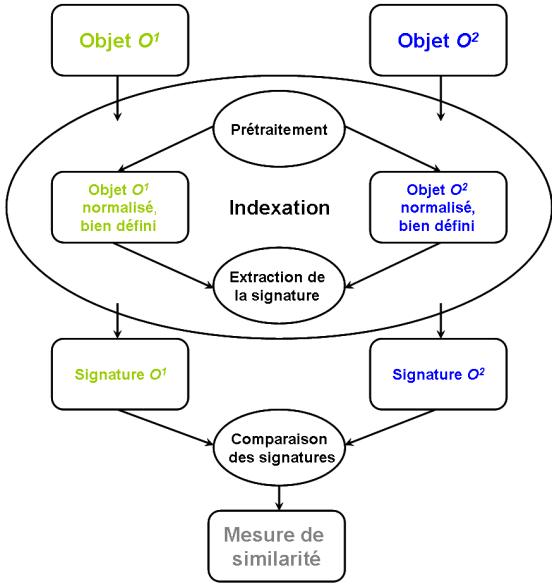



FIG. 1.1 – Processus comparant deux objets 3D.

### 1.2.2 Recherche par similarité

La recherche par le contenu visuel d'objets 3D consiste généralement à rechercher les modèles visuellement similaires à l'objet-requête. On parle de recherche par similarité, fondamentalement différente de la recherche classique par mots-clés. La recherche par similarité permet de retrouver les objets les plus proches d'un objet-requête suivant une mesure de similarité. Il existe deux types de recherche par similarité : la recherche à $\epsilon$-près et la recherche des $k$-plus proches voisins.

Dans la recherche à $\epsilon$-près, il s'agit de retrouver les objets situés à une distance d'au plus $\epsilon$ de l'objet requête, dans le but d'éliminer les objets trop éloignés de la requête. Le principal inconvénient de cette recherche est la difficulté à contrôler la taille du résultat en fonction de la valeur de $\epsilon$. Cette recherche nécessite une bonne connaissance des données pour être performante.

La recherche des $k$-plus proches voisins ($k$pp voisins) consiste à retrouver les $k$ objets les plus proches de l'objet-requête au sens de la mesure de similarité associée aux descripteurs de forme. Elle garantit automatiquement l'obtention de $k$ objets dans la liste des résultats mais certains objets de la liste peuvent être trop éloignés de la requête pour être considérés comme similaires. En pratique, la recherche des $k$-plus proches voisins est la plus courante, car elle ne nécessite pas de connaissances a priori sur la base d'objets, sur la distribution des descripteurs et des mesures de similarité. Dans notre étude, nous considérons uniquement la recherche des $k$-plus proches voisins.

### 1.2.3 Architecture d'un système de recherche par le contenu d'objets 3D

Une recherche par le contenu d'objets 3D est basée sur un calcul de similarités entre un objet exemple et les objets d'une base de données. Plus précisément, les objets résultant d'une requête présentée à l'utilisateur sont classés en fonction d'une mesure de similarité entre le descripteur de l'objet exemple et les descripteurs des objets de la base. La figure 1.2 illustre l'architecture générale d'un système de recherche par le contenu.

Un système général de recherche d'objets 3D par l'exemple est constitué de deux phases indépendantes :

**Une phase hors ligne (Offline) :** dans laquelle est réalisée l'indexation de la base de données. Durant cette phase, l'utilisateur n'est pas encore connecté au système de la recherche par le contenu. Cette phase peut donc prendre le temps nécessaire à l'extraction des descripteurs.
*L'indexation offline* consiste à extraire les signatures associées aux modèles 3D de la base de données. Celles-ci sont enregistrées dans une base de données organisée comme un dictionnaire inverse (nom du fichier et signature), qui permet de retrouver rapidement le modèle 3D associé à une signature donnée.

**Une phase en ligne (Online) :** où l'utilisateur interroge la base à l'aide de l'objet exemple. Durant cette seconde phase, le temps de réponse du système est crucial, il faut le réduire au maximum.
- *L'indexation online* concerne seulement l'objet requête. Dès que la requête est

spécifiée par l'utilisateur, la signature associée à l'objet requête est calculée.
- *La recherche online* correspond à la partie requête du moteur, qui restitue les résultats. Dans un premier temps, les similarités entre l'objet requête et tous les modèles de la base de données sont mesurées à l'aide d'une distance associée au descripteur. Ces mesures font appel à la signature de l'objet requête (indexation online) et à la base des signatures stockées (indexation offline). Puis, un algorithme de tri est appliqué sur l'ensemble des mesures de similarité pour fournir une liste complète de modèles classés selon leur ordre de ressemblance. Pour une requête donnée, les résultats de la recherche sont présentés par les $k$ premiers modèles de la liste (les $k$ plus proches voisins), $k$ étant la taille de la fenêtre de la requête, fixée par l'utilisateur.

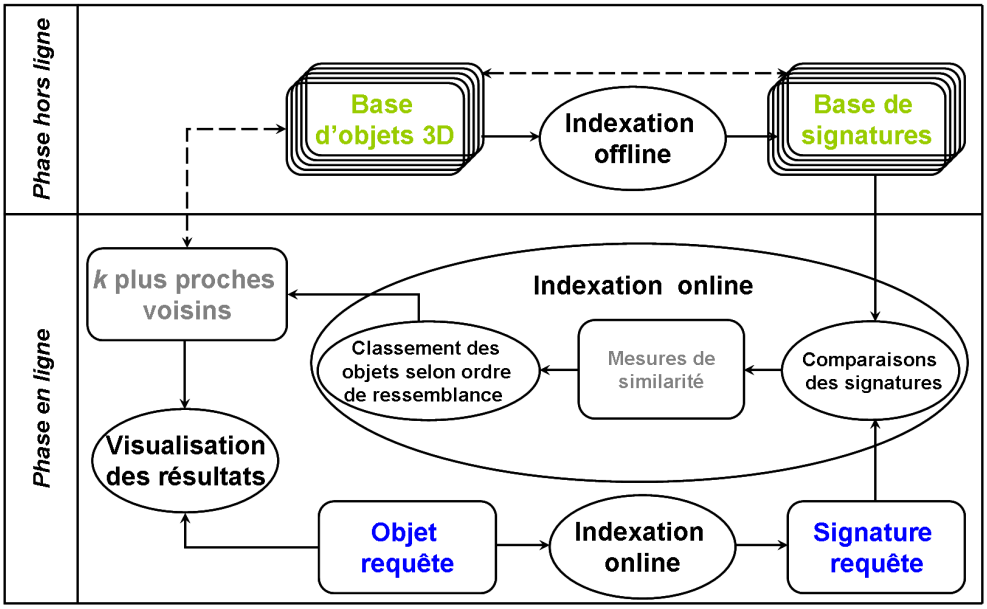



FIG. 1.2 – Architecture générale d'un système de recherche par le contenu d'objets 3D.

À notre connaissance, il n'existe pas actuellement de moteurs de recherche d'objets 3D entièrement opérationnels sur le Web. Les systèmes existants disponibles sur Internet sont plutôt envisagés comme des démonstrateurs et sont développés par des équipes de recherche.

### 1.2.4 Critères de la recherche d'objets 3D

Dans cette section, nous présentons les critères que doivent vérifier une recherche par le contenu d'objets 3D. En effet, nous ne pouvons pas conclure sur l'efficacité d'un processus de recherche, cette notion étant subjective. C'est pourquoi nous proposons d'estimer l'efficacité relativement aux besoins et aux objectifs de l'utilisateur.

Le travail bibliographique nous a permis de constater que l'efficacité d'un processus de recherche d'objets 3D peut être estimée en fonction de deux principaux objectifs qui sont la précision et la rapidité. Le but de toute requête faite sur une base d'objets 3D est de retrouver de manière précise des objets similaires à l'objet requête si la base en contient. La recherche doit être aussi rapide que possible afin de permettre une visualisation interactive des résultats. Idéalement, un processus de recherche par le contenu d'objets 3D doit pouvoir offrir des résultats à la fois rapides et précis.

Toutefois, dans la majorité des méthodes, les critères de rapidité et précision s'opposent. La communauté scientifique a proposé jusqu'à présent un compromis entre ces deux critères à cause de la complexité rencontrée dans la description de la forme et dans la définition de la similarité. D'ailleurs, entre ces deux choix, il y a toujours un juste milieu efficace, un compromis qui dépend fortement du contexte et de l'application visée. Il est à souligner que dans la plupart des méthodes, la précision est souvent mise en avant, mais la rapidité reste un critère de jugement.

#### 1.2.4.1 Critères de précision

Dans un processus de recherche par le contenu, la signature, la mesure de similarité et leur association doivent être efficacement exploitées pour permettre une meilleure précision. Pour cela, elles doivent satisfaire les critères suivants :

*Robustesse à la définition de l'objet*

- **Invariances aux transformations affines :** les objets de forme similaire doivent être retrouvés de la même façon quelles que soient leurs positions, orientations ou tailles. En particulier, la mesure de similarité entre un ob-

jet $O$ et sa transformée $\tau(O)$ doit être nulle, $\tau$ étant une concaténation de translations, de rotations, de réflexions et de changement d'échelle.

- **Invariances à la finesse et à la connectivité du maillage :** les signatures extraites des maillages (cf. Figure 1.3) représentant un même objet 3D, mais de résolutions (niveaux de détail) différentes, doivent être très proches. D'autre part, deux maillages représentant la même surface d'un même objet 3D, mais de connectivités différentes, doivent avoir la même signature.
- **Robustesse face aux bruits :** la signature ne doit pas être sensible aux bruits, aux perturbations, aux déformations locales de la surface de l'objet, à la présence de facettes dégénérés et d'auto-intersection de la surface intervenant au niveau du maillage.

*Qualité de la recherche par similarité*

- **Pertinence de la description :** la description de l'objet doit être suffisamment efficace et robuste pour fournir une signature fidèle à la forme 3D.
- **Pertinence de la similarité :** la mesure de similarité doit être adaptée au descripteur de forme pour caractériser au mieux la similarité entre deux objets 3D.
- **Discrimination de la forme :** est le critère fondamental de la précision pour que la recherche par similarité soit pertinente. Le pouvoir discriminant d'une méthode de recherche par similarité permet d'établir une distinction entre les objets similaires et non similaires à l'objet-requête. Quand l'ensemble descripteur - mesure de similarité possède un haut pouvoir discriminant, on dit qu'il a de bonnes performances de recherche.

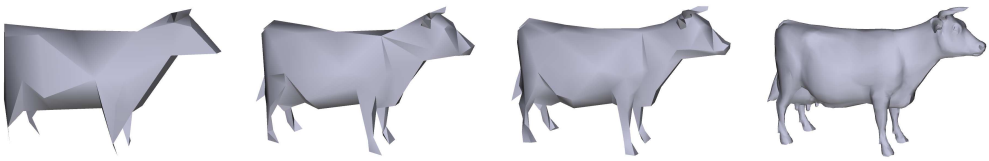


FIG. 1.3 – Maillages de résolutions différentes (100, 250, 400 et 5800 facettes) représentant le même modèle de "vache".

#### 1.2.4.2 Critères de rapidité

Le temps de réponse d'un système de recherche par le contenu à une requête donnée, c'est-à-dire, le temps écoulé entre le moment où la requête est spécifiée et le moment où des modèles similaires sont retrouvés, doit être de l'ordre de la seconde. Pour vérifier cela, trois critères seront considérés :

- **Compacité de la représentation :** la représentation de la signature doit être concise et compacte. Une signature décrivant l'objet sous forme d'un vecteur de dimension fixe est généralement plus compacte qu'une signature sous forme d'un graphe ou d'une structure plus sophistiquée. Toutefois, la dimension d'une signature doit être de taille raisonnable pour permettre d'effectuer des requêtes rapides.
- **Rapidité de l'extraction de la signature :** l'extraction de la signature, incluant l'étape de prétraitement, doit être rapide. Pour un système de recherche par le contenu, les signatures des objets de la base existante sont calculées à l'avance et seule celle de l'objet requête est calculée au moment de la requête. Néanmoins, il est souhaitable que le temps d'extraction soit le plus petit possible.
- **Rapidité de la mesure de la similarité :** le calcul de la similarité entre deux objets doit être extrêmement rapide, en particulier si la base de données est de grande taille.

## 1.3 Objet 3D

Nous pouvons distinguer différents types de représentations d'objets 3D, parmi lesquelles :

- *Les représentations surfaciques :* du type maillage 3D. Dans cette catégorie, l'objet est représenté par sa frontière. Dans le cas d'un objet polyédrique, elle est composée d'un ensemble de facettes polygonales planes. Dans la littérature, le maillage triangulaire constitue la forme la plus populaire de surfaces polyédriques.
- *Les représentations volumiques :* du type voxel. C'est une représentation de l'objet par une union de volumes unitaires et élémentaires disjointes, appelées voxels. À la différence de la représentation surfacique, elle est particulièrement utile pour représenter les données volumiques. La finesse de cette représentation dépend de la taille des voxels.

– *Les représentations algébriques :* Dans la catégorie des surfaces algébriques, on trouve notamment les surfaces implicites qui sont décrites par une équation implicite de type $f(x, y, z) = 0$.

Dans nos travaux, nous nous intéresserons uniquement aux objets 3D représentés par des maillages polygonaux. Nous en rappelons les principales caractéristiques dans la section suivante.

## 1.3.1 Maillage polyédrique

La frontière d'un objet polyédrique est représentée par une surface (ou maillage), composée d'un ensemble de facettes polygonales planes. Ces polygones sont limités par une suite d'arêtes, chacune définie par ses deux extrémités.

### 1.3.1.1 Définition d'un maillage

Un maillage d'un objet 3D est défini par ses sommets, arêtes et facettes polygonales. La plupart des standards de représentation des modèles polyédriques consistent en une énumération des sommets, numérotés implicitement par leur ordre d'apparition, puis une énumération des facettes décrites par une liste ordonnée des indices de sommets. Les arêtes sont représentées implicitement dans la description des facettes. Les positions des sommets dans l'espace 3D sont exprimées par leurs coordonnées dans un repère cartésien. Pour un maillage polygonal,

| | | | |
|---|---|---|---|
| $\mathbf{p}_1$ | $0$ | $\varphi$ | $\varphi^3$ |
| $\mathbf{p}_2$ | $0$ | $-\varphi$ | $\varphi^3$ |
| $\mathbf{p}_3$ | $\varphi^2$ | $\varphi^2$ | $\varphi^2$ |
| $\mathbf{p}_4$ | $-\varphi^2$ | $\varphi^2$ | $\varphi^2$ |
| $\mathbf{p}_5$ | $-\varphi^2$ | $-\varphi^2$ | $\varphi^2$ |
| $\mathbf{p}_6$ | $\varphi^2$ | $-\varphi^2$ | $\varphi^2$ |
| $\mathbf{p}_7$ | $\varphi^3$ | $0$ | $\varphi$ |
| $\mathbf{p}_8$ | $-\varphi^3$ | $0$ | $\varphi$ |
| $\mathbf{p}_9$ | $\varphi$ | $\varphi^3$ | $0$ |
| $\mathbf{p}_{10}$ | $-\varphi$ | $\varphi^3$ | $0$ |
| $\mathbf{p}_{11}$ | $-\varphi$ | $-\varphi^3$ | $0$ |
| $\mathbf{p}_{12}$ | $\varphi$ | $-\varphi^3$ | $0$ |
| $\mathbf{p}_{13}$ | $\varphi^3$ | $0$ | $-\varphi$ |
| $\mathbf{p}_{14}$ | $-\varphi^3$ | $0$ | $-\varphi$ |
| $\mathbf{p}_{15}$ | $\varphi^2$ | $\varphi^2$ | $-\varphi^2$ |
| $\mathbf{p}_{16}$ | $-\varphi^2$ | $\varphi^2$ | $-\varphi^2$ |
| $\mathbf{p}_{17}$ | $-\varphi^2$ | $-\varphi^2$ | $-\varphi^2$ |
| $\mathbf{p}_{18}$ | $\varphi^2$ | $-\varphi^2$ | $-\varphi^2$ |
| $\mathbf{p}_{19}$ | $0$ | $\varphi$ | $-\varphi^3$ |
| $\mathbf{p}_{20}$ | $0$ | $-\varphi$ | $-\varphi^3$ |

| | | | | | |
|---|---|---|---|---|---|
| $F_1$ | 1 | 2 | 5 | 8 | 4 |
| $F_2$ | 1 | 3 | 7 | 6 | 2 |
| $F_3$ | 1 | 4 | 10 | 9 | 3 |
| $F_4$ | 2 | 5 | 11 | 12 | 6 |
| $F_5$ | 3 | 7 | 13 | 15 | 9 |
| $F_6$ | 4 | 10 | 16 | 14 | 8 |
| $F_7$ | 5 | 8 | 14 | 17 | 11 |
| $F_8$ | 6 | 12 | 18 | 13 | 7 |
| $F_9$ | 9 | 15 | 19 | 16 | 10 |
| $F_{10}$ | 11 | 17 | 20 | 18 | 12 |
| $F_{11}$ | 13 | 18 | 20 | 19 | 15 |
| $F_{12}$ | 14 | 16 | 19 | 20 | 17 |

$\varphi = \frac{1+\sqrt{5}}{2}$ le nombre d'or
$\varphi^2 = \varphi + 1$
$\varphi^3 = 2\varphi + 1$

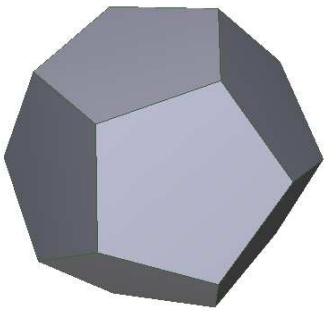


FIG. 1.4 – Représentation d'un dodécaèdre par un maillage polygonal.

les sommets donnent une information géométrique de la surface maillée alors que les arêtes et les facettes apportent une information topologique sur la connexité du maillage.

#### 1.3.1.2 Maillage triangulaire

Un maillage 3D est dit triangulaire quand toutes ses facettes sont des triangles. Un maillage polygonal est converti en un maillage triangulaire par une triangulation de chacune de ses facettes. La triangulation la plus simple consiste à partager un polygone à $k$ sommets en $(k-2)$ triangles, ayant un sommet commun et chacun une arête ne contenant pas ce sommet. Dans le cadre de cette thèse, nous avons considéré des maillages triangulaires.

| | | | |
|---|---|---|---|
| $\mathbf{p}_1$ | $0$ | $\varphi$ | $\varphi^3$ |
| $\mathbf{p}_2$ | $0$ | $-\varphi$ | $\varphi^3$ |
| $\mathbf{p}_3$ | $\varphi^2$ | $\varphi^2$ | $\varphi^2$ |
| $\mathbf{p}_4$ | $-\varphi^2$ | $\varphi^2$ | $\varphi^2$ |
| $\mathbf{p}_5$ | $-\varphi^2$ | $-\varphi^2$ | $\varphi^2$ |
| $\mathbf{p}_6$ | $\varphi^2$ | $-\varphi^2$ | $\varphi^2$ |
| $\mathbf{p}_7$ | $\varphi^3$ | $0$ | $\varphi$ |
| $\mathbf{p}_8$ | $-\varphi^3$ | $0$ | $\varphi$ |
| $\mathbf{p}_9$ | $\varphi$ | $\varphi^3$ | $0$ |
| $\mathbf{p}_{10}$ | $-\varphi$ | $\varphi^3$ | $0$ |
| $\mathbf{p}_{11}$ | $-\varphi$ | $-\varphi^3$ | $0$ |
| $\mathbf{p}_{12}$ | $\varphi$ | $-\varphi^3$ | $0$ |
| $\mathbf{p}_{13}$ | $\varphi^3$ | $0$ | $-\varphi$ |
| $\mathbf{p}_{14}$ | $-\varphi^3$ | $0$ | $-\varphi$ |
| $\mathbf{p}_{15}$ | $\varphi^2$ | $\varphi^2$ | $-\varphi^2$ |
| $\mathbf{p}_{16}$ | $-\varphi^2$ | $\varphi^2$ | $-\varphi^2$ |
| $\mathbf{p}_{17}$ | $-\varphi^2$ | $-\varphi^2$ | $-\varphi^2$ |
| $\mathbf{p}_{18}$ | $\varphi^2$ | $-\varphi^2$ | $-\varphi^2$ |
| $\mathbf{p}_{19}$ | $0$ | $\varphi$ | $-\varphi^3$ |
| $\mathbf{p}_{20}$ | $0$ | $-\varphi$ | $-\varphi^3$ |

$\varphi = \frac{1+\sqrt{5}}{2}$ le nombre d'or
$\varphi^2 = \varphi + 1$
$\varphi^3 = 2\varphi + 1$

| | | | |
|---|---|---|---|
| $T_1$ | 1 | 2 | 5 |
| $T_2$ | 1 | 5 | 8 |
| $T_3$ | 1 | 8 | 4 |
| $T_4$ | 1 | 3 | 7 |
| $T_5$ | 1 | 7 | 6 |
| $T_6$ | 1 | 6 | 2 |
| $T_7$ | 1 | 4 | 10 |
| $T_8$ | 1 | 10 | 9 |
| $T_9$ | 1 | 9 | 3 |
| $T_{10}$ | 2 | 5 | 11 |
| $T_{11}$ | 2 | 11 | 12 |
| $T_{12}$ | 2 | 12 | 6 |
| $T_{13}$ | 3 | 7 | 13 |
| $T_{14}$ | 3 | 13 | 15 |
| $T_{15}$ | 3 | 15 | 9 |
| $T_{16}$ | 4 | 10 | 16 |
| $T_{17}$ | 4 | 16 | 14 |
| $T_{18}$ | 4 | 14 | 8 |
| $T_{19}$ | 5 | 8 | 14 |
| $T_{20}$ | 5 | 14 | 17 |
| $T_{21}$ | 5 | 17 | 11 |
| $T_{22}$ | 6 | 12 | 18 |
| $T_{23}$ | 6 | 18 | 13 |
| $T_{24}$ | 6 | 13 | 7 |
| $T_{25}$ | 9 | 15 | 19 |
| $T_{26}$ | 9 | 19 | 16 |
| $T_{27}$ | 9 | 16 | 10 |
| $T_{28}$ | 11 | 17 | 20 |
| $T_{29}$ | 11 | 20 | 18 |
| $T_{30}$ | 11 | 18 | 12 |
| $T_{31}$ | 13 | 18 | 20 |
| $T_{32}$ | 13 | 20 | 19 |
| $T_{33}$ | 13 | 19 | 15 |
| $T_{34}$ | 14 | 16 | 19 |
| $T_{35}$ | 14 | 19 | 20 |
| $T_{36}$ | 14 | 20 | 17 |

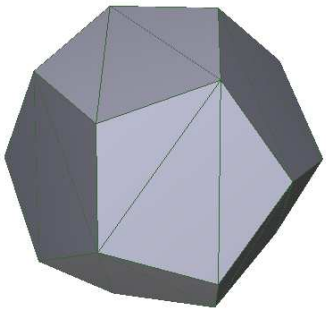


FIG. 1.5 – Représentation du dodécaèdre de la figure 1.4 par un maillage triangulaire.

Dans la suite de cette section, nous rappelons quelques notations et définitions utiles pour l'ensemble de nos développements ultérieurs.
La surface de l'objet 3D, notée $\mathfrak{S}$, est représentée par un maillage triangulaire

défini par l'ensemble des triangles

$$\mathfrak{T} = \{T_1, ..., T_{N_T}\}, \qquad T_i \subset \mathbb{R}^3, \tag{1.1}$$

donnés par l'ensemble de sommets

$$\mathfrak{P} = \{\mathbf{p}_i | \mathbf{p}_i = (x_i, y_i, z_i) \in \mathbb{R}^3, 1 \leq i \leq N_P\}. \tag{1.2}$$

Chaque triangle $T_i$ est déterminé par ses trois sommets $\mathbf{a_i}$, $\mathbf{b_i}$ et $\mathbf{c_i}$ de $\mathfrak{P}$ et défini comme un ensemble infini de points représentés par les coordonnées barycentriques de ses sommets :

$$T_i = \{\mathbf{p} | \mathbf{p} = \alpha \mathbf{a}_i + \beta \mathbf{b}_i + (1 - \alpha - \beta)\mathbf{c}_i,\ \alpha, \beta \in \mathbb{R}_+, \alpha + \beta \leq 1\}. \tag{1.3}$$

La surface $\mathfrak{S}$ est donnée par les ensembles infinis de points, définissant les triangles $T_i$, selon la formule suivante :

$$\mathfrak{S} = \bigcup_{i=1}^{N_T} T_i = \{\mathbf{p} | \mathbf{p} \in T_i, 1 \leq i \leq N_T\}. \tag{1.4}$$

Nous utiliserons cette formulation dans tous les calculs d'intégrale qui suivent. Dans la suite, le centre de gravité et l'aire de la surface du triangle $T_i$ seront notés $\mathbf{g}_i$ et $A_i$,

$$\mathbf{g}_i = \frac{\mathbf{a_i} + \mathbf{b_i} + \mathbf{c_i}}{3}, \qquad A_i = \frac{1}{2}|(\mathbf{c_i} - \mathbf{a_i}) \times (\mathbf{b_i} - \mathbf{a_i})|. \tag{1.5}$$

L'aire $\mathcal{A}$ de la surface totale du maillage $\mathfrak{S}$ est donnée par :

$$\mathcal{A} = \iint_{\mathbf{p} \in \mathfrak{S}} ds = \sum_{i=1}^{N_T} \iint_{\mathbf{p} \in T_i} ds = \sum_{i=1}^{N_T} A_i. \tag{1.6}$$

### 1.3.2 Formats d'un maillage 3D

Grâce aux nombreux outils de création et de modélisation tridimensionnelle, les maillages 3D sont aujourd'hui représentés sous une multitude de formats standards. Parmi les formats les plus répandus, on peut citer : *Object File Format* (OFF), *Virtual Reality Modeling Language* (WRL/VRML), *Wavefront Object* (OBJ), *DirectX Model* (X), *Autodesk 3D Studio* (3DS), *Maya* (MA), *AutoCAD Drawing Object* (DWG), *LightWave Object* (LWO), *Google SketchUp Document* (SKP) et *Adobe Portable Document 3D* (PDF). La plupart des modeleurs 3D proposent d'importer et d'exporter les objets sous différents formats. De plus, de

nombreux outils de conversion ont été développés au cours des dernières années pour permettre un échange de données aisé entre les différents formats.

Les standards les plus communément admis dans les systèmes de recherche par le contenu d'objets 3D sont :

- *Object File Format* (OFF) développé par *Digital Equipment Corporation's Workstation Systems Engineering* en 1986 pour l'échange et l'archivage d'objets 3D. Ce format décrit uniquement des objets ou des scènes 3D statiques. Outre la description des sommets et des facettes du maillage, il peut représenter l'attribut couleur.

- *Virtual Reality Modeling Language* (VRML) développé par le consortium Web3D pour la description d'univers statiques et interactifs 3D virtuels. Il représente une scène 3D sous forme d'un arbre hiérarchique dont les nœuds décrivent des objets ou des propriétés de la scène (maillages 3D, formes élémentaires, sons, sources de lumières, couleurs). La version du standard introduit entre autre les interpolateurs qui, attachés à un nœud animable (maillages 3D, transformations géométriques, etc.), permettent de modifier ses paramètres en fonction du temps. Plus précisément, un interpolateur est défini par un ensemble d'instants temporels et de valeurs associées à ces instants. Les players VRML génèrent l'animation en mettant à jour, à chaque instant temporel, les champs animés (position ou des normales associées aux sommets d'un maillage) grâce à des techniques d'interpolation (linéaire pour les positions et les translations, sphérique pour les normales, etc.). Des logiciels VRML permettent également d'extraire des objets 3D d'une scène tridimensionnelle. Ces outils sont nécessaires pour la recherche d'objets 3D.

Les différentes expérimentations réalisées au cours de cette thèse ont été effectuées sur quatre bases de données. Les modèles de la base de Princeton (cf. section 1.4.2) et de la base de SHREC'09 (cf. section 1.4.3) sont représentés sous le format OFF, alors que ceux de la base de MPEG'7 (cf. section 1.4.1) sont représentés sous le format VRML.

## 1.4 Bases d'Objets 3D

La plus grande base d'objets 3D est actuellement Google 3D Warehouse. Elle est disponible sur Internet et estimée à quelques millions d'objets. Pour l'alimenter, l'équipe de recherche Google a développé des outils logiciels permettant de

partager, de stocker et de rechercher par mots-clés des modèles 3D. Or jusqu'à présent, la scalabilité (passage à l'échelle) n'a pas été abordée dans le cadre de la recherche par le contenu d'objets 3D. C'est un sujet qui suscitera peut être plus d'intérêts dans le futur.

Les bases d'objets 3D peuvent être classées en deux grandes catégories pour la recherche et l'indexation :

- **Les bases généralistes** sont des bases d'objets de sujets très variés comprenant des familles d'objets très différents (par exemple personnages, véhicules, avions, animaux, arbres, bâtiments, chaises, tables, etc.)
- **Les bases spécialistes** sont des bases contenant des objets d'un seul type ou d'un domaine particulier (modèles de visages, modèles articulés, modèles de CAO, modèles archéologiques)

Bien qu'il y ait une distinction entre bases généralistes et bases spécialisées, la notion de similitude de forme dans les bases spécialisées est souvent plus difficile à interpréter que dans les bases généralistes. Dans le cas typique des bases de modèles articulés, l'information de structure sous jacente est primordiale dans la caractérisation de la forme. Ceci réduit l'espace de la description et rend plus difficile la tâche de caractérisation.

Le choix de la description d'une forme 3D dépend du type de base de données considérée et de l'utilisation qui va en être faite. Par exemple, les descripteurs adaptés à la détection d'une face dans une base de visages ne sont pas les mêmes que ceux utilisés dans la recherche d'une pièce mécanique dans une base de tôlerie. Un des challenges de la recherche d'objets 3D est de pouvoir définir ou choisir le descripteur approprié à la base de données. Orienter ainsi l'utilisateur dans le processus de recherche nécessite un minimum de connaissance sur les types d'objets présents dans la base.

Tous les systèmes de recherche par le contenu ont été testés sur des bases d'objets de petite taille. Le Tableau 1.1 récapitule les principales bases d'objets 3D privées et publiques utilisées par la communauté scientifique. Les bases généralistes contenant des familles d'objets différentes sont les plus courantes. La plupart des modèles de ces bases ont été collectés sur internet (généralement des sites www.3dcafe.com et www.viewpoint.com). C'est la raison pour laquelle de

mêmes objets sont présents dans plusieurs bases. Rappelons que, récemment, le commerce en ligne a touché le monde 3D, d'où l'apparition de sites spécialisés dans la vente des modèles 3D.

| Nature des modèles | Base d'objets 3D | ♯ modèles | accès |
|---|---|---|---|
| Modèles généralistes | NTU Database [123] | 10911 | public |
| | Princeton Shape Benchmark [104] | 1814 | public |
| | Konstanz Database [46] | 1841 | public |
| | SHREC'09 Generic Database [87] | 800 | public |
| | Utrecht Database [43] | 684 | public |
| | ITI Database [59] | 544 | public |
| | Carnegie Mellon Database [41] | 2000 | privé |
| | MPEG7 Database [55] | 227 | privé |
| Modèles de visages | York University Database [89] | 5000 | public |
| | SHREC'07 3D Face Database [88] | 1000 | public |
| | Gavab Database [38] | 549 | public |
| Modèles articulés | McGill 3D Shape Benchmark [45] | 445 | public |
| | SHREC'07 Watertight Database [86] | 400 | public |
| Modèles de CAO | Purdue Engineering Shape Benchmark [44] | 865 | public |
| | Renault Semantic Database [39] | 5000 | privé |
| Modèles archéologiques | Sculpteur Database [110] | 567 | privé |

TAB. 1.1 – Les principales bases d'objets 3D.

Une base de test est généralement associée à une classification qui constitue la vérité terrain. Elle sert à évaluer le bon comportement des méthodes de recherche en vérifiant l'appartenance des objets retournés à la classe de l'objet requête. La base peut être entièrement ou partiellement classée. Il est à souligner qu'il n'existe pas de classification unique pour une base de données car la notion de similarité reste subjective ; on peut par exemple souhaiter classer des modèles de type "animal" à un niveau très grossier (tétrapode, insecte, oiseau, poisson) ou bien à un plus petit niveau de détail (cheval, chien, papillon, abeille, etc.). Dans le premier cas, une classe peut représenter des animaux de formes très différentes, mais associés au même mot, contrairement à la deuxième classification plus fine où chaque classe ne peut représenter que des objets de formes similaires. Dans cette thèse, nous proposons que le critère de la classification d'une base de test est la forme 3D.

De plus, les outils d'évaluation des méthodes de recherche par le contenu, définis dans la section 6.1 sont fortement liés à la classification fournie. Donc, si les tests sont effectués sur une seule base de données avec une seule classification, les

résultats obtenus dépendront de la nature des objets 3D collectés et du critère de classification.

Quatre bases généralistes ont été utilisées dans l'évaluation des performances des approches proposées pour la recherche par le contenu d'objets 3D : la base de MPEG7 [55], les deux bases de Princeton [104] et la base de SHREC'09 [87]. Les modèles de ces bases sont tous des polyèdres à maillage triangulaire.

### 1.4.1 Base de MPEG7

La base de MPEG7, première base introduite, est composée de 227 modèles classés en 15 classes. La plus petite classe comporte 7 modèles, alors que la plus grande en contient 35. Les modèles diffèrent de par leur forme, leur topologie ou leur résolution de maillage. Une particularité de cette base est la présence de 50 modèles correspondant aux lettres "A", "B", "C", "D", et "E". Chaque classe de lettre contient 10 modèles de résolutions différentes, ce qui permet d'évaluer la robustesse des méthodes de la recherche par le contenu par rapport aux variations topologiques. Notons que certaines classes regroupent des objets sémantiquement et géométriquement différents : citons par exemple la classe "aérodynamique" qui contient les avions commerciaux et militaires, les planeurs, les hélicoptères, les requins et les dauphins, et la classe "tétrapodes " qui comporte les dinosaures, les crocodiles, les mammifères domestiques et les humanoïdes avec des positions de bras et jambes différentes.

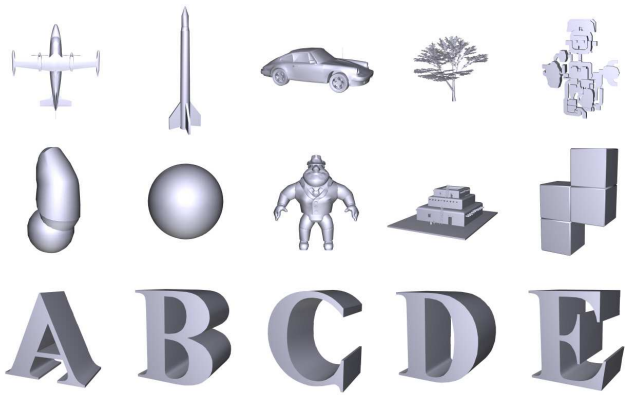



FIG. 1.6 – Modèles 3D représentant les classes de la base de MPEG7.

### 1.4.2 Princeton Shape Benchmark

L'équipe de recherche de Princeton (PSRAG) a proposé le benchmark le plus fréquemment utilisé dans le domaine de la recherche par le contenu. Il est formé de deux bases de modèles 3D : le Test Princeton Shape Benchmark (PSB Test) et le Train Princeton Shape Benchmark (PSB Train). Chaque base comporte 907 modèles entièrement classés. La base de Test est constituée de 92 classes, tandis que la base de Train est composée de 90 classes. Les classes sont de tailles inégales. La plus petite contient 4 modèles, alors que la plus grande en comporte 50. Les deux bases sont suffisamment cohérentes et sont, pour la plupart, homogènes en forme. Néanmoins, un petit nombre de classes regroupe des objets par proximité sémantique plutôt que par similarité de forme. Par exemple, les classes des satellites et des escaliers contiennent des modèles de formes très différentes.

| Classe | ♯ modèles | Classe | ♯ modèles |
|---|---|---|---|
| Biplane airplane | 14 | Cabinet furniture | 9 |
| Commercial airplane | 11 | School desk furniture | 4 |
| Fighter jet airplane | 50 | Bench seat | 11 |
| Glider airplane | 19 | Dining chair | 11 |
| Stealth bomber airplane | 5 | Desk chair seat | 15 |
| Hot air balloon balloon vehicle | 9 | Shelves furniture | 13 |
| Helicopter aircraft | 18 | Rectangular table | 25 |
| Enterprise like spaceship | 11 | Single leg round | 6 |
| Satellite spaceship | 7 | Geographic map | 12 |
| Flying saucer spaceship | 13 | Handgun gun | 10 |
| Tie fighter spaceship | 5 | Hat | 6 |
| Ant insect | 5 | Hourglass | 6 |
| Butterfly insect | 7 | Ladder | 4 |
| Human biped | 50 | Streetlight lamp | 8 |
| Human arms out | 20 | Glass with stem liquid container | 9 |
| Walking human | 8 | Pail liquid container | 4 |
| Flying bird | 14 | Vase liquid container | 11 |
| Standing bird | 7 | Mailbox | 7 |
| Dog quadruped | 7 | Electrical guitar guitar | 13 |
| Horse quadruped | 6 | Newtonian toy | 4 |
| Rabbit quadruped | 4 | Bush plant | 9 |
| Snake animal | 4 | Flowers plant | 4 |
| Sea turtle underwater creature | 6 | Potted plant plant | 26 |
| Fish underwater creature | 17 | Barren tree | 11 |
| Axe blade | 4 | Conical tree | 10 |
| Knife blade | 7 | Satellite dish | 4 |
| Sword blade | 16 | Large sail boat sailboat | 6 |
| Face body part | 16 | Ship sea vessel | 11 |
| Hand body part | 17 | Submarine sea vessel | 9 |
| Head body part | 16 | Billboard sign | 4 |
| Skull body part | 6 | Sink | 4 |
| Book | 4 | Slot machine | 4 |
| Barn building | 5 | Staircase | 7 |
| Church building | 4 | Hammer tool | 4 |
| Gazebo building | 5 | Shovel tool | 6 |
| One story home building | 14 | Umbrella | 6 |
| Skyscraper building | 5 | Race car | 14 |
| One peak tent tent | 4 | Sedan car | 10 |
| Two story home building | 10 | Covered wagon vehicle | 5 |
| Chess set | 9 | Motorcycle cycle | 6 |
| City | 10 | Monster truck vehicle | 5 |
| Desktop computer | 11 | Semi vehicle | 7 |
| Computer monitor display device | 13 | Jeep suv | 5 |
| Door | 18 | Train car | 5 |
| Eyeglasses | 7 | Wheel | 4 |
| Fireplace | 6 | Gear wheel | 9 |

TAB. 1.2 – Les différentes classes de la base Test de Princeton Shape Benchmark

### 1.4.3 Base de SHREC'09

La base de SHREC'09 est composée de 800 modèles classés en 40 classes. Contrairement aux autres bases d'objets, chaque classe comporte le même nombre (20) de modèles. Ceci permet de rendre l'évaluation objective et non biaisée par rapport à la variabilité de la taille des classes. Notons que cette catégorisation est très fine et purement sémantique. Toutefois, une variabilité importante de la forme persiste au sein de quelques classes comme les classes des "insectes", des "bâtiments", des "lampes" et des "téléphones". Par contre, certaines classes, bien que différentes comportent des objets de forme similaire. Il y a, par exemple, les "motos" et les "vélos", les "oiseaux" et les "insectes ailés", les "chaises de cuisine" et les "chaises de bureau", les "tables rondes" et les "tables rectangulaires", les "avions" et les "planeurs", les "arbres" et les" plantes en pot". Cette base de données, est, nous semble-t-il, la plus "propre" des bases d'objets 3D généralistes.

FIG. 1.7 – Modèles 3D représentant les classes de la base de SHREC'09.

# Deuxième partie

# Normalisation

# Chapitre 2

# Normalisation : Problématique & Etat de l'art

Un objet 3D est généralement donné dans l'espace 3D, $\mathbb{R}^3$, sous une position, orientation et échelle arbitraires. Les méthodes d'indexation 3D ne sont pas forcément robustes aux translations, rotations et aux variations d'échelle. Une étape de prétraitement, appelée étape de normalisation, consiste à dimensionner et positionner l'objet 3D dans un repère intrinsèque à celui-ci. Cette étape permet d'assurer l'invariance aux transformations affines du processus de recherche par le contenu d'objets 3D.

Ce chapitre présente un état de l'art en normalisation d'objets 3D. Les principales méthodes sur lesquelles se base la phase de normalisation, que ce soit pour centrer, aligner ou dimensionner les objets 3D, sont présentées. La méthode d'alignement proposée dans le chapitre 3 a été élaborée à la suite de cette étude.

## 2.1 Problématique

Un objet tridimensionnel est, a priori, donné dans un repère et à une échelle arbitraires. Le but de la normalisation est de définir un repère et une taille intrinsèques à l'objet 3D. Après normalisation, deux objets de formes similaires doivent avoir des tailles proches et doivent être positionnés et orientés quasiment de la même façon. Dans le cadre de la recherche par le contenu, l'étape de normalisation permet d'assurer l'invariance aux transformations affines suivantes :

- translation ;
- rotation et réflexion ;
- changement d'échelle.

Par conséquent, un processus de normalisation doit être :
- précis et robuste par rapport aux petites déformations de la surface de l'objet,
- invariant aux transformations appliquées sur l'objet.

Une normalisation complète d'un modèle 3D consiste en un calcul du repère et de la taille associés au modèle. Elle peut s'avérer nécessaire quand les calculs effectués sur l'objet sont sensibles aux changements de taille, d'orientation et de position. Dans d'autres cas, une normalisation partielle (en orientation, en position ou en taille) peut suffire.

Le processus de normalisation se décompose en trois phases :

- **Le centrage** (normalisation de la position) consiste à positionner l'objet par rapport à l'origine du repère. Il permet d'apporter à la méthode d'indexation l'invariance aux translations.

- **L'alignement** (normalisation de l'orientation) consiste à orienter l'objet dans l'espace 3D. Il permet d'assurer l'invariance aux transformations orthogonales (rotations et réflexions).

- **La mise à échelle** (normalisation de la taille) consiste à dimensionner l'objet. Elle permet d'apporter à la méthode d'indexation l'invariance aux changements d'échelle.

Nous présentons dans la suite de ce chapitre les principales techniques de normalisation d'objets 3D.

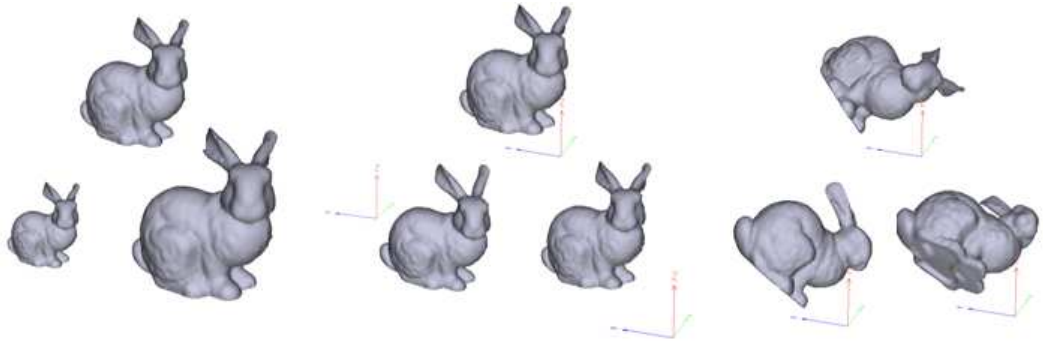


FIG. 2.1 – Un modèle 3D représenté dans l'espace 3D, sous trois positions (au centre) , trois orientations (à droite) et trois échelles différentes (à gauche).

## 2.2 Etat de l'art

### 2.2.1 Centrage

Le centrage permet de rendre l'objet robuste aux translations 3D. Il consiste à déplacer un centre particulier de l'objet (calculé à partir des données de l'objet) à l'origine du repère. Dans la littérature, on distingue :

#### 2.2.1.1 Le centre de gravité

Connu aussi sous le nom de centre d'inertie, il est le barycentre des points qui composent la surface du maillage $\mathfrak{S}$ (1.4) :

$$\mathbf{g} = \iint_{\mathbf{p}\in\mathfrak{S}} \mathbf{p}\, ds = \frac{1}{\mathcal{A}} \sum_{i=1}^{N_T} \iint_{\mathbf{p}\in T_i} \mathbf{p}\, ds = \frac{1}{\mathcal{A}} \sum_{i=1}^{N_T} A_i \mathbf{g}_i, \tag{2.1}$$

$A_i$ et $\mathbf{g}_i$ (1.5) étant respectivement l'aire et le centre de gravité de la face triangulaire $T_i$ de $\mathfrak{T}$ (1.1). Le centre d'inertie ne dépend que de la masse surfacique et de la forme de l'objet. C'est donc une caractéristique intrinsèque de l'objet.

#### 2.2.1.2 Le centre de la boîte englobante

Il est défini par les coordonnées suivants :

$$\mathbf{c}_{BB} = \left( \frac{x_{max} - x_{min}}{2}, \frac{y_{max} - y_{min}}{2}, \frac{z_{max} - z_{min}}{2} \right). \tag{2.2}$$

$x_{max}$, $y_{max}$, $z_{max}$, et $x_{min}$, $y_{min}$, $z_{min}$, sont respectivement les coordonnées maxima et minima des sommets de l'objet 3D. Ce centre n'est pas invariant aux rotations. Il est donc plus intéressant de centrer l'objet après alignement, si l'on choisit le centre de la boîte englobante comme centre.

#### 2.2.1.3 Le centre de symétrie

Il est donné par l'intersection des trois plans orthogonaux de symétrie miroir maximale. Ce centre proposé par Podolak et al. [101] est calculé après une mesure des symétries miroir de l'objet dans l'espace des plans 3D. La limitation de cette méthode est son coût de calcul et l'ambiguïté du choix du centre de symétrie quand l'objet présente des symétries locales mais pas de symétries globales.

### 2.2.2 Alignement

L'alignement a pour but de rendre l'objet robuste aux transformations orthogonales (aux rotations et aux réflexions 3D). L'invariance aux rotations est obtenue en calculant les trois directions ou les trois axes du repère et en fixant l'ordre de ces axes. L'invariance aux réflexions est obtenue par la suite en fixant les sens des trois axes.

#### 2.2.2.1 Analyse en Composantes Principales :

L'analyse en composante principale (ACP) est la méthode d'alignement la plus répandue dans la littérature. La technique consiste à calculer les axes principaux de l'objet et à les étiqueter par ordre croissant ou décroissant des valeurs propres. Elle est simple, rapide et peut être appliquée de manière automatique pour tout objet 3D. Plusieurs variantes d'ACP ont été développées dans le cadre de la recherche par le contenu.

Couramment utilisée dans l'analyse des données, l'ACP discrète est la première technique utilisée pour aligner un objet 3D. Elle est calculée sur un échantillonnage de points pris sur la surface de l'objet. Les sommets ou les centres des faces triangulaires du maillage sont généralement utilisés dans ce type d'approche. L'inconvénient majeur de cette technique est sa sensibilité aux changements de résolution du maillage. Deux techniques d'ACP ont été alors proposées pour pallier cette limitation. Paquet et al. [99] pondèrent les centres de gravité des triangles par leur surface. Cette idée a été étendue par Vranic et al. [127] dans le cas continu pour aboutir à l'ACP continue (ACPC), connue aussi sous le nom de la *Continuous Principal Component Analysis*. L'ACPC a été utilisée dans le cadre de nos travaux car elle paraît être la plus stable et la plus précise de toutes les variantes de l'ACP. Elle est détaillée dans la suite de cette section.

Contrairement à l'ACP discrète calculée seulement sur les sommets du maillage, l'approche "Continue" permet l'application de l'ACP sur un ensemble "infini" de points appartenant à une union de triangles. Cette approche est plus précise que l'ACP discrète mais elle est légèrement plus coûteuse. Comme toute décomposition en composantes principales, l'ACPC se base sur le calcul de la matrice de covariance sur $\mathfrak{S}$ (1.4). Elle peut être vue comme la matrice des moments géométriques d'ordre deux et peut s'écrire :

$$\begin{aligned} C &= \frac{1}{\mathcal{A}} \iint_{\mathbf{p}\in\mathfrak{S}} (\mathbf{p}-\mathbf{g})\cdot(\mathbf{p}-\mathbf{g})^T \; ds \\ &= \frac{1}{\mathcal{A}} \sum_{i=1}^{N_T} \iint_{\mathbf{p}\in T_i} (\mathbf{p}-\mathbf{g})\cdot(\mathbf{p}-\mathbf{g})^T \; ds \\ &= \frac{1}{\mathcal{A}} \sum_{i=1}^{N_T} \iint_{\mathbf{p}\in T_i} m(\mathbf{p})\cdot m(\mathbf{p})^T \; ds, \end{aligned}$$

avec $m(\mathbf{p}) = (\mathbf{p}-\mathbf{g})$, ce qui donne après le calcul de l'intégrale et en utilisant l'équation (1.3) :

$$C = \frac{1}{12\mathcal{A}} \sum_{i=1}^{N_T} A_i(m(\mathbf{a}_i)\cdot m(\mathbf{a}_i)^T + m(\mathbf{b}_i)\cdot m(\mathbf{b}_i)^T + m(\mathbf{c}_i)\cdot m(\mathbf{c}_i)^T + 9\; m(\mathbf{g}_i)\cdot m(\mathbf{g}_i)^T), \tag{2.3}$$

$\mathbf{a}_i$, $\mathbf{b}_i$, $\mathbf{c}_i$ et $\mathbf{g}_i$ étant respectivement les sommets et le centre de gravité du triangle $T_i$ de $\mathfrak{T}$ (1.1).

La partie restante de cette étape suit l'ACP standard. Puisque $C$ est une matrice symétrique définie positive, elle est alors diagonalisable, ses valeurs propres sont des réels positifs et ses vecteurs propres orthogonaux sont les composantes principales cherchées.

Les valeurs propres de la matrice $C$ sont calculées et triées par ordre décroissant. Puis, les vecteurs propres correspondants sont calculés et normés de façon à ce que la matrice orthogonale $(\mathbf{v}_1; \mathbf{v}_2; \mathbf{v}_3)$ soit directe (déterminant égal à 1). Nous obtenons ainsi la matrice rotation,

$$R = (\mathbf{v}_1; \mathbf{v}_2; \mathbf{v}_3), \tag{2.4}$$

qui a les vecteurs propres de $C$ comme lignes et $\triangle$ la matrice diagonale des valeurs propres correspondantes tel que $\triangle(1,1) \geq \triangle(2,2) \geq \triangle(3,3)$ et

$$C = R^T \triangle R.$$

En procédant à la rotation $R$ de tous les points de l'ensemble de départ $\mathfrak{P}$ (1.2) centré, nous définissons un nouvel ensemble de sommets ayant les axes principaux alignés avec ceux du repère.

$$\forall \mathbf{p} \in \mathfrak{P}, \qquad r(\mathbf{p}) = R \cdot (\mathbf{p} - \mathbf{g}). \tag{2.5}$$

L'ordre des axes $x$, $y$ et $z$ suit l'ordre des valeurs propres : le vecteur propre associé à la plus grande valeur propre est aligné avec le premier axe du repère $x$, le deuxième avec l'axe $y$ et le troisième avec l'axe $z$.

Afin d'obtenir un alignement plus complet, Vranic et al. [127] proposent une méthode qui assure l'invariance aux réflexions. Elle consiste à calculer les distances signées de la surface de l'objet 3D par rapport aux trois plans principaux de coordonnées $(xy)$, $(yz)$ et $(zx)$. Ces distances permettent de déterminer les sens des axes. Cette technique est précédée d'un centrage et d'un alignement assurant l'invariance aux translation et aux rotations, c'est-à-dire, la transformation affine $r$ des points de l'ensemble $\mathfrak{P}$ (cf. équation (2.5)). Une matrice diagonale définit la matrice de réflexion,

$$F = \mathrm{diag}(\mathrm{sign}(f_x), \mathrm{sign}(f_y), \mathrm{sign}(f_z)), \tag{2.6}$$

où $f_x$, $f_y$ et $f_z$ sont calculés comme suit :

pour t = x, y, z
$$\begin{aligned} f_t &= \frac{1}{\mathcal{A}} \iint_{\mathbf{p}\in\mathfrak{S}} sign(t_{r(\mathbf{p})}) \cdot t^2_{r(\mathbf{p})}\, ds \\ &= \frac{1}{\mathcal{A}} \sum_{i=1}^{N_T} \iint_{\mathbf{p}\in T_i} sign(t_{r(\mathbf{p})}) \cdot t^2_{r(\mathbf{p})}\, ds, \end{aligned}$$

ce qui donne après calcul de l'intégrale et en utilisant l'équation (1.3) :

$$f_t = \frac{1}{6\mathcal{A}} \sum_{i=1}^{N_T} A_i \cdot F_i^t, \quad \text{pour t = x, y, z ,} \tag{2.7}$$

$$F_i^t = \begin{cases} sign(t_{r(\mathbf{a}_i)}) J_i^t, & si\ sign(t_{r(\mathbf{a}_i)}) = sign(t_{r(\mathbf{b}_i)}) = sign(t_{r(\mathbf{c}_i)}) \\ sign(t_{r(\mathbf{a}_i)})(-J_i^t + 2L_i^{ta}), & si\ sign(t_{r(\mathbf{a}_i)}) \neq sign(t_{r(\mathbf{b}_i)}) = sign(t_{r(\mathbf{c}_i)}) \\ sign(t_{r(\mathbf{b}_i)})(-J_i^t + 2L_i^{tb}), & si\ sign(t_{r(\mathbf{b}_i)}) \neq sign(t_{r(\mathbf{c}_i)}) = sign(t_{r(\mathbf{a}_i)}) \\ sign(t_{r(\mathbf{c}_i)})(-J_i^t + 2L_i^{tc}), & si\ sign(t_{r(\mathbf{c}_i)}) \neq sign(t_{r(\mathbf{a}_i)}) = sign(t_{r(\mathbf{b}_i)}) \end{cases}$$

$$J_i^t = (t_{r(\mathbf{a}_i)})^2 + (t_{r(\mathbf{b}_i)})^2 + (t_{r(\mathbf{c}_i)})^2 + t_{r(\mathbf{a}_i)} t_{r(\mathbf{b}_i)} + t_{r(\mathbf{a}_i)} t_{r(\mathbf{c}_i)} + t_{r(\mathbf{b}_i)} t_{r(\mathbf{c}_i)},$$

$$L_i^{ta} = \frac{(t_{r(\mathbf{a}_i)})^4}{(t_{r(\mathbf{b}_i)} - t_{r(\mathbf{a}_i)})(t_{r(\mathbf{c}_i)} - t_{r(\mathbf{a}_i)})},$$

$$L_i^{tb} = \frac{(t_{r(\mathbf{b}_i)})^4}{(t_{r(\mathbf{a}_i)} - t_{r(\mathbf{b}_i)})(t_{r(\mathbf{c}_i)} - t_{r(\mathbf{b}_i)})},$$

$$L_i^{tc} = \frac{(t_{r(\mathbf{c}_i)})^4}{(t_{r(\mathbf{a}_i)} - t_{r(\mathbf{c}_i)})(t_{r(\mathbf{b}_i)} - t_{r(\mathbf{c}_i)})},$$

$\mathbf{a}_i$, $\mathbf{b}_i$, $\mathbf{c}_i$ étant les sommets des triangles $T_i$ de $\mathfrak{T}$ (1.1).
L'invariance aux réflexions est assurée en appliquant la transformation affine $f$ aux points de départ constituant l'ensemble $\mathfrak{P}$ (1.2). Un nouvel ensemble de points est créé selon la formule suivante :

$$\forall \mathbf{p} \in \mathfrak{P}, \qquad f(\mathbf{p}) = F \cdot R \cdot (\mathbf{p} - \mathbf{g}), \tag{2.8}$$

$F$, $R$ et $\mathbf{g}$ étant respectivement la matrice de réflexion, la matrice de rotation et le centre de gravité.

*Limitations de l' Analyse en Composantes Principales :*

L'application de l'ACP, qu'elle soit discrète ou continue, présente certaines limites conduisant à des alignements erronés. Parmi les limitations de l'ACP, nous avons :

- Des objets similaires ou appartenant à la même classe sémantique peuvent avoir un alignement très différents via l'ACP. Comme le montre la Figure 2.2, les trois modèles de tasses n'ont pas la même direction verticale.

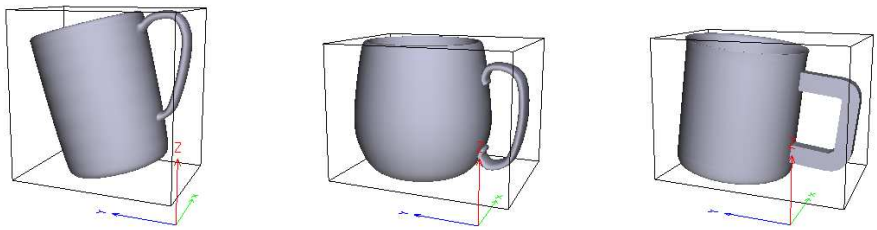


FIG. 2.2 – Alignements différents pour des objets appartenant à la même classe.

- Les directions et la pose "naturelles" de l'objet pour la perception humaine peuvent ne pas correspondre à celles calculées par l'ACP. Par exemple dans la Figure 2.3 la pose du modèle humain est inversée, la position au sol n'est pas celle du pot de la plante ou des roues du wagon.

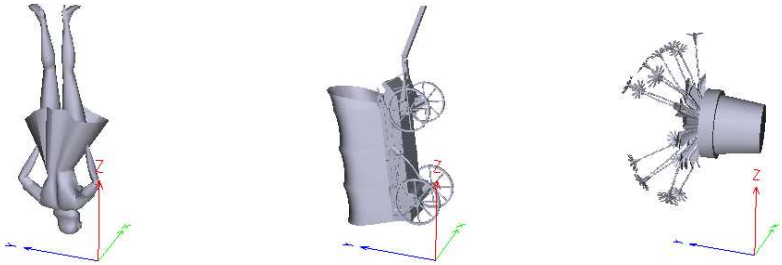


FIG. 2.3 – Exemples d'alignements calculés par l'ACP qui sont différents de la pose naturelle.

– L'étiquetage des trois axes principaux par ordre décroissant des valeurs propres de la matrice de covariance se révèle bien souvent fragile. En effet, cet ordre peut être instable pour des objets ayant des tailles similaires selon deux ou trois directions principales. Plus précisément, si les valeurs propres sont proches, les axes principaux peuvent commuter, sans affecter les valeurs propres. Pour lever ce problème, il faut générer les six copies possibles en intervertissant X, Y et Z sur les trois axes d'inertie. Les repères à considérer sont $(\mathbf{v}_1; \mathbf{v}_2; \mathbf{v}_3)$, $(\mathbf{v}_1; \mathbf{v}_3; \mathbf{v}_2)$, $(\mathbf{v}_2; \mathbf{v}_1; \mathbf{v}_3)$, $(\mathbf{v}_2; \mathbf{v}_3; \mathbf{v}_1)$, $(\mathbf{v}_3; \mathbf{v}_1; \mathbf{v}_2)$ et $(\mathbf{v}_3; \mathbf{v}_2; \mathbf{v}_1)$.

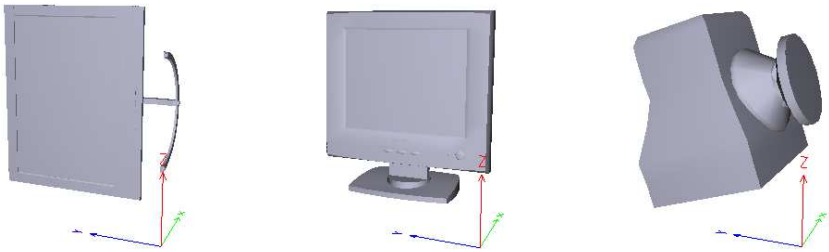


FIG. 2.4 – Exemples d'alignements erronés liés aux étiquetages des axes principaux selon l'ordre des valeurs propres.

– La méthode proposée par Vranic, pour assurer une invariance aux réflexions, ne permet pas dans tous les cas de fixer les orientations des axes. En effet, quand la distance signée de la surface 3D à un plan de base est très proche de zéro, nous ne pouvons pas choisir l'orientation de façon univoque. Un changement d'échelle de la partie gauche ou droite de l'objet en question (séparées par le plan de base) provoquera une inversion de sens. On a, de ce fait une ambiguïté d'orientation pour chaque axe. Pour lever ce problème, il faut générer les huit configurations correspondant aux huit repères pos-

sibles ayant les mêmes axes principaux mais de sens contraires : $(\mathbf{v}_1; \mathbf{v}_2; \mathbf{v}_3)$, $(-\mathbf{v}_1; \mathbf{v}_2; \mathbf{v}_3)$, $(\mathbf{v}_1; -\mathbf{v}_2; \mathbf{v}_3)$, $(\mathbf{v}_1; \mathbf{v}_2; -\mathbf{v}_3)$, $(-\mathbf{v}_1; -\mathbf{v}_2; -\mathbf{v}_3)$, $(\mathbf{v}_1; -\mathbf{v}_2; -\mathbf{v}_3)$, $(-\mathbf{v}_1; \mathbf{v}_2; -\mathbf{v}_3)$ et $(-\mathbf{v}_1; -\mathbf{v}_2; \mathbf{v}_3)$.

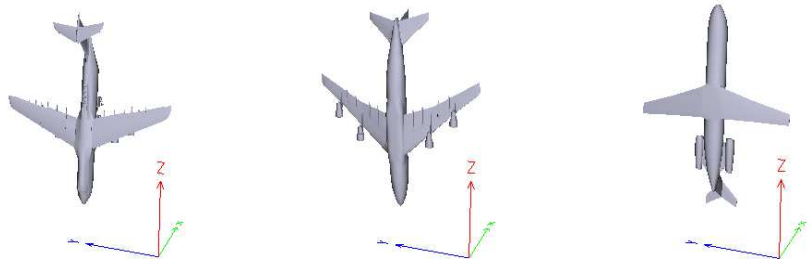



FIG. 2.5 – Exemples d'alignements erronés liés aux changements de sens des axes principaux.

#### 2.2.2.2 Alignement selon les axes de symétrie :

Podolak et al. [101] proposent d'aligner l'objet 3D en utilisant les symétries de réflexion dominantes. Dans un premier temps, l'objet 3D est représenté par des fonctions sphériques (cf. section 4.1.3.7) calculées par la transformée de symétrie de réflexion planaire (PRST) introduite par les auteurs. Cette représentation mesure la symétrie de réflexion par rapport aux plans intersectant l'objet 3D, et ceci se fait directement sur la surface de l'objet. Connaissant les valeurs PRST de chaque plan, la normale du plan ayant la symétrie maximale fournit le premier axe de symétrie $\mathbf{s}_1$. Le deuxième axe de symétrie $\mathbf{s}_2$ est sélectionné en cherchant le plan ayant la symétrie maximale parmi les plans parallèles au premier axe de symétrie $\mathbf{s}_1$. De la même manière, le troisième axe de symétrie $\mathbf{s}_3$ est donné par le plan maximisant la symétrie parmi les plans ayant une normale perpendiculaire à la fois aux $\mathbf{s}_1$ et $\mathbf{s}_2$. Comparé à l'ACP, cette méthode permet généralement des alignements plus naturels et plus cohérents à l'intérieur des classes d'objets 3D. La limitation de cette méthode est son coût de calcul.

### 2.2.3 Mise à échelle

La mise à échelle permet de rendre l'objet robuste aux changements d'échelle. Cette technique consiste à définir pour chaque objet un facteur d'échelle $s$ (calculé à partir des données de l'objet), de manière à déterminer une taille spécifique normalisée. L'invariance à l'échelle est alors réalisée en multipliant l'ensemble

$\mathfrak{P}$ (1.2) par l'inverse de $s$. Dans la littérature, plusieurs variantes de facteur d'échelle ont été développées :

#### 2.2.3.1 La distance moyenne de tous les points de l'objet au centre de gravité :

Cette normalisation d'échelle est la méthode la plus répandue pour obtenir une taille normalisée des objets 3D. Le calcul du facteur d'échelle peut se faire dans le cas continu selon la formule :

$$\begin{aligned} d_c &= \frac{1}{\mathcal{A}} \iint_{\mathbf{p}\in\mathfrak{S}} \|\mathbf{p}-\mathbf{g}\| \, ds \\ &= \frac{1}{\mathcal{A}} \sum_{i=1}^{N_T} \iint_{\mathbf{p}\in T_i} \|\mathbf{p}-\mathbf{g}\| ds \\ &= \frac{1}{\mathcal{A}} \sum_{i=1}^{N_T} \iint_{\mathbf{p}\in T_i} \|m(\mathbf{p})\| \, ds, \end{aligned}$$

avec $m(\mathbf{p}) = (\mathbf{p}-\mathbf{g})$, ce qui donne en utilisant l'équation (1.3) :

$$d_c = \frac{2}{\mathcal{A}} \sum_{i=1}^{N_T} A_i \int_0^1 \int_0^{1-\alpha} \|\alpha \; m(\mathbf{a}_i) + \beta \; m(\mathbf{b}_i) + (1-\alpha-\beta) \; m(\mathbf{c}_i)\| \; d\alpha d\beta, \quad (2.9)$$

$\mathbf{a}_i$, $\mathbf{b}_i$, $\mathbf{c}_i$ étant les sommets des triangles $T_i$ de $\mathfrak{T}$ (1.1).
Or, le calcul de ces intégrales est trop complexe et coûteux. On ne peut pas en déduire une formulation explicite comme dans le cas de l'ACP continue. Il existe plusieurs méthodes pour obtenir une approximation discrète. L'approche la plus stable, introduite par Vranic dans [131], consiste à calculer la moyenne des distances entre le centre de gravité de l'objet et les centres de gravité des triangles prélevés uniformément sur la surface du maillage, pondérées par les surfaces des triangles. Cette approximation est basée sur une subdivision pseudo uniforme de chaque triangle centré $m(T_i)$ (de sommets m($\mathbf{a}_i$), m($\mathbf{b}_i$) et m($\mathbf{c}_i$)) en $n_i$ triangles d'aire presque égales. Elle est notée $d_{moy}$ et calculée comme suit :

$$d_{moy} = \frac{1}{\mathcal{A}} \sum_{i=1}^{N_T} \sum_{j=1}^{n_i} \frac{A_i}{n_i} \|m(\mathbf{g}_i^j)\|, \quad (2.10)$$

$\mathbf{g}_i^j$ étant les centres de gravité des triangles uniformément répartis sur le triangle centré $m(T_i)$. Le nombre $p_i$ est calculé selon la formule suivante :

$$n_i = \left[ \frac{A_i}{\mathcal{A}} \; n_{min} \right],$$

où $A_i$ (1.5) et $\mathcal{A}$ (1.6) sont respectivement l'aire du triangle $T_i$ et l'aire de la surface totale de l'objet 3D, et $n_{min}$ est le nombre minimal de triangles à considérer pour tout objet 3D lors de la subdivision, $\sum_{i=1}^{N_T} n_i \geq n_{min}$.

#### 2.2.3.2 Le facteur d'échelle utilisant les distances moyennes des points de l'objet aux trois hyperplans :

Ce facteur n'est pas invariant aux translations et aux rotations. Pour le calculer, il faut d'abord centrer l'objet et aligner ses axes principaux avec ceux du repère. Il est défini comme suit :

$$d_{xyz} = \sqrt{\frac{{d_x}^2 + {d_y}^2 + d_z^2}{3}}, \tag{2.11}$$

où $d_x$, $d_y$ et $d_z$ sont respectivement les distances moyennes des points de l'objet centré aux plans principaux de coordonnées (yz), (zx) et (xy). En reprenant la transformation $r$ définie dans l'équation (2.5), les expressions de $d_x$, $d_y$ et $d_z$ peuvent s'écrire :

$$\text{pour t = x, y, z,} \qquad \begin{aligned} d_t &= \frac{1}{\mathcal{A}} \iint_{\mathbf{p}\in\mathfrak{S}} |t_{r(\mathbf{p})}| \, ds \\ &= \frac{1}{\mathcal{A}} \sum_{i=1}^{N_T} \iint_{\mathbf{p}\in T_i} |t_{r(\mathbf{p})}| \, ds. \end{aligned}$$

#### 2.2.3.3 La distance entre le centre de gravité de l'objet et le point de la surface le plus éloigné de celui-ci :

Ce facteur est invariant aux rotations, et le fait de dépendre du centre de gravité $\mathbf{g}$ lui confère une invariance aux translations. Il est donné par la formule suivante :

$$d_{max} = \max_{1\leq i\leq N_P} \|\mathbf{g} - \mathbf{p}_i\|, \tag{2.12}$$

$\mathbf{p}_i \in \mathfrak{P}$ (1.2), $1 \leq i \leq N_P$, étant les sommets du maillage. Cette distance correspond au rayon de la sphère qui englobe l'objet et de centre, le centre de gravité (cf. Annexe A.1.2).

### 2.2.4 Normalisation complète

La plupart des méthodes de recherche par le contenu ne satisfont à aucune des invariances par rapport aux transformations géométriques citées précédemment. Afin d'obtenir des résultats invariants à la fois aux translations, rotations,

réflexion et aux changements d'échelle, ces méthodes font appel à une étape de normalisation complète, alors nécessaire avant l'extraction de la signature.
En utilisant tout ce qui précède, nous définissons l'application $\tau$ permettant une normalisation complète de l'objet 3D comme suit :

$$\forall \mathbf{p} \in \mathfrak{P}, \qquad \tau(\mathbf{p}) = s^{-1} \cdot F \cdot R \cdot (\mathbf{p} - \mathbf{c}), \tag{2.13}$$

où $\mathbf{c}$ est le point particulier utilisé pour le centrage, R la matrice de rotation, F la matrice de réflexion et s le facteur d'échelle. Pour que la représentation de l'objet tridimensionnel soit indépendante des transformations affines, $\tau$ est appliqué à tous les points de départ qui constituent l'ensemble $\mathfrak{P}$ (1.2).

## 2.3 Conclusion

Dans ce chapitre, une synthèse bibliographique avec principes, avantages et limitations des méthodes les plus représentatives de la normalisation d'objets 3D a été présentée. Trois phases de la normalisation ont été identifiées et décrites en détails : le centrage, l'alignement et la mise à échelle. Alors que des solutions simples et robustes ont été proposées pour le centrage et la mise à échelle, l'alignement reste un problème complexe.

La méthode d'alignement [101] basée sur les symétries de réflexion dominantes permet d'obtenir des poses naturelles et assez cohérentes à l'intérieur d'une même classe d'objets. Elle nécessite toutefois des évaluations nombreuses de symétries miroirs dans l'espace des plans, ce qui implique un temps de calcul important. Les techniques d'analyse en composantes principales [99, 127] assurent une complexité linéaire en fonction du nombre des facettes. Elles sont extrêmement rapides, mais, restent instables. Notons que ces méthodes assurent un alignement de haute qualité pour certaines classes d'objets représentant des symétries de réflexion.

Il ressort de cette étude qu'il n'existe pas actuellement de méthode satisfaisant à la fois aux contraintes d'invariance aux rotations, d'efficacité d'alignement et de faible complexité.

La méthode d'alignement que nous proposons dans le chapitre suivant essaie de répondre à ces critères et s'attaque à cet ambitieux objectif. S'appuyant sur les propriétés de symétries, elle offre des alignements naturels et pertinents, et

permet l'invariance aux rotations pour les approches de recherche par le contenu. Pour atteindre une faible complexité, notre méthode reprend l'ACPC afin de retrouver d'éventuelles directions optimales parmi les directions principales de l'objet 3D.

# Chapitre 3

# Alignement

Dans ce chapitre nous présentons une nouvelle méthode d'alignement d'objets 3D. Elle est basée sur les propriétés de symétrie : la symétrie de réflexion et la symétrie de translation locale. Une première étape consiste à retrouver des éventuels axes d'alignement optimaux parmi les axes principaux de l'objet en s'appuyant sur des propriétés intéressantes de l'ACP vis-à-vis des symétries de réflexion. Les plans de symétrie de l'objet 3D sont utilisés comme critère de sélection. Ce prétraitement transforme la résolution du problème en une étude de plusieurs cas basée sur le nombre des axes principaux retenus. Pour évaluer la symétrie locale de translation d'une forme 3D le long d'une direction donnée, nous introduisons une nouvelle mesure que nous appelons coût d'invariance locale de translation ($CILT$).
L'approche proposée permet de rendre les méthodes d'indexation robustes aux rotations 3D, et de fournir les directions optimales pour que la caractérisation de la forme 3D soit compacte et pertinente. Les résultats expérimentaux sur les bases de données montrent que la méthode proposée calcule des alignements cohérents à l'intérieur des classes d'objets. De plus, un tel alignement d'objets 3D peut être efficacement exploité dans d'autres cadres, telles que la visualisation 3D, la comparaison de modèles, la classification ou la reconnaissance d'objets 3D. Notre méthode d'alignement a été publié dans *International Conference on Shape Modeling and Applications 2008* [31].

Les trois directions optimales calculées par notre algorithme d'alignement sont données dans un ordre et des sens arbitraires. Il y a 48 manières différentes de définir un repère de coordonnées canoniques à partir de ces directions. Dans la section 3.4.2, nous présentons une perspective de ce travail pour retrouver d'une

façon semi-automatique la pose naturelle. La méthode proposée est basée sur des hypothèses déduites de la perception humaine qui permet de réduire le nombre de cas possibles. Elle propose à l'utilisateur un ensemble de repères candidats, qui contient la solution optimale. L'ensemble des travaux présentés dans ce chapitre ont fait l'objet d'une publication dans le journal *Graphical Models* [32].

Après une brève présentation en section 3.1 des objectifs, nous décrivons dans la section 3.2 une nouvelle classification des objets polyédriques en termes de symétrie miroir. Ensuite, nous présentons notre stratégie de sélection automatique d'éventuelles directions optimales via une analyse en composantes principales. Dans la section 3.3, nous détaillons notre nouvel algorithme d'alignement d'objets 3D. Nous introduisons également les différentes techniques utilisées pour évaluer les symétries considérées. Nous abordons dans la section 3.4 les problèmes liés à l'étiquetage et aux choix de sens des trois axes d'alignement. Dans la section 3.5, les résultats d'alignement obtenus avec notre méthode sont présentés, discutés et comparés aux méthodes existantes.

## 3.1 Objectif

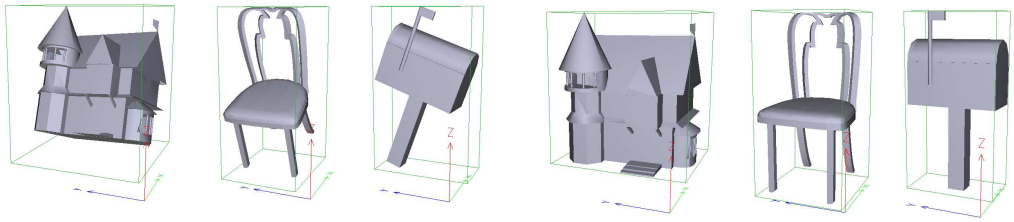
Comme énoncé dans le chapitre précédent, la nature des objets 3D nécessite la mise en place d'un processus d'alignement. Il faut alors calculer trois axes intrinsèques à l'objet et les aligner avec ceux du repère. Pour l'indexation 3D, ceci peut être considéré comme une étape de pré-description sur laquelle reposera l'extraction de la signature de l'objet 3D. Comme toute caractérisation d'objet 3D, l'alignement peut être efficacement exploité dans la recherche par le contenu s'il vérifie l'ensemble des critères suivants :

- **Invariance aux rotations 3D :** les objets de forme similaire doivent être alignés de la même façon quelles que soient leurs orientations initiales.
- **Invariance aux transformations anisotropiques 3D :** un objet aligné qui a subi un rétrécissement ou un allongement d'une taille raisonnable suivant les directions d'alignement doit conserver le même alignement.
- **Pertinence des directions calculées :** la caractérisation des directions doit être pertinente pour que les signatures calculées dans le nouveau repère le soient aussi.
- **Rapidité des calculs :** le calcul de l'alignement doit être rapide, car c'est une étape de prétraitement avant l'extraction de la signature dans le processus de recherche d'objets 3D.

Remarquons que l'alignement communément effectué par un être humain satisfait ces critères. Nous nous basons sur des observations de la perception humaine. Pour ce faire, nous utilisons le fait qu'un utilisateur cherche généralement à aligner un objet selon des axes particuliers, caractérisant des symétries, qu'elles soient locales ou globales. Cette approche permet à la fois de trouver la pose la plus naturelle d'un objet et d'aligner deux objets visuellement similaires de la même manière.

Le processus d'alignement mis en place calcule les directions support de symétries. Pour les définir, nous utilisons deux analyses de symétrie différentes : l'une caractérise les symétries miroir (globales et locales), l'autre décrit les symétries locales de translation. Trouver toutes les symétries d'une forme 3D est bien plus difficile et plus coûteux que de vérifier si une transformation est une symétrie. Une nouveauté du processus proposé est de combiner l'analyse en composantes principales et la symétrie de réflexion pour détecter les éventuels plans miroirs,

puis d'utiliser si nécessaire la symétrie locale de translation. Ceci permet de réduire considérablement le temps de calcul, et d'obtenir des solutions adéquates à nos attentes.



Alignements faits avec l'ACPC. Alignements faits avec notre méthode.

FIG. 3.1 – Alignement de modèles 3D avec deux méthodes différentes. Les flèches rouges, vertes et bleues représentent les axes principaux pour l'ACPC et les axes d'alignement optimaux pour notre méthode.

## 3.2 Réflexions & Objets 3D

La caractérisation d'un objet 3D par des symétries de réflexion a suscité beaucoup de travaux. S'appuyant essentiellement sur des études de la perception humaine [42], ces travaux ont motivé notre choix et nous ont conduit à considérer en premier lieu la symétrie de réflexion. Il est important de noter ici que ce type de symétrie, qu'elle soit globale ou locale, est présente dans la plupart des objets 3D.

Rappelons qu'une symétrie miroir est une symétrie orthogonale par rapport à un plan $\pi$ de $\mathbb{R}^3$. Elle est également appelée réflexion ou symétrie planaire. Elle est notée $\gamma_\pi$, et $\pi$ est appelé plan miroir. Celui-ci peut être caractérisé par sa normale $\mathbf{n}$ et sa distance à l'origine $\delta$. Cette symétrie associe à chaque point $\mathbf{p}$ de $\mathfrak{S}$ (1.4) un point image $\mathbf{q}$ de $\mathfrak{S}$ défini par :

$$\mathbf{q} = \mathbf{p} - 2\,(\mathbf{n}^T \cdot \mathbf{p} - \delta)\,\mathbf{n}. \tag{3.1}$$

La recherche des réflexions d'un objet 3D a fait l'objet de nombreux travaux [24, 66, 77, 80, 101, 108, 114, 116]. Notant les limites de ces approches quant au temps de calcul, nous proposons une méthode rapide se basant sur l'analyse en composantes principales. Notre but étant de calculer un alignement, c'est-à-dire les trois directions orthogonales définissant le repère intrinsèque à l'objet 3D,

il n'est pas nécessaire de calculer toutes les réflexions caractérisant l'objet 3D. Notre système fournit seulement les réflexions qui peuvent être caractérisées par les directions principales.

La recherche rapide des plans de réflexion à travers une ACP s'est basée sur deux études. Dans un premier temps, nous avons étudié les groupes de symétrie polyédriques [36] et introduit une nouvelle classification des polyèdres basée sur les réflexions. Puis, nous avons étudié, pour chaque classe, les propriétés particulières de l'ACP vis-à-vis des réflexions. Nous allons dans cette section passer en revue les différentes classes de symétrie miroir dont relèvent les objets polyédriques 3D. Nous présenterons ensuite la discussion qui a guidé ces travaux, qui consiste à évaluer la position des composantes principales par rapport aux plans miroir.

### 3.2.1 Classes de symétrie miroir

Dans cette partie, nous nous sommes intéressés aux groupes de symétrie polyédriques, et particulièrement aux comportements des réflexions à l'intérieur de chaque groupe. Une telle étude nous a amené à introduire une nouvelle classification des polyèdres en termes de symétrie miroir. Pour les objets polyédriques, le nombre de classes ponctuelles est limité à cinq. Les opérations de symétrie constituent une classe ponctuelle et peuvent être construites à partir des réflexions par rapport aux plans.

En se basant sur les groupes de symétrie polyédriques [36], on peut classer les objets polyédriques existants en cinq familles :

1. $\mathbf{G_C}$ : les objets 3D ayant une symétrie cyclique appelée aussi symétrie pyramidale à $n$ plis. Ils ont $n$ ($n > 1$) plans miroirs passant par un axe particulier, par exemple une $n$-pyramide régulière, une table allongée ($n = 2$) et une table carrée ($n = 4$).
2. $\mathbf{G_D}$ : les objets 3D ayant une symétrie diédrique appelée aussi symétrie prismatique à $n$ plis. Ils ont $n$ ($n > 1$) plans miroirs passant par un axe particulier et un plan miroir perpendiculaire à cet axe, par exemple un $n$-prisme régulier ou un $n$-bipyramide régulier.
3. $\mathbf{G_R}$ : les objets 3D ayant une symétrie polyédrique comme les cinq polyèdres réguliers convexes, appelés aussi solides de Platon. Ce groupe comporte trois

sous-groupes : $\mathbf{G_T}$ symétrie tétraédrique (6 plans miroirs), $\mathbf{G_O}$ symétrie octaédrique (9 plans miroirs) et $\mathbf{G_I}$ symétrie icosaédrique (15 plans miroirs).

4. $\mathbf{G_U}$ : les objets 3D ayant une seule symétrie miroir. C'est le cas de nombreux objets naturels comme les avions, les animaux, les humains, les chaises, les voitures, etc.
5. $\mathbf{G_Z}$ : les objets 3D n'ayant aucune symétrie miroir, comme les plantes et les arbres.

Cette classification est valable pour les symétries de réflexion parfaites. Nous l'étendrons aux symétries de réflexion quasi-parfaites dans la section 3.3.1.1.

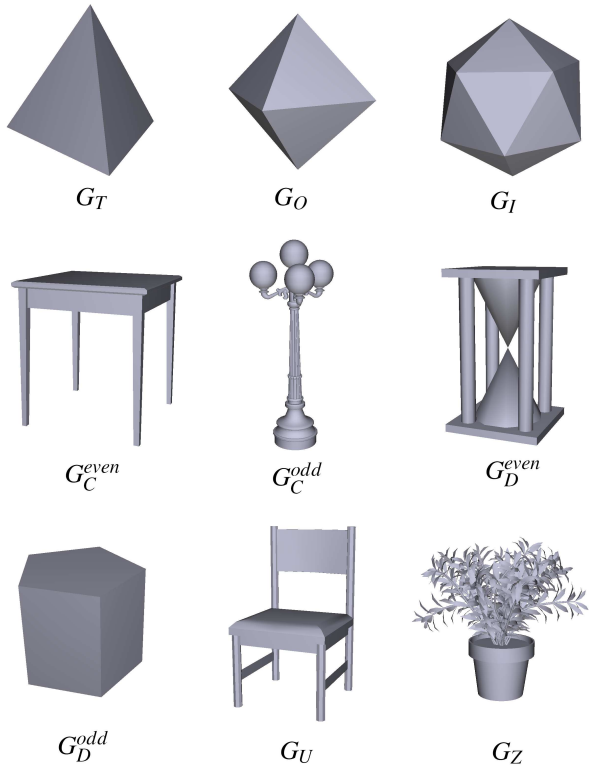



FIG. 3.2 – Exemples de modèles 3D appartenant aux différentes classes de symétrie miroir : Tétraèdre $\in G_T$, Octaèdre $\in G_O$, Icosaèdre $\in G_I$, Table $\in G_C^{even}$, Lampadaire $\in G_C^{odd}$, Sablier $\in G_D^{even}$, 5-prisme $\in G_D^{odd}$, Chaise $\in G_U$, Plante $\in G_Z$.

### 3.2.2 Réflexions & Composantes principales

Il existe un lien très étroit entre les composantes principales et les plans miroir pour tout modèle ayant des réflexions. Pour nos preuves, nous avons retenu l'ACP

Continue (ACPC) [127], car elle semble être la plus complète et la plus stable de toutes les variantes d'ACP que nous avons étudiées. Rappelons que l'ACPC calcule les trois vecteurs propres de la matrice de covariance $C$.

Selon Minovic et al. [79], si $\pi$ est un plan miroir de $\mathfrak{S}$ (1.4) et $\mathbf{n}$ est le vecteur normal de $\pi$, alors $\pi$ contient le centre de gravité de $\mathfrak{S}$ et $\mathbf{n}$ est un vecteur propre de la matrice de covariance $C$, et par la même, une composante principale de $\mathfrak{S}$. Dans la suite, nous montrons ces résultats dans le cas continu.

**Lemme 3.1.** *Soit* $\mathbf{g}$ *(2.1) le centre de gravité de* $\mathfrak{S}$ *(1.4). Si* $\pi$ *est un plan miroir de* $\mathfrak{S}$ *alors* $\mathbf{g} \in \pi$ *.*

*Démonstration.*
Soit $\mathbf{g}$ le centre de gravité de $\mathfrak{S}$. Alors, on peut écrire :

$$\mathbf{g} = \frac{1}{\mathcal{A}} \iint_{\mathbf{p}\in\mathfrak{S}} \mathbf{p}\ ds.$$

Supposons que $\pi = \{\mathbf{u} \in \mathbb{R}^3 | \mathbf{n}^T \cdot \mathbf{u} = \delta\}$ est un plan miroir de $\mathfrak{S}$. Alors,

$$\forall \mathbf{p} \in \mathfrak{S}, \exists (\mathbf{p}', d_p) \in (\mathfrak{S}, \mathbb{R}) \text{ tel que } \mathbf{n}^T \cdot \mathbf{p} = \delta + d_p \text{ et } \mathbf{n}^T \cdot \mathbf{p}' = \delta - d_p.$$

En utilisant les deux formules précédentes, la projection orthogonale de $\mathbf{g}$ sur $\pi$ peut se calculer en intégrant :

$$\begin{aligned}
\mathbf{n}^T \cdot \mathbf{g} &= \frac{1}{\mathcal{A}} \iint_{\mathbf{p}\in\mathfrak{S}} \mathbf{n}^T \cdot \mathbf{p}\ ds \\
&= \frac{1}{2\mathcal{A}} \left[ \iint_{\mathbf{p}\in\mathfrak{S}} \mathbf{n}^T \cdot \mathbf{p}\ ds + \iint_{\mathbf{p}'\in\mathfrak{S}} \mathbf{n}^T \cdot \mathbf{p}'\ ds \right] \\
&= \frac{1}{2\mathcal{A}} \left[ \iint_{\mathbf{p}\in\mathfrak{S}} (\delta + d_p)\ ds + \iint_{\mathbf{p}\in\mathfrak{S}} (\delta - d_p)\ ds \right] \\
&= \frac{1}{2\mathcal{A}} \iint_{\mathbf{p}\in\mathfrak{S}} 2\delta\ ds \quad = \quad \frac{\delta}{\mathcal{A}} \iint_{\mathbf{p}\in\mathfrak{S}} ds \quad = \quad \delta
\end{aligned}$$

Donc, $\mathbf{g} \in \pi$. □

**Lemme 3.2.** *Si* $\pi$ *est un plan miroir de* $\mathfrak{S}$ *(1.4) et* $\mathbf{n}$ *est la normale de* $\pi$*, alors* $\mathbf{n}$ *est un axe principal de* $\mathfrak{S}$*.*

*Démonstration.*
Soit $\mathbf{g}$ (2.1) le centre de gravité de $\mathfrak{S}$. Le calcul de la matrice de covariance $\mathfrak{S}$ se fait selon la formule :

$$C = \frac{1}{\mathcal{A}} \iint_{\mathbf{p}\in\mathfrak{S}} (\mathbf{p} - \mathbf{g}) \cdot (\mathbf{p} - \mathbf{g})^T\ ds$$

Supposons que $\mathfrak{S}$ a un plan miroir $\pi = \{\mathbf{u} \in \mathbb{R}^3 | \mathbf{n}^T \cdot \mathbf{u} = \delta\}$. Ceci implique que $\forall\ \mathbf{p} \in \mathfrak{S}$, $\exists\ (\mathbf{p}', \mathbf{p}_\pi, d_p) \in (\mathfrak{S}, \pi, \mathbb{R})$ tel que $\mathbf{p} = \mathbf{p}_\pi + d_p\ \mathbf{n}$ et $\mathbf{p}' = \mathbf{p}_\pi - d_p\ \mathbf{n}$. L'expression générale de la matrice de covariance à intégrer peut s'écrire :

$$\begin{aligned}
C &= \frac{1}{2\mathcal{A}} \iint_{\mathbf{p}\in\mathfrak{S}} (\mathbf{p}-\mathbf{g}) \cdot (\mathbf{p}-\mathbf{g})^T\ ds + \frac{1}{2\mathcal{A}} \iint_{\mathbf{p}'\in\mathfrak{S}} (\mathbf{p}'-\mathbf{g}) \cdot (\mathbf{p}'-\mathbf{g})^T\ ds \\
&= \frac{1}{2\mathcal{A}} \iint_{\mathbf{p}\in\mathfrak{S}} (\mathbf{p}_\pi - \mathbf{g} + d_p\ \mathbf{n}) \cdot (\mathbf{p}_\pi - \mathbf{g} + d_p\ \mathbf{n})^T\ ds \\
&\qquad + \frac{1}{2\mathcal{A}} \iint_{\mathbf{p}\in\mathfrak{S}} (\mathbf{p}_\pi - \mathbf{g} - d_p\ \mathbf{n}) \cdot (\mathbf{p}_\pi - \mathbf{g} - d_p \mathbf{n})^T\ ds \\
&= \frac{1}{\mathcal{A}} \left[ \iint_{\mathbf{p}\in\mathfrak{S}} (\mathbf{p}_\pi - \mathbf{g}) \cdot (\mathbf{p}_\pi - \mathbf{g})^T\ ds + \iint_{\mathbf{p}\in\mathfrak{S}} d_p^2\ \mathbf{n} \cdot \mathbf{n}^T\ ds \right]
\end{aligned}$$

Montrons que $\mathbf{n}$ est un vecteur propre de la matrice de covariance $C$ de $\mathfrak{S}$, si $\exists$ $\lambda \neq 0$ tel que $C \cdot \mathbf{n} = \lambda\ \mathbf{n}$.

$$\begin{aligned}
C \cdot \mathbf{n} &= \frac{1}{\mathcal{A}} \left[ \iint_{\mathbf{p}\in\mathfrak{S}} (\mathbf{p}_\pi - \mathbf{g}) \cdot (\mathbf{p}_\pi - \mathbf{g})^T\ ds \right] \cdot \mathbf{n} + \frac{1}{\mathcal{A}} \left[ \iint_{\mathbf{p}\in\mathfrak{S}} d_p^2\ \mathbf{n} \cdot \mathbf{n}^T\ ds \right] \cdot \mathbf{n} \\
&= \frac{1}{\mathcal{A}} \iint_{\mathbf{p}\in\mathfrak{S}} (\mathbf{p}_\pi - \mathbf{g}) \cdot (\mathbf{p}_\pi - \mathbf{g})^T \cdot \mathbf{n}\ ds + \frac{1}{\mathcal{A}} \iint_{\mathbf{p}\in\mathfrak{S}} d_p^2\ \mathbf{n} \cdot \mathbf{n}^T \cdot \mathbf{n}\ ds
\end{aligned}$$

En utilisant le fait que $\mathbf{g} \in \pi$ (selon le lemme précédent) et que $\mathbf{p}_\pi \in \pi$ (projection orthogonale de $\mathbf{p}$ sur $\pi$), on obtient :

$$\begin{aligned}
(\mathbf{p}_\pi - \mathbf{g})^T \cdot \mathbf{n} &= \mathbf{n}^T \cdot (\mathbf{p}_\pi - \mathbf{g}) \\
&= \mathbf{n}^T \cdot \mathbf{p}_\pi - \mathbf{n}^T \cdot \mathbf{g} \\
&= \delta - \delta \quad = \quad 0
\end{aligned}$$

En prenant $\mathbf{n}$ unitaire, on obtient :

$$\mathbf{n}^T \cdot \mathbf{n} = 1$$

Ce qui donne en combinant les trois équations :

$$C \cdot \mathbf{n} = \left[ \frac{1}{\mathcal{A}} \iint_{\mathbf{p}\in\mathfrak{S}} d_p^2\ ds \right] \mathbf{n} = \lambda\ \mathbf{n}$$

Donc, la normale $\mathbf{n}$ au plan miroir $\pi$ est un vecteur propre de $\mathfrak{S}$ et $\frac{1}{\mathcal{A}} \iint_{\mathbf{p}} d_p^2$ est la valeur propre associée. $\square$

De plus, si $\mathfrak{S}$ a $n$ plans miroirs passant par un axe particulier (dans les cas $\mathbf{G_C}, \mathbf{G_D}, \mathbf{G_R}$ de la section 3.2.1), on obtient $n$ vecteurs propres différents associés à une même valeur propre. Dans ce cas, $\mathfrak{S}$ a une symétrie de rotation d'ordre $n$ ($n > 1$) autour du même axe. Remarquons, de plus, que si $\mathfrak{S}$ a un ensemble de plans miroir passant par un même axe et répartis en couples de plans orthogonaux deux à deux, l'ACPC détecte parmi cet ensemble un couple de plans orthogonaux.

Dans la suite, pour chaque classe décrite dans la section 3.2.1, nous examinons les éventuels plans miroirs détectés via l'ACPC :
- Si $\mathcal{M} \in \mathbf{G_C}$ et $n$ est pair ($\mathcal{M} \in \mathbf{G_C^{even}}$), alors l'ACPC détecte deux normales associées à deux plans miroir orthogonaux et l'axe de la symétrie cyclique (intersection des plans miroir). Si $\mathcal{M} \in \mathbf{G_C}$ et $n$ est impair ($\mathcal{M} \in \mathbf{G_C^{odd}}$), l'ACPC ne fournit qu'une seule normale associée à un plan miroir.
- Si $\mathcal{M} \in \mathbf{G_D}$, alors l'ACPC détecte au moins deux normales. La première est associée à un des $n$ plans miroir et la deuxième est associée au plan miroir orthogonal aux $n$ plans miroir. Si, de plus, $n$ est pair, alors l'ACPC donne une troisième normale associée au plan miroir (orthogonal au premier plan miroir).
- Si $\mathcal{M} \in \mathbf{G_O}$, alors l'ACPC détecte trois normales orthogonales associées aux trois plans miroir.
- Si $\mathcal{M} \in \mathbf{G_T} \bigcup \mathbf{G_I}$, alors l'ACPC détecte une seule normale associée à un plan miroir.
- Si $\mathcal{M} \in \mathbf{G_U}$, alors l'ACPC détecte la normale associée au seul plan miroir.

| Symétrie | Classe | ♯ plans miroirs | ♯ axes principaux retenus |
|---|---|---|---|
| cyclique | $\mathbf{G_C^{even}}$ | $n$ avec $n > 1$ | 2 |
| | $\mathbf{G_C^{odd}}$ | $n$ avec $n > 1$ | 1 |
| diédrique | $\mathbf{G_D^{even}}$ | $n+1$ avec $n > 1$ | 3 |
| | $\mathbf{G_C^{odd}}$ | $n+1$ avec $n > 1$ | 2 |
| polyédrique | $\mathbf{G_T}$ | 6 | 1 |
| | $\mathbf{G_O}$ | 9 | 3 |
| | $\mathbf{G_I}$ | 15 | 1 |
| une seule réflexion | $\mathbf{G_U}$ | 1 | 1 |
| aucune réflexion | $\mathbf{G_Z}$ | 0 | 0 |

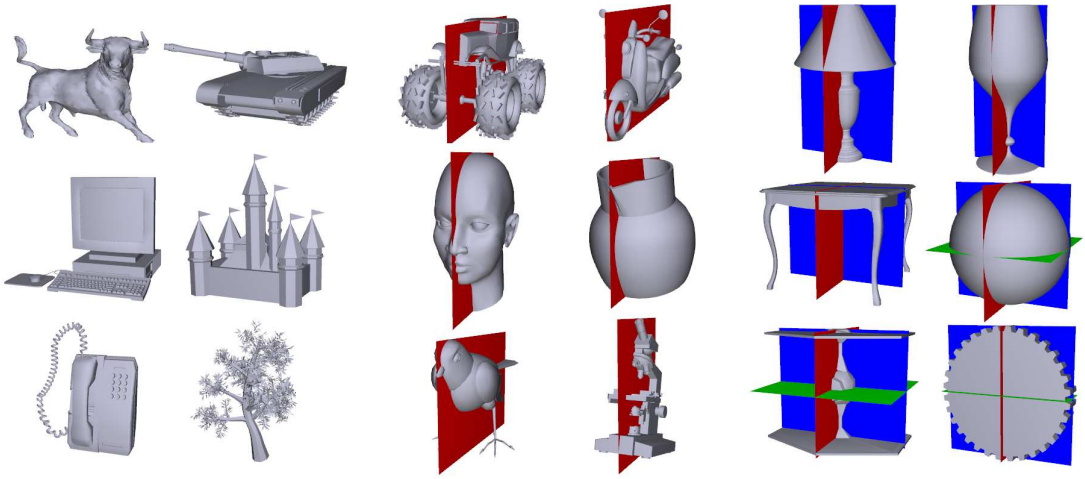
TAB. 3.1 – Les axes principaux retenus par l'ACPC pour chaque classe de symétrie miroir.

Les droites support des normales associées aux plans miroirs détectés par l'ACPC sont les premiers axes d'alignement retournés par notre méthode.

Donc, quand $\mathcal{M} \in \mathbf{G_C^{even}} \bigcup \mathbf{G_D} \bigcup \mathbf{G_O}$, l'ACPC détecte au moins deux axes d'alignement, quand $\mathcal{M} \in \mathbf{G_C^{odd}} \bigcup \mathbf{G_T} \bigcup \mathbf{G_I} \bigcup \mathbf{G_U}$, l'ACPC donne seulement un seul axe d'alignement, et quand $\mathcal{M} \in \mathbf{G_Z}$, l'ACPC ne fournit aucun axe d'alignement (cf. Figure 3.3).
Résumons cette étude en utilisant la fonction $N_{GA}(\mathcal{M})$ qui calcule le nombre d'axes d'alignement optimaux obtenus via l'ACPC. Elle est définie comme suit :

$$N_{GA}(\mathcal{M}) = \begin{cases} 2^+ & si \quad \mathcal{M} \in \mathbf{G_C^{even}} \bigcup \mathbf{G_D} \bigcup \mathbf{G_O} \\ 1 & si \quad \mathcal{M} \in \mathbf{G_C^{odd}} \bigcup \mathbf{G_T} \bigcup \mathbf{G_I} \bigcup \mathbf{G_U} \\ 0 & si \quad \mathcal{M} \in \mathbf{G_Z} \end{cases}$$

La valeur $N_{GA}$ d'un objet $\mathcal{M}$ donné est calculée en testant la réflexion de ses trois plans principaux. Cette fonction fait l'objet du test principal de notre algorithme d'alignement.



$N_{GA} = 0$ $N_{GA} = 1$ $N_{GA} = 2^+$

FIG. 3.3 – Exemples de modèles 3D regroupés selon les valeurs de $N_{GA}$.

Le calcul de $N_{GA}$ s'avère particulièrement précieux pour guider les calculs de notre algorithme d'alignement. En effet, quand $N_{GA}(\mathcal{M}) = 2^+$, le problème de l'alignement est résolu. C'est l'ACPC qui fournit l'alignement optimal. Pour les objets ayant $N_{GA}(\mathcal{M}) = 1$, un seul axe d'alignement est retourné, ce qui résout une partie du problème. Cependant, pour les objets ayant $N_{GA}(\mathcal{M}) = 0$, les directions fournies par l'ACPC ne peuvent pas être des axes d'alignement. Il nous reste maintenant à compléter l'étude dans le cas où $N_{GA}(\mathcal{M}) \in \{0, 1\}$. Ceci fait l'objet de la section suivante.

## 3.3 Alignement

Étant donné $\mathcal{M}$ un objet 3D, les grandes lignes de notre algorithme d'alignement se résument ainsi :

**Algorithme 3.1** Alignement

**1.** Centrer $\mathcal{M}$ en déplaçant son centre de gravité $\mathbf{g}$ (2.1) à l'origine du repère (cf. section 2.2.1.1) et normaliser la taille de $\mathcal{M}$ en utilisant le facteur d'échelle $d_{moy}$ (2.10) (cf. section 2.2.3.1).

**2.** Calculer les trois vecteurs propres $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ de la matrice de covariance $C$ (2.3) et transformer l'objet centré dans le nouveau système de coordonnées en appliquant une rotation $R$ $(\mathbf{v}_1; \mathbf{v}_2; \mathbf{v}_3)$ (2.4) qui a les vecteurs propres comme lignes (cf. section 2.2.2.1)

**3.** Tester la réflexion des trois plans principaux c'est-à-dire les trois nouveaux hyperplans de coordonnées $(xy)$, $(yz)$ et $(zx)$ et déduire $N_{GA}(\mathcal{M})$. Cette étape est détaillée dans la section 3.3.1.1.

**4.** Sélectionner les axes d'orientation pertinents suivant la valeur de $N_{GA}(\mathcal{M})$ :

**Si** $N_{GA}(\mathcal{M}) = 2^+$ **alors** Retourner les trois axes d'alignement $R_{ga}$ $(\mathbf{n}_1; \mathbf{n}_2; \mathbf{n}_3) = (\mathbf{v}_1; \mathbf{v}_2; \mathbf{v}_3)$.

**Si** $N_{GA}(\mathcal{M}) = 1$ **alors** Retourner le premier axe d'alignement $\mathbf{n}_1$, la normale du plan miroir détecté par l'ACPC, $\mathbf{n}_1 \in \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, et transformer l'objet dans le nouveau repère en appliquant une rotation $R'$.
- **Si** $\mathbf{n}_1 = \mathbf{v}_1$ **alors** $R'$ $(\mathbf{n}_1; \mathbf{v}_2; \mathbf{v}_3)$
- **Si** $\mathbf{n}_1 = \mathbf{v}_2$ **alors** $R'$ $(\mathbf{n}_1; \mathbf{v}_3; \mathbf{v}_1)$
- **Si** $\mathbf{n}_1 = \mathbf{v}_3$ **alors** $R'$ $(\mathbf{n}_1; \mathbf{v}_1; \mathbf{v}_2)$

**Si** $N_{GA}(\mathcal{M}) = 0$ **alors** Retourner le premier axe d'alignement $\mathbf{n}_1$, la normale du plan associé à la réflexion maximale (détaillée dans la section 3.3.1.2) et transformer l'objet dans le nouveau repère en appliquant une rotation $R'$ $(\mathbf{n}_1; \mathbf{v'}_2; \mathbf{v'}_3)$.

**5. Si** $N_{GA}(\mathcal{M}) \in \{0, 1\}$ **alors** Chercher la direction $\mathbf{n}_2$ qui maximise le coût d'invariance locale de translation $CILT$ comme le montre l'algorithme 3.2 et retourner les trois axes d'alignement $R_{ga}$ $(\mathbf{n}_1; \mathbf{n}_2; \mathbf{n}_3)$.

La Figure 3.4 illustre l'exécution de l'algorithme sur quatre modèles ayant des valeurs $N_{GA}$ différentes. Remarquons que pour les deux premiers modèles, la montgolfière et le sablier, où $N_{GA} = 2^+$, la solution est donnée par l'ACPC.

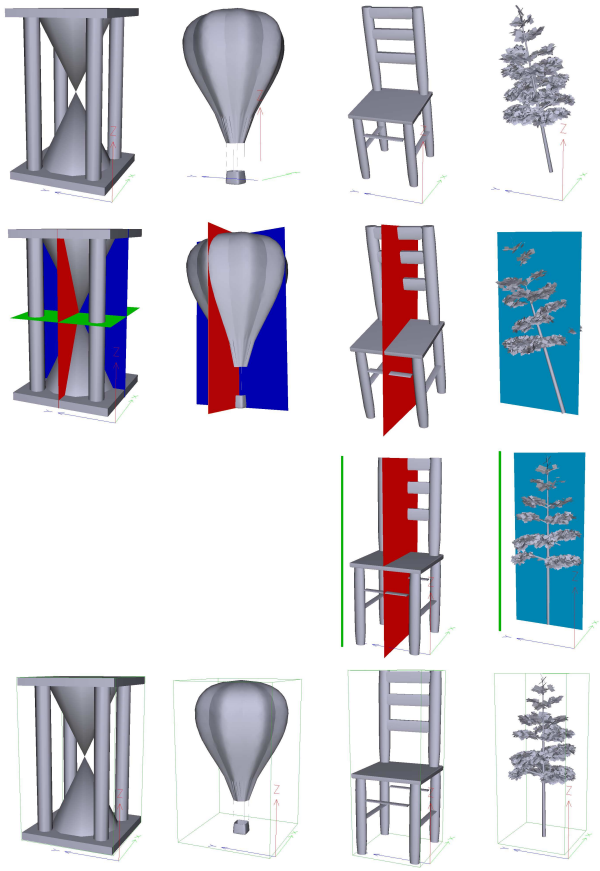


FIG. 3.4 – Les différentes étapes de notre algorithme d'alignement appliqué sur quatre modèles : montgolfière et sablier $N_{GA} = 2^+$, chaise $N_{GA} = 1$, arbre $N_{GA} = 0$. ***Ligne 1*** : Alignements obtenus via l'ACPC, ***Ligne 2*** : Test de la réflexion par rapport aux trois plans principaux de coordonnées $(xy)$, $(yz)$ et $(zx)$, le plan en bleu clair dans l'exemple de l'arbre étant le plan de réflexion maximale, ***Ligne 3*** : Sélection de la direction qui maximise le coût d'invariance locale de translation $CILT$,***Ligne 4*** : Résultats de notre algorithme d'alignement.

### 3.3.1 Estimation de la symétrie

L'estimation de la symétrie peut s'effectuer directement sur la surface de l'objet 3D, ou utiliser une représentation discrète de l'objet. On appellera :
- *Distance de symétrie* d'une forme 3D donnée, la distance minimale entre la surface de départ et son image par rapport à la symétrie. Cette mesure a l'avantage de procéder directement sur les points de la surface 3D, permettant d'obtenir une précision et une robustesse dans l'estimation de la symétrie. Par contre, elle est coûteuse pour les maillages tridimensionnels de grande taille.

- *Mesure de symétrie* d'un descripteur 3D donné, la distance de similarité entre le descripteur de la forme de départ et celui de son image par rapport à la symétrie. Cette mesure a été bien utile pour approximer la distance de symétrie. La plupart des algorithmes d'estimation de symétrie dans l'espace 3D font usage de cette mesure à cause de sa rapidité.

La suite de cette section est consacrée à ces deux méthodes d'estimation utilisées dans le cadre de la symétrie miroir.

#### 3.3.1.1 Distance de symétrie continue

Considérons $\mathfrak{S}_\gamma$ l'image de $\mathfrak{S}$ (1.4) par une symétrie orthogonale par rapport au plan $\gamma$. Elle est caractérisée par un ensemble de sommets $\mathfrak{P}_\gamma = \{\mathbf{p}'_1, ..., \mathbf{p}'_{N_P}\}$, images des sommets de $\mathfrak{S}$ (cf. $\mathfrak{P}$ (1.2) par la symétrie miroir. En se basant sur la distance entre deux surfaces 3D [10, 35] et sur la distance de symétrie [134], nous proposons une nouvelle distance de symétrie Continue $DSC_\gamma$ de $\mathfrak{S}$ par rapport à un plan miroir $\gamma$ :

$$CSD_\gamma(\mathfrak{S}) = \frac{1}{\mathcal{A}} \iint_{\mathbf{p}\in\mathfrak{S}} d(\mathbf{p}, \mathfrak{S}_\gamma)\ ds, \tag{3.2}$$

où $\mathcal{A}$ (1.6) représente l'aire de $\mathfrak{S}$ et $d$ est la distance entre le point $\mathbf{p}$ de $\mathfrak{S}$ et $\mathfrak{S}_\gamma$, définie comme suit :

$$d(\mathbf{p}, \mathfrak{S}_\gamma) = \min_{\mathbf{p}'\in\mathfrak{S}_\gamma} \|\mathbf{p} - \mathbf{p}'\|_2,$$

$\|.\|_2$ étant la distance Euclidienne.
La distance de symétrie est calculée en intégrant $d(\mathbf{p}, \mathfrak{S}_\gamma)$ sur toute la surface de l'objet 3D. Nous obtenons ainsi une distance plus précise que la distance discrète proposée dans [134], mais en un temps légèrement plus long. Toutefois, pour obtenir une évaluation correcte de distance entre un point de $\mathfrak{S}$ et la surface $\mathfrak{S}_\gamma$, chaque facette triangulaire est uniformément échantillonnée, conduisant à un échantillonnage de la surface en $N_S$ sommets. L'intégrale sur chaque triangle $T_i \subset \mathfrak{T}$(1.1) est approximativement donnée par la somme des intégrales sur les triangles $T_i^j$ obtenus après échantillonnage de $T_i$.
Pour calculer la distance minimale de chaque sommet $\mathbf{p}$ d'un triangle $T_i^j$ à la surface $\mathfrak{S}_\gamma$, il faut évaluer la distance de $\mathbf{p}$ à chaque facette de $\mathfrak{S}_\gamma$. Ceci conduit à une complexité de $\mathcal{O}(N_T N_S)$, pénalisante pour les modèles de grande taille.

Cette complexité a été réduite dans [35] en diminuant le nombre d'évaluations de distances point-triangle. L'idée consiste à partitionner la boîte englobant $\mathfrak{S}_\gamma$ en cellules cubiques et d'utiliser les cellules dans un processus d'indexation pour une recherche rapide du triangle le plus proche de $\mathfrak{S}_\gamma$ au sommet du triangle $T_i^j$.

$\gamma$ est un plan miroir parfait de $\mathfrak{S}$, si et seulement si $CSD_\gamma(\mathfrak{S})$ est nul. Dans l'étape 3 de notre algorithme d'alignement 3.1, nous proposons de considérer aussi les plans miroir quasi-parfaits, ce qui nous amène à étendre cette définition ainsi : $\gamma$ est un plan miroir de $\mathfrak{S}$ si et seulement si $SD_\gamma(\mathfrak{S}) < \epsilon$ ($\epsilon \simeq 0$).
Ceci sera testé sur les plans principaux pour calculer la valeur de $N_{GA}$.

#### 3.3.1.2 Descripteurs de symétrie

Un descripteur de symétrie représente les symétries d'un objet donné dans l'espace 3D. Il est généralement associé à un descripteur de forme basé sur une fonction sphérique ou une fonction 3D qui tourne autour du modèle. L'approche la plus commune a été proposée par Kazhdan et al. [68] et consiste à décrire les réflexions des plans passant par le centre de l'objet 3D. Podolak et al. [101] l'ont étendue en considérant tous les plans passant par le volume englobant l'objet.

Cependant le coût de calcul de ces méthodes reste élevé. Pour permettre une estimation plus rapide, nous avons repris l'idée développée dans [68], et nous nous sommes limités aux mesures de symétries par rapport aux plans passant par le centre de l'objet et qui sont dans un voisinage angulaire des plans principaux. Contrairement à [68], notre méthode d'estimation de symétrie est locale.
Nous l'utilisons dans l'étape 4 de l'algorithme 3.1 pour retrouver le premier axe d'alignement, et ceci dans le cas où $N_{GA}(\mathcal{M}) = 0$. Après une représentation des symétries dans un espace de plans, le processus de recherche parcourt les différentes valeurs du descripteur de symétrie et sélectionne le plan associé à la réflexion maximale, ce qui permet d'obtenir le premier axe d'alignement.

### 3.3.2 Invariance locale de translation

Par définition, en géométrie, la symétrie de translation désigne l'invariance d'un objet infini par rapport à une translation. Nous étendons cette notion à un objet fini, en particulier, à un objet 3D. Cette nouvelle symétrie qu'on appellera symétrie locale de translation implique que le modèle 3D a les mêmes propriétés

géométriques dans différentes régions le long d'une direction donnée. Des descripteurs de formes appropriés sont proposés pour évaluer l'invariance des propriétés géométriques.

Comme nous l'avons déjà énoncé dans notre algorithme, cette nouvelle symétrie sera évaluée dans les cas où $N_{GA}(\mathcal{M}) \in \{0, 1\}$. Trouver la direction qui maximise l'invariance locale de translation est la dernière étape de notre algorithme d'alignement. Plus précisément, nous étudions les symétries locales de translation le long des directions perpendiculaires au premier axe d'alignement $\mathbf{n}_1$ retourné dans l'étape 4 de l'algorithme 3.1.
Dans ce processus, nous sommes amenés à quantifier les invariances locales de translation. Ce calcul se fait en deux étapes. Dans un premier temps, nous calculons un descripteur de forme $f$ défini sur un intervalle à une dimension pour représenter l'objet le long d'une direction donnée. Dans un deuxième temps, une mesure de symétrie de $f$ permet de caractériser les invariances locales de translation le long de cette direction.

#### 3.3.2.1 Descripteur de forme $f$

Soient $\mathbf{d} \in \mathbb{R}^3$ un vecteur unité et $\pi_{\mathbf{d}}(\rho)$, $\rho \in \mathbb{R}$ une famille de plans perpendiculaires à $\mathbf{d}$ et à des distances signées $\rho$ de l'origine du repère. En notant $I_{\mathbf{d}}$ l'intervalle défini par les limites de la surface $\mathfrak{S}$ (1.4) selon la direction $\mathbf{d}$, $\mathfrak{S}$ peut être représentée comme suit :

$$\mathfrak{S} = \bigcup_{\rho \in I_{\mathbf{d}}} \mathcal{S}_{\mathbf{d}}(\rho),$$

$\mathcal{S}_{\mathbf{d}}(\rho)$ étant une partie ou une bande de la surface $\mathfrak{S}$ délimitée par les deux plans $\pi_{\mathbf{d}}(\rho - \delta)$ et $\pi_{\mathbf{d}}(\rho + \delta)$, $\delta \in \mathbb{R}$.
Dans la suite, nous considérons la fonction $f_{\mathbf{d}}$ définie sur l'intervalle $I_{\mathbf{d}}$, ayant des valeurs scalaires ou des valeurs vectorielles tels que $f_{\mathbf{d}}(\rho)$ soit le descripteur de forme de $\mathcal{S}_{\mathbf{d}}(\rho)$ pour tout $\rho \in I_{\mathbf{d}}$. Différents modèles de descripteurs de forme $f_{\mathbf{d}}$ sont proposés dans la section 3.3.2.4.

#### 3.3.2.2 Coût d'Invariance Locale de Translation - CILT

**Définition 3.3.** Soient $f_{\mathbf{d}}$ un descripteur de forme défini sur un intervalle $I_{\mathbf{d}}$ et $\mathbf{d}$ un vecteur unité, on dit que $f_{\mathbf{d}}$ a une *invariance locale de translation le long*

*de* $\mathbf{d}$ *sur l'intervalle* $I \subset I_{\mathbf{d}}$ si pour tout $\rho, \rho' \in I$, $f_{\mathbf{d}}(\rho) = f_{\mathbf{d}}(\rho')$, c'est-à-dire si $f_{\mathbf{d}}$ est constant sur l'intervalle $I$.

Pour mesurer l'invariance locale de translation du descripteur de forme $f_{\mathbf{d}}$, nous calculons les sous-intervalles maximaux $I_i$ de $I_d$ où $f_{\mathbf{d}}$ a une invariance locale de translation le long de $\mathbf{d}$ sur $I_i$. Le coût de cette symétrie est défini comme suit :

**Définition 3.4.** Soit $f_{\mathbf{d}}$ un descripteur de forme défini sur un intervalle $I_{\mathbf{d}}$, le *coût d'invariance locale de translation* $(CILT)$ *de* $f_{\mathbf{d}}$ *le long de* $\mathbf{d}$ est la somme des longueurs des sous-intervalles maximaux $I_i$ de $I_d$ où $f_{\mathbf{d}}$ a une invariance locale de translation le long de $\mathbf{d}$ sur $I_i$.

$$CILT(f_{\mathbf{d}}) = \sum_{I_i \in \mathcal{I}} \mathcal{L}(I_i), \tag{3.3}$$

avec $\mathcal{L}(I_i)$ la longueur de $I_i$ et $\mathcal{I} = \{I_i \subset I_d | I_i \text{ maximal} ; f_{\mathbf{d}} \text{ a une invariance locale de translation le long de } \mathbf{d} \text{ sur } I_i\}$.

#### 3.3.2.3 CILT dans le calcul de l'alignement

Dans cette section, la symétrie locale de translation est utilisée pour retrouver le deuxième axe d'alignement. Connaissant le premier axe d'alignement $\mathbf{n}_1$ retourné dans l'étape 4 de l'algorithme 3.1, le deuxième axe d'alignement sera celui dont la direction maximise le coût d'invariance local de translation parmi les directions perpendiculaires à $\mathbf{n}_1$.

Le calcul des CILT le long des directions perpendiculaires à $\mathbf{n}_1$ se fait d'une façon discrète, en effectuant des rotations de l'objet autour de $\mathbf{n}_1$ comme l'illustre la Figure 3.5. Pour évaluer les symétries locales de translations le long de ces directions, nous définissons un ensemble de transformations $\mathcal{R}_{\mathbf{n}}^K$ donné par les rotations $R_k$ d'angles $\theta_k = \frac{\pi k}{K}$ autour de l'axe $\mathbf{n}_1$ avec $0 \leq k < K$ :

$$R_k = R_{(1,0,0)}^{\theta_k} \cdot R', \tag{3.4}$$

$$\text{avec} \qquad R_{(1,0,0)}^{\theta_k} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos(\theta_k) & -sin(\theta_k) \\ 0 & sin(\theta_k) & cos(\theta_k) \end{bmatrix}$$

et $R'$ la matrice rotation ayant les coordonnées de $\mathbf{n}_1$ comme première ligne (cf. Algorithme 3.1) .

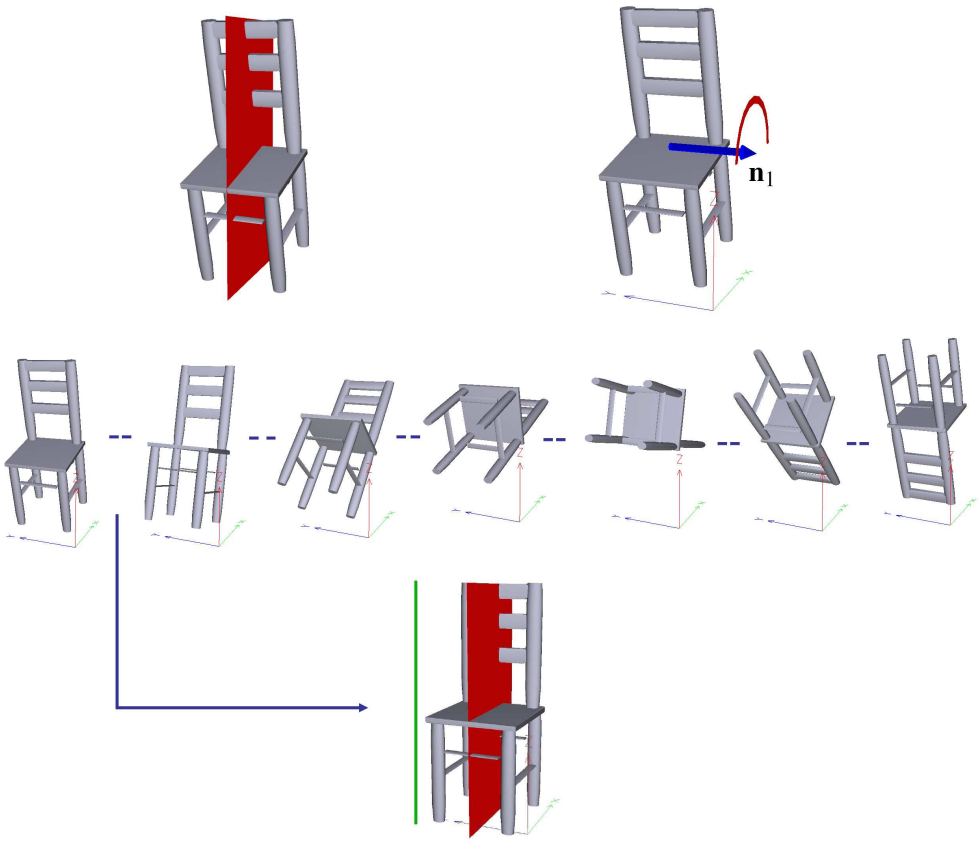



FIG. 3.5 – Calcul des axes d'alignement pertinents d'une chaise $N_{GA}(\mathcal{M}) = 1$. ***Ligne 1*** : Sélectionner $\mathbf{n}_1$ (en bleu) la normale du plan miroir détecté par l'ACPC comme premier axe d'alignement, ***Ligne 2*** : Tourner l'objet autour de $\mathbf{n}_1$ et calculer le CILT pour chaque rotation, ***Ligne 3*** : Trouver le deuxième axe d'alignement $\mathbf{n}_2$ (en vert) la direction qui maximise le CILT. La chaise est alors alignée selon les directions $\mathbf{n}_1$ et $\mathbf{n}_2$.

Dans la suite, nous associons à chaque $R_k \in \mathcal{R}^K_{\mathbf{n}}$ un vecteur directeur unité $\mathbf{d}_k$ qui est donné par les coordonnées de la deuxième ligne de la matrice $R_k$. Pour chaque $k$, un descripteur de forme $f_k$ est calculé selon la direction $\mathbf{d}_k$ pour en déduire son coût d'invariance locale de translation. Calculer le deuxième axe d'alignement revient à trouver la direction $\mathbf{d}_{ga}$ ou la rotation associée $R_{ga}$, qui maximise le $CILT(f_k)$ :

$$R_{ga} = \underset{R_k \in \mathcal{R}^K_{\mathbf{n}}}{\operatorname{argmax}} \quad CILT(f_k). \tag{3.5}$$

L'algorithme calculant la deuxième direction pertinente $\mathbf{d}_{ga}$ (connaissant la première $\mathbf{n}_1$) est le suivant :

**Algorithme 3.2** Calcul de la matrice rotation $R_{ga}$ maximisant le $CILT$

**Pour tout** $\theta_k$, $0 \leq k < K$ **faire**
Calculer la rotation $R_k$ (3.4) et la direction associée $\mathbf{d}_k$.
Appliquer la rotation $R_k$ sur l'objet transformé (étape 1), pour obtenir $\mathfrak{S}_k$.
Calculer l'intervalle $I_k$ de longueur $\mathcal{L}_{I_k}$ et le descripteur $f_k$ de $\mathfrak{S}_k$ défini sur $I_k$.
Mesurer $CILT(f_k)$ (3.3).
**Fin pour**
Retourner $R_{ga}$ (3.5) associée à $f_{ga}$ qui maximise $CILT$.

Connaissant la transformation $R'$, cet algorithme permet de trouver la rotation $R_{ga}$ maximisant le coût d'invariance locale de translation dans l'espace $\mathcal{R}^K_{\mathbf{n}}$. Par construction, la première ligne de la matrice $R_{ga}$ est le premier axe d'alignement $\mathbf{n}_1$. La deuxième ligne est le deuxième axe d'alignement $\mathbf{n_2} = \mathbf{d}_{ga}$. La troisième ligne de $R_{ga}$ correspondra par conséquent au troisième axe d'alignement $\mathbf{n_3}$.

#### 3.3.2.4 Exemples de descripteur de forme $f_k$

Comme nous l'avons déjà énoncé dans l'algorithme précédent, $\mathfrak{S}_k$ est la surface de l'objet donnée dans le système de coordonnées défini par $(\mathbf{n}_1, \mathbf{d}_k, \mathbf{n}_1 \wedge \mathbf{d}_k)$. $\mathcal{S}_k(\rho)$ et $I_k$ sont définis au début de cette section en remplaçant $\mathbf{d} = \mathbf{d}_k$.
Pour extraire l'information des invariances locales de translation, trois descripteurs de formes $G_k$, $E_k$ et $F_k$ ont été proposés : le premier caractérise les régions en fonction de leur gravité, le deuxième les caractérise en fonction de leur extremum et le troisième s'attache à la forme d'un point de vue plus local. Le processus de description mis en place caractérise la surface 3D $\mathfrak{S}_k$ le long d'une seule direction $\mathbf{n}_1 \wedge \mathbf{d}_k$. Nous ne considérons pas les directions $\mathbf{n}_1$ et $\mathbf{d}_k$ car la première est déjà sélectionnée comme axe optimal et la deuxième est l'axe du paramètre $\rho$ définissant les $\mathfrak{S}_k(\rho)$.

- *Moyenne globale $G_k$ :*

Ce descripteur est représenté par une courbe donnée par les projections des centres de gravité des sous-parties $\mathcal{S}_k(\rho)$ sur la direction $\mathbf{n}_1 \wedge \mathbf{d}_k$ :

$$G_k(\rho) = \frac{1}{\mathcal{A}_k(\rho)} \iint_{\mathbf{p} \in \mathcal{S}_k(\rho)} (\mathbf{n}_1 \wedge \mathbf{d}_k)^T \cdot \mathbf{p} \; ds, \tag{3.6}$$

$\mathcal{A}_k(\rho) = \iint_{\mathbf{p} \in \mathcal{S}_k(\rho)} ds$ étant l'aire de $\mathcal{S}_k(\rho)$. Ce descripteur permet de mesurer l'invariance locale de translation de la gravité le long de la direction $\mathbf{d}_k$.

- *Extremum global $E_k$ :*

Ce descripteur est représenté par deux courbes données par les projections des extremums des sous-parties $\mathcal{S}_k(\rho)$ sur la direction $\mathbf{n}_1 \wedge \mathbf{d}_k$ :

$$E_k(\rho) = \left( \min_{\mathbf{p} \in \mathcal{S}_k(\rho)} ((\mathbf{n}_1 \wedge \mathbf{d}_k)^T \cdot \mathbf{p}) \ , \ \max_{\mathbf{p} \in \mathcal{S}_k(\rho)} ((\mathbf{n}_1 \wedge \mathbf{d}_k)^T \cdot \mathbf{p}) \right). \tag{3.7}$$

Ce descripteur permet de mesurer l'invariance locale de translation du couple formé par les coordonnées maxima et minima le long de la direction $\mathbf{d}_k$.

- *Moyenne vectorielle $F_k$ :*

Soit $J_{\mathbf{k}}$ l'intervalle défini par les limites de la surface $\mathfrak{S}_k$ selon la direction $\mathbf{n}_1 \wedge \mathbf{d}_k$. L'espace de $\mathfrak{S}_k$ est subdivisé en $M_k$ cellules le long de la direction $\mathbf{n}_1 \wedge \mathbf{d}_k$ en échantillonnant uniformément l'intervalle $J_k$ :

$$\mathcal{S}_k(\rho) = \bigcup_{j=1}^{M_k} \mathcal{S}_k^j(\rho),$$

$\mathcal{S}_k^j(\rho)$ étant l'intersection de $\mathcal{S}_k(\rho)$ et la $j^{ème}$ cellule. Le descripteur $F_k$ représentant $\mathcal{S}_k$ est donné par les gravités et les aires des surfaces $\mathcal{S}_k^j(\rho)$, $1 \leq j \leq M_k$ :

$$F_k(\rho) = \left(a_k^j(\rho), g_k^j(\rho)\right)_{M_k}, \tag{3.8}$$

$$\text{avec} \qquad a_k^j(\rho) = \begin{cases} \iint_{\mathbf{p} \in \mathcal{S}_k^j(\rho)} ds & \text{si } \mathcal{S}_k^j(\rho) \neq \emptyset, \\ \emptyset & \text{sinon.} \end{cases}$$

$$g_k^j(\rho) = \begin{cases} \frac{1}{a_k^j(\rho)} \iint_{\mathbf{p} \in \mathcal{S}_k^j(\rho)} (\mathbf{n}_1 \wedge \mathbf{d}_k)^T \cdot \mathbf{p} \ ds & \text{si } \mathcal{S}_k^j(\rho) \neq \emptyset, \\ \emptyset & \text{sinon.} \end{cases}$$

Ce descripteur permet de mesurer de manière plus précise l'invariance locale de la forme le long de la direction $\mathbf{d}_k$.

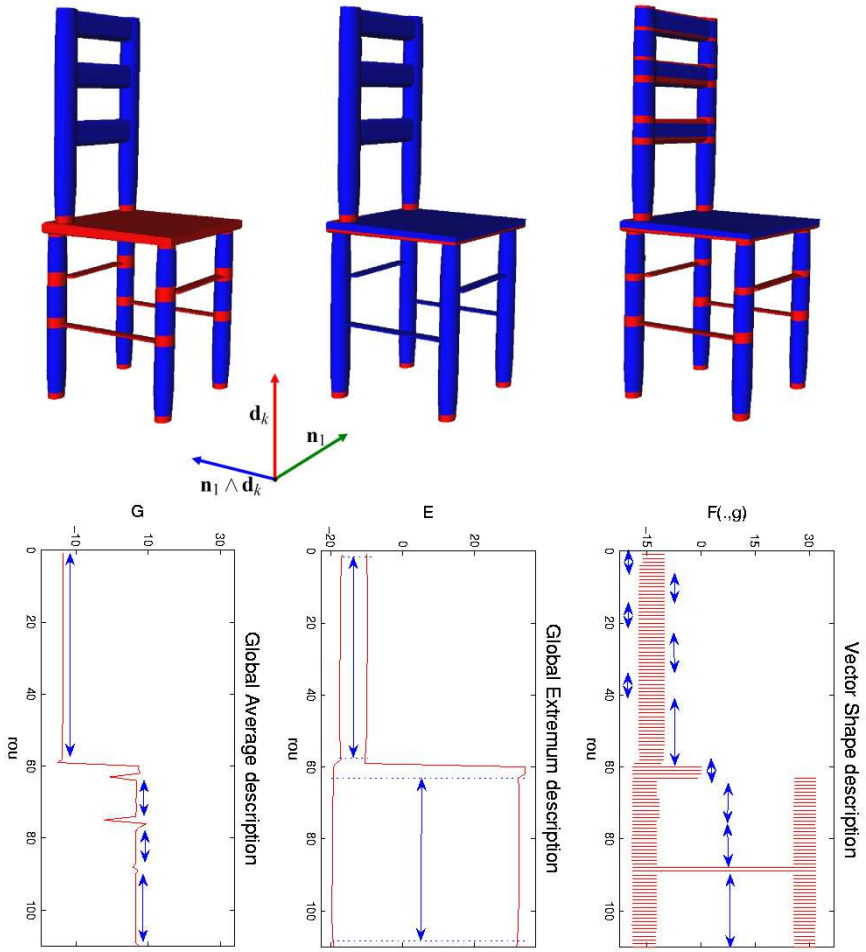



FIG. 3.6 – Trois descripteurs de formes $G$, $E$ et $F$ calculés sur une chaise le long de la direction verticale $\mathbf{d}_k$ et utilisés pour extraire l'information des invariances locales de translation. Dans ce cas, les trois descripteurs correspondent à la même direction qui maximise le CILT. Les régions de la surface associées aux intervalles ayant des invariances locales de translation sont colorées en bleu et les autres sont colorées en rouge.

#### 3.3.2.5 Distances entre composantes $f_k(i)$ et $f_k(i')$

Après avoir proposé trois méthodes pour calculer le descripteur de forme $f_k$, nous détaillons les paramètres de sa version discrète. Pour une direction $d_k$ donnée, $f_k(\rho)$ $\rho \in I_{\mathbf{d}}$, sera représenté par l'ensemble discret, $f_k(i)$ $1 \leq i \leq N_k$, défini sur $N_k$ points répartis régulièrement sur l'intervalle $I_k$. Pour obtenir une discrétisation uniforme sur toutes les directions $d_k$, chaque partition de $I_k$ doit avoir une longueur fixe $\frac{\mathcal{L}(I_k)}{N_k} = 2\delta$ (cf. section 3.3.2.1), et ceci pour toutes les directions $d_k$. Nous proposons de définir une unité de mesure $N = \frac{s}{2\delta}$, $s$ étant la dimension de l'objet 3D. La mise à l'échelle effectuée lors de la première étape de l'algorithme 3.1 permet de fixer $N = \frac{1}{2\delta}$ pour tous les objets 3D ($N = 32, 64, 128$). Nous définissons le nombre de points $N_k$ comme suit :

$$N_k = \lfloor N \ \mathcal{L}(I_k) \rfloor. \tag{3.9}$$

De la même façon, nous prenons $M_k = \lfloor N \ \mathcal{L}(J_k) \rfloor$ dans le calcul de $F_k$ (3.8).

Pour appliquer la définition 3.3 de l'invariance locale de translation, une distance $dist(f_k(i), f_k(i'))$ (avec $f_k(i)$, $1 \leq i \leq N_k$) est associée à chaque descripteur $f_k$ proposé. Dans le cas discret, nous considérons que la relation d'égalité $f_k(i) = f_k(i')$ est vérifiée quand $dist(f_k(i), f_k(i')) < \epsilon_N$, $\epsilon_N \simeq 0$ étant une erreur normalisée fixée pour tous les objets 3D. Nous utilisons la distance Euclidienne $dist(,) = l_1$ pour la *moyenne globale* $G_k$ (3.6) et l'*extremum global* $E_k$ (3.7), et nous proposons une distance plus appropriée à la *moyenne vectorielle* $F_k$ (3.8) :

$$dist(F_k(i), F_k(i')) = \frac{d_{F_k}(i, i') + d_{F_k}(i', i)}{\mathcal{A}_k(i) + \mathcal{A}_k(i')},$$

$$\text{avec} \quad d_{F_k}(i, i') = \sum_{1 \leq j \leq M_k}^{g_k^j(i) \neq \emptyset} a_k^j(i) \min_{1 \leq j' \leq M_k}^{g_k^{j'}(i') \neq \emptyset} \|g_k^j(i) - g_k^{j'}(i')\|_1.$$

La distance $d_{F_k}(i, i')$ a une complexité de l'ordre $\mathcal{O}(M_k M_k)$. Pour améliorer le temps de calcul, nous réduisons le nombre des évaluations des distances $\|g_k^j(i) - g_k^{j'}(i')\|_1$, $1 \leq j, j' \leq M_k$. Pour ce faire, nous supposons une cohérence spatiale : l'indice $j_m = \text{argmin}_{1 \leq j' \leq M_k} \|g_k^j(i) - g_k^{j'}(i')\|_1$ est proche de l'indice $j$. Ainsi, nous testons d'abord la valeur de $g_k^j(i')$ et nous nous arrêtons quand $g_k^j(i') \neq \emptyset$. Sinon, les points adjacents sont traités, dans l'ordre croissant de la distance de $g_k^j(i)$ et nous nous arrêtons quand la condition $g_k^{j_m}(i') \neq \emptyset$ est vérifiée. Notons que tous les points non-évalués sont plus éloignés que le point trouvé.

## 3.4 Ordre et sens des axes d'alignements

Les trois axes d'alignement $\mathbf{n_1}, \mathbf{n_2}, \mathbf{n_3}$ de $R_{ga}$ calculés par l'algorithme 3.1 sont donnés dans un ordre et des sens arbitraires. En considérant l'étiquetage $(x, y, z)$ des trois axes d'alignement avec deux sens possibles pour chaque axe, il y a 48 manières différentes de définir un repère de coordonnées canoniques.

Dans ce travail, nous avons utilisé la méthode basée sur des moments pour fixer un seul repère parmi les 48 repères. Cette méthode garantit généralement le même ordre et les mêmes sens des axes d'alignement pour les objets de forme similaire. Cependant, ce choix reste un problème délicat pour la perception humaine : certains objets ont une position naturelle qui n'est pas forcément la solution donnée par la méthode basée sur des moments.

Dans la section 3.4.2, nous présentons une perspective de ce travail : une méthode alternative basée sur des hypothèses déduites de la perception humaine qui permet de réduire le nombre de cas possibles. Elle propose à l'utilisateur un ensemble de repères candidats, qui contient la solution optimale. Dans certains cas, une seule solution est retournée (cf. cas (a) de la Figure 3.8).

### 3.4.1 Méthode basée sur les moments

Une utilisation des moments permet de déterminer l'ordre et les sens des axes d'alignement. La méthode s'inspire des travaux de Vranic et al. [127] (cf. section 2.2.2.1). On procède alors comme suit :
- Pour fixer l'ordre des axes d'alignement, nous calculons les projections moyennes des points de $\mathfrak{S}_{ga}$ sur les hyperplans de coordonnées $(yz)$, $(zx)$ et $(xy)$, $\mathfrak{S}_{ga}$ étant la surface du modèle dans le nouveau repère $R_{ga}$.

$$c_x = \frac{1}{\mathcal{A}} \iint_{\mathbf{p} \in \mathfrak{S}_{ga}} p_x.p_x \; ds, \;\; c_y = \frac{1}{\mathcal{A}} \iint_{\mathbf{p} \in \mathfrak{S}_{ga}} p_y.p_y \; ds, \;\; c_z = \frac{1}{\mathcal{A}} \iint_{\mathbf{p} \in \mathfrak{S}_{ga}} p_z.p_z \; ds,$$

avec $\mathbf{p} = (p_x, p_y, p_z)$.
Nous étiquetons les trois axes d'alignement par ordre décroissant des valeurs $c_x, c_y$ et $c_z$ pour en déduire une matrice de rotation $\mathbf{A}$ dont les lignes sont des vecteurs unitaires $(1, 0, 0)$, $(0, 1, 0)$ et $(0, 0, 1)$. Cette rotation $\mathbf{A}$ sera appliquée sur l'ensemble des sommets de $\mathfrak{S}_{ga}$ pour donner une nouvelle surface $\mathfrak{S}'_{ga}$.
- Pour déterminer les sens des axes d'alignement, nous calculons la distance signée

de la nouvelle surface $\mathfrak{S}'_{ga}$ aux trois hyperplans de coordonnées $(yz)$, $(zx)$ et $(xy)$ définie comme suit :

$$f_x = \frac{1}{\mathcal{A}} \iint_{\mathbf{p} \in \mathfrak{S}'_{ga}} sign(p_x)|p_x|^2 ds \ , \ \ \mathbf{p} = (p_x, p_y, p_z) \quad \text{et similairement pour } f_y, f_z.$$

La matrice diagonale $\mathbf{F} = diag(sign(f_x), sign(f_y), sign(f_z))$ sera appliquée sur l'ensemble des sommets de $\mathfrak{S}'_{ga}$.
Nous obtenons finalement la transformation affine définie par le produit $\mathbf{F}.\mathbf{A}$ qui donnera un seul repère intrinsèque à l'objet 3D avec un seul ordre des axes et un seul sens pour chaque axe.

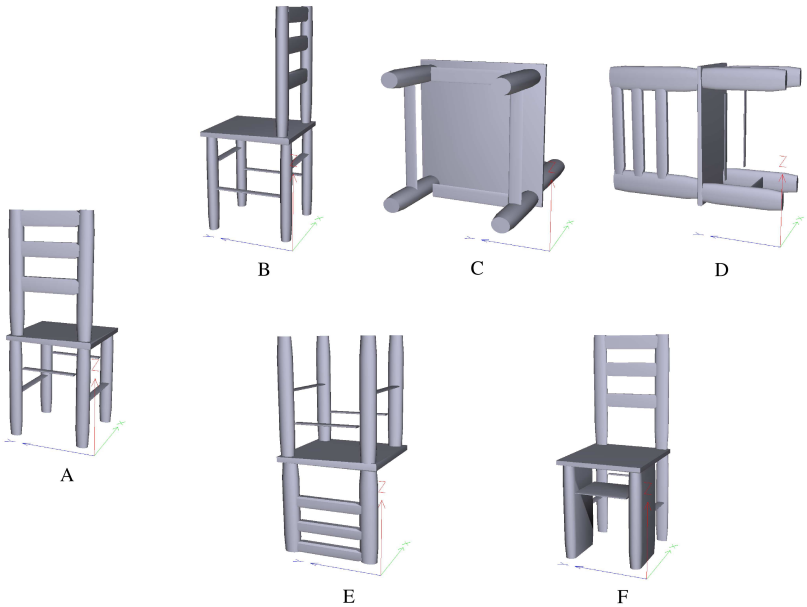



FIG. 3.7 – Différents systèmes de coordonnées obtenus avec la méthode basée sur les moments après déformations d'une chaise.

Cette méthode n'est pas robuste aux petites déformations de l'objet 3D. Pour illustrer ce propos, considérons le modèle de chaise et ses déformations présentés dans la Figure 3.7. Dans un premier temps, la chaise (A) a été transformée en (B), (C) et (D) en effectuant respectivement une élongation de la profondeur par un facteur 1.1, de la largeur par un facteur 1.9 et, de la largeur et de la profondeur par un facteur 2. La méthode basée sur les moments donne quatre ordres différents des axes d'alignement pour les quatre chaises. D'autre part, la

chaise (A) a été transformée en (E) et (F) en changeant les dimensions du dos de la chaise, et en élargissant les pieds avant de la chaise. Il en résulte des sens différents des axes d'alignement.

### 3.4.2 Méthode alternative

Cette méthode alternative est fondée sur une discussion utilisant des hypothèses et des propriétés de la pose naturelle d'un objet 3D (la direction verticale et les deux directions horizontales qu'un homme donne intuitivement). Nous considérons la valeur de $N_{GA}$ et le groupe de symétrie d'un objet, détaillés dans la section 3.2.1. Le but de cette méthode est de proposer parmi l'ensemble des 48 repères possibles un sous-ensemble de repères candidats pour une représentation naturelle de l'objet 3D. Cette méthode permet de réduire considérablement le nombre des représentations possibles, jusqu'à la proposition, dans certains cas, d'un seul repère qui est la solution optimale.

Notre travail sera guidé par l'observation suivante : en étudiant la perception humaine, Ferguson [42] a observé que la symétrie miroir est plus pertinente et mieux perçue quand elle est à la verticale. Cette observation peut être confirmée en examinant les modèles 3D de la base de Princeton et de la base de MPEG7. Par exemple, les modèles de la Figure 3.9 ayant une symétrie de réflexion ont un plan miroir à la verticale.

Ainsi, nous nous basons sur :
les hypothèses suivantes pour fixer les axes d'alignement selon la verticale ou l'horizontale,

- Quand l'objet 3D a une symétrie diédrique, l'axe de symétrie est vertical.
- Quand l'objet 3D a une symétrie cyclique, l'axe de symétrie est vertical.
- Quand l'objet 3D a un plan miroir de normale $\mathbf{n}$ qui n'est pas un axe d'une symétrie diédrique, $\mathbf{n}$ est horizontal.

la propriété de la symétrie cyclique suivante,

- Quand l'objet 3D a une symétrie cyclique d'ordre $n$ multiple de 4, les deux axes d'alignement perpendiculaires à l'axe de symétrie jouent le même rôle.

et la propriété de la symétrie de réflexion pour fixer le sens des axes,

- Quand l'objet 3D a un plan miroir de normale $\mathbf{n}$, le choix entre les deux sens de $\mathbf{n}$ est le même.

Dans la suite, nous aurons besoin d'un processus qui calcule pour un objet 3D l'ordre d'une éventuelle symétrie cyclique autour d'un axe $\mathbf{n}$ donné. Les descripteurs de symétrie calculés sur une distribution de plans passant par $\mathbf{n}$ [64, 68] peuvent être exploités pour détecter ce type de symétrie. Ces aspects ne seront cependant pas abordés ici.

Soient $\mathcal{M}$ l'objet 3D et $R_{ga} = (\mathbf{n_1}; \mathbf{n_2}; \mathbf{n_3})$ le repère donné par l'algorithme 3.1. Considérant $N_{GA}$ le nombre des axes principaux retenus par notre méthode, nous discutons l'ordre et les sens des axes d'alignement. Nous calculons pour chaque cas un ensemble de repères $(\mathbf{x}; \mathbf{y}; \mathbf{z}) = (\pm\mathbf{n_i}; \pm\mathbf{n_j}; \pm\mathbf{n_k})$ avec $i, j, k \in \{1, 2, 3\}$, $\mathbf{z}$ étant la direction verticale et, $\mathbf{x}$ et $\mathbf{y}$ étant les directions horizontales. L'utilisateur sera amené à choisir la solution optimale parmi les repères proposés. Selon la section 3.2, nous avons :

- Si $N_{GA} = 3$, alors $\mathcal{M} \in \mathbf{G_D^{even}} \bigcup \mathbf{G_O}$
  Dans ce cas, $\mathcal{M}$ a trois symétries cycliques autour des trois axes principaux. Le groupe $\mathbf{G_D^{even}}$ présente un cas particulier, quand l'ordre de la symétrie diédrique est égal à 2, où tout axe principal peut être l'axe vertical $\mathbf{z}$ du repère. Pour différencier les trois cas, $\mathbf{G_D^{even}}(n = 2)$, $\mathbf{G_D^{even}}(n > 2)$ et $\mathbf{G_O}$, nous évaluons l'ordre de la symétrie cyclique autour de chaque axe principal :

  - Si l'ordre de la symétrie cyclique autour de chaque axe principal est égal à 2, alors $\mathcal{M} \in \mathbf{G_D^{even}}(n = 2)$. Aucune décision ne peut être prise sur le choix de l'axe vertical. Six repères seront alors proposés en permutant les axes principaux.

  - Si l'ordre de la symétrie cyclique autour d'un des trois axes principaux est supérieur à 2 et l'ordre des symétries cycliques autour de deux autres axes principaux est égal à 2, alors $\mathcal{M} \in \mathbf{G_D^{even}}(n > 2)$. L'axe de la symétrie diédrique d'ordre $(n > 2)$ sera vertical $\mathbf{z}$ du repère. De plus, si $n$ est multiple de 4, alors les deux axes horizontaux $\mathbf{x}$ et $\mathbf{y}$ jouent le même rôle et nous aurons un seul repère (cf. cas (a) de la Figure 3.8), sinon nous aurons deux repères différents à proposer.

  - Si l'ordre de la symétrie cyclique autour de chaque axe principal est égal à 4, alors $\mathcal{M} \in \mathbf{G_O}$. Dans ce cas, toute permutation ou toute inversion des axes principaux aboutira à la même représentation du modèle $\mathcal{M}$. Un seul repère sera proposé à l'utilisateur.

Tous les cas sont détaillés dans le Tableau 3.2.

| $N_{GA}$ | Ordre de la symétrie cyclique autour | | | Classe | n multiple de 4 | Axe vertical | Repères candidats proposés à l'utilisateur |
|---|---|---|---|---|---|---|---|
| | $\mathbf{n_1}$ | $\mathbf{n_2}$ | $\mathbf{n_3}$ | | | | |
| 3 | 2 | 2 | 2 | $\mathbf{G_D^{even}}$ | | | $(\mathbf{n_1};\mathbf{n_2};\mathbf{n_3})$ $(\mathbf{n_2};\mathbf{n_1};\mathbf{n_3})$ $(\mathbf{n_1};\mathbf{n_3};\mathbf{n_2})$ $(\mathbf{n_3};\mathbf{n_1};\mathbf{n_2})$ $(\mathbf{n_2};\mathbf{n_3};\mathbf{n_1})$ $(\mathbf{n_3};\mathbf{n_2};\mathbf{n_1})$ |
| 3 | 2 | 2 | $n > 2$ | $\mathbf{G_D^{even}}$ | non | $\mathbf{n_3}$ | $(\mathbf{n_1};\mathbf{n_2};\mathbf{n_3})$ $(\mathbf{n_2};\mathbf{n_1};\mathbf{n_3})$ |
| 3 | 2 | 2 | $n > 2$ | $\mathbf{G_D^{even}}$ | oui | $\mathbf{n_3}$ | $(\mathbf{n_1};\mathbf{n_2};\mathbf{n_3})$ |
| 3 | $n > 2$ | 2 | 2 | $\mathbf{G_D^{even}}$ | non | $\mathbf{n_1}$ | $(\mathbf{n_2};\mathbf{n_3};\mathbf{n_1})$ $(\mathbf{n_3};\mathbf{n_2};\mathbf{n_1})$ |
| 3 | $n > 2$ | 2 | 2 | $\mathbf{G_D^{even}}$ | oui | $\mathbf{n_1}$ | $(\mathbf{n_2};\mathbf{n_3};\mathbf{n_1})$ |
| 3 | 2 | $n > 2$ | 2 | $\mathbf{G_D^{even}}$ | non | $\mathbf{n_2}$ | $(\mathbf{n_3};\mathbf{n_1};\mathbf{n_2})$ $(\mathbf{n_1};\mathbf{n_3};\mathbf{n_2})$ |
| 3 | 2 | $n > 2$ | 2 | $\mathbf{G_D^{even}}$ | oui | $\mathbf{n_2}$ | $(\mathbf{n_3};\mathbf{n_1};\mathbf{n_2})$ |
| 3 | 4 | 4 | 4 | $\mathbf{G_0}$ | | $\mathbf{n_3}$ | $(\mathbf{n_1};\mathbf{n_2};\mathbf{n_3})$ |

TAB. 3.2 – Les repères candidats pour une représentation naturelle de l'objet 3D quand $N_{GA} = 3$.

– Si $N_{GA} = 2$, alors $\mathcal{M} \in \mathbf{G_C^{even}} \bigcup \mathbf{G_D^{odd}}$.
Dans ce cas, l'ACPC détecte deux plans miroir de normales respectives $\mathbf{n_i}$ et $\mathbf{n_j}$ avec $i, j \in \{1, 2, 3\}$ et $i \neq j$. Si $\mathcal{M} \in \mathbf{G_D^{odd}}$, $\mathcal{M}$ a exactement deux symétries cycliques autour de $\mathbf{n_i}$ et $\mathbf{n_j}$. Par contre, si $\mathcal{M} \in \mathbf{G_C^{even}}$, $\mathcal{M}$ a une seule symétrie cyclique autour du troisième axe principal $\mathbf{n_k}$ avec $k \in \{1, 2, 3\} \backslash \{i, j\}$. Donc, pour différencier les deux cas, $\mathbf{G_C^{even}}$ et $\mathbf{G_D^{odd}}$, il suffit de tester la présence d'une symétrie cyclique autour d'un seul axe principal. Nous choisissons l'axe $\mathbf{n_i}$ :

- S'il existe une symétrie cyclique autour de $\mathbf{n_i}$, alors $\mathcal{M} \in \mathbf{G_D^{odd}}$. Si l'ordre $n$ de cette symétrie est égal à 2, alors $\mathbf{n_j}$ est l'axe de la symétrie diédrique et sera l'axe vertical. Sinon, $\mathbf{n_i}$ est l'axe de la symétrie diédrique et sera l'axe vertical. Dans tous les cas, nous aurons quatre configurations de repères (cf. cas (b) de la Figure 3.8).

- S'il n'existe pas de symétrie cyclique autour de $\mathbf{n_i}$, alors $\mathcal{M} \in \mathbf{G_C^{even}}$ et l'axe de la symétrie cyclique qui est le troisième axe principal (différent de $\mathbf{n_i}$ et de $\mathbf{n_j}$) sera l'axe vertical $\mathbf{z}$. De plus, si l'ordre $n$ de la symétrie cyclique est multiple de 4, alors $\mathbf{n_i}$ et $\mathbf{n_j}$ jouent le même rôle et nous aurons deux repères différents (cf. cas (c) de la Figure 3.8). Sinon, nous aurons quatre repères (cf. cas (d) de la Figure 3.8).

Soulignons que ce test peut être effectué de façon similaire pour l'axe $\mathbf{n_j}$ et en inversant les cas pour $\mathbf{n_k}$. Tous les cas sont détaillés dans le Tableau 3.3.

| $N_{GA}$ | Ordre de la symétrie cyclique autour | | | Classe | n multiple de 4 | Axe vertical | Repères candidats proposés à l'utilisateur |
|---|---|---|---|---|---|---|---|
| | $\mathbf{n_1}$ | $\mathbf{n_2}$ | $\mathbf{n_3}$ | | | | |
| 2 | 2 | $n > 2$ | - | $\mathbf{G_D^{odd}}$ | | $\mathbf{n_2}$ | $(\mathbf{n_3}; \pm\mathbf{n_1}; \mathbf{n_2})$ $(\pm\mathbf{n_1}; \mathbf{n_3}; \mathbf{n_2})$ |
| 2 | 2 | - | $n > 2$ | $\mathbf{G_D^{odd}}$ | | $\mathbf{n_3}$ | $(\mathbf{n_2}; \pm\mathbf{n_1}; \mathbf{n_3})$ $(\pm\mathbf{n_1}; \mathbf{n_2}; \mathbf{n_3})$ |
| 2 | - | 2 | $n > 2$ | $\mathbf{G_D^{odd}}$ | | $\mathbf{n_3}$ | $(\mathbf{n_1}; \pm\mathbf{n_2}; \mathbf{n_3})$ $(\pm\mathbf{n_2}; \mathbf{n_1}; \mathbf{n_3})$ |
| 2 | $n > 2$ | 2 | - | $\mathbf{G_D^{odd}}$ | | $\mathbf{n_1}$ | $(\mathbf{n_3}; \pm\mathbf{n_2}; \mathbf{n_1})$ $(\pm\mathbf{n_2}; \mathbf{n_3}; \mathbf{n_1})$ |
| 2 | $n > 2$ | - | 2 | $\mathbf{G_D^{odd}}$ | | $\mathbf{n_1}$ | $(\mathbf{n_2}; \pm\mathbf{n_3}; \mathbf{n_1})$ $(\pm\mathbf{n_3}; \mathbf{n_2}; \mathbf{n_1})$ |
| 2 | - | $n > 2$ | 2 | $\mathbf{G_D^{odd}}$ | | $\mathbf{n_2}$ | $(\mathbf{n_1}; \pm\mathbf{n_3}; \mathbf{n_2})$ $(\pm\mathbf{n_3}; \mathbf{n_1}; \mathbf{n_2})$ |
| 2 | $n \geq 2$ | - | - | $\mathbf{G_C^{even}}$ | non | $\mathbf{n_1}$ | $(\mathbf{n_2}; \mathbf{n_3}; \pm\mathbf{n_1})$ $(\mathbf{n_3}; \mathbf{n_2}; \pm\mathbf{n_1})$ |
| 2 | $n \geq 2$ | - | - | $\mathbf{G_C^{even}}$ | oui | $\mathbf{n_1}$ | $(\mathbf{n_2}; \mathbf{n_3}; \pm\mathbf{n_1})$ |
| 2 | - | $n \geq 2$ | - | $\mathbf{G_C^{even}}$ | non | $\mathbf{n_2}$ | $(\mathbf{n_3}; \mathbf{n_1}; \pm\mathbf{n_2})$ $(\mathbf{n_1}; \mathbf{n_3}; \pm\mathbf{n_2})$ |
| 2 | - | $n \geq 2$ | - | $\mathbf{G_C^{even}}$ | oui | $\mathbf{n_2}$ | $(\mathbf{n_3}; \mathbf{n_1}; \pm\mathbf{n_2})$ |
| 2 | - | - | $n \geq 2$ | $\mathbf{G_C^{even}}$ | non | $\mathbf{n_3}$ | $(\mathbf{n_1}; \mathbf{n_2}; \pm\mathbf{n_3})$ $(\mathbf{n_2}; \mathbf{n_1}; \pm\mathbf{n_3})$ |
| 2 | - | - | $n \geq 2$ | $\mathbf{G_C^{even}}$ | oui | $\mathbf{n_3}$ | $(\mathbf{n_1}; \mathbf{n_2}; \pm\mathbf{n_3})$ |

TAB. 3.3 – Les repères candidats pour une représentation naturelle de l'objet 3D quand $N_{GA} = 2$.

- Si $N_{GA} = 1$, alors $\mathcal{M} \in \mathbf{G_C^{odd}} \bigcup \mathbf{G_T} \bigcup \mathbf{G_I} \bigcup \mathbf{G_U}$
  Le seul axe d'alignement $\mathbf{n_1}$ détecté par l'ACPC sera horizontal. Si $\mathcal{M} \in \mathbf{G_C^{odd}} \bigcup \mathbf{G_T} \bigcup \mathbf{G_I}$, $\mathcal{M}$ a une seule symétrie cyclique qui est soit autour de $\mathbf{n_2}$ soit autour de $\mathbf{n_3}$. Par contre, si $\mathcal{M} \in \mathbf{G_U}$, $\mathcal{M}$ n'a aucune symétrie cyclique. Pour différencier les deux cas, il est nécessaire de tester la présence des symétries cycliques autour de $\mathbf{n_2}$ et de $\mathbf{n_3}$.
  - S'il existe une symétrie cyclique autour de $\mathbf{n_2}$ ou autour de $\mathbf{n_3}$, alors $\mathcal{M} \in \mathbf{G_C^{odd}} \bigcup \mathbf{G_T} \bigcup \mathbf{G_I}$ et l'axe de la symétrie cyclique sera vertical. L'utilisateur aura à choisir entre huit repères donnés dans le Tableau 3.4.
  - S'il n'existe pas de symétrie cyclique autour de $\mathbf{n_2}$ et de $\mathbf{n_3}$, alors $\mathcal{M} \in \mathbf{G_U}$. L'utilisateur aura à choisir entre seize repères donnés dans le Tableau 3.4.

  En pratique, il est judicieux de conserver les seize représentations possibles proposées dans le cas de $\mathbf{G_U}$ pour tous les groupes. Tester l'existence de la symétrie cyclique autour de $\mathbf{n_2}$ et $\mathbf{n_3}$ conduirait à une complexité trop élevée. D'une part, la majorité des modèles ayant $N_{GA} = 1$ appartiennent à $\mathbf{G_U}$, et risquent de subir le test. D'autre part, différencier l'union $\mathbf{G_C^{odd}} \bigcup \mathbf{G_T} \bigcup \mathbf{G_I}$ de $\mathbf{G_U}$ ne permet pas de réduire considérablement le nombre de repères.
- Si $N_{GA} = 0$, alors $\mathcal{M} \in \mathbf{G_Z}$.
  L'axe d'alignement $\mathbf{n_1}$ qui est la normale du plan associé à la réflexion maximale (cf. section 3.3.1.2) sera horizontal. 32 repères résumés dans le

Tableau 3.4 seront proposés à l'utilisateur.

| $N_{GA}$ | Classe | Axe Horizontal | Axe vertical | Repères candidats proposés à l'utilisateur |
|---|---|---|---|---|
| 1 | $\mathbf{G_C^{odd}} \bigcup \mathbf{G_T} \bigcup \mathbf{G_I}$ | $\mathbf{n_1}$ | $\mathbf{n_2}$ | $(\mathbf{n_1}; \pm\mathbf{n_3}; \pm\mathbf{n_2})$ $(\pm\mathbf{n_3}; \mathbf{n_1}; \pm\mathbf{n_2})$ |
| 1 | $\mathbf{G_C^{odd}} \bigcup \mathbf{G_T} \bigcup \mathbf{G_I}$ | $\mathbf{n_1}$ | $\mathbf{n_3}$ | $(\mathbf{n_1}; \pm\mathbf{n_2}; \pm\mathbf{n_3})$ $(\pm\mathbf{n_2}; \mathbf{n_1}; \pm\mathbf{n_3})$ |
| 1 | $\mathbf{G_U}$ | $\mathbf{n_1}$ | $\mathbf{n_2}$ or $\mathbf{n_3}$ | $(\mathbf{n_1}; \pm\mathbf{n_3}; \pm\mathbf{n_2})$ $(\pm\mathbf{n_3}; \mathbf{n_1}; \pm\mathbf{n_2})$ $(\mathbf{n_1}; \pm\mathbf{n_2}; \pm\mathbf{n_3})$ $(\pm\mathbf{n_2}; \mathbf{n_1}; \pm\mathbf{n_3})$ |
| 0 | $\mathbf{G_Z}$ | $\mathbf{n_1}$ | $\mathbf{n_2}$ or $\mathbf{n_3}$ | $(\pm\mathbf{n_1}; \pm\mathbf{n_3}; \pm\mathbf{n_2})$ $(\pm\mathbf{n_3}; \pm\mathbf{n_1}; \pm\mathbf{n_2})$ $(\pm\mathbf{n_1}; \pm\mathbf{n_2}; \pm\mathbf{n_3})$ $(\pm\mathbf{n_2}; \pm\mathbf{n_1}; \pm\mathbf{n_3})$ |

TAB. 3.4 – Les repères candidats pour une représentation naturelle de l'objet 3D quand $N_{GA} \leq 1$.

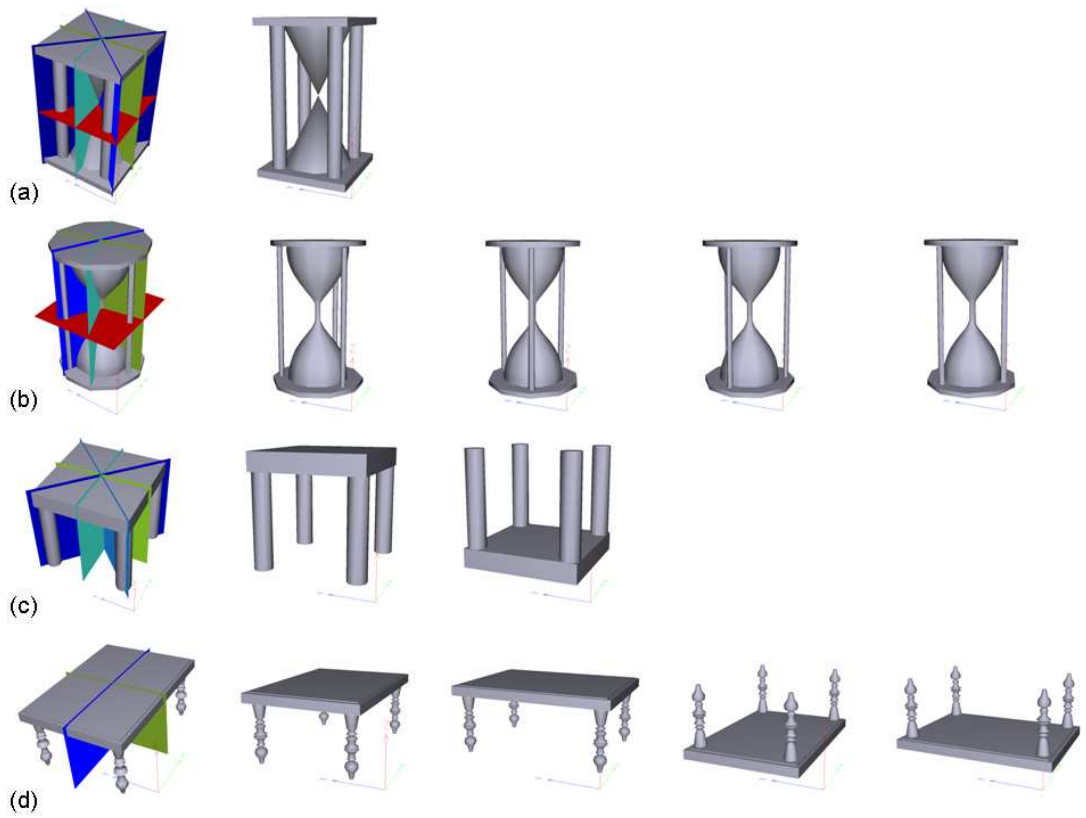



FIG. 3.8 – Les repères candidats pour une représentation naturelle de l'objet 3D : (a) $N_{GA} = 3$, $\mathcal{M} \in \mathbf{G_D^{even}}$ avec $n = 4$, (b) $N_{GA} = 2$, $\mathcal{M} \in \mathbf{G_C^{odd}}$ avec $n = 3$, (c) $N_{GA} = 2$, $\mathcal{M} \in \mathbf{G_C^{even}}$ avec $n = 4$, (d) $N_{GA} = 2$, $\mathcal{M} \in \mathbf{G_C^{even}}$ avec $n = 2$.

Un outil interactif qui propose à l'utilisateur un ensemble de repères candidats sur la base de cette discussion est en cours de développement dans le cadre du stage de master de Skander El Fekih. Dans le contexte de la recherche par le

contenu, cette méthode peut être utilisée pour réduire le nombre de cas à envisager lors du calcul de la similarité entre deux objets 3D. Plus précisément, en considérant $\mathbf{F}$ une signature décrivant un objet 3D dans un repère donné et $d$ une distance associée, nous proposons de comparer deux objets $O^1$ et $O^2$ comme suit : d'abord, générer, pour chacun, les représentations de la signature correspondant seulement à l'ensemble des repères candidats, $(\mathbf{F}_i^1)_{1\leq i\leq N_1}$ et $(\mathbf{F}_i^2)_{1\leq i\leq N_2}$, ensuite, calculer la similarité entre $O^1$ et $O^2$ selon la formule suivante :

$$\Delta(O^1, O^2) = \begin{cases} \min_{1\leq i\leq N_1} d(\mathbf{F}_i^1, \mathbf{F}_1^2) & \text{si } N_1 > N_2, \\ \min_{1\leq i\leq N_2} d(\mathbf{F}_1^1, \mathbf{F}_i^2) & \text{sinon,} \end{cases} \tag{3.10}$$

$N_1$ et $N_2$ étant respectivement le nombre de repères considérés de $O^1$ et $O^2$.
Il est intéressant de remarquer que la mesure de similarité ainsi définie est donnée en calculant $\max(N_1, N_2)$ distances entre représentations, contrairement à la plupart des méthodes optimisées de la littérature qui consistent à calculer 48 distances correspondant aux 48 repères possibles.

## 3.5 Résultats expérimentaux

Cette section présente les résultats de notre méthode sur la base de Test de Princeton. Rappelons que notre algorithme d'alignement calcule seulement les trois directions optimales qui sont données dans un ordre et des sens arbitraires. Pour plus de lisibilité dans les Figures 3.9, 3.10, 3.11 et 3.12, nous avons choisi, pour chacun des modèles présentés, le sens et l'ordre les plus naturels.

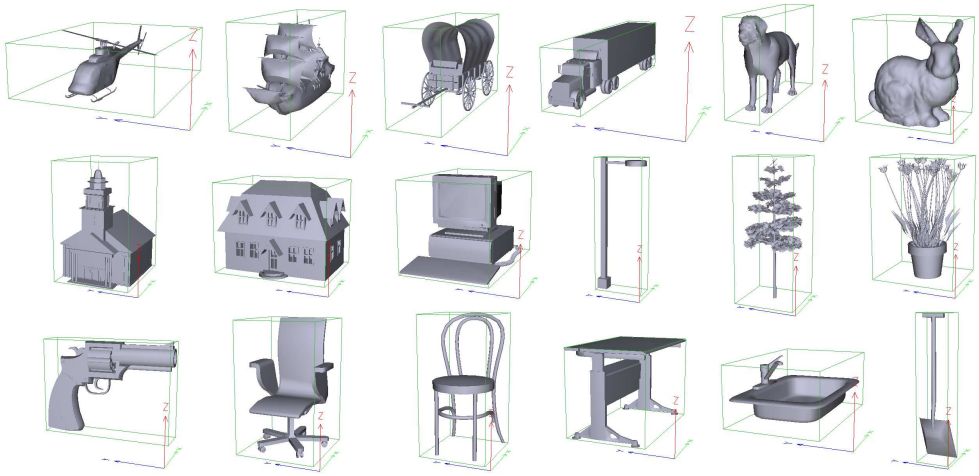


FIG. 3.9 – Alignements de modèles 3D utilisant notre méthode.

Nous constatons que notre algorithme retrouve les directions pertinentes pour la plupart des modèles 3D. La Figure 3.9 présente quelques modèles de la base de Princeton alignés avec notre méthode. Comparée à l'ACPC, les résultats montrent, pour notre alignement, une amélioration de la précision (cf. Figure 3.10), et un alignement plus stable à l'intérieur de chaque classe. Comme le montrent les Figures 3.11 et 3.12, notre approche fournit des alignements similaires pour des modèles appartenant à la classe "boîte à lettre" et à la classe "chaise de cuisine".

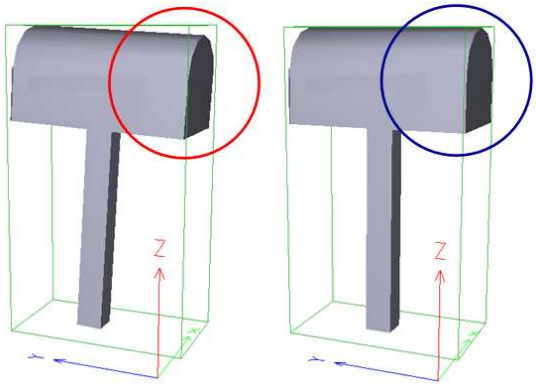



FIG. 3.10 – Comparaison de la précison de l'alignement d'une boîte à lettre : l'ACPC (à gauche) et notre méthode (à droite).

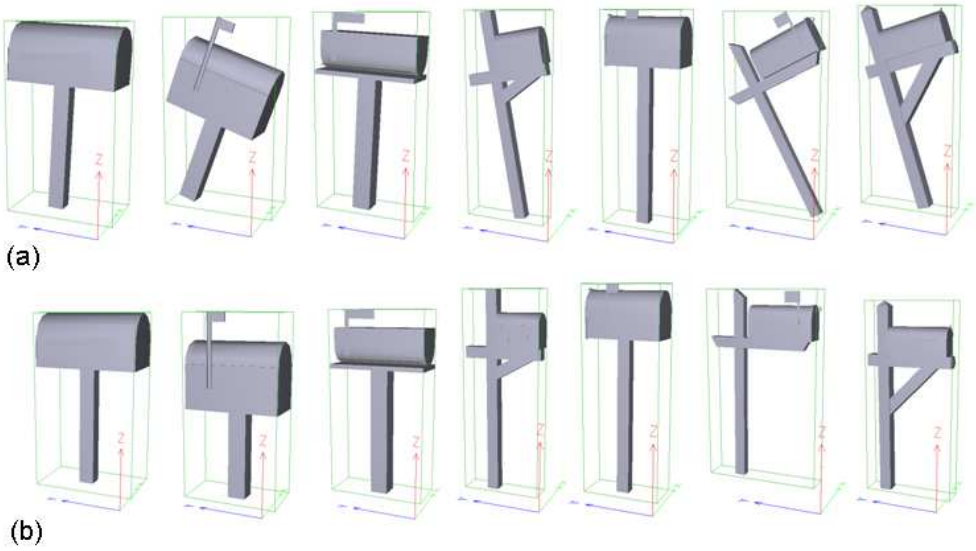



FIG. 3.11 – Alignement des modèles de la classe "boîte à lettre" utilisant l'ACPC (a) et notre méthode (b).

FIG. 3.12 – Alignement des modèles de la classe “chaise de cuisine” utilisant l'ACPC (a) et notre méthode (b).

Le Tableau 3.5 résume les tests effectués sur 33 classes différentes, généralement mal alignées avec l'ACPC. Il donne, pour chaque classe, le pourcentage d'objets parfaitement alignés (c'est-à-dire similaire à celui retourné par le système visuel humain) et compare les résultats de l'ACPC à ceux de notre méthode pour les trois descripteurs $G$, $E$ et $F$ décrits dans la section 3.3.2.4. Pour calculer les pourcentages, nous avons demandé à des experts de modélisation 3D de regrouper les modèles de chaque classe en deux sous-classes : “bien aligné” ou “mal aligné”. Par exemple, tous les modèles de la Figure 3.9 (qui font partie de l'ensemble de test) ont été considérés comme des modèles “bien alignés”. Les pourcentages d'objets parfaitement alignés figurant dans le Tableau 3.5 indiquent la moyenne des modèles “bien alignés” à l'intérieur de chaque classe.

Notons qu'avec les trois descripteurs ($G$, $E$ et $F$), notre méthode est beaucoup plus précise que l'ACPC, et permet d'atteindre des pourcentages d'alignement proches de 100%. Nous constatons que le descripteur de forme $F$ obtient des résultats légèrement meilleurs que les descripteurs $G$ et $E$. Cette différence s'explique simplement par le fait que le descripteur $F$ caractérise la symétrie locale de translation d'un point de vue plus local.

| **Classe** | ♯ ($\mathcal{M}$) | **ACPC** | **Notre méthode** | | |
|---|---|---|---|---|---|
| | | | $G$ | $E$ | $F$ |
| Helicopter aircraft | 18 | 77.7% | 94.4% | 100% | 100% |
| Enterprise spaceship | 11 | 36.4% | 100% | 100% | 100% |
| Dog quadruped | 7 | 00.0% | 14.3% | 28.6% | 85.7% |
| Horse quadruped | 6 | 16.7% | 66.7% | 66.7% | 83.3% |
| Rabbit quadruped | 4 | 00.0% | 25.0% | 75.0% | 75.0% |
| Head body part | 16 | 62.5% | 56.2% | 81.2% | 100% |
| Skull body part | 6 | 00.0% | 16.7% | 16.7% | 100% |
| Barn building | 5 | 40.0% | 80.0% | 80.0% | 80.0% |
| Church building | 4 | 00.0% | 100% | 100% | 75.0% |
| One story building | 14 | 35.7% | 85.7% | 92.9% | 92.9% |
| Two story building | 10 | 10.0% | 80.0% | 100% | 100% |
| Chess set | 9 | 66.7% | 100% | 100% | 100% |
| Desktop computer | 11 | 00.0% | 63.6% | 81.8% | 81.8% |
| Computer monitor | 13 | 00.0% | 92.3% | 92.3% | 100% |
| Fireplace | 6 | 00.0% | 83.3% | 83.3% | 83.3% |
| Cabinet furniture | 9 | 66.7% | 100% | 100% | 100% |
| School desk furniture | 4 | 00.0% | 100% | 100% | 100% |
| Bench seat | 11 | 00.0% | 100% | 100% | 100% |
| Dining chair | 11 | 00.0% | 100% | 100% | 100% |
| Desk chair seat | 15 | 00.0% | 100% | 100% | 100% |
| Rectangular table | 25 | 72.0% | 100% | 100% | 100% |
| Handgun gun | 10 | 00.0% | 80% | 90% | 100% |
| Ladder | 4 | 50.0% | 100% | 100% | 100% |
| Streetlight lamp | 8 | 75.0% | 100% | 100% | 100% |
| Mailbox | 7 | 14.3% | 100% | 100% | 100% |
| Potted plant | 26 | 53.8% | 92.3% | 88.5% | 100% |
| Conical tree | 10 | 70.0% | 90.0% | 80.0% | 90.0% |
| Large sail boat | 6 | 00.0% | 50.0% | 100% | 100% |
| Sink | 4 | 25.0% | 75.0% | 100% | 100% |
| Slot machine | 4 | 25.0% | 75.0% | 50.0% | 100% |
| Covered wagon vehicle | 5 | 00.0% | 60.0% | 60.0% | 100% |
| Semi vehicle | 7 | 14.3% | 57.1% | 100% | 100% |
| Train car | 5 | 40.0% | 100% | 100% | 100% |

TAB. 3.5 – Comparaison des performances d'alignement de l'ACPC et de notre méthode avec les trois descripteurs $G$, $E$ et $F$, en termes de pourcentages d'objets parfaitement alignés pour 33 classes différentes (311 modèles).

La méthode d'alignement proposée est robuste vis-à-vis des petites déformations de l'objet 3D, plus précisément vis-à-vis des transformations anisotropiques, qu'elles soient globales ou locales. La chaise présentée à droite de la Figure 3.13 a été alignée avec notre méthode, puis a subi un ensemble de déformations suivant les directions d'alignement.

- élongation de la hauteur par un facteur 0.6,
- rétrécissement de la hauteur du dos de la chaise,
- élargissement des pieds avant de la chaise,
- élongation de la largeur par un facteur 1.9,
- élongation de la largeur et de la profondeur par un facteur 2,

donnant ainsi cinq nouveaux objets 3D.

La Figure 3.13 montre les résultats d'alignement des différents objets créés. Dans tous les cas, notre approche fournit des alignements pertinents et retourne les mêmes directions que l'objet initial. Ces résultats soulignent l'intérêt de l'utilisation des symétries qui sont invariantes par rapport à ce type de déformations.

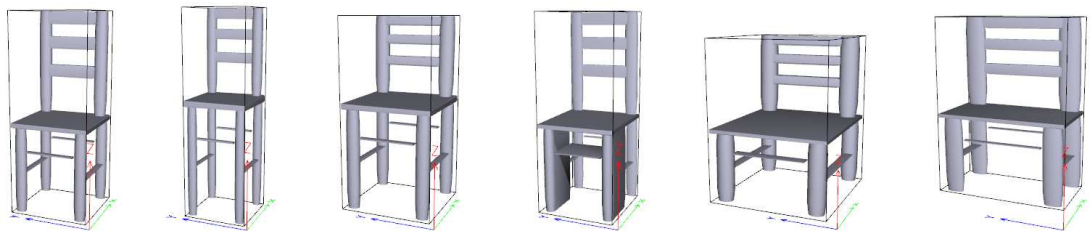


FIG. 3.13 – Robustesse de la méthode d'alignement face aux transformations anisotropiques globales et locales.

Dans le cadre de la recherche par le contenu d'objets 3D, notre méthode d'alignement (AL) est intégrée à la phase de normalisation. Pour comparer les performances de recherche, avec les deux méthodes d'alignement (AL et ACPC), une série d'expérimentations ont été réalisées sur les approches d'indexation proposées dans le chapitre 5. Les Figures 6.6, 6.8, 6.10, 6.14, 6.17, 6.20 et les Tableaux 6.6, 6.8, 6.10, 6.14, 6.17, 6.20 montrent les résultats, en termes de courbes précision-rappel et mesures de performance, NN, FT, ST et DCG (cf. section 6.1). Dans tous les cas, notre méthode d'alignement se révèle plus performante que celle fondée sur l'ACPC. Notons que notre méthode parvient à mieux aligner des objets d'une même classe donnant ainsi une similarité plus pertinente et invariante aux rotations.

## Complexité :

L'utilisation de l'ACPC permet à notre méthode d'alignement d'atteindre des temps de calculs nettement inférieurs aux méthodes d'alignement basées sur la caractérisation de la symétrie dans l'espace des rotations 3D. Rappelons que la complexité de l'ACPC est en $\mathcal{O}(N_T)$, $N_T$ étant le nombre des facettes du maillage. Comme le montre le Tableau 3.6, les modèles ayant $N_{GA} = 2^+$ représentent en moyenne entre 25 et 28.5% des objets de la base. Dans ces cas, le coût d'alignement est faible car les trois axes optimaux sont obtenus via l'ACPC. Le temps de calculs reste compétitif pour les modèles ayant $N_{GA} = 1$ qui constituent plus

de la moitié des modèles dans les différentes bases d'objets (en moyenne entre 50 et 58%). La complexité comprend celle de l'ACPC qui retourne le premier axe d'alignement et celle de l'algorithme 3.2 qui calcule les autres axes d'alignement. Cette complexité augmente pour les modèles ayant $N_{GA} = 0$ car le calcul du plan associé à la réflexion maximale est l'étape la plus coûteuse.

| Base d'objets 3D | $N_{GA}$ | | |
|---|---|---|---|
| | 0 | 1 | $2^+$ |
| psbTest | 181 | 467 | 259 |
| | 20.0% | 51.5% | 28.5% |
| psbTrain | 190 | 462 | 255 |
| | 21.0% | 50.9% | 28.1% |
| mpeg7 | 56 | 114 | 57 |
| | 24.7% | 50.2% | 25.1% |
| shrec09 | 116 | 463 | 221 |
| | 14.5% | 57.9% | 27.6% |

TAB. 3.6 – Répartition des modèles 3D selon les valeurs de $N_{GA}$ pour les quatre bases d'objets 3D.

## 3.6 Conclusion

Dans ce chapitre, nous avons présenté une nouvelle méthode d'alignement d'objets 3D. L'idée générale est motivée par une observation de la perception humaine, le fait qu'un utilisateur cherche généralement à aligner un objet 3D selon des axes particuliers, caractérisant des symétries intrinsèques à l'objet. Pour ce faire, deux types de symétrie ont été considérés, la symétrie de réflexion et la symétrie locale de translation. Notre algorithme d'alignement repose sur quatre étapes successives :

1. le centrage et la mise à l'échelle de l'objet 3D,
2. le test de la réflexion par rapport aux trois plans principaux de l'objet 3D,
3. la sélection des directions d'alignement qui caractérisent des symétries de réflexion par rapport aux plans principaux et, le cas échéant, la sélection de la direction qui correspond à la réflexion maximale,
4. l'extraction des directions naturelles en utilisant une nouvelle mesure de symétrie locale de translation pour les objets dont l'alignement ne pourrait être obtenu uniquement avec des plans de symétrie miroir.

Les considérations de symétries miroirs par rapport aux plans principaux -qui sont faites sur la fonction $N_{GA}$- s'avèrent particulièrement précieuses dans notre algorithme d'alignement car elles permettent de simplifier considérablement le problème et d'obtenir des solutions de manière rapide.

Nous avons vu dans la partie expérimentale que cette méthode produit des alignements naturels et intuitifs pour la perception humaine. Elle est rapide et efficace pour aligner des objets 3D de forme similaire. Par ailleurs, nous avons montré que cette solution apporte des résultats qualitatifs et quantitatifs plus précis que la méthode à base d'une analyse en composantes principales. Pour valider notre méthode d'alignement dans le cadre de l'indexation 3D, nous l'avons intégré dans la phase de normalisation. Elle a été ensuite évaluée et comparée à la méthode fondée sur l'analyse en composantes principales. Dans tous les processus d'indexation proposés dans cette thèse, notre méthode d'alignement se révèle plus performante en termes de courbes précision-rappel et mesures de performance. Ceci confirme que notre méthode parvient à mieux aligner des objets à l'intérieur de chaque classe donnant ainsi une description plus pertinente et invariante aux rotations.

# Troisième partie

# Recherche par le contenu d'objets 3D

# Chapitre 4

# Etat de l'art

La troisième partie de cette thèse est consacrée à la recherche par le contenu d'objets 3D. Ce chapitre dresse un état de l'art des différentes approches proposées dans la littérature. Nous nous intéressons plus particulièrement à la caractérisation de la similarité entre deux objets 3D qui est la tâche fondamentale de la recherche par le contenu. Elle consiste à définir une ou plusieurs signatures de formes et les mesures de similarité appropriées de façon à trouver rapidement les $k$ premiers objets les plus proches d'un objet requête.

L'analyse de la littérature montre qu'il existe une variété de descripteurs de forme 3D qui dérivent des différents aspects des objets 3D. Certains d'entre eux reposent directement sur les propriétés géométriques et topologiques de l'objet 3D, alors que d'autres se fondent sur les projections 2D de l'objet 3D. Nous avons fait le choix, pour notre étude bibliographique, de regrouper les descripteurs de forme en deux familles d'approches :
- **les Approches 3D**,
Plusieurs grandes voies se distinguent dans les différentes approches 3D proposées. Nous pouvons les séparer en quatre groupes : les approches statistiques, les approches basées sur une partition de l'espace, les approches par transformées, et les approches structurales.
- **les Approches 2D/3D**.
Différents types d'images peuvent être extraites d'un objet 3D pour le caractériser : les silhouettes, les images de profondeur, les coupes et les images de spin. Les deux premières sont les plus répandues dans les approches 2D/3D et sont déduites des projections de l'objet 3D sous différents angles de vues. C'est pourquoi nous les appellerons dans la suite les approches multi-vues.

Nous ne pourrons pas citer exhaustivement toutes les approches évoquées dans la littérature. Nous invitons le lecteur à se référer au récents livre [40] et état de l'art [119], ainsi qu'aux études comparatives entre systèmes de recherche par le contenu, reportées dans [23, 112, 137, 138]. Ces études consistent à évaluer et comparer les performances des méthodes existantes sur une base d'objets 3D généralement classé (base de MPEG7 [137, 138], base de PSB [112] et la base Konstanz [23]). Le lecteur peut aussi consulter les résultats du concours annuel organisé dans le cadre du réseau d'excellence AIM@SHAPE, appelé *3D SHape REtrieval Contests* (SHREC) ([90, 125] pour des évaluations réalisées sur des bases de modèles généralistes, et [15, 53, 81, 120, 126] pour des évaluations faites sur des bases plus spécifiques telles que les bases de modèles articulés, de visages et de CAO).

La combinaison de plusieurs descripteurs de forme augmente généralement la robustesse de la recherche par le contenu donnant ainsi des résultats plus satisfaisants. Des solutions proposées dans [5, 20, 22, 131] permettent de combiner en un descripteur hybride, différents aspects de caractérisation de la forme. Cet état de l'art ne traite pas ce type d'approches.

Dans cette thèse, nous nous intéressons à l'approche globale de recherche par le contenu : à partir d'un objet 3D exemple sélectionné par l'utilisateur, le système retourne les objets de la base qui lui sont visuellement proches. Ici, la similarité est mesurée sur les objets entiers. La recherche globale est alors "approximative" car elle prend en compte toutes les parties de l'objet, y compris celles qui ne sont pas pertinentes pour la recherche. Au cours des cinq dernières années, différentes approches [48, 50, 52, 84] de requêtes partielles ont été proposées pour identifier et comparer des parties d'objets dans un contexte d'indexation. Ces approches permettent à l'utilisateur de spécifier l'objet ou la partie qui l'intéresse dans un objet 3D et de retrouver des parties similaires dans d'autres objets de la base. Ce mode de recherche permet à l'utilisateur d'exprimer plus précisément ses préférences. La segmentation de l'objet 3D en parties est la tâche la plus complexe et la plus coûteuse de ce type d'approches. La recherche par partie ne sera pas abordée dans cet état de l'art.

## 4.1 Approches 3D

Notre étude bibliographique a fait ressortir deux grandes familles d'approches. Dans cette section, nous nous intéressons à la première famille qui groupe les approches 3D cherchant à extraire la signature de l'objet directement à partir de l'objet 3D. Nous présentons ces approches selon quatre classes :

- les approches statistiques,
- les approches basées sur une partition de l'espace,
- les approches par transformées,
- les approches structurales.

### 4.1.1 Approches statistiques

L'objectif de ces approches est de caractériser un objet 3D par une ou plusieurs distributions stockées sous forme d'histogrammes. Généralement, un descripteur statistique consiste à estimer la distribution de la mesure d'une primitive géométrique donnée (cordes, segments, triangles, tétraèdres) ou d'une caractéristique mathématique donnée (courbure, orientation). Notons que ces mesures sont calculées directement sur la surface de l'objet 3D.

#### 4.1.1.1 Distributions à base de cordes

Paquet et Rioux [97] sont les premiers à avoir proposé une méthode d'indexation d'objets 3D qui repose sur des statistiques construites à partir des segments (ou cordes) reliant les centres de gravité des triangles du maillage et le centre de gravité de l'objet 3D. Le descripteur de forme est défini par trois distributions : une distribution des longueurs des cordes et, deux distributions des angles entre les cordes et les deux premiers axes principaux de l'objet 3D. Une importante limitation de cette méthode réside dans sa dépendance à la connectivité et à la granularité du maillage. Par contre, elle a l'avantage d'être extrêmement rapide et assez discriminante pour les objets uniformément échantillonnés.

#### 4.1.1.2 Distributions de formes

Osada et al. [95, 96] proposent cinq distributions de forme pour indexer un objet 3D. Les mesures considérées sont la distance entre le centre de gravité et un point de la surface, la distance entre deux points (cf. Figure 4.1), l'angle entre trois

points, la racine carrée de l'aire du triangle formé par trois points et la racine cubique du volume du tétraèdre formé par quatre points. Tous les points considérés sont pris au hasard sur des faces triangulaires de l'objet, les faces triangulaires étant elles-mêmes prises au hasard. Pour obtenir une invariance par rapport à l'échelle, les auteurs proposent de normaliser les distributions de formes.

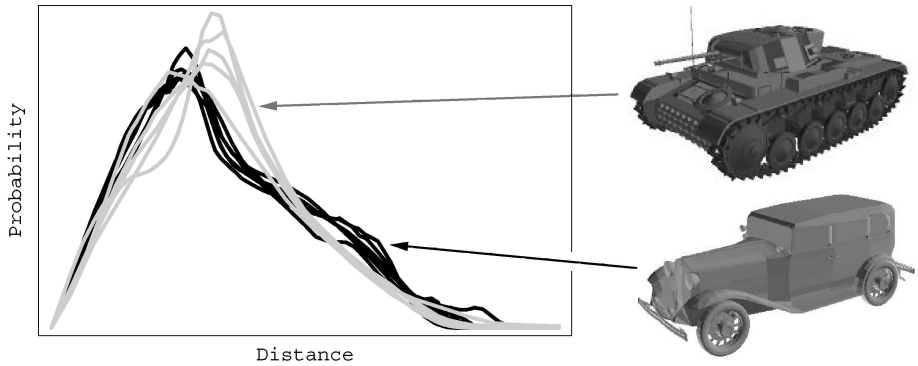



FIG. 4.1 – Distributions de forme de différents véhicules : cinq chars (courbes grises) et six voitures (courbes noires) [96]

Les principaux avantages de cette approche sont la compacité du descripteur, la rapidité de calcul, l'invariance aux transformations géométriques et la robustesse aux bruits. Toutefois, trop élémentaire pour caractériser la forme, son utilisation semble plus adaptée aux recherches d'objets de forme globalement similaires sans exigence trop forte de précision.
Ip et al. [60] ont enrichi la distribution de la distance entre deux points pris au hasard en proposant trois nouvelles distributions dérivées. Ils classent pour cela les paires de points selon trois critères : (*in*) le segment reliant les points se trouve complètement à l'intérieur de l'objet, (*out*) complètement à l'extérieur de l'objet, (*mixed*) à l'intérieur et à l'extérieur de l'objet. La comparaison des objets se fait en calculant une somme pondérée des distances entre les quatre histogrammes. Cette méthode exige une définition correcte du maillage due au critère de classification utilisé dans la construction des histogrammes.
Ohbuchi et al. [91] proposent une autre extension appelée *Absolute Angle-Distance histogram*. Cette nouvelle distribution est paramétrée par la distance entre deux points pris au hasard et l'angle entre les facettes auxquelles appartiennent les points. Des expériences faites par les auteurs montrent qu'avec une légère augmentation des coûts de calcul, leur distribution dépasse la distribution de forme classique en termes de performance.

#### 4.1.1.3 Images gaussiennes étendues

Les images gaussiennes étendues connues sous le nom des EGI (*Extended Gaussian Images*) est une fonction sphérique définie sur la sphère de Gauss (sphère unité), caractérisant l'information d'orientation des points d'une surface 3D. La description d'une forme 3D par les EGI a été définie par Horn [58] et utilisée par Kazhdan et al. [112] et récemment par Wang et al. [132] dans la recherche par le contenu d'objets 3D. Le calcul du descripteur consiste à construire un histogramme défini sur un ensemble discret d'orientations couvrant la sphère unité. Chaque composante de l'histogramme est donnée par l'aire des facettes ayant des orientations appartenant à l'intervalle angulaire associé. Une importante limitation de cette représentation réside dans sa grande dépendance à l'information d'orientation. Par exemple, les EGI d'une pyramide en escalier et une pyramide lisse sont complètement différentes. Pour pallier cette limitation, plusieurs extensions des EGI peuvent être utilisées. La plus connue est les CEGI (*Complex EGI*) [63] qui apportent l'information de distance des facettes à l'origine du repère. Dans ce cas, les composantes habituelles des EGI sont incrémentées d'un nombre complexe dont l'amplitude est l'aire de la facette et la phase est sa distance au centre de la sphère.
Les EGI et ses différentes versions ne sont pas robustes vis-à-vis des multiples représentations topologiques (résolutions différentes) d'un même objet. De plus, elles ne sont pas invariantes aux transformations géométriques. Cela implique que les méthodes de recherche à base d'EGI doivent intégrer une étape de normalisation ou un mécanisme de mise en correspondance dans la mesure de similarité.

#### 4.1.1.4 Distributions de spectre de forme 3D

Le descripteur de spectre de forme 3D (SF3D) proposé par Zaharia et Prêteux [135, 139] a été retenu comme descripteur normalisé MPEG-7. Il est défini comme la distribution de l'index de forme de Koenderink [70] caractérisant localement les courbures de la surface de l'objet 3D (cf. Figure 4.2).
Soient $\mathbf{p}$ un point sur une surface régulière et $k^1_{\mathbf{p}}$ et $k^2_{\mathbf{p}}$ avec $k^1_{\mathbf{p}} \geq k^2_{\mathbf{p}}$ les courbures principales au point $\mathbf{p}$. L'indice de forme $I_{\mathbf{p}}$ s'exprime par :

$$I_{\mathbf{p}} = \frac{1}{2} - \frac{1}{\pi} \arctan \frac{k^1_{\mathbf{p}} + k^2_{\mathbf{p}}}{k^1_{\mathbf{p}} - k^2_{\mathbf{p}}}. \tag{4.1}$$

$I_{\mathbf{p}}$ n'est pas défini sur les surfaces planes, et son estimation sur les facettes de bord n'est pas fiable en raison du nombre réduit des facettes adjacentes.

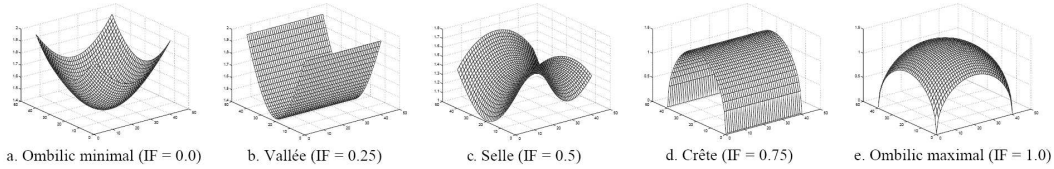



FIG. 4.2 – Exemples d'indices de forme (IF) calculés sur cinq formes élémentaires [139].

Pour construire le descripteur de spectre de forme 3D, L'intervalle $[0, 1]$ définissant le domaine des valeurs de l'index de forme est subdivisé uniformément en $N$ intervalles $\Delta_k$, $k \in \{1, ..., N\}$. Le SF3D est représenté comme un histogramme à $N$ composantes, la composante $k$ cumulant l'aire normalisée relative de toutes les facettes ayant un index de forme appartenant à $\Delta_k$. Deux composantes supplémentaires sont ajoutées au descripteur, la première quantifie l'aire normalisée des surfaces planes et la deuxième quantifie l'aire normalisée des facettes de bord. Intrinsèquement invariant aux transformations géométriques, le SF3D n'est en revanche pas robuste aux changements topologiques (représentations différentes, simplification, remaillage, etc.). De plus, cette méthode requiert des maillages réguliers avec des normales bien orientées.

#### 4.1.1.5 Carte de courbures

J. Assfalg et al. [11] introduisent une méthode basée sur l'idée que la forme d'un objet 3D peut être décrite par une carte de courbures de sa surface. Après une phase de prétraitement durant laquelle les propriétés différentielles de la surface l'objet 3D sont calculées, la surface de l'objet 3D est déformée pour épouser la surface d'une sphère, tout en gardant les informations de ses courbures. La projection du maillage sur la sphère avec la donnée de courbure associée aux sommets forment la carte de courbure de l'objet 3D. La similarité entre deux objets est calculée en comparant leurs cartes de courbures via des descripteurs qui tiennent compte de la surface des régions et de leur arrangement spatial. Les descripteurs sont des histogrammes de courbures définis sur trois niveaux de détail, calculés du plus fin au plus grossier.

### 4.1.2 Approches basées sur une partition de l'espace

L'objectif de ces méthodes est de caractériser l'objet 3D dans son espace. La première étape consiste à partitionner l'espace 3D englobant l'objet en cel-

lules. Cette décomposition résulte soit d'un échantillonnage uniforme d'un seul paramètre (axial, radial, angulaire), soit d'un échantillonnage uniforme de l'espace des paramètres, le plus souvent cartésien $(x, y, z)$ ou sphérique $(r, \mathbf{u})$. Sur chaque cellule, un descripteur de forme locale est calculé, et l'ensemble constitue la signature de l'objet 3D.

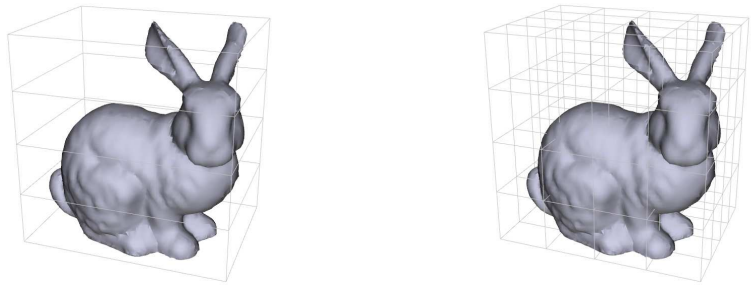


FIG. 4.3 – Décomposition de l'espace englobant l'objet 3D : à gauche selon le paramètre $z$, à droite selon les paramètres cartésiens $(x, y, z)$.

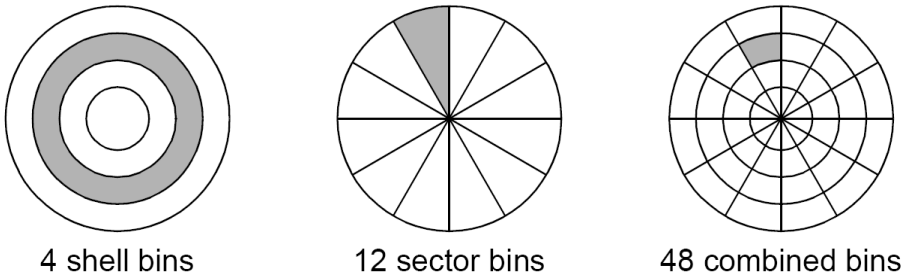



FIG. 4.4 – Trois décompositions classiques de l'espace 2D : coquilles, secteurs et combinés [6].

#### 4.1.2.1 Distribution spatiale

L'idée présentée par Ankerst et al. [6] consiste à prélever un ensemble de points uniformément répartis sur la surface de l'objet 3D, et à caractériser la forme par leur distribution dans l'espace 3D. Pour ce faire, les auteurs proposent de partitionner l'espace en $N$ cellules selon trois types de décomposition (cf. Figure 4.4). La décomposition angulaire est donnée par les faces d'un polyèdre régulier rétroprojecté sur la sphère unité. Cette méthode permet d'obtenir des partitions de taille homogène contrairement à la décomposition classique selon les paramètres $(\theta, \varphi)$. Le descripteur de forme est défini par un histogramme de taille $N$, chaque

composante étant égale au nombre de points d'échantillonnage appartenant à la cellule correspondante. Durant les phases de recherche, la similarité entre objets est calculée par une fonction quadratique de distance entre histogrammes :

$$d_A(\mathbf{h}, \mathbf{h}') = (\mathbf{h} - \mathbf{h}')\mathbf{A}(\mathbf{h} - \mathbf{h}')^T = \sum_{i=1}^{N}\sum_{j=1}^{N} a_{ij}(h_i - h'_i)(h_j - h'_j), \tag{4.2}$$

$a_{ij} = e^{-\sigma d(i,j)}$ étant le poids de la similarité avec $\sigma \in [1, 10]$. Cette méthode s'avère être performante pour retrouver des formes similaires dans une base de molécules 3D.

#### 4.1.2.2 Partition de l'espace selon les axes principaux

Ohbuchi et al. [94] proposent de partitionner régulièrement l'espace de l'objet 3D en $N$ cellules de taille égale le long d'un axe principal, et d'attribuer trois valeurs statistiques pour chaque cellule : le moment d'inertie autour de l'axe, la distance moyenne de la surface à l'axe, et la variance de la distance de la surface à l'axe. Le descripteur de forme est défini par neuf histogrammes, trois histogrammes normalisés sur chacun des trois axes principaux (cf. Figure 4.5). La mesure de similarité proposée est invariante aux changements de sens des axes. Deux distances ont été utilisées : la distance euclidienne et une distance élastique basée sur la programmation dynamique.

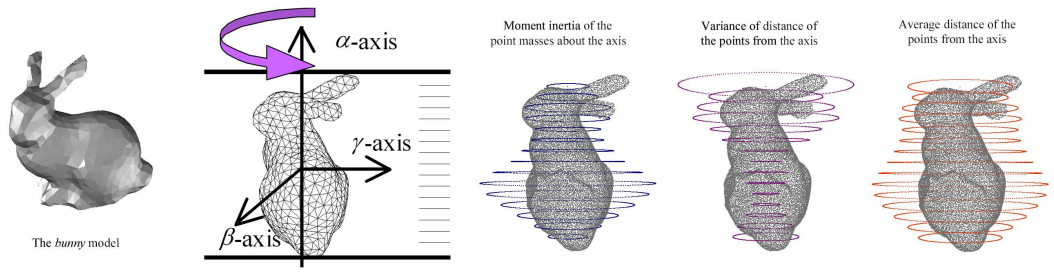



FIG. 4.5 – Exemples de trois histogrammes calculés autour d'un axe principal $N = 15$ [94].

La méthode proposée peut être vue comme une méthode statistique paramétrée par les axes principaux, dont les principaux avantages sont la rapidité de calculs, l'invariance aux transformations géométriques et la robustesse aux déplacements et étirements au sein de la forme (grâce à la programmation dynamique). La

troisième mesure statistique, la variance, semble être inutile car elle peut être calculée à partir des deux autres mesures. Comme pour d'autres descripteurs, cette méthode est dépendante de l'ACP, sur laquelle sont construites les paramétrisations. Elle pourrait être améliorée en utilisant des axes d'alignement plus robustes.

#### 4.1.2.3 Points saillants

Tangelder et Veltkamp [118] proposent une approche fondée sur des représentations par points saillants. Le volume de l'objet 3D est régulièrement partitionné en $N^3$ cellules de taille égale. Trois méthodes sont proposées pour attribuer à chacune des cellules non vides une valeur caractéristique et un poids :
1. le sommet ayant la plus forte courbure gaussienne, et sa courbure gaussienne normalisée,
2. la position moyenne des sommets appartenant à la cellule, et la variation moyenne des normales des facettes appartenant à la cellule,
3. le centre de masse des sommets appartenant à la cellule, et un poids égal à 1.
Pour comparer deux ensembles de points saillants, les auteurs proposent la mesure *Proportionnal Transportation Distance* inspirée de la distance *Earth Mover's Distance*. Cette mesure permet d'effectuer l'appariement entre différentes représentations et de garantir des invariances par rapport aux petites déformations locales. Elle est, par contre, très coûteuse.

#### 4.1.2.4 Contextes de forme 3D

Initialement utilisés dans la reconnaissance de forme 2D [13], les contextes de forme (*Shape Context*) ont été étendus en 3D par Körtgen et al. [71] pour l'indexation 3D. Les calculs de contextes de forme 3D consistent à prélever $N$ points uniformément répartis sur la surface de l'objet 3D et à partitionner l'espace en $C$ cellules obtenues par une superposition de coquilles concentriques centrées et de secteurs angulaires partant du centre de l'objet. Sur chaque point $\mathbf{p}_i$, un histogramme de taille $C$ est calculé sur les positions relatives des $(N-1)$ points $\mathbf{p}_j$ : chaque vecteur $\mathbf{p}_i - \mathbf{p}_j$ est localisé par la cellule dans laquelle il se trouve, et la composante associée à cette cellule est incrémentée de 1. En pratique, les auteurs choisissent d'utiliser $N$ histogrammes de $C$ cellules avec $N = 200$ et $C = 432$ ($N_\theta * N_\varphi = 6 * 12$ secteurs et $N_r = 6$ $\log_2$-coquilles), soit un descripteur de très grande taille égale à 86400 coefficients. Pour comparer deux descrip-

teurs, les auteurs proposent de calculer le meilleur appariement entre les deux ensembles d'histogrammes en minimisant un coût de correspondance normalisé. Cette méthode requiert des mesures de similarité entre toutes les paires possibles des contextes de forme des deux objets, soit une complexité de l'ordre de $\mathcal{O}(N^3)$. Comparés à de nombreuses méthodes [119], les contextes de forme 3D sont peu discriminants et trop coûteux en temps et en mémoire, les limitations principales étant le choix aléatoire des points sur lesquels est défini et calculé le descripteur, et le nombre $N$ qui reste très grand pour une recherche globale, mais peut être efficace pour une recherche partielle. Par ailleurs, les contextes de forme n'étant pas invariants aux transformations géométriques, cette méthode requiert un alignement spatial préalable.

#### 4.1.2.5 Voxelisation de forme 3D

Le processus de voxelisation d'une forme 3D définie par un maillage se décompose en trois étapes : discrétisation, échantillonnage et représentation. La première consiste à discrétiser l'espace des paramètres $(x, y, z)$ pour partitionner ensuite la boîte englobant l'objet 3D en $N^3$ cellules (ou voxels) de taille égale. La dernière étape attribue une valeur numérique à chaque cellule n'intersectant pas la surface de l'objet 3D. Dans le cas binaire, cette valeur est définie par

$$v_{ijk} = \begin{cases} 1 & \text{si } \mathfrak{S} \cap \mu_{ijk} \neq \emptyset, \\ 0 & \text{si } \mathfrak{S} \cap \mu_{ijk} = \emptyset, \end{cases} \tag{4.3}$$

avec $\mu_{ijk}$ le volume définissant la cellule d'indices $ijk$ et $\mathfrak{S}$ (1.4) la surface de l'objet 3D. Dans [131], Vranic propose de représenter cette valeur par

$$v_{ijk} = \frac{Aire(\mathfrak{S} \cap \mu_{ijk})}{\mathcal{A}}, \quad \mathcal{A} \text{ (1.6) étant l'aire de la surface } \mathfrak{S}.$$

Dans le cadre d'une indexation 3D par voxelisation, le descripteur de forme est défini par le vecteur caractéristique $[v_{ijk}]_{0 \leq i,j,k < N}$. Ce type d'approches fait appel à une méthode d'alignement pour assurer une invariance aux rotations.

#### 4.1.2.6 Descripteurs sphériques 3D

L'objectif de ces descripteurs est de caractériser l'objet 3D par une ou plusieurs fonctions sphériques calculées sur sa surface. On distingue deux types de partitions sphériques construites à partir d'un échantillonnage uniforme de l'espace des paramètres $(\theta, \varphi)$ ou d'un échantillonnage uniforme de la surface de la

boule unité en un ensemble de points $\{\mathbf{u}_i, 0 \leq i < N\}$ . La fonction sphérique est donnée par l'ensemble des attributs descriptifs caractérisant localement les cellules. Plusieurs méthodes de calcul d'attributs sont proposées dans [67, 131].

### 4.1.3 Approches par transformées

Les représentations de forme 2D fondées sur des transformées ont été souvent proposées dans la littérature comme descripteurs de forme 2D. On peut citer : le descripteur de Fourier, le descripteur de Hough, les moments géométriques et les moments de Zernike. Par analogie 2D-3D, la plupart de ces transformées ont été étendues et utilisées en indexation 3D. La représentation de l'objet 3D par une transformée est généralement définie par la projection de la fonction caractérisant l'objet 3D (surfacique ou volumique) sur une famille de fonctions caractéristiques. Dans cette section, nous nous sommes principalement intéressés aux méthodes basées sur les moments, les transformées et les harmoniques sphériques.

#### 4.1.3.1 Moments géométriques

Les moments géométriques sont définis par une projection d'une fonction caractéristique $f$ de l'objet 3D sur la famille des fonctions $x^p y^q z^r$, $(p,\ q,\ r) \in \mathbb{N}^3$. Les moments géométriques dits d'ordre $(p+q+r)$ sont notés par $M_{pqr}$ et peuvent être calculés par :

$$M_{pqr} = \iint_{\mathbf{p} \in \mathfrak{S}} f(\mathbf{p})\ x^p y^q z^r\ ds. \tag{4.4}$$

En particulier, les moments géométriques d'ordre un et deux sont utilisés pour calculer les paramètres de la normalisation de l'objet 3D (le centre de gravité et les trois axes principaux). Dans le contexte de l'indexation, la description d'une forme par les moments géométriques a été introduite dans l'approche par moments statistiques centrés 3D de Paquet et Rioux [98]. Le calcul des moments se fait de façon discrète sur les points du maillage de l'objet en appliquant la formule suivante :

$$M_{pqr} = \sum_{i=1}^{N_T} A_i\ (x_i - x_{\mathbf{g}})^p (y_i - y_{\mathbf{g}})^q (z_i - z_{\mathbf{g}})^r\ ds,$$

$N_T$ (1.1) étant le nombre de triangles du maillage, $A_i$ et $\mathbf{g}_i = (x_i, y_i, z_i)$ (1.5) étant respectivement l'aire et le centre de gravite du triangle d'indice $i$, et $\mathbf{g}$ (2.1) le centre de gravité de l'objet. Cette approche est sensible aux remaillages.

Saupe et Vranic [109] proposent d'utiliser une fonction sphérique $r(\mathbf{u})$ définie sur la boule unité. L'objet 3D est centré, mis à l'échelle, et aligné via une analyse en composantes principales continue. Les moments géométriques d'ordre $(p+q+r)$ sont calculés comme suit :

$$M_{pqr} = \sum_{a=0}^{N} \sum_{b=0}^{N} \Delta_{s_a} \; r(\mathbf{u}_{ab}) \; x_{ab}^{p} \; y_{ab}^{q} \; z_{ab}^{r},$$

$r(\mathbf{u}_{ab})$ étant la distance maximale entre la surface de l'objet et son centre, suivant la direction $\mathbf{u}_{ab} = \mathbf{u}(\theta_a, \varphi_b) = (\cos\varphi_b \sin\theta_a,\ \sin\varphi_b \sin\theta_a,\ \cos\theta_a) = (x_{ab},\ y_{ab},\ z_{ab})$, et $\Delta_{s_a} = \frac{4\pi}{N}(\cos(\theta_a - \frac{\pi}{N}) - \cos(\theta_a - \frac{\pi}{N}))$ l'aire de la surface sur la boule unité correspondant au point $\mathbf{u}_{ab}$ pour compenser l'échantillonnage angulaire non uniforme. Les moments d'ordre $1 \leq p+q+r \leq k$ forment le vecteur caractéristique de l'objet 3D. Les meilleurs résultats sont obtenus pour $k \geq 11$. L'utilisation des moments géométriques ne semble pas être à conseiller. En effet, une étude comparative en termes de performance faite par Vranic [131] sur différentes bases d'objets 3D place ce descripteur à la dernière place.

#### 4.1.3.2 Moments de Zernike 3D

Les moments de Zernike 3D définis par Canterakis [25] est une extension des moments de Zernike 2D, très utilisés dans la reconnaissance de forme 2D. Ces moments tridimensionnels sont calculés en projetant une fonction définissant l'objet 3D, $f$, sur un système complet de fonctions orthonormées définies sur la boule unité. Ces fonctions sont appelées polynômes de Zernike 3D et notées $Z_{nl}^{m}$. Il est possible d'approximer la fonction $f$ par un nombre fini de ses moments de Zernike $\Omega_{nl}^{m}$ :

$$\hat{f}(\mathbf{p}) = \sum_{n} \sum_{l} \sum_{m} \Omega_{nl}^{m} \cdot Z_{nl}^{m}(\mathbf{p}), \tag{4.5}$$

avec $n \in [0, N]$, $l \in [0, n]$ tel que $(n-l)$ est paire et $m \in [-l, l]$. $\Omega_{nl}^{m}$ sont donnés par une combinaison linéaire des moments géométriques $M_{rst}$ d'ordre $(r+s+t)$ définis dans la section précédente :

$$\Omega_{nl}^{m} = \frac{3}{4\pi} \sum_{r+s+t \leq n} X_{rst}^{nlm} M_{rst}, \tag{4.6}$$

$X_{rst}^{nlm}$ étant la fonction définissant la combinaison linéaire.
Novotni et Klein [85] ont repris cette décomposition pour définir le descripteur de Zernike 3D. Le calcul du descripteur se fait de façon discrète en quatre étapes :

1. Voxeliser l'objet centré et mis à l'échelle, pour définir la fonction $f$,
2. Calculer tous les moments géométriques d'ordre $(p+q+r) \leq N$,
3. Calculer les moments de Zernike $\Omega_{nl}^{m}$ pour $m \geq 0$ et en déduire les $\Omega_{nl}^{m}$ pour $m < 0$ en utilisant la relation de symétrie $\Omega_{nl}^{-m} = (-1)^m \overline{\Omega_{nl}^{m}}$,
4. Calculer la norme du vecteur $F_{nl} = \|\Omega_{nl}\|$. L'ensemble $F_{nl}$, $n \in [0, N]$ et $l \in [0, n]$, constitue le descripteur de Zernike 3D de taille $\frac{(N+2)*(N+1)}{2}$.

La méthode proposée est invariante aux translations, aux rotations, aux symétries et aux changements d'échelle. Les expérimentations comparant la description par moments de Zernike 3D à celle des harmoniques sphériques aboutissent à des résultats similaires pour la classe des avions et une supériorité du descripteur de Zernike 3D pour la classe des chaises.

#### 4.1.3.3 Transformée de Fourier 3D

La transformée de Fourier a été utilisée comme descripteur de forme 2D pour caractériser les formes extraites d'une image 2D. Vranic et Saupe [128] l'ont étendu sur les objets 3D et ont proposé d'utiliser la transformée de Fourier 3D dans le cadre de l'indexation par le contenu. Sur l'objet voxelisé préalablement normalisé par une analyse en composante principale continue, les coefficients de Fourier sont calculés selon la formule :

$$\hat{f}_{pqr} = \frac{1}{\sqrt{N^3}} \sum_{a=-\frac{N}{2}}^{\frac{N}{2}-1} \sum_{b=-\frac{N}{2}}^{\frac{N}{2}-1} \sum_{c=-\frac{N}{2}}^{\frac{N}{2}-1} f_{pqr} \; e^{-j2\pi(pa+qb+rc)/N}, \tag{4.7}$$

où $-N/2 \leq a, b, c, p, q, r < N/2$ et $f_{pqr}$ est l'ensemble des voxels caractérisant l'objet 3D. Les normes des coefficients $|\hat{f}_{pqr}|$ vérifiant $1 \leq |p| + |q| + |r| \leq k \leq N/2$ constituent le descripteur de Fourier. En effet, le terme $|p| + |q| + |r|$ permet de limiter une partie des coefficients à cause la symétrie hermitienne de la transformée de Fourier et la borne $k$ permet de retenir seulement les basses fréquences pour leur pouvoir discriminant.

#### 4.1.3.4 Transformée de Hough 3D

Le descripteur de Hough 3D (DH3D) introduit par Zaharia et Prêteux dans [135, 136] est fondé sur un principe d'accumulation des points de la surface de l'objet 3D sur des plans de $\mathbb{R}^3$. Tout plan est représenté en coordonnées sphériques par le triplet $(r, \theta, \varphi)$, où $r \geq 0$ est sa distance à l'origine, $\theta \in [0, 2\pi[$ l'angle d'azimut

et $\varphi \in [-\frac{\pi}{2}, \frac{\pi}{2}[$ l'angle d'élévation. En échantillonnant uniformément l'espace des paramètres sphériques, un ensemble de $N_r \times N_\theta \times N_\varphi$ composantes est créé pour définir l'histogramme de Hough 3D. La contribution d'une facette sur un plan d'orientation $(\theta_j, \varphi_k)$ est donnée par l'aire de sa projection sur le plan considéré. La composante $h(r_i, \theta_j, \varphi_k)$ est incrémentée par la contribution de chaque facette ayant une distance $r_{j,k}^{\mathbf{g}_l}$ quantifiée à la plus proche valeur de $r_i$, $r_{j,k}^{\mathbf{g}_l} = \mathbf{g}_l \cdot \mathbf{n}_{jk}$ étant la distance signée du plan d'orientation $(\theta_j, \varphi_k)$ passant par le centre de gravité $\mathbf{g}_l$ de la facette. La construction de ce descripteur est effectuée dans le cadre d'un repère intrinsèque défini par le centre de gravité et les axes principaux de l'objet 3D. Pour s'affranchir des problèmes d'étiquetage et d'inversion des axes principaux, les auteurs proposent une version optimisée (DH3DO) en considérant les partitions de la sphère unité obtenues par projection des sommets d'un polyèdre régulier sur la sphère. Comparé au SF3D, le DH3D s'est révélé plus performant sur la base de MPEG-7 et intrinsèquement plus stable topologiquement.

#### 4.1.3.5 Transformée radiale-angulaire 3D

Ricard et al. [105, 106] proposent une nouvelle méthode d'indexation d'objets 3D basée sur la transformée radiale-angulaire 3D (ART 3D). Cette transformation unitaire orthogonale complexe est définie sur la boule unité et donnée par le produit de trois fonctions séparables suivant la direction radiale et les deux directions angulaires. Le descripteur de forme est donné par un vecteur contenant les coefficients ART 3D d'ordre $(n, m_\theta, m_\varphi)$ définis comme suit :

$$F_{nm_\theta m_\varphi} = \int_0^{2\pi} \int_0^{\pi} \int_0^{1} f(\rho, \theta, \varphi) \overline{R_n(\rho) A_{m_\theta}(\theta) A_{m_\varphi}(\varphi)} \, \rho d\rho d\theta d\varphi, \tag{4.8}$$

$f$ étant une fonction sphérique représentant la surface de l'objet en coordonnées sphériques.
Comme le descripteur obtenu est seulement invariant aux rotations autour de l'axe $z$, les auteurs proposent un centrage, une mise à l'échelle et un alignement du premier axe principal suivant l'axe $z$. Les résultats expérimentaux montrent une supériorité des harmoniques sphériques sur la base de Princeton et une équivalence de performances sur la base de Renault.

#### 4.1.3.6 Harmoniques sphériques

Les harmoniques sphériques notées $Y_l^m$ $(l \geq 0, |m| \leq l)$ sont des fonctions harmoniques particulières qui constituent une base orthonormée de l'espace de

Hilbert $L^2(S^2)$ des fonctions de carré intégrables sur la sphère $S^2$. La description d'un objet 3D par les harmoniques sphériques requiert une étape de représentation de l'objet par une ou plusieurs fonctions sphériques. Les harmoniques sphériques sont particulièrement utiles pour s'affranchir du problème d'invariance aux rotations.

La théorie des harmoniques sphériques consiste à décomposer une fonction sphérique $f(\theta, \varphi)$ définie sur la sphère $S^2$ en une somme de ses coefficients harmoniques :

$$f(\theta, \varphi) = \sum_{l \geq 0} \sum_{|m| \leq l} \hat{f}_{l,m} \, Y_l^m(\theta, \varphi), \tag{4.9}$$

$$\hat{f}_{l,m} =< f(\theta, \varphi), Y_l^m(\theta, \varphi) >= \int_0^{\pi} \int_0^{2\pi} f(\theta, \varphi) \overline{Y_l^m(\theta, \varphi)} \, \sin\theta d\varphi d\theta.$$

Cette technique peut être vue comme une généralisation de la transformée de Fourier aux fonctions sphériques.

La décomposition en harmoniques sphériques a été utilisée par Funkhouser et al. [49] sur des fonctions définies par l'intersection de l'objet voxelisé avec un ensemble de sphères concentriques centrées (cf. Figure 4.6). L'inconvénient de cette approche est sa dépendance par rapport à la résolution de la voxelisation, entraînant une perte de détails dans la description de la forme 3D.

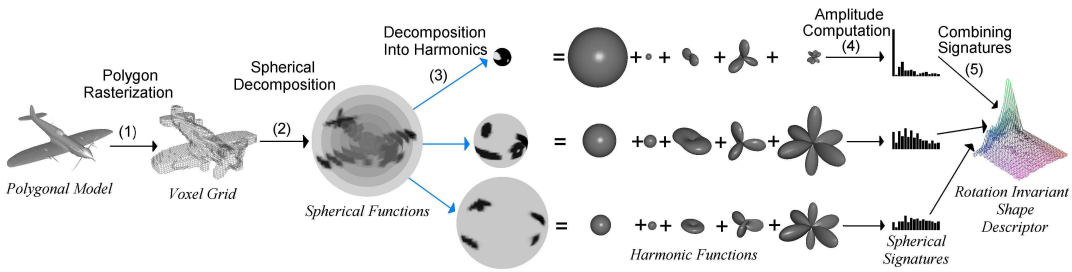



FIG. 4.6 – La décomposition en harmoniques sphériques de la fonction sphérique voxelisée [49].

Vranic et al. [127, 129, 130] proposent d'appliquer la méthode sur des fonctions sphériques directement calculées sur la surface de l'objet 3D. Quatre fonctions ont été considérées : $r(\mathbf{u})$ la distance maximale entre la surface de l'objet et son centre, utilisée dans leur méthode basée sur les moments géométriques, $s(\mathbf{u})$ la norme du produit scalaire de $\mathbf{u}$ et la normale unitaire du triangle contenant le point $r(\mathbf{u})\mathbf{u}$, $r(\mathbf{u}) + js(\mathbf{u})$ la fonction complexe, et $r_k$ $k \in \{1, ..., R\}$ ensemble

de fonctions sphériques, $r_k(\mathbf{u})$ étant la distance maximale entre la surface de l'objet prise dans la coquille $k$ et le centre de l'objet. L'étude comparative [131] des quatre méthodes en termes de performance place la décomposition en harmoniques sphériques de la quatrième méthode basée sur les fonctions $r_k$ à la première place.
Kazhdan et al. [64, 67] ont appliqué la décomposition harmonique sur les fonctions sphériques d'un certain nombre de descripteurs existants, pour avoir des représentations invariantes aux rotations. Parmi ces fonctions, on cite les EGI [58], la fonction sphérique étendue $r(\mathbf{u})$ [129], la distribution radiale, les distributions surfaciques [6] par secteurs et par combinée coquilles-secteurs, et les fonctions sphériques d'une représentation voxelisé [49]. À travers les résultats obtenus sur leur base de données, les auteurs montrent deux avantages principaux de la décomposition en harmoniques sphériques. D'une part, la nouvelle représentation améliore la performance du descripteur initialement aligné par une ACP. D'autre part, elle réduit la taille de la signature conduisant à une représentation plus compacte et un calcul de similarité plus rapide.

#### 4.1.3.7 Descripteurs de symétrie

Kazhdan et al. [65, 66] introduisent le descripteur de symétrie. C'est une fonction sphérique qui mesure la symétrie de réflexion par rapport à tout plan passant par le centre de l'objet. L'ensemble des plans est paramétré par $(\theta, \varphi)$ en prenant les normales $(\cos\varphi\sin\theta,\ \sin\varphi\sin\theta,\ \cos\theta)$. L'estimation de la symétrie de réflexion se fait indirectement sur une ou plusieurs fonctions sphériques définissant l'objet 3D. La Figure 4.7 montre des exemples de descripteurs de symétrie calculés sur différents objets 3D. Les pointes correspondent aux plans proches d'une symétrie miroir et les vallées correspondent aux plans proches d'une antisymétrie.

Kazhdan et al. [64, 68] généralisent ce concept de symétrie de réflexion aux symétries à $k$-plis (*$k$-fold symmetry*). Pour calculer l'autocorrélation des fonctions sphériques dans l'espace des rotations, les auteurs proposent une méthode efficace et rapide, fondée sur la décomposition en harmoniques sphériques qui a des propriétés intéressantes vis-à-vis de ces symétries. Les résultats expérimentaux montrent que la combinaison - descripteur de symétrie et un des descripteurs sphériques - permet d'améliorer les performances de la recherche.
Podolak et al. [101] proposent une nouvelle représentation, fondée sur une trans-

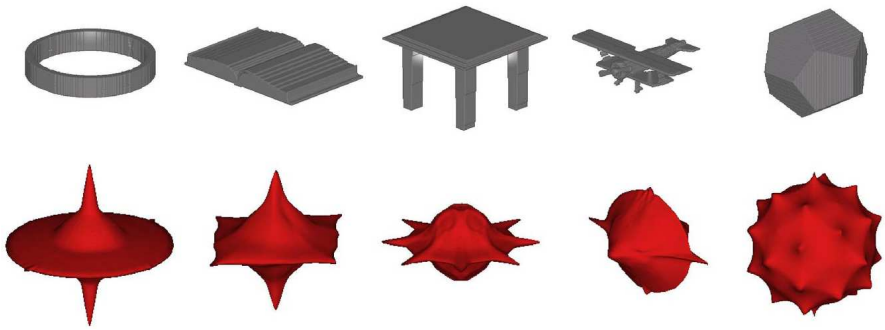


FIG. 4.7 – Cinq modèles 3D et leurs descripteurs de symétrie miroir respectives [65].

formée de symétrie de réflexion planaire (PRST). Contrairement au descripteur proposé dans [66], le nouveau descripteur mesure la symétrie de réflexion par rapport à tout plan intersectant l'objet 3D, et ceci se fait directement sur la surface de l'objet 3D. Les auteurs utilisent une estimation de Monte Carlo pour réduire la complexité de leur méthode et un algorithme de raffinement itératif pour augmenter la précision de la transformée. Dans un autre contexte, cette transformée est utile pour l'alignement, la segmentation et la sélection automatique des meilleurs points de vue. L'aspect limitatif de cette méthode reste le coût de calcul.

### 4.1.4 Approches structurelles

Ces approches consistent à caractériser les objets 3D par une ou plusieurs représentations structurelles, généralement sous forme de graphes. La comparaison des objets se fait en utilisant des algorithmes de mise en correspondance de graphes pour estimer les variations et les différences structurelles entre objets 3D. Ce type de représentation permet de décrire de manière précise la structure de l'objet 3D et d'extraire des informations de haut niveau, intrinsèques à l'objet 3D. Cependant, ces méthodes restent coûteuses, très sensibles aux bruits et très dépendantes de la topologie.

#### 4.1.4.1 Graphe de formes élémentaires 3D

Les méthodes de ce type représentent l'objet 3D comme une suite de formes élémentaires 3D. Plusieurs types de primitives volumiques ont été utilisés pour représenter les objets 3D : les cubes, les cylindres, les sphères, les ellipsoïdes, les cônes, les prismes, etc.

Medioni et François [78] proposent d'utiliser une modélisation par sous-parties volumiques des objets 3D à base de formes géométriques élémentaires appelées Geons (*GEometric iONS*) [18]). Après une décomposition de l'objet 3D en geons, ces éléments sont hiérarchiquement organisés au sein d'un arbre de description, intégrant également des relations d'adjacence. Cette structure hiérarchique permet en outre de définir et de calculer plus efficacement une mesure de similarité fondée sur un coût de transition issu d'une méthode d'appariement par graphes. Irani et Ware [61] transforment le graphe de geons 3D en un diagramme UML et comparent par la suite les diagrammes UML entre eux.

#### 4.1.4.2 Squelette d'un objet 3D

Sundar et al. [117] proposent de représenter l'objet 3D par un squelette sous forme d'un graphe dont les nœuds sont décrits par des attributs géométriques et topologiques. Après une voxelisation de l'objet 3D, les points du squelette sont obtenus en appliquant l'algorithme d'amincissement basé sur une transformée de distance [51]. L'amincissement de l'objet consiste à enlever successivement des couches selon un paramètre d'amincissement. Les points du squelette sont connectés dans un graphe de forme acyclique par application d'un algorithme de l'arbre couvrant de poids minimum. En utilisant différents paramètres d'amincissement, une structure hiérarchique de graphes est obtenue. Ces graphes contiennent en chacun des nœuds des informations géométriques sur la distribution radiale autour du segment du squelette original associé au nœud, et des informations topo-

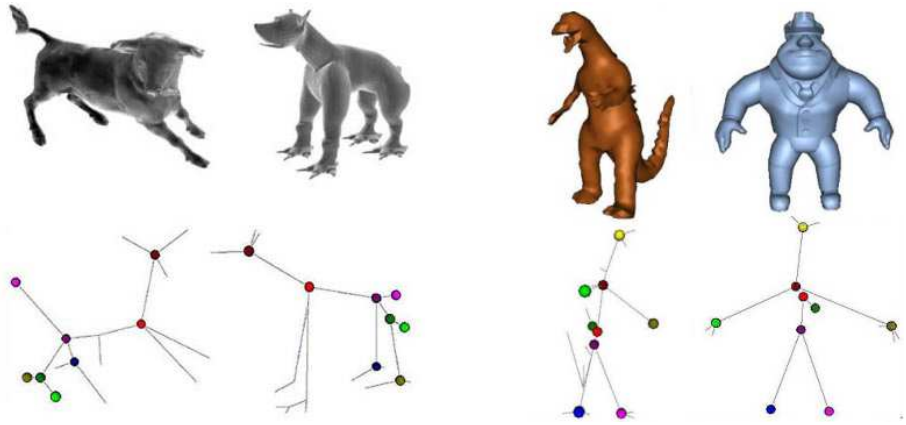


FIG. 4.8 – Exemples d'appariements entre deux graphes de squelettes : deux animaux (à gauche), un dinosaure et un humain (à droite). Les couleurs des nœuds montrent les appariements nœud-à-nœud des squelettes [117].

logiques sur les sous-graphes communs au nœud utilisant les valeurs propres de ses matrices d'adjacence. Pour mesurer la similarité entre objets 3D, les auteurs effectuent une comparaison approximative entre leurs graphes hiérarchiques, basée sur un algorithme glouton qui calcule la cardinalité maximale et la correspondance de poids minimal dans un graphe bipartie. Pour parvenir à des appariements optimaux entre objets articulés, il est préférable de ne considérer que la signature topologique.

#### 4.1.4.3 Graphe de Reeb

Le graphe de Reeb d'un objet 3D se présente sous forme d'un squelette dont sa construction repose sur la théorie de Morse [113] : Une fonction continue $\mu$ définie sur une surface fermée caractérise la topologie de la surface en ses points critiques. Selon les valeurs de la fonction $\mu$, la surface de l'objet est segmentée en régions auxquelles sont associés les nœuds du graphe. Le squelette est ensuite obtenu en reliant les nœuds dont les régions sont connexes. Il existe différentes fonctions continues $\mu$ [16] qui peuvent être utilisées dans la construction du graphe de Reeb (cf. Figure 4.9). Comme l'aspect du graphe résultant est entièrement lié au choix de la fonction $\mu$, il est important de choisir soigneusement cette fonction.

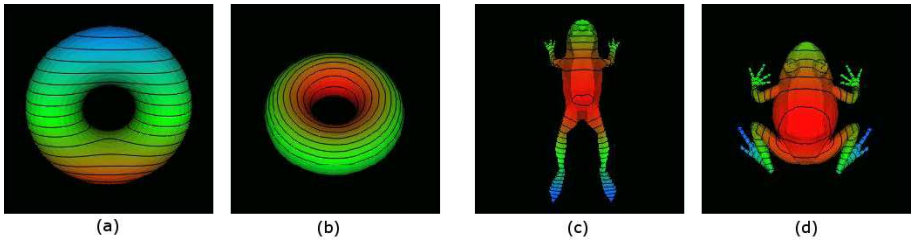



FIG. 4.9 – Exemples de fonctions $\mu$ : $h$ hauteur (a), $g$ intégrale de distance géodésique (b, c et d) [57]. Dans le cas des tores, (a et b), la fonction basée sur $g$ caractérise de manière plus intrinsèque la topologie de la forme, comparée à la fonction basée sur $h$ fortement liée à l'information de la hauteur. Dans le cas des grenouilles, (c et d), les différentes parties conservent la même valeur de $\mu$ indépendamment de la position des bras et des pieds et aussi des dimensionnements des objets, ce qui rend la représentation avec la fonction $g$ plus intuitive.

La méthode d'appariement de graphes de Reeb multirésolution introduite par Hilaga et al. [57] utilise la distance géodésique $g$ dans le calcul de $\mu$. $\mathbf{v}$ étant un point de la surface $\mathfrak{S}$, $\mu$ est définie par :

$$\mu(\mathbf{v}) = \int_{\mathbf{p}\in\mathfrak{S}} g(\mathbf{v},\mathbf{p})\ ds, \tag{4.10}$$

avec $g(\mathbf{v}, \mathbf{p})$ les distances géodésiques du point $\mathbf{v}$ aux autres points $\mathbf{p}$ de $\mathfrak{S}$. Cette fonction a la propriété d'invariance aux rotations, aux bruits et petites ondulations ou déformations. Pour avoir une invariance par rapport à l'échelle, une version normalisée de $\mu$ est proposée.
Durant la phase de l'indexation, les auteurs proposent de représenter l'objet 3D par des graphes de Reeb à plusieurs niveaux de résolution. La similarité entre deux objets est estimée par une mise en correspondance des nœuds des graphes de Reeb en partant des diagrammes les plus grossiers et en propageant la mise en correspondance jusqu'aux graphes les plus fins. L'aspect multirésolution est adapté aux recherches dans les grandes bases de données. Elles se limitent à des comparaisons à faibles niveaux de résolutions afin d'accélérer le calcul, comme elles peuvent s'étendre à la résolution maximale pour obtenir des résultats plus fins.

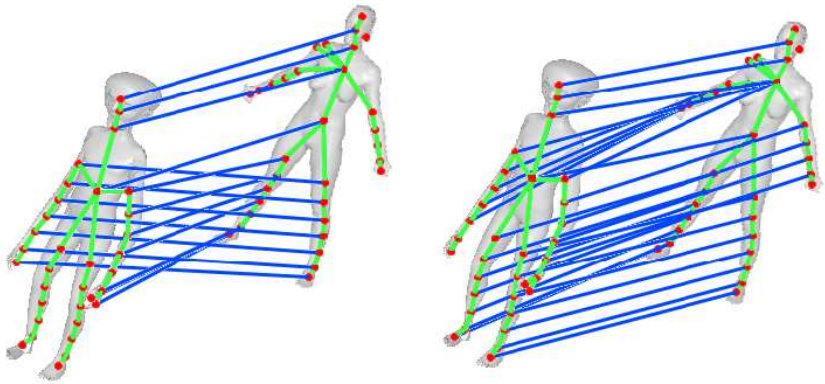


FIG. 4.10 – Deux exemples d'appariements entre deux graphes de Reeb multirésolution, sans et avec l'information géométrique [122]. Sans l'information géométrique (à gauche), les jambes peuvent être appariées aux bras car ils sont topologiquement équivalents. En ajoutant les attributs géométriques (à droite), tous les nœuds sont naturellement appariés.

Insuffisante pour obtenir des appariements satisfaisants, cette approche a été étendue par Biasotti et al. [17] et par Tung et Schmitt [121, 122] (cf. Figure 4.10). Ils proposent d'enrichir ces graphes par un ensemble d'attributs géométriques pour améliorer leur mise en correspondance et l'estimation de la similarité entre modèles.
Les graphes de Reeb basés sur une distance géodésique sont efficaces dans la recherche des objets articulés. Cependant, ils nécessitent de disposer un maillage propre. Donc, ils ne peuvent pas être appliqués à tout maillage car le moindre

problème topologique, par exemple des faces oubliées, peut perturber le calcul de la distance géodésique. Dans certains cas, une étape de prétraitement peut résoudre ces problèmes.

## 4.2 Approches 2D/3D

Dans cette section, nous nous intéressons aux méthodes qui caractérisent l'objet 3D par un ensemble de projections 2D. Nous présentons ces approches selon deux classes :

- les approches multi-vues : à base de silhouette ou d'images de profondeur,
- les autres approches : à base de coupes ou d'images de spin.

### 4.2.1 Approches multi-vues

L'idée générale des approches multi-vues est basée sur le fait que deux objets 3D sont similaires ont des vues similaires, et vice versa. En pratique, ces approches 2D/3D consistent à associer à l'objet 3D un ensemble d'images de projections 2D prises sous différents angles de vue. La forme 3D est alors indirectement représentée par différents descripteurs de forme 2D associés à ces images de projection. La similarité entre deux objets 3D s'obtient à partir des mesures de similarités calculées entre les descripteurs de forme 2D. Notons que ce principe peut être utilisé pour retrouver des objets 3D à partir d'une requête 2D [9, 37, 49, 73, 75, 105], et vice versa. Deux paramètres sont à prendre en compte dans les approches multi-vues : le nombre de vues à considérer et la position des vues dans l'espace tridimensionnel.

Les deux principales projections rencontrées dans les approches multi-vues sont les silhouettes [8, 34, 76, 131] et les images de profondeur [56, 92, 93, 100, 124]. Bien que les silhouettes soient des images faciles à stocker et très utiles dans certaines approches, leur principal inconvénient est le peu d'information provenant de l'objet 3D contenu dans chaque image. Pour obtenir une description plus complète de la forme 3D et capter plus les propriétés topologiques et structurelles de l'objet, il s'avère intéressant d'utiliser les images de profondeur.

#### 4.2.1.1 Approches à base de silhouettes

Mahmoudi et Daoudi [76] proposent une méthode 2D/3D à base de silhouettes. L'objet 3D est aligné par une analyse en composantes principales et représenté par un ensemble de 7 vues : les trois premières sont déterminées par les directions principales et les quatre autres sont déduites des principales vues. Pour indexer les silhouettes caractérisant l'objet 3D, les auteurs proposent d'utiliser un descripteur promu dans MPEG-7, le CSS (*Curvature Scale Space*) organisé autour d'une structure arborescente *M-Tree*. Ce descripteur caractérise le contour en exploitant les maxima de courbure, détectés à travers une analyse multi-échelle.

Vranic décrit dans [131] une approche 2D/3D qui consiste à caractériser l'objet 3D par seulement trois silhouettes. Après un centrage, une mise à échelle et un alignement par une ACP de l'objet 3D, les trois silhouettes sont prises dans les directions x, y et z du nouveau repère. Pour caractériser la silhouette, Vranic utilise le descripteur de Fourier unidimensionnel sur une séquence de points obtenus avec un échantillonnage angulaire uniforme ou un échantillonnage uniforme du contour. Comparée aux autres méthodes proposées, cette approche se positionne parmi les meilleures. Elle sera détaillée dans la section 5.2.3.1.

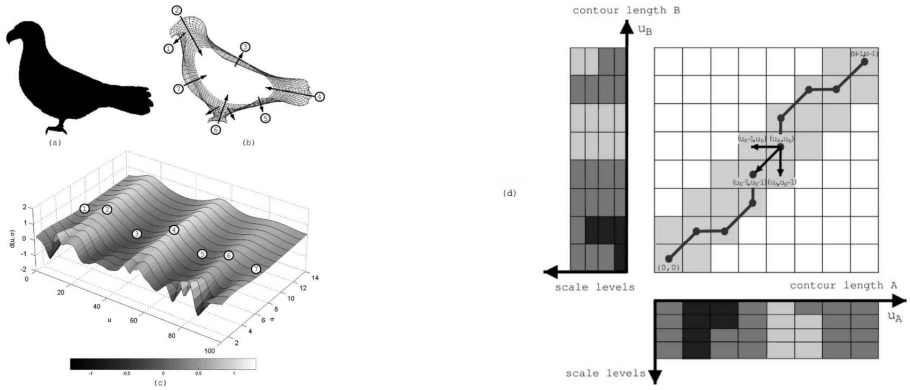



FIG. 4.11 – Représentation d'une silhouette avec MCC et la distance associée : (a) image originale de la silhouette, (b) versions filtrées du contour sur différents niveaux d'échelle, (c) la représentation finale MCC pour 100 points de contours sur 14 niveaux d'échelle, (d) correspondance entre deux représentations MCC par une programmation dynamique [3].

De la même manière, Napoléon et al. [82] utilisent trois silhouettes prises dans les directions principales de l'objet 3D. La nouveauté consiste à représenter

chaque silhouette par un descripteur multi-échelle du contour (MCC, *Multi-Scale Convexity/Concavity representation* [3]) et à utiliser la programmation dynamique dans la comparaison des descripteurs pour trouver la meilleure correspondance entre silhouettes (cf. Figure 4.11). Cette méthode d'indexation 2D est invariante aux transformations affines 2D (translation, rotation et changement d'échelle) et robuste aux petites distorsions et déformations géométriques. Comme dans [136], pour comparer deux objets 3D, les auteurs proposent d'utiliser une mesure de similarité optimale calculée sur les 48 configurations possibles de poses afin de s'affranchir des problèmes d'étiquetage et d'inversion des axes principaux.

Dans [33, 34, 111], Chen et al. proposent un descripteur 2D/3D appelé *Light-Field Descriptor* (LFD). L'objet 3D est représenté par un ensemble de $N$ *Light-Fields* $(L_i)_{i=\{0...N\}}$, $L_i$ étant un ensemble de dix vues prises sur les dix premiers sommets d'un dodécaèdre centré autour de l'objet. Les positions des caméras des $N$ *LightFields* sont répartis uniformément sur la sphère unité pour couvrir différents angles de vues. Les vues sont représentées par les silhouettes de l'objet 3D notées $(\mathbf{s}_{i,k})_{k=\{0...10\}}$. Dans ce mode de projection, les silhouettes créées sur deux sommets opposés sont identiques, c'est pourquoi dix sommets sont considérés au lieu de vingt. La silhouette $\mathbf{s}_{i,k}$ est décrite par un vecteur $(\mathbf{C}_{i,k})_{l=\{0...45\}}$ défini par ses 35 moments de Zernike et ses 10 coefficients de Fourier.

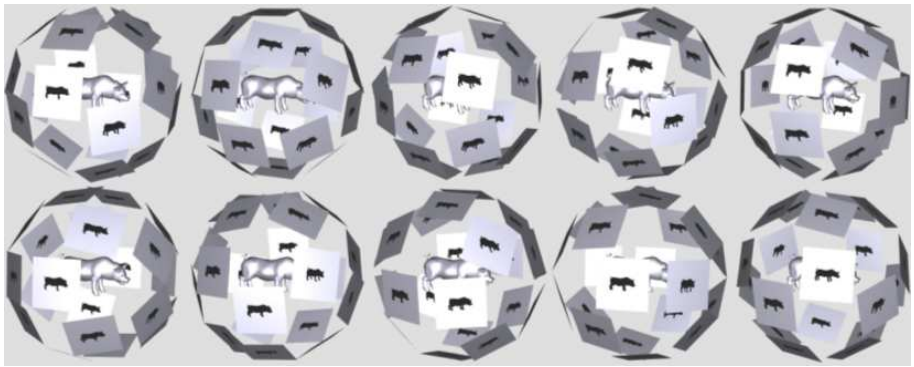


FIG. 4.12 – Un ensemble de *LightFields* pour un modèle 3D, N = 10 [34].

Pour comparer deux objets, les auteurs proposent une mesure de similarité invariante aux rotations. Leur idée consiste à faire tourner les objets autour du centre jusqu'à la similarité globale la plus élevée (corrélation croisée) entre deux

ensembles de vues soit atteinte, en prenant en compte des rotations à l'intérieur de chaque *LightField* et des positions de l'ensemble des $N$ *LightFields*. La mesure de similarité entre deux objets, $O^1$ et $O^2$ représentés respectivement par $L_i^1$ et $L_j^2$, est donnée comme suit :

$$\Delta(O^1, O^2) = \min_{1 \leq i,j \leq N} d_L(L_i^1, L_j^2), \tag{4.11}$$

$d_L(L_i^1, L_j^2)$ étant la mesure de similarité entre deux *LightFields* défini par :

$$d_L(L_i^1, L_j^2) = \min_{1 \leq r \leq 60} \left( \sum_{l=1}^{10} d(\mathbf{s}_{i,k}^1, \mathbf{s}_{j,r(k)}^2) \right),$$

$$\text{avec} \quad d(\mathbf{s}_{i,k}^1, \mathbf{s}_{j,r(k)}^2) = \sum_{l=1}^{45} |\mathbf{C}_{i,k,l}^1, \mathbf{C}_{j,r(k),l}^2|.$$

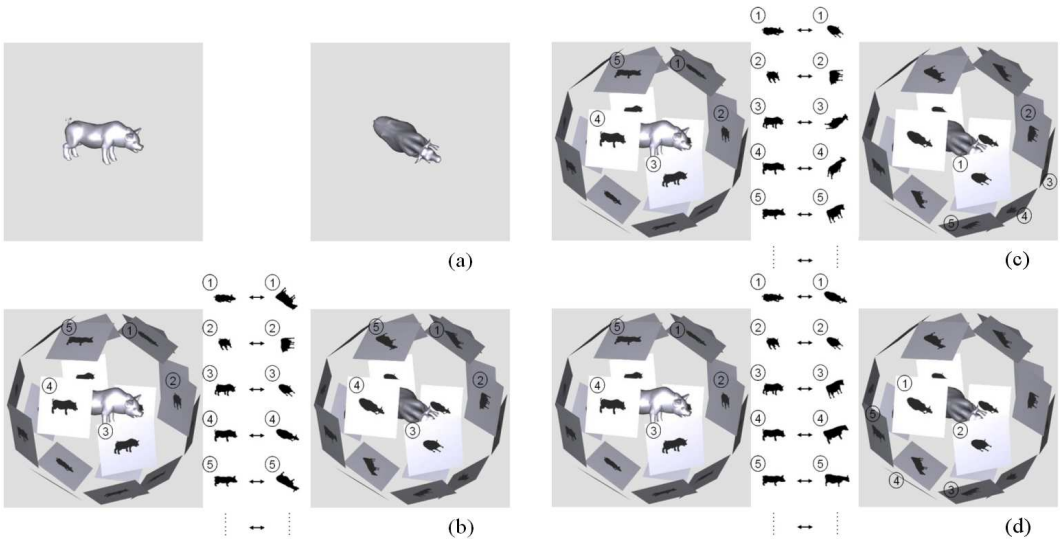



FIG. 4.13 – Comparaison de deux *LightFields* associés à deux modèles 3D (a), un cochon et une vache. Pour chaque objet, dix silhouettes sont créées sur les sommets d'un dodécaèdre. Les deux ensembles sont comparés en permutant l'ordre de mise en correspondance des vues (b, c, d, etc.) jusqu'à trouver l'appariement optimal [34].

La mesure de similarité $d_L$ entre deux *LightFields* est calculée en recherchant le meilleur appariement antre les vingt points de vues. Considérant les points de vues donnés par les vingt sommets d'un dodécaèdre (chaque sommet connecté à trois arêtes), il existe soixante rotations possibles $r$ entre les deux ensembles de points, comme le montre la Figure 4.13.

L'étude comparative [112] faite sur la base de PSB Test montre les bonnes performances du LFD dans la recherche et le place à la première position. Le point limitant ce descripteur est le coût de la comparaison entre deux objets, estimé à $60 \times N^2$ (comparaisons + $10 \times 45$ (additions + multiplications)). Pour palier ces problèmes, les auteurs proposent des mesures de similarités moins coûteuses en mettant en place une série d'études permettant d'élaguer la base de données au fur et à mesure. Une première étude, avec un descripteur de forme réduit et une mesure de similarité de faible coût, permet de définir un premier ensemble d'objets de la base, sur lequel seront exécutées les études suivantes, et ainsi de suite.

Filali Ansary et al. ont étendu leur méthode de recherche d'objets 3D par vue [9], appelée *Adaptive Views Clustering*, à la recherche par l'exemple [7, 8]. Pour réduire l'espace de la description, un sous ensemble de vues dites caractéristiques est sélectionné parmi 320 points de vue répartis uniformément sur la sphère et caractérisés par leurs moments de Zernike 2D. La sélection est basée sur un algorithme de classification adaptatif et un modèle statistique fixant le nombre optimal de vues. Partant du fait que toutes les vues d'un objet 3D ne contiennent pas la même quantité d'information, ils proposent d'utiliser une approche bayésienne dans le processus de l'indexation. Leur étude comparative réalisée sur la base de PSB Test place leur méthode juste derrière le LFD [34].

#### 4.2.1.2 Approches à base d'images de profondeur

Heczko et al. [56] ont été les premiers à proposer une approche multi-vues pour la recherche par le contenu d'objets 3D. Le même descripteur, *Depth Buffer-Based Descriptor* (DBD), a été utilisé et détaillé par Vranic dans le cadre de sa thèse [131]. Pour assurer un comportement invariant par rapport aux transformations géométriques, chaque objet 3D est tout d'abord centré, mis à l'échelle et aligné par une analyse en composantes principales continue. Par projection de l'objet sur les six faces de sa boîte englobante, six images de profondeur sont calculées, puis transformées dans l'espace de Fourier via la transformée de Fourier rapide 2D. La signature de l'objet 3D est donnée par les coefficients basse-fréquence des images correspondantes. Comparée à de nombreux autres descripteurs 3D [21, 131], cette approche montre une meilleure capacité de discrimination

des objets 3D. Elle sera détaillée dans la section 5.2.3.2.

Inspirée du descripteur DBD [131], Passalis et al. [100] proposent une version améliorée appelée PTK. Elle s'appuie sur trois idées principales : D'abord, la symétrie de l'objet est utilisée pour corriger l'alignement donné par une ACP. Ensuite, les deux images de profondeur associées aux faces opposées sont remplacées par leur différence et leur somme pour conférer à la signature une invariance aux changements de sens des axes. Enfin, des poids définis par les valeurs propres sont associés aux coefficients de Fourier. Les résultats obtenus par cette approche sur la base de PSB Test sont meilleurs que la méthode fondée sur le DBD dont ils se sont inspiré.

L'approche proposée par Vajramushti et al. [124] consiste à caractériser l'objet 3D par les six images de profondeur calculées sur les faces de sa boîte englobante. L'aire et le volume déduits de ces projections constituent la signature de l'objet 3D. Les mêmes caractéristiques sont utilisées dans un algorithme itératif qui fait varier les six degrés de liberté caractérisant la position et l'orientation de l'objet 3D. Ceci est utilisé pour estimer les erreurs de similarité entre images de profondeur. Cette approche permet de retrouver le meilleur appariement de formes 2D pour tenter de remédier aux problèmes de l'ACP conduisant à des poses erronées. Cependant, la description est trop élémentaire pour caractériser la forme globale.

Dans [92], Ohbuchi et al. proposent un descripteur 2D/3D appelé *Multiple Orientation Depth Fourier Descriptor* (MODFD). Sur un objet 3D centré et mis à l'échelle, des images de profondeur sont calculées à partir de 42 points de vue différents, répartis uniformément sur la sphère unité (les 12 sommets et les milieux des 30 arêtes d'un icosaèdre). En utilisant la méthode de Zhang [140], chaque image de profondeur dans le système de coordonnées cartésiennes $(x,\ y)$ est transformée en image de profondeur dans le système de coordonnées polaires $(r,\ \theta)$. Les basses fréquences de la transformée de Fourier de chaque image résultante représente le vecteur caractéristique de la vue, et l'ensemble des vecteurs des 42 vues constitue la signature de l'objet 3D. La similarité entre deux objets 3D, $O^1$ et $O^2$, est mesurée en minimisant la distance entre toutes les $42^2$ combinaisons

possibles des vecteurs des deux ensembles :

$$\Delta(O^1, O^2) = \frac{1}{42} \sum_{i=1}^{42} \left( \min_{1 \leq k \leq 42} \left( \sum_{j=1}^{N} |f_{ij}^1 - f_{kj}^2| \right) \right), \tag{4.12}$$

les matrices $(f^1)_{ij}$ et $(f^2)_{ij}$ étant respectivement les signatures des objets $O^1$ et $O^2$, avec $1 \leq i \leq 42$ l'indice de l'image de profondeur, $1 \leq j \leq N$ l'indice du coefficient de Fourier et $N$ la taille du vecteur caractéristique. Le fait d'utiliser à la fois cette mesure de similarité et les coordonnées polaires des images de profondeur permet d'avoir une pseudo invariance par rapport aux rotations. Cependant, cette méthode est très coûteuse en temps de calcul. Les évaluations faites par les auteurs montrent que leur méthode a les meilleurs résultats de recherche comparée aux méthodes fondées sur les distributions de forme.

Ohbuchi et al. [93] utilisent les caractéristiques visuelles locales multi-échelle pour caractériser un objet 3D. Comme dans [92], les images de profondeur sont extraites à partir des points de vue placés uniformément sur la sphère englobant l'objet 3D préalablement centré et mis à l'échelle. Des points d'intérêt sont calculés sur toutes les images de profondeur en utilisant l'algorithme de SIFT (*Scale Invariant Feature Transform*). L'ensemble des points est quantifié par des mots visuels d'un alphabet de taille $N$. Cet alphabet (*codebook*) est donné par une classification non-supervisée *k-means* des points d'intérêt extraits des modèles 3D de la base. La signature de l'objet 3D est représentée par un histogramme de taille $N$, chaque composante $k$ étant le nombre de points d'intérêt quantifiés par le mot d'indice $k$. Pour comparer deux signatures, les auteurs proposent d'utiliser la divergence de Kullback-Leibler [72], définie comme une mesure de dissimilarité entre deux distributions de probabilités. Des expériences faites sur deux bases d'objets 3D montrent que ce descripteur avec 42 points de vue, est plus performant que le LFD [34] sur la base de McGill (modèles articulés) et a des résultats comparables au LFD sur la base de Princeton (modèles généralistes).

### 4.2.2 Autres approches

D'autres approches utilisent des informations provenant des coupes [103] ou des images de spin [12], construites à partir des objets 3D. Elles sont moins répandues que les silhouettes et les images de profondeur dans les approches 2D/3D.

#### 4.2.2.1 Approches à base de coupes

Pu et al. [103] représentent un objet 3D par un ensemble de coupes 2D (*slice*) prises le long des directions du repère (cf. Figure 4.14). Pour avoir la meilleure représentation, ils considèrent deux repères intrinsèques à l'objet 3D, le premier défini par les directions principales et le deuxième par les directions d'orientation maximale, et choisissent parmi les deux celui qui minimise le volume de la boîte englobante. Le descripteur de forme 2D utilisé pour caractériser chaque coupe est la distribution de la distance entre deux points pris au hasard. Ainsi, la mesure de similarité entre deux objets 3D, $O^1$ et $O^2$, est calculée selon la formule :

$$\Delta(O^1, O^2) = \sum_{d=1}^{3} \sum_{i=1}^{N} \sum_{j=1}^{M} (h^1_{dij} - h^2_{dij})^2, \tag{4.13}$$

$\mathbf{h}^1_{di}$ et $\mathbf{h}^2_{di}$ étant les distributions de la coupe d'indice $i$ le long de la direction $d$, $N$ le nombre de coupes le long de chaque direction, $M$ la taille de l'histogramme représentant chaque coupe et $j$ l'indice de la composante de l'histogramme.

FIG. 4.14 – Trois modèles 3D et leurs coupes respectives le long d'une direction [103].

#### 4.2.2.2 Approches à base d'images de spin

Assfalg et al. [12] proposent de caractériser un objet 3D par un ensemble d'images de spin [62]. Pour chaque sommet du maillage $\mathbf{p}$, une image de spin est construite en projetant les autres sommets $\mathbf{p}_i$ (1.2) dans un espace bidimensionnel selon la formule suivante :

$$(\alpha, \beta) = (\sqrt{\|\mathbf{p}_i - \mathbf{p}\|^2 - (\mathbf{n} \cdot (\mathbf{p}_i - \mathbf{p}))^2}, \mathbf{n} \cdot (\mathbf{p}_i - \mathbf{p})), \tag{4.14}$$

$\mathbf{n}$ étant la normale de la surface sur le sommet $\mathbf{p}$. En quantifiant l'espace des paramètres $(\alpha, \beta)$, chaque pixel de l'image de spin $S_{\mathbf{p}}(i, j)$ est donné par le nombre

des sommets qui ont la même projection $(\alpha_i, \beta_j)$. La Figure 4.15 illustre quatre exemples d'images de spin calculées sur quatre sommets d'un modèle 3D. Pour représenter l'image de spin, l'espace est partitionné selon trois types de décompositions : (1) six secteurs angulaires avec $(\beta > 0)$, (2) six secteurs angulaires avec $(\beta < 0)$ et (3) six coquilles. La densité des sommets dans chaque cellule fournit trois distributions.

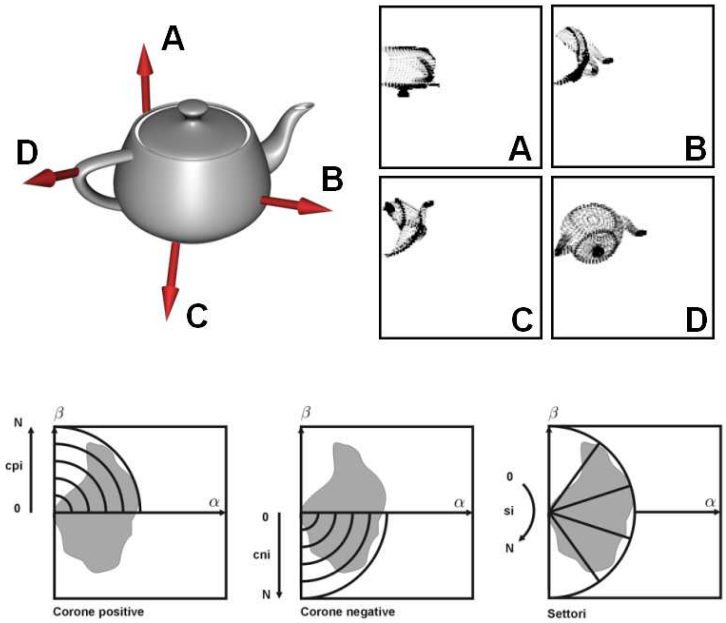



FIG. 4.15 – Exemples d'images de spin construites sur quatre points de la surface de l'objet 3D. En bas, les trois décompositions utilisées pour calculer la signature d'une spin image [12].

Pour réduire l'espace de description, les auteurs proposent d'utiliser une classification floue [14] sur l'ensemble des vecteurs caractéristiques décrivant les images de spin, et de définir la signature de l'objet 3D par les centres des classes résultantes. La similarité entre deux objets est donnée par la correspondance minimisant la somme des distances entre les centres des classes.

## 4.3 Conclusion

Dans ce chapitre, nous avons présenté les principales approches de recherche par le contenu d'objets 3D. Cette étude approfondie nous a permis d'avoir une vision claire du domaine de la recherche, et de formuler quelques remarques sur les différentes approches 3D et 2D/3D :

- La plupart des descripteurs proposés, qu'ils soient 3D ou 2D/3D, ne sont pas invariants aux transformations géométriques. D'autre part, ils sont plus ou moins sensibles au niveau de modélisation de l'objet 3D (niveau de maillage par exemple), aux petites modifications (ajout ou retrait de facettes), et plus généralement aux bruits divers. Des procédés (comme le filtrage, le débruitage, le remaillage adaptatif, la normalisation...) sont généralement mis en œuvre pour remédier à ces problèmes. Ces étapes de prétraitement présentent en revanche un comportement assez instable. Il est donc préférable de sélectionner les plus performantes et d'utiliser des signatures satisfaisant intrinsèquement ces propriétés d'invariance.

- La plupart des approches statistiques offrent l'avantage de la compacité, d'un faible coût de calcul, et d'une invariance par rapport aux transformations géométriques. En revanche, elles sont très sensibles aux modifications du maillage, à l'exception des histogrammes de cordes [97] et des distributions de forme [60, 91, 96]. Toutefois, trop élémentaires pour caractériser la notion de forme, ces distributions ne constituent pas des outils de recherche pertinents.

- Les descripteurs 3D proposés dans les approches basées sur une partition de l'espace ne vérifient pas les contraintes d'invariance géométrique. Dans ce type d'approche, un alignement par une analyse en composantes principales est calculé pour s'affranchir de ce problème. Ces méthodes ont l'avantage de fournir une représentation compacte de l'objet dans son espace. De plus, la forme de l'objet est caractérisée par une signature rapide à calculer et à comparer, ce qui permet d'utiliser ces approches sur des grandes bases et de garantir des temps de recherche courts.

- En ce qui concerne les approches par transformée, il est généralement difficile d'interpréter le lien entre la forme 3D et sa signature. Mais, comme la plupart des descripteurs sont déduits de la projection d'une représentation de l'objet 3D sur un ensemble de fonctions mathématiques, l'information extraite peut être vue comme une caractéristique de ces fonctions. Ces méthodes ne gardent aucune information topologique et se limitent à une comparaison de vecteurs caractérisant des formes transformées dans un autre espace. Toutefois, elles ont un coût de calcul de descripteur et de mesure de similarité faibles.

- Le choix de la discrétisation de l'espace de description est un problème délicat dans l'indexation des objets 3D. Certaines méthodes utilisent des descripteurs à base d'une représentation sphérique selon les paramètres $(\theta, \varphi)$ qui peut conduire à des calculs erronés, dus à la surreprésentation de l'information aux pôles et au favoritisme de l'axe des $z$.

- Contrairement aux autres approches 3D, les approches structurelles visent à décrire la notion de forme de manière plus intuitive et plus compréhensive. Comme nous l'avons montré dans la section 4.1.4, une représentation élaborée par squelette ou graphe permet de capter à la fois des informations topologiques et géométriques de la forme, conduisant à une description complète de haut niveau. En revanche, ces méthodes ne peuvent pas être appliquées sur tous les objets, en particulier les objets surfaciques mal définis, car l'extraction du graphe est trop sensible aux perturbations de la surface de l'objet 3D. D'autre part, ces méthodes ne peuvent pas être appliquées sur des grandes bases d'objets 3D, car l'extraction des graphes et la mesure de similarité associée sont très coûteuses en temps de calcul, ce qui les rend difficilement exploitables dans le cadre des objectifs spécifiques de l'indexation. Notons que la plupart des approches structurelles peuvent être utilisées dans la recherche par partie.

- Les approches multi-vues à base de silhouettes, [76] (7 silhouettes) et [82, 131] (3 silhouettes), reposent sur le même principe en utilisant un alignement par une ACP, avec des descripteurs de forme 2D différents. Dans [8, 34], l'ACP est remplacée par une discrétisation angulaire plus fine de l'espace de vues : le premier propose une mesure de similarité quasi-invariante aux rotations alors que le deuxième fait appel à un algorithme de classification pour extraire les vues caractéristiques et un processus bayésien dans la mesure de similarité. Les deux méthodes se sont avérées très coûteuses en temps de calcul.

- En ce qui concerne les approches à base d'images de profondeur, le descripteur DBD [56] et sa version améliorée PTK [100] se révèlent être puissants pour retrouver les objets de formes similaires en un temps raisonnable. L'algorithme de régulation du positionnement de l'objet 3D introduit par Vajramushti et al. [124] permet d'effectuer le meilleur appariement entre deux ensembles d'images de profondeur, mais leur descripteur 2D présente des faiblesses dans la caractérisation de la forme pour discerner les classes d'objets 3D. Les deux méthodes proposées

par Ohbuchi ne font pas appel à un prétraitement de l'objet 3D par une ACP, mais à une mesure de similarité optimisée dans [93] et à une signature globale sous forme d'histogramme de points d'intérêts quantifiés dans [92], les deux permettant d'obtenir une invariance par rapport aux rotations.

Pour ces raisons, nos études se sont portées essentiellement sur les approches 2D/3D. Une première idée simple venant du fait que toutes les vues n'ont pas la même pertinence dans la description de la forme 3D, nous ont poussé à mettre en place une nouvelle technique pouvant s'appliquer à la plupart des approches 2D/3D. D'autre part, vu les limitations des méthodes d'indexation par vues, nous avons proposé une nouvelle approche à base d'images de profondeur. La méthode est fondée sur une extraction d'information des lignes de profondeur fournissant une représentation de plus haut-niveau, et sur une mesure de similarité permettant une invariance aux petites déformations locales.

Un nouveau descripteur basé sur une partition de l'espace, appelé *3D Gaussian Descriptor*, a été également proposé. Fondé sur la transformée de Gauss, ce descripteur offre une caractérisation compacte et pertinente de l'objet 3D. Il est associé à une étape de normalisation complète lui conférant un comportement d'invariance géométrique.

# Chapitre 5

# Indexation d'objets 3D

L'indexation 3D est considérée comme le cœur du processus de la recherche par le contenu d'objets 3D. Le principe consiste à caractériser la forme des objets de manière compacte et pertinente pour en déduire une signature. La similarité entre objets 3D est alors mesurée en comparant leurs signatures. Dans ce chapitre, nous présentons les descripteurs de forme introduits dans cette thèse à savoir : un descripteur extrait directement de la géométrie de l'objet, un deuxième caractérisant les images de projection, et un troisième décrivant les lignes de profondeur.

La section 5.1 porte sur l'indexation d'objets à l'aide d'un nouveau descripteur de forme 3D, appelé *3D Gaussian Descriptor*. Dérivé de la transformée gaussienne, ce descripteur fournit une représentation définie sur une grille 3D, en exploitant les caractéristiques géométriques locales de la surface du modèle 3D.

Dans la section 5.2, nous nous intéressons plus particulièrement à l'indexation 2D/3D. Pour pallier les limitations de la représentation multi-vues que procure les descripteurs de forme, nous proposons une amélioration du processus général de la recherche par le contenu. L'approche améliorée consiste à calculer des vecteurs caractéristiques avec un descripteur de forme 2D/3D, puis à associer des pertinences aux différentes projections 2D, qui seront par la suite injectées dans l'estimation des similarités. Ceci permet d'obtenir une description plus équilibrée et de parvenir à des calculs de similarité plus fins.

La section 5.3 est consacrée aux approches 2D/3D basées sur les lignes de profondeur. Trois approches calculant des lignes de profondeur sont proposées : (1) *depth line approach*, (2) *multi-view depth line approach* et (3) *cylindrical depth*

*line approach*. Une méthode de transcription de lignes de profondeur, permettant une indexation compacte et pertinente, est ensuite décrite. Plusieurs mesures de similarité adaptées aux séquences sont présentées.

Toutes les méthodes d'indexation proposées dans ce chapitre ne sont pas invariantes aux transformations affines. C'est pourquoi nous faisons appel à l'étape complète de normalisation. Pour ce faire, le centre de gravité $\mathbf{g}$ (2.1) de l'objet 3D est déplacé à l'origine du repère (cf. section 2.2.1.1), les directions fournies par notre méthode d'alignement (cf. algorithme 3.3 ) sont alignées avec ceux du repère, et la taille de l'objet est normalisée en utilisant le facteur d'échelle $d_{moy}$ (2.10) (cf. section 2.2.3.1).

Après avoir spécifié les mesures de similarité associées aux descripteurs de forme, utilisées lors des applications de requête, les performances de la recherche par le contenu sont évaluées et comparées objectivement sur les bases de données décrites dans la section 1.4. Ces résultats seront présentés dans le chapitre 6.

## 5.1 Approche 3D basée sur le descripteur gaussien

Un nouveau descripteur de forme 3D appelé *3D Gaussian Descriptor* (3DGD) est proposé dans cette section. Il fait partie de la famille des descripteurs basés sur une partition de l'espace (cf. section 4.1.2.2). Le principe de ce descripteur est de caractériser et d'amplifier localement le voisinage de la surface 3D. Pour cela, nous introduisons des fonctions gaussiennes qui mesurent l'influence des points de la surface sur des points régulièrement répartis dans l'espace englobant l'objet 3D. Ce descripteur offre une caractérisation compacte, robuste et attachée à la forme 3D. Combinée à une mesure de similarité adéquate et performante, nous montrons que la fidélité de description améliore la précision des recherches par rapport aux approches 3D existantes.

### 5.1.1 *3D gaussian descriptor*

#### 5.1.1.1 Définition de descripteur gaussien 3D

Dérivée de l'expression de la loi gaussienne, la transformée gaussienne $g$ est une application réelle définie sur un ensemble de points de l'espace, et obtenue par une sommation sur la surface de l'objet 3D. Soient $\mathfrak{S}$ la surface de l'objet et $\sigma$ la bande. Nous définissons la transformée gaussienne sur un point $\mathbf{q}$ de l'espace par la formule suivante :

$$g(\mathbf{q}, \mathfrak{S}, \sigma) = \iint_{\mathbf{p} \in \mathfrak{S}} e^{-d^2(\mathbf{p},\mathbf{q})/\sigma^2} \, ds, \tag{5.1}$$

où $d$ est la distance euclidienne du point $\mathbf{q}$ au point $\mathbf{p}$ de $\mathfrak{S}$ (1.4).

La transformée gaussienne est utilisée pour définir un descripteur 3D défini comme suit : Soient $\mathcal{M}$ un objet 3D préalablement centré, aligné et mis à l'échelle et $B$ une boîte englobant $\mathcal{M}$, définie par l'ensemble des sommets $\{(x, y, z) | x \in \{x_m, x_p\}, y \in \{y_m, y_p\}, z \in \{z_m, z_p\}\}$. Nous calculons un échantillonnage volumique de la boîte $B$ selon les paramètres ($x$, $y$, $z$), pour obtenir $N^3$ points $\mathbf{q}_{ijk} = (x_i, y_j, z_k)$, centres des cellules de taille égale. Ils sont regroupés dans les ensembles suivants :

- $\mathcal{X} = \{x_i = x_m + (i + 0.5)\Delta_x\}$, où $\Delta_x = (x_p - x_m)/N$,
- $\mathcal{Y} = \{y_j = y_m + (j + 0.5)\Delta_y\}$, où $\Delta_y = (y_p - y_m)/N$,
- $\mathcal{Z} = \{z_k = z_m + (k + 0.5)\Delta_z\}$, où $\Delta_z = (z_p - z_m)/N$,

avec $i$, $j$ et $k \in \{0, ..., N-1\}$. Quatre types de boîtes (cf. Annexe A.1.1) ont été évalués lors des expérimentations.
Soit $g$ l'application donnée par l'équation (5.1), définie sur l'ensemble $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ dans $\mathbb{R}$. Nous définissons les coefficients du descripteur gaussien 3D par :

$$g_{ijk} = g(\mathbf{q}_{ijk}, \mathfrak{S}, \sigma) = \iint_{\mathbf{p} \in \mathfrak{S}} e^{-d^2(\mathbf{p}, \mathbf{q}_{ijk})/\sigma^2} \, ds, \tag{5.2}$$

Le coefficient $g_{ijk}$ peut être vu comme la somme des contributions des points de la surface dans la caractérisation locale autour du point $\mathbf{q}_{ijk}$.

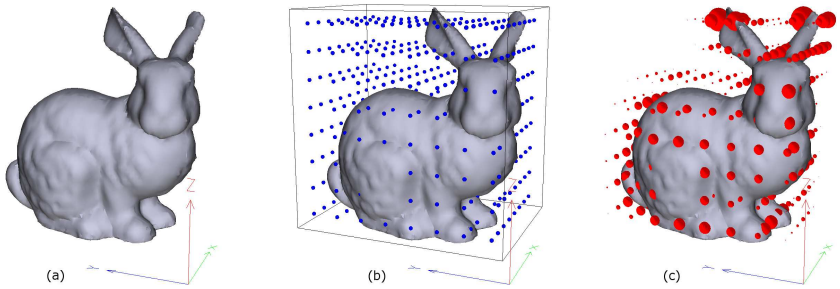



FIG. 5.1 – Exemple de calcul de descripteur 3DGA : (a) modèle "bunny" normalisé, (b) échantillonnage de la boîte englobante (N = 8), les points $\mathbf{q}_{ijk}$ sont colorés en bleu, (c) les rayons des boules rouges sont proportionnels aux coefficients $g_{ijk}$.

Dans la suite, la contribution d'un point $\mathbf{p}$ de $\mathfrak{S}$ dans la description locale $g_{ijk}$ sera notée $e^{\mathbf{p}}_{ijk}$ et exprimée par :

$$e^{\mathbf{p}}_{ijk} = e^{-d^2(\mathbf{p}, \mathbf{q}_{ijk})/\sigma^2}. \tag{5.3}$$

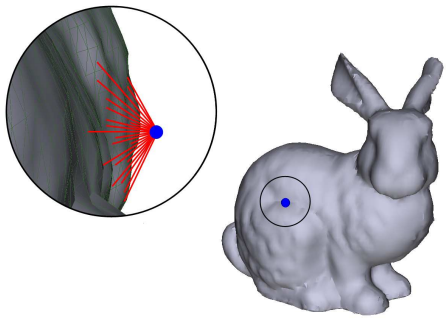


FIG. 5.2 – Contributions des points de la surface $\mathfrak{S}$, voisins à $\mathbf{q}_{ijk}$, dans la description locale $g_{ijk}$

#### 5.1.1.2 Approximation discrète

Le descripteur gaussien 3D repose sur un principe d'accumulation de points de la surface $\mathfrak{S}$ de l'objet 3D sur un ensemble fini de points dans $\mathbb{R}^3$, donné par l'équation (5.2). Le calcul de l'intégrale étant trop complexe, nous proposons une approximation discrète de $\mathfrak{S}$ par un ensemble $E$ fini de points, $\mathbf{p}$, appartenant à $\mathfrak{S}$. Dans ce cas, les coefficients du descripteur sont calculés comme suit :

$$g_{ijk} = \sum_{\mathbf{p} \in E} e_{ijk}^{\mathbf{p}}. \tag{5.4}$$

Si l'on prend comme ensemble $E_T = \bigcup_{i=1}^{N_T} \mathbf{g}_i$ un nuage de points de $\mathfrak{S}$ défini par les centres de gravité $\mathbf{g}_i$ (1.5) de toutes les facettes $T_i$ de $\mathfrak{T}$ (1.1), l'approximation est sensible au changement de résolution du maillage et à la taille des triangles et de leur répartition sur le maillage. Pour pallier cet inconvénient, nous avons cherché à améliorer l'approximation sur deux points : l'échantillonnage uniforme sur la surface et la pondération des contributions par l'aire des facettes.

*Échantillonnage uniforme :*
En pratique, les maillages des modèles 3D sont constitués de facettes de taille variable. De plus, les arêtes peuvent avoir le même ordre de grandeur que les pas de quantification, $\Delta_x$ $\Delta_y$ et $\Delta_z$, qui définissent la grille 3D sur laquelle est construit le descripteur. Cela peut conduire à des estimations erronées des coefficients $g_{ijk}$. Pour une meilleure précision, nous proposons d'effectuer un rééchantillonnage du maillage 3D. La solution la plus adaptée consiste à prélever des points répartis uniformément sur la surface du maillage, et associés à des facettes suffisamment petites. Cette approximation de la surface $\mathfrak{S}$ est obtenue par une subdivision pseudo uniforme de chaque triangle $T_i$ de $\mathfrak{T}$ (1.1) en $n_i$ triangles $\{T_i^j\}_{1..n_i}$ d'aire quasiment identiques. Le nouvel ensemble $E_U$ est défini comme suit :

$$E_U = \bigcup_{i=1}^{N_T} \bigcup_{j=1}^{n_i} \mathbf{g}_i^j, \tag{5.5}$$

$\mathbf{g}_i^j$ étant le centre de gravité du triangle $T_i^j$.
Si $n_{min}$ est le nombre minimal d'échantillons pris sur un maillage, l'aire du triangle $T_i^j$ est alors inférieure à $\mathcal{A}/n_{min}$ ($n_{min}$ est une valeur suffisamment grande). Pour chaque facette $T_i$, $n_i$ est calculé selon la formule suivante :

$$n_i = \left[ \frac{A_i}{\mathcal{A}} \, n_{min} \right],$$

$A_i$ (1.5) et $\mathcal{A}$ (1.6) étant respectivement l'aire de $T_i$ et de $\mathfrak{S}$.

*Pondération par la surface :*
La deuxième amélioration consiste à choisir une contribution proportionnelle à l'aire de la facette considérée pour pouvoir gérer des modèles irrégulièrement maillés et en particulier, des modèles contenant des faces dont l'aire est inférieure à $\mathcal{A}/n_{min}$. Nous proposons donc de définir la contribution $e^{\mathbf{p}}_{ijk}$ dans le cas discret comme suit :

$$e^{\mathbf{p}}_{ijk} = A_{\mathbf{p}}\ e^{-d^2(\mathbf{p},\mathbf{q}_{ijk})/\sigma^2}, \tag{5.6}$$

où $A_{\mathbf{p}}$ est l'aire de la facette associée au centre de gravité $\mathbf{p}$.

Le descripteur $\mathbf{G} = [g_{ijk}]$ peut être construit en estimant successivement les contributions des points $\mathbf{p}$ de $E$ dans les descriptions locales autour des points $\mathbf{q}_{ijk}$ définis précédemment. Ceci peut s'avérer coûteux, de l'ordre $\mathcal{O}(|E|N^3)$. Une étude du terme gaussien et un réordonnancement du calcul permettent d'accélérer le processus.

#### 5.1.1.3 Optimisation de la construction du descripteur

Avant de donner les grandes lignes de l'algorithme de calcul du descripteur, examinons le terme gaussien de la contribution d'un point $\mathbf{p}$ de $E$ dans la description locale $g_{ijk}$ (5.4). Si $\mathbf{p}$ est dans un voisinage très proche de $\mathbf{q}_{ijk}$, $e^{\mathbf{p}}_{ijk}$ (5.6) prend une valeur positive non nulle, et l'accumulateur $g_{ijk}$ est incrémenté. En revanche, si $\mathbf{p}$ est éloigné de $\mathbf{q}_{ijk}$, $e^{-d^2(\mathbf{p},\mathbf{q}_{ijk})/\sigma^2}$ est presque nul (c'est-à-dire $e^{-d^2(\mathbf{p},\mathbf{q}_{ijk})/\sigma^2} < \epsilon$, $\epsilon$ un seuil très proche de zéro à fixer). Dans ce cas, $\mathbf{p}$ ne contribue pas à la description autour du point $\mathbf{q}_{ijk}$, et l'accumulateur $g_{ijk}$ reste inchangé. Ce test sur l'appartenance au voisinage de $\mathbf{q}_{ijk}$ est une solution qui permet d'éviter de calculer tous les $e^{\mathbf{p}}_{ijk}$ et d'accélérer considérablement les calculs.

Supposons le seuil $\epsilon > 0$ fixé et considérons la boule voisinage de centre $\mathbf{q}_{ijk}$ :

$$\mathcal{B}_{ijk} = \{\mathbf{q} \in \mathbb{R}^3 | d(\mathbf{q}, \mathbf{q}_{ijk}) \leq \sigma\sqrt{-\ln(\epsilon)}\}. \tag{5.7}$$

Les points $\mathbf{p}$ de $E$ appartenant à cette boule vérifient l'équation $e^{\mathbf{p}}_{ijk} \geq \epsilon$, et la contribution du point $\mathbf{p}$ dans la description locale $g_{ijk}$ peut donc s'exprimer comme suit :

$$e^{\mathbf{p}}_{ijk} = \delta_{\mathcal{B}_{ijk}}(\mathbf{p})\ A_{\mathbf{p}}\ e^{-d^2(\mathbf{p},\mathbf{q}_{ijk})/\sigma^2}, \tag{5.8}$$

avec $\delta_{\mathcal{B}_{ijk}}$ la fonction d'appartenance d'un point $\mathbf{p}$ à l'ensemble $\mathcal{B}_{ijk}$, définie par :

$$\delta_{\mathcal{B}_{ijk}}(\mathbf{p}) = \begin{cases} 1 & \text{si } \mathbf{p} \in \mathcal{B}_{ijk}, \\ 0 & \text{si } \mathbf{p} \notin \mathcal{B}_{ijk}. \end{cases} \tag{5.9}$$

Donc, si nous introduisons un seuil $\epsilon$, tout point extérieur à la boule $\mathcal{B}_{ijk}$ ne contribue pas à $g_{ijk}$. Il est donc judicieux d'échanger l'ordre des calculs et d'examiner d'abord les points de l'ensemble $E$ pour calculer leurs contributions dans les $g_{ijk}$, en fonction de leur localisation.

Pour calculer les contributions d'un point $\mathbf{p}$ de $E$, nous localisons d'abord $\mathbf{p}$ par rapport aux boules $\mathcal{B}_{ijk}$ en comparant successivement les coordonnées en $x$, $y$ et $z$ de $\mathbf{p}$ par rapport à ceux des $\mathbf{q}_{ijk}$. Quand $\mathbf{p} = (x, y, z) \in \mathcal{B}_{ijk}$, nous avons :

$$|x - x_i|^2 \leq r^2,$$

$$|y - y_j|^2 \leq r^2 - |x - x_i|^2,$$

$$|z - z_k|^2 \leq r^2 - |x - x_i|^2 - |y - y_i|^2,$$

$r = \sigma\sqrt{-\ln(\epsilon)}$ étant le rayon de la boule.
Donc, pour sélectionner les boules candidates contenant le point $\mathbf{p}$, il suffit de calculer les bornes inférieures et supérieures limitant l'espace de ces boules, ce qui donne, en considérant les indices $i$, $j$ et $k$ des $\mathbf{q}_{ijk}$, la propriété suivante :
Le point $\mathbf{p}$ intervient dans le calcul des $g_{ijk}$ quand

$$i \in \{i^d, ..., i^f\}, \text{ avec } \begin{array}{l} x_{i^f-1} < x + r \leq x_{i^f} \text{ et} \\ x_{i^d} \leq x - r < x_{i^d+1}, \end{array}$$

$$j \in \{j_i^d, ..., j_i^f\}, \text{ avec } \begin{array}{r} y_{j_i^f-1} < y + \sqrt{r^2 - |x - x_i|^2} \leq y_{j_i^f} \text{ et} \\ y_{j_i^d} \leq y - \sqrt{r^2 - |x - x_i|^2} < y_{j_i^d+1}, \end{array}$$

$$k \in \{k_{ij}^d, ..., k_{ij}^f\}, \text{ avec } \begin{array}{r} z_{k_{ij}^f-1} < z + \sqrt{r^2 - |x - x_i|^2 - |y - y_i|^2} \leq z_{k_{ij}^f} \text{ et} \\ z_{k_{ij}^d} \leq z - \sqrt{r^2 - |x - x_i|^2 - |y - y_i|^2} < z_{k_{ij}^d+1}. \end{array}$$

Ce qui revient, dans la construction algorithmique, à définir pour tout $\mathbf{p}$ de $\mathfrak{S}$ une plage d'indices $(i, j, k) \in \{i^d, ..., i^f\} \times \{j_i^d, ..., j_i^f\} \times \{k_{ij}^d, ..., k_{ij}^f\}$ correspondant aux $g_{ijk}$ qui sont influencés par $\mathbf{p}$.

#### 5.1.1.4 Normalisation du descripteur

La formulation de la contribution dans le cas discret est trop fortement influencée par les aires des facettes, car il existe des formes 3D qui présentent les mêmes caractéristiques locales, mais pour des aires de facettes très différentes. Pour limiter ce biais, nous proposons de sommer les aires des facettes associées aux points de $\mathcal{B}_{ijk}$ (5.7) contribuant à $g_{ijk}$ et normaliser ce dernier par rapport à cette somme :

$$A_{ijk} = \sum_{\mathbf{p} \in E} \delta_{\mathcal{B}_{ijk}}(\mathbf{p})\ A_{\mathbf{p}}\ , \tag{5.10}$$

$$g_{ijk} = \frac{1}{A_{ijk}} \sum_{\mathbf{p} \in E} e_{ijk}^{\mathbf{p}}\ ,$$

$\delta_{\mathcal{B}_{ijk}}$ étant la fonction d'appartenance définie dans l'équation (5.9).

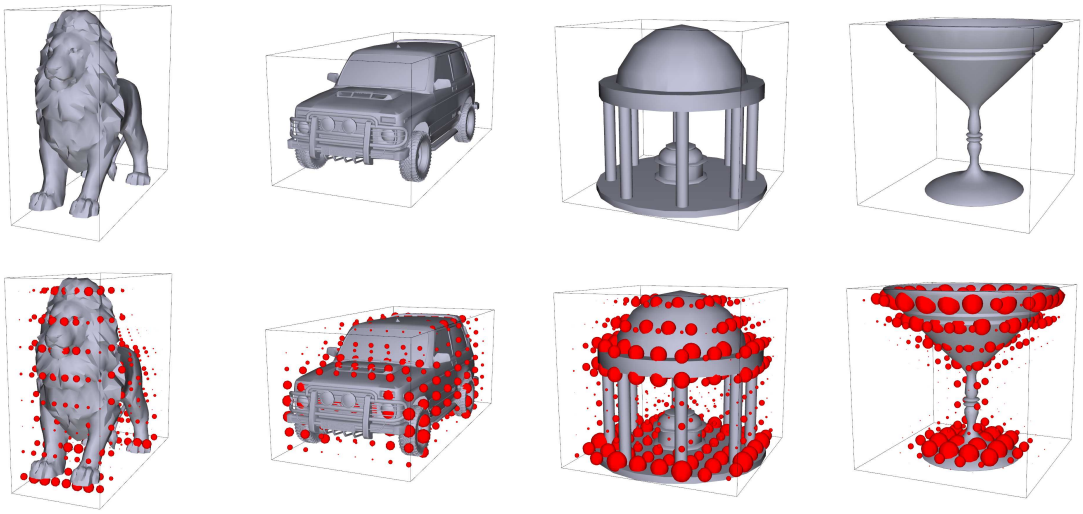


FIG. 5.3 – Exemples de descripteurs 3DGA calculés sur quatre modèles différents. Les rayons des boules rouges sont proportionnels aux coefficients du descripteur.

#### 5.1.1.5 Calcul du descripteur

Étant donné $E$ l'ensemble de points pris sur la surface de l'objet 3D (cf. section 5.1.1.2), les grandes lignes de l'algorithme calculant le descripteur gaussien 3D se résument ainsi :

**Algorithme 5.1** 3D Gaussian Descriptor

$r = \sigma\sqrt{-\ln(\epsilon)}$
**Pour tout** $i, j, k$, $0 \leq i, j, k < N - 1$ **faire**
  initialiser $g_{ijk} = A_{ijk} = 0$
**Fin pour**
**Pour tout** $\mathbf{p} \in E$ **faire**
  calculer $i^d$ et $i^f$
  **Pour tout** $i$, $i^d \leq i \leq i^f$ **faire**
    $d^2_{x_i} = (x - x_i)^2$
    **Si** $d^2_{x_i} \leq r^2$ **alors**
      calculer $j^d_i$ et $j^f_i$
      **Pour tout** $j$, $j^d_i \leq j \leq j^f_i$ **faire**
        $d^2_{y_j} = (y - y_j)^2$
        **Si** $d^2_{x_i} + d^2_{y_j} \leq r^2$ **alors**
          calculer $k^d_{ij}$ et $k^f_{ij}$
          **Pour tout** $k$, $k^d_{ij} \leq k \leq k^f_{ij}$ **faire**
            $d^2_{z_k} = (z - z_k)^2$
            $d^2(\mathbf{p}, \mathbf{q}_{ijk}) = d^2_{x_i} + d^2_{y_j} + d^2_{z_k}$
            $A_{ijk} += A_{\mathbf{p}}$
            $g_{ijk} += A_{\mathbf{p}}\ e^{-d^2(\mathbf{p}, \mathbf{q}_{ijk})/\sigma^2}$
          **Fin pour**
        **Fin si**
      **Fin pour**
    **Fin si**
  **Fin pour**
**Fin pour**
**Pour tout** $i, j, k$, $0 \leq i, j, k < N - 1$ **faire**
  Normaliser $g_{ijk} / = A_{ijk}$
**Fin pour**

### 5.1.2 Mesure de similarité adoptée dans 3DGA

Pour comparer deux objets $O^1$ et $O^2$, nous générons, pour chacun, le descripteur gaussien associé, $\mathbf{G}_1 = [g^1_{ijk}]$ et $\mathbf{G}_2 = [g^2_{ijk}]$. La mesure de similarité peut être calculée en utilisant une distance euclidienne dans l'espace des coefficients de Gauss comme suit :

$$\Delta(O^1, O^2) = d(\mathbf{G}_1, \mathbf{G}_2)\ , \tag{5.11}$$

$d$ étant la distance euclidienne $l_1$ ou $l_2$.

Pour être moins sensible aux petits déplacements ou aux petites variations, nous introduisons une nouvelle mesure de similarité qui minimise la distance entre

les paires de composantes voisines :

$$\begin{aligned}\Delta(O^1, O^2) &= d_v(\mathbf{G}_1, \mathbf{G}_2) \\ &= \sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\sum_{k=0}^{N-1} \min \begin{cases} |g^1_{ijk} - g^2_{ijk}|, \\ \min |g^1_{ijk} - g^2_{i'j'k'}| \ _{pour|i-i'|+|j-j'|+|k-k'|=1}, \\ \min |g^2_{ijk} - g^1_{i'j'k'}| \ _{pour|i-i'|+|j-j'|+|k-k'|=1}. \end{cases}\end{aligned} \tag{5.12}$$

Il est intéressant de remarquer que cette mesure de similarité ainsi définie est symétrique.

## 5.2 Amélioration d'approches 2D/3D

Nous présentons dans cette section une technique pouvant s'appliquer à la plupart des approches 2D/3D qui caractérisent l'objet 3D à partir de ses vues 2D. Les méthodes multi-vues décrites dans la section 4.2.1 ne prennent pas en compte la différence de quantité d'information provenant des différentes vues d'un même objet. Ceci conduit à des calculs de similarité biaisés. C'est pourquoi nous proposons d'introduire des indices de pertinence dans ces approches 2D/3D pour améliorer l'estimation des similarités des modèles. Cette approche a fait l'objet d'une publication dans *International Conference on Shape Modeling and Applications 2006* [26].

### 5.2.1 Principe

En examinant les vues associées à un objet 3D, nous remarquons que certaines vues contiennent plus d'informations pertinentes que d'autres. Il convient toutefois de nuancer ce propos en rappelant que la reconnaissance d'une forme 3D est affectée en changeant le point de vue de l'observateur. Par exemple, dans la Figure 5.4, le modèle humain est reconnu plus rapidement et assurément par le système visuel via la vue de face, un peu moins via la vue de profil. Par contre, il est plus difficile de reconnaître le modèle à partir de sa vue ascendante. Ainsi, toutes les vues d'un objet 3D ne sont pas équivalentes dans la reconnaissance. Ceci est étroitement lié aux différences d'informations pertinentes contenues dans les vues. Cette problématique se pose dans le cadre d'une similarité entre objets 3D par une description multi-vues.

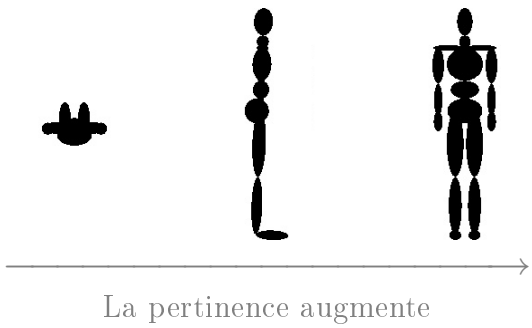



FIG. 5.4 – Différence de la pertinence entre les vues dans la description de la forme 3D. Il est plus intuitif d'identifier l'homme à partir de la deuxième ou la troisième image que de la première.

Dans la plupart des approches 2D/3D, toutes les projections (qu'elles soient des silhouettes ou des images de profondeur) sont traitées de la même façon lors de l'évaluation de la similarité. Dans cette section, nous présentons un processus permettant d'évaluer différemment les images en fonction de leurs pertinences. Pour ce faire, nous avons introduit un indice de pertinence qui reflète l'importance des informations contenues dans l'image : plus l'image contient d'informations pertinentes, plus son indice de pertinence est élevé. Un indice de pertinence est alors attribué à chaque image de projection caractérisant le modèle 3D. Remarquons que la problématique traitée ici se rapproche de celle de la recherche de critères permettant de choisir le meilleur point de vue associée à un objet 3D [102]. Dans cet article, les auteurs montrent que chaque critère considéré est insuffisant pour résoudre le problème dans tous les cas.

Plusieurs méthodes peuvent être envisagées pour définir la pertinence d'une image de projection dans la description de l'objet 3D. Ceci dépend essentiellement de la complexité et de la nature des informations extraites de la forme 2D comme la surface, le contour, la courbure, la profondeur, la structure, la connectivité, etc. Pour quantifier la pertinence, nous pouvons utiliser des mesures statistiques telles que la moyenne, la variance ou l'entropie. Toutefois, il est possible de définir une pertinence hybride intégrant différents aspects. Il suffit alors de combiner différents indices de pertinence.

Notre technique d'amélioration s'intègre dans le processus classique de la recherche par le contenu. Elle intervient dans la caractérisation des vues de l'objet 3D et dans la mesure de similarité entre deux objets 3D. Sur l'ensemble des images de projections représentant l'objet, des indices de pertinence sont calculés. En couplant ces indices avec les caractéristiques des images associées, nous obtenons une description homogène et équilibrée de la forme 3D. Comparer deux objets 3D consiste à utiliser une distance pondérée par les indices de pertinence représentant leurs poids dans la description de la forme 3D.

Comme nous l'avons déjà mentionné, notre technique peut s'appliquer à la plupart des méthodes multi-vues qui consistent à caractériser l'objet 3D par ses projections 2D. Le calcul d'indice de pertinence est introduit pour deux signatures, la première basée sur une indexation par silhouette et la deuxième sur une indexation par images de profondeur.

### 5.2.2 Indices de pertinence

L'indice de pertinence d'une image de projection indique son degré ou son niveau de pertinence dans la description de l'objet 3D. Pour le caractériser, nous exploitons des heuristiques utilisant des propriétés géométriques de la forme projetée. Dans cette section, nous proposons des modèles d'indices de pertinence associés aux silhouettes et aux images de profondeur, les deux principaux types d'images déduites d'une projection 2D. Chaque modèle vise à caractériser la pertinence de l'image selon un critère donné.

#### 5.2.2.1 Indices de pertinence associés aux silhouettes

Plusieurs indices de pertinence peuvent être associés aux silhouettes :

*1. Surface :* Un premier modèle évalue la taille de la surface projetée. Il s'agit de calculer le nombre de pixels noirs de la silhouette afin d'estimer approximativement la surface de la projection de l'objet 3D sur l'image :

$$R_a = card\{s_{ij} | s_{ij} = 0, 0 \leq i, j \leq N-1\}, \tag{5.13}$$

$s_{ij}$ étant la valeur du pixel à la position $(i, j)$ de l'image silhouette $\mathbf{s}$.
L'image-silhouette et l'image de profondeur présentent la même surface lors de la projection de l'objet 3D. Cet indice de pertinence peut être retenu pour les images de profondeur.

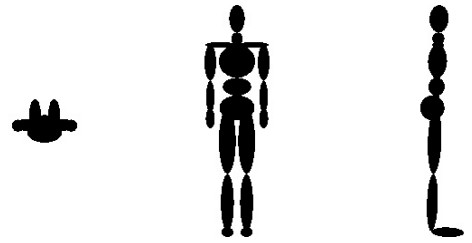
FIG. 5.5 – Limitation de l'indice de pertinence fondé sur la surface.

Cette méthode simple s'est révélée inappropriée pour mesurer la pertinence dans certains cas. L'ensemble des résultats expérimentaux présentés sur la Figure 5.5 montre que les indices de pertinence les plus élevés correspondent aux images de la

première ligne. Mais, comme nous pouvons le constater, les images de la deuxième ligne sont plus pertinentes que les précédentes. Pour contourner ces limitations, nous proposons une version dérivée de $R_a$ en prenant sa racine carrée :

$$R_{sa} = \sqrt{R_a}. \tag{5.14}$$

Cette nouvelle pertinence atténue l'influence de la taille de la surface projetée dans l'évaluation de la pertinence. Pour modérer l'influence du secteur, (voir les valeurs de $R_{sa}$ pour le modèle d'un humain sur la Figure 5.6).



21.79 46.81 31.40

FIG. 5.6 – Exemples d'indices de pertinence $R_{sa}$ (%) associés aux trois silhouettes obtenus après projection du modèle sur les faces de sa boîte englobante.

*2. Contour :* Le contour est un attribut important de la forme décrite par la silhouette. Sa longueur $L$ pourrait être utilisée dans le calcul de pertinence. Cependant, les résultats montrent que le contour et sa longueur sont très sensibles à des perturbations locales de la forme 2D. Comme nous pouvons le constater sur la Figure 5.7, des formes semblables peuvent avoir des contours de longueurs différentes. Donc, il n'est pas toujours possible de représenter la pertinence d'une forme par la longueur de son contour.

Pour définir un indice de pertinence plus robuste et qui dépend du contour, nous choisissons la corde moyenne, c'est-à-dire la moyenne des longueurs de toutes les cordes possibles reliant deux points contour.

$$R_c = \frac{1}{L(L-1)} \sum_{p=0}^{L-1} \sum_{q=0}^{L-1} \sqrt{|i_p - i_q|^2 + |j_p - j_q|^2}, \tag{5.15}$$

où $(i_p, j_p)$ appartiennent à l'ensemble $\mathbf{S}_L$ (5.20) défini par les positions de ces points contour, et $L$ est la longueur du contour.

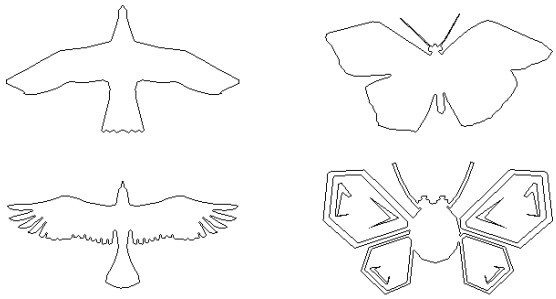


FIG. 5.7 – Limitation du contour de la silhouette dans la définition de la pertinence. Deux exemples de modèles de la même classe ayant des contours de longueurs différentes.

Notons enfin qu'il est possible de définir une pertinence calculée sur des points particuliers détectés sur la forme 2D. Par exemple, des points d'intérêt [74] ou des points de forte courbure peuvent être exploités pour éviter de considérer tous les points du contour. Toutefois, de telles représentations fondées sur ce type d'information se révèlent également imprécises pour certains modèles 3D, en raison de l'instabilité de comportement de ces points. Ces aspects ne seront cependant pas abordés ici.

#### 5.2.2.2 Indices de pertinence associés aux images de profondeur

Cette partie présente trois méthodes que nous envisageons pour calculer une pertinence associée à une image de profondeur :

*1. Surface :* L'indice de pertinence $R_a$ défini dans l'équation 5.13 peut être retenu dans le cas des images de profondeur. Il caractérise la surface de la projection de l'objet 3D sur le plan. Il est donné par le nombre des pixels ayant des valeurs positives :

$$R_a = card\{v_{ij} | v_{ab} > 0, 0 \leq a, b \leq N-1\}, \tag{5.16}$$

$v_{ab}$ étant la valeur du pixel à la position $(a, b)$ de l'image de profondeur $\mathbf{v}$.

*2. Profondeur :* La deuxième méthode introduit la notion de profondeur. Elle est la somme de tous les pixels non nuls sur l'image. En d'autres termes, elle correspond au volume compris entre les parties visibles de l'objet 3D et la face

de projection :

$$R_d = \sum_{a=0}^{N-1} \sum_{b=0}^{N-1} v_{ab}, \tag{5.17}$$

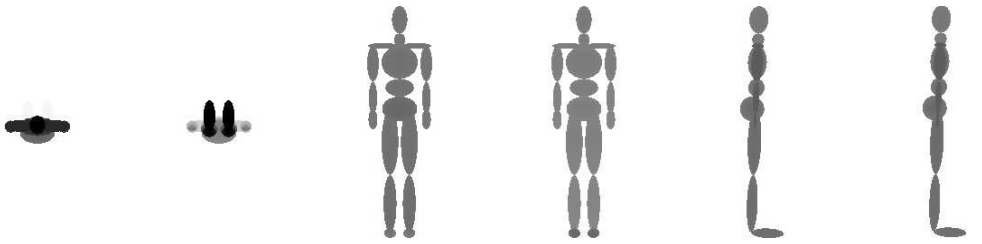
$v_{ab}$ étant la valeur du pixel à la position $(a, b)$ de l'image de profondeur $\mathbf{v}$.

*3. Global :* La troisième pertinence est plus globale. Elle est égale à la somme des distances des points de l'objet visibles sur l'image de profondeur avec le centre de la boîte englobante. Cet indice peut être calculé directement à partir de l'image de profondeur :

$$R_g = \frac{1}{2w} \sum_{a=0}^{N-1} \sum_{b=0}^{N-1} d_{ab}, \tag{5.18}$$

$$d_{ab} = \sqrt{|a - N/2|^2 + |b - N/2|^2 + w|v_{ab} - 1/2|^2},$$

où $\delta$ et $\delta'$ sont les distances signées des plans de projection $\pi$ et $\pi'$ au centre de l'objet. Les valeurs numériques des indices de pertinence $R_g$ associés au modèle humain 3D sont données sur la Figure 5.8.



8.99 8.34 27.61 26.60 14.23 14.23

FIG. 5.8 – Exemples d'indices de pertinence $R_g$ (%) associés aux six images de profondeur obtenus après projection du modèle sur les six faces de sa boîte englobante.

#### 5.2.2.3 Normalisation des indices de pertinence

Pour pouvoir introduire les pertinences dans l'estimation des similarités, les indices de pertinence doivent être préalablement normalisés. Les indices de pertinence normalisés $\bar{R}^i$ sont définis comme suit :

$$\bar{R}^i = \frac{R^i}{\sum_{i=1}^{V} R^i}, \tag{5.19}$$

$R^i$ étant les indices de pertinence associés aux projections 2D (silhouettes ou images de profondeur), définis précédemment.

### 5.2.3 Approches 2D/3D améliorées

Nous avons décrit dans la section précédente les différentes méthodes pour calculer les indices de pertinence associés aux silhouettes et aux images de profondeur. Nous allons maintenant les intégrer dans le processus de calcul 2D/3D afin d'améliorer la description et de parvenir à des mesures de similarités plus fins. Notre choix s'est porté sur deux signatures, la première fondée sur une indexation par silhouettes et la deuxième sur une indexation par images de profondeur. Comparés à de nombreux autres descripteurs, ces descripteurs montrent une meilleure efficacité et capacité de discrimination des objets 3D. Nous présentons dans cette section les deux approches en détail, ainsi que nos choix des paramètres.

#### 5.2.3.1 *Enhanced silhouette based approach* - ESA

Après normalisation, l'objet 3D est projeté orthogonalement sur les faces de sa boîte englobante pour extraire les silhouettes. Dans ce mode de projection, une silhouette et son opposé sont identiques, ce qui réduit le nombre des silhouettes de six à trois. Pour chaque silhouette, un indice de pertinence est calculé, et un contour associé à la forme est créé et indexé par un descripteur de forme 2D.

Le descripteur de forme 2D proposé dans [131] pour indexer une silhouette est basé sur quatre éléments qui vont être explicités dans la suite de cette section. Tout d'abord, le contour associé à la silhouette est détecté par un filtre d'opérations logiques. Il s'agira ensuite d'effectuer une sélection de points contours les plus éloignés selon une discrétisation angulaire uniforme. Une transformation discrète unidimensionnelle est alors appliquée sur cet ensemble en utilisant l'algorithme de transformée rapide. Seule une partie des coefficients de Fourier caractérisera la forme de la silhouette.

*Détection du contour :*
Une image silhouette **s** est donnée par une matrice carrée $[s_{ij}]$, où les éléments $s_{ij}$ $(1 < i, j < N)$ dénotent les valeurs des pixels de l'image monochrome : 1 dénote les pixels du fond (blanc) et 0 dénote les pixels de la silhouette (noir). L'élément $s_{ij}$ est un point contour si sa valeur est égale à 0 et au moins un de ses quatre proches voisins (nord, ouest, sud et est pixels) appartient au fond de l'image.

L'ensemble de ces points forment le contour $\mathbf{S}_L$ associé à l'image silhouette $\mathbf{s}$ :

$$\mathbf{S}_L = \{(i_0, j_0), \dots, (i_{L-1}, j_{L-1})\}, \qquad L \in \mathbb{N} \tag{5.20}$$

où $L$ est la longueur du contour. Dans le cas où l'objet possède des trous ou est composé de parties disjointes, il y aura plus qu'un contour dans l'image silhouette. Ici $\mathbf{S}_L$ sera considéré comme l'union de ces contours disjoints.

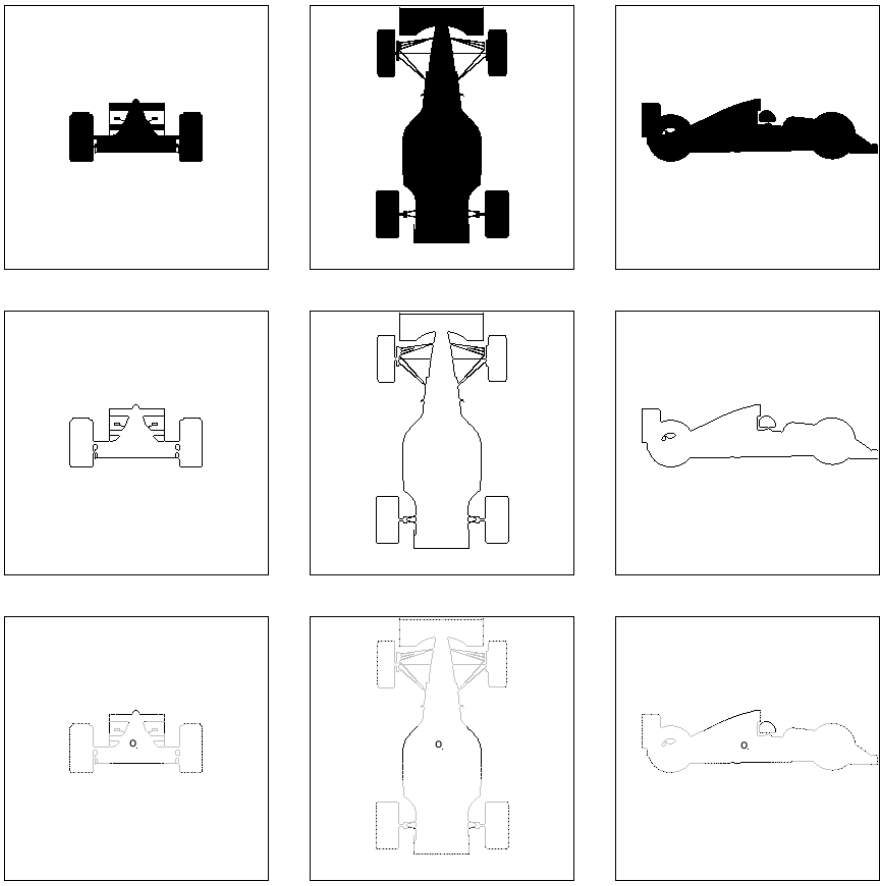


FIG. 5.9 – Extraction des séquences $\mathbf{C}_K$. **Ligne 1** : Trois images silhouettes extraites en projetant le modèle 3D sur sa boîte englobante, **Ligne 2** : Les trois contours obtenus avec un filtre d'opérations logiques, **Ligne 3** : Quantification angulaire des trois contours. Chaque contour est représenté par les $K$ points noirs.

*Quantification angulaire :*

Cependant, la longueur du contour $L$ variant d'une forme à une autre, un échantillonnage uniforme des points est donc nécessaire. Nous avons choisi ci-dessous une méthode permettant de sélectionner les $K$ points les plus éloignés du centre

de l'image $o = (N/2, N/2)$ selon une discrétisation angulaire uniforme. Les points choisis sont définis dans un système de coordonnées polaires et mis à échelle en les multipliant par $a_{max}/N$. Le contour est représenté par la nouvelle séquence $\mathbf{C}_K$ définie comme suit :

$$\mathbf{C}_K = \{c_0, \dots, c_{K-1}\},$$

$$c_i = (\rho_i, \varphi_i) = \begin{cases} (0,0), & \Psi_i = \emptyset \\ \frac{a_{max}}{N}(\rho_{max}, \varphi_{max}), & \text{sinon,} \end{cases}$$

$$\Psi_i = \left\{ (\rho_q, \varphi_q) = (i_q, j_q)_{\in \mathbf{S}_L} - o | \frac{(2i-1)\pi}{K} \leq \rho_q < \frac{(2i+1)\pi}{K} \right\},$$

$$(\rho_{max}, \varphi_{max}) = \underset{\Psi_i}{\operatorname{argmax}}\, \rho_q.$$

*Description par Fourier :*
La représentation de la silhouette par la transformée de Fourier discrète est naturelle et facile d'utilisation. Elle consiste à appliquer une transformation de Fourier à une dimension sur des données discrètes de $\mathbf{C}_K$. Nous obtenons un vecteur de même taille contenant les coefficients de Fourier. La transformation discrète correspond à une sommation sur l'espace de paramètre $\rho$, donnée par la formule suivante :

$$\hat{f}_p = \frac{1}{\sqrt{K}} \sum_{i=0}^{K-1} \rho_i e^{-j\frac{2\pi}{K}i.p}, \qquad p = 0, ..., K-1, \tag{5.21}$$

où $j$ est l'unité imaginaire et $\hat{f}_p \in \mathbb{C}$ sont les coefficients de Fourier. Si K est une puissance de 2, l'algorithme de la transformée de Fourier rapide peut être appliquée, ce qui permet d'accélérer les calculs et de réduire la complexité de $\mathcal{O}(K^2)$ en $\mathcal{O}(K \log(K))$. Les modules des coefficients obtenus, $|\hat{f}_p|$, servent à décrire le contour de la silhouette.

*Vecteur caractéristique :*
Pour des données réelles ($\rho_i \in \mathbb{R}$), la transformée a une propriété intéressante dite propriété de symétrie hermitienne. Les coefficients de Fourier déduits de l'équation (5.21) satisfont la relation suivante :

$$\forall i = 1, ..., K-1 \;\; \rho_i \in \mathbb{R} \Rightarrow \hat{f}_p = \overline{\hat{f}_{K-p}}, \quad p = 1, ..., K/2 - 1. \tag{5.22}$$

$\hat{f}_p$ et $\hat{f}_{K-p}$ sont des complexes conjugués, et ils ont le même module, $|\hat{f}_p| = |\hat{f}_{K-p}|$. Toute l'information spectrale est contenue dans l'intervalle de fréquences $\{0, ..., K/2\}$. Comme la silhouette sera représentée par les normes des coefficients

de Fourier, il sera intéressant de prendre seulement les $k$ premières fréquences, $\{\hat{f}_0, ..., \hat{f}_{k-1}\}$, tel que $k \leq K/2$. Ce qui nous amène à avoir $k$ valeurs numériques pour chaque silhouette. Trois silhouettes sont extraites sur la boîte englobante au lieu de six car les silhouettes sur les faces opposées sont identiques à un miroir près. Ce qui nous donne, pour chaque objet, trois vecteurs $\mathbf{F}^i = (|\hat{f}_0^i|, ..., |\hat{f}_{k-1}^i|)$ associés aux trois images silhouettes $\mathbf{s}^i$, $1 \leq i \leq 3$.

#### 5.2.3.2 *Enhanced depth-buffer based approach* - EDBA

Pour que sa signature soit robuste aux invariances géométriques, l'objet 3D est tout d'abord aligné, normalisé par rapport à une boîte englobante d'axes parallèles à ceux du repère intrinsèque à l'objet 3D. En projetant le modèle sur les six faces de cette boîte, des images de profondeur sont calculées, et les indices de pertinence associés sont déduits. Chaque image de profondeur sera représentée par un descripteur de forme 2D et un indice de pertinence.

Le descripteur de forme 2D proposé dans [56] pour indexer une image de profondeur se base sur une transformation de celle-ci dans l'espace de Fourier. La signature de l'objet 3D est établie en stockant, pour chaque image de profondeur, les coefficients basse-fréquence de l'image Fourier associée.

*Description par Fourier :*
La signature d'une image de profondeur est extraite en corrélant ses données dans le domaine spectral. Il s'agit d'appliquer une transformée de Fourier discrète bidimensionnelle (2D-FT) sur l'image de profondeur pour obtenir une image de Fourier de même taille. Les coefficients de Fourier $\hat{f}_{pq} \in \mathbb{C}$ pour une image $v_{ab} \in \mathbb{R}$, $a = 0, ..., N-1$, $b = 0, ..., N-1$, sont calculés comme suit :

$$\hat{f}_{pq} = \frac{1}{N} \sum_{a=0}^{N-1} \sum_{b=0}^{N-1} v_{ab} e^{-j2\pi(pa+qb)/N}, \tag{5.23}$$

où j est l'unité imaginaire et $p = 0, ..., N-1$, $q = 0, ..., N-1$. La complexité de cette 2D-FT est de l'ordre de $\mathcal{O}(N^4)$. Dans cette étape, nous avons cherché à réduire au maximum la complexité de la FT classique afin d'assurer une implémentation efficace.
En décomposant cette formule, il est possible de diminuer cette complexité pour parvenir à $\mathcal{O}(N^3)$. Il suffit de passer d'une opération à deux dimensions à deux

opérations à une dimension (1D-FT) :

$$\hat{f}_{pq} = \frac{1}{\sqrt{N}} \sum_{a=0}^{N-1} \left( \frac{1}{\sqrt{N}} \sum_{b=0}^{N-1} v_{ab} e^{-j2\pi pb/N} \right) e^{-j2\pi qa/N}, \tag{5.24}$$

$$\hat{f}_{pq} = \frac{1}{\sqrt{N}} \sum_{a=0}^{N-1} \hat{f}_{aq} e^{-j2\pi qa/N}, \ avec \ \ \hat{f}_{aq} = \frac{1}{\sqrt{N}} \sum_{b=0}^{N-1} v_{ab} e^{-j2\pi pb/N}. \tag{5.25}$$

Si $N$ est une puissance de 2, on peut appliquer la 2D-FFT qui permet d'accélérer les calculs et de réduire la complexité de $\mathcal{O}(N^3)$ en $\mathcal{O}(N^2 \log(N))$. Avant la 2D-FFT, nous décalons cycliquement tous les pixels de l'image d'origine comme suit :

$$v_{a'b'} \leftarrow v_{ab}, \quad a = a'[N/2], \ b = b'[N/2], \tag{5.26}$$

avec $0 \leq a', b' \leq N-1$. Ce décalage permet après l'application de la 2D-FFT de caractériser l'image de profondeur par les coefficients basse-fréquence situés au centre de l'image de Fourier.

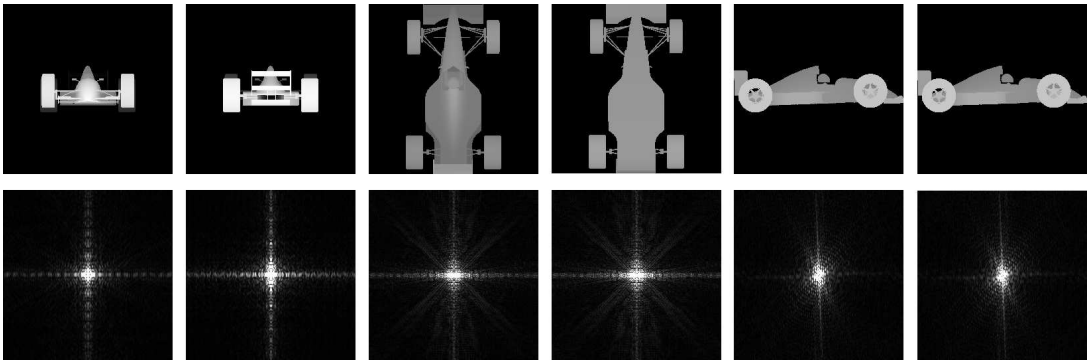


FIG. 5.10 – Six images de profondeur extraites en projetant le modèle 3D sur sa boîte englobante et leurs images de Fourier respectives.

*Vecteur caractéristique :*

Nous retrouvons une propriété intéressante de la transformée de Fourier qui permet de ne considérer que la moitié de l'image de Fourier. Il s'agit de la propriété de symétrie hermitienne de la transformée de Fourier quand les données d'entrée sont réelles ($v_{ab} \in [0,1]$). Les coefficients obtenus après la 2D-FFT $\hat{v}_{pq}$ satisfont alors la relation suivante :

$$v_{ab} \in \mathbb{R} \Rightarrow \hat{f}_{pq} = \overline{\hat{f}_{p'q'}}, \quad (p+p') = 0[N], \ (q+q') = 0[N], \tag{5.27}$$

où $\overline{\hat{f}_{p'q'}}$ est le conjugué complexe de $\hat{f}_{pq}$, $0 \leq a, b, p, q, p', q' \leq N-1$. Comme les images de profondeur seront représentées seulement par les normes des coefficients de Fourier $\|\hat{f}_{pq}\|$, il sera intéressant de prendre la moitié de ces normes d'après la formule précédente.
Le principal avantage de la FFT est que l'information contenue dans l'image de profondeur est concentrée essentiellement au niveau des coefficients basse-fréquence. Sur les images de Fourier, nous pouvons constater que les coefficients importants, c'est-à-dire les pixels à forte luminance sont les coefficients basse-fréquence, ceux du centre de l'image. Donc, pour extraire les signatures, nous prenons seulement des coefficients inclus dans une demi-boule prise au centre de l'image de Fourier. Le rayon de cette boule est à fixer et la taille de la signature en dépendra.

### 5.2.4 Mesure de similarité adoptée dans ESA et EDBA

Pour comparer deux objets $O^1$ et $O^2$, nous générons, pour chacun, les coefficients basse-fréquence de Fourier, $\mathbf{F}_1^i$ et $\mathbf{F}_2^i$, $1 \leq i \leq V$ des projections 2D (silhouettes ou images de profondeur). Nous calculons ensuite les indices de pertinence associés aux projections extraites pour chaque objet, $\bar{R}_1^i$ et $\bar{R}_2^i$, $1 \leq i \leq V$. Enfin, nous prenons la somme des distances entre les vecteurs de Fourier pondérés par les indices de pertinence normalisés comme mesure de similarité entre les objets $O^1$ et $O^2$ :

$$\Delta(O^1, O^2) = \sum_{i=1}^{V} d(\bar{R}_V^i \mathbf{F}_1^i, \bar{R}_2^i \mathbf{F}_2^i), \tag{5.28}$$

$d$ étant la distance euclidienne $l_1$ ou $l_2$ et $V$ étant le nombre de vues ($V = 3$ pour ESA et $V = 6$ pour EDBA).

## 5.3 Approches 2D/3D basées sur les lignes de profondeur

Un des objectifs de cette thèse a été de proposer un nouveau descripteur adapté au contenu spécifique des images de profondeur. Toutes les approches multi-vues proposées dans la littérature reposent sur des descripteurs de forme 2D classiques (la transformée de Fourier 2D [56, 92, 100], les points d'intérêt [93], l'aire et le volume [124]). Elles ne dépassent pas le cadre des méthodes directement dérivées de l'analyse d'images.

Une image de profondeur peut être envisagée comme un ensemble de lignes qui représentent l'évolution de la surface 3D le long de directions parallèles. Chaque ligne de profondeur apporte une information liée à une forme ou à l'aspect d'une partie de l'objet. Pour caractériser une image de profondeur, il apparaît donc intéressant de la décomposer d'abord en lignes, et d'indexer ces lignes par la suite. Pour décrire une ligne de profondeur d'une manière compacte, notre idée est de proposer une méthode de transcription qui utilise les pentes des profondeurs. Comparer deux lignes de profondeurs revient donc à comparer leurs séquences d'états. Il convient d'introduire une mesure de similarité fondée sur la programmation dynamique pour atteindre la correspondance maximale entre formes 3D ayant subi des déplacements ou des déformations locales.

### 5.3.1 Principe

Le principe de cette nouvelle approche 2D/3D consiste à décrire l'objet 3D à partir de ses lignes de profondeur. Après une étape de normalisation conférant à la méthode une invariance aux transformations géométriques, des images de profondeurs sont calculées. Sur chaque image, des lignes de profondeur sont extraites, puis représentées par des séquences d'états. Cette phase d'indexation s'appuie sur une méthode de transcodage qui permet une description compacte et pertinente. L'ensemble des séquences résultantes forment la signature de l'objet 3D. Pour comparer les objets, nous proposons une mesure de similarité utilisant des distances adaptées aux séquences d'états. Une partie de ce travail a été publiée dans ICME'07 [29] et ICIP'07 [30].

### 5.3.2 Descripteurs à base de lignes de profondeur

Les descripteurs proposés consistent à caractériser un objet 3D par un ensemble de séquences d'états décrivant l'information de la profondeur. Cet ensemble est construit sur les lignes extraites des images de profondeur. La signature de l'objet est calculée en deux étapes, l'extraction des lignes de profondeur et la transcription des lignes de profondeur.

#### 5.3.2.1 Extraction de lignes de profondeur

Pour extraire les lignes de profondeur, nous introduisons trois méthodes. Elles diffèrent par le type et le nombre des projections utilisées pour construire les images de profondeur. Les deux premières méthodes utilisent des projections sur un ensemble donné de plans et la troisième utilise des projections sur les faces latérales des cylindres englobants.

#### 5.3.2.2 *Depth Line Approach* - DLA

Pour garantir des signatures de petites tailles et par là même une recherche rapide, la première approche consiste à représenter l'objet 3D par seulement six vues. Dans la projection, nous considérons les faces de la boîte englobant l'objet 3D, de directions réparties uniformément sur la sphère unité. Le caractère englobant permet une description compacte.

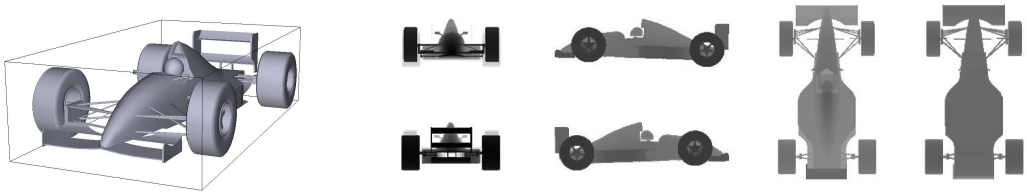


FIG. 5.11 – Six images de profondeur obtenues en projetant le modèle 3D sur sa boîte engblobante.

Durant la phase de l'indexation, l'objet 3D est projeté sur les faces de sa boîte englobante pour extraire six images de profondeur $\mathbf{v_i}$ $i \in \{1, ..., 6\}$, de résolution $N \times N$ (cf. Figure 5.11). Les lignes et les colonnes de chaque image de profondeur forment un ensemble de $2 \times N$ lignes de profondeur de taille $N$. Ainsi, le descripteur à base de lignes de profondeur sera représenté par $6 \times 2 \times N$ séquences de taille $N$, dérivées des lignes de profondeur créées.

#### 5.3.2.3 *Multi-view Depth Line Approach* - MDLA

Nous venons de voir le principe de la première méthode d'indexation à base de lignes de profondeur. Pour l'améliorer, nous proposons le '*Multi-view Depth Line Approach* utilisant plus de vues.

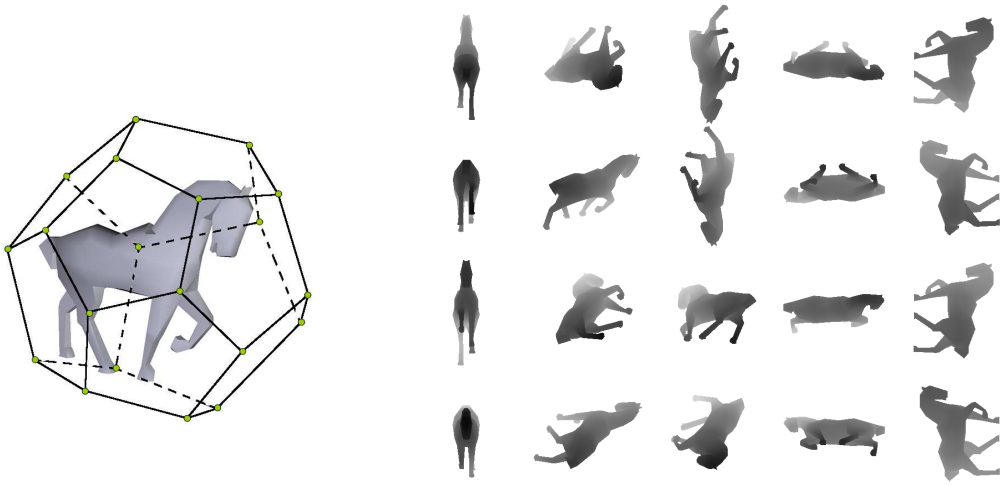


FIG. 5.12 – Vingt images de profondeur obtenues en projetant le modèle 3D suivant les directions des sommets d'un dodécaèdre.

Pour ce faire, nous choisissons de placer les points de vue sur les directions des sommets d'un dodécaèdre centré autour de l'objet, car c'est le polyèdre régulier ayant le plus de sommets répartis uniformément sur la sphère de vues. En projetant l'objet 3D sur les plans limitant celui-ci dans les directions des sommets, vingt images de profondeur $\mathbf{v_i}$, $i \in \{1, ..., 20\}$ de taille $N \times N$ pixels sont créées (cf. Figure 5.12). De chaque image de profondeur, sont extraites $2 \times N$ lignes de profondeur qui constituent les lignes et les colonnes. Le descripteur est ainsi représenté par $20 \times 2 \times N$ séquences de taille $N$, dérivées des lignes de profondeur extraites.

#### 5.3.2.4 *Cylindrical Depth Line Approach* - CDLA

Contrairement aux approches précédentes, l'espace projectif utilisé dans cette méthode est le cylindre englobant. Cela permet de créer une nouvelle image de profondeur dite image cylindrique. Cette image peut être vue comme la collection de bandes centrales des images de projection reconstruites localement en tournant l'objet autour de l'axe du cylindre.

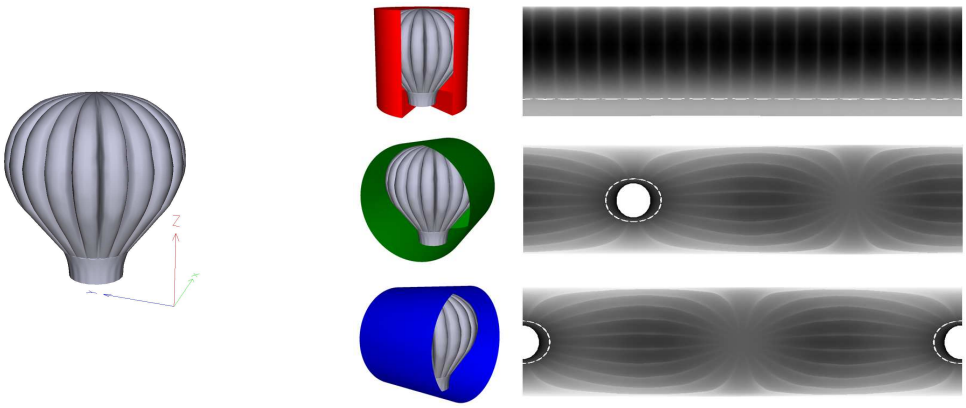


FIG. 5.13 – Trois images de profondeur obtenues en projetant le modèle 3D sur les faces latérales des cylindres englobants.

Pour capturer aux mieux la forme 3D sur tous les côtés, nous considérons les trois cylindres englobants d'axes alignés avec ceux du repère intrinsèque à l'objet (cf. Annexe A.1.3). Le modèle est projeté orthogonalement sur les faces latérales afin de générer les trois images de profondeur $\mathbf{v_i}$, $i \in \{1, 2, 3\}$ (cf. Figure 5.13). Chaque pixel a comme valeur la profondeur du premier point visible de l'objet 3D dans la direction orthogonale à l'axe du cylindre. La taille des images est de $N \times M$ pixels. En ne considérant que les lignes des images, on obtient un nombre de $3 \times N$ lignes de profondeur possibles. Dans notre implémentation, une ligne de profondeur a une longueur de $M = 4 \times N$ pixels. Le descripteur à base de lignes de profondeurs sera représenté par $3 \times N$ séquences de taille $M$, dérivées des lignes de profondeur créées.

| | ♯ images de profondeur | Taille de l'image de profondeur | Extraction ligne/colonne | Taille du descripteur | Taille de la séquence |
|---|---|---|---|---|---|
| DLA | 6 | $N \times N$ | r/c | $12 \times N$ | $N$ |
| MDLA | 20 | $N \times N$ | r/c | $40 \times N$ | $N$ |
| CDLA | 3 | $N \times 4N$ | r | $3 \times N$ | $4N$ |

TAB. 5.1 – Paramètres de l'extraction des lignes de profondeur pour les trois approches. La DLA et la CDLA ont une même taille de descripteur, contrairement à la MDLA qui a une taille quatre fois plus grande.

#### 5.3.2.5 Transcription d'une ligne de profondeur

Une fois les lignes de profondeur extraites, des signatures associées doivent être définies. Pour en obtenir une représentation compacte, nous allons proposer

une méthode de transcription qui caractérise une ligne de profondeur par une séquence d'états. Cette séquence sera représentée par une suite de symboles d'un alphabet préalable. Il est important que les états soient soigneusement définis pour permettre une description robuste et efficace de la forme.

Il est possible d'assimiler une ligne de profondeur à une courbe 2D. Dans ce cas, la transcription décrira l'évolution de la courbe le long de la ligne. Toutefois, cette technique présente un inconvénient majeur. Lorsque la ligne de profondeur est discontinue (ce qui correspond à des points de fond), il ne sera pas toujours possible de parvenir à des appariements optimaux. Il est alors intéressant d'étudier, analyser et décrire au mieux le contenu de la ligne de profondeur.

Notre méthode de transcodage se base essentiellement sur une étude de régions. Nous avons cherché à savoir quelles sont les différentes régions présentes sur une ligne de profondeur. Deux catégories de régions ont été retenues :

- *Régions fond* : elles représentent le fond de l'image. Deux types de points peuvent être alors distingués : point extérieur et point intérieur, c'est-à-dire, appartenant ou non aux extrémités de la ligne de profondeur. Les points intérieurs sont délimités par des points de projection. Nous les trouvons par exemple dans le cas où l'objet possède des trous ou est composé de parties disjointes.
- *Régions projection* : elles contiennent les projections des points visibles du modèle 3D. Pour ces régions, nous adoptons une description qui tient compte de la pente de la courbe en calculant sa dérivée. Bien qu'il existe plusieurs méthodes pour estimer la dérivée, nous utilisons l'expression de la première dérivée pour des raisons de simplicité $D(p) = dl(p) - dl(p+1)$, $\mathbf{dl} = [dl(p)]_{1,..,N}$ étant la ligne de profondeur. C'est la pente du segment joignant deux points de profondeur, d'indices successifs. Pour représenter cette estimation à chaque point, nous présentons trois états, état décroissant, état croissant et état constant, selon le signe de sa dérivée (négatif ou positif ou nul). Nous obtenons ainsi des informations sur la forme en examinant le comportement de la dérivée de la ligne de profondeur.
  L'utilisation de la dérivée nous a paru particulièrement intéressante car elle permet de garder l'information sur la forme sans s'attacher trop à la profondeur.

Chaque ligne de profondeur est codée en une séquence de $N$ états. Chaque état est représenté par un des cinq caractères (symboles), o, c, /, -, \, correspondant respectivement à : fond-intérieur, fond-extérieur, profondeur-décroissant, profondeur-constant et profondeur-croissant (cf. Figure (c) 5.14). Ces symboles définissent l'alphabet de notre méthode de transcodage.

| Caractère | Régions | Etat |
|---|---|---|
| *o* | Fond | Extérieur |
| *c* | Fond | Intérieur |
| / | Profondeur | Croissant |
| – | Profondeur | Constant |
| \ | Profondeur | Décroissant |

TAB. 5.2 – Les cinq états proposés pour représenter une ligne de profondeur.

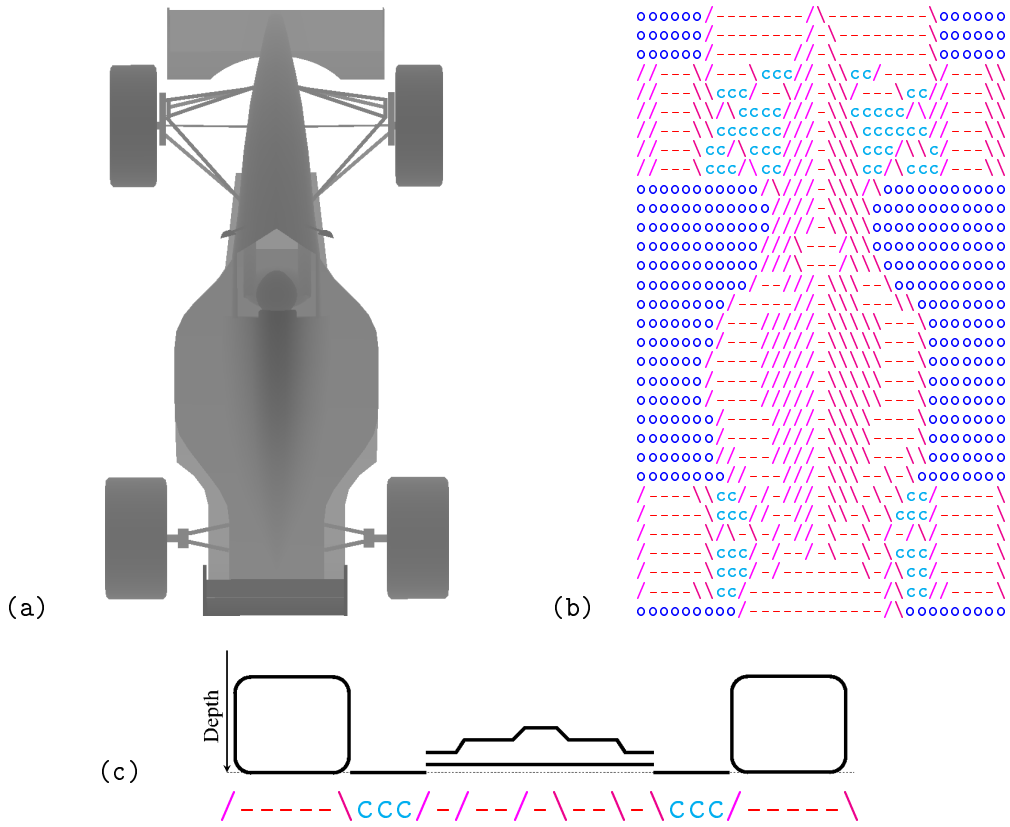



FIG. 5.14 – (a). Image de profondeur, (b). Séquences d'états associées aux lignes de profondeur, (c). Transcription d'une ligne de profondeur.

La méthode de transcription est la même pour toutes les lignes de profondeur extraites par les différentes approches proposées précédemment.

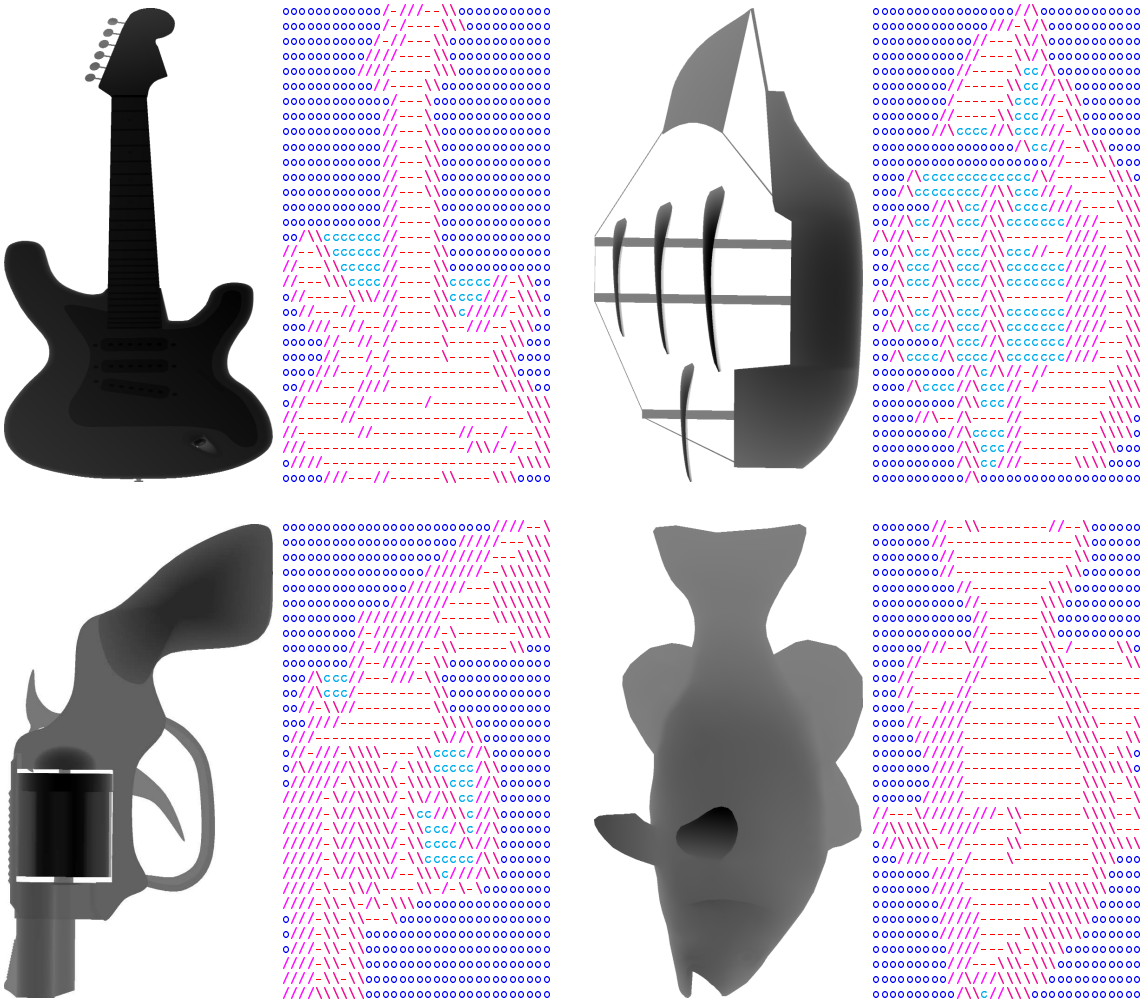


FIG. 5.15 – Exemples de descripteurs à base de lignes de profondeur. Chaque image de profondeur est caractérisée par un ensemble de 32 séquences d'états.

### 5.3.3 Mesure de similarité adoptée dans DLA, MDLA et CDLA

Le processus d'indexation mis en place caractérise l'objet 3D sous un format particulier (séquences d'états) différent des formats classiques (vecteurs ou graphes). Nous devons alors proposer une mesure de similarité adaptée au descripteur de forme pour caractériser au mieux la similarité entre objets 3D. Cette section sera organisée comme suit : dans un premier temps, nous détaillons les distances utilisées pour comparer deux séquences d'états, puis nous présentons la mesure de similarité que nous associons aux descripteurs à base de lignes de profondeur.

#### 5.3.3.1 Mesure de similarité entre deux séquences

Comme nous venons de le voir, chaque ligne de profondeur est décrite par une séquence d'états représentée sous forme d'un "mot". Pour comparer deux lignes de profondeur, nous étudions différentes mesures de similarité qui peuvent être associées à ces mots : une première mesure basée sur la distance de Hamming et deux mesures fondées sur la programmation dynamique qui semblent être des solutions plus prometteuses que la première.

Soit l'alphabet $\mathfrak{A} = \{o, c, /, -, \backslash\}$ défini précédemment et $\mathcal{S}$ l'ensemble des séquences de longueur $N$ à valeur dans $\mathfrak{A}$. Dans la suite, notons $A = (a_0, a_1, ..., a_{N-1})$ et $B = (b_0, b_1, ..., b_{N-1})$ les deux séquences de $\mathcal{S}$ à comparer.

**Distance de Hamming**

La distance de Hamming calcule le nombre de positions binaires de valeurs différentes séparant deux séquences de symboles (cf. Figure 5.16). Formellement, la distance de Hamming entre deux séquences $A$ et $B$ est le cardinal de l'ensemble des images de $A$ qui diffèrent de celles de $B$ :

$$d_H(A, B) = card\{\ i \mid a_i \neq bi\ \}. \tag{5.29}$$

Cette première métrique est très simple et rapide. Cependant, elle n'est pas optimale pour mesurer la similarité entre signatures d'objets 3D. En effet, certaines séquences représentent les mêmes formes 2D qui ne sont pas forcément bien alignées sur l'axe de projection.

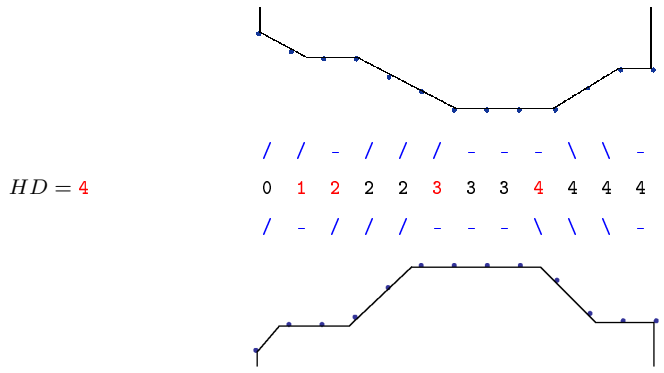



FIG. 5.16 – Comparaison de deux lignes de profondeurs utilisant la distance de Hamming entre leurs séquences respectives, $N = 12$ et $(C, S) = (0, 1)$.

Comme nous pouvons le constater sur la Figure 5.17, la distance de Hamming ne tolère pas un déplacement ou un étirement d'une partie ou de la totalité de la forme. Il en résulte que deux formes similaires qui ne sont pas parfaitement alignées ne seront pas considérées de la même manière lors de la comparaison. Il est possible d'éviter ce problème en utilisant des algorithmes d'appariement de séquences, qui permettent d'établir la correspondance maximale entre deux séquences, en repérant leurs sous-séquences communes ou similaires. Pour ce faire, nous allons utiliser la programmation dynamique.

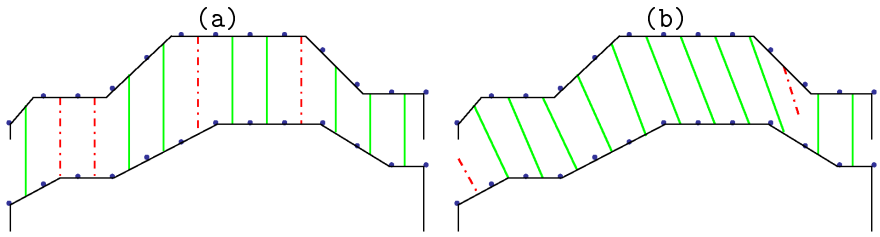



FIG. 5.17 – Comparaison entre la distance de Hamming et la distance utilisant la programmation dynamique. Les deux lignes à comparer représentent deux formes similaires, mais ils ne sont pas parfaitement alignés le long de l'axe des abscisses. (a) 4 substitutions, la distance de Hamming calcule une mesure de similarité pessimiste. (b) 2 substitutions, la programmation dynamique permet de trouver le meilleur alignement optimisant la mesure de similarité.

### Distance basée sur une programmation dynamique

La programmation dynamique s'applique le plus souvent aux problèmes d'optimisation qui satisfont au principe d'optimalité de Bellman : "Dans une séquence optimale, chaque sous-séquence doit être aussi optimale". Des solutions élégantes apportées au problème de l'alignement optimal entre deux séquences ont été proposées par Levenshtein (1966), Needleman & Wunsch (1970) et Smith & Waterman (1981). La distance de Levenshtein, connue sous le nom de "distance d'édition", a été adoptée pour l'alignement de courbes [133] et l'appariement de formes 2D [3, 69]. Les méthodes de Needleman-Wunsch [83] et de Smith-Waterman [115] ont été utilisées dans la comparaison des séquences d'états, en particulier dans la bioinformatique. Notons que pour mesurer la similarité entre deux séquences d'états pouvant subir des déplacements et des déformations (rétrécissement ou étirements), des métriques sophistiquées données par les scores des alignements optimaux sont les plus adaptées.

Dans ce travail, nous proposons d'utiliser l'algorithme de Needleman-Wunsch (NW) et l'algorithme de Smith-Waterman (SW) pour comparer deux séquences. Le premier trouve le meilleur alignement global entre deux séquences, tandis que le second trouve des paires de sous-séquences similaires via des alignements locaux. Les deux méthodes calculent un coût d'alignement qui peut être considéré comme une mesure de similarité entre séquences.

*Alignement des séquences :*
Mesurer la similarité entre deux séquences nécessite d'aligner ces séquences, c'est-à-dire de mettre en correspondance leurs régions similaires pour estimer le coût de la transformation d'une séquence à une autre. À chaque position dans l'alignement correspond une des trois opérations suivantes :

- un match (appariement), quand le même caractère apparaît dans les deux séquences,
- une substitution, quand il existe deux caractères différents appartenant chacun à une séquence,
- un gap, c'est-à-dire une insertion d'un caractère dans seulement une séquence, ou une délétion (suppression) dans une des deux séquences.

Pour aligner les séquences $A$ et $B$ utilisant la programmation dynamique, une matrice carrée $N \times N$ dite matrice de correspondance $\mathbf{M}$ est construite. Chaque cellule $M_{i,j}$ contient la similarité entre les éléments $a_i$ et $b_j$ (un match $C$ et une substitution $S$).

**Algorithme 5.2** Calcul de la matrice de correspondance $\mathbf{M}$

**Pour tout** $i$, $0 \leq i < N$ **faire**
  **Pour tout** $j$, $0 \leq j < N$ **faire**
    **Si** $a_i = b_j$ **alors**
      $M_{i,j} = C$
    **Sinon**
      $M_{i,j} = S$
    **Fin si**
  **Fin pour**
**Fin pour**

Puis, une matrice carrée $(N+1) \times (N+1)$ appelé matrice de programmation dynamique $\mathbf{H}$ est calculée utilisant la matrice de correspondance $\mathbf{M}$ et les opérations du gap (une insertion $I$ ou une délétion $D$). Chaque élément $H_{i,j}$ représente le score de similarité maximum entre les deux sous-séquences $(a_0, a_1, ..., a_{i-1})$ et

$(b_0, b_1, ...b_{j-1})$. La construction de la matrice **H** dépend de l'équation de programmation dynamique associée à l'algorithme considéré.

*Algorithme de Needleman-Wunch :*
La matrice de programmation dynamique **H** est calculée par une récursion à deux dimensions en utilisant l'algorithme suivant :

**Algorithme 5.3** Algorithme de Needleman-Wunch

calculer **M** (cf. Algorithme 5.2)
$H_{0,0} = 0$
**Pour tout** $j$, $0 \leq j \leq N$ **faire**
  $H_{0,j} = H_{0,j-1} + D$
**Fin pour**
**Pour tout** $i$, $1 \leq i \leq N$ **faire**
  $H_{i,0} = H_{i-1,0} + I$
  **Pour tout** $j$, $1 \leq j \leq N$ **faire**

$$H_{i,j} = \max \begin{cases} H_{i-1,j-1} + M_{i-1,j-1} \\ H_{i-1,j} + I \\ H_{i,j-1} + D \end{cases}$$

  **Fin pour**
**Fin pour**
$d_{NW} = C - H_{N,N}/N$

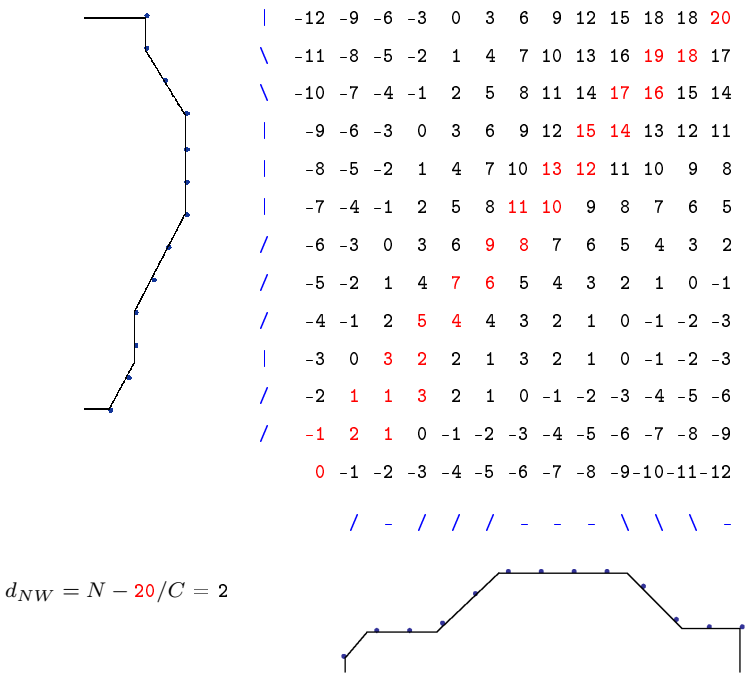



FIG. 5.18 – Exemple d'alignement de séquences d'états utilisant l'algorithme de Needleman-Wunsch, $N = 12$ et $(C, S, I, D) = (2, -1, -1, -1)$.

Dans cet algorithme, la propagation des scores dans la matrice de programmation dynamique est uniquement locale : chaque cellule intérieure reçoit les scores des trois cellules précédentes et envoie son résultat aux trois cellules suivantes. La valeur $d_{NW}(A, B) = C - H_{N,N}/N$ obtenue en fin de calcul représente la similarité globale entre $A$ et $B$. Le temps d'exécution de l'algorithme est de l'ordre $\mathcal{O}(N^2)$. La Figure 5.18 donne un exemple de calcul de matrice de programmation dynamique avec l'algorithme de Needleman-Wunsch.

Smith et Waterman ont étendu cette méthode pour la recherche d'alignements optimaux locaux. La modification majeure de l'algorithme se situe au niveau de l'équation de récurrence qui force la valeur de $H_{i,j}$ à être supérieure ou égale à zéro. Autrement dit, si le meilleur alignement jusqu'à la position $(i, j)$ aboutit à un score négatif, l'alignement est interrompu. Un nouvel alignement local est engagé à partir de cette position.

*Algorithme de Smith-Waterman :*
La matrice de programmation dynamique **H** est calculée récursivement en utilisant l'algorithme de Smith-Waterman suivant :

**Algorithme 5.4** Algorithme de Smith-Waterman

calculer **M** (cf. Algorithme 5.2)
$d_{SW} = 0$
**Pour tout** $j$, $0 \leq j \leq N$ **faire**
  $H_{0,j} = 0$
**Fin pour**
**Pour tout** $i$, $1 \leq i \leq N$ **faire**
  $H_{i,0} = 0$
  **Pour tout** $j$, $1 \leq j \leq N$ **faire**

$$H_{i,j} = \max \begin{cases} H_{i-1,j-1} + M_{i-1,j-1} \\ \max\limits_{1 \leq l < i} \{H_{i-l,j} + I_l\} \\ \max\limits_{1 \leq k < j} \{H_{i,j-k} + D_k\} \\ 0 \end{cases}$$

    $d_{SW} = \max(d_{SW}, H_{i,j})$
  **Fin pour**
**Fin pour**
$d_{SW} = C - d_{SW}/N$

La propagation des scores dans la matrice de programmation dynamique est globale : tous les chemins menant à une cellule donnée sont évalués. Ces chemins peuvent avoir n'importe quelle longueur et comporter des gaps de longueur

variable. La quantité $d_{SW}(A, B) = C - \max(H_{i,j})/N$ représente la meilleure similarité locale entre $A$ et $B$. La Figure 5.19 donne un exemple de calcul de matrice de programmation dynamique avec l'algorithme de Smith-Waterman.

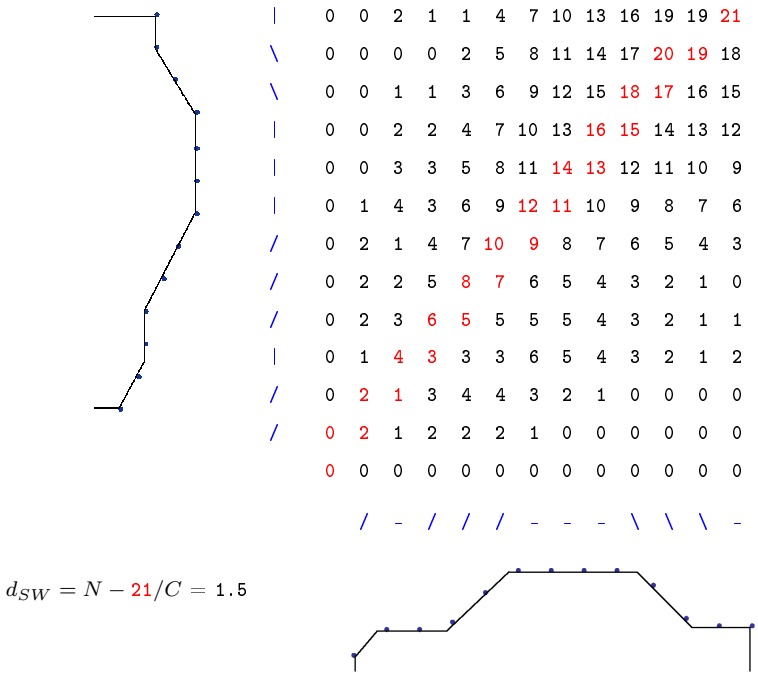



FIG. 5.19 – Exemple d'alignement de séquences d'états utilisant l'algorithme de Smith-Waterman, $N = 12$ et $(C, S, I_l, D_k) = (2, -1, -l, -k)$.

Dans la construction de la matrice de programmation dynamique $\mathbf{H}$, les délétions $D_k$ de longueur $k$ et les insertions $I_l$ de longueur $l$ sont données par des fonctions de gap. On utilise souvent la fonction affine de type $G_k = -(o + e \cdot k)$, composée de deux parties : un gap à pénalité ouverte $o$ et un gap à pénalité étendue $e$. Dans ce cas, les relations de programmation dynamique sont calculées en utilisant des nouvelles matrices $\mathbf{P}$ et $\mathbf{Q}$, comme l'a montré Gotoh dans [54]. La complexité de l'algorithme devient $\mathcal{O}(N^2)$, $N$ étant la taille de la séquence (cf. Algorithme 5.5).

Dans notre implémentation, nous utilisons un gap de $k$ positions successives ayant une pénalité globale $G_k = -k$. Cette pénalité peut être vue comme une fonction linéaire, $D_k = -k$ et $I_l = -l$.

**Algorithme 5.5** Algorithme de Smith-Waterman optimisé

calculer $\mathbf{M}$ (cf. Algorithme 5.2)
$d_{SW} = 0$
**Pour tout** $j$, $0 \leq j \leq N$ **faire**
  $H_{0,j} = P_{0,j} = 0$
**Fin pour**
**Pour tout** $i$, $1 \leq i \leq N$ **faire**
  $H_{i,0} = Q_{i,0} = 0$
  **Pour tout** $j$, $1 \leq j \leq N$ **faire**
    $P_{i,j} = \max(H_{i-1,j} - o, P_{i-1,j} - e)$
    $Q_{i,j} = \max(H_{i,j-1} - o, Q_{i,j-1} - e)$
    $H_{i,j} = \max \begin{cases} H_{i-1,j-1} + M_{i-1,j-1} \\ P_{i,j} \\ Q_{i,j} \\ 0 \end{cases}$
    $d_{SW} = \max(d_{SW}, H_{i,j})$
  **Fin pour**
**Fin pour**
$d_{SW} = C - d_{SW}/N$

#### 5.3.3.2 Mesure de similarité entre deux descripteurs à base de lignes de profondeur

Pour comparer deux objets $O^1$ et $O^2$, nous générons, pour chacun, les séquences $\mathbf{S^1}$ et $\mathbf{S^2}$ déduites des lignes de profondeur. Nous calculons ensuite les scores de similarité entre chaque séquence de l'objet $O^1$ et la séquence correspondante de l'objet $O^2$. Enfin, la mesure de similarité entre les objets $O^1$ et $O^2$ est définie comme la somme des scores :

- *Depth Line Approach / Multi-view Depth Line Approach* :

$$\Delta(O^1, O^2) = \sum_{i=1}^{V} \sum_{j=r,c} \sum_{k=0}^{N-1} d(\mathbf{S^1_{i,j,k}}, \mathbf{S^2_{i,j,k}}), \tag{5.30}$$

  où $i$ dénote l'indice de l'image, $j$ indique si la ligne de profondeur est une ligne ou une colonne, et $k$ dénote l'ordre de la ligne ou la colonne dans l'image de profondeur. $V$ est le nombre de vues : $V = 6$ pour DLA et $V = 20$ pour MDLA.

- *Cylindrical Depth Line Approach* :

$$\Delta(O^1, O^2) = \sum_{i=1}^{3} \sum_{j=0}^{N-1} d(\mathbf{S^1_{i,j}}, \mathbf{S^2_{i,j}}), \tag{5.31}$$

  où $i$ et $j$ dénotent respectivement l'indice de l'image et l'ordre de la ligne dans l'image de profondeur.

Les distances de similarité $d$ considérées sont la distance de Hamming $d_H$, la distance de Needleman-Wunch $d_{NW}$ et la distance de Smith-Waterman $d_{SW}$.

## 5.4 Conclusion

Dans ce chapitre, nous avons abordé la problématique de comparaison d'objets polyédriques 3D, visant des applications de requête par similarité de forme. Nous avons proposé six approches originales d'indexation d'objets 3D, dont une approche fondée sur la transformée de Gauss 3D, deux approches multi-vues améliorées et trois à base de lignes de profondeur. En outre, des mesures de similarité associées aux différentes signatures sont également proposées.

Dans le prochain chapitre, nous allons étudier et évaluer les performances des approches proposées et les comparer à d'autres approches, à l'aide d'outils quantitatifs classiquement utilisés pour les applications de recherche d'information dans les bases de données.

# Chapitre 6

# Résultats expérimentaux

Dans ce chapitre, nous évaluons les performances des différentes approches de recherche par le contenu d'objets 3D présentées dans le chapitre 5 sur quatre bases de données généralistes (cf. section 1.4) de taille et de complexité différentes. La section 6.1 présente les différentes techniques utilisées pour évaluer les méthodes proposées. À l'aide de ces outils, nous fixons en section 6.2, pour chaque processus de recherche, les paramètres de description, distances de similarité et méthodes d'alignement permettant d'atteindre les meilleures performances. Une étude comparative des méthodes d'indexation 3D est détaillée dans la section 6.3.

## 6.1 Techniques d'évaluation

Soient $\mathfrak{B}$ une base de modèles 3D entièrement classée (classifié) et $C_1, ..., C_c$ ses classes d'objets similaires :

$$\mathfrak{B} = \bigcup_{i=1}^{c} C_i \ , \qquad \forall i \neq i', \quad C_i \cap C_{i'} = \emptyset \tag{6.1}$$

$$C_i = \{O_1^i, ..., O_{n_i}^i\} \ , \quad O_j^i \in \mathfrak{B} \ (1 \leq j \leq n_i), \tag{6.2}$$

$$\text{et} \quad n = |\mathfrak{B}| = \sum_{i=1}^{c} n_i. \tag{6.3}$$

Pour évaluer les performances de recherche d'un système donné, la mesure de similarité entre chaque paire d'objets de la base $\mathfrak{B}$ est calculée. Ceci permet de fournir, pour chaque objet de la base, un classement des objets selon leurs ordres de ressemblance.

Soient $O \in C_i$ l'objet requête et $R_k(O)$ le résultat de la requête de taille $k$ :

$$R_k(O) = \{Q_1, Q_2, ..., Q_k\} \ , \quad Q_j \in \mathfrak{B} \setminus \{O\} \ (1 \leq j \leq k). \tag{6.4}$$

$R_k(O)$ représente la liste ordonnée des $k$ plus proches objets de $\mathfrak{B} \setminus \{O\}$ retournés dans un ordre croissant de similarité. L'objet $Q_1$ est le plus proche voisin, $Q_2$ est le deuxième voisin, etc. Dans la suite, pour une requête $R_k(O)$, nous appellerons pertinents les objets appartenant à la classe de l'objet requête $O$. L'ensemble des objets pertinents et son cardinal sont donnés respectivement par :

$$P_k(O) = C_i \cap R_k(O) \ , \quad p_k(O) = |P_k(O)|. \tag{6.5}$$

### 6.1.1 Mesures de performance

Pour comparer quantitativement et objectivement les performances des approches proposées, quatre mesures d'évaluation sont calculées :

- Le *Nearest Neighbour* (NN) : caractérise le plus proche voisin. Il est défini pour chaque objet de la base comme un score binaire, égal à 1 si le plus proche voisin retourné appartient à la même classe de la requête et à 0 sinon.

$$\mathrm{NN}(O) = p_1(O),$$

$$\mathrm{NN}(C_i) = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathrm{NN}(O_j^i),$$

$$\mathrm{NN}(\mathfrak{B}) = \frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n_i}\mathrm{NN}(O_j^i). \tag{6.6}$$

– Le *First Tier* (FT) : connu aussi sous le nom de *R-precision*. Il est défini pour un objet requête par le pourcentage des réponses pertinentes obtenues au rang $(n_i - 1)$, $n_i$ étant le nombre d'éléments de la classe de la requête.

$$\mathrm{FT}(O) = \frac{p_{(n_i-1)}(O)}{n_i - 1},$$

$$\mathrm{FT}(C_i) = \frac{1}{n_i}\sum_{j=1}^{n_i}\mathrm{FT}(O_j^i),$$

$$\mathrm{FT}(\mathfrak{B}) = \frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n_i}\mathrm{FT}(O_j^i). \tag{6.7}$$

– Le *Second Tier* (ST) : connu aussi sous le nom de *Bull's Eye Performance*. C'est le pourcentage des objets pertinents retournés parmi les $2(n_i-1)$ plus proches voisins, $n_i$ étant le nombre d'objets de la classe de la requête.

$$\mathrm{ST}(O) = \frac{p_{2(n_i-1)}(O)}{n_i - 1},$$

$$\mathrm{ST}(C_i) = \frac{1}{n_i}\sum_{j=1}^{n_i}\mathrm{ST}(O_j^i),$$

$$\mathrm{ST}(\mathfrak{B}) = \frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n_i}\mathrm{ST}(O_j^i). \tag{6.8}$$

– Le *Discounted Cumulative Gain* (DCG) : consiste à cumuler les contributions de tous les objets de la base, avec des poids dépendant du rang des objets retournés par la requête $R_{(n-1)}(O)$ (6.4), $n$ étant le nombre d'objets dans la base.

$$\mathrm{DCG}(O) = \frac{1 + \sum_{k=1}^{n-1}\frac{G_k}{\log_2(k+1)}}{1 + \sum_{k=1}^{n_i-1}\frac{1}{\log_2(k+1)}}, \quad \text{avec} \quad G_k = \begin{cases} 1 & \text{si } Q_k \in C_i \\ 0 & \text{sinon} \end{cases}.$$

$$\mathrm{DCG}(C_i) = \frac{1}{n_i}\sum_{j=1}^{n_i}\mathrm{DCG}(O_j^i),$$

$$\mathrm{DCG}(\mathfrak{B}) = \frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n_i}\mathrm{DCG}(O_j^i). \tag{6.9}$$

Ces mesures permettent d'analyser et de comparer rapidement les performances des différentes méthodes. Elles ont été utilisées dans [112].

### 6.1.2 Courbe précision-rappel

Pour une analyse plus fine des performances, une évaluation en termes de courbe précision-rappel sera présentée. Deux mesures sont calculées pour définir cette courbe : le rappel qui représente le nombre d'objets pertinents retournés $p_k(O)$ (6.5) au rang $k$ sur le nombre d'éléments de la classe de l'objet requête $(n_i - 1)$, et la précision qui représente le nombre d'objets pertinents retournés $p_k(O)$ sur le nombre total de réponses obtenues $k$. Formellement, pour un objet requête $O \in C_i$, ce couple à la position $k$ est défini comme suit :

$$\text{Rappel}_k(O) = \frac{p_k(O)}{n_i - 1}, \qquad \text{Précision}_k(O) = \frac{p_k(O)}{k}. \tag{6.10}$$

Le rappel mesure la capacité du système à retrouver les modèles pertinents, alors que la précision mesure la capacité du système à rejeter tous les modèles non pertinents. Notons que les valeurs de NN, FT et ST peuvent être déduites des mesures rappel/précision, $\text{NN}(O) = \text{Précision}_1(O)$, $\text{FT}(O) = \text{Rappel}_{(n_i-1)}(O)$, $\text{ST}(O) = \text{Rappel}_{2(n_i-1)}(O)$.
La courbe précision-rappel de l'objet $O$ est donnée par les $n_i - 1$ points suivants :

$$\Big(\text{Rappel}_{k_j}(O), \text{Précision}_{k_j}(O)\Big) = \left(\frac{j}{n_i - 1}, \frac{j}{k_j}\right), \qquad 1 \leq j \leq (n_i - 1),$$

$k_j$ étant la position du $j^{ème}$ objet pertinent retourné dans $R_k(O)$ (6.4), $Q_{k_j} \in C_i$. Comme nous pouvons le constater, la définition du $\text{Rappel}_{k_j}(O)$ dépend de la taille de la classe. Afin de calculer une moyenne de la précision sur des classes de tailles différentes, nous proposons d'estimer la précision sur des valeurs fixes de rappel

$$\text{rappel}_l = \frac{l}{L}, \quad 1 \leq l \leq L, \tag{6.11}$$

en appliquant une interpolation linéaire :

$$\text{Rappel}_{k_j}(O) \leq \text{rappel}_l \leq \text{Rappel}_{k_{j+1}}(O),$$

$$\Delta_{p/r}(O) = \frac{\text{Précision}_{k_{j+1}}(O) - \text{Précision}_{k_j}(O)}{\text{Rappel}_{k_{j+1}}(O) - \text{Rappel}_{k_j}(O)},$$

$$\text{précision}_l(O) = \text{Précision}_{k_j}(O) + \Delta_{p/r}(O) \cdot (\text{rappel}_l - \text{Rappel}_{k_j}(O)).$$

Dans la suite, la courbe précision-rappel pour chaque objet $O$ de la base de données sera représentée en abscisse par les $(\text{rappel}_l)_{1 \leq l \leq L}$ et en ordonnée par les $(\text{précision}_l(O))_{1 \leq l \leq L}$. La moyenne des rappels peut alors être faite sur une classe

de la base et sur la base d'objets. Les courbes précision-rappel associées seront représentées en abscisse par les $(\text{rappel}_l)_{1\leq l\leq L}$ et respectivement en ordonnée par les :

$$\text{précision}_l(C_i) = \frac{1}{n_i}\sum_{j=1}^{n_i}\text{précision}_l(O_j^i), \quad 1 \leq l \leq L,$$

$$\text{précision}_l(\mathfrak{B}) = \frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n_i}\text{précision}_l(O_j^i), \quad 1 \leq l \leq L, . \tag{6.12}$$

## 6.2 Évaluation des approches proposées

Dans cette section, les méthodes présentées dans le chapitre 5 sont évaluées en fonction des différents paramètres de description, distances de similarité et méthodes d'alignement sur les quatre bases généralistes PSB Test, PSB Train, MPEG7 et SHREC'09. Ceci nous permettra de fixer les paramètres optimisant les performances de recherche pour chaque méthode.

### 6.2.1 *3D Gaussian Descriptor based Approach*

Comme nous avons vu dans la section 5.1.1.2, lors de la construction du descripteur gaussien 3D, différents paramètres entrent dans l'approximation discrète de la transformée de Gauss. Dans cette partie, nous avons étudié les possibilités d'utiliser deux ensembles finis de points pris sur la surface de l'objet 3D :
- $E_T$ l'ensemble contenant les centres de gravité des facettes décrivant le maillage,
- $E_U$ l'ensemble donné par un échantillonnage de points répartis uniformément sur la surface.
En outre, nous avons proposé de pondérer la contribution par les tailles des facettes pour pouvoir gérer des modèles irrégulièrement maillés.

L'expérimentation menée compare plusieurs variantes : avec et sans pondération, pour les ensembles $E_T$ et $E_U$. L'évaluation de $n_{min}$, qui correspond à la finesse de l'échantillonnage dans le cas de $E_U$, est réalisée en comparant trois valeurs $n_{min} \in \{10^3, 10^4, 10^5\}$. Les Figures 6.1 et les Tableaux 6.1 contiennent respectivement les résultats en termes de courbes précision-rappel et de critères de performance (NN, FT, ST et DCG).

PSB Test Database

| $E$ | Pondé | NN | FT | ST | DCG |
|---|---|---|---|---|---|
| $E_U 10^5$ | avec | 60.75 | 34.55 | 43.66 | 60.18 |
| $E_U 10^4$ | avec | 60.09 | 34.10 | 43.71 | 59.83 |
| $E_U 10^3$ | avec | 57.00 | 32.42 | 42.14 | 58.56 |
| $E_T$ | avec | 57.00 | 29.86 | 38.23 | 56.37 |
| $E_U 10^4$ | sans | 52.81 | 28.32 | 36.95 | 54.49 |
| $E_T$ | sans | 41.90 | 16.06 | 21.10 | 43.26 |

PSB Train Database

| $E$ | Pondé | NN | FT | ST | DCG |
|---|---|---|---|---|---|
| $E_U 10^5$ | avec | 65.05 | 37.26 | 48.20 | 63.59 |
| $E_U 10^4$ | avec | 64.72 | 37.08 | 47.71 | 63.43 |
| $E_U 10^3$ | avec | 63.29 | 35.18 | 46.15 | 62.02 |
| $E_T$ | avec | 62.40 | 33.83 | 42.93 | 60.50 |
| $E_U 10^4$ | sans | 56.78 | 29.84 | 39.25 | 56.96 |
| $E_T$ | sans | 40.57 | 17.20 | 22.53 | 44.44 |

MPEG7 Database

| $E$ | Pondé | NN | FT | ST | DCG |
|---|---|---|---|---|---|
| $E_U 10^5$ | avec | 84.58 | 62.62 | 73.62 | 83.99 |
| $E_U 10^4$ | avec | 84.58 | 61.68 | 73.33 | 83.51 |
| $E_U 10^3$ | avec | 84.14 | 61.03 | 73.50 | 83.15 |
| $E_T$ | avec | 82.38 | 59.82 | 69.64 | 82.06 |
| $E_U 10^4$ | sans | 84.14 | 61.05 | 70.04 | 81.93 |
| $E_T$ | sans | 80.18 | 51.19 | 60.19 | 76.48 |

SHREC09 Database

| $E$ | Pondé | NN | FT | ST | DCG |
|---|---|---|---|---|---|
| $E_U 10^5$ | avec | 85.50 | 53.25 | 64.00 | 72.95 |
| $E_U 10^4$ | avec | 84.13 | 52.63 | 63.88 | 72.51 |
| $E_U 10^3$ | avec | 83.25 | 50.78 | 62.70 | 71.01 |
| $E_T$ | avec | 83.38 | 49.76 | 60.34 | 69.51 |
| $E_U 10^4$ | sans | 76.63 | 36.95 | 46.59 | 57.21 |
| $E_T$ | sans | 62.25 | 24.70 | 31.02 | 40.79 |

TAB. 6.1 – Mesures de performance (%) de l'approche 3DGA pour différentes méthodes approximant discrètement la transformée de Gauss, avec $N = 8$, $\sigma = 0.23$, la boîte englobante BB et la distance $d_v$.

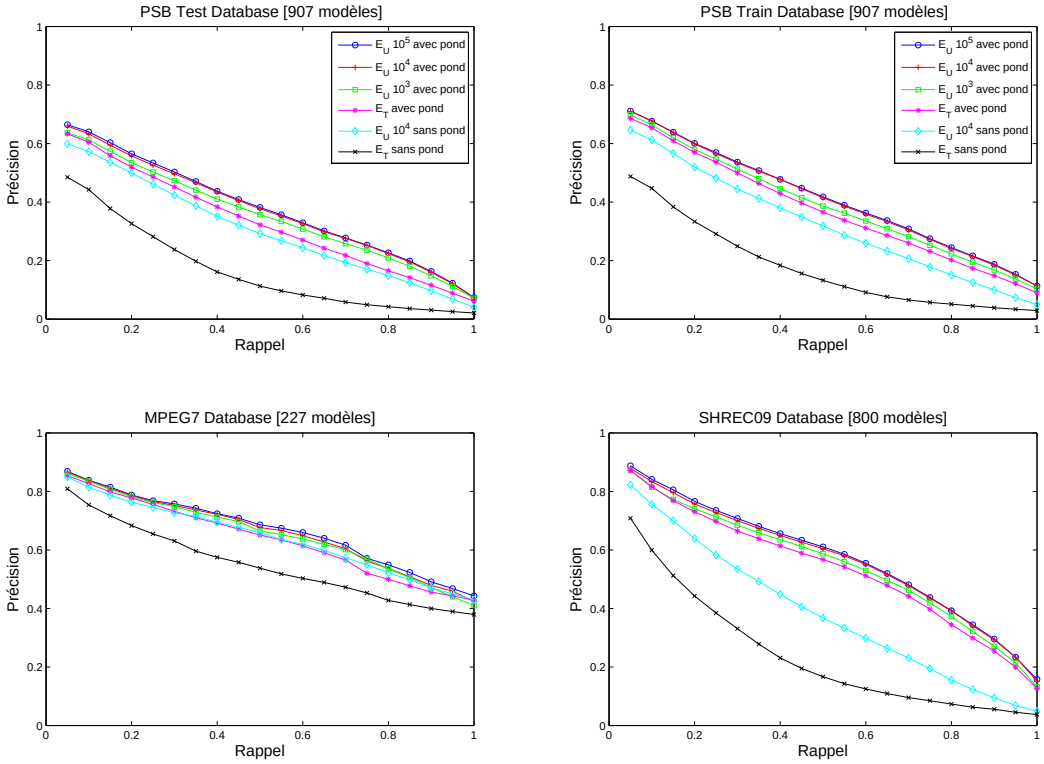



FIG. 6.1 – Courbes de précision-rappel de l'approche 3DGA pour différentes méthodes approximant discrètement la transformée de Gauss, avec $N = 8$, $\sigma = 0.23$, la boîte englobante BB et la distance $d_v$.

Ces résultats montrent la nette supériorité des descripteurs construits après un remaillage uniforme, et calculés avec un mécanisme de pondération par la surface. Cette performance est étroitement liée à un échantillonnage plus uniforme

et à une représentation plus complète de la surface. Mentionnons également que l'utilisation de l'échantillonnage $E_T$ conduit à des résultats nettement inférieurs à ceux produits avec $E_U$. Nous considérons dans la suite que la surface est approximée par $E_U$ ($n_{min} = 10^5$) et que la représentation de la surface est pondérée.

Une fois le choix de la méthode d'approximation discrète effectué, des séries d'expériences ont été réalisées. Dans un premier temps, des tests ont été mis en place pour fixer les paramètres du descripteur : la taille du descripteur $N^3$, la boîte englobante $B$ et la variance $\sigma$ (cf. section 5.1.1.1). Puis nous avons évalué les différentes mesures de similarité associées au descripteur gaussien 3D, décrites dans la section 5.1.2 et enfin, nous avons comparé la robustesse de la description obtenue avec notre méthode d'alignement à celle avec une ACPC. Ces tests ont été faits sur quatre bases d'objets présentées dans la section 1.4.

L'évaluation du paramètre $dim = N^3$, qui correspond à la taille de la discrétisation volumique et, par là même, à la taille du descripteur, est faite en comparant cinq valeurs de $N$, $N \in \{4, 6, 8, 10, 12\}$. Ce qui donne respectivement les dimensions suivantes $dim \in \{64, 216, 512, 1000, 1728\}$. Les Figures 6.2 et les Tableaux 6.2 présentent respectivement les courbes précision-rappel et les mesures de performance obtenues sur les quatre bases de données.

PSB Test Database

| $N$ | $dim$ | NN | FT | ST | DCG |
|---|---|---|---|---|---|
| 12 | 1728 | 61.19 | 35.37 | 44.55 | 61.08 |
| 10 | 1000 | 61.41 | 35.23 | 44.51 | 60.83 |
| 8 | 512 | 60.75 | 34.55 | 43.66 | 60.18 |
| 6 | 216 | 56.12 | 32.36 | 42.40 | 58.36 |
| 4 | 64 | 43.88 | 23.46 | 32.59 | 50.54 |

PSB Train Database

| $N$ | $dim$ | NN | FT | ST | DCG |
|---|---|---|---|---|---|
| 12 | 1728 | 67.14 | 37.69 | 48.34 | 64.19 |
| 10 | 1000 | 65.93 | 37.65 | 48.39 | 64.18 |
| 8 | 512 | 65.05 | 37.26 | 48.20 | 63.59 |
| 6 | 216 | 59.87 | 35.11 | 46.28 | 61.86 |
| 4 | 64 | 46.75 | 25.90 | 35.83 | 53.79 |

MPEG7 Database

| $N$ | $dim$ | NN | FT | ST | DCG |
|---|---|---|---|---|---|
| 12 | 1728 | 85.46 | 62.96 | 74.84 | 84.46 |
| 10 | 1000 | 85.02 | 62.48 | 74.29 | 84.32 |
| 8 | 512 | 84.58 | 62.62 | 73.62 | 83.99 |
| 6 | 216 | 82.38 | 61.55 | 73.68 | 83.06 |
| 4 | 64 | 82.38 | 54.65 | 72.63 | 80.51 |

SHREC09 Database

| $N$ | $dim$ | NN | FT | ST | DCG |
|---|---|---|---|---|---|
| 12 | 1728 | 86.25 | 53.58 | 64.54 | 73.59 |
| 10 | 1000 | 85.25 | 53.65 | 64.56 | 73.49 |
| 8 | 512 | 85.50 | 53.25 | 64.00 | 72.95 |
| 6 | 216 | 82.00 | 49.97 | 61.70 | 70.19 |
| 4 | 64 | 66.63 | 35.32 | 47.74 | 56.42 |

TAB. 6.2 – Mesures de performance (%) de l'approche 3DGA pour cinq valeurs du paramètre $N$, avec $\sigma = 0.23$, la boîte englobante BB et la distance $d_v$.

Les performances de tous les descripteurs de tailles supérieures à 216 sont équivalentes. Cependant, il faut garder un $N = 8$ relativement petit pour rester compétitif, en termes de taille du descripteur et donc en termes de coût de stockage et de calcul de la mesure de similarité.

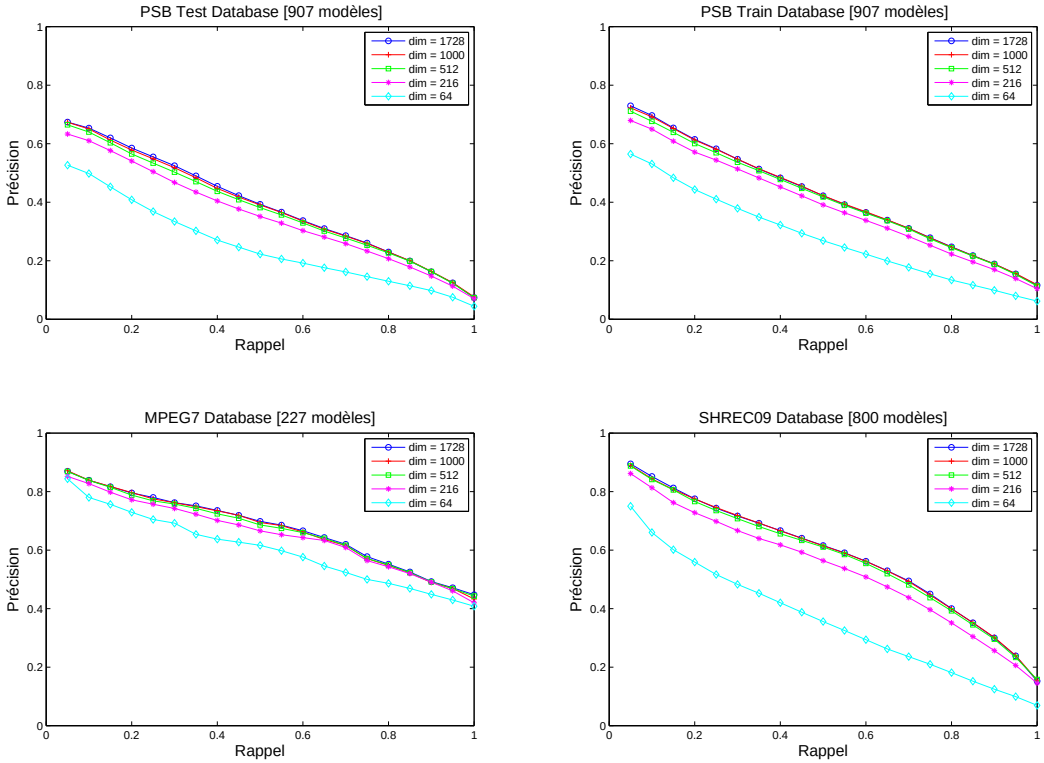



FIG. 6.2 – Courbes de précision-rappel de l'approche 3DGA pour cinq valeurs du paramètre $N$, avec $\sigma = 0.23$, la boîte englobante BB et la distance $d_v$.

La boîte englobante constitue le support de la description gaussienne 3D. Le choix de la boîte englobante est un problème intéressant dans la construction du descripteur. Pour ce faire, nous avons mené une expérimentation comparant les quatre boîtes englobantes : $CC$, $BB$, $EBB$ et $CBC$ (cf. Annexe A.1.1). Les courbes précision-rappel et les mesures de performance, illustrant les résultats avec ces boîtes, sont données sur les Figures 6.3 et les Tableaux 6.3.

PSB Test Database

| boîte | NN | FT | ST | DCG |
|---|---|---|---|---|
| BB | 60.75 | 34.55 | 43.66 | 60.18 |
| EBB | 46.31 | 24.00 | 31.96 | 51.23 |
| CC | 45.65 | 23.30 | 31.94 | 50.63 |
| CBC | 44.87 | 22.27 | 30.54 | 49.92 |

PSB Train Database

| boîte | NN | FT | ST | DCG |
|---|---|---|---|---|
| BB | 65.05 | 37.26 | 48.20 | 63.59 |
| EBB | 51.05 | 24.01 | 31.99 | 52.35 |
| CC | 43.88 | 21.29 | 29.47 | 49.24 |
| CBC | 43.00 | 20.55 | 28.69 | 48.63 |

MPEG7 Database

| boîte | NN | FT | ST | DCG |
|---|---|---|---|---|
| BB | 84.58 | 62.62 | 73.62 | 83.99 |
| EBB | 79.74 | 58.29 | 70.66 | 80.48 |
| CC | 78.41 | 51.52 | 64.76 | 78.02 |
| CBC | 75.33 | 55.19 | 67.36 | 78.76 |

SHREC09 Database

| boîte | NN | FT | ST | DCG |
|---|---|---|---|---|
| BB | 85.50 | 53.25 | 64.00 | 72.95 |
| EBB | 69.50 | 34.19 | 44.29 | 54.30 |
| CC | 68.50 | 31.61 | 42.00 | 51.40 |
| CBC | 62.63 | 29.80 | 39.59 | 49.03 |

TAB. 6.3 – Mesures de performance (%) de l'approche 3DGA pour quatre types de boîte englobante, avec $N = 8$, $\sigma = 0.23$ et la distance $d_v$.

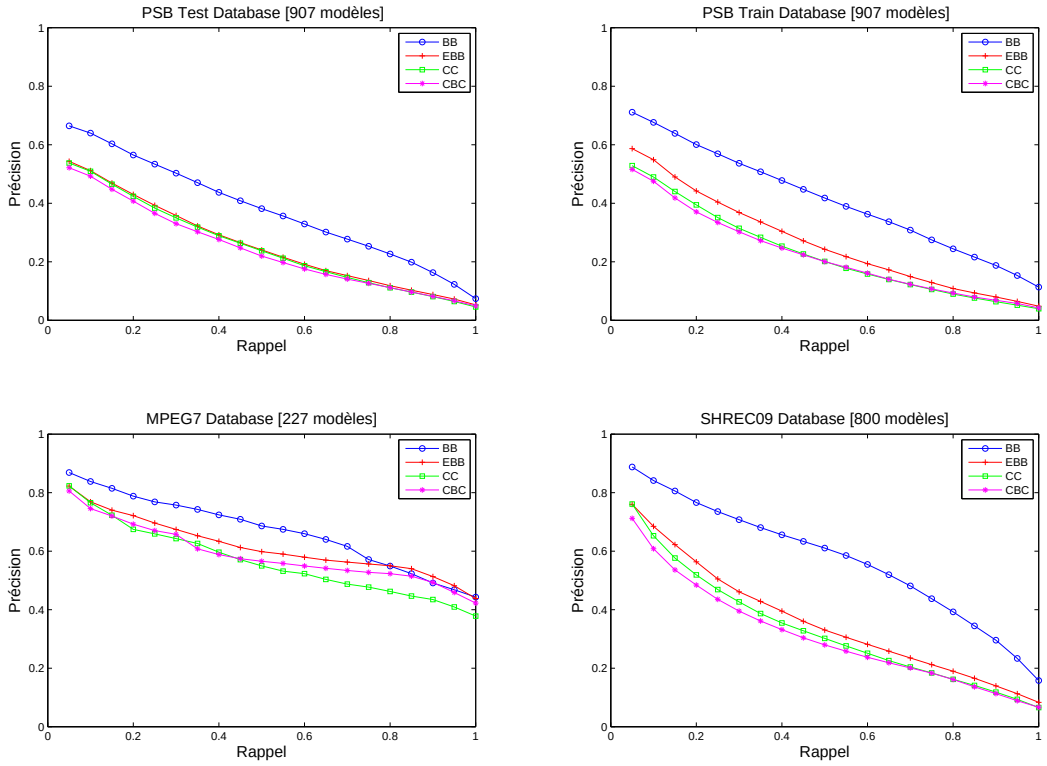



FIG. 6.3 – Courbes de précision-rappel de l'approche 3DGA pour quatre types de boîte englobante, avec $N = 8$, $\sigma = 0.23$ et la distance $d_v$.

Les meilleurs résultats sont obtenus avec le descripteur construit sur la boîte englobante $BB$, la plus petite boîte englobant l'objet 3D. Notre approche se veut l'exploitation d'une représentation multi-échelle alliant les informations de directions et de courbures pour l'extraction de primitives pertinentes aux frontières des formes. Notre choix s'est porté sur la boîte englobante $BB$ qui présente un potentiel très intéressant pour la caractérisation des points de l'ensemble $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$.

Les mêmes tests ont été réalisés pour évaluer la variance $\sigma$ intervenant dans le calcul des coefficients du descripteur gaussien 3D définis dans l'équation (5.2). Nous avons cherché à optimiser la valeur $\sigma$ en termes de performance. Les expérimentations ont été conduites en deux phases : une première est faite en comparant vingt valeurs réparties uniformément sur l'intervalle $[0, 1]$, puis, connaissant la valeur optimale $\sigma = 0.25$, une deuxième est réalisée sur les huit valeurs voisines appartenant à l'intervalle $]0.2, 0.3[$ privé de la valeur optimale calculée.

PSB Test Database

| $\sigma$ | NN | FT | ST | DCG |
|---|---|---|---|---|
| 1.00 | 42.34 | 21.96 | 31.06 | 49.65 |
| 0.75 | 48.40 | 25.96 | 35.54 | 52.96 |
| 0.50 | 54.14 | 30.04 | 39.51 | 56.46 |
| 0.25 | 59.65 | 34.36 | 43.72 | 60.09 |
| 0.23 | 60.75 | 34.55 | 43.66 | 60.18 |
| 0.05 | 48.07 | 24.24 | 32.33 | 51.22 |

PSB Train Database

| $\sigma$ | NN | FT | ST | DCG |
|---|---|---|---|---|
| 1.00 | 42.01 | 22.09 | 31.23 | 50.24 |
| 0.75 | 47.41 | 26.13 | 36.04 | 53.89 |
| 0.50 | 56.56 | 30.31 | 40.54 | 57.96 |
| 0.25 | 64.06 | 37.00 | 47.84 | 63.50 |
| 0.23 | 65.05 | 37.26 | 48.20 | 63.59 |
| 0.05 | 50.50 | 25.60 | 34.06 | 53.20 |

MPEG7 Database

| $\sigma$ | NN | FT | ST | DCG |
|---|---|---|---|---|
| 1.00 | 77.53 | 54.95 | 65.37 | 78.73 |
| 0.75 | 84.14 | 58.26 | 66.84 | 80.68 |
| 0.50 | 84.14 | 61.05 | 70.07 | 82.48 |
| 0.25 | 85.02 | 62.85 | 73.46 | 83.81 |
| 0.23 | 84.58 | 62.62 | 73.62 | 83.99 |
| 0.05 | 79.74 | 56.82 | 71.85 | 80.06 |

SHREC09 Database

| $\sigma$ | NN | FT | ST | DCG |
|---|---|---|---|---|
| 1.00 | 63.00 | 32.62 | 44.74 | 52.83 |
| 0.75 | 68.38 | 37.16 | 49.06 | 57.72 |
| 0.50 | 75.62 | 41.99 | 54.64 | 63.19 |
| 0.25 | 85.00 | 52.76 | 63.88 | 72.65 |
| 0.23 | 85.50 | 53.25 | 64.00 | 72.95 |
| 0.05 | 70.00 | 36.49 | 47.76 | 56.75 |

TAB. 6.4 – Mesures de performance (%) de l'approche 3DGA pour six valeurs du paramètre $\sigma$, avec $N = 8$, la boîte englobante BB et la distance $d_v$.

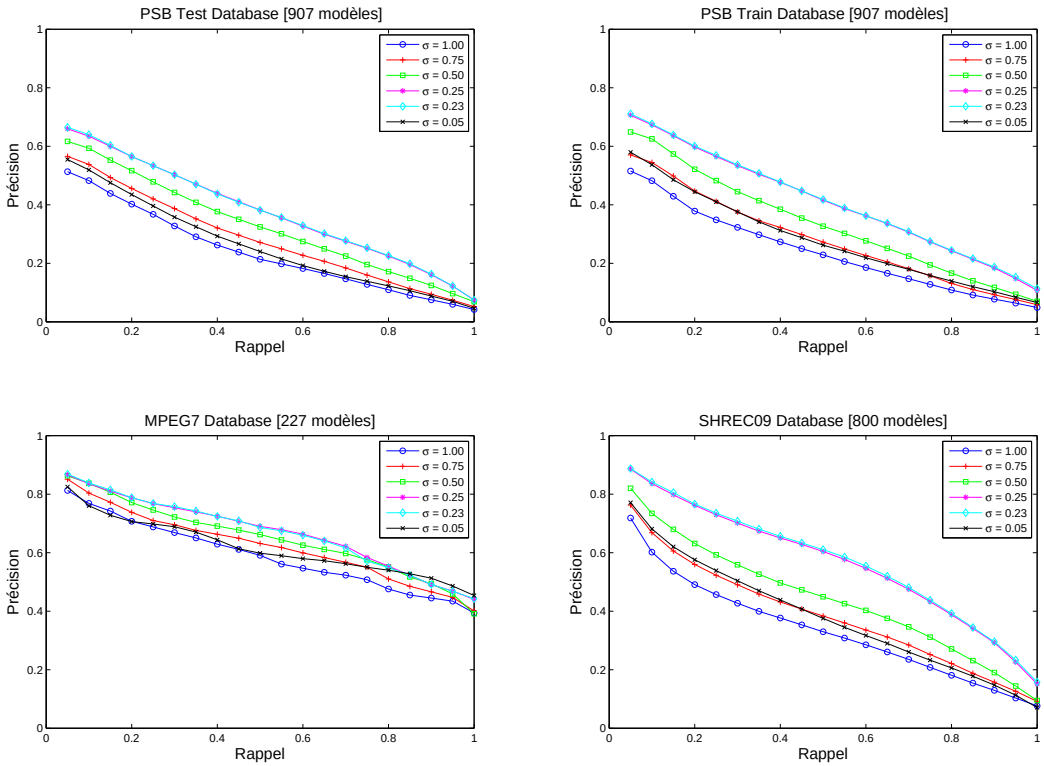



FIG. 6.4 – Courbes de précision-rappel de l'approche 3DGA pour six valeurs du paramètre $\sigma$, avec $N = 8$, la boîte englobante BB et la distance $d_v$.

Les résultats avec $\sigma \in \{0.05, 0.23, 0.25, 0.5, 0.75, 1\}$ sont présentés sur les Figures 6.4 et les Tableaux 6.4. Nous pouvons remarquer que les courbes de précision-rappel et les mesures de performance qui représentent $\sigma = 0.23$ sont les meilleures.

Afin d'associer la meilleure mesure de similarité au descripteur, nous avons étudié et comparé la performance des distances définies dans la section 5.1.2. Les Figures 6.5 et les Tableaux 6.5 montrent les résultats, en termes de courbes précision-rappel et en termes de critères de performance obtenus avec le *3D Gaussian Descriptor* de taille 512 pour toutes les distances $d = l_1$, $l_2$ et $d_v$.

PSB Test Database

| Distance | NN | FT | ST | DCG |
|---|---|---|---|---|
| $d_v$ | 60.75 | 34.55 | 43.66 | 60.18 |
| $l_1$ | 59.43 | 32.43 | 41.53 | 58.81 |
| $l_2$ | 55.79 | 30.07 | 38.17 | 56.28 |

PSB Train Database

| Distance | NN | FT | ST | DCG |
|---|---|---|---|---|
| $d_v$ | 65.05 | 37.26 | 48.20 | 63.59 |
| $l_1$ | 63.73 | 35.80 | 45.99 | 62.15 |
| $l_2$ | 61.41 | 33.03 | 41.83 | 59.41 |

MPEG7 Database

| Distance | NN | FT | ST | DCG |
|---|---|---|---|---|
| $d_v$ | 84.58 | 62.62 | 73.62 | 83.99 |
| $l_1$ | 84.58 | 58.93 | 70.18 | 81.15 |
| $l_2$ | 81.94 | 57.38 | 66.60 | 80.86 |

SHREC09 Database

| Distance | NN | FT | ST | DCG |
|---|---|---|---|---|
| $d_v$ | 85.50 | 53.25 | 64.00 | 72.95 |
| $l_1$ | 83.13 | 49.68 | 60.81 | 70.27 |
| $l_2$ | 82.38 | 44.72 | 54.94 | 65.19 |

TAB. 6.5 – Mesures de performance (%) de l'approche 3DGA pour les différentes distances, avec $N = 8$, $\sigma = 0.23$ et la boîte englobante BB.

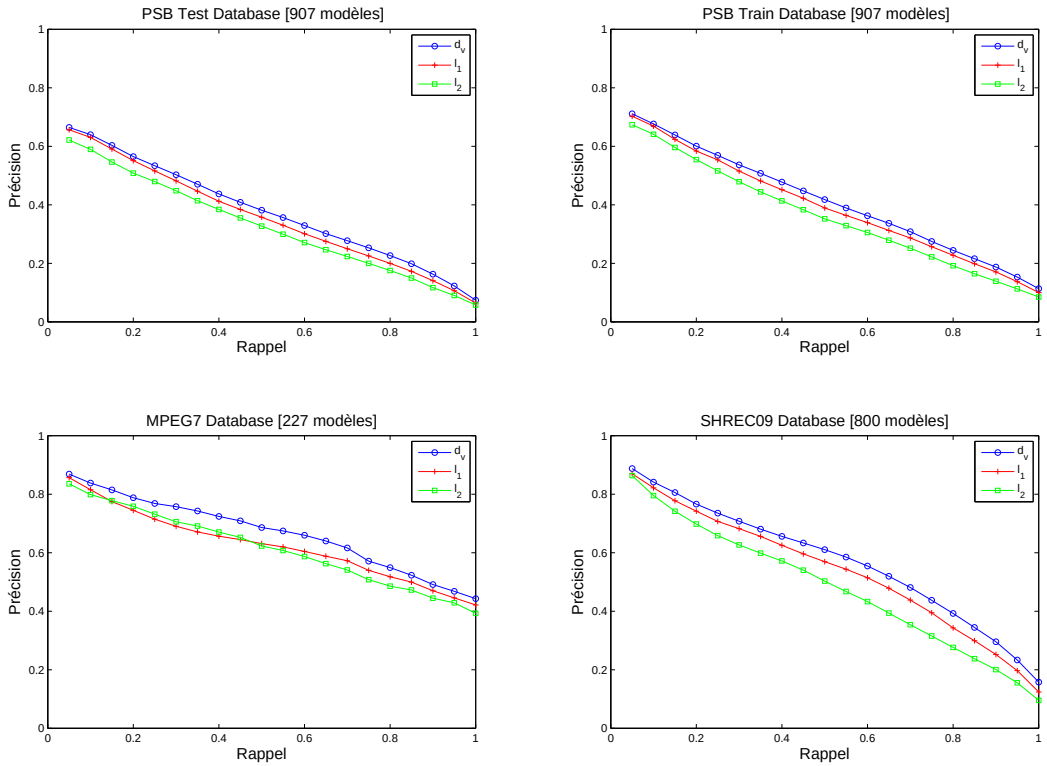



FIG. 6.5 – Courbes de précision-rappel de l'approche 3DGA pour les différentes distances, avec $N = 8$, $\sigma = 0.23$ et la boîte englobante BB.

Les résultats obtenus avec une variance $\sigma = 0.23$ mettent en valeur la performance de la méthode utilisant la distance $d_v$. La supériorité de la distance $d_v$ par rapport aux distances euclidiennes est en revanche pénalisée par une complexité de calcul treize fois plus importante. Nous pouvons également faire remarquer que cette mesure de similarité semble a priori mal adaptée pour les descripteurs de taille inférieure à 64. Une étude comparative sur des expérimentations obtenues avec $N = 4$ et les différentes distances montre une supériorité des distances euclidiennes. En effet, la représentation grossière avec une petite valeur de $N$ diminue la performance de la distance $d_v$.

Enfin, nous avons effectué des tests sur les méthodes de normalisation retenues dans le processus général de la recherche par le contenu de modèles 3D. Les résultats les comparant sont résumés dans les Figures 6.6 et les Tableaux 6.6.

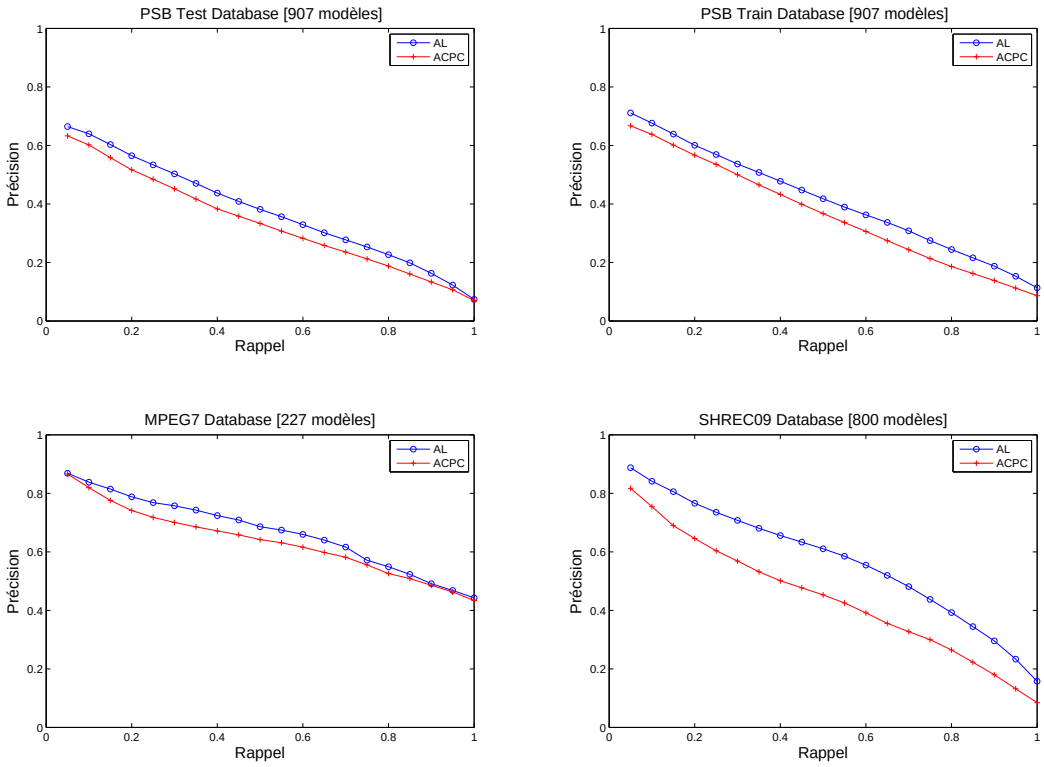



FIG. 6.6 – Courbes de précision-rappel de l'approche 3DGA pour les méthodes d'alignement, avec $N = 8$, $\sigma = 0.23$, la boîte englobante BB et la distance $d_v$.

Notre méthode d'alignement obtient des résultats supérieurs à l'ACPC, ce qui montre l'intérêt de notre méthode dans la normalisation et dans le processus gé-

néral de la recherche par le contenu d'objets 3D.

PSB Test Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 60.75 | 34.55 | 43.66 | 60.18 |
| ACPC | 56.89 | 30.50 | 39.64 | 56.85 |

PSB Train Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 65.05 | 37.26 | 48.20 | 63.59 |
| ACPC | 59.54 | 33.01 | 43.45 | 60.12 |

MPEG7 Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 84.58 | 62.62 | 73.62 | 83.99 |
| ACPC | 85.02 | 60.06 | 71.95 | 82.05 |

SHREC09 Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 85.50 | 53.25 | 64.00 | 72.95 |
| ACPC | 76.25 | 41.18 | 52.05 | 62.37 |

TAB. 6.6 – Mesures de performance (%) de l'approche 3DGA pour les deux méthodes d'alignement, avec $N = 8$, $\sigma = 0.23$, la boîte englobante BB et la distance $d_v$.

Ainsi, les résultats nous ont permis de fixer les paramètres optimaux du descripteur gaussien 3D. Nous recommandons d'utiliser une signature de dimension 512 ($N = 8$), déduite après une étape complète de normalisation dont notre méthode d'alignement fait partie. Nous suggérons d'utiliser les paramètres suivants : échantillonnage uniforme - $E_U$ avec pondération par la taille des facettes, le nombre minimal de points échantillonnés $n_{min}$ - $10^5$, la boîte englobante - $BB$, et la variance $\sigma$ - 0.23. Pour la mesure de similarité, la nouvelle distance $d_v$ (5.13) se révèle être utile dans la mesure où $N \geq 6$.

### 6.2.2 *Enhanced Silhouette based Approach*

Afin d'étudier les performances de notre approche améliorée proposée dans la section 5.2.3.1, nous avons appliqué notre technique au descripteur fondé sur le calcul des silhouettes, proposé par Vranic [131]. Dans la suite de cette section, nous reprenons les mêmes paramètres de ce descripteur qui ont été optimisés en termes de performances par l'auteur. Les résultats présentés dans la suite sont obtenus avec : les images de silhouettes de taille $256 \times 256$ ($N = 256$), la boîte de projection étant le cube canonique englobant $CBC$, le nombre de points contours sélectionnés ($K = 256$), et la taille du vecteur caractéristique ($dim = 300$). La mesure de similarité associée à ce descripteur est la distance euclidienne $l_1$.

Dans un premier temps, nous avons comparé les différents modèles d'indices de pertinence définis pour les silhouettes (cf. section 5.2.2.1). Les Figures 6.7 et les Tableaux 6.7 présentent respectivement les courbes précision-rappel et les mesures de performance, NN, FT, ST et DCG, obtenues avec les indices de per-

tinence normalisés $\overline{R} = R_a$, $R_{sa}$ et $R_c$.

PSB Test Database

| Pertinence | NN | FT | ST | DCG |
|---|---|---|---|---|
| $R_c$ | 57.44 | 32.24 | 43.32 | 59.91 |
| $R_a$ | 54.36 | 32.02 | 43.69 | 59.22 |
| $R_{sa}$ | 57.77 | 32.62 | 44.42 | 60.24 |

PSB Train Database

| Pertinence | NN | FT | ST | DCG |
|---|---|---|---|---|
| $R_c$ | 53.69 | 31.19 | 41.87 | 59.01 |
| $R_a$ | 53.14 | 30.69 | 41.68 | 58.50 |
| $R_{sa}$ | 54.47 | 31.98 | 43.18 | 59.80 |

MPEG7 Database

| Pertinence | NN | FT | ST | DCG |
|---|---|---|---|---|
| $R_c$ | 87.67 | 66.07 | 74.45 | 85.54 |
| $R_a$ | 88.55 | 66.76 | 75.58 | 86.19 |
| $R_{sa}$ | 87.23 | 67.12 | 75.62 | 86.20 |

SHREC09 Database

| Pertinence | NN | FT | ST | DCG |
|---|---|---|---|---|
| $R_c$ | 73.75 | 40.22 | 54.93 | 63.88 |
| $R_a$ | 71.88 | 39.43 | 54.45 | 62.93 |
| $R_{sa}$ | 73.50 | 41.05 | 56.07 | 64.65 |

TAB. 6.7 – Mesures de performance (%) de l'approche ESA pour les différents modèles d'indice de pertinence.

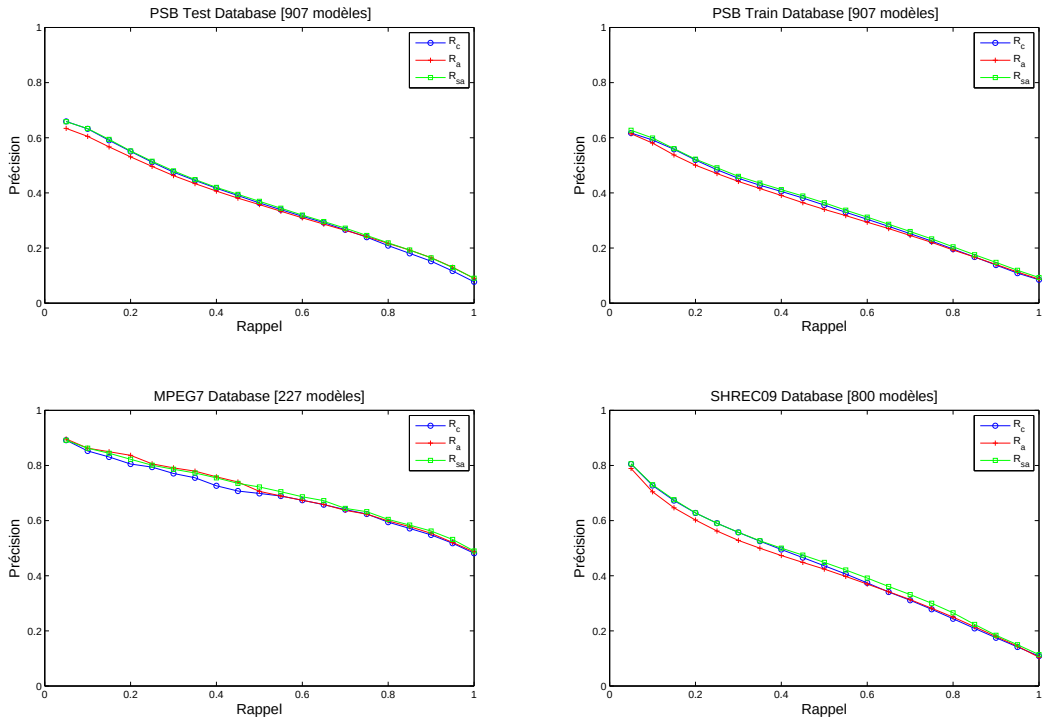



FIG. 6.7 – Courbes de précision-rappel de l'approche ESA pour les différents modèles d'indice de pertinence.

Les courbes et les mesures de performances sur les quatre bases d'objets sont homogènes, avec une légère supériorité de l'indice de pertinence $R_{sa}$. Les autres indices de pertinence donnent des résultats équivalents à 1% près en termes de scores de performance.

Dans une autre série d'expérimentation, nous avons examiné l'influence des

différentes méthodes d'alignement sur les performances de la recherche. Les Figures 6.8 et les Tableaux 6.8 résument les résultats, en termes de courbes précision-rappel et en termes des mesures d'évaluation NN, FT, ST, et DCG, obtenus avec l'approche améliorée et le modèle d'indice de pertinence retenu $R_{sa}$.

PSB Test Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 57.77 | 32.62 | 44.42 | 60.24 |
| ACPC | 55.79 | 30.44 | 41.16 | 58.17 |

PSB Train Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 54.47 | 31.98 | 43.18 | 59.80 |
| ACPC | 54.91 | 31.27 | 42.34 | 59.08 |

MPEG7 Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 87.23 | 67.12 | 75.62 | 86.20 |
| ACPC | 87.67 | 65.98 | 75.57 | 85.30 |

SHREC09 Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 73.50 | 41.05 | 56.07 | 64.65 |
| ACPC | 71.00 | 38.88 | 53.21 | 61.83 |

TAB. 6.8 – Mesures de performance (%) de l'approche ESA pour les deux méthodes d'alignement, avec l'indice de pertinence $R_{sa}$.

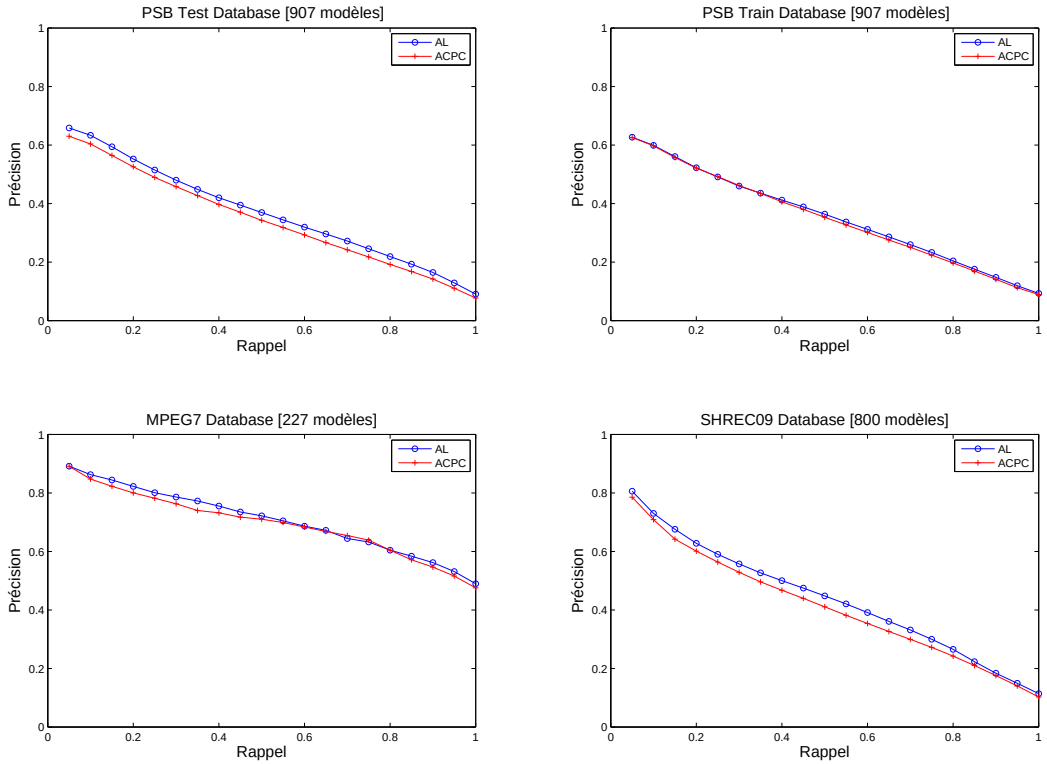



FIG. 6.8 – Courbes de précision-rappel de l'approche ESA pour les deux méthodes d'alignement, avec l'indice de pertinence $R_{sa}$.

Notre méthode d'alignement proposée dans le chapitre 3 s'est encore révélée plus performante que celle à base d'une ACPC.

### 6.2.3 *Enhanced Depth-Buffer based Approach*

Comme énoncé dans la section 5.2.3.2, notre technique d'amélioration est appliquée à l'approche multi-vues à base d'images de profondeur, proposée par Heczko et al. [56]. Notre choix des attributs du descripteur retenu a été guidé par l'étude comparative faite par Vranic dans [131]. Le choix des valeurs des paramètres du descripteur et la mesure de similarité associée a été établi de manière empirique. Les meilleurs résultats sont donnés avec une description de Fourier sur des images de profondeur de taille $256 \times 256$, extraites des projections de l'objet 3D sur la boîte englobante étendue $EBB$, la taille du vecteur caractéristique étant $dim = 438$. La similarité entre deux objets est mesurée avec la distance euclidienne $l_1$ entre les vecteurs caractéristiques. L'évaluation détaillée dans la suite portera seulement sur les modèles d'indice de pertinence et les méthodes d'alignement.

Afin de trouver la meilleure pertinence associée à ce descripteur, nous avons comparé les performances des indices de pertinence définis pour les images de profondeur dans la section 5.2.2.2. Les Figures 6.9 et les Tableaux 6.9 montrent les résultats, en termes de courbes précision-rappel et en termes de mesures de performance obtenues avec la méthode 2D/3D améliorée pour tous les modèles d'indices normalisés $\overline{R} = R_a$, $R_{sa}$, $R_d$ et $R_g$.

PSB Test Database

| Pertinence | NN | FT | ST | DCG |
|---|---|---|---|---|
| $R_g$ | 65.38 | 38.32 | 49.67 | 64.08 |
| $R_d$ | 65.71 | 38.19 | 49.40 | 64.05 |
| $R_a$ | 65.93 | 37.72 | 49.02 | 63.98 |
| $R_{sa}$ | 64.61 | 37.91 | 48.99 | 63.69 |

PSB Train Database

| Pertinence | NN | FT | ST | DCG |
|---|---|---|---|---|
| $R_g$ | 64.06 | 36.11 | 46.58 | 63.25 |
| $R_d$ | 64.39 | 36.67 | 47.13 | 63.44 |
| $R_a$ | 63.95 | 35.60 | 45.75 | 62.76 |
| $R_{sa}$ | 65.16 | 36.28 | 46.56 | 63.26 |

MPEG7 Database

| Pertinence | NN | FT | ST | DCG |
|---|---|---|---|---|
| $R_g$ | 89.87 | 67.24 | 77.10 | 87.42 |
| $R_d$ | 89.87 | 66.74 | 76.89 | 87.19 |
| $R_a$ | 91.19 | 67.56 | 77.63 | 87.62 |
| $R_{sa}$ | 91.63 | 66.91 | 77.62 | 87.41 |

SHREC09 Database

| Pertinence | NN | FT | ST | DCG |
|---|---|---|---|---|
| $R_g$ | 81.38 | 44.32 | 56.45 | 66.09 |
| $R_d$ | 81.50 | 44.49 | 56.45 | 66.33 |
| $R_a$ | 80.00 | 42.77 | 55.39 | 65.04 |
| $R_{sa}$ | 81.38 | 43.11 | 54.76 | 64.97 |

TAB. 6.9 – Mesures de performance (%) de l'approche EDBA pour les différents modèles d'indice de pertinence.

Les performances de tous les indices de pertinence sont équivalentes pour les quatre bases d'objets 3D, avec une très légère supériorité pour les indices de pertinence $R_g$ et $R_d$.

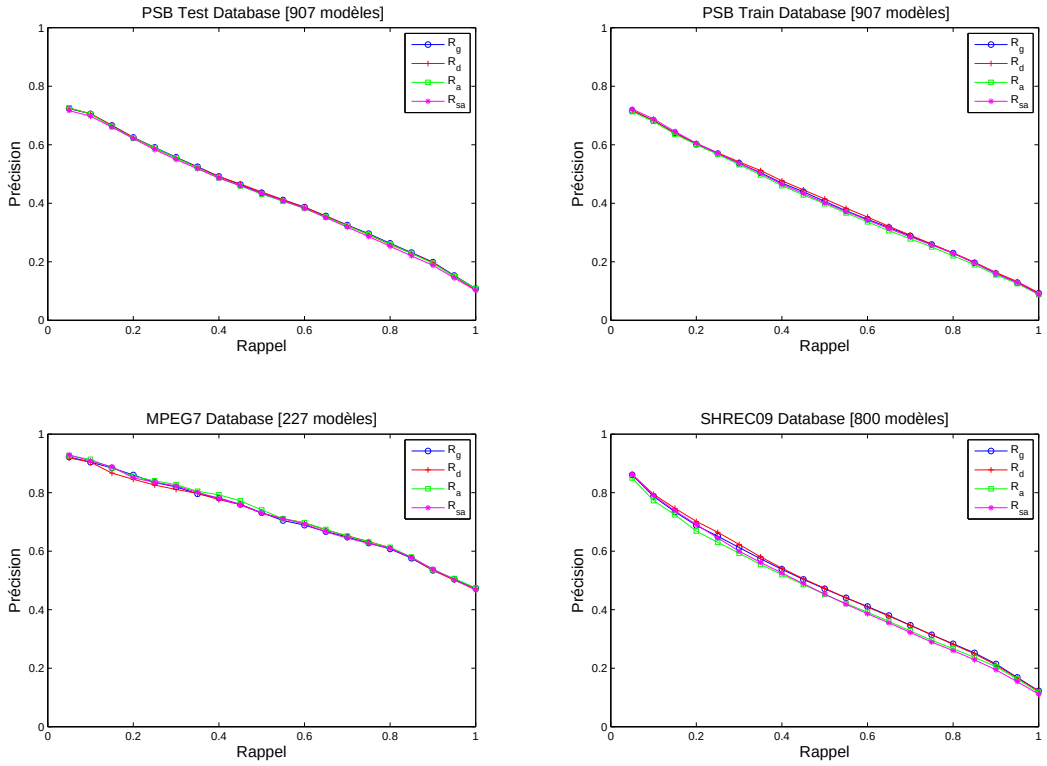



FIG. 6.9 – Courbes de précision-rappel de l'approche EDBA pour les différents modèles d'indice de pertinence.

La deuxième évaluation menée sur le même descripteur concerne les deux principales techniques d'alignement retenues dans l'étape de la normalisation des modèles 3D. Les résultats comparatifs sont résumés dans les Figures 6.10 et les Tableaux 6.10.

PSB Test Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 65.38 | 38.32 | 49.67 | 64.08 |
| ACPC | 60.53 | 34.28 | 44.15 | 60.28 |

PSB Train Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 64.06 | 36.11 | 46.58 | 63.25 |
| ACPC | 62.96 | 34.16 | 44.35 | 61.38 |

MPEG7 Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 89.87 | 67.24 | 77.10 | 87.42 |
| ACPC | 89.43 | 66.55 | 76.98 | 86.54 |

SHREC09 Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 81.38 | 44.32 | 56.45 | 66.09 |
| ACPC | 74.50 | 39.76 | 51.38 | 60.61 |

TAB. 6.10 – Mesures de performance (%) de l'approche EDBA pour les deux méthodes d'alignement, avec l'indice de pertinence $R_g$.

Ces résultats confirment les conclusions précédemment énoncées sur la robustesse de notre méthode d'alignement (cf. section 3.3) et sa capacité à améliorer consi-

dérablement les performances de la recherche par rapport à la méthode classique fondée sur une analyse en composante principales.

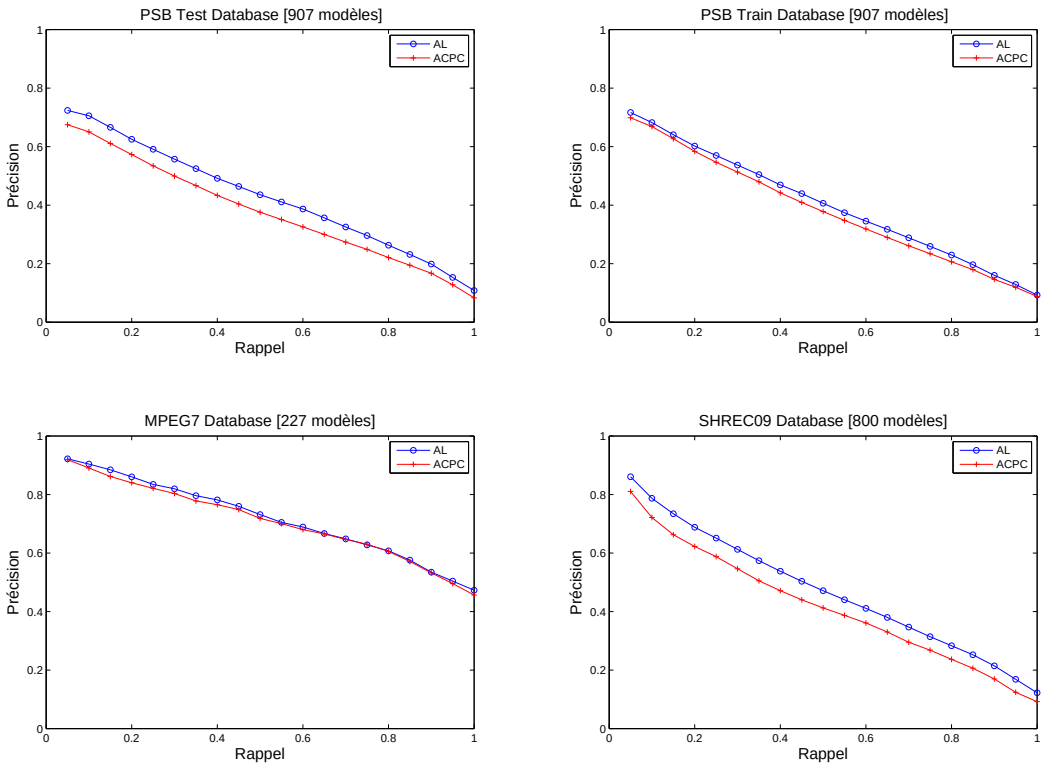



FIG. 6.10 – Courbes de précision-rappel de l'approche EDBA pour les deux méthodes d'alignement, avec l'indice de pertinence $R_g$.

### 6.2.4 *Depth Line Approach*

Des séries d'expériences ont été réalisées pour optimiser les performances du *Depth Line Approach*. Dans un premier temps, nous avons évalué les paramètres du descripteur : la résolution de l'image de profondeur $N$ et la boîte englobante. Puis, nous avons testé les différentes mesures de similarité associées aux séquences d'états, décrites dans la section 5.3.3.1 et enfin, nous avons comparé la robustesse de la description obtenue avec notre méthode d'alignement à celle avec une analyse en composantes principales continue.

La taille de la signature varie en fonction de la résolution $N$ choisie pour définir les images de profondeur. Nous avons évalué l'influence de ce paramètre avec les valeurs suivantes $N = 8, 16, 32, 64$. Les Figures 6.11 et les Tableaux 6.11

contiennent respectivement les résultats en termes de courbes précision-rappel et de mesures de performance.

PSB Test Database

| $N$ | NN | FT | ST | DCG |
|---|---|---|---|---|
| 64 | 66.26 | 38.67 | 50.75 | 65.17 |
| 32 | 65.38 | 38.79 | 50.77 | 65.26 |
| 16 | 65.05 | 38.32 | 49.75 | 64.64 |
| 8 | 60.86 | 35.92 | 46.47 | 62.52 |

PSB Train Database

| $N$ | NN | FT | ST | DCG |
|---|---|---|---|---|
| 64 | 67.81 | 41.38 | 52.99 | 67.37 |
| 32 | 67.70 | 41.53 | 53.01 | 66.99 |
| 16 | 66.04 | 40.50 | 51.43 | 66.16 |
| 8 | 63.51 | 38.14 | 48.84 | 64.11 |

MPEG7 Database

| $N$ | NN | FT | ST | DCG |
|---|---|---|---|---|
| 64 | 90.31 | 66.15 | 78.16 | 87.04 |
| 32 | 88.55 | 64.95 | 76.01 | 86.03 |
| 16 | 85.46 | 62.47 | 74.13 | 84.41 |
| 8 | 83.70 | 60.42 | 72.11 | 83.08 |

SHREC09 Database

| $N$ | NN | FT | ST | DCG |
|---|---|---|---|---|
| 64 | 87.75 | 61.91 | 74.27 | 81.22 |
| 32 | 87.38 | 62.54 | 74.69 | 81.42 |
| 16 | 88.00 | 61.97 | 74.63 | 81.34 |
| 8 | 88.75 | 58.15 | 71.51 | 78.76 |

TAB. 6.11 – Mesures de performance (%) de l'approche DLA pour quatre valeurs du paramètre $N$, avec la boîte englobante BB et la distance de Hamming $d_H$.

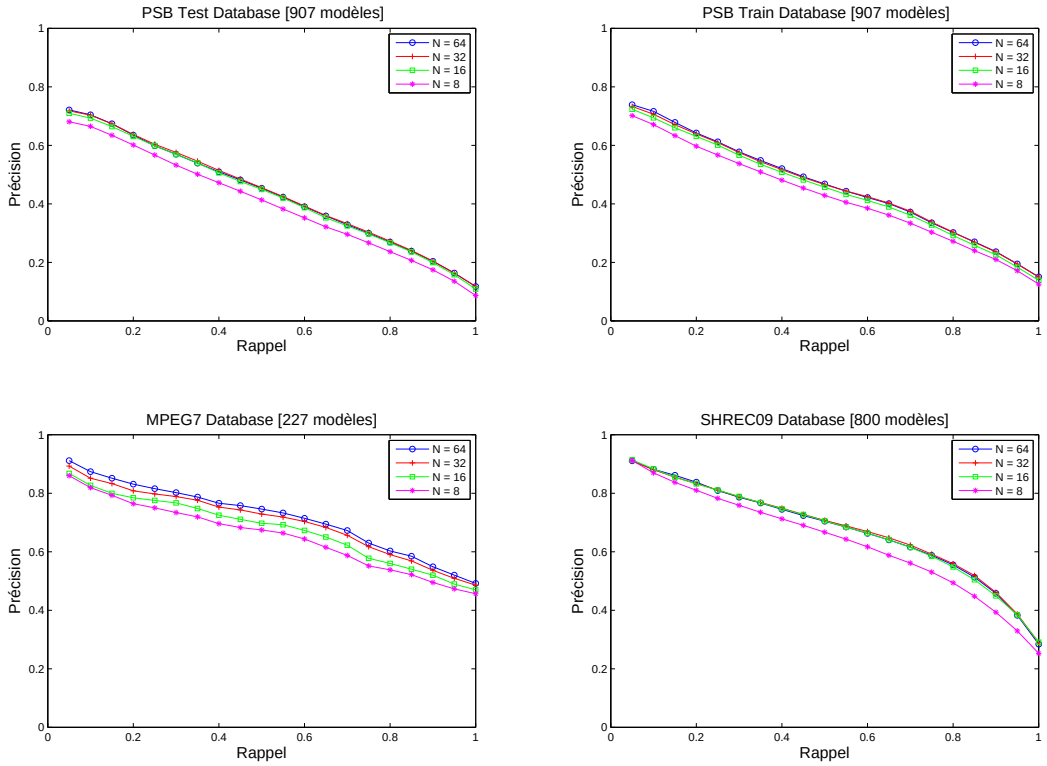



FIG. 6.11 – Courbes de précision-rappel de l'approche DLA pour quatre valeurs du paramètre $N$, avec la boîte englobante BB et la distance de Hamming $d_H$.

Nous notons que les résultats sont homogènes et assez stables à partir de $N = 16$. Les meilleurs résultats sont obtenus avec $N = 32$ pour la base de Princeton et la base de SHREC, et avec $N = 64$ pour la base de MPEG7. Pour conserver un bon compromis performance-complexité, nous fixerons $N = 32$ dans la suite.

Dans le cadre d'une deuxième expérimentation, nous avons évalué les boîtes englobantes suivantes : BB, EBB, CC, CBC (cf. Annexe A.1.1). Les résultats sont présentés sur les courbes précision-rappel de la Figure 6.12 et les Tableaux 6.12 des mesures de performance.

PSB Test Database

| boîte | NN | FT | ST | DCG |
|---|---|---|---|---|
| BB | 65.38 | 38.79 | 50.77 | 65.26 |
| EBB | 59.43 | 33.44 | 42.57 | 59.50 |
| CC | 58.77 | 32.96 | 42.61 | 60.02 |
| CBC | 58.55 | 32.72 | 41.77 | 58.82 |

PSB Train Database

| boîte | NN | FT | ST | DCG |
|---|---|---|---|---|
| BB | 67.70 | 41.53 | 53.01 | 66.99 |
| EBB | 61.85 | 34.74 | 43.40 | 60.89 |
| CC | 61.08 | 33.31 | 43.28 | 60.94 |
| CBC | 58.99 | 33.15 | 42.15 | 59.31 |

MPEG7 Database

| boîte | NN | FT | ST | DCG |
|---|---|---|---|---|
| BB | 88.55 | 64.95 | 76.01 | 86.03 |
| EBB | 87.67 | 64.18 | 74.33 | 84.57 |
| CC | 84.58 | 61.19 | 69.97 | 81.98 |
| CBC | 88.55 | 61.46 | 76.00 | 84.76 |

SHREC09 Database

| boîte | NN | FT | ST | DCG |
|---|---|---|---|---|
| BB | 87.38 | 62.54 | 74.69 | 81.42 |
| EBB | 80.88 | 44.20 | 55.04 | 64.97 |
| CC | 79.13 | 44.45 | 56.61 | 66.42 |
| CBC | 77.25 | 41.62 | 52.19 | 62.16 |

TAB. 6.12 – Mesures de performance (%) de l'approche DLA pour quatre types de boîte englobante, avec $N = 32$ et la distance de Hamming $d_H$.

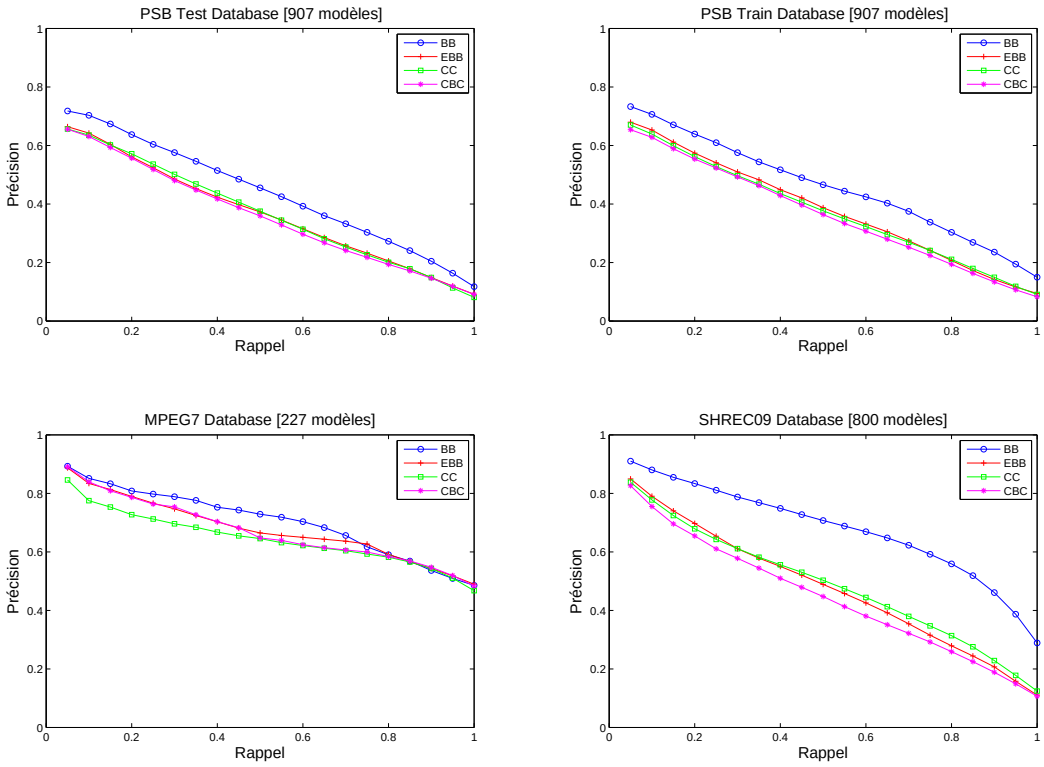



FIG. 6.12 – Courbes de précision-rappel de l'approche DLA pour quatre types de boîte englobante, avec $N = 32$ et la distance de Hamming $d_H$.

Les résultats montrent une nette supériorité de la boîte englobante BB sur toutes les bases d'objets. Les descripteurs calculés sur les autres boîtes obtiennent des performances moyennement faibles et assez proches sur les bases de Princeton et la base de SHREC. Cependant, les résultats sur la base de MPEG7 montrent une faiblesse du descripteur calculé sur la boîte CC et une équivalence de performance, à 1.5% près, des descripteurs calculés sur les autres boîtes.

Les performances de la méthode avec les distances, $d_H$, $d_{NW}$ et $d_{SW}$, utilisées dans la mesure de similarité sont exprimées en termes de courbes précision-rappel (cf. Figures 6.13) et mesures de performance NN, FT, ST et DCG (cf. Tableaux 6.13).

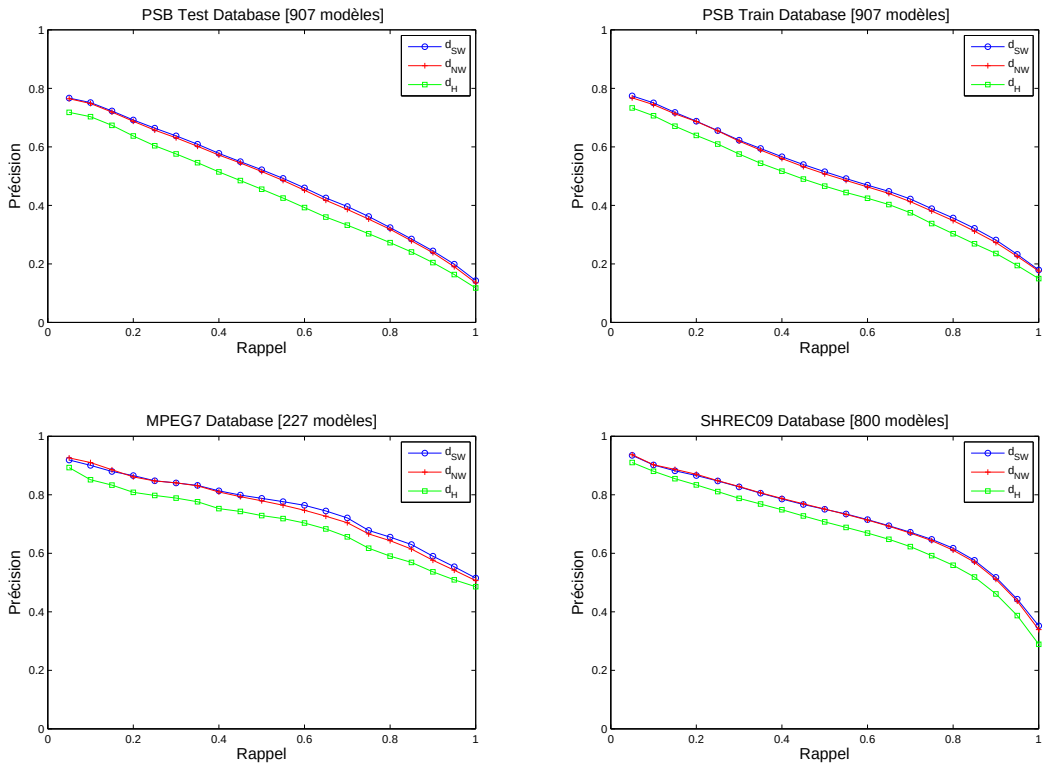



FIG. 6.13 – Courbes de précision-rappel de l'approche DLA pour les différentes distances, avec $N = 32$ et la boîte englobante BB.

Nous observons que les distances de Needleman-Wunsch et Smith-Waterman ont donné des résultats nettement supérieurs à la distance de Hamming. Les résultats confirment que les métriques fondées sur la programmation dynamique sont les plus adaptées pour comparer les séquences d'états. Elles permettent d'établir la

correspondance maximale entre deux séquences, en repérant leurs sous-séquences identiques ou similaires.

PSB Test Database

| Distance | NN | FT | ST | DCG |
|---|---|---|---|---|
| $d_{SW}$ | 70.78 | 43.41 | 55.63 | 69.24 |
| $d_{NW}$ | 71.33 | 42.94 | 55.27 | 68.77 |
| $d_H$ | 65.38 | 38.79 | 50.77 | 65.26 |

PSB Train Database

| Distance | NN | FT | ST | DCG |
|---|---|---|---|---|
| $d_{SW}$ | 71.56 | 45.42 | 57.71 | 70.40 |
| $d_{NW}$ | 70.67 | 45.04 | 57.25 | 70.01 |
| $d_H$ | 67.70 | 41.53 | 53.01 | 66.99 |

MPEG7 Database

| Distance | NN | FT | ST | DCG |
|---|---|---|---|---|
| $d_{SW}$ | 90.75 | 69.95 | 82.00 | 89.25 |
| $d_{NW}$ | 92.07 | 69.24 | 80.92 | 89.03 |
| $d_H$ | 88.55 | 64.95 | 76.01 | 86.03 |

SHREC09 Database

| Distance | NN | FT | ST | DCG |
|---|---|---|---|---|
| $d_{SW}$ | 91.00 | 66.21 | 78.37 | 84.10 |
| $d_{NW}$ | 91.00 | 66.20 | 78.05 | 84.12 |
| $d_H$ | 87.38 | 62.54 | 74.69 | 81.42 |

TAB. 6.13 – Mesures de performance (%) de l'approche DLA pour les différentes distances, avec $N = 32$ et la boîte englobante BB.

Pour conclure notre évaluation, nous avons effectué des tests sur les méthodes d'alignement (AL et ACPC) utilisées lors de la phase de normalisation. Les résultats comparatifs sont présentés dans les Figures 6.14 et les Tableaux 6.14.

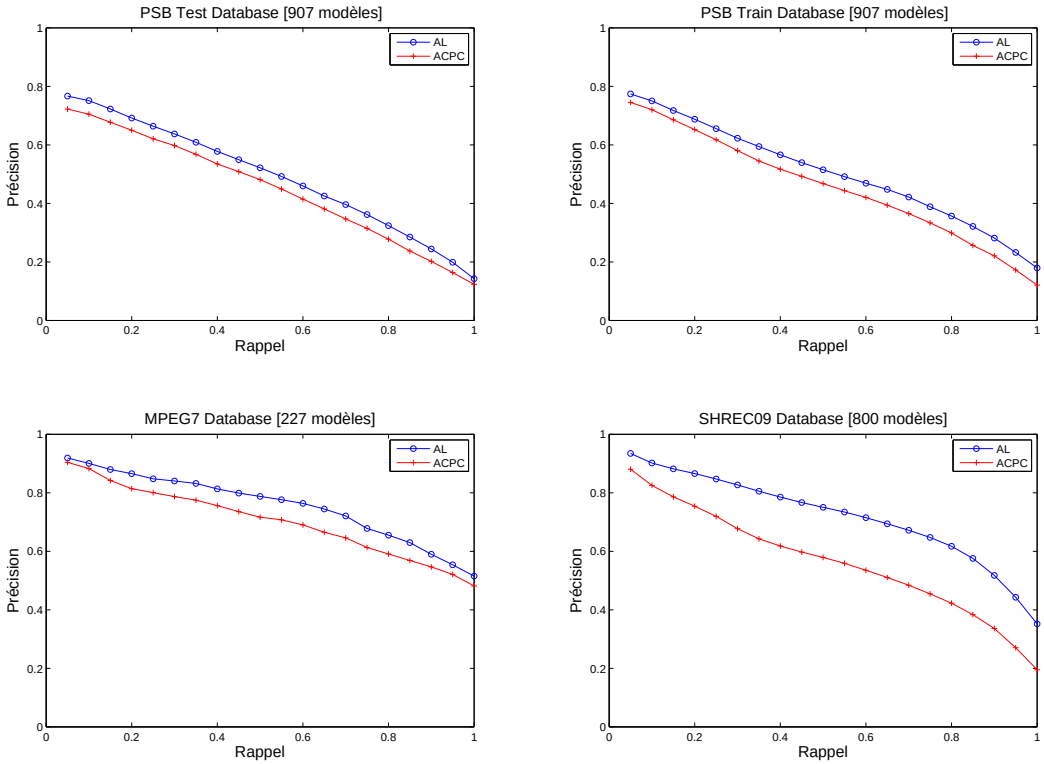



FIG. 6.14 – Courbes de précision-rappel de l'approche DLA pour les deux méthodes d'alignement, avec $N = 32$, la boîte englobante BB et la distance de Smith-Waterman $d_{SW}$.

PSB Test Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 70.78 | 43.41 | 55.63 | 69.24 |
| ACPC | 65.93 | 40.46 | 51.18 | 65.79 |

PSB Train Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 71.56 | 45.42 | 57.71 | 70.40 |
| ACPC | 67.14 | 40.94 | 52.93 | 67.10 |

MPEG7 Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 90.75 | 69.95 | 82.00 | 89.25 |
| ACPC | 88.99 | 64.93 | 78.13 | 85.71 |

SHREC09 Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 91.00 | 66.21 | 78.37 | 84.10 |
| ACPC | 84.50 | 52.07 | 63.40 | 72.01 |

TAB. 6.14 – Mesures de performance (%) de l'approche DLA pour les deux méthodes d'alignement, avec $N = 32$, la boîte englobante BB et la distance de Smith-Waterman $d_{SW}$.

Les résultats montrent une supériorité de notre méthode d'alignement par rapport à l'ACPC. Contrairement aux mêmes expérimentations réalisées sur la base de SHREC avec les approches 3DGA, ESA et EDBA (cf. Figures 6.6, 6.8, 6.10 et Tableaux 6.6, 6.8, 6.10), nous enregistrons un gain important allant jusqu'à 16% en termes de scores, alors que le gain n'est que d'environ 5% sur les autres bases d'objets.

Ces tests nous ont permis de définir les paramètres optimaux de l'approche DLA : Nous recommandons d'utiliser des images de profondeur de résolution $32 \times 32$ projetées sur les faces de la plus petite boîte englobant l'objet 3D, $BB$. Pour la mesure de similarité, nous optons pour les distances, $d_{NW}$ et $d_{SW}$, basées sur la programmation dynamique qui se révèlent être plus performantes que la distance de Hamming.

### 6.2.5 *Multi-view Depth Line Approach*

Trois ensembles d'expérimentations ont été faits pour optimiser les performances du *Multi-view Depth Line Approach*. Le premier a porté sur le paramètre $N$ caractérisant la taille des images de profondeur. Le deuxième a été réalisé avec les différentes distances associées aux séquences d'états, décrites dans la section 5.3.3.1. Le troisième a porté sur les méthodes d'alignement adoptées dans la phase de normalisation.

Les performances de la méthode avec différentes valeurs du paramètre $N = 8$, 16, 32, 64 sur les quatre bases d'objets sont exprimées en termes de courbes précision-rappel (cf. Figures 6.15) et mesures performance NN, FT, ST et DCG (cf. Tableaux 6.15).

PSB Test Database

| $N$ | NN | FT | ST | DCG |
|---|---|---|---|---|
| 64 | 68.91 | 43.41 | 54.21 | 68.20 |
| 32 | 67.48 | 43.22 | 53.91 | 67.55 |
| 16 | 67.37 | 42.49 | 53.20 | 67.04 |
| 8 | 67.59 | 41.84 | 52.31 | 66.27 |

PSB Train Database

| $N$ | NN | FT | ST | DCG |
|---|---|---|---|---|
| 64 | 72.33 | 45.29 | 56.20 | 70.26 |
| 32 | 72.55 | 44.58 | 55.96 | 69.86 |
| 16 | 71.33 | 43.80 | 54.57 | 68.93 |
| 8 | 70.23 | 41.90 | 52.47 | 67.55 |

MPEG7 Database

| $N$ | NN | FT | ST | DCG |
|---|---|---|---|---|
| 64 | 89.87 | 69.48 | 79.52 | 88.05 |
| 32 | 88.99 | 68.59 | 78.18 | 87.49 |
| 16 | 88.11 | 66.95 | 75.71 | 86.38 |
| 8 | 88.11 | 64.80 | 73.26 | 85.04 |

SHREC09 Database

| $N$ | NN | FT | ST | DCG |
|---|---|---|---|---|
| 64 | 93.38 | 63.19 | 75.51 | 82.89 |
| 32 | 92.63 | 63.66 | 75.84 | 82.99 |
| 16 | 91.75 | 63.05 | 75.71 | 82.65 |
| 8 | 91.00 | 61.33 | 73.88 | 81.39 |

TAB. 6.15 – Mesures de performance (%) de l'approche MDLA pour quatre valeurs du paramètre $N$, avec la distance de Hamming $d_H$.

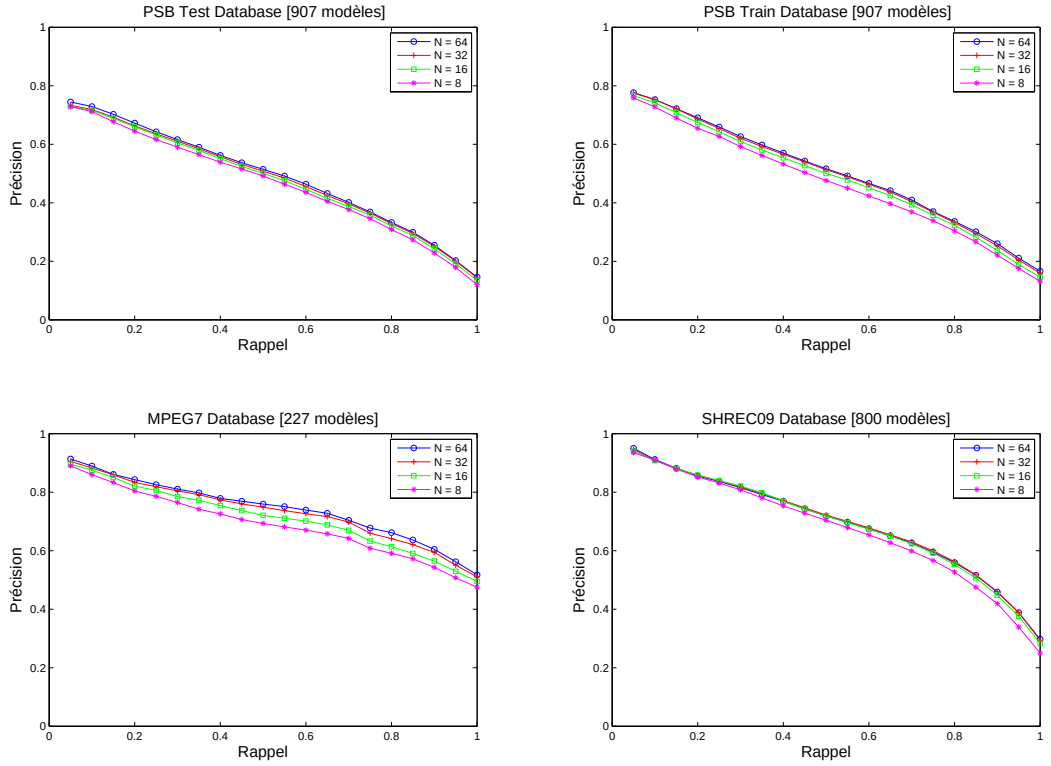



FIG. 6.15 – Courbes de précision-rappel de l'approche MDLA pour quatre valeurs du paramètre $N$, avec la distance de Hamming $d_H$.

A l'instar des résultats obtenus avec la DLA, nous notons des performances homogènes et équivalentes à partir de $N = 16$ pour la MDLA. Les résultats à $N = 32$ et $N = 64$ étant très proches, nous avons opté pour la plus petite résolution ($N = 32$) car les temps de mesures de similarité sont plus courts et la taille de la signature est plus petite.

Les Figures 6.16 et les Tableaux 6.16 présentent respectivement les résultats en termes de courbes précision-rappel et de critères de performance, obtenus avec les distances $d_H$, $d_{NW}$ et $d_{SW}$.

PSB Test Database

| Distance | NN | FT | ST | DCG |
|---|---|---|---|---|
| $d_{SW}$ | 72.77 | 48.32 | 59.92 | 71.77 |
| $d_{NW}$ | 72.11 | 47.39 | 59.23 | 71.16 |
| $d_H$ | 67.48 | 43.22 | 53.91 | 67.55 |

PSB Train Database

| Distance | NN | FT | ST | DCG |
|---|---|---|---|---|
| $d_{SW}$ | 77.51 | 49.43 | 60.87 | 73.57 |
| $d_{NW}$ | 76.41 | 48.96 | 60.78 | 73.23 |
| $d_H$ | 72.55 | 44.58 | 55.96 | 69.86 |

MPEG7 Database

| Distance | NN | FT | ST | DCG |
|---|---|---|---|---|
| $d_{SW}$ | 91.63 | 73.18 | 83.40 | 89.85 |
| $d_{NW}$ | 92.07 | 72.58 | 81.68 | 89.82 |
| $d_H$ | 88.99 | 68.59 | 78.18 | 87.49 |

SHREC09 Database

| Distance | NN | FT | ST | DCG |
|---|---|---|---|---|
| $d_{SW}$ | 94.25 | 67.72 | 79.28 | 85.87 |
| $d_{NW}$ | 94.75 | 67.47 | 79.57 | 85.71 |
| $d_H$ | 92.63 | 63.66 | 75.84 | 82.99 |

TAB. 6.16 – Mesures de performance (%) de l'approche MDLA pour les différentes distances, avec $N = 32$.

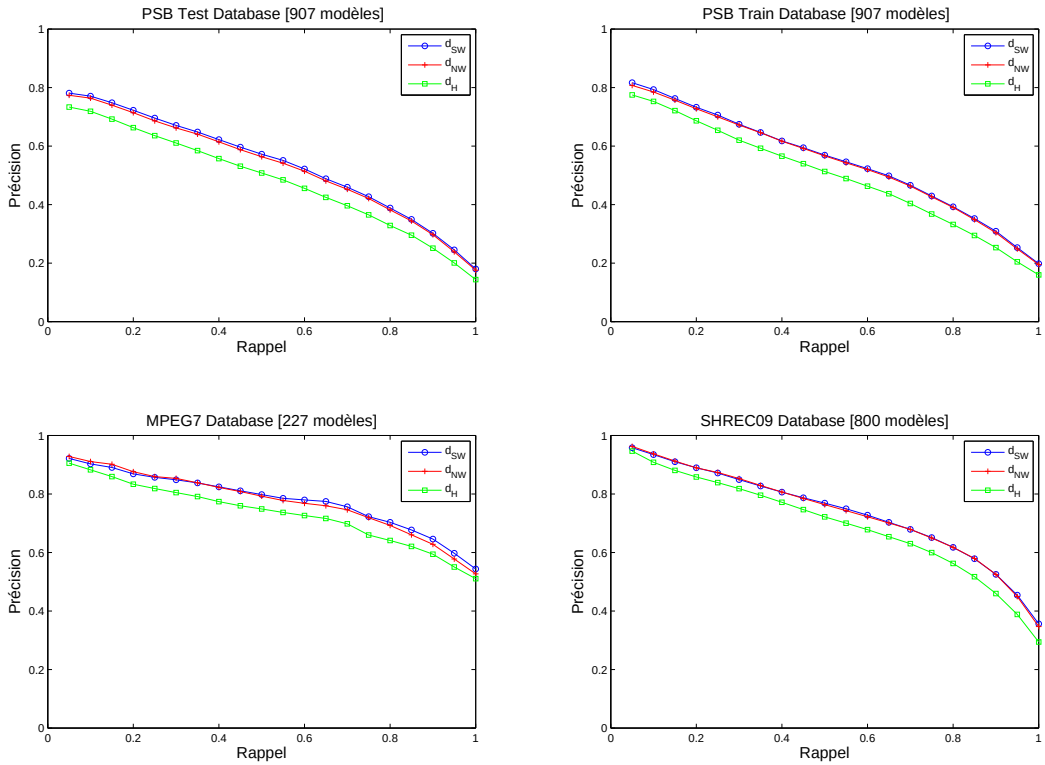



FIG. 6.16 – Courbes de précision-rappel de l'approche MDLA pour les différentes distances, avec $N = 32$.

Les résultats obtenus avec le descripteur MDLA montrent une nette supériorité des distances fondées sur la programmation dynamique. Nous observons un léger avantage de la distance de Smith-Waterman qui calcule le meilleur alignement

local entre deux séquences.

Les deux méthodes d'alignement ont été testées sur les différentes bases d'objets 3D. Les résultats sont reportés sur les Figures 6.17 et les Tableaux 6.17 qui présentent respectivement les courbes précision-rappel et les mesures de performance, NN, FT, ST et DCG.

PSB Test Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 72.77 | 48.32 | 59.92 | 71.77 |
| ACPC | 68.58 | 44.06 | 54.33 | 68.11 |

PSB Train Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 77.51 | 49.43 | 60.87 | 73.57 |
| ACPC | 72.33 | 44.55 | 55.53 | 69.83 |

MPEG7 Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 91.63 | 73.18 | 83.40 | 89.85 |
| ACPC | 89.87 | 69.13 | 81.00 | 87.65 |

SHREC09 Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 94.25 | 67.72 | 79.28 | 85.87 |
| ACPC | 87.50 | 52.26 | 63.41 | 72.98 |

TAB. 6.17 – Mesures de performance (%) de l'approche MDLA pour les deux méthodes d'alignement, avec $N = 32$ et la distance de Smith-Waterman $d_{SW}$.

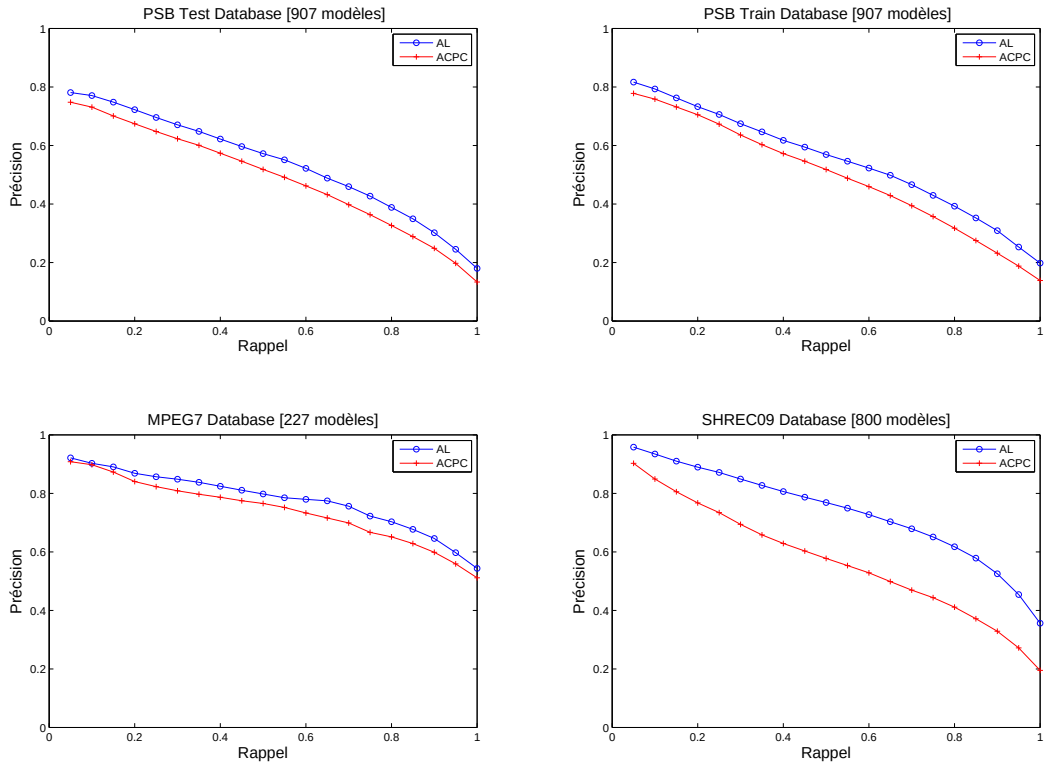



FIG. 6.17 – Courbes de précision-rappel de l'approche MDLA pour les deux méthodes d'alignement, avec $N = 32$ et la distance de Smith-Waterman $d_{SW}$.

Comparée à l'ACPC, notre méthode d'alignement se révèle plus performante. Celle-ci parvient à mieux aligner des objets de la même classe permettant ainsi une similarité plus pertinente et invariante aux rotations.

Les résultats des tests aux différents niveaux de résolution $N$, avec les trois distances associées aux MDLA, nous ont permis de faire ces choix optimisant les performances de la recherche : la résolution $N = 32$ et la distance de Smith-Waterman $d_{SW}$. Nous recommandons d'utiliser notre méthode d'alignement dans l'étape de normalisation.

### 6.2.6 *Cylindrical Depth Line Approach*

Afin d'optimiser les performances du *Cylindrical Depth Line Approach*, une série de tests ont été réalisés sur les bases de données décrites dans la section 1.4. Pour ce faire, nous avons évalué le paramètre $N$ caractérisant la taille de l'image de profondeur cylindrique, puis, nous avons testé les différentes distances associées aux séquences d'états et enfin, nous avons comparé la robustesse de la description obtenue avec notre méthode d'alignement à celle avec une analyse en composantes principales.

Comme énoncé dans la section 5.3.2.4, le paramètre $N$ fixe la taille de la séquence d'états ($M = 4 \times N$) et, par là même, la taille du descripteur. Pour garantir une recherche rapide, nous avons choisi d'évaluer trois valeurs de $N$, $N \in \{8, 16, 32\}$. Elles donnent respectivement des séquences de taille raisonnable $M \in \{32, 64, 128\}$. Les Figures 6.18 et les Tableaux 6.18 présentent respectivement les courbes précision-rappel et les mesures de performance, NN, FT, ST et DCG.

PSB Test Database

| $N$ | NN | FT | ST | DCG |
|---|---|---|---|---|
| 32 | 55.90 | 31.58 | 42.59 | 58.20 |
| 16 | 58.66 | 35.27 | 46.93 | 61.74 |
| 8 | 60.86 | 35.30 | 46.73 | 61.65 |

PSB Train Database

| $N$ | NN | FT | ST | DCG |
|---|---|---|---|---|
| 32 | 56.89 | 33.73 | 44.71 | 60.43 |
| 16 | 63.18 | 37.76 | 48.11 | 64.01 |
| 8 | 62.85 | 36.20 | 47.23 | 63.01 |

MPEG7 Database

| $N$ | NN | FT | ST | DCG |
|---|---|---|---|---|
| 32 | 77.09 | 58.85 | 73.96 | 81.55 |
| 16 | 83.26 | 65.75 | 76.59 | 84.66 |
| 8 | 84.14 | 64.17 | 74.78 | 83.71 |

SHREC09 Database

| $N$ | NN | FT | ST | DCG |
|---|---|---|---|---|
| 32 | 80.13 | 50.36 | 63.34 | 70.32 |
| 16 | 85.75 | 54.88 | 68.94 | 76.26 |
| 8 | 85.63 | 54.67 | 67.97 | 76.06 |

TAB. 6.18 – Mesures de performance (%) de l'approche CDLA pour trois valeurs du paramètre $N$, avec la distance de Hamming $d_H$.

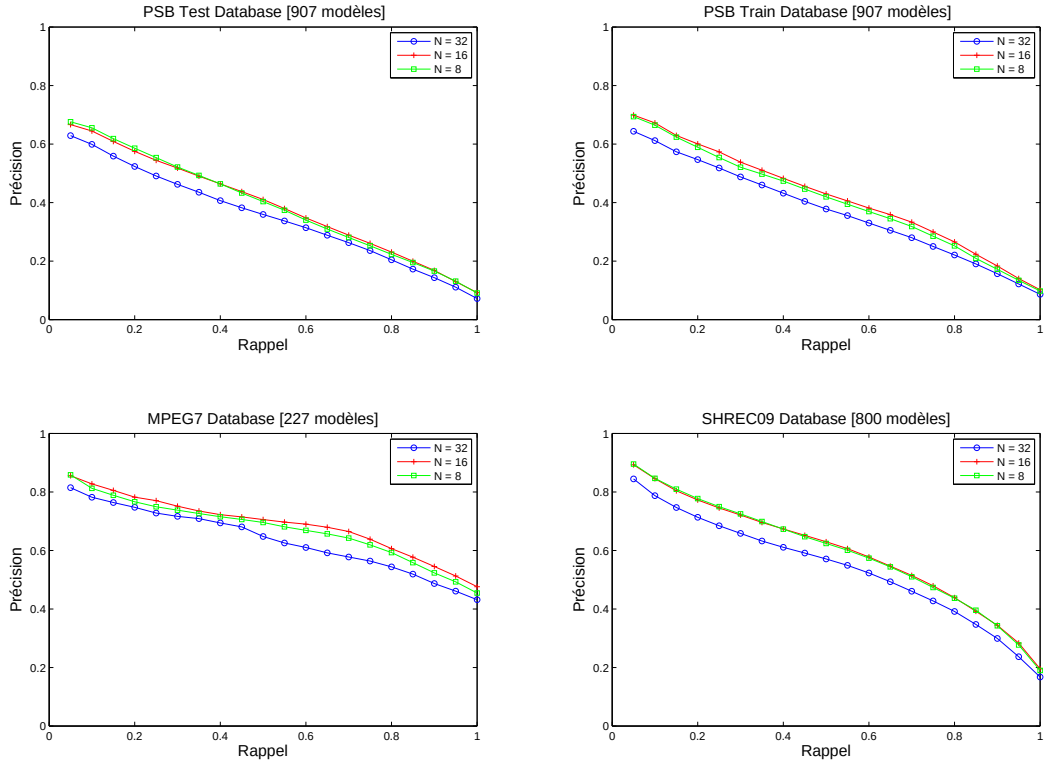



FIG. 6.18 – Courbes de précision-rappel de l'approche CDLA pour trois valeurs du paramètre $N$, avec la distance de Hamming $d_H$.

Les résultats sont homogènes sur les quatre bases d'objets. Les meilleures performances sont obtenues pour $N = 16$.

Afin d'associer la meilleure mesure de similarité au descripteur CDLA, nous avons comparé les performances des distances, $d_{SW}$, $d_{NW}$ et $d_H$, définies dans la section 5.3.3.1. Elles sont illustrées dans les Figures 6.20 et les Tableaux 6.20.

PSB Test Database

| Distance | NN | FT | ST | DCG |
|---|---|---|---|---|
| $d_{SW}$ | 65.49 | 37.57 | 49.69 | 64.47 |
| $d_{NW}$ | 64.39 | 37.57 | 49.30 | 64.06 |
| $d_H$ | 58.66 | 35.27 | 46.93 | 61.74 |

PSB Train Database

| Distance | NN | FT | ST | DCG |
|---|---|---|---|---|
| $d_{SW}$ | 69.02 | 40.77 | 52.56 | 67.24 |
| $d_{NW}$ | 68.47 | 40.50 | 52.16 | 66.73 |
| $d_H$ | 63.18 | 37.76 | 48.11 | 64.01 |

MPEG7 Database

| Distance | NN | FT | ST | DCG |
|---|---|---|---|---|
| $d_{SW}$ | 86.78 | 67.98 | 79.24 | 86.36 |
| $d_{NW}$ | 88.11 | 68.02 | 78.72 | 86.51 |
| $d_H$ | 83.26 | 65.75 | 76.59 | 84.66 |

SHREC09 Database

| Distance | NN | FT | ST | DCG |
|---|---|---|---|---|
| $d_{SW}$ | 87.88 | 57.95 | 71.79 | 78.94 |
| $d_{NW}$ | 88.63 | 58.07 | 71.60 | 78.65 |
| $d_H$ | 85.75 | 54.88 | 68.94 | 76.26 |

TAB. 6.19 – Mesures de performance (%) de l'approche CDLA pour les différentes distances, avec $N = 16$.

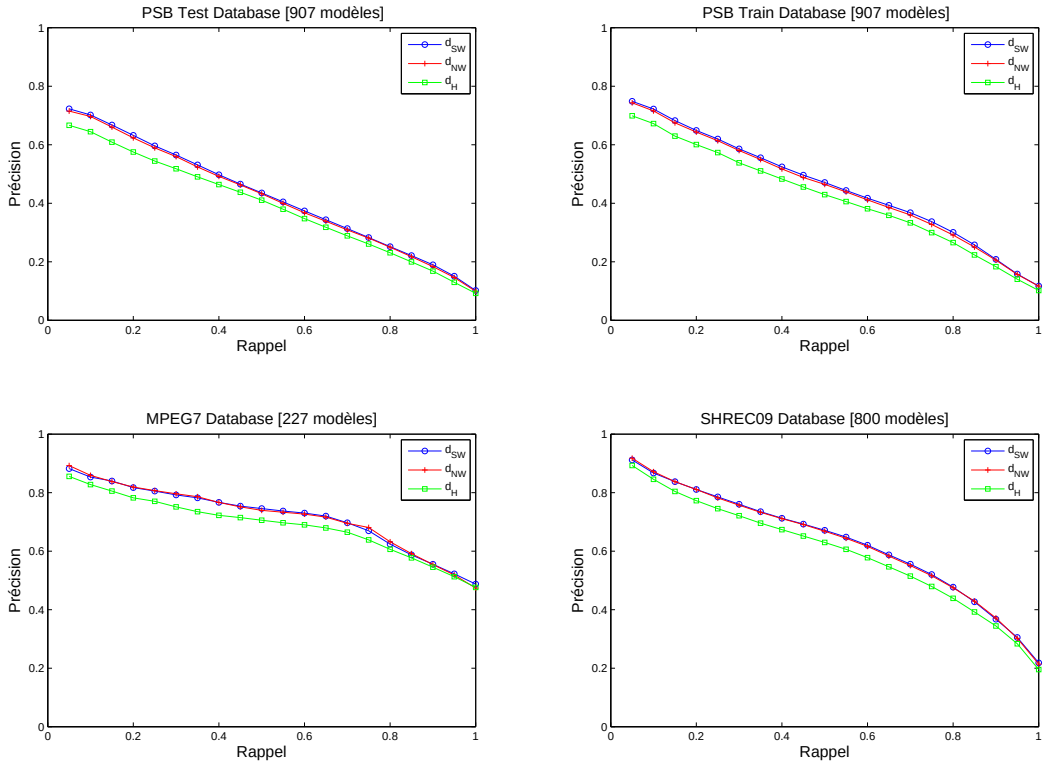



FIG. 6.19 – Courbes de précision-rappel de l'approche CDLA pour les différentes distances, avec $N = 16$.

La distance de Hamming, bien que relativement rapide, fournit des résultats nettement inférieurs aux distances fondées sur la programmation dynamique. Comme dans les approches précédentes à base de lignes de profondeur, les performances des distances de Needleman-Wunsch et Smith-Waterman sont équivalentes à 0.5% près en termes de scores.

PSB Test Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 65.49 | 37.57 | 49.69 | 64.47 |
| ACPC | 63.40 | 35.61 | 45.78 | 62.45 |

PSB Train Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 69.02 | 40.77 | 52.56 | 67.24 |
| ACPC | 65.16 | 37.66 | 48.43 | 64.34 |

MPEG7 Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 86.78 | 67.98 | 79.24 | 86.36 |
| ACPC | 84.58 | 66.29 | 77.70 | 85.59 |

SHREC09 Database

| Alignement | NN | FT | ST | DCG |
|---|---|---|---|---|
| AL | 87.88 | 57.95 | 71.79 | 78.94 |
| ACPC | 81.63 | 46.90 | 59.03 | 67.96 |

TAB. 6.20 – Mesures de performance (%) de l'approche CDLA pour les deux méthodes d'alignement, avec $N = 16$ et la distance de Smith-Waterman $d_{SW}$.

Les performances de l'approche CDLA avec les méthodes d'alignement ACPC et AL utilisées dans la phase de la normalisation des modèles 3D sont exprimées

en termes de courbes précision-rappel (cf. Figures 6.20) et mesures de performance NN, FT, ST et DCG (cf. Tableaux 6.20).

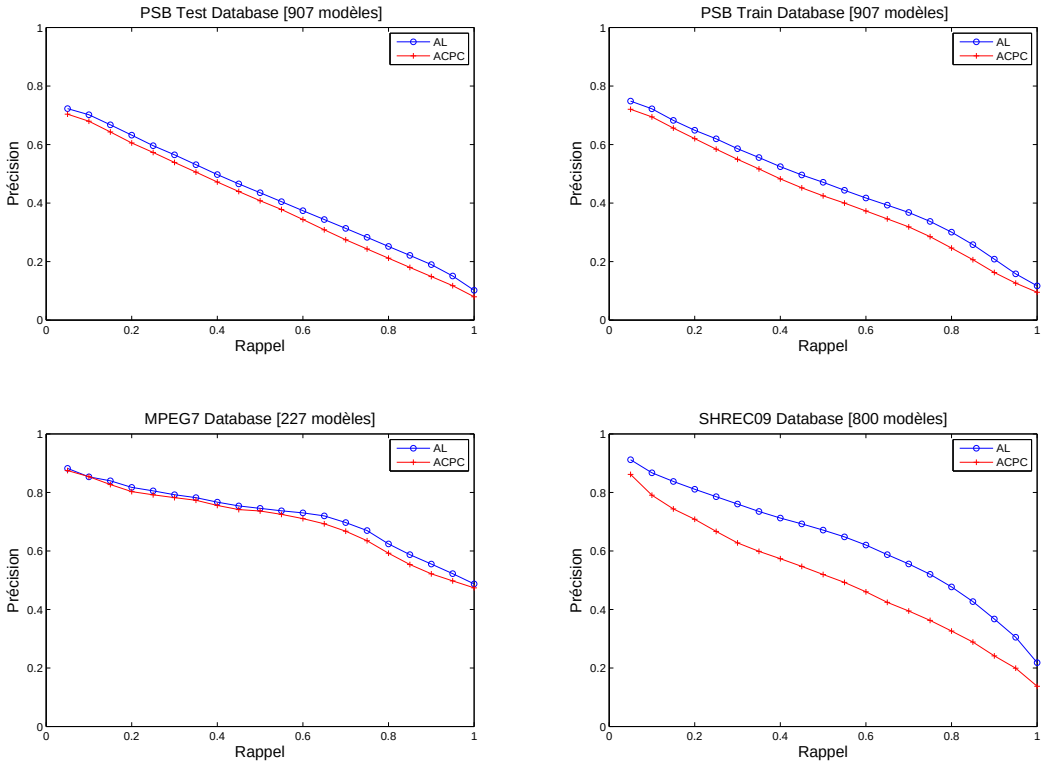



FIG. 6.20 – Courbes de précision-rappel de l'approche CDLA pour les deux méthodes d'alignement, avec $N = 16$ et la distance de Smith-Waterman $d_{SW}$.

À l'instar des résultats obtenus avec les autres approches, notre méthode d'alignement proposée dans le chapitre 3 est meilleure que celle obtenue via une ACPC.

À l'aide des différents tests menés, nous fixons les paramètres permettant d'atteindre les meilleures performances de CDLA. Nous préconisons d'utiliser à la fois la résolution $N = 16$, une des distances fondées sur la programmation dynamique $d_{NW}$ ou $d_{SW}$ et notre méthode d'alignement.

## 6.3 Étude comparative

En s'appuyant sur les résultats expérimentaux obtenus dans la section précédente, nous pouvons à présent comparer les différentes approches proposées. Pour cela, nous avons choisi pour chacune d'entre elles les paramètres et la distance

qui ont fourni les meilleurs résultats sur les quatre bases d'objets 3D. Tous les résultats présentés dans la suite sont obtenus avec notre méthode d'alignement. Les performances des méthodes de recherche par le contenu sont exprimées en termes de courbes précision-rappel (cf. Figures 6.21) et mesures de performance NN, FT, ST et DCG (cf. Tableaux 6.21 et Figures 6.22, 6.23, 6.24, 6.25).

PSB Test Database

| Approche | NN | FT | ST | DCG |
|---|---|---|---|---|
| MDLA | 72.77 | 48.32 | 59.92 | 71.77 |
| DLA | 70.78 | 43.41 | 55.63 | 69.24 |
| CDLA | 65.49 | 37.57 | 49.69 | 64.47 |
| EDBA | 65.38 | 38.32 | 49.67 | 64.08 |
| ESA | 57.77 | 32.62 | 44.42 | 60.24 |
| 3DGA | 60.75 | 34.55 | 43.66 | 60.18 |

PSB Train Database

| Approche | NN | FT | ST | DCG |
|---|---|---|---|---|
| MDLA | 77.51 | 49.43 | 60.87 | 73.57 |
| DLA | 71.56 | 45.42 | 57.71 | 70.40 |
| CDLA | 69.02 | 40.77 | 52.56 | 67.24 |
| EDBA | 64.06 | 36.11 | 46.58 | 63.25 |
| ESA | 54.47 | 31.98 | 43.18 | 59.80 |
| 3DGA | 65.05 | 37.26 | 48.20 | 63.59 |

MPEG7 Database

| Approche | NN | FT | ST | DCG |
|---|---|---|---|---|
| MDLA | 91.63 | 73.18 | 83.40 | 89.85 |
| DLA | 90.75 | 69.95 | 82.00 | 89.25 |
| CDLA | 86.78 | 67.98 | 79.24 | 86.36 |
| EDBA | 89.87 | 67.24 | 77.10 | 87.42 |
| ESA | 87.23 | 67.12 | 75.62 | 86.20 |
| 3DGA | 84.58 | 62.62 | 73.62 | 83.99 |

SHREC09 Database

| Approche | NN | FT | ST | DCG |
|---|---|---|---|---|
| MDLA | 94.25 | 67.72 | 79.28 | 85.87 |
| DLA | 91.00 | 66.21 | 78.37 | 84.10 |
| CDLA | 87.88 | 57.95 | 71.79 | 78.94 |
| EDBA | 81.38 | 44.32 | 56.45 | 66.09 |
| ESA | 73.50 | 41.05 | 56.07 | 64.65 |
| 3DGA | 85.50 | 53.25 | 64.00 | 72.95 |

TAB. 6.21 – Mesures de performance (%) des approches proposées.

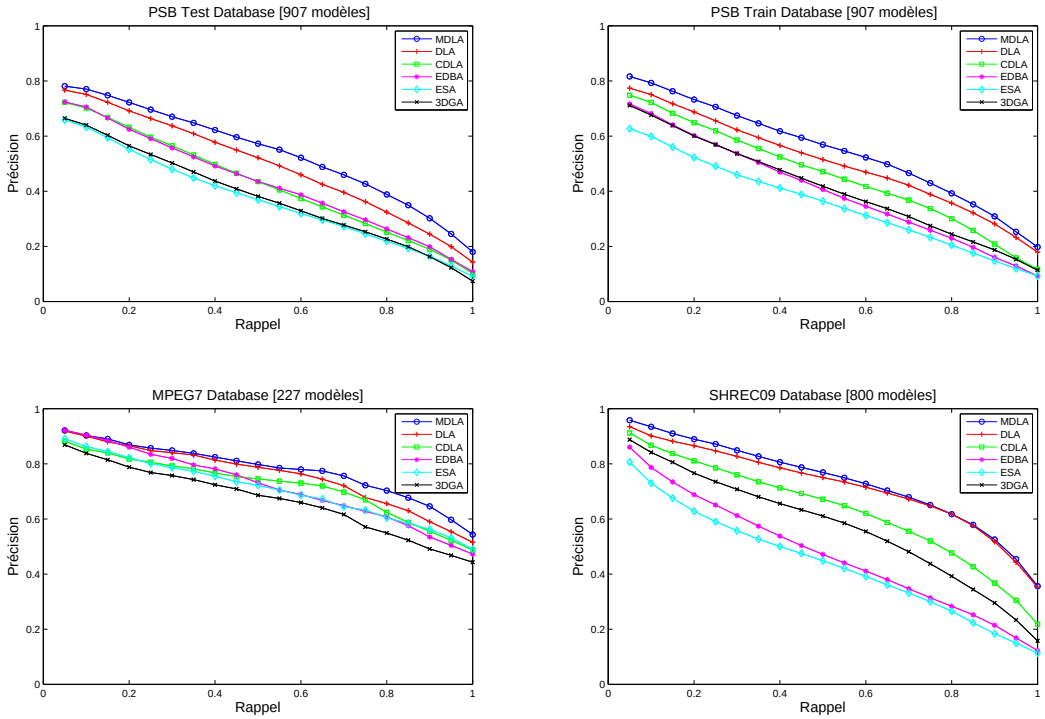



FIG. 6.21 – Courbes de précision-rappel des approches proposées.

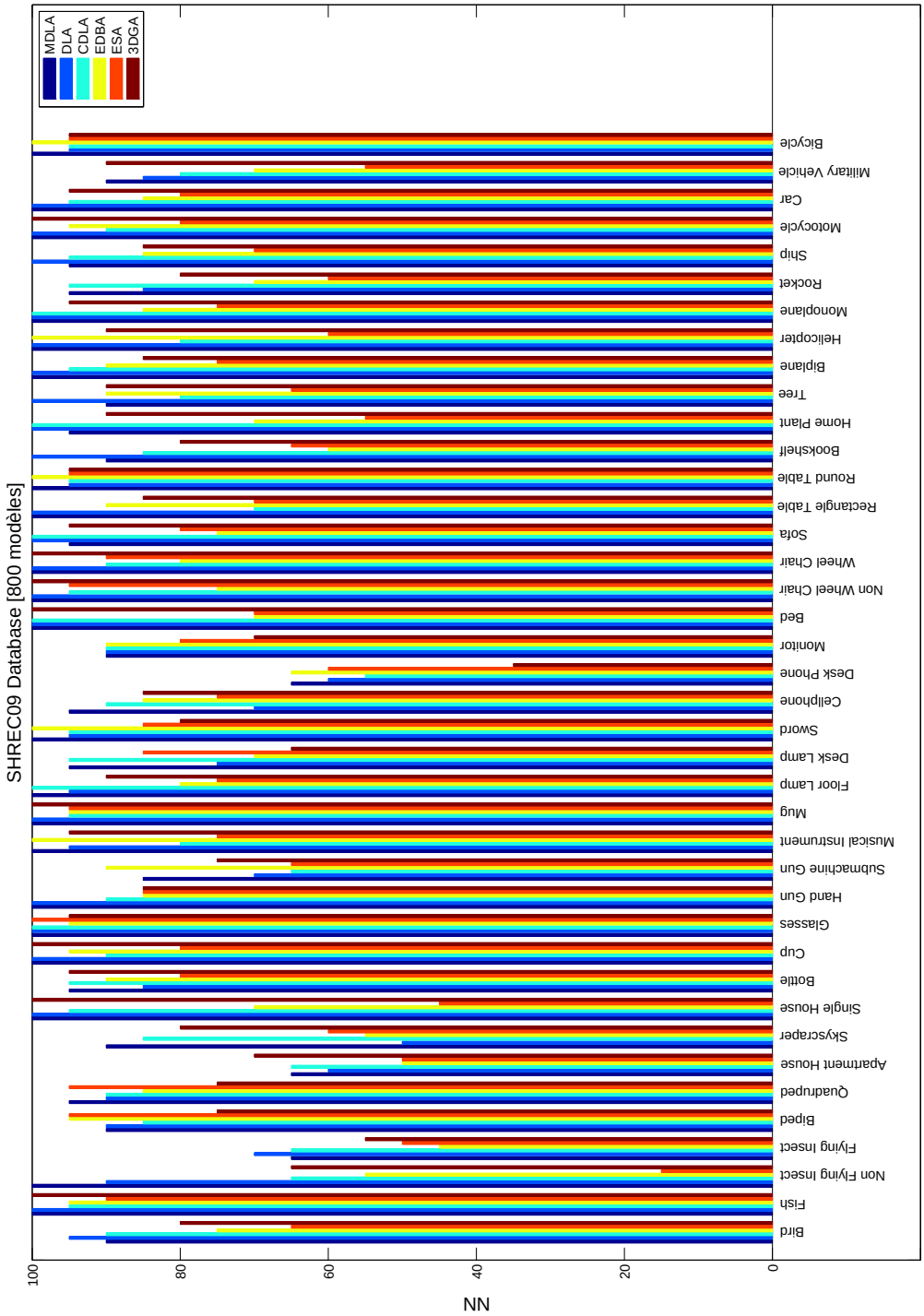



FIG. 6.22 – NN (%) des approches proposées calculés sur les 40 classes de la base de SHREC09.

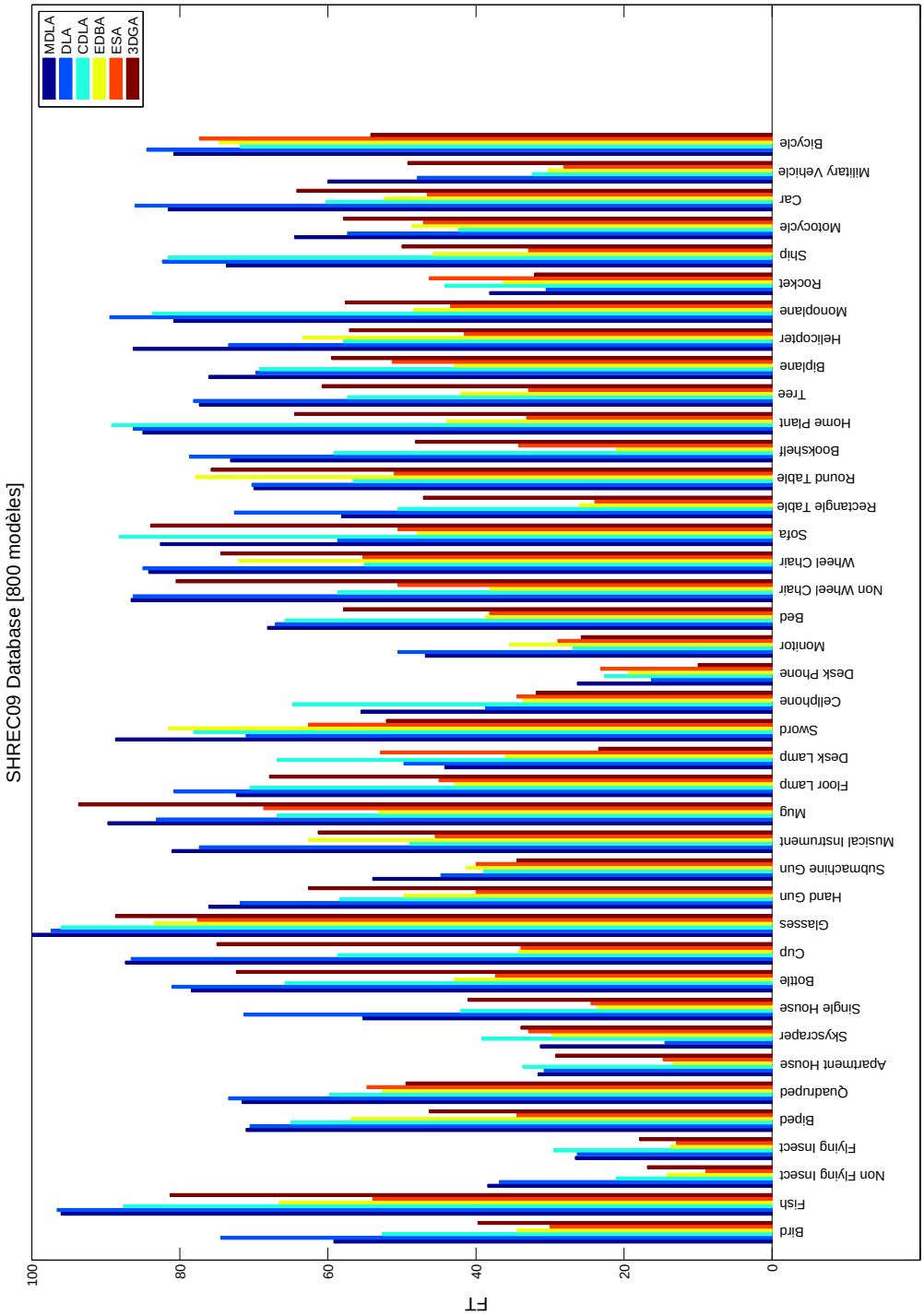



FIG. 6.23 – FT (%) des approches proposées calculés sur les 40 classes de la base de SHREC09.

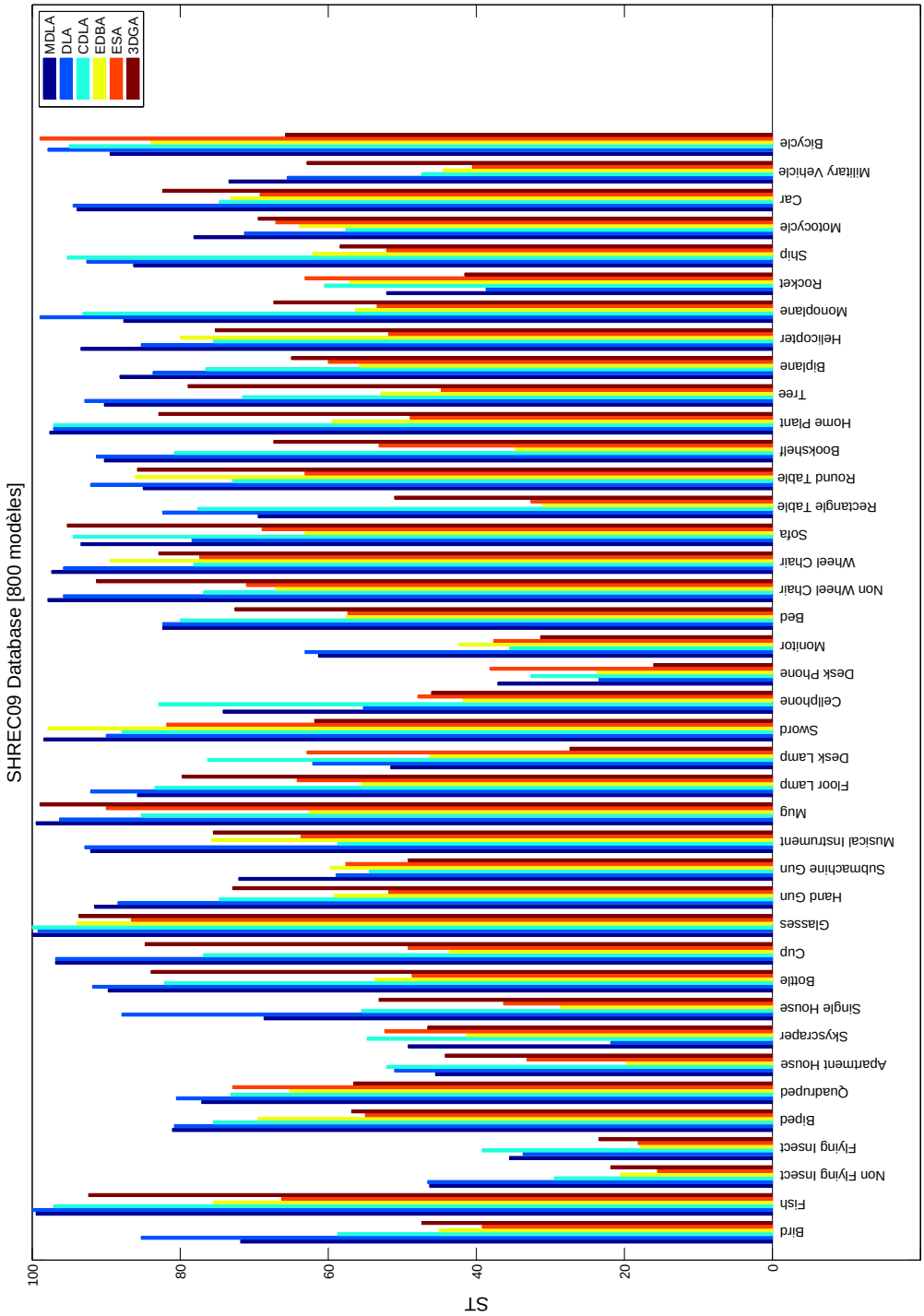



FIG. 6.24 – ST (%) des approches proposées calculés sur les 40 classes de la base de SHREC09.

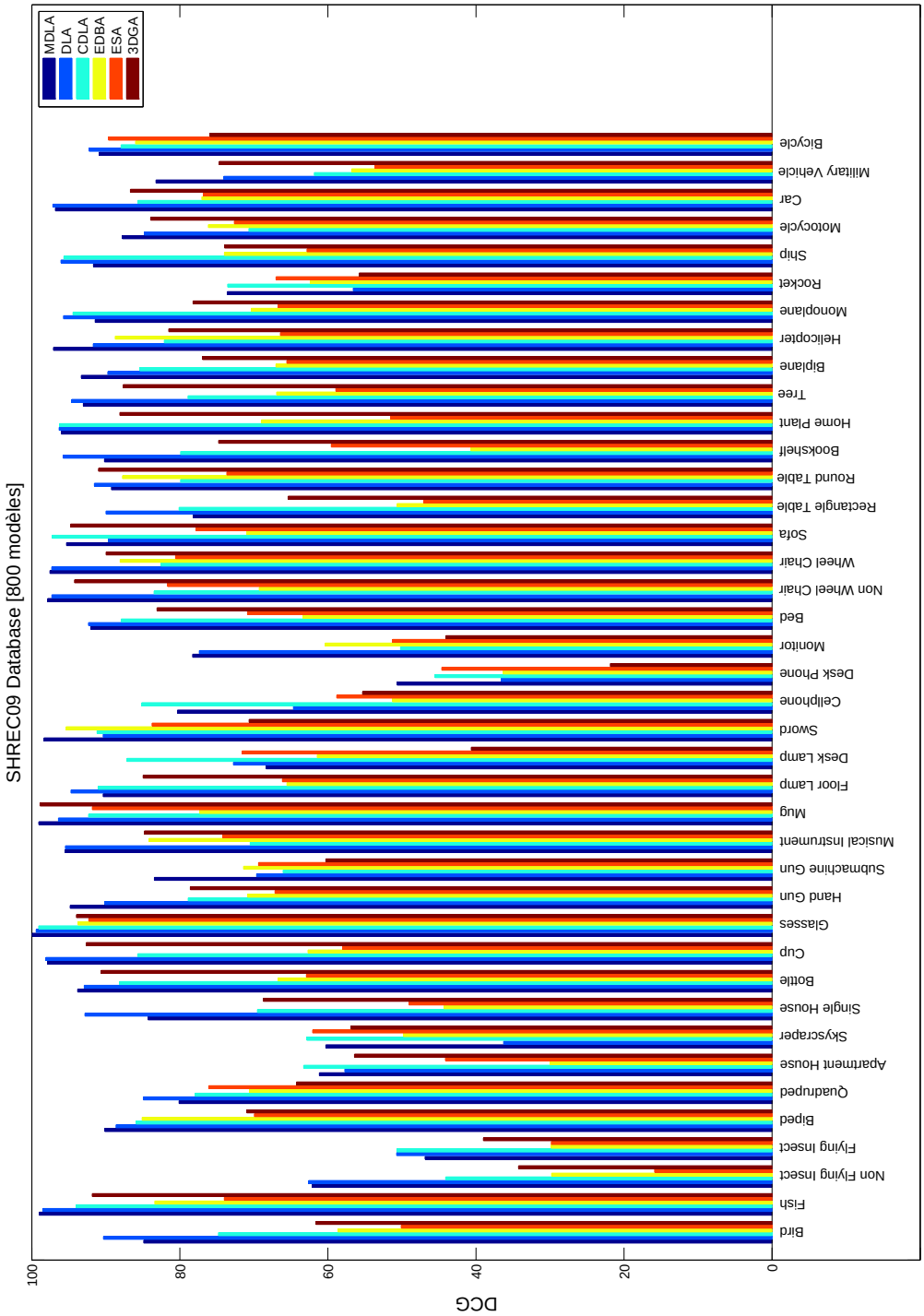



FIG. 6.25 – DCG (%) des approches proposées calculés sur les 40 classes de la base de SHREC09.

Les approches fondées sur les lignes de profondeur utilisant la programmation dynamique et notre méthode d'alignement ont les meilleures performances pour toutes les bases d'objets 3D. Comme on pouvait le prévoir, la *multi-view depth line approach* se classe première vu qu'elle possède un nombre de vues le plus élevé et représente l'information maximum portée par les vues.

Le Tableau comparatif 6.22 présente les mesures de performance (NN, FT, ST et DCG) des méthodes proposées dans cette thèse et des méthodes les plus représentatives de l'état de l'art, calculées sur la base généraliste PSB Test. Ce tableau englobe nos évaluations et celles de [7, 26, 112].

| Approche | cf. section | Famille | Type | NN | FT | ST | DCG |
|---|---|---|---|---|---|---|---|
| **MDLA** | 5.3.2.3 | **2D/3D** | **Pro** | **72.8** | **48.3** | **59.9** | **71.8** |
| **DLA** | 5.3.2.2 | **2D/3D** | **Pro** | **70.8** | **43.4** | **55.6** | **69.2** |
| **CDLA** | 5.3.2.4 | **2D/3D** | **Pro** | **65.5** | **37.6** | **49.7** | **64.5** |
| LFD | 4.2.1.1 [34] | 2D/3D | Sil | 65.7 | 38.0 | 48.7 | 64.3 |
| **EDBA** | 5.2.3.2 | **2D/3D** | **Pro** | **65.4** | **38.3** | **49.8** | **64.1** |
| AVC | 4.2.1.1 [8] | 2D/3D | Sil | 62.0 | 35.5 | 45.5 | 63.0 |
| **3DGA** | 5.1.1 | **3D** | **Par** | **60.8** | **34.6** | **43.7** | **60.2** |
| **ESA** | 5.2.3.1 | **2D/3D** | **Sil** | **57.8** | **32.6** | **44.4** | **60.2** |
| REXT | 4.1.3.6 [130] | 3D | Tra | 60.2 | 32.7 | 43.2 | 60.1 |
| DBD | 4.2.1.2 [56] | 2D/3D | Pro | 59.2 | 32.9 | 41.8 | 58.9 |
| SHD | 4.1.3.6 [67] | 3D | Tra | 55.6 | 30.9 | 41.1 | 58.4 |
| GEDT | 4.1.3.6 [67] | 3D | Tra | 60.3 | 31.3 | 40.7 | 58.4 |
| SIL | 4.2.1.1 [131] | 2D/3D | Sil | 52.8 | 28.5 | 38.8 | 56.3 |
| EXT | 4.1.3.6 [109] | 3D | Tra | 54.9 | 28.6 | 37.9 | 56.2 |
| SECSHEL | 4.1.2.1 [6] | 3D | Par | 54.6 | 26.7 | 35.0 | 54.5 |
| VOXEL | 4.1.2.5 [112] | 3D | Par | 54.0 | 26.7 | 35.3 | 54.3 |
| SECTORS | 4.1.2.1 [6] | 3D | Par | 50.4 | 24.9 | 33.4 | 52.9 |
| CEGI | 4.1.1.3 [63] | 3D | Sta | 42.0 | 21.1 | 28.7 | 47.9 |
| EGI | 4.1.1.3 [58] | 3D | Sta | 37.7 | 19.7 | 27.7 | 47.2 |
| D2 | 4.1.1.2 [95] | 3D | Sta | 31.1 | 15.8 | 23.5 | 43.4 |
| SHELLS | 4.1.2.1 [6] | 3D | Par | 22.7 | 11.1 | 17.3 | 38.6 |

TAB. 6.22 – Mesures de performance (%) calculées sur la base généralistes PSB Test. Les méthodes sont classées suivant les valeurs de DCG. Sta : Approche statistique, Par : Approche basée sur une partition de l'espace, Tran : Approches par transformées, Pro : Approches à base d'images de profondeur, Sil : Approches à base de silhouettes.

Ces résultats montrent la bonne performance des méthodes d'indexation d'objets 3D basées sur les vues, et en particulier, des méthodes fondées sur les lignes de profondeur. Ces méthodes 2D/3D, bien que relativement simples fournissent des résultats meilleurs que les méthodes 3D basées sur les transformées ou sur une partition de l'espace. Dans cette évaluation globale, la MDLA se place en première position. En la comparant à la meilleure méthode du *Princeton Shape Benchmark* [112] (LFD), nous enregistrons des gains importants de 7.1%, 10.3%, 11.2% et 7.5%, en termes de NN, FT, ST et DCG. Cette bonne performance est aussi confirmée dans l'évaluation de SHREC'09 [4] où notre méthode est la première en termes de mesures de performance et courbe précision-rappel (cf. Annexe B.3).

Remarquons que la DLA est meilleure que la EDBA qui, elle même, améliore l'approche DBD. Toutes ces approches utilisent les mêmes vues 2D. Ceci met en évidence l'efficacité de la description par transcription de lignes de profondeur par rapport à la description de Fourier 2D.

Les résultats montrent également que les approches multi-vues, EDBA et ESA, améliorent respectivement DBD et SIL dont elles se sont issues. Nous notons des gains de performance allant de 5.2% à 8% pour EDBA et de 3.9% à 5.6% pour ESA. Ceci est dû à la fois à la méthode d'alignement et à la nouvelle mesure de similarité qui permet de pondérer correctement les contributions de chaque vue par des indices de pertinence.

Contrairement à nos attentes, la comparaison conduite sur les approches 3D place la 3DGA devant les approches par transformées, ce qui est généralement pas le cas quand on considère l'ensemble des approches par transformées et des approches par partition de l'espace. Ceci peut s'expliquer par le fait que la 3DGA utilise notre alignement pour normaliser les modèles contrairement aux approches par transformées. En reprenant les mesures de performance de la 3DGA avec la méthode d'alignement basée sur une ACPC (NN = 56.8, FT = 30.5, ST = 39.6, DCG = 56.9, cf. Tableau 6.6), la 3DGA se placera derrière GEDT. Par contre, quelle que soit la méthode d'alignement utilisée (AL ou ACPC), la 3DGA est la meilleure des approches 3D basées sur une partition de l'espace.

## 6.4 Conclusion

Pour répondre aux objectifs de cette thèse, nous avons proposé un ensemble de méthodes de recherche par le contenu d'objets 3D. La précision d'indexation et de recherche a été le critère principal qui a guidé leur mise en place. Nous avons fixé, pour chaque processus de recherche, les paramètres de description, distances de similarités et méthodes d'alignement permettant d'atteindre les meilleures performances. Une étude comparative des méthodes d'indexation 3D a été détaillée par la suite.

Toutes les approches d'indexation 3D proposées dans cette thèse font appel à une étape de normalisation, à cause de leurs non-invariance aux transformations affines. Les résultats expérimentaux ont permis de mettre en évidence l'efficacité de la nouvelle méthode d'alignement présentée dans le chapitre 3 et sa supériorité par rapport à la méthode fondée sur l'analyse en composantes principales continue.

Les nouveaux descripteurs à base de lignes de profondeur ont montré au cours des différents tests réalisés leur grande précision sur les quatre bases d'objets par rapport aux autres descripteurs de forme proposés. En effet, l'originalité de ces descripteurs réside dans la méthode de transcription qui permet une caractérisation compacte et pertinente. Combinée à une mesure de similarité adéquate et performante (fondée sur la programmation dynamique), nous avons montré que la fidélité de description améliore la précision des recherches d'objets 3D et dépasse les approches qui utilisent une description de Fourier sur les mêmes images de profondeur. De plus, une seconde amélioration significative des résultats a été observée en augmentant le nombre de vues de six à vingt. L'approche MDLA permet une description plus globale de l'objet 3D donnant ainsi des résultats fortement satisfaisants. Enfin, la comparaison conduite sur la base de Test de Princeton établit objectivement la supériorité de la méthode MDLA par rapport aux principaux méthodes existantes.

# Conclusion générale

Dans cette thèse, nous nous sommes intéressés à la recherche par le contenu d'objets 3D, et plus particulièrement aux deux principaux aspects de l'amélioration des performances de la recherche. Ces aspects sont liés au temps de recherche et à la qualité des résultats. L'objectif de notre travail a été de proposer des méthodes rapides et efficaces permettant de retrouver des objets similaires à un objet requête dans une base généraliste. Deux problématiques ont été étudiées : l'alignement et l'indexation d'objets 3D. Nous allons détailler plus précisément nos contributions dans la suite.

## Alignement d'objets 3D

Nous avons proposé une nouvelle méthode calculant les axes d'alignement d'un modèle 3D. Pour ce faire, nous avons montré comment les symétries peuvent être utiles pour aligner un objet 3D, et en particulier pour retrouver ses directions naturelles. Deux types de symétries ont été considérés : la symétrie de réflexion et la symétrie locale de translation. Pour atteindre une faible complexité, notre méthode retourne les directions normales aux éventuelles réflexions caractérisées par les plans principaux de l'objet 3D. L'introduction d'une nouvelle mesure de symétrie locale de translation nous a permis de retrouver les axes d'alignement pour les objets dont l'alignement ne pourrait être obtenu uniquement avec des plans de symétrie miroir.
En conclusion, notre méthode d'alignement a les propriétés suivantes :

**- Intuitive pour la perception humaine**

Nous avons montré que notre méthode calcule, dans la plupart des cas, des alignements correspondant à ceux donnés par un être humain. Les résultats obtenus sont très satisfaisants sur un grand nombre d'objets 3D en termes de précision et de rapidité.

**- Cohérente à l'intérieur d'une classe d'objets**

Notre méthode fournit des alignements cohérents, plus robustes et plus

précis que la méthode à base d'une analyse en composantes principales à l'intérieur d'une même classe d'objets.

**- Étape efficace de normalisation pour l'indexation 3D**

Notre méthode d'alignement a été intégrée dans la phase de normalisation, puis a été évaluée et comparée à la méthode fondée sur l'analyse en composantes principales. Dans tous les processus d'indexation proposés dans cette thèse, notre méthode d'alignement se révèle plus performante en termes de courbes précision-rappel et mesures de performance. Ceci confirme que notre méthode parvient à mieux aligner des objets à l'intérieur de chaque classe, produisant ainsi une description plus pertinente et une similarité plus discriminante.

## Indexation d'objets 3D

L'état de l'art nous a permis d'étudier les principes, avantages et limitations des approches les plus représentatives d'indexation d'objets 3D et nous a conduit vers de nouvelles solutions :

**- Indexation 3D par descripteur gaussien**

Nous avons proposé une nouvelle méthode d'indexation fondée sur une partition de l'espace englobant l'objet 3D. En utilisant des fonctions gaussiennes, notre méthode extrait de façon précise des informations locales caractérisant la surface de l'objet. La signature est compacte, rapide à calculer et à comparer, ce qui permet d'utiliser cette méthode sur des grandes bases de données et de garantir des temps de recherche courts.

**- Pertinence des vues 2D dans l'indexation 2D/3D**

Constatant que l'information est inégalement répartie sur les vues 2D correspondants aux projections d'un objet 3D, nous avons introduit une technique d'amélioration pouvant s'appliquer aux approches multi-vues. Elle part du principe que les vues doivent contribuer à la description de la forme 3D proportionnellement aux informations contenues dans celles-ci. Pour ce faire, des indices de pertinence ont été définis, associés aux images de projection et intégrés dans la mesure de similarité. Cette technique permet une description non biaisée de la forme 3D et un calcul de similarité plus fin.

**- Indexation 2D/3D par lignes de profondeur**

Nous avons proposé une famille d'approches multi-vues à base d'images

de profondeur. Nous avons introduit un nouveau descripteur adapté au contenu spécifique des images de profondeur. Il est fondé sur une extraction des lignes de profondeur et sur une technique de transcription de ceux-ci. En raisonnant directement sur les pentes des lignes de profondeur, nous obtenons une caractérisation compacte et de plus haut niveau. Combiné à une mesure de similarité adéquate fondée sur la programmation dynamique, l'ensemble assure une invariance aux petites déformations locales.

### Recherche par le contenu d'objets 3D

Après avoir spécifié les mesures de similarité, les approches d'indexation proposées ont été évaluées sur quatre bases de données généralistes. Les résultats expérimentaux ont mis en évidence l'efficacité des approches multi-vues à base d'images de profondeur, et en particulier celles fondées sur les lignes de profondeur. Les réponses aux requêtes sont homogènes et cohérentes. Ces approches permettent de discerner les différentes classes d'objets et de regrouper les objets de formes similaires. Les tests comparatifs effectués nous ont permis de valider la robustesse de l'approche MDLA et de confirmer sa supériorité par rapport à toutes les autres méthodes sur des bases de données généralistes. Ces résultats ont été confortés par l'évaluation *SHREC 2009 - Generic Shape Retrieval Contest* où elle est classée première [4].
En conclusion, tous les algorithmes développés dans le cadre de cette thèse ont permis d'atteindre les objectifs d'efficacité et de rapidité. Nous préconisons l'utilisation de l'approche MDLA précédée de notre algorithme d'alignement pour toute recherche par le contenu dans une base d'objets généraliste.

## Perspectives

### Pour l'alignement

**- Amélioration de la méthode d'alignement**

La nouvelle méthode d'alignement a montré son fort potentiel à calculer les directions naturelles, mais peut être améliorée de différentes façons. Un premier travail consisterait à proposer une solution plus rapide pour les objets dont l'alignement ne comprend aucune direction principale. Un deuxième reposerait sur l'introduction de nouveaux types de symétries qu'elles soient globales ou locales pouvant caractériser des directions par-

ticulières ou conforter les solutions de notre méthode d'alignement. Il serait aussi intéressant d'étudier d'autres attributs ou caractéristiques géométriques de la forme qui ont un lien avec les directions naturelles de l'objet 3D.

**- Pose naturelle et position au sol**

Une extension de notre travail sur l'alignement est en cours de réalisation. C'est un outil interactif qui permet à un utilisateur de choisir parmi une liste optimale de repères candidats le repère correspondant à la pose naturelle. Il serait intéressant de valider par les résultats de cet outil les hypothèses de la perception humaine sur lesquelles s'est basée notre discussion. Dans le prolongement de cette étude, une perspective intéressante serait d'étudier la problématique du choix de la position au sol (*upright orientation*) abordée par Fu et al. [47]. En effet, on pourrait résoudre ce problème en procédant comme dans la recherche de la pose naturelle. Il s'agirait de considérer les directions de notre alignement et d'établir une discussion similaire à celle de la section 3.4.2.

**- Alignement à l'intérieur d'une scène 3D**

Le domaine de la modélisation et de la création des scènes 3D pourrait bénéficier directement des contributions apportées dans cette thèse pour l'alignement d'objets 3D. Les solutions que nous proposons peuvent être appliquées pour aligner des objets dans une scène 3D ou aligner les objets entre eux. Notre méthode d'alignement peut être efficacement exploitée dans d'autres cadres, telles que la visualisation 3D, la classification de modèles ou la reconnaissance d'objets 3D.

## Pour l'indexation 3D

**- Nouveau descripteur basé sur l'image gradient**

La méthodologie proposée pour la caractérisation de l'image de profondeur a été adaptée aux lignes de profondeur. Il serait intéressant de pouvoir extraire à partir de l'image de profondeur une image gradient 2D qui contienne plus d'informations sur le voisinage. En s'inspirant de notre technique de transcription, une perspective pourrait être de proposer une méthode similaire appropriée aux régions présentes dans cette nouvelle image. Nous avons souligné l'intérêt de la programmation dynamique dans la comparaison des séquences d'états. Celle-ci ne pourrait pas être appliquée dans le cas

d'une matrice d'états 2D. Mais, il serait toutefois intéressant d'introduire une distance appropriée qui permet une invariance aux petites déformations locales, comme par exemple la "*Earth Mover's distance*" [107]. Ainsi, une nouvelle méthode de recherche par le contenu d'objets 3D pourrait être proposée.

**- Approcher au mieux la reconnaissance humaine**

Nous avons montré que l'approche MDLA est actuellement la plus appropriée pour effectuer une recherche par le contenu dans une base de données généraliste. Cependant, elle n'atteint pas les performances de la reconnaissance d'un être humain. Pour améliorer cette approche, on pourrait l'intégrer dans des méthodes de recherche hybrides ou dans des méthodes d'apprentissage semi-supervisé, ou encore élaborer des stratégies de recherche plus sophistiquées (recherche en cascade, recherche globales et locales, etc.).

**- Évoluer vers des descriptions plus spécifiques**

Il serait fructueux d'étudier, avec une méthodologie similaire et une même exigence de précision et rapidité, le cas de la recherche d'objets 3D dans des bases spécialistes et pour d'autres secteurs applicatifs. L'idéal serait de pouvoir disposer, dans chaque cas, de listes de descripteurs et de mesures de similarité les plus efficaces en termes de précision et de rapidité. Il est à remarquer que la notion de similarité entre objets 3D peut varier suivant les applications et les types de données.

# Quatrième partie

# Annexes

# Annexe A

# Outils pour la description d'un objet 3D

## A.1 Volumes englobants

Un volume englobant d'un objet 3D est un volume fermé qui contient complètement celui-ci. Dans le cadre de recherche par le contenu, ces volumes sont utilisés dans l'extraction de la signature de l'objet 3D. Dans ce qui suit, nous définissons trois types élémentaires de volumes englobants.

### A.1.1 Boîtes englobantes

Une boîte englobante peut être parallélépipédique ou cubique. Dans cette section, quatre boîtes sont présentées :

**Définition A.1.** La boîte englobante (*bounding box - BB*) d'un objet 3D définie par l'ensemble des sommets :

$$\{(x,y,z)|x \in \{x_{min}, x_{max}\}, y \in \{y_{min}, y_{max}\}, z \in \{z_{min}, z_{max}\}\},$$

$$\begin{aligned} x_{min} &= \min_{1 \leq i \leq N_P} x_i, \quad y_{min} = \min_{1 \leq i \leq N_P} y_i, \quad z_{min} = \min_{1 \leq i \leq N_P} z_i, \\ x_{max} &= \max_{1 \leq i \leq N_P} x_i, \quad y_{max} = \max_{1 \leq i \leq N_P} y_i, \quad z_{max} = \max_{1 \leq i \leq N_P} z_i, \end{aligned} \tag{A.1}$$

$\mathbf{p}_i = (x_i, y_i, z_i) \in \mathfrak{P}$ (1.2), $1 \leq i \leq N_P$, étant les sommets de l'objet.

**Définition A.2.** La boîte englobante étendue (*Extended bounding box - EBB*) d'un objet 3D définie par l'ensemble des sommets :

$$\{(x,y,z)|x \in \{c_x - w, c_x + w\}, y \in \{c_y - w, c_y + w\}, z \in \{c_z - w, c_z + w\}\},$$

$$w = \max\left\{\frac{x_{max} - x_{min}}{2}, \frac{y_{max} - y_{min}}{2}, \frac{z_{max} - z_{min}}{2}\right\},$$
$$c_x = \frac{x_{max} + x_{min}}{2},\ c_y = \frac{y_{max} + y_{min}}{2},\ c_z = \frac{z_{max} + z_{min}}{2}, \tag{A.2}$$

où $x_{max}$, $x_{min}$, $y_{max}$, $y_{min}$, $z_{max}$ et $z_{min}$ sont donnés par la définition.

**Définition A.3.** Le cube canonique englobant (*Canonical bounding cube - CBC*) d'un objet 3D défini par l'ensemble des sommets :

$$\{(x, y, z)|x, y, z \in \{-a_{max}, a_{max}\}\},$$
$$a_{max} = \max\left\{\max_{1\leq i\leq N_P} |x_i|, \max_{1\leq i\leq N_P} |y_i|, \max_{1\leq i\leq N_P} |z_i|\right\}, \tag{A.3}$$

$\mathbf{p}_i = (x_i, y_i, z_i) \in \mathfrak{P}$ (1.2), $1 \leq i \leq N_P$, étant les sommets de l'objet.

**Définition A.4.** Le cube canonique (*Canonical cube - CC*) d'un objet 3D défini par l'ensemble des sommets :

$$\{(x, y, z)|x, y, z \in \{-w, w\}\}, \tag{A.4}$$

$w$ étant la longueur de l'arête à fixer.

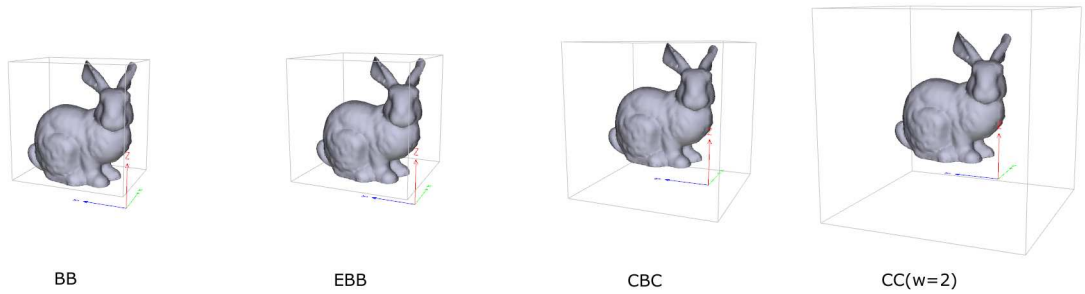



FIG. A.1 – Les quatre types de boîtes englobantes : BB, EBB, CBC et CC.

### A.1.2 Sphère englobante

**Définition A.5.** La sphère englobante (*bounding sphere*) est la sphère la plus petite qui contient l'objet 3D et de centre le centre de gravité $\mathbf{g}$ de celui-ci. Elle est définie par un rayon $r_{max}$ donnée par la distance maximale du centre de la sphère aux points de la surface de l'objet 3D.

$$r_{max} = \max_{1\leq i\leq N_P} \|\mathbf{g} - \mathbf{p}_i\|, \tag{A.5}$$

$\mathbf{p}_i \in \mathfrak{P}$ (1.2), $1 \leq i \leq N_P$, étant les sommets du maillage.

### A.1.3 Cylindre englobant

**Définition A.6.** Le cylindre englobant (*bounding cylinder*) est le cylindre contenant le modèle entier. L'axe du cylindre est aligné sur une des trois directions du repère intrinsèque à l'objet. Il est défini par une hauteur $h$, une coordonnée tridimensionnelle représentant le centre du cylindre $c$, et un rayon scalaire $r_{max}$ qui définit la distance maximale de l'axe du cylindre aux points de l'objet 3D.

$$\begin{array}{ccc}
\text{cylindre englobant} & \text{cylindre englobant} & \text{cylindre englobant} \\
\text{suivant l'axe des x} & \text{suivant l'axe des y} & \text{suivant l'axe des z} \\
r^x_{max} = \max\limits_{1\leq i\leq N_v} \sqrt{y_i^2 + z_i^2} & r^y_{max} = \max\limits_{1\leq i\leq N_v} \sqrt{z_i^2 + x_i^2} & r^z_{max} = \max\limits_{1\leq i\leq N_v} \sqrt{x_i^2 + y_i^2} \\
c^x = (\frac{x_{max}+x_{min}}{2}, 0, 0) & c^y = (0, \frac{y_{max}+y_{min}}{2}, 0) & c^z = (0, 0, \frac{z_{max}+z_{min}}{2}) \\
h^x = x_{max} - x_{min} & h^y = y_{max} - y_{min} & h^z = z_{max} - z_{min}
\end{array} \tag{A.6}$$

$\mathbf{p}_i = (x_i, y_i, z_i) \in \mathfrak{P}$ (1.2), $1 \leq i \leq N_P$, étant les sommets de l'objet.

## A.2 Projections d'objet 3D

Les deux principales projections utilisées dans les approches multi-vues sont les silhouettes et les images de profondeur.

### A.2.1 Image Silhouette

Une image silhouette est calculée de la manière suivante : Soient $\pi$ le plan de projection et $\nu \in \pi$ la surface sur laquelle est projeté l'objet. Tout d'abord, $\nu$ est subdivisée en $N \times N$ partitions de taille égale. Chaque partition correspond à un pixel noir ou blanc qui indique s'il existe ou non des points de l'objet 3D dans la projection orthogonale sur $\nu$.

Soit $\mathbf{s}$ une image silhouette associée au plan $\pi = \{\mathbf{p} \in \mathbb{R}^3 | \mathbf{p}^T \cdot \mathbf{x}' = \delta\}$ qui n'intersecte pas l'objet 3D. L'image est représentée en 2D selon deux directions données par $\mathbf{y}'$ et $\mathbf{z}'$, avec $\mathcal{R}'(\mathbf{x}', \mathbf{y}', \mathbf{z}')$ un repère orthonormé de $\mathbb{R}^3$. $s_{ij}$ est la valeur du pixel à la position $(i, j)$, $0 \leq i, j \leq N-1$. Elle représente la partition $\nu_{ij}$ définie comme suit :

$$\nu_{ij} = \left\{ (x, y, z)_{\mathcal{R}'} | \begin{array}{l} x = \delta, \\ y'_{min} + \frac{y'_{max}-y'_{min}}{N} i \leq y < y'_{min} + \frac{y'_{max}-y'_{min}}{N}(i+1), \\ z'_{min} + \frac{z'_{max}-z'_{min}}{N} j \leq z < z'_{min} + \frac{z'_{max}-z'_{min}}{N}(j+1). \end{array} \right\}$$

Initialement, tous les pixels de l'image silhouette sont mis à 1, représentant la valeur du fond (blanc). Chaque point $\mathbf{p} = (x, y, z)_{\mathcal{R}'}$ se trouvant sur la surface du modèle 3D est projeté orthogonalement sur le plan $\pi$. La projection orthogonale $\mathbf{p}_\pi$ est déterminée par les coordonnées $(\delta, y, z)_{\mathcal{R}'}$. Le pixel $s_{ij}$ est mis à 0 (noir) quand il existe un point $\mathbf{p}$ de l'objet 3D tel que sa projection orthogonale sur le plan $\pi$ appartient à la partition $\nu_{ij}$.

### A.2.2 Image de profondeur

Une image de profondeur, appelée aussi image 2.5 D, est une image à niveaux de gris où les valeurs indiquent les profondeurs des premiers points visibles de l'objet 3D dans la projection orthogonale au plan de l'image. L'image est subdivisée en $N \times N$ partitions de taille égale. Chaque partition correspond à un pixel de niveau de gris où est stockée sa distance avec le premier point de l'objet 3D dans la direction perpendiculaire au plan de l'image.

Dans la plupart des approches fondées sur les images de profondeur, les caméras sont placées sur les directions orthogonales aux faces de la boîte englobante ou sur les directions des sommets d'un dodécaèdre. Ces choix permettent une répartition uniforme sur la sphère de vues. Dans les deux cas, l'objet 3D est caractérisé par un ensemble de paires de vues parallèles. Pour accélérer l'étape d'extraction, il serait intéressant de calculer simultanément deux images de profondeur correspondant à deux plans parallèles, limitant l'objet 3D de part et d'autre.

Soient $\mathbf{v}$ et $\mathbf{v}'$ deux images de profondeur associées à deux plans parallèles $\pi = \{\mathbf{p} \in \mathbb{R}^3 | \mathbf{p}^T \cdot \mathbf{x}' = \delta\}$ et $\pi' = \{\mathbf{p} \in \mathbb{R}^3 | \mathbf{p}^T \cdot \mathbf{x}' = \delta'\}$ limitant l'objet 3D. Les deux images sont représentées en 2D selon deux directions données par $\mathbf{y}'$ et $\mathbf{z}'$, avec $\mathcal{R}'(\mathbf{x}', \mathbf{y}', \mathbf{z}')$ un repère orthonormé de $\mathbb{R}^3$. $\mathbf{v}_{ij}$ et $\mathbf{v}'_{ij}$ sont les valeurs des pixels à la position $(i, j)$, $0 \leq i, j \leq N - 1$. Ils déterminent respectivement les ensembles $\nu_{ij}$ et $\nu'_{ij}$ :

$$\nu_{ij} = \left\{ (x, y, z)_{\mathcal{R}'} | \begin{array}{l} x = \delta, \\ y'_{min} + \frac{y'_{max} - y'_{min}}{N} i \leq y < y'_{min} + \frac{y'_{max} - y'_{min}}{N}(i+1), \\ z'_{min} + \frac{z'_{max} - z'_{min}}{N} j \leq z < z'_{min} + \frac{z'_{max} - z'_{min}}{N}(j+1). \end{array} \right\}$$

$$\nu'_{ij} = \left\{ (x', y', z')_{\mathcal{R}'} | \begin{array}{l} x = \delta', \\ y'_{min} + \frac{y'_{max} - y'_{min}}{N} i \leq y < y'_{min} + \frac{y'_{max} - y'_{min}}{N}(i+1), \\ z'_{min} + \frac{z'_{max} - z'_{min}}{N} j \leq z < z'_{min} + \frac{z'_{max} - z'_{min}}{N}(j+1). \end{array} \right\}$$

Initialement, tous les pixels de l'image de profondeur sont mis à 0, représentant la valeur du fond. La valeur 1 est mise quand l'objet 3D intersecte le plan de vue associé à l'image de profondeur. Chaque point $\mathbf{p} = (x, y, z)_{\mathcal{R}'}$ se trouvant sur la surface du modèle 3D est projeté orthogonalement sur les deux plans parallèles. $\pi$ et $\pi'$, les projections correspondant respectivement à la face avant et à la face arrière, sont déterminées par les coordonnées $(x_{min}, y_P, z_P)$ et $(x_{max}, y_P, z_P)$.

Si $P_v \in \nu_{ab}$, alors $v_{ab}$ est mis à jour : $v_{ab} \leftarrow \max\left(v_{ab}, \frac{x_{max}-x_p}{2w}\right)$,
Si $P_u \in \upsilon_{ab}$, alors $u_{ab}$ est mis à jour : $u_{ab} \leftarrow \max\left(u_{ab}, \frac{x_p-x_{min}}{2w}\right)$.

# Annexe B

# 3D SHape REtrieval Contests - SHREC

Le *3D SHape REtrieval Contests* - SHREC est un concours annuel organisé par le réseau d'excellence AIM@SHAPE dans le cadre de la recherche par le contenu d'objets 3D. L'objectif principal consiste à comparer et évaluer les algorithmes de l'état de l'art participant à ce concours.

Nous présentons ici les résultats comparatifs dans les campagnes de SHREC auxquelles nous avons participé à savoir :
- SHREC'06 (base de données généraliste) en section B.1,
- SHREC'07 (base de données articulée) en section B.2,
- SHREC'09 (base de données généraliste) en section B.3.

Nous précisons, pour chaque évaluation, la méthode et les paramètres utilisés.

## B.1 Résultats de SHREC'06

Dans cette première campagne, la base de données contient 1814 modèles de Princeton Shape Benchmark [104], et l'ensemble de test comporte 30 objets requêtes extérieurs à la base, choisis parmi des objets proposés par les participants (cf. rapport technique [125] et évaluation de SHREC'06 [1]). Nous avons participé [27] avec les deux méthodes multi-vues améliorées présentées dans la section 5.2 :

- **EDBA** (cf. section 5.2.3.2) avec l'indice de pertinence $R_g$ et la méthode d'alignement fondée sur l'ACPC. Cette approche est notée C1.

- **ESA** (cf. section 5.2.3.1) avec l'indice de pertinence $R_{sa}$ et la méthode d'alignement fondée sur l'ACPC. Cette approche est notée C2.

| Rang | Participant | Run code | MADR value |
|---|---|---|---|
| 1 | Makadia et al. | M2 | 0.54986260 |
| 2 | Makadia et al. | M1 | 0.54084843 |
| 3 | Daras et al. | D1 | 0.52424060 |
| 4 | **Chaouch et al.** | **C1** | **0.50018275** |
| 5 | Papadakis et al. | P1 | 0.49523294 |
| 6 | Shilane et al. | S3 | 0.49371490 |
| 7 | Zaharia et al. | Z1 | 0.49247277 |
| 8 | Shilane et al. | S2 | 0.48770607 |
| 9 | **Chaouch et al.** | **C2** | **0.42156762** |
| 10 | Shilane et al. | S1 | 0.39706558 |
| 11 | Makadia et al. | M3 | 0.39249521 |
| 12 | Makadia et al. | M4 | 0.37667266 |
| 13 | Laga et al. | L1 | 0.32631385 |
| 14 | Laga et al. | L2 | 0.30619973 |
| 15 | Jayanti et al. | J2 | 0.26785165 |
| 16 | Jayanti et al. | J3 | 0.23702210 |
| 17 | Jayanti et al. | J1 | 0.23020707 |

TAB. B.1 – *Mean average dynamic recall* des algorithmes participant à SHREC'06.

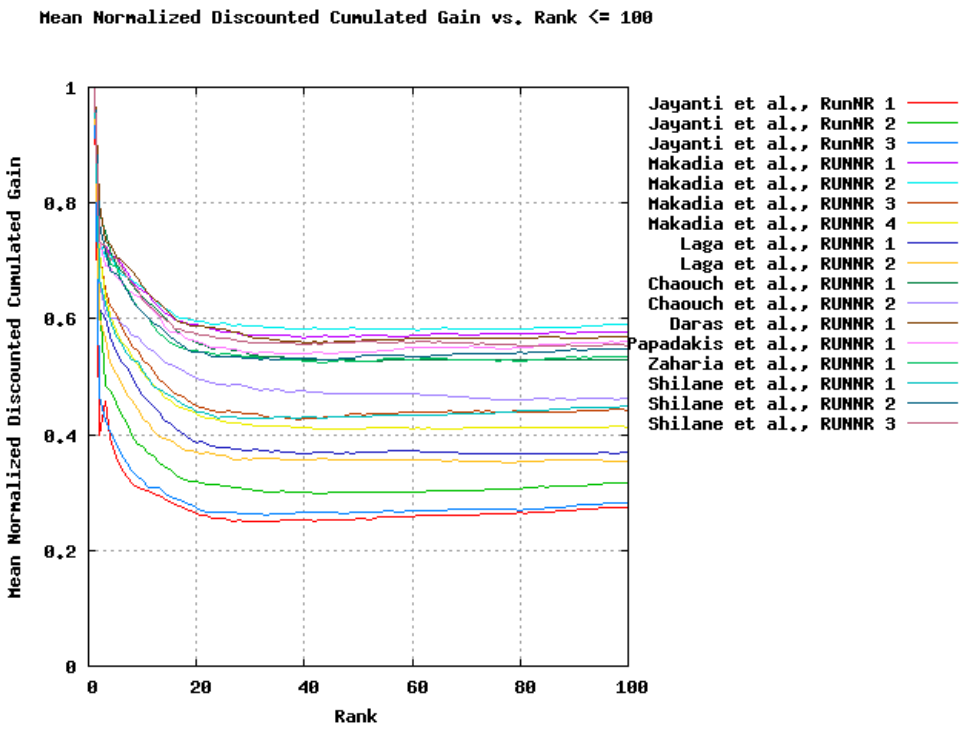



FIG. B.1 – Courbes de DCG normalisé des algorithmes participant à SHREC'06.

Les performances des approches testées regroupant les résultats de toutes les requêtes se répartissent grossièrement en trois groupes en considérant le *mean*

*average dynamic recall* (cf. Tableau B.1) :
- premier ensemble de huit méthodes, MADR de 0.549 à 0.487,
- deuxième ensemble de quatre méthodes, MADR de 0.421 à 0.376,
- dernier ensemble de cinq méthodes, MADR de 0.326 à 0.230.

Les classements des méthodes différaient en fonction des requêtes, mais globalement, les trois groupes restaient homogènes. L'approche EDBA se situe dans le premier groupe alors que l'approche ESA se place dans le deuxième groupe. Les résultats montrent un comportement correct pour les deux approches.

## B.2 Résultats de SHREC'07

Dans cette évaluation, la base de données [86] est fortement orientée "objets articulés". Elle comporte 400 modèles 3D répartis sur 20 classes. La même base constitue l'ensemble de test, c'est-à-dire, chaque objet de la base a fait l'objet d'une requête (cf. rapport technique [126]). Nous avons participé [28] avec la méthode multi-vues à base de lignes de profondeurs, présentée dans la section 5.3 :

- **DLA** (cf. section 5.3.2.2) avec $N = 32$, la boîte englobante BB et la méthode d'alignement fondée sur l'ACPC.

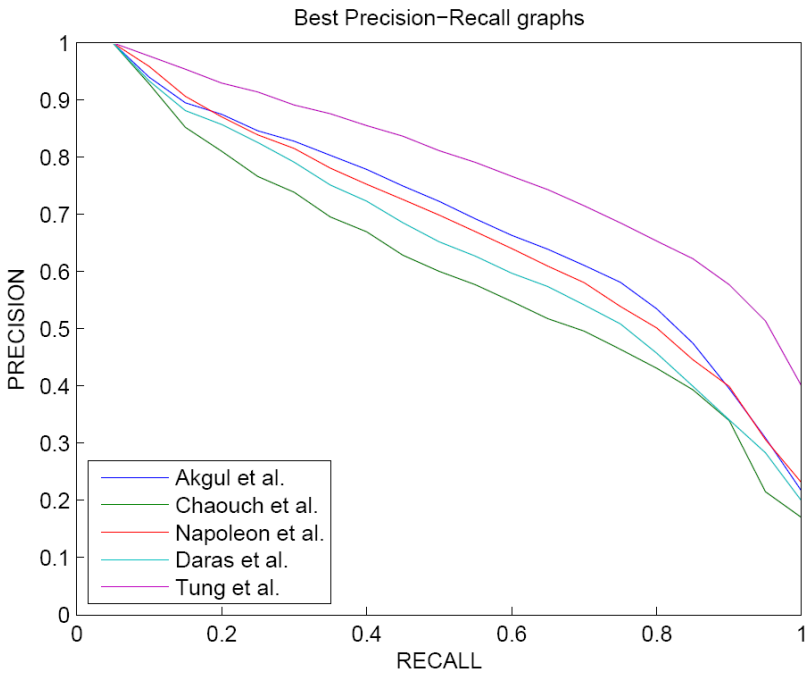



FIG. B.2 – Courbes de précision-rappel des algorithmes participant à SHREC'07.

Comme on pouvait le prévoir, notre algorithme s'est nettement mieux comporté dans les classes d'objets généralistes. Les courbes précision-rappel calculées sur les classes "table", "tenaille" et "vase" (cf. Figures B.3.b, B.3.c et B.3.d), montrent une supériorité de notre méthode par rapport aux méthodes participantes.
Par contre, ce type de méthode n'est pas adapté aux classes d'objets articulés. Comme nous pouvons le constater sur les courbes précision-rappel de la classe "armadillo" (cf. Figure B.3.a), notre méthode d'indexation a un comportement médiocre. Notons une forte variabilité dans cette classe due aux mouvements des membres du personnage. Les classes des objets articulés étant majoritaires, ceci s'est traduit par un mauvais classement en moyenne de notre méthode dans cette évaluation (cf. Figure B.2).

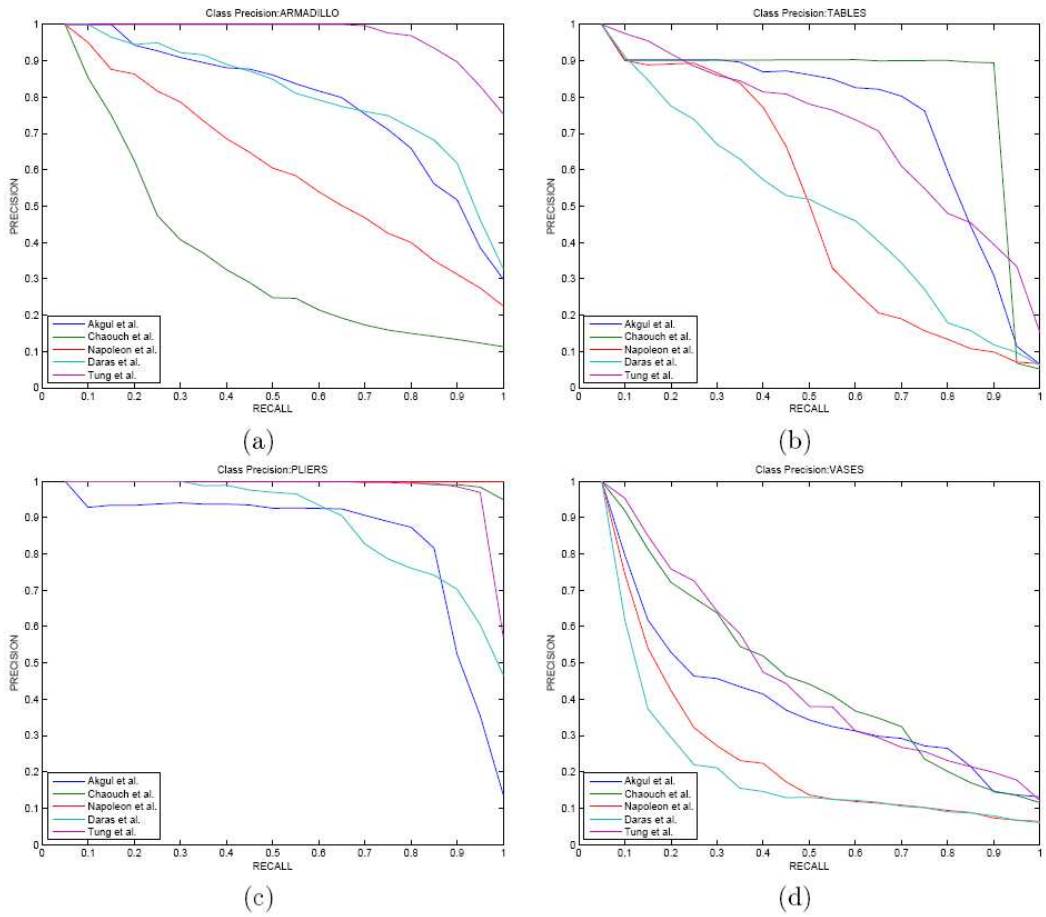



FIG. B.3 – Courbes de précision-rappel des algorithmes participant à SHREC'07, calculées sur quatre classes d'objets.

## B.3 Résultats de SHREC'09

Dans la récente campagne de SHREC, la base de données [87] contient 720 modèles 3D équi-répartis en 40 classes, et l'ensemble de test comporte 80 objets requêtes extérieurs de la base, deux objets pour chaque classe (cf. section 1.4.3 pour plus de détails sur la base). Cette base de données, est, nous semble-t-il, la plus "propre" des bases d'objets 3D généralistes (cf. évaluation de SHREC'09 [2, 4]). Nous avons participé avec la méthode multi-vues à base de lignes de profondeurs, présentée dans la section 5.3 :

- **MDLA** (cf. section 5.3.2.3) avec $N = 32$ et notre méthode d'alignement présentée dans le chapitre 3.

La plupart des méthodes testées sont récentes et semblent être adaptées au cas des objets généralistes. Notons aussi la participation de beaucoup de descripteurs hybrides dans cette évaluation. La Figure B.5 et le Tableau B.2 présentent respectivement les courbes précision-rappel et les mesures de performance, NN, FT, ST, E et DCG, obtenues avec les différentes méthodes.

| Participant | Méthode | NN | FT | ST | E | DCG |
|---|---|---|---|---|---|---|
| Akgül, Ceyhun Burak | DBFc8 | 82.5 | 43.3 | 55.0 | 38.3 | 74.8 |
| | DBFc10 | 82.5 | 44.3 | 57.4 | 39.8 | 75.7 |
| | DBFc12 | 81.3 | 44.9 | 57.8 | 40.6 | 75.9 |
| Bustos, Benjamin | DSR-segment | 86.3 | 56.1 | 69.6 | 49.0 | 82.5 |
| | DSR-nosegment | 85.0 | 54.6 | 69.1 | 47.9 | 81.9 |
| | Entropy-123-6-segment | 83.8 | 52.6 | 66.3 | 46.4 | 80.3 |
| | Entropy-6789-6-segment | 83.8 | 52.8 | 66.8 | 46.7 | 80.5 |
| | W1-segment | 83.8 | 52.8 | 66.6 | 46.6 | 80.6 |
| **Chaouch, Mohamed** | **MDLA** | **96.3** | **73.0** | **84.8** | **60.2** | **91.7** |
| Daras, Petros | 3D-shape-impact | 80.0 | 44.7 | 56.7 | 39.6 | 74.9 |
| | Compact-multiview | 80.0 | 49.0 | 62.6 | 43.7 | 77.1 |
| | Compound-SID-CMVD | 87.5 | 55.8 | 69.0 | 48.7 | 83.0 |
| Furuya, Takahiko | BF-SIFT | 85.0 | 48.3 | 62.4 | 43.3 | 77.7 |
| | MR-SPRH-UDR | 87.5 | 55.0 | 70.3 | 49.1 | 82.4 |
| Lian, Zhouhui | SHD+GSMD | 87.5 | 59.7 | 73.3 | 51.4 | 85.0 |
| | RECT+SHD+GSMD | 92.5 | 63.3 | 77.8 | 54.2 | 87.5 |
| | RECT+SHD+GSMD+MR | 92.5 | 72.4 | 84.4 | 59.5 | 90.4 |
| Napoléon, Thibault | Run1 | 90.0 | 52.2 | 66.5 | 45.9 | 81.4 |
| | Run2 | 95.0 | 61.5 | 70.1 | 50.2 | 86.4 |
| | Run3 | 95.0 | 63.9 | 77.1 | 54.0 | 88.2 |
| | Run4 | 90.0 | 55.0 | 66.2 | 46.5 | 82.6 |
| | Run5 | 88.7 | 57.0 | 70.9 | 49.7 | 83.8 |

TAB. B.2 – Mesures de performance (%) des algorithmes participant à SHREC'09.

La comparaison conduite montre une nette supériorité de notre méthode. En termes de toutes les mesures de performance calculées, elle obtient les meilleurs résultats. En particulier, avec NN égal à 96.3%, la MDLA n'a pas retrouvé d'objet pertinent en première place pour seulement trois requêtes (cf. Figure B.4).

Notons une forte similarité de forme dans les trois cas : insecte ailé et planeur, mitraillette et fusil, insectes avec et sans ailes. Ayant un FT égal à 73%, la MDLA permet en moyenne de retrouver entre 13 et 14 objets appartenant à la classe de la requête parmi les 18 premiers objets retournés. En termes de requêtes, notre méthode se montre la plus efficace pour retourner les résultats les plus intuitivement homogènes.

FIG. B.4 – Les plus proches voisins retournés par notre méthode (cas des trois réponses n'appartenant pas à la classe de l'objet requête), les objets requêtes étant Q16, Q43 et Q38.

Il ressort aussi de cette évaluation que les deux meilleures méthodes font appel à un alignement différent de l'ACP, ce qui met en avant l'importance d'un bon alignement dans la recherche par le contenu d'objets 3D.

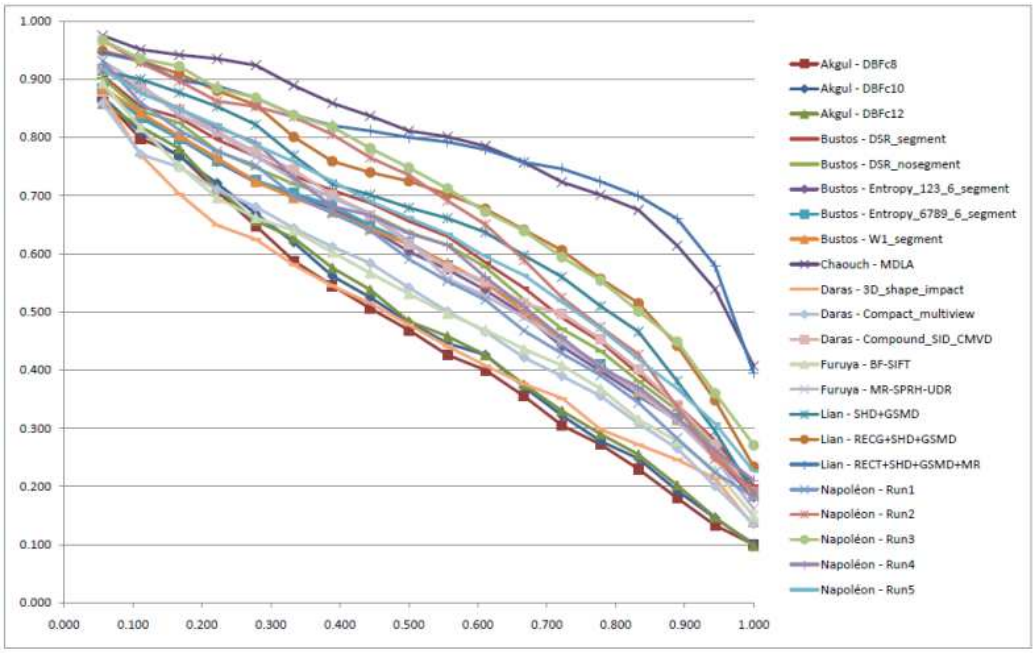



FIG. B.5 – Courbes de précision-rappel des algorithmes participant à SHREC'09.

# Annexe C

# Moteur de recherche

Pour pouvoir lancer des requêtes, nous avons intégré les résultats de notre système de recherche par le contenu d'objets 3D dans le moteur de recherche Ikona-Maestro [19] développé au sein de l'équipe IMEDIA-INRIA. Une requête est exprimée sous la forme d'un objet exemple issu de la base d'objets. Notre démonstrateur est disponible sur http://www-rocq.inria.fr/cgi-bin/imedia/circario.cgi/3d.

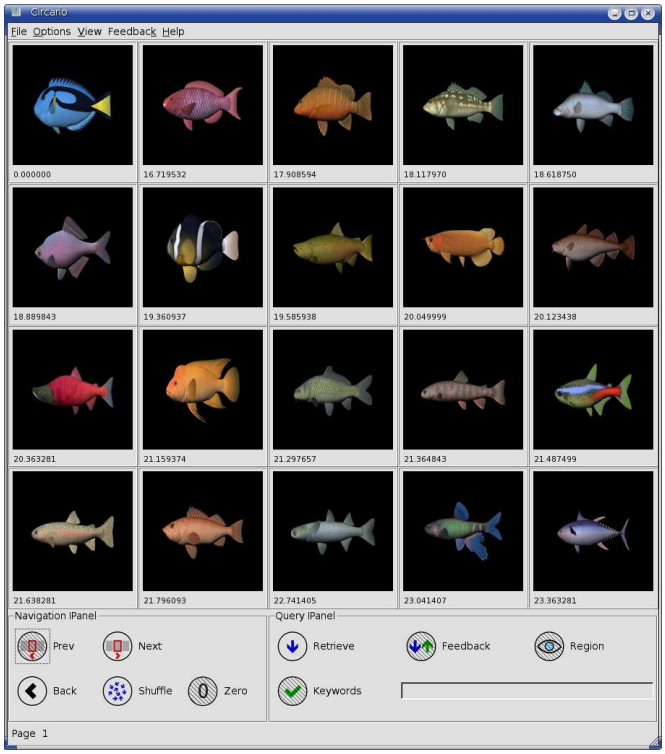



FIG. C.1 – Modèles retrouvés par notre système pour une requête de la classe "*Fish*" (20). La recherche est effectuée sur la base de SHREC'09.

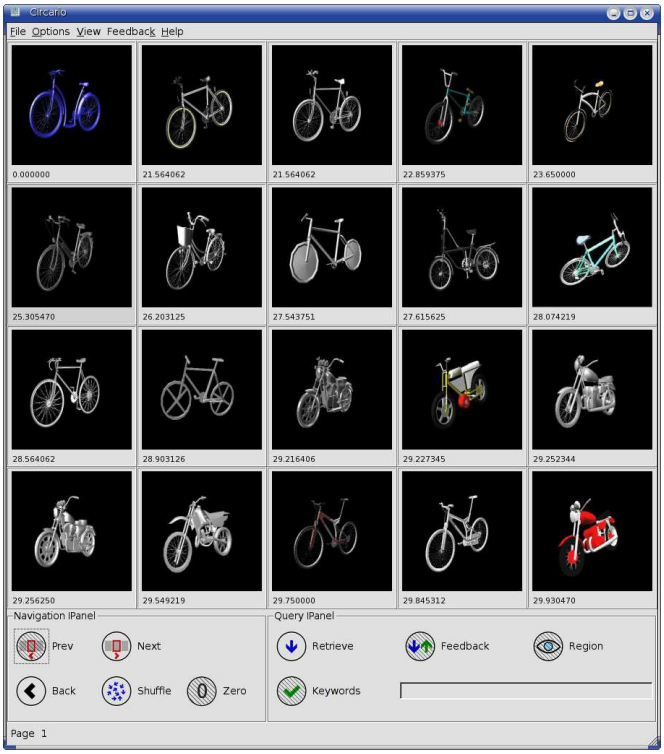



FIG. C.2 – Modèles retrouvés par notre système pour une requête de la classe "*Bicycle*" (20). La recherche est effectuée sur la base de SHREC'09.

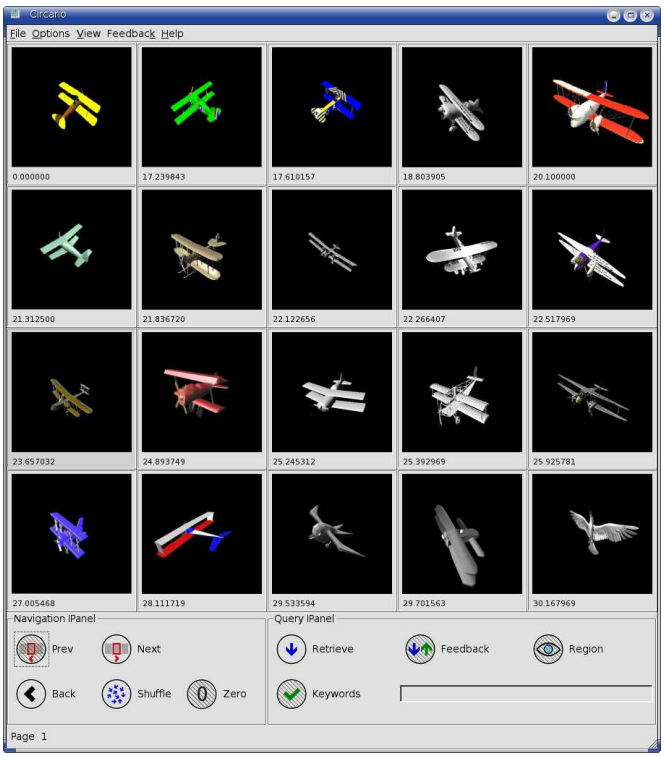



FIG. C.3 – Modèles retrouvés par notre système pour une requête de la classe "*Biplane*" (20). La recherche est effectuée sur la base de SHREC'09.

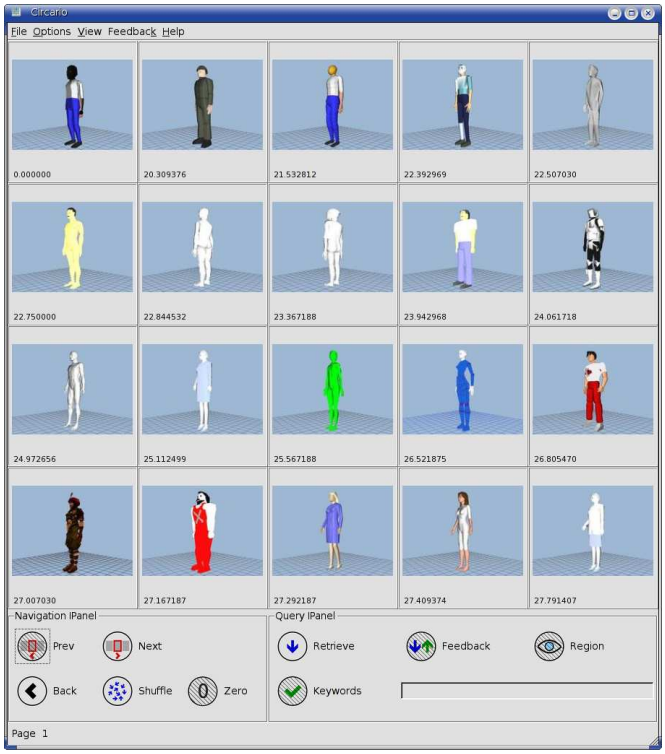



FIG. C.4 – Modèles retrouvés par notre système pour une requête de la classe "*Human biped*" (50). La recherche est effectuée sur la base de PSB Test.

FIG. C.5 – Modèles retrouvés par notre système pour une requête de la classe "*Potted plant*" (26). La recherche est effectuée sur la base de PSB Test.

FIG. C.6 – Modèles retrouvés par notre système pour une requête de la classe "*Fighter jet airplane*" (50). La recherche est effectuée sur la base de PSB Test.

# Table des figures


1.1 Processus comparant deux objets 3D. . . . . 15

1.2 Architecture générale d'un système de recherche par le contenu d'objets 3D. . . . . 17

1.3 Maillages de résolutions différentes (100, 250, 400 et 5800 facettes) représentant le même modèle de "vache". . . . . 19

1.4 Représentation d'un dodécaèdre par un maillage polygonal. . . . . 21

1.5 Représentation du dodécaèdre de la figure 1.4 par un maillage triangulaire. . . . . 22

1.6 Modèles 3D représentant les classes de la base de MPEG7. . . . . 27

1.7 Modèles 3D représentant les classes de la base de SHREC'09. . . . 29

2.1 Un modèle 3D représenté dans l'espace 3D, sous trois positions (au centre) , trois orientations (à droite) et trois échelles différentes (à gauche). . . . . 35

2.2 Alignements différents pour des objets appartenant à la même classe. 39

2.3 Exemples d'alignements calculés par l'ACP qui sont différents de la pose naturelle. . . . . 40

2.4 Exemples d'alignements erronés liés aux étiquetages des axes principaux selon l'ordre des valeurs propres. . . . . 40

2.5 Exemples d'alignements erronés liés aux changements de sens des axes principaux. . . . . 41

3.1 Alignement de modèles 3D avec deux méthodes différentes. Les flèches rouges, vertes et bleues représentent les axes principaux pour l'ACPC et les axes d'alignement optimaux pour notre méthode. 50

3.2 Exemples de modèles 3D appartenant aux différentes classes de symétrie miroir : Tétraèdre $\in G_T$, Octaèdre $\in G_O$, Icosaèdre $\in G_I$, Table $\in G_C^{even}$, Lampadaire $\in G_C^{odd}$, Sablier $\in G_D^{even}$, 5-prisme $\in G_D^{odd}$, Chaise $\in G_U$, Plante $\in G_Z$. . . . . 52

3.3 Exemples de modèles 3D regroupés selon les valeurs de $N_{GA}$. . . . 56

3.4 Les différentes étapes de notre algorithme d'alignement appliqué sur quatre modèles : montgolfière et sablier $N_{GA} = 2^+$, chaise $N_{GA} = 1$, arbre $N_{GA} = 0$. ***Ligne 1*** : Alignements obtenus via l'ACPC, ***Ligne 2*** : Test de la réflexion par rapport aux trois plans principaux de coordonnées $(xy)$, $(yz)$ et $(zx)$, le plan en bleu clair dans l'exemple de l'arbre étant le plan de réflexion maximale, ***Ligne 3*** : Sélection de la direction qui maximise le coût d'invariance locale de translation $CILT$,***Ligne 4*** : Résultats de notre algorithme d'alignement. . . . . 58

3.5 Calcul des axes d'alignement pertinents d'une chaise $N_{GA}(\mathcal{M}) = 1$. ***Ligne 1*** : Sélectionner $\mathbf{n}_1$ (en bleu) la normale du plan miroir détecté par l'ACPC comme premier axe d'alignement, ***Ligne 2*** : Tourner l'objet autour de $\mathbf{n}_1$ et calculer le CILT pour chaque rotation, ***Ligne 3*** : Trouver le deuxième axe d'alignement $\mathbf{n}_2$ (en vert) la direction qui maximise le CILT. La chaise est alors alignée selon les directions $\mathbf{n}_1$ et $\mathbf{n}_2$. . . . . 63

3.6 Trois descripteurs de formes $G$, $E$ et $F$ calculés sur une chaise le long de la direction verticale $\mathbf{d}_k$ et utilisés pour extraire l'information des invariances locales de translation. Dans ce cas, les trois descripteurs correspondent à la même direction qui maximise le CILT. Les régions de la surface associées aux intervalles ayant des invariances locales de translation sont colorées en bleu et les autres sont colorées en rouge. . . . . 66

3.7 Différents systèmes de coordonnées obtenus avec la méthode basée sur les moments après déformations d'une chaise. . . . . 69

3.8 Les repères candidats pour une représentation naturelle de l'objet 3D : (a) $N_{GA} = 3$, $\mathcal{M} \in \mathbf{G_D^{even}}$ avec $n = 4$, (b) $N_{GA} = 2$, $\mathcal{M} \in \mathbf{G_C^{odd}}$ avec $n = 3$, (c) $N_{GA} = 2$, $\mathcal{M} \in \mathbf{G_C^{even}}$ avec $n = 4$, (d) $N_{GA} = 2$, $\mathcal{M} \in \mathbf{G_C^{even}}$ avec $n = 2$. . . . . 74

3.9 Alignements de modèles 3D utilisant notre méthode. . . . . 75

3.10 Comparaison de la précison de l'alignement d'une boîte à lettre : l'ACPC (à gauche) et notre méthode (à droite). . . . . 76
3.11 Alignement des modèles de la classe "boîte à lettre" utilisant l'ACPC (a) et notre méthode (b). . . . . 76
3.12 Alignement des modèles de la classe "chaise de cuisine" utilisant l'ACPC (a) et notre méthode (b). . . . . 77
3.13 Robustesse de la méthode d'alignement face aux transformations anisotropiques globales et locales. . . . . 79

4.1 Distributions de forme de différents véhicules : cinq chars (courbes grises) et six voitures (courbes noires) [96] . . . . . 88
4.2 Exemples d'indices de forme (IF) calculés sur cinq formes élémentaires [139]. . . . . 90
4.3 Décomposition de l'espace englobant l'objet 3D : à gauche selon le paramètre $z$, à droite selon les paramètres cartésiens $(x, y, z)$. . . . 91
4.4 Trois décompositions classiques de l'espace 2D : coquilles, secteurs et combinés [6]. . . . . 91
4.5 Exemples de trois histogrammes calculés autour d'un axe principal $N = 15$ [94]. . . . . 92
4.6 La décomposition en harmoniques sphériques de la fonction sphérique voxelisée [49]. . . . . 99
4.7 Cinq modèles 3D et leurs descripteurs de symétrie miroir respectives [65]. . . . . 101
4.8 Exemples d'appariements entre deux graphes de squelettes : deux animaux (à gauche), un dinosaure et un humain (à droite). Les couleurs des nœuds montrent les appariements nœud-à-nœud des squelettes [117]. . . . . 102
4.9 Exemples de fonctions $\mu$ : $h$ hauteur (a), $g$ intégrale de distance géodésique (b, c et d) [57]. Dans le cas des tores, (a et b), la fonction basée sur $g$ caractérise de manière plus intrinsèque la topologie de la forme, comparée à la fonction basée sur $h$ fortement liée à l'information de la hauteur. Dans le cas des grenouilles, (c et d), les différentes parties conservent la même valeur de $\mu$ indépendamment de la position des bras et des pieds et aussi des dimensionnements des objets, ce qui rend la représentation avec la fonction $g$ plus intuitive. . . . . 103

4.10 Deux exemples d'appariements entre deux graphes de Reeb multi-résolution, sans et avec l'information géométrique [122]. Sans l'information géométrique (à gauche), les jambes peuvent être appariées aux bras car ils sont topologiquement équivalents. En ajoutant les attributs géométriques (à droite), tous les nœuds sont naturellement appariés. . . . . 104
4.11 Représentation d'une silhouette avec MCC et la distance associée : (a) image originale de la silhouette, (b) versions filtrées du contour sur différents niveaux d'échelle, (c) la représentation finale MCC pour 100 points de contours sur 14 niveaux d'échelle, (d) correspondance entre deux représentations MCC par une programmation dynamique [3]. . . . . 106
4.12 Un ensemble de *LightFields* pour un modèle 3D, N = 10 [34]. . . . 107
4.13 Comparaison de deux *LightFields* associés à deux modèles 3D (a), un cochon et une vache. Pour chaque objet, dix silhouettes sont créées sur les sommets d'un dodécaèdre. Les deux ensembles sont comparés en permutant l'ordre de mise en correspondance des vues (b, c, d, etc.) jusqu'à trouver l'appariement optimal [34]. . . . . 108
4.14 Trois modèles 3D et leurs coupes respectives le long d'une direction [103]. . . . . 112
4.15 Exemples d'images de spin construites sur quatre points de la surface de l'objet 3D. En bas, les trois décompositions utilisées pour calculer la signature d'une spin image [12]. . . . . 113

5.1 Exemple de calcul de descripteur 3DGA : (a) modèle "bunny" normalisé, (b) échantillonnage de la boîte englobante (N = 8), les points $\mathbf{q}_{ijk}$ sont colorés en bleu, (c) les rayons des boules rouges sont proportionnels aux coefficients $g_{ijk}$. . . . . 120
5.2 Contributions des points de la surface $\mathfrak{S}$, voisins à $\mathbf{q}_{ijk}$, dans la description locale $g_{ijk}$ . . . . . 120
5.3 Exemples de descripteurs 3DGA calculés sur quatre modèles différents. Les rayons des boules rouges sont proportionnels aux coefficients du descripteur. . . . . 124
5.4 Différence de la pertinence entre les vues dans la description de la forme 3D. Il est plus intuitif d'identifier l'homme à partir de la deuxième ou la troisième image que de la première. . . . . 127

5.5 Limitation de l'indice de pertinence fondé sur la surface. . . . . 129

5.6 Exemples d'indices de pertinence $R_{sa}$ (%) associés aux trois silhouettes obtenus après projection du modèle sur les faces de sa boîte englobante. . . . . 130

5.7 Limitation du contour de la silhouette dans la définition de la pertinence. Deux exemples de modèles de la même classe ayant des contours de longueurs différentes. . . . . 131

5.8 Exemples d'indices de pertinence $R_g$ (%) associés aux six images de profondeur obtenus après projection du modèle sur les six faces de sa boîte englobante. . . . . 132

5.9 Extraction des séquences $\mathbf{C}_K$. **Ligne 1** : Trois images silhouettes extraites en projetant le modèle 3D sur sa boîte englobante, **Ligne 2** : Les trois contours obtenus avec un filtre d'opérations logiques, **Ligne 3** : Quantification angulaire des trois contours. Chaque contour est représenté par les $K$ points noirs. . . . . 134

5.10 Six images de profondeur extraites en projetant le modèle 3D sur sa boîte englobante et leurs images de Fourier respectives. . . . . 137

5.11 Six images de profondeur obtenues en projetant le modèle 3D sur sa boîte engblobante. . . . . 140

5.12 Vingt images de profondeur obtenues en projetant le modèle 3D suivant les directions des sommets d'un dodécaèdre. . . . . 141

5.13 Trois images de profondeur obtenues en projetant le modèle 3D sur les faces latérales des cylindres englobants. . . . . 142

5.14 (a). Image de profondeur, (b). Séquences d'états associées aux lignes de profondeur, (c). Transcription d'une ligne de profondeur. 144

5.15 Exemples de descripteurs à base de lignes de profondeur. Chaque image de profondeur est caractérisée par un ensemble de 32 séquences d'états. . . . . 145

5.16 Comparaison de deux lignes de profondeurs utilisant la distance de Hamming entre leurs séquences respectives, $N = 12$ et $(C, S) = (0, 1)$. . . . . 146

5.17 Comparaison entre la distance de Hamming et la distance utilisant la programmation dynamique. Les deux lignes à comparer représentent deux formes similaires, mais ils ne sont pas parfaitement alignés le long de l'axe des abscisses. (a) 4 substitutions, la distance de Hamming calcule une mesure de similarité pessimiste. (b) 2 substitutions, la programmation dynamique permet de trouver le meilleur alignement optimisant la mesure de similarité. 147
5.18 Exemple d'alignement de séquences d'états utilisant l'algorithme de Needleman-Wunsch, $N = 12$ et $(C, S, I, D) = (2, -1, -1, -1)$. . 149
5.19 Exemple d'alignement de séquences d'états utilisant l'algorithme de Smith-Waterman, $N = 12$ et $(C, S, I_l, D_k) = (2, -1, -l, -k)$. . . 151

6.1 Courbes de précision-rappel de l'approche 3DGA pour différentes méthodes approximant discrètement la transformée de Gauss, avec $N = 8$, $\sigma = 0.23$, la boîte englobante BB et la distance $d_v$. . . . . 160
6.2 Courbes de précision-rappel de l'approche 3DGA pour cinq valeurs du paramètre $N$, avec $\sigma = 0.23$, la boîte englobante BB et la distance $d_v$. . . . . 162
6.3 Courbes de précision-rappel de l'approche 3DGA pour quatre types de boîte englobante, avec $N = 8$, $\sigma = 0.23$ et la distance $d_v$. . . . 163
6.4 Courbes de précision-rappel de l'approche 3DGA pour six valeurs du paramètre $\sigma$, avec $N = 8$, la boîte englobante BB et la distance $d_v$. . . . . 164
6.5 Courbes de précision-rappel de l'approche 3DGA pour les différentes distances, avec $N = 8$, $\sigma = 0.23$ et la boîte englobante BB. 165
6.6 Courbes de précision-rappel de l'approche 3DGA pour les méthodes d'alignement, avec $N = 8$, $\sigma = 0.23$, la boîte englobante BB et la distance $d_v$. . . . . 166
6.7 Courbes de précision-rappel de l'approche ESA pour les différents modèles d'indice de pertinence. . . . . 168
6.8 Courbes de précision-rappel de l'approche ESA pour les deux méthodes d'alignement, avec l'indice de pertinence $R_{sa}$. . . . . 169
6.9 Courbes de précision-rappel de l'approche EDBA pour les différents modèles d'indice de pertinence. . . . . 171
6.10 Courbes de précision-rappel de l'approche EDBA pour les deux méthodes d'alignement, avec l'indice de pertinence $R_g$. . . . . 172

6.11 Courbes de précision-rappel de l'approche DLA pour quatre valeurs du paramètre $N$, avec la boîte englobante BB et la distance de Hamming $d_H$. . . . . 173
6.12 Courbes de précision-rappel de l'approche DLA pour quatre types de boîte englobante, avec $N = 32$ et la distance de Hamming $d_H$. . 174
6.13 Courbes de précision-rappel de l'approche DLA pour les différentes distances, avec $N = 32$ et la boîte englobante BB. . . . . 175
6.14 Courbes de précision-rappel de l'approche DLA pour les deux méthodes d'alignement, avec $N = 32$, la boîte englobante BB et la distance de Smith-Waterman $d_{SW}$. . . . . 176
6.15 Courbes de précision-rappel de l'approche MDLA pour quatre valeurs du paramètre $N$, avec la distance de Hamming $d_H$. . . . . 178
6.16 Courbes de précision-rappel de l'approche MDLA pour les différentes distances, avec $N = 32$. . . . . 179
6.17 Courbes de précision-rappel de l'approche MDLA pour les deux méthodes d'alignement, avec $N = 32$ et la distance de Smith-Waterman $d_{SW}$. . . . . 180
6.18 Courbes de précision-rappel de l'approche CDLA pour trois valeurs du paramètre $N$, avec la distance de Hamming $d_H$. . . . . 182
6.19 Courbes de précision-rappel de l'approche CDLA pour les différentes distances, avec $N = 16$. . . . . 183
6.20 Courbes de précision-rappel de l'approche CDLA pour les deux méthodes d'alignement, avec $N = 16$ et la distance de Smith-Waterman $d_{SW}$. . . . . 184
6.21 Courbes de précision-rappel des approches proposées. . . . . 185
6.22 NN (%) des approches proposées calculés sur les 40 classes de la base de SHREC09. . . . . 186
6.23 FT (%) des approches proposées calculés sur les 40 classes de la base de SHREC09. . . . . 187
6.24 ST (%) des approches proposées calculés sur les 40 classes de la base de SHREC09. . . . . 188
6.25 DCG (%) des approches proposées calculés sur les 40 classes de la base de SHREC09. . . . . 189

A.1 Les quatre types de boîtes englobantes : BB, EBB, CBC et CC. . 202

B.1 Courbes de DCG normalisé des algorithmes participant à SHREC'06. 208
B.2 Courbes de précision-rappel des algorithmes participant à SHREC'07. 209
B.3 Courbes de précision-rappel des algorithmes participant à SHREC'07, calculées sur quatre classes d'objets. . . . . 210
B.4 Les plus proches voisins retournés par notre méthode (cas des trois réponses n'appartenant pas à la classe de l'objet requête), les objets requêtes étant Q16, Q43 et Q38. . . . . 212
B.5 Courbes de précision-rappel des algorithmes participant à SHREC'09. 212

C.1 Modèles retrouvés par notre système pour une requête de la classe "*Fish*" (20). La recherche est effectuée sur la base de SHREC'09. . 213
C.2 Modèles retrouvés par notre système pour une requête de la classe "*Bicycle*" (20). La recherche est effectuée sur la base de SHREC'09. 214
C.3 Modèles retrouvés par notre système pour une requête de la classe "*Biplane*" (20). La recherche est effectuée sur la base de SHREC'09. 214
C.4 Modèles retrouvés par notre système pour une requête de la classe "*Human biped*" (50). La recherche est effectuée sur la base de PSB Test. . . . . 215
C.5 Modèles retrouvés par notre système pour une requête de la classe "*Potted plant*" (26). La recherche est effectuée sur la base de PSB Test. . . . . 215
C.6 Modèles retrouvés par notre système pour une requête de la classe "*Fighter jet airplane*" (50). La recherche est effectuée sur la base de PSB Test. . . . . 216

# Liste des tableaux


1.1 Les principales bases d'objets 3D. . . . . 26

1.2 Les différentes classes de la base Test de Princeton Shape Benchmark 28

3.1 Les axes principaux retenus par l'ACPC pour chaque classe de symétrie miroir. . . . . 55

3.2 Les repères candidats pour une représentation naturelle de l'objet 3D quand $N_{GA} = 3$. . . . . 72

3.3 Les repères candidats pour une représentation naturelle de l'objet 3D quand $N_{GA} = 2$. . . . . 73

3.4 Les repères candidats pour une représentation naturelle de l'objet 3D quand $N_{GA} \leq 1$. . . . . 74

3.5 Comparaison des performances d'alignement de l'ACPC et de notre méthode avec les trois descripteurs $G$, $E$ et $F$, en termes de pourcentages d'objets parfaitement alignés pour 33 classes différentes (311 modèles). . . . . 78

3.6 Répartition des modèles 3D selon les valeurs de $N_{GA}$ pour les quatre bases d'objets 3D. . . . . 80

5.1 Paramètres de l'extraction des lignes de profondeur pour les trois approches. La DLA et la CDLA ont une même taille de descripteur, contrairement à la MDLA qui a une taille quatre fois plus grande. 142

5.2 Les cinq états proposés pour représenter une ligne de profondeur. 144

6.1 Mesures de performance (%) de l'approche 3DGA pour différentes méthodes approximant discrètement la transformée de Gauss, avec $N = 8$, $\sigma = 0.23$, la boîte englobante BB et la distance $d_v$. . . . . 160

6.2 Mesures de performance (%) de l'approche 3DGA pour cinq valeurs du paramètre $N$, avec $\sigma = 0.23$, la boîte englobante BB et la distance $d_v$. . . . . 161
6.3 Mesures de performance (%) de l'approche 3DGA pour quatre types de boîte englobante, avec $N = 8$, $\sigma = 0.23$ et la distance $d_v$. . . . . 162
6.4 Mesures de performance (%) de l'approche 3DGA pour six valeurs du paramètre $\sigma$, avec $N = 8$, la boîte englobante BB et la distance $d_v$. . . . . 164
6.5 Mesures de performance (%) de l'approche 3DGA pour les différentes distances, avec $N = 8$, $\sigma = 0.23$ et la boîte englobante BB. 165
6.6 Mesures de performance (%) de l'approche 3DGA pour les deux méthodes d'alignement, avec $N = 8$, $\sigma = 0.23$, la boîte englobante BB et la distance $d_v$. . . . . 167
6.7 Mesures de performance (%) de l'approche ESA pour les différents modèles d'indice de pertinence. . . . . 168
6.8 Mesures de performance (%) de l'approche ESA pour les deux méthodes d'alignement, avec l'indice de pertinence $R_{sa}$. . . . . 169
6.9 Mesures de performance (%) de l'approche EDBA pour les différents modèles d'indice de pertinence. . . . . 170
6.10 Mesures de performance (%) de l'approche EDBA pour les deux méthodes d'alignement, avec l'indice de pertinence $R_g$. . . . . 171
6.11 Mesures de performance (%) de l'approche DLA pour quatre valeurs du paramètre $N$, avec la boîte englobante BB et la distance de Hamming $d_H$. . . . . 173
6.12 Mesures de performance (%) de l'approche DLA pour quatre types de boîte englobante, avec $N = 32$ et la distance de Hamming $d_H$. . 174
6.13 Mesures de performance (%) de l'approche DLA pour les différentes distances, avec $N = 32$ et la boîte englobante BB. . . . . 176
6.14 Mesures de performance (%) de l'approche DLA pour les deux méthodes d'alignement, avec $N = 32$, la boîte englobante BB et la distance de Smith-Waterman $d_{SW}$. . . . . 177
6.15 Mesures de performance (%) de l'approche MDLA pour quatre valeurs du paramètre $N$, avec la distance de Hamming $d_H$. . . . . 178
6.16 Mesures de performance (%) de l'approche MDLA pour les différentes distances, avec $N = 32$. . . . . 179

6.17 Mesures de performance (%) de l'approche MDLA pour les deux méthodes d'alignement, avec $N = 32$ et la distance de Smith-Waterman $d_{SW}$. . . . . 180
6.18 Mesures de performance (%) de l'approche CDLA pour trois valeurs du paramètre $N$, avec la distance de Hamming $d_H$. . . . . 181
6.19 Mesures de performance (%) de l'approche CDLA pour les différentes distances, avec $N = 16$. . . . . 182
6.20 Mesures de performance (%) de l'approche CDLA pour les deux méthodes d'alignement, avec $N = 16$ et la distance de Smith-Waterman $d_{SW}$. . . . . 183
6.21 Mesures de performance (%) des approches proposées. . . . . 185
6.22 Mesures de performance (%) calculées sur la base généralistes PSB Test. Les méthodes sont classées suivant les valeurs de DCG. Sta : Approche statistique, Par : Approche basée sur une partition de l'espace, Tran : Approches par transformées, Pro : Approches à base d'images de profondeur, Sil : Approches à base de silhouettes. . . 190

B.1 *Mean average dynamic recall* des algorithmes participant à SHREC'06. 208
B.2 Mesures de performance (%) des algorithmes participant à SHREC'09. 211

# Liste des algorithmes


3.1 Alignement . . . . . 57
3.2 Calcul de la matrice rotation $R_{ga}$ maximisant le $CILT$ . . . . . 64
5.1 3D Gaussian Descriptor . . . . . 125
5.2 Calcul de la matrice de correspondance $\mathbf{M}$ . . . . . 148
5.3 Algorithme de Needleman-Wunch . . . . . 149
5.4 Algorithme de Smith-Waterman . . . . . 150
5.5 Algorithme de Smith-Waterman optimisé . . . . . 152

# Bibliographie


[1] Evaluation of SHREC'06 - Shape retrieval contest. http://give-lab.cs.uu.nl/shrec/shrec2006/index.html. 207

[2] Evaluation of SHREC'09 - Generic shape retrieval contest. http://www.itl.nist.gov/iad/vug/sharp/benchmark/shrecGeneric/results.html. 211

[3] T. Adamek and N. E. O'Connor. A multiscale representation method for nonrigid shapes with a single closed contour. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(5) :742–753, 2004. 106, 107, 147, 220

[4] C. Akgül, A. Axenopoulos, B. Bustos, M. Chaouch, P. Daras, H. Dutagaci, T. Furuya, A. Godil, S. Kreft, Z. Lian, T. Napoléon, A. Mademlis, R. Ohbuchi, P. L. Rosin, B. Sankur, T. Schreck, X. Sun, M. Tezuka, Y. Yemez, A. Verroust-Blondet, and M. Walter. SHREC 2009 - Generic Shape Retrieval Contest. In *EUROGRAPHICS Workshop on 3D Object Retrieval*, 2009. 191, 195, 211

[5] C. B. Akgül, B. Sankur, F. Schmitt, and Y. Yemez. Multivariate density-based 3D shape descriptors. In *IEEE International Conference on Shape Modeling and Applications (SMI'07)*, pages 3–12, Lyon, France, June 2007. 86

[6] M. Ankerst, G. Kastenmüller, H.-P. Kriegel, and T. Seidl. 3D shape histograms for similarity search and classification in spatial databases. In *6th International Symposium on Large Spatial Databases (SSD'99)*, volume 1651 of *Lecture Notes in Computer Science*, pages 207–226, Hong Kong, China, July 1999. Springer. 91, 100, 190, 219

[7] T. Filali Ansary. *Indexation de modèles 3D à partir de vues 2D*. PhD thesis, Institut National des Télécommunications, Université des Sciences et Technologies de Lille, October 2006. 109, 190

[8] T. Filali Ansary, M. Daoudi, and J.P. Vandeborre. A bayesian 3D search engine using adaptive views clustering. *IEEE Transactions on Multimedia*, 9(1) :78–88, January 2007. 105, 109, 115, 190

[9] T. Filali Ansary, J.P. Vandeborre, and M. Daoudi. A bayesian framework for 3d models retrieval based on characteristic views. In *2nd IEEE International Symposium on 3D Data Processing Visualization Transmission (3DPVT'04)*, Thessaloniki, Greece, September 2004. 105, 109

[10] N. Aspert, D. Santa-Cruz, and T. Ebrahimi. MESH : Measuring errors between surfaces using the Hausdorff distance. In *IEEE International Conference on Multimedia and Expo (ICME 02)*, pages 705–708, Lausanne, Switzerland, August 2002. 59

[11] J. Assfalg, A. Del Bimbo, and P. Pala. Curvature Maps for 3D CBR. In *International Conference on Multimedia and Expo (ICME'03)*, volume 10, Baltimore, July 2003. 90

[12] J. Assfalg, A. Del Bimbo, and P. Pala. Spin images for retrieval of 3D objects by local and global similarity. In *17th IEEE International Conference on Pattern Recognition (ICPR'04)*, volume 3, pages 906–909, Cambridge, UK, August 2004. 111, 112, 113, 220

[13] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 24 :509–522, 2002. 93

[14] J. C. Bezdek, M. R. Pal, J. Keller, and R. Krisnapuram. *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*. Kluwer Academic Publishers, Norwell, MA, USA, 1999. 113

[15] S. Biasotti and M. Attene. SHape REtrieval Contest 2008 : Stability of watertight models. In *IEEE International Conference on Shape Modeling and Applications (SMI'08)*, pages 217–218, Stony Brook, NY, USA, June 2008. 86

[16] S. Biasotti, D. Giorgi, M. Spagnuolo, and B. Falcidieno. Reeb graphs for shape analysis and applications. *Theoretical Computer Science*, 392(1-3) :5–22, 2008. 103

[17] S. Biasotti, S. Marini, M. Mortara, G. Patanè, M. Spagnuolo, and B. Falcidieno. 3D shape matching through topological structures. In *11th Interna-*

*tional Conference, Discrete Geometry for Computer Imagery (DGCI'03)*, pages 194–203, Naples, Italy, November 2003. 104

[18] I. Biederman. Human image understanding : recent research and a theory. In *Computer Vision, Graphics and Image Processing*, volume 32, pages 29–73, October 1985. 102

[19] N. Boujemaa, J. Fauqueur, M. Ferecatu, F. Fleuret, V. Gouet, B. Lesaux, and H. Sahbi. Ikona : Interactive specific and generic image retrieval. In *International workshop on Multimedia ContentBased Indexing and Retrieval (MMCBIR'2001)*, 2001. 213

[20] B. Bustos, D. Keim, D. Saupe, T. Schreck, and D. Vranic. Automatic selection and combination of descriptors for effective 3D similarity search. In *IEEE International Symposium on Multimedia Software Engineering*, pages 514–521, Miami, Florida, USA, December 2004. 86

[21] B. Bustos, D. A. Keim, D. Saupe, T. Schreck, and D. Vranic. An experimental comparison of feature-based 3D retrieval methods. In *Second International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'04)*, Thessaloniki, Greece, September 2004. 109

[22] B. Bustos, D. A. Keim, D. Saupe, T. Schreck, and D. Vranic. Using entropy impurity for improved 3D object similarity search. In *IEEE International Conference on Multimedia and Expo (ICME'04)*, pages 1303–1306, Taipei, Taiwan, June 2004. 86

[23] B. Bustos, D. A. Keim, D. Saupe, T. Schreck, and D. Vranic. Feature-based similarity search in 3D object databases. *ACM Computing Surveys*, 37(4) :345–387, December 2005. 86

[24] D. Caillière, F. Denis, D. Pelé, and A. Baskurt. 3D mirror symmetry detection using hough transform. In *IEEE International Conference on Image Processing (ICIP'08)*, pages 1772–1775, October 2008. 50

[25] N. Canterakis. 3D zernike moments and zernike affine invariants for 3D image analysis and recognition. In *11th Scandinavian Conference on Image Analysis*, Kangerlussuaq, Greenland, 1999. 96

[26] M. Chaouch and A. Verroust-Blondet. Enhanced 2D/3D approaches based on relevance index for 3D-shape retrieval. In *International Conference on Shape Modeling and Applications (SMI'06)*, page 36, Matsushima, Japan, June 2006. 127, 190

[27] M. Chaouch and A. Verroust-Blondet. Enhanced silhouette and depth-buffer based approaches for 3D-shape retrieval. Technical report, June 2006. 207

[28] M. Chaouch and A. Verroust-Blondet. 2D/3D descriptor based on depth line encoding. Technical report, June 2007. 209

[29] M. Chaouch and A. Verroust-Blondet. 3D model retrieval based on depth line descriptor. In *IEEE International Conference on Multimedia & Expo (ICME'07)*, pages 599–602, Beijing, China, July 2007. 139

[30] M. Chaouch and A. Verroust-Blondet. A new descriptor for 2D depth image indexing and 3D model retrieval. In *IEEE International Conference on Image Processing (ICIP'07)*, volume VI, pages 373–376, San Antonio, Texas, USA, September 2007. 139

[31] M. Chaouch and A. Verroust-Blondet. A novel method for alignment of 3D models. In *IEEE International Conference on Shape Modeling and Applications (SMI'08)*, pages 187–195, Stony Brook, NY, USA, June 2008. 47

[32] M. Chaouch and A. Verroust-Blondet. Alignment of 3D models. *Graphical Models*, 71 :63–76, 2009. 48

[33] D.Y. Chen. *Three-Dimensional Model Shape Description and Retrieval Based on LightField Descriptors*. PhD thesis, National Taiwan University, June 2003. 107

[34] D.Y. Chen, X.P. Tian, Y.T. Shen, and M. Ouhyoung. On visual similarity based 3D model retrieval. *Computer graphics forum*, 22(3) :223–232, September 2003. 105, 107, 108, 109, 111, 115, 190, 220

[35] P. Cignoni, C. Rocchini, and R. Scopigno. Metro : Measuring error on simplified surfaces. *Computer Graphics Forum*, 17(2) :167–174, June 1998. 59, 60

[36] P. R. Cromwell. *Polyhedra*. Cambridge University Press, 1999. 51

[37] C.M. Cyr and B.B. Kimia. 3D object recognition using shape similarity-based aspect graph. In *International Conference on Computer Vision (ICCV'01)*, pages 254–261, 2001. 105

[38] Escuela Superior de Ciencias Experimentales y Tecnología, Gavab Departament of Computing. Gavab Database. http://gavab.escet.urjc.es/recursos_en.html. 26

[39] Semantic Project (3D Service d'Echange et de MANipulation Tatouage, Indexation et Compression). Renault Semantic Database. http://liris.cnrs.fr/semantic-3d/SEMANTIC-3D.htm. 26

[40] J. L. Dugelay, A. Baskurt, and M. Daoudi. *3D Object Processing : Compression, Indexing and Watermarking*. April 2008. John Wiley & Sons, ISBN : 978-0-470-06542-6, April 2008. 86

[41] Carnegie Mellon University, Electrical and Computer Engineering. Carnegie Mellon Database. http://amp.ece.cmu.edu/projects/3DModelRetrieval/. 26

[42] R. W. Ferguson. Modeling orientation effects in symmetry detection : The role of visual structure. In *22nd Conference of the Cognitive Science Society*, pages 125–130, New Jersey, 2000. 50, 70

[43] Utrecht University, Center for Geometry, Imaging and Virtual Environments. Utrecht Database. http://www.cs.uu.nl/centers/give/multimedia/3Drecog/3Dmatching.html. 26

[44] Purdue University, Center for Information Sciences in Engineering. Purdue Engineering Shape Benchmark. http://shapelab.ecn.purdue.edu/Benchmark.aspx. 26

[45] McGill University, Shape Analysis Group, Centre for Intelligent Machines and School of Computer Science. McGill 3D Shape Benchmark. http://www.cim.mcgill.ca/~shape/benchMark/. 26

[46] Universität Leipzig, Intitut für Informatik. Konstanz Database. http://merkur01.inf.uni-konstanz.de/CCCC/. 26

[47] H. Fu, D. Cohen-Or, G. Dror, and A. Sheffer. Upright orientation of man-made objects. *ACM Trans. on Graphics*, 27(3), 2008. 196

[48] T. Funkhouser, M. Kazhdan, P. Shilane, P. Min, W. Kiefer, A. Tal, S. Rusinkiewicz, and D. Dobkin. Modeling by example. In *SIGGRAPH 2004 Conference Proceedings*, Los Angeles, August 2004. ACM SIGGRAPH, Addison Wesley. 86

[49] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman, D. Dobkin, and D. Jacobs. A search engine for 3D models. *ACM Transactions on Graphics*, January 2003. 99, 100, 105, 219

[50] T. Funkhouser and P. Shilane. Partial matching of 3D shapes with priority-driven search. In *Eurographics Symposium on Geometry Processing*, Sardinia, Italy, June 2006. 86

[51] N. Gagvani and D. Silver. Parameter-controlled volume thinning. *Graphical Models and Image Processing*, 61(3) :149–164, 1999. 102

[52] R. Gal and D. Cohen-Or. Salient geometric features for partial shape matching and similarity. *ACM Transactions on Graphics (TOG)*, 25(1) :130–150, 2006. 86

[53] D. Giorgi and S. Marini. SHape REtrieval Contest 2008 : Classification of watertight models. In *IEEE International Conference on Shape Modeling and Applications (SMI'08)*, pages 219–220, Stony Brook, NY, USA, June 2008. 86

[54] O. Gotoh. An improved algorithm for matching biological sequences. *Journal of Molecular Biology*, 162(3) :705–708, 1982. 151

[55] MPEG7 Video Group. MPEG7 Database. http://merkur01.inf.uni-konstanz.de/CCCC/. 26, 27

[56] M. Heczko, D. A. Keim, D. Saupe, and D. V. Vranic. Verfahren zur Ähnlichkeitssuche auf 3D-objekten (methods for similarity search on 3D databases. *Datenbank-Spektrum*, 2(2) :54–63, 2002. 105, 109, 115, 136, 139, 170, 190

[57] M. Hilaga, Y. Shinagawa, T. Kohmura, and T. L. Kunii. Topology matching for fully automatic similarity estimation of 3D shapes. In *SIGGRAPH 2001 Conference Proceedings*, pages 203–212, Los Angeles,USA, August 2001. ACM SIGGRAPH, Addison Wesley. 103, 219

[58] BKP. Horn. Extended Gaussian Images. In *Proc. of the IEEE*, volume 72, pages 1671–1686, December 1984. 89, 100, 190

[59] Informatics & Telematics Institute. ITI Database. http://3d-search.iti.gr/3DSearch/index.html. 26

[60] C. Yiu Ip, D. Lapadat, L. Sieger, and W. C. Regli. Using shape distributions to compare solid models. In *7th ACM Symposium on Solid Modeling and Applications*, pages 273–280, Saarbrücken, June 2002. 88, 114

[61] P. Irani and C. Ware. Diagramming information structures using 3D perceptual primitives. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 10(1) :1–19, 2003. 102

[62] A. E. Johnson and M. Hebert. Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5) :433–449, 1999. 112

[63] S. B. Kang and K. Ikeuchi. The complex EGI : a new representation for 3D pose determination. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 15(7) :707–721, July 1993. 89, 190

[64] M. Kazhdan. *Shape representation and Algorithm for 3D Model Retrieval.* PhD thesis, Princeton University, June 2004. 71, 100

[65] M. Kazhdan, B. Chazelle, D. Dobkin, A. Finkelstein, and T. Funkhouser. A reflective symmetry descriptor. In *European Conference on Computer Vision (ECCV)*, pages 642–656, May 2002. 100, 101, 219

[66] M. Kazhdan, B. Chazelle, D. Dobkin, T. Funkhouser, and S. Rusinkiewicz. A reflective symmetry descriptor for 3D models. *Algorithmica*, 38(2) :201–225, November 2003. 50, 100, 101

[67] M. Kazhdan, T. Funkhouser, and S. Rusinkiewicz. Rotation invariant spherical harmonic representation of 3D shape descriptors. In *Proceedings of the 2003 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing (SGP'03)*, pages 156–164, Aire-la-Ville, Switzerland, June 2003. Eurographics Association. 95, 100, 190

[68] M. Kazhdan, T. Funkhouser, and S. Rusinkiewicz. Symmetry descriptors and 3D shape matching. In *Eurographics Symposium on Geometry Processing*, Nice, France, July 2004. 60, 71, 100

[69] E. J. Keogh, L. Wei, X. Xi, S.-H. Lee, and M. Vlachos. Lb_keogh supports exact indexing of shapes under rotation invariance with arbitrary representations and distance measures. In *International Conference on Very large Data Bases (VLDB'06)*, pages 882–893. VLDB Endowment, 2006. 147

[70] J.J Koenderink and A.J. van Doorn. Surface shape and curvature scales. *Image and Vision Computing*, 10(8) :557–565, 1992. 89

[71] M. Körtgen, G. J. Park, M. Novotni, and R. Klein. 3D shape matching with 3D shape contexts. In *Central European Seminar on Computer Graphics*, Slovakia, April 2003. 93

[72] S. Kullback and R.A. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22 :79–86, 1951. 111

[73] J. Löffler. Content-based retrieval of 3D models in distributed web databases by visual shape information. In *International Conference on Information Visualisation*, pages 82–87, London, England, UK, July 2000. 105

[74] D. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2) :91–110, 2004. 131

[75] D. Macrini, A. Shokoufandeh, S. J. Dickinson, K. Siddiqi, and S. W. Zucker. View-based 3D object recognition using shock graphs. In *The International Conference on Pattern Recognition (ICPR'02)*, pages 24–28, Québec, Canada, August 2002. 105

[76] S. Mahmoudi and M. Daoudi. 3D models retrieval by using characteristic views. In *The International Conference on Pattern Recognition (ICPR'02)*, pages 11–15, Québec, Canada, August 2002. 105, 106, 115

[77] A. Martinet, C. Soler, N. Holzschuch, and F. X. Sillion. Accurate detection of symmetries in 3D shapes. *ACM Trans. on Graphics*, 25(2) :439–464, 2006. 50

[78] G. Medioni and A. François. 3D structures for generic object recognition. In *15th International Conference on Pattern Recognition*, volume 1, pages 30–37, September 2000. 102

[79] P. Minovic, S. Ishikawa, and K. Kato. Symmetry identification of a 3-D object represented by octree. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 15(5) :507–514, 1993. 53

[80] N. J. Mitra, L. Guibas, and M. Pauly. Partial and approximate symmetry detection for 3D geometry. *ACM Transactions on Graphics*, 25(3) :560–568, 2006. 50

[81] R. Muthuganapathy and K. Ramani. SHape REtrieval Contest 2008 : CAD models. In *IEEE International Conference on Shape Modeling and Applications (SMI'08)*, pages 221–222, Stony Brook, NY, USA, June 2008. 86

[82] T. Napoleon, T. Adamek, F. Schmitt, and N. E. O'Connor. SHREC'08 entry : Multi-view 3D retrieval using multi-scale contour representation. In *IEEE International Conference on Shape Modeling and Applications (SMI'08)*, pages 227–228, Stony Brook, NY, USA, June 2008. 106, 115

[83] S. Needleman and C. Wunsch. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, 48(3) :443–453, 1970. 147

[84] M. Novotni, P. Degener, and R. Klein. Correspondence generation and matching of 3D shape subparts. Technical report, Universität Bonn, Germany, 2005. 86

[85] M. Novotni and R. Klein. 3D Zernike descriptors for content based shape retrieval. In *Proceedings of the eighth ACM symposium on Solid modeling and applications*, pages 216–225. ACM Press, 2003. 96

[86] Network of Excellence Project AIM@SHAPE, IMATI-Ge CNR Shape Modelling Group. SHREC'07 Watertight Database. http://watertight.ge.imati.cnr.it/. 26, 209

[87] Network of Excellence Project AIM@SHAPE, National Institut of Standards and Technology. SHREC'09 Generic Database. http://www.itl.nist.gov/iad/vug/sharp/benchmark/shrecGeneric/data.html. 26, 27, 211

[88] Network of Excellence Project AIM@SHAPE, Utrecht University, Center for Geometry, Imaging and Virtual Environments. SHREC'07 3D Face Database. http://give-lab.cs.uu.nl/SHREC/shrec2007. 26

[89] The University of York, Department of Computer Science. York University Database. http://www-users.cs.york.ac.uk/~nep/research/3Dface/tomh/3DFaceDatabase.html. 26

[90] R. Ohbuchi. SHape REtrieval contest 2008 : Generic models. In *IEEE International Conference on Shape Modeling and Applications (SMI'08)*, pages 223–224, Stony Brook, NY, USA, June 2008. 86

[91] R. Ohbuchi, T. Minamitani, and T. Takei. Shape-similarity search of 3D models by using enhanced shape functions. *International Journal of Computer Applications in Technology (IJCAT)*, 23(3/4/5) :70–85, 2005. 88, 114

[92] R. Ohbuchi, M. Nakazawa, and T. Takei. Retrieving 3D shapes based on their appearance. In *5th ACM SIGMM Workshop on Multimedia Information Retrieval (MIR 2003)*, Berkeley, Californica, USA, November 2003. 105, 110, 111, 116, 139

[93] R. Ohbuchi, K. Osada, T. Furuya, and T. Banno. Salient local visual featuers for shape-based 3D model retrieval. In *IEEE International Conference*

*on Shape Modeling and Applications (SMI'08)*, pages 93–102, Stony Brook, NY, USA, June 2008. 105, 111, 116, 139

[94] R. Ohbuchi, T. Otagiri, M. Ibato, and T. Takei. Shape-similarity search of three-dimensional models using parameterized statistics. In *Pacific Graphics 2002*, Beijing, China, October 2002. 92, 219

[95] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin. Matching 3D models with shape distributions. In *International Conference on Shape Modeling and Applications*, pages 154–166, Genova, May 2001. 87, 190

[96] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin. Shape distributions. *ACM Transactions on Graphics*, 21(4) :807–832, October 2002. 87, 88, 114, 219

[97] E. Paquet and M. Rioux. Nefertiti : A query by content software for three-dimensional models databases management. In *International Conference on Recent Advances in 3-D Digital Imaging and Modeling (3DIM '97)*, pages 345–352, Ottawa, Ontario, Canada, May 1997. 87, 114

[98] E. Paquet and M. Rioux. The mpeg-7 standard and the content-based management of three-dimensional data : A case study. In *IEEE International Conference on Multimedia Computing and Systems (ICMCS'99)*, volume 1, pages 375–380, Florence, Italy, June 1999. 95

[99] E. Paquet, M. Rioux, A. Murching, T. Naveen, and A. Tabatabai. Description of shape information for 2-D and 3-D objects. *Signal Processing : Image Communication*, 16 :103–122, 2000. 36, 44

[100] G. Passalis1, T. Theoharis, and I. A. Kakadiaris. PTK : a novel depth buffer-based shape descriptor for three-dimensional object retrieval. *The Visual Computer*, 23(1) :5–14, January 2006. 105, 110, 115, 139

[101] J. Podolak, P. Shilane, A. Golovinskiy, S. Rusinkiewicz, and T. Funkhouser. A planar-reflective symmetry transform for 3D shapes. *ACM Transactions on Graphics*, 25(3) :549–559, 2006. 36, 41, 44, 50, 60, 100

[102] O. Polonsky, G. Patanè, S. Biasotti, C. Gotsman, and M. Spagnuolo. What's in an image ? *The Visual Computer*, 21(8-10) :840–847, 2005. 128

[103] J. Pu, Y. Liu, X. Gu, H. Zha, W. Liu, and Y. Uehara. 3D model retrieval based on 2D slice similarity measurements. In *The second International Symposium on 3D Data Processing, Visualization, and Transmis-*

*sion(3DPVT'04)*, pages 95–101, Thessaloniki,Greece, September 2004. 111, 112, 220

[104] Princeton Shape Retrieval and Analysis Group. Princeton Shape Benchmark.
http://shape.cs.princeton.edu/benchmark/. 26, 27, 207

[105] J. Ricard. *Indexation et recherche d'objets 3D à partir de requêtes 2D et 3D*. PhD thesis, Université Claude Bernard Lyon 1, December 2005. 98, 105

[106] J. Ricard, D. Coeurjolly, and A. Baskurt. ART extension for description, indexing and retrieval of 3d objects. In *International Conference on Pattern Recognition (ICPR'04)*, volume 3, pages 79–82, Cambridge, UK, August 2004. 98

[107] Y. Rubner, C. Tomasi, and L.J. Guibas. The earth mover's distance as a metric for image retrieval. *International Journal of Computer Vision*, 40(2) :99–121, November 2000. 197

[108] R.M. Rustamov. Augmented symmetry transforms. In *IEEE International Conference on Shape Modeling and Applications (SMI'07)*, pages 13–20, Lyon, France, June 2007. 50

[109] D. Saupe and D. V. Vranic. 3D model retrieval with spherical harmonics and moments. In *Proceedings of the 23rd DAGM-Symposium on Pattern Recognition*, pages 392–397, Munich, Germany, September 2001. Springer-Verlag. 96, 190

[110] Sculpteur R&D Project (Semantic and content-based multimedia exploitation for European benefit). Sculpteur Database.
http://www.sculpteurweb.org/. 26

[111] Y.T. Shen, D.Y. Chen, X.P. Tian, and M. Ouhyoung. 3D model search engine based on lightfield descriptors. In *Eurographics 2003*, Granada, Spain, September 2003. 107

[112] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser. The Princeton shape benchmark. In *Shape Modeling and Applications Conference (SMI'04)*, pages 167–178, Genova, Italy, June 2004. IEEE. 86, 89, 109, 157, 190, 191

[113] Y. Shinagawa, T. L. Kunii, and Y. L. Kergosien. Surface coding based on morse theory. *IEEE Computer Graphics and Applications*, 11(5) :66–78, 1991. 103

[114] P. Simari, E. Kalogerakis, and K. Singh. Folding meshes : Hierarchical mesh segmentation based on planar symmetry. In *Fourth Eurographics Symposium on Geometry Processing*, pages 111–120, June 2006. 50

[115] T. F. Smith and M. S. Waterman. Identification of common molecular subsequences. *J Mol Biol*, 147(1) :195–197, March 1981. 147

[116] C. Sun and J. Sherrah. 3D symmetry detection using the extended gaussian image. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(2) :164–168, 1997. 50

[117] H. Sundar, D. Silver, N. Gagvani, and S. Dickinson. Skeleton based shape matching and retrieval. In *Shape Modeling and Applications Conference (SMI'03)*, Seoul, Korea, May 2003. IEEE. 102, 219

[118] J.W.H. Tangelder and R.C. Veltkamp. Polyhedral model retrieval using weighted point sets. In *Shape Modeling and Applications Conference (SMI'03)*, Seoul, Korea, May 2003. IEEE. 93

[119] J.W.H. Tangelder and R.C. Veltkamp. A survey of content based 3D shape retrieval methods. *Multimedia Tools and Applications*, 39(3) :441–471, September 2008. 86, 94

[120] F. B. ter Haar, M. Daoudi, and R. C. Veltkamp. SHape REtrieval Contest 2008 : 3D face scans. In *IEEE International Conference on Shape Modeling and Applications (SMI'08)*, pages 225–226, Stony Brook, NY, USA, June 2008. 86

[121] T. Tung. *Indexation 3D de bases de données d'objets par graphes de Reeb améliorés*. PhD thesis, TSI, ENST Paris, June 2005. 104

[122] T Tung and F. Schmitt. Augmented Reeb graphs for content-based retrieval of 3D mesh models. In *Shape Modeling and Applications Conference (SMI'04)*, pages 157–166, Genova, Italy, June 2004. IEEE. 104, 220

[123] National Taiwan University. NTU Database. http://3d.csie.ntu.edu.tw/~dynamic/database/index.html. 26

[124] N. Vajramushti, I. A. Kakadiaris, T. Theoharis, and G. Papaioanno. Efficient 3D object retrieval using depth images. In *Proceedings of the 6th ACM SIGMM International Workshop on Multimedia Information Retrie-*

*val (MIR'04)*, pages 189–196, New York, USA, October 2004. ACM Press. 105, 110, 115, 139

[125] R. C. Veltkamp, R. Ruijsenaars, M. Spagnuolo, R. van Zwol, and Frank ter Haar. SHREC2006 : 3D shape retrieval contest. Technical Report UU-CS-2006-030, Utrecht University, 2006. 86, 207

[126] R. C. Veltkamp and Frank ter Haar. SHREC 2007 : 3D shape retrieval contest. Technical Report UU-CS-2007-015, Utrecht University, 2007. 86, 209

[127] D. Vranic, D. Saupe, and J. Richter. Tools for 3D-object retrieval : Karhune-Loeve transform and spherical harmonics. In *2001 Workshop Multimedia Signal Processing*, Cannes, France, October 2001. 36, 38, 44, 53, 68, 99

[128] D. V. Vranic and D. Saupe. 3D shape descriptor based on 3D Fourier transform. In K. Fazekas, editor, *EURASIP Conference on Digital Signal Processing for Multimedia Communications and Services (ECMCS'01)*, pages 271–274, Budapest, September 2001. 97

[129] D. V. Vranic and D. Saupe. Description of 3D-shape using a complex function on the sphere. In *IEEE International Conference on Multimedia and Expo (ICME'02)*, pages 177–180, Lausanne, August 2002. 99, 100

[130] D.V. Vranic. An improvement of rotation invariant 3D-shape descriptor based on functions on concentric spheres. In *IEEE International Conference on Image Processing (ICIP'03)*, pages 757–760, Barcelona, Spain, September 2003. 99, 190

[131] D.V. Vranic. *3D Model Retrieval*. PhD thesis, University of Leipzig, 2004. 42, 86, 94, 95, 96, 100, 105, 106, 109, 110, 115, 133, 167, 170, 190

[132] D. Wang, J. Zhang, H. Wong, and Y. Li. 3D model retrieval based on multi-shell extended gaussian image. In *Advances in Visual Information Systems, 9th International Conference, VISUAL,*, pages 426–437, Shanghai, China, 2007. Springer. 89

[133] K. Wang and T. Gasser. Alignement of curves by dynamic time warping. *The Annals of Statistics*, 25(3) :1251–1276, 1997. 147

[134] H. Zabrodsky, S. Peleg, and D. Avnir. Symmetry as a continuous feature. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 17(12) :1154–1166, 1995. 59

[135] T. Zaharia. *Indexation de vidéos et de maillages 3D dans le contexte MPEG-7*. PhD thesis, Université René Descartes - Paris V, December 2001. 89, 97

[136] T. Zaharia and F. Prêteux. Hough transform-based 3D mesh retrieval. In *SPIE Conference 4476 on Vision Geometry X, San Diego, CA*, pages 175–185, August 2001. 97, 107

[137] T. Zaharia and F. Prêteux. 3D versus 2D/3D shape descriptors : A comparative study. In *Proceedings SPIE Conference on Image Processing : Algorithms and Systems III - IS & T / SPIE Symposium on Electronic Imaging, Science and Technology '03, San Jose, CA*, volume 5298, January 2004. 86

[138] T. Zaharia and F. Prêteux. Descripteurs de forme : Etude comparée des approches 3D et 2D/3D. In *14ème Congrès Francophone AFRIF-AFIA Reconnaissance des Formes et Intelligence Artificielle (RFIA'04), Toulouse, France*, January 2004. 86

[139] T. Zaharia, F. Prêteux, and M. Preda. 3D shape spectrum descriptor. In *MPEG-7 ISO/IEC JTC1/SC29/WG11 MPEG99/M5242*, Melbourne, Australia, October 1999. 89, 90, 219

[140] D. S. Zhang and G. Lu. Shape-based image retrieval using generic Fourier descriptor. *Signal Processing : Image Communication*, 17(10) :825–848, November 2002. 110

# Liste des publications

## Article de journal

M. Chaouch and A. Verroust-Blondet. Alignment of 3D models. *Graphical Models*, 71 :63–76, 2009.

## Articles de conférences internationales

M. Chaouch and A. Verroust-Blondet. 3D Gaussian Descriptor for 3D Shape Retrieval. In *IEEE International Conference on Multimedia & Expo (ICME'09)*, Cancun, Mexico, July 2009.

M. Chaouch and A. Verroust-Blondet. A novel method for alignment of 3D models. In *IEEE International Conference on Shape Modeling and Applications (SMI'08)*, pages 187–195, Stony Brook, NY, USA, June 2008.

M. Chaouch and A. Verroust-Blondet. A new descriptor for 2D depth image indexing and 3D model retrieval. In *IEEE International Conference on Image Processing (ICIP'07)*, volume VI, pages 373–376, San Antonio, Texas, USA, September 2007.

M. Chaouch and A. Verroust-Blondet. 3D model retrieval based on depth line descriptor. In *IEEE International Conference on Multimedia & Expo (ICME'07)*, pages 599–602, Beijing, China, July 2007.

M. Chaouch and A. Verroust-Blondet. Enhanced 2D/3D approaches based on relevance index for 3D-shape retrieval. In *International Conference on Shape Modeling and Applications (SMI'06)*, page 36, Matsushima, Japan, June 2006.

C. Akgül, A. Axenopoulos, B. Bustos, M. Chaouch, P. Daras, H. Dutagaci, T. Furuya, A. Godil, S. Kreft, Z. Lian, T. Napoléon, A. Mademlis, R. Ohbuchi, P. L. Rosin, B. Sankur, T. Schreck, X. Sun, M. Tezuka, Y. Yemez, A. Verroust-Blondet,

and M. Walter. SHREC 2009 - Generic Shape Retrieval Contest. In *EUROGRAPHICS Workshop on 3D Object Retrieval*, 2009.

## Rapport techniques

M. Chaouch and A. Verroust-Blondet. 2D/3D descriptor based on depth line encoding. In *SHREC 2007 : 3D shape retrieval contest. Technical Report UU-CS-2007-015, Utrecht University*, R. C. Veltkamp, and Frank ter Haar, June, 2007.

M. Chaouch and A. Verroust-Blondet. Enhanced silhouette and depth-buffer based approaches for 3D-shape retrieval. In *SHREC 2006 : 3D shape retrieval contest. Technical Report UU-CS-2006-030, Utrecht University*, R. C. Veltkamp, R. Ruijsenaars, M. Spagnuolo, R. van Zwol, and Frank ter Haar, June, 2006.