



**HAL**  
open science

# Indexation 3D de bases de données d'objets par graphes de Reeb améliorés

Tony Tung

► **To cite this version:**

Tony Tung. Indexation 3D de bases de données d'objets par graphes de Reeb améliorés. Interface homme-machine [cs.HC]. Télécom ParisTech, 2005. Français. NNT: . tel-00009843

**HAL Id: tel-00009843**

**<https://pastel.hal.science/tel-00009843>**

Submitted on 28 Jul 2005

**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'Ecole Nationale Supérieure des  
Télécommunications

Spécialité : **Signal et Images**

**Tony TUNG**

Indexation 3D de bases de données d'objets  
par graphes de Reeb améliorés

Soutenue le 10 juin 2005 devant le jury composé de :

Françoise Prêteux	Présidente
Atilla Baskurt	Rapporteur
Mohamed Daoudi	Rapporteur
Frédéric Jurie	Examineur
Francis Schmitt	Directeur de thèse

---

# Remerciements

Je dédie ce travail de thèse à **ma mère**, Mme Kinh Thang, pour la remercier d'être toujours aussi fière de moi. Cette thèse représente beaucoup pour moi et je suis donc d'autant plus heureux de pouvoir la lui dédier.

Tout n'a pas toujours été facile pendant ces trois années, et je tiens donc à remercier grandement mon directeur de thèse Francis Schmitt, sans qui je n'aurais jamais réussi à faire cette thèse. Je le remercie pour sa gentillesse, son soutien, et toute l'attention qu'il m'a porté tout au long de la thèse.

Je tiens aussi à remercier les membres de mon jury de thèse, d'avoir accepté de s'intéresser à nos travaux. Un grand merci à :

- Mme Françoise Prêteux, professeur à l'INT, d'avoir m'avoir fait l'honneur d'accepter de présider le jury de cette thèse,
- MM. les professeurs Atilla Baskurt et Mohamed Daoudi, experts ès indexation 3D, d'avoir accepté de rapporter cette thèse. Je tiens d'ailleurs à leur témoigner toute ma gratitude pour leur gentillesse et pour la pertinence de leurs remarques.
- M. Frédéric Jurie, chercheur CNRS, d'avoir accepté d'être examinateur de ma thèse et pour ses encouragements.

Merci aux partenaires du projet Européen Sculpteur, Christian Lahanier, Ruven Pillay, Denis Pitzalis, Fabrizio Giorgini, Matthew Addis, Kirk Martinez, Paul Lewis, James Stevenson, Joe Padfield, Simon Goodall, Jean-Luc Dufresne, Raffaella, Marco, Barbara et Benjamin pour la belle réussite du projet.

J'ai adoré travailler au département TSI de Télécom Paris! Tous les permanents sont excellents et sympathiques. Je leur remercie donc pour le bon accueil qu'ils m'ont fait dès le premier jour, et pour avoir été tous aussi disponible. Merci merci merci à Isabelle Bloch, Yann Gousseau, Dominique Asselineau, Michel Roux, Florence Tupin, Jean-Marie Nicolas, Hans Brettel, Elsa Angelini, Jamal Atif :op, Marine Campedel, et au chef Henri Maître. (Gros respect à tous!) Je tiens à remercier Sophie-Charlotte Barrière sans qui je n'aurai jamais pu envoyer mon manuscrit à temps! Un grand merci aussi à Pat, Patricia Friedrich, pour sa bonne humeur et son aide :-)

Evidemment, ces trois années à TSI n'auront pas pu être aussi *merveilleuses* sans certaines personnes... J'adresse donc un merci sans limite à mes compagnons de la première heure : Antonio de m'avoir toujours soutenu à tous les moments (besito gua-

---

peton! beso Anton!), au Dr. Alejandro qui toujours là pour moi aussi, à l'école, au taekwondo ou au bar (jajajajaaaa y beso à Milena), Dr. Gaspaaar (il n'a pas l'air mais il est drôle y beso à Wietke), Dr. Carlos le plus parisien des anglais (ouééé ton viewer il déchire, merci à lui aussi, y beso à Arantxa), Dr. Oscar el cabronito del amor (y beso à Maite), Dr. radioactiv Sveta :-), les sirènes Dr. Céline (la meilleure) et Dr. Ferdaous (moi aussi je t'aime bien), Pau el rey de la noche (que hombre! y beso a Fernanda), Daniel et ses inoubliables soirées Magenta (allez-y), Florent trop fort (merci!), Gero zu viiiel stark, Saeid trop gentil, Dr. Najib toujours là, François et ses tatouages, Dr. Dalila et ses blagues, Hassan aussi(!), Yong (xie xie), Eve :o), Dr. Thomas le mélomane polyglotte, et MERCI aussi la nouvelle génération, très prometteuse : Thomas (vive l'art painting), Ceyhun (la relève :op), Céline (bisou :o\*), Jean-François (glouglou), François (vive le bois), Julien (stresse pas), Cédric, Adel, Olivier, Luo, Yvan, Mihai. Et merci aussi aux 45851 stagiaires, en particulier Pierre-Alexandre Pont, Nicolas et Thanh Ha, Clément, Céleste et Farah qui ont participé au projet, Houda et Marie pour leur sourire.

Bien sûr, un énorme merci au Dr. Said Ladjal, le plus fort (et pas qu'en math). Merci pour ton aide très très précieuse, merci de m'avoir soutenu, et merci de m'avoir laissé gagner au bras de fer (!).

Non loin de moi, mes amis, que je remercie de m'avoir soutenu et encouragé. Merci à vous d'avoir été là et d'être toujours là. Franckyyy t'es BON, Ricou (allez je le dis : vive la montagne! ) et Nadège (bisou), Dr. Chen (merci d'être toujours là), Jérôme (vive la nuit d'hiver), Raphiot l'homme de la night, Eric et Steph toujours en vacances. Merci aussi à Emeline from Le Mans, Marie et Stéphane (kiss), Dr. Yassine (meuah), Asher-chouchou, la guapa Marta, a mi amigo Genaro (beso), y a los mexicanos muy calientes Fausto, Adriana y Esme!

J'embrasse et remercie très fort ma soeur Karine. Merci pour tes encouragements et toute ta tendresse. Merci de m'avoir donné l'exemple pendant toute nos jeunes années que je n'oublierai jamais. Merci aussi Mama (orkoun) pour tes prières et ta gentillesse. Merci A-Kong, j'espère que tu es fier du travail de ton petit-fils! Merci aussi Tri d'être toujours là pour nous aider (bisou à Pascal). Merci aux cnoilles et aux ctiouks, Petit B, Bessorrou, Boulgoma.

Sans égal, je remercie du fond du cœur ma loutre Virginie sans qui je n'aurais jamais pu commencer cette thèse. Cette thèse est aussi la tienne (et à Nouski). Merci au soutien de la famille Vagné (bises à tous Tony, Marie-Thérèse, Jean-Marie, Pépé, Mémé, Nénette, Alain, Guillaume, Vivien, Larry et Eliott).

Mi Penélope corazoncita! Gracias para todo! Ahora he terminado y es gracias de ti! C'est grâce à toi que j'ai pu finir cette thèse! Muchas gracias para tu amor y para estar siempre a lado de mi. Te envio mucho mucho mucho besos :o\*

# Résumé

La rapide avancée de la technologie numérique a permis d'améliorer les méthodes d'acquisition et de rendu de modèles 3D. On peut constater qu'aujourd'hui les bases de données d'objets 3D sont présentes dans beaucoup de domaines, qu'ils soient ludiques (jeux, multimédia) ou scientifiques (applications médicales, industrielles, héritage culturel, etc.). La facilité d'acquisition et de reconstruction des modèles 3D, ainsi que leur modélisation permettent de créer de grandes bases de données, et il devient difficile de naviguer dans ces bases pour retrouver des informations. L'indexation des objets 3D apparaît donc comme une solution nécessaire et prometteuse pour gérer ce nouveau type de données. Notre étude s'insérant dans le cadre du projet européen SCULPTEUR IST-2001-35372 dont des partenaires étaient des musées, nous avons donc travaillé avec des bases de données de modèles 3D muséologiques. L'indexation des éléments d'une base de données consiste à définir une méthode permettant d'effectuer des comparaisons parmi les composants de cette base. Actuellement, une des principales applications consiste à effectuer des requêtes de similarité : étant donné une "clé" de recherche, on extrait de la base de données les éléments ayant la clé la plus similaire.

Nous présentons dans ce mémoire une méthode d'indexation de modèles 3D appliquée aux recherches par similarité de forme et d'aspect dans des bases de données d'objets 3D. L'approche repose sur la méthode d'appariement de graphes de Reeb multirésolution proposée par [Hilaga et al., 2001]. Dans le cadre de notre étude, nous travaillons avec des maillages de modèles 3D de géométrie plus ou moins complexes, à différents niveaux de résolution, et parfois texturés. L'approche originale, basée sur la topologie des objets 3D, s'est avérée insuffisante pour obtenir des appariements satisfaisants. C'est pourquoi nous proposons d'étendre les critères de cohérence topologique pour les appariements et de fusionner au graphe des informations géométriques et visuelles pour améliorer leur mise en correspondance et l'estimation de la similarité entre modèles. Ces attributs sont librement pondérables afin de s'adapter au mieux aux requêtes d'un utilisateur. Nous obtenons une représentation souple, multicritère et multirésolution que nous nommons graphe de Reeb multirésolution augmenté (aMRG). Nous comparons cette approche à un ensemble varié de méthodes d'indexation. Elle se révèle être très performante pour retrouver les objets de formes similaires et discerner les différentes classes de formes 3D.

---

# Abstract

The strong development of numerical technologies has led to efficient 3D acquisition of real objects and rendering of 3D methods. Nowadays 3D object databases appear in various areas for leisure (games, multimedia) as well as for scientific applications (medical, industrial part catalogues, cultural heritage, etc.). Large database can be nowadays quickly populated using 3D mesh acquisition and reconstruction tools which have become easy to use, and with new ergonomic 3D design tools which have become very popular. As database size is growing, tools to retrieve information become more and more important. 3D object indexing appears to be a useful and very promising way to manage this new kind of data. As our study took place in the framework of the European project SCULPTEUR IST-2001-35372 which involved museums, we worked with museological 3D model databases. Database indexing consists on defining a method able to perform comparisons between the database elements. Similarity retrieval is one of the main application : using a research “key”, a subset of elements with the most similar keys are extracted from the database.

This manuscript presents a 3D shape matching method for 3D mesh models applied to content-based search in database of 3D objects. The approach is based on the multiresolution Reeb graph (MRG) proposed by [Hilaga et al., 2001]. MRG provides a rich representation of shapes able in particular to embed the object topology. In our framework, we consider 3D mesh models of various geometrical complexity, of different resolution, and when available with color texture map. The original approach, mainly based on the 3D object topology, is not accurate enough to obtain satisfying matching. Therefore we propose to reinforce the topological consistency conditions of the matching and to merge within the graph geometrical and visual information to improve matching and calculation of shape similarity between models. Besides, all these new attributes can be freely weighted to fit the user requirements for object retrieval. We obtain a flexible multiresolutional and multicriteria representation that we called augmented multiresolution Reeb graph (aMRG). The approach has been tested and compared with other methods. It reveals very performant for the retrieval and the classification of similar 3D shapes.



# Table des Matières

<b>Remerciements</b>	<b>iii</b>
<b>Résumé</b>	<b>v</b>
<b>Abstract</b>	<b>vii</b>
<b>Introduction</b>	<b>1</b>
<b>1 Etat de l'art de l'indexation 3D</b>	<b>5</b>
1.1 L'indexation 3D : un domaine récent de recherche . . . . .	5
1.2 Normalisation des données . . . . .	8
1.3 Méthodes globales . . . . .	10
1.3.1 Aire et volume . . . . .	10
1.3.2 Histogrammes de cordes . . . . .	10
1.3.3 Distribution de forme . . . . .	12
1.3.4 Descripteur de Hough 3D . . . . .	13
1.4 Méthodes locales . . . . .	14
1.4.1 Descripteur de spectre de forme 3D . . . . .	14
1.4.2 "Extended Gaussian Images" . . . . .	16
1.5 Approches par transformées . . . . .	16
1.5.1 Transformée de Fourier 3D . . . . .	16
1.5.2 Harmoniques Sphériques . . . . .	17
1.5.3 Calculs de moments . . . . .	18
1.5.4 Transformation Radiale Angulaire 3D . . . . .	19
1.6 Partition de l'espace . . . . .	19
1.6.1 Ensemble de points pondérés . . . . .	19
1.6.2 Contextes de forme 3D . . . . .	20
1.6.3 Structures volumétriques élémentaires 3D . . . . .	21
1.7 Approches par graphes . . . . .	21
1.7.1 Squelette par axe médian . . . . .	21
1.7.2 Extraction de graphes . . . . .	22
1.7.3 Graphe de Reeb . . . . .	22
1.8 Mesure de similarité . . . . .	23
1.8.1 Définitions . . . . .	24
1.8.2 Normes de Minkowski . . . . .	24

1.8.3	Distance de Bhattacharyya . . . . .	25
1.8.4	Distance de Earth Mover . . . . .	25
1.8.5	La divergence de Kullback-Leibler . . . . .	26
1.8.6	Divergence du $\chi^2$ . . . . .	27
1.8.7	Distance de Hausdorff . . . . .	27
1.8.8	Intersection d'histogrammes . . . . .	27
1.9	Approches 2D . . . . .	28
1.9.1	Histogramme de couleur . . . . .	29
1.9.2	Histogramme monochrome . . . . .	30
1.9.3	Région de couleur CCV . . . . .	31
1.9.4	Histogramme monochrome multi-échelle . . . . .	32
1.9.5	Région de couleur multi-échelle MCCV . . . . .	32
1.9.6	Transformée en ondelettes pyramidales (PWT) . . . . .	33
1.9.7	Requête par fax (QBF) . . . . .	33
1.10	Bilan . . . . .	34
<b>2</b>	<b>Graphe de Reeb multirésolution</b> . . . . .	<b>37</b>
2.1	Théorie de Morse . . . . .	37
2.2	Graphe de Reeb . . . . .	39
2.2.1	Définition du graphe de Reeb . . . . .	39
2.2.2	Exemple . . . . .	39
2.2.3	Choix de la fonction $\mu$ . . . . .	40
2.2.4	Discrétisation de la fonction $\mu$ . . . . .	42
2.2.5	Choix des bases . . . . .	43
2.2.6	Algorithme de Dijkstra . . . . .	44
2.2.7	Normalisation de $\mu$ . . . . .	45
2.2.8	Construction du graphe de Reeb . . . . .	46
2.3	Graphe de Reeb Multirésolution . . . . .	47
2.3.1	Définition . . . . .	47
2.3.2	Exemple . . . . .	49
2.3.3	Propriétés . . . . .	49
2.4	Discussion sur les graphes MRG . . . . .	50
2.4.1	Choix de la fonction $\mu$ . . . . .	50
2.4.2	Résolution du graphe . . . . .	50
2.4.3	Maillage . . . . .	52
2.5	Bilan . . . . .	54
<b>3</b>	<b>Indexation 3D avec les graphes de Reeb multirésolution augmentés</b> . . . . .	<b>55</b>
3.1	Appariements de graphes aMRG . . . . .	55
3.1.1	Attributs des nœuds . . . . .	56
3.1.2	Stratégie . . . . .	57
3.1.3	Critères de cohérence topologique . . . . .	57
3.1.4	Critère de cohérence topologique supplémentaire . . . . .	59
3.1.5	Discrimination analytique pour l'appariement des nœuds . . . . .	60
3.1.6	Ajout d'attributs topologiques . . . . .	61

3.1.7	Ajout d'attributs géométriques . . . . .	64
3.2	Procédure d'appariements des nœuds . . . . .	66
3.3	Calcul de similarité de nœuds . . . . .	67
3.4	Ajout d'attributs géométriques et visuels . . . . .	68
3.4.1	Motivations . . . . .	68
3.4.2	Calcul du volume . . . . .	69
3.4.3	Mesures des cordes . . . . .	70
3.4.4	Mesure des courbures . . . . .	71
3.4.5	Descripteur de Hough 3D . . . . .	73
3.4.6	Quantification des couleurs . . . . .	74
3.5	Similarité d'objets . . . . .	75
3.6	Mise en correspondance d'ensemble de nœuds . . . . .	76
3.7	Indexation des graphes . . . . .	77
3.8	Bilan . . . . .	79
<b>4</b>	<b>Résultats expérimentaux</b>	<b>81</b>
4.1	Bases de données . . . . .	81
4.2	Techniques d'évaluation . . . . .	85
4.3	Etude des méthodes implémentées . . . . .	91
4.3.1	Le rapport aire/volume . . . . .	92
4.3.2	Histogrammes de cordes . . . . .	94
4.3.3	Distribution de forme D2 . . . . .	102
4.3.4	Les harmoniques sphériques . . . . .	105
4.3.5	Les EGI complexes . . . . .	106
4.3.6	Le descripteur de spectre de forme 3D . . . . .	109
4.3.7	Le descripteur de Hough 3D . . . . .	112
4.3.8	Bilan . . . . .	115
4.4	Résultats des aMRG . . . . .	116
4.5	Bilan . . . . .	123
<b>5</b>	<b>Applications</b>	<b>127</b>
5.1	Recherche d'objets 3D texturés . . . . .	127
5.2	Recherche de sous-graphes . . . . .	130
5.3	Recherche 2D/3D . . . . .	134
<b>6</b>	<b>Conclusion et perspectives</b>	<b>137</b>
6.1	Conclusion . . . . .	137
6.2	Perspectives de recherche . . . . .	138
<b>A</b>	<b>Bases de données</b>	<b>141</b>
<b>B</b>	<b>Interface développée</b>	<b>147</b>
	<b>Bibliographie</b>	<b>149</b>
	<b>Publications</b>	<b>157</b>

## TABLE DES MATIÈRES

---

# Introduction

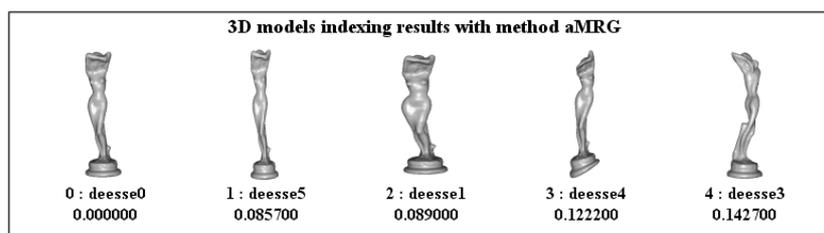


FIG. 1 – Résultat de la recherche de l’objet déesse (n° 0) dans une base de données de 567 objets.

La rapide avancée des multiples technologies numériques a permis d’améliorer les méthodes d’acquisition et de rendu de modèles 3D. On peut constater qu’aujourd’hui les bases de données d’objets 3D sont présentes dans beaucoup de domaines, qu’ils soient ludiques (jeux <sup>1</sup>, multimédia <sup>2</sup>) ou scientifiques (applications médicales <sup>3</sup> industrielles <sup>4</sup>, héritage culturel <sup>5</sup>, etc.). La plus grande facilité d’acquisition et de reconstruction des modèles 3D permettent aujourd’hui de créer de grandes bases de données où il devient difficile de naviguer pour retrouver des informations. L’indexation des objets 3D apparaît donc comme une solution nécessaire et prometteuse pour gérer ce nouveau type de données.

L’indexation d’une base de données consiste à définir une méthode permettant de déterminer pour chaque élément de la base une “clé” de recherche avec laquelle les éléments vont pouvoir être comparés et classés. Actuellement, une des principales applications consiste à effectuer des requêtes de similarité : étant donné un objet particulier sur lequel se porte l’intérêt, on détermine la clé de recherche à partir de son modèle 3D et on extrait de la base de données les éléments ayant les clés les plus similaires (cf. Figure 1). Le principe général commun à toutes les méthodes de requête

<sup>1</sup>Modèles 3D du jeu Wing Commander (<http://www.wcnews.com/wc3d/mesh.shtml>).

<sup>2</sup>Visite virtuelle (<http://www.visite-virtuelle.com>).

<sup>3</sup>3D medical illustration and artworks (<http://www.med-ars.it>), modèles 3D de molécules d’Audrey Sanderson (<http://www.dist214.k12.il.us/users/asanders/chemhome2.html>), base de données de protéines (<http://3d.csie.ntu.edu.tw/ProteinRetrieval>).

<sup>4</sup>Projet RNRT SEMANTIC-3D avec base de données RENAULT (<https://www.semantic-3d.net>).

<sup>5</sup>Base de données anthropologiques Nefertiti (<http://www.cleopatra.nrc.ca>), projet IST SCULPTEUR (<http://www.sculpteurweb.org>).

---

par similarité de forme 3D est de produire pour chaque élément de la base un *vecteur caractéristique* de la forme, et de proposer une mesure de similarité de forme permettant de comparer deux vecteurs caractéristiques.

C'est dans cet esprit qu'a vu le jour le projet Européen IST-2001-35372 SCULPTEUR (Semantic and Content-based multimedia exploitation for European benefit) qui regroupe des musées, des laboratoires de recherches universitaires et des industriels d'Angleterre, d'Italie et de France afin de mettre en place un système permettant de partager les ressources culturelles des partenaires du réseau. Dans le cadre de ce projet, nos travaux ont contribué à mettre en place un protocole d'acquisition et de constitution de bases de données d'objets 3D texturés pour les musées et, parallèlement, à développer des outils de recherches basées sur le contenu et la forme de ces objets. Nous avons essentiellement consacré nos efforts à ce dernier point, et nous en exposons les résultats dans ce manuscrit.

La majorité des approches proposées dans la littérature scientifique pour l'indexation 3D repose sur des représentations à l'aide d'histogrammes de mesures statistiques caractérisant un aspect géométrique des objets 3D. Ces différentes mesures sont calculées à partir de maillages triangulés de la surface des objets, ce type de maillage étant la représentation la plus simple et la plus fréquemment rencontrée sur Internet (avec le format VRML - Virtual Reality Modeling Language). Différents types d'approches existent. Les plus anciennes sont des applications directes de méthodes 2D : descripteur de Fourier, calculs de Moments invariants, extraction d'axe médian, etc. La littérature en matière de description de forme et d'indexation 2D est très riche. Les approches 3D dites globales permettent généralement de caractériser "grossièrement" la forme des objets. Elles s'avèrent souvent efficaces pour discriminer les formes simples mais insuffisamment discriminantes pour les formes plus complexes. A l'inverse, les approches 3D locales caractérisent des propriétés locales de la surface des objets. Elles sont donc dépendantes de la "qualité de la représentation de la surface" et sensibles aux bruits sur la surface des objets typiquement dus à l'échantillonnage et à la connectivité du maillage. Afin de répondre au cahier des charges proposé par le projet SCULPTEUR, c'est-à-dire trouver un descripteur de forme 3D efficace pour des modèles 3D hautement résolus provenant de numérisation d'objets d'art muséologiques, nous avons étudié et testé les approches les plus répandues.

En particulier, nous nous sommes intéressés à une approche consistant à évaluer la similarité de graphes <sup>6</sup> extraits d'objets 3D proposée par [Hilaga et al., 2001]. Elle consiste à calculer des graphes de Reeb <sup>7</sup> multirésolution (ou MRG) à partir des maillages des modèles 3D, et à appliquer un processus d'appariement topologique des graphes, puis à évaluer la similarité des formes des modèles en utilisant une fonc-

---

<sup>6</sup>Un graphe est composé de nœuds et d'arêtes reliant les nœuds. Un nœud de graphe correspond à une partie de l'objet.

<sup>7</sup>Les graphes de Reeb ont été proposés par le mathématicien Georges Reeb (1920-1992), professeur à l'Université de Strasbourg, suite à ses travaux de recherche sur la topologie des surfaces [Reeb, 1946].

tion de coût. La méthode, robuste et sophistiquée, nous a séduit par la richesse des informations capturées par la représentation multirésolution des MRG. En effet, les graphes MRG capturent des informations liées à la topologie des objets et décrivent bien la forme des objets de manière intuitive. Leur construction repose sur la théorie mathématique de Morse, et donc sur une fonction  $\mu$  calculée sur la surface des objets. Un choix adéquat de la fonction  $\mu$  permet d'obtenir des propriétés d'invariance géométrique de la représentation par graphe vis-à-vis de l'orientation des modèles et des propriétés de robustesse en cas de bruit sur la représentation utilisées pour les modèles.

L'étude des différentes méthodes d'indexation 3D et des travaux originaux de [Hilaga et al., 2001] nous a permis de connaître les avantages et limites des MRG et nous a conduit vers les solutions que nous proposons pour améliorer certains aspects de la méthode. Ainsi, nous avons renforcé la robustesse de la procédure d'appariement des graphes MRG et introduit des informations (ou attributs) topologiques dans les nœuds des graphes. Nous avons également ajouté des attributs géométriques globaux et locaux ainsi que des attributs colorimétriques, tout en profitant des aspects multirésolutions des MRG, pour améliorer la discrimination des formes 3D similaires. Chaque attribut apporte un type d'information lié à la forme ou à l'aspect d'une partie de l'objet. Notre idée principale a été d'utiliser les graphes de Reeb multirésolution comme support et de fusionner un ensemble de données, choisies par la pertinence de leur pouvoir descriptif, afin d'obtenir une caractérisation complète et précise des objets. Ces graphes ainsi enrichis sont appelés graphes de Reeb multirésolution augmentés (ou aMRG) (cf. Section Publications p. 157).

Afin d'évaluer l'efficacité des aMRG nous avons développé un banc de tests comprenant trois bases de données d'objets 3D de complexités différentes et mis en œuvre divers outils visuels et qualitatifs. Plusieurs autres méthodes d'indexation 3D ont été implémentés à titre informatif et comparatif. De plus nous avons étudié les résultats des calculs de similarité avec différentes mesures de distance.

Les résultats expérimentaux obtenus s'avèrent favorables à la méthode des aMRG que nous proposons et sont très prometteurs. Les aMRG possèdent un fort pouvoir discriminant et selon nos tests, permettent de retrouver efficacement les objets et classes d'objets similaires. De plus, nous proposons trois applications exploitant l'habileté des aMRG : la recherche basée sur la forme et la couleur, la recherche de parties d'objets, et la recherche 2D/3D.

Nous présentons dans le **Chapitre 1** les principaux travaux qui ont été effectués en matière de reconnaissance d'objets 3D dans des bases de données. L'étude des différentes approches (globales, locales, par squelette, etc.) nous a permis de bien comprendre les limites de leurs possibilités et de mieux cerner nos axes de recherches.

Le **Chapitre 2** présente l'approche des graphes de Reeb multirésolution (MRG)

proposée par [Hilaga et al., 2001] que nous avons retenue. Les MRG offrent une représentation intuitive de la forme des objets 3D et sont bien adaptés au problème de reconnaissance de forme qui est développé dans les chapitres suivants.

Nous proposons une méthode par graphes de Reeb multirésolution augmentés (aMRG) dans le **Chapitre 3**. Nous avons étendu l’approche originale de [Hilaga et al., 2001] en améliorant le procédé d’appariement de graphes et en enrichissant les graphes d’informations topologiques, géométriques et colorimétriques afin de proposer un outil précis de recherche de formes similaires 3D basée sur le contenu (cf. Figure 2).

L’étude qualitative au **Chapitre 4** permet de mettre en évidence les atouts et faiblesses des aMRG comparé aux différentes approches testées. Nous utilisons pour cela diverses méthodes d’évaluation inspirées d’outils de classification. Nous présentons les résultats que nous avons obtenus en appliquant la méthode sur différentes bases de données que nous avons constituées.

Le **Chapitre 5** présente les trois applications spécifiques très prometteuses déjà évoquées ci-dessus que permettent les graphes aMRG. En effet, cette méthode par graphes multirésolution est particulièrement satisfaisante pour la richesse des informations qu’elle capture et la qualité de la représentation par graphe obtenue.

Nous concluons au **Chapitre 6** avec une discussion sur la qualité de la méthode que nous proposons, ainsi que les possibles évolutions futures.

En annexes de nos travaux sont présentés le détail des bases qui ont été constituées spécialement pour tester les méthodes d’indexation 3D, ainsi qu’un aperçu de l’interface graphique de l’outil réalisé qui nous a permis d’obtenir ces résultats.

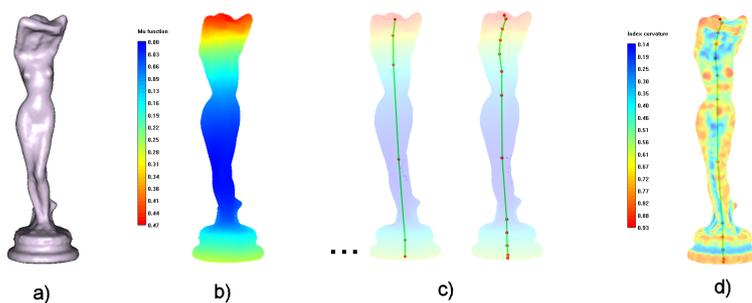


FIG. 2 – Construction du graphe de Reeb multirésolution augmenté du modèle déesse à la résolution 5 : a) modèle déesse ; b) calcul d’une fonction sur la surface du modèles ; c) construction des graphes de Reeb à différents niveaux de résolution ; d) calcul d’attributs topologiques, géométriques et colorimétriques.

## CHAPITRE 1

# Etat de l'art de l'indexation 3D

L'indexation d'objets 3D dans des bases de données est un thème de recherche très récent. Ce chapitre présente un état de l'art des différents descripteurs proposés dans la littérature. Nous présentons tout d'abord à la section 1.1 les données du problème, les enjeux, le format des objets 3D et les bases de données. La section 1.2 présente l'étape de normalisation des données (prétraitement par l'analyse en composantes principales) qui est requise par beaucoup de méthodes. Nous présentons ensuite dans les sections suivantes les diverses méthodes d'indexation 3D : les approches peuvent être globales (section 1.3), locales (section 1.4), par transformées (section 1.5), reposant sur des partitions de l'espace (section 1.6) ou par graphes (section 1.7). Aussi, nous présentons diverses formules de calcul de distances (section 1.8) pour estimer la similarité des vecteurs caractéristiques produits par ces descripteurs. Nous présentons à la section 1.9 les approches dérivées de méthodes 2D qui peuvent être appliquées, puis en section 1.10 nous présentons les méthodes que nous avons choisies d'implémenter pour nos tests.

## 1.1 L'indexation 3D : un domaine récent de recherche

L'indexation d'objets 3D est un axe de recherche encore très récent qui a débuté aux alentours de 1997. C'est un sujet qui suscite un intérêt croissant attirant de plus en plus d'équipe de recherche. On constate qu'environ 80% des publications datent des cinq dernières années comme indiqué dans l'étude bibliographique effectuée dans [Tangelder and Veltkamp, 2004], et que des outils de nature très différentes sont développés [Shilane et al., 2004, Goodall et al., 2004].

La problématique de l'indexation 3D consiste à décrire de manière compacte la forme d'un objet 3D (cf. Figure 1.1), le but étant principalement la reconnaissance de forme dans de très grosses bases de données. Pour cela, on utilise des *descripteurs de forme* afin d'obtenir des *vecteurs caractéristiques* ou *signatures* des objets 3D qui servent de clés dans les bases de données.

Nous ne considérons dans notre étude que les modèles 3D définis par leur surface, qui est la forme la plus répandue pour représenter les objets 3D. La surface d'un objet 3D est généralement définie par un maillage 3D composé de triangles. Si un objet

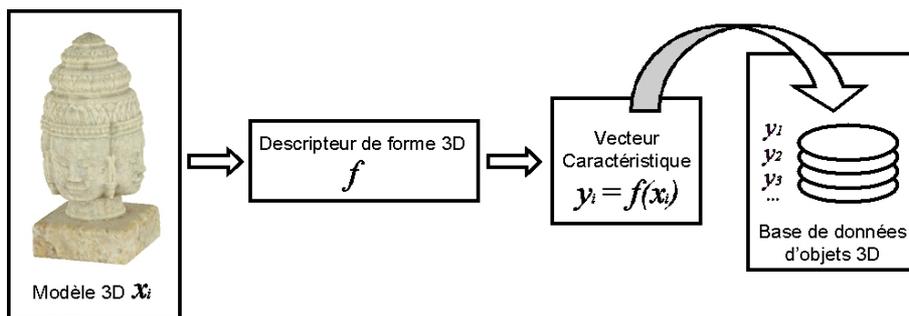


FIG. 1.1 – L'indexation 3D consiste à coder de manière compacte l'information visuelle relative à un modèle 3D afin d'optimiser les recherches dans les bases de données d'objets 3D.

est initialement défini au moyen d'une autre représentation, comme par exemple des surfaces NURBS provenant de logiciels de CAO, ou des nuages de points 3D obtenus avec un scanner laser, la représentation doit alors être convertie en un maillage triangulaire, qui est la représentation la plus simple, la plus commode et la plus répandue pour représenter les surfaces. Les maillages 3D sont stockés dans la base de données sous un format spécifique qui est usuellement le VRML (Virtual Reality Modeling Language), et la plupart des bases de données ne contiennent que des modèles à faible résolution des objets. Les bases de données d'objets 3D contenant des maillages 3D avec une carte de texture associée sont encore peu courantes, en particulier sur Internet [Esteban and Schmitt, 2004]. Le tableau 1.1 récapitule la composition des bases de données utilisées par les auteurs des articles cités dans les sections suivantes [Shilane et al., 2004]. La plupart des modèles de ces bases proviennent du site [3dcafe, 2005] et sont donc de faible résolution, avec un maillage de mauvaise qualité (“soupe de triangles”).

Base de données d'objets 3D	# modèles	# classes	# objets classés	Taille de la plus grande classe	Résolution des maillages
MPEG-7 [Zaharia and Prêteux, 2002]	1300	23	362	14%	faible
MRG [Hilaga et al., 2001]	230	32	230	15%	faible
CCCC [Vranic, 2003]	1841	54	416	13%	faible
Utrecht [Tangelder and Veltkamp, 2003]	684	6	512	47%	faible
PSB [Shilane et al., 2004]	6670	161	1814	6%	faible
SCULPTEUR & ENST [cf. Chapitre 4]	567	27	266	8%	haute

TAB. 1.1 – Bases de données d'objet 3D disponibles sur d'internet.

Les signatures des modèles 3D peuvent être efficacement exploitées pour diverses applications prometteuses, telles que la recherche par le contenu, la comparaison de modèles ou la classification, d'autant plus si elles vérifient l'ensemble des critères suivants :

**Compacité :** les signatures de petites tailles permettent d'effectuer des requêtes plus

rapides, et rendent la base de données plus transportables.

**Rapidité et robustesse des calculs :** la fabrication des vecteurs caractéristiques et le calcul de leur similarité ne doit pas être trop long, en particulier si la base de données est destinée à un grand public.

**Invariance aux transformations affines :** les objets doivent pouvoir être retrouvés quelque soit leur position, orientation ou taille (invariance aux rotation, translation et facteur d'échelle).

**Invariance de la connectivité du maillage :** les signatures de modèles de formes similaires doivent rester très proches et indépendant de la connectivité du maillage.

**Robustesse face aux petites perturbations du maillage :** le descripteur ne doit pas être trop sensible aux bruits (déformations, décimations) intervenant sur la position des sommets ou des arêtes du maillage.

**Pertinence de l'information codée :** la caractérisation doit être suffisamment efficace pour que les objets puissent être bien retrouvés.

Nous présentons dans la suite de ce chapitre les principales techniques d'indexation 3D. En particulier, nous analysons leur efficacité relativement aux besoins. En effet, une communauté non scientifique peut être plus intéressées par une recherche rapide basée sur la forme globale des objets plutôt que par des détails. L'utilisation de descripteurs globaux, semblent adaptés à ce type de requête. Ils permettent d'obtenir des résultats très rapidement mais avec une précision restant assez grossière. D'un autre côté, des scientifiques peuvent accorder plus d'attention sur des propriétés particulières de la surface des objets et dans ce cas une approche basée sur des descripteurs locaux semble plus adaptée. Pour les applications plus scientifiques la rapidité n'est pas forcément le critère discriminant, la précision des résultats étant souvent la plus désirée. Il y a donc un compromis à trouver entre rapidité et précision qui dépend beaucoup du contexte et de l'application visée. Pour ce compromis, l'utilisation de descripteurs multi-échelles semblent particulièrement bien adaptée et a constitué un de nos axes principaux de recherche pour l'indexation 3D.

Le travail bibliographique nous a permis de constater que le processus de recherche d'un objet 3D se compose généralement de trois phases qui sont la normalisation, l'extraction des *vecteurs caractéristiques*, et la mesure de leur similarité :

1. La normalisation consiste à orienter les objets dans un repère canonique. Cette étape permet d'apporter aux descripteurs l'invariance en rotation, en translation, et au facteur d'échelle de manière à ce que les objets puissent être considérés indépendamment de leur taille et de leur position dans l'espace. La technique d'alignement spatiale la plus répandue est l'ACP (Analyse par Composantes Principales). Elle est décrite dans la section 1.2.
2. L'extraction des *vecteurs caractéristiques* est ensuite effectuée au moyen de *descripteurs de forme* dans lesquels on distingue les méthodes globales (cf. Section 1.3), les méthodes locales (cf. Section 1.4), les décompositions spatiales (cf.

Section 1.5), les partitions de l'espace (cf. Section 1.6), les approches par graphes (cf. Section 1.7) et les approches 2D (cf. Section 1.9).

3. Le calcul de similarité de deux modèles est obtenu en mesurant la distance entre leurs *vecteurs caractéristiques*. Différentes distances ont été définies (cf. Section 1.8) : les normes de Minkowski  $L_N$ , la distance de Bhattacharyya, la distance de Earth Mover, la divergence de Kullback-Leibler, la divergence du  $\chi^2$ , la distance de Hausdorff, et les intersections d'histogrammes.

Les sections suivantes présentent l'ACP, différents *descripteurs de forme*, et différentes mesures de similarité, en prenant soin de mettre en avant leurs avantages et faiblesses. Un bilan récapitulatif est présenté à la fin de ce chapitre.

## 1.2 Normalisation des données

L'Analyse en Composantes Principales (ACP) est aussi connue sous le nom de transformée de Karhunen-Loeve ou de transformée d'Hotelling. La technique permet de définir un repère de coordonnées canoniques pour chaque objet, de manière à déterminer une pose spécifique normalisée. Une transformation rigide est alors appliquée à la représentation de l'objet pour le décrire dans son repère canonique.

Si on considère un maillage de triangles, le principe se base sur le calcul des valeurs propres et des vecteurs propres de la matrice de covariance calculée avec les centres de gravité des triangles. Le barycentre de l'objet devient alors le centre de l'espace 3D, et les directions  $P_1$ ,  $P_2$  et  $P_3$  du repère canonique correspondent aux directions des vecteurs propres de la matrice de covariance, et donc aux axes principaux de l'ellipsoïde d'inertie (cf. Figure 1.2).

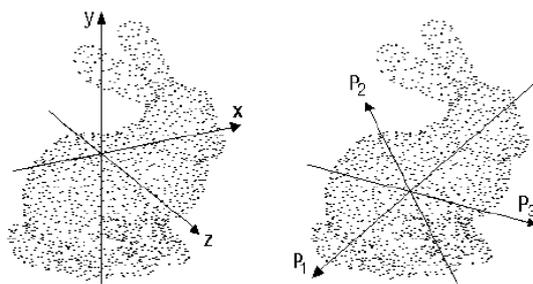


FIG. 1.2 – L'Analyse en Composantes Principales permet d'obtenir un repère canonique lié à l'objet.

Une importante limitation de l'ACP réside dans son intolérance au changement de résolution du maillage pour un modèle donné. Plusieurs variantes de l'ACP ont été développées pour pallier cette limitation : les auteurs de [Zaharia and Prêteux, 2002] ont proposé de pondérer les triangles par leur surface, et des travaux plus poussés de [Vranic et al., 2001] ont permis d'aboutir à l'“ACP continue” (Continuous PCA ou

CPCA) qui donne de meilleurs résultats concernant la pondération des triangles.

La CPCA a été appliquée dans le cadre de nos travaux comme étape de pré-traitement afin de normaliser la pose des objets. Soit  $I = \{v\}$  l'ensemble des points des triangles du maillage de l'objet 3D, et  $S$  la surface totale de l'objet :

- L'*invariance par translation* est obtenue en translatant le centre de gravité  $\mathbf{G}$  de l'objet à l'origine du repère. On obtient alors  $I_1 = I - \mathbf{G} = \{\mathbf{u} | \mathbf{u} = \mathbf{v} - \mathbf{G}, \mathbf{v} \in I\}$ .
- L'*invariance par rotation* est obtenue en calculant la matrice de covariance  $M = \frac{1}{S} \int_{I_1} \mathbf{v} \cdot \mathbf{v}^T dv$ .  $M$  est réelle symétrique. Ses valeurs propres sont donc réelles. Elles sont classées par ordre décroissant. On forme alors la matrice de rotation  $R$ , dont les colonnes sont les vecteurs propres unitaires classés selon l'ordre des valeurs propres associées, et on obtient par rotation de  $I_1$ , l'ensemble des points  $I_2 = R \cdot I_1 = \{\mathbf{w} | \mathbf{w} = R \cdot \mathbf{u}, \mathbf{u} \in I_1\}$ .
- L'*invariance par réflexion* est obtenue en multipliant l'ensemble des points  $\mathbf{w} = (w_x, w_y, w_z)$  de  $I_2$  par la matrice diagonale  $D = \text{diag}(\text{sign}(d_x), \text{sign}(d_y), \text{sign}(d_z))$  où  $d_x = \frac{1}{S} \int_{I_2} \text{sign}(w_x) w_x^2 dw$  (et similairement pour  $d_y$  et  $d_z$ ). Dans la pratique, nous déterminons le signe de  $d_z$  par produit vectoriel des deux premiers axes afin d'obtenir un repère direct.
- L'*invariance à l'échelle* est assurée en effectuant le produit des points  $I_2$  par l'inverse de  $s = \left( (s_x^2 + s_y^2 + s_z^2) / 3 \right)^{1/2}$ , où  $s_x, s_y$  et  $s_z$  représentent les distances moyennes des points  $\mathbf{w} \in I_2$  aux hyperplans de coordonnées  $(yz), (xz)$  et  $(xy)$  respectivement, et avec  $s_x = \frac{1}{S} \int_{I_2} |w_x| dw$  (et similairement pour  $s_y$  et  $s_z$ ).

On obtient finalement la transformation affine  $\tau$  définie par :

$$\tau(\mathbf{v}) = s^{-1} \cdot D \cdot R \cdot (\mathbf{v} - \mathbf{G}), \mathbf{v} \in I. \quad (1.1)$$

L'ACP est la méthode la plus répandue pour obtenir une pose normalisée des objets 3D. La méthode est simple, rapide et peut être appliquée de manière automatique à une base d'objets 3D.

Dans [Kazhdan et al., 2004] les auteurs proposent d'appliquer une remise à l'échelle anisotrope en transformant  $I_1$  par  $M^{-1/2}$ . Le calcul est ensuite réitéré plusieurs fois afin d'obtenir un modèle isotropique dont la variance est alors égale à 1 dans toutes les directions.

Cependant le choix des axes d'orientation reste un problème délicat. En effet, certains objets ont une position naturelle qui n'est pas forcément celle donnée par l'ACP. Par exemple dans la figure 1.3 la position debout d'un modèle humain ou d'un animal ou la position au sol d'un avion privilégie les directions naturelles bas-haut, avant-arrière, gauche-droite. On devra donc porter une attention particulière à cette phase de pré-traitement lorsqu'elle est requise par les descripteurs de forme. Les descripteurs de forme géométriquement invariants ont donc l'avantage d'éviter ce problème.

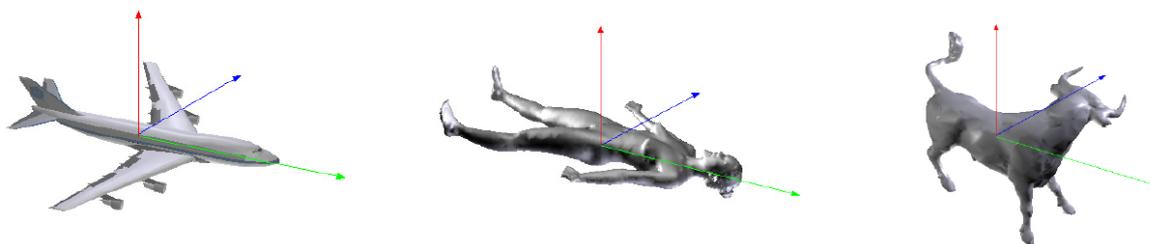


FIG. 1.3 – Certains modèles possèdent une position naturelle et ne requièrent pas de rotation d'axes.

## 1.3 Méthodes globales

Un descripteur global tient compte des caractéristiques géométriques globales d'un objet. Cette section présente les principales méthodes globales existantes et leurs avantages et faiblesses respectifs.

### 1.3.1 Aire et volume

L'aire, le volume, et les rapports du type aire/volume peuvent être utilisés comme vecteurs caractéristiques (ou métadonnées vu la concision de l'information) pour décrire les objets [Zhang and Chen, 2001]. Ces informations ne sont pas assez discriminantes pour effectuer des recherches précises basées sur la forme des objets mais peuvent être de précieuses informations pour les utilisateurs, en particulier si une échelle peut leur être associée (aire et volume). On pourrait cependant utiliser ces informations pour effectuer un élagage rapide des données afin d'accélérer une requête 3D spécifique.

L'aire et le volume sont rapides à calculer mais nécessitent un maillage de bonne qualité : le maillage de la surface du modèle 3D doit être une variété <sup>1</sup> et pour le calcul du volume, la surface doit être aussi fermée et les normales orientées.

### 1.3.2 Histogrammes de cordes

Le CNRC (Canadian National Research Council) d'Ottawa est le premier centre de recherche à avoir proposé une méthode d'indexation 3D, publiée en 1997 [Paquet and Rioux, 1997]. Leur base de données est constituée d'un grand nombre d'objets digitalisés avec les scanners laser 3D qu'ils ont développés. Un des modèles est un scanner laser blanc composé de trois raies dans les parties rouge, verte et bleue du spectre. Il permet d'acquérir simultanément la forme 3D et la couleur d'un objet.

L'approche d'indexation 3D proposée repose sur des statistiques sur les cordes de l'objets. Les cordes étant définies ici comme les segments reliant le centre de gravité de l'objet aux centres de chaque triangle du maillage. (Il aurait peut-être été plus

---

<sup>1</sup>Chaque arête a au plus deux faces adjacentes. (En anglais : manifold.)

approprié de dénommer ces segments comme des rayons plutôt que comme des cercles.)  
Trois histogrammes de cordes sont donc construits :

- un histogramme des longueurs des cordes
- un histogramme des angles entre les cordes et le 1<sup>er</sup> axe principal
- un histogramme des angles entre les cordes et le 2<sup>e</sup> axe principal

Un alignement spatial ACP est donc requis comme phase de pré-traitement.

Dans [Paquet and Rioux, 1997], des histogrammes caractérisant la couleur RVB sont aussi construits pour permettre des recherches utilisant des cartes de texture associées aux objets.

Dans le cadre du projet Nefertiti, un moteur de recherche d'objets 3D de forme similaire a été développé. Il utilise les objets de leur base de données et les requêtes sont basées sur les différents critères 3D et de couleur (cf. Figure 1.4).

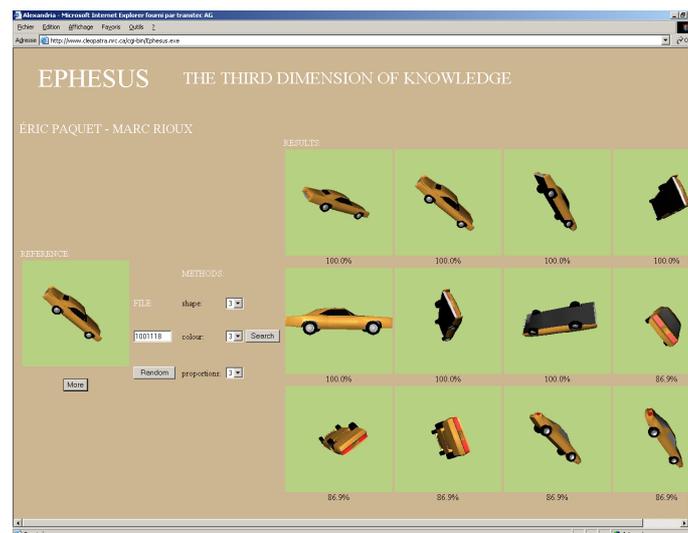


FIG. 1.4 – Moteur de recherche 3D du CNRC d'Ottawa (<http://www.cleopatra.nrc.ca>).

Les vecteurs caractéristiques associés aux modèles 3D sont donc des histogrammes de cordes et de couleurs, les caractéristiques des cordes étant calculées sur tout le maillage. Selon les auteurs, la taille optimale des histogrammes seraient de 64 intervalles. La technique n'est pas très discriminante pour des objets de forme complexe, et n'est pas robuste vis-à-vis de perturbations sur la connectivité du maillage. Par contre, cette approche a l'avantage d'être rapide à exécuter, discriminante pour les formes simples, et les informations sur les cordes (longueurs maximum, minimum, etc.) peuvent être utiles pour les utilisateurs.

### 1.3.3 Distribution de forme

Le groupe de recherche “Shape Retrieval and Analysis” de l’université de Princeton, USA, est l’équipe qui a le plus contribué à l’indexation 3D en terme de quantité de descripteurs de forme 3D. Leur descripteur le plus connu est la distribution de forme D2 (D2 Shape distribution) [Osada et al., 2002]. Une distribution de forme représente une distribution de probabilité d’une fonction de forme mesurant des propriétés géométriques globales de l’objet. La distribution de forme D2 d’un objet 3D représente, sous la forme d’un histogramme normalisé, les probabilités d’occurrence de distance (Euclidienne) entre des paires de points choisis aléatoirement sur la surface de l’objet (cf. Figure 1.5).

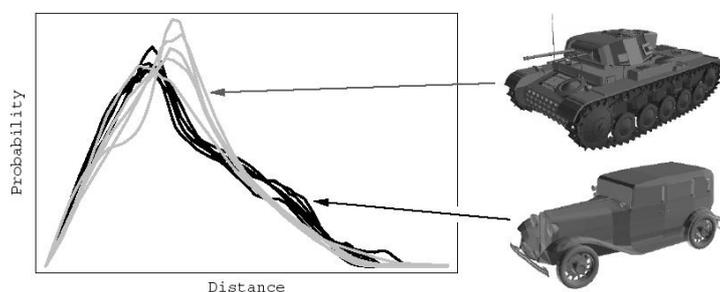


FIG. 1.5 – *Distribution de forme D2 de différents véhicules [Osada et al., 2002].*

Les auteurs préconisent des signatures de 64 entiers, et la norme de Minkowski  $L1$  pour calculer les distances entre signatures. L’approche proposée est invariante aux transformations euclidiennes, et est robuste aux déformations de la connectivité du maillage.

De plus, les auteurs proposent un outil de recherche 2D-3D (cf Figure 1.6) où la requête se fait par un graphique 2D dessiné par l’utilisateur et censé correspondre à une esquisse de l’objet recherché.

L’efficacité de cette interface reste cependant à être prouvée, tant il est difficile pour un utilisateur de dessiner des croquis à main levée.

La distribution de forme D2 proposée par l’université de Princeton est donc une méthode probabiliste, dont les principaux avantages sont la facilité d’implémentation, la rapidité de calculs, l’invariance aux transformations géométriques et la robustesse aux perturbations du maillage (connectivité, décimation). Les signatures caractérisent la forme globale des objets mais non les détails. La méthode semble donc être plus adaptée aux recherches d’objets similaires dans des bases de données d’objets de formes très différentes.

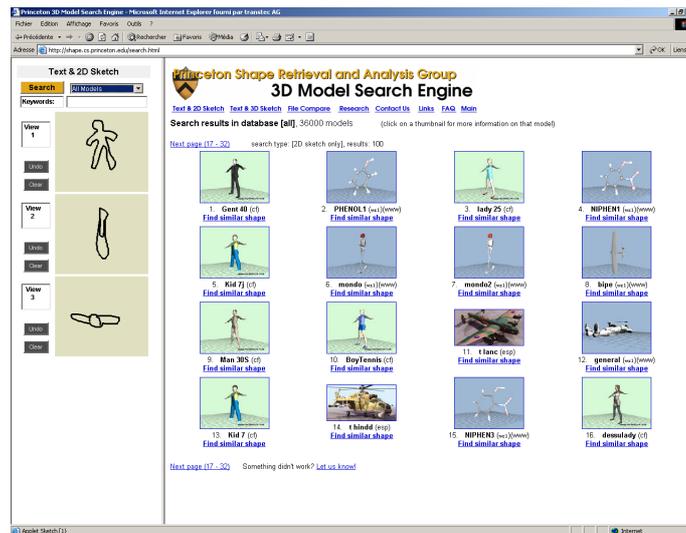


FIG. 1.6 – Moteur de recherche 2D-3D de l'université de Princeton (<http://shape.cs.princeton.edu/search.html>).

### 1.3.4 Descripteur de Hough 3D

Le descripteur de Hough 3D (DH3D) a été proposé par Zaharia et Prêteux [Zaharia and Prêteux, 2002] pour succéder au descripteur de spectre de forme 3D qu'ils ont développé pour MPEG-7 (cf. les descripteurs locaux, en section 1.4.1). La transformée de Hough 3D d'un maillage triangulaire consiste à accumuler les paramètres représentatifs des plans définis par les triangles du maillage de l'objet. En coordonnées sphériques un plan est décrit par le triplet  $(r, \theta, \phi)$ , où  $r \geq 0$  est la distance de  $\Pi$  à l'origine,  $\theta \in [0, 2\pi[$  est l'angle d'azimut et  $\phi \in [-\pi/2, \pi/2[$  l'angle d'élévation ;  $(\theta, \phi)$  exprimant la direction de la normale du plan  $\Pi$  (cf. Figure 1.7).

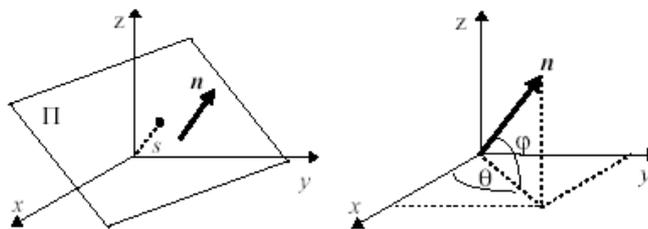


FIG. 1.7 – Passage en coordonnées sphériques pour la transformée de Hough. Un plan est décrit par 3 paramètres  $(r, \theta, \phi)$ .

Chacun des axes du repère sphérique est généralement uniformément échantillonné. On crée alors un histogramme 3D pour les différents triplets discrétisés  $(r, \theta, \phi)$  où chaque triangle apportera une contribution proportionnelle à son aire. L'accumulation

des contributions dans l’histogramme 3D permettra de caractériser la surface d’un objet 3D.

Pour pallier au problème d’invariance aux translations et rotations, une ACP doit être appliquée en premier lieu. Pour s’affranchir du problème de sur-représentation des triplets situés au voisinage des pôles produit par un échantillonnage uniforme des coordonnées sphériques, les auteurs proposent de projeter les normales des faces sur un octèdre régulier subdivisé et rétroprojeté sur la sphère unité.

Le DH3D s’est révélé performant sur la base de donnée MPEG-7, et a la propriété d’être stable quant à d’éventuelles variations de connectivité du maillage. Les auteurs utilisent deux niveaux de décomposition pour l’octaèdre (soit 128 facettes) et  $N_s = 20$  ( $N_s$  étant le nombre d’intervalles pour quantifier  $s$ ). Le descripteur est alors composé de 2560 éléments. Afin de réduire la taille du descripteur les mêmes auteurs [Zaharia and Prêteux, 2004] proposent une nouvelle version du DH3D avec quantification vectorielle, le DH3DO-QV, qui retourne sensiblement les mêmes résultats que leur précédent descripteur mais avec des vecteurs caractéristiques de 128 éléments. Cette dernière version a été implémentée dans le cadre du projet RNRT SEMANTIC-3D (<http://www.semantic-3d.net>) dans le but d’indexer les modèles 3D de pièces de voiture de la base de données de RENAULT.

## 1.4 Méthodes locales

Un descripteur local tient compte des caractéristiques géométriques locales d’un objet, par opposition aux méthodes globales qui ont tendance à décrire de manière trop grossière la forme des objets. Cette section décrit les avantages et faiblesses des principales méthodes existantes.

### 1.4.1 Descripteur de spectre de forme 3D

Le descripteur de spectre de forme 3D (3D shape spectrum descriptor ou 3DSSD) proposé par Zaharia et Prêteux [Zaharia and Prêteux, 2002] a été adopté comme descripteur de forme pour le format standard MPEG-7. Il s’agit d’un descripteur caractérisant les courbures locales de la surface des objets 3D au moyen de l’indice de forme de Koenderink [Koenderink, 1990].

Soit  $\mathbf{p}$  un point sur une surface 3D régulière, et  $k_p^1$  et  $k_p^2$  les courbures principales de la surface associées à  $p$ . L’indice de forme  $I_p$  s’écrit :

$$I_p = \frac{1}{2} - \frac{1}{\pi} \arctan \frac{k_p^1 + k_p^2}{k_p^1 - k_p^2}, \text{ with } k_p^1 \geq k_p^2$$

$I_p$  est à valeur dans  $[0,1]$  et n’est pas défini pour des surfaces planes pour lesquelles on a :  $k_p^1 = k_p^2 = 0$  (cf. Figure 1.8).

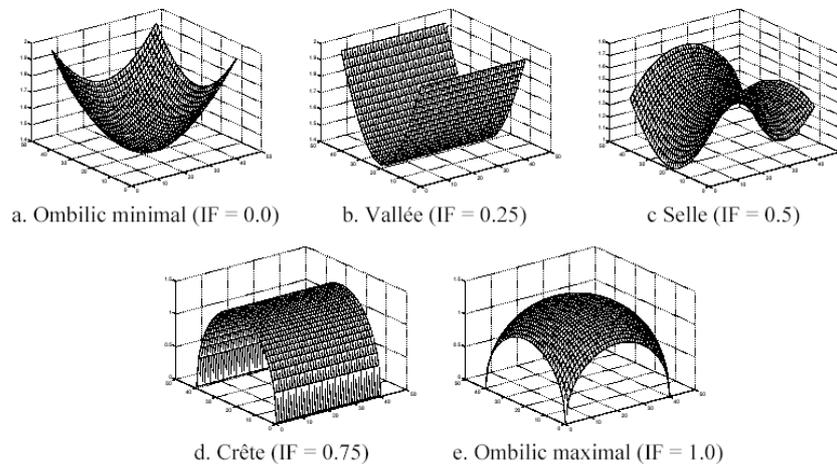


FIG. 1.8 – *L'indice de forme permet de caractériser localement la courbure de la surface [Zaharia and Prêteux, 2002].*

Les courbures principales sont estimées directement depuis le maillage triangulé, préalablement lissé avec un filtre passe-bas (par exemple de type Laplacien) pour éliminer le bruit sur la position des sommets du maillage. Les courbures principales sont estimées en calculant la surface polynômiale du second degré qui approxime localement le maillage 3D. En chaque sommet du maillage la normale à la surface polynomiale est égale à la somme des normales des triangles adjacents au sommet, les triangles étant pondérés par leur surface.

La méthode est donc invariante aux transformations rigides, mais est très dépendante de la connectivité du maillage, et très sensible à sa qualité (simplification, remaillage, etc). Parmi les autres approches visant à caractériser les objets 3D en estimant leurs courbures locales, on peut citer [Adán et al., 2003] qui déterminent les courbures à partir de cônes formés par les sommets et leurs N-couronnes<sup>2</sup> associées, et [Assfalg et al., 2003] qui utilisent une projection des courbures sur une carte 2D. Ces méthodes requièrent des maillages très réguliers avec des normales bien orientées. De plus une phase de pré-traitement est nécessaire pour garantir que le maillage est bien une variété et pour lisser la surface au moyen d'un filtre passe-bas.

Le Descripteur de Spectre de Forme 3D est donc sensible aux détails des objets et par conséquent la pertinence des signatures dépend grandement de la nature des modèles. Les études menées sur la base de données MPEG-7 ont révélés par exemple que le descripteur retrouvait mieux les humains dont la forme est plus marquée, que

<sup>2</sup>La 1-couronne d'un sommet  $v$  est l'ensemble des sommets du maillage directement connectés à  $v$  par une arête. La N-couronne de  $v$  est l'ensemble des sommets connectés à  $v$  par un chemin de longueur minimum contenant N arêtes.

les voitures qui sont plus compactes [Zaharia and Prêteux, 2002].

### 1.4.2 “Extended Gaussian Images”

Le principe des *Extended Gaussian Images (EGI)* est de projeter une fonction synthétisant des informations concernant le maillage 3D triangulé sur une sphère de Gauss partitionnée en facettes. Chaque triangle contribue à la facette correspondant à la direction de sa normale avec un poids égal à l’aire du triangle [Horn, 1984]. Variante des EGI, les *Complex Extended Gaussian Images (CEGI)* apportent comme contributions aux facettes, pour chaque triangle, un nombre complexe dont l’amplitude est égal à l’aire du triangle et la phase est égale à la distance du centre du triangle au centre de la sphère [Kang and Ikeuchi, 1994]. Cette représentation permet de discriminer les formes primitives. En outre, elle permet aussi d’obtenir diverses informations utiles comme les propriétés de symétrie ou la longueur des cordes.

Les inconvénients des EGI sont leur dépendance à la connectivité du maillage, les sur-représentations de l’information aux pôles dues à la discrétisation en coordonnées sphériques, la non-invariance aux transformations géométriques, et aussi le fait que la méthode est mal adaptée aux objets non homéomorphes à une sphère. Cependant, il est possible d’utiliser les faces d’un polyèdre régulier rétroprojeté sur la sphère de Gauss pour obtenir des partitions de l’espace de taille homogène. Dans la pratique, à l’instar du descripteur de Hough 3D (cf. Section 1.3.4), nous avons implémenté la décomposition sur un octaèdre subdivisé deux fois, soit  $8 * 4 * 4 = 128$  facettes (cf. Chapitre 4).

## 1.5 Approches par transformées

Les méthodes par transformées telles les descripteurs de Fourier ou les calculs de moments ont souvent été proposés dans la littérature comme descripteurs de forme 2D. L’aspect multi-résolution des vecteurs caractéristiques et le gain en détails lorsque l’on prend en compte un plus grand nombre de coefficients sont le principal avantage de ce type d’approche. Elle a été reprise dans plusieurs travaux sur l’indexation 3D : transformée de Fourier, harmoniques sphériques, calculs de moments, et transformée radiale angulaire.

### 1.5.1 Transformée de Fourier 3D

La transformée de Fourier 3D d’un objet tridimensionnel permet de caractériser ses informations dans l’espace des fréquences [Vranic and Saupe, 2001, Dutagaci and Sankur, 2005].

Vranic et Saupe, dans [Vranic and Saupe, 2001] propose d’utiliser la transformée de Fourier 3D sur des modèles 3D voxelisés (les auteurs proposent d’appliquer une 3D

DCT sur tous les voxels). Ils ont développé un moteur de recherche d'objet 3D utilisant une base de données composée de modèles d'MPEG-7 afin de tester leurs algorithmes (cf. Figure 1.9).

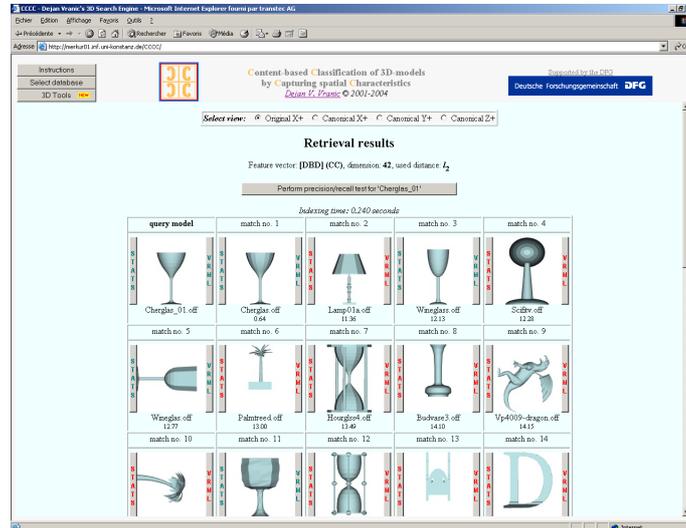


FIG. 1.9 – Moteur de recherche CCC développé par l'équipe de l'université de Leipzig (<http://merkur01.inf.uni-konstanz.de/CCC>).

Cependant, le problème majeur est inhérent à la transformée de Fourier elle-même. En effet l'espace d'application doit être continu, ce qui implique que les maillages doivent être réguliers, et donc les objets complexes ne peuvent pas être bien décrits. Aussi, les descripteurs de Fourier ne sont pas très discriminants lorsque le nombre de coefficients est trop faible, et de plus les coefficients des hautes fréquences sont très sensibles aux bruits ou aux petites variations de la connectivité du modèle 3D de l'objet, et à la finesse de la voxelisation dans le cas présenté ci-dessus. Par ailleurs, les descripteurs n'étant pas invariant aux transformations, cette approche requiert un alignement spatial par ACP.

## 1.5.2 Harmoniques Sphériques

Dans la continuité de leur travaux, l'équipe de Leipzig a également proposé d'appliquer une transformation de Fourier sur la sphère  $S^2$  [Saupe and Vranic, 2001] en appliquant les formules d'harmoniques sphériques proposées dans [Healy et al., 1998]. Puis, afin de s'affranchir du problème d'invariance aux rotations, l'équipe de Princeton proposa d'appliquer la décomposition en harmoniques sphériques de fonctions sphériques définies par l'intersection de la surface de l'objet 3D avec un ensemble de sphères concentriques [Kazhdan and Funkhouser, 2002], (cf. Figure 1.10).

La méthode des harmoniques sphériques de Princeton a donné de meilleurs résultats sur leur base de données que leur précédent descripteur (distribution de forme D2).

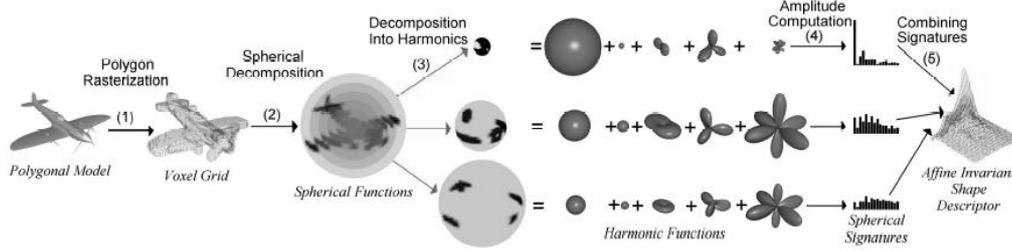


FIG. 1.10 – *Décomposition en harmoniques sphériques [Kazhdan and Funkhouser, 2002].*

Cependant, elle se base sur une voxelisation des modèles 3D et dépend donc du niveau de résolution de la voxelisation, entraînant ainsi une perte de détails dans la description de l'objet. C'est pourquoi [Vranic, 2003] proposent d'appliquer la méthode directement sur les maillages 3D avec de nouvelles fonctions sphériques. Les résultats qu'ils ont obtenus sur leur base de données avec leur méthode sont supérieurs à ceux basés sur le modèle voxelisé, et moins coûteux en temps. Cependant, ces résultats montrent aussi que l'information encodée ne permet pas vraiment d'effectuer des requêtes précises sur les formes, la limitation principale étant le nombre de sphères concentriques et le nombre de coefficients harmoniques qui restent peut-être trop faibles. Les auteurs choisissent en pratiques 32 sphères concentriques et 16 harmoniques par sphères, soit un descripteur de  $32 * 16 = 512$  coefficients.

### 1.5.3 Calculs de moments

Les calculs de moments dans le contexte de l'analyse des formes consistent à projeter une fonction  $f$  (de carré intégrable)  $\in \mathcal{L}^2$  définissant un objet 3D, sur une famille de fonctions complexes  $\Phi = \{\varphi_i\}_{i \in \mathbb{N}}$  dans un domaine  $\Omega$  :

$$\mu_i = \langle f, \varphi_i \rangle = \int_{\Omega} f(\mathbf{x}) \cdot \overline{\varphi_i(\mathbf{x})} d\mathbf{x}.$$

Les coefficients  $\mu_i$  permettent de caractériser les objets à différents niveaux de précision selon l'ordre choisi pour la décomposition (i.e. le nombre de coefficients). Différentes représentations de moments existent (ayant différentes propriétés) et dépendent de la famille de fonctions  $\Phi$  utilisée pour la décomposition.

On a par exemple :

- les moments statistiques centrés 3D  $\mathcal{M}_{qrs}$  calculés sur la surface des objets, en chaque triangle [Paquet and Rioux, 1999] :

$$\mathcal{M}_{qrs} = \sum_{i=1}^n m_i (x_i - x_0)^q (y_i - y_0)^r (z_i - z_0)^s, \quad (1.2)$$

$n$  étant le nombre de triangles du maillage,  $(x_i, y_i, z_i)$  son barycentre,  $m_i$  le poids associé au triangle d'indice  $i$  et habituellement pris égal à son aire, et  $(x_0, y_0, z_0)$  le centre de l'objet,

- les moments géométriques d'ordre  $(q, r, s)$  avec une fonction sphérique  $r(\mathbf{u})$  introduits dans [Saupe and Vranic, 2001] (ray-moments) :

$$\mathcal{M}_{qrs} = \sum_{i,j=0}^{n-1} r(\mathbf{u}_{ij}) \Delta s_{ij} x_{ij}^q y_{ij}^r z_{ij}^s, \quad (1.3)$$

$r(\mathbf{u}_{ij})$  étant distance maximale entre la surface de l'objet et son centre, dans les  $n^2$  directions  $\mathbf{u}_{ij} = (x_{ij}, y_{ij}, z_{ij}) = (\cos\phi_i \sin\theta_j, \sin\phi_i \sin\theta_j, \cos\theta_j)$ ,  $i, j = 0, \dots, n-1$ , et  $\Delta s_{ij}$  l'aire de la surface sur la sphère correspondant au point  $\mathbf{u}_{ij}$  pour compenser l'échantillonnage non uniforme de la sphère,

- les moments de Zernike 3D, invariants à l'échelle, aux symétries, aux rotations et aux translations [Canterakis, 1999, Novotni and Klein, 2003].

### 1.5.4 Transformation Radiale Angulaire 3D

La Transformation Radiale Angulaire 3D (3D Angular Radial Transform, ou 3D ART), est une extension de la 2D ART qui a été proposée dans le cadre du standard MPEG-7 pour indexer les images à l'aide de descripteur de forme se basant sur les régions [Ricard et al., 2004]. Soient  $(\rho, \theta, \phi)$  les coordonnées sphériques des objets 3D, les coefficients 3D ART  $F_{nm_\theta m_\phi}$  d'ordre  $(n, m_\theta, m_\phi)$  s'écrivent :

$$F_{nm_\theta m_\phi} = \int_0^{2\pi} \int_0^\pi \int_0^1 V_{nm_\theta m_\phi}^*(\rho, \theta, \varphi) f(\rho, \theta, \varphi) d\rho d\theta d\varphi, \quad (1.4)$$

$f(\rho, \theta, \varphi)$  étant une fonction sphérique définissant la surface discrétisée de l'objet et  $V_{nm_\theta m_\phi}^*(\rho, \theta, \varphi)$  la base de fonction 3D ART définie dans [Ricard et al., 2004]. Les auteurs obtiennent leurs meilleurs résultats en discrétisant le modèles 3D avec une grille de  $64 * 64 * 64$  voxels et en choisissant  $n = 3$ ,  $m_\theta = 5$  et  $m_\phi = 5$ . Les descripteurs obtenus sont compacts mais non invariants aux transformations (les auteurs proposent l'ACP comme méthode d'alignement spatial). Les résultats expérimentaux obtenus par les auteurs sont, en performance, équivalents aux Harmoniques Sphériques de l'équipe de Princeton.

## 1.6 Partition de l'espace

### 1.6.1 Ensemble de points pondérés

Tangelder et Veltkamp proposent dans [Tangelder and Veltkamp, 2003] de partitionner régulièrement l'espace en  $25 * 25 * 25$  cellules de taille égale et d'attribuer un poids à chacune des cellules non vides de points. Trois méthodes sont proposées pour déterminer la valeur associée à une cellule (la valeur du point saillant) :

- la courbure Gaussienne des points la plus forte dans chaque secteur,

- la variation moyenne des normales,
- la position moyenne des points dans chaque secteur.

La taille des vecteurs caractéristiques dépend donc de la finesse de la partition de l'espace, et de la taille des données utilisées. De plus, les auteurs proposent une variante de la distance de Earth Mover décrite en section 1.8.4, la distance “Proportional Transportation”, pour évaluer les distances entre les vecteurs étant donné que ces distances sont bien adaptées aux ensembles de points pondérés.

Cependant l'utilisation de ces distances est très coûteuse en temps. Les descripteurs proposés restent peu discriminants (on perd la notion de connectivité entre les points), et la méthode nécessite que les objets aient une pose normalisée par ACP.

### 1.6.2 Contextes de forme 3D

L'idée présentée par les auteurs de [Körtgen et al., 2004] consiste à effectuer  $N$  tirages aléatoires de points auxquels on associe des distributions caractérisant la forme de l'objet (appelés contextes de forme 3D ou 3D Shape Context), à l'instar des distributions de forme de Princeton [Osada et al., 2002]. Les points sont choisis sur le modèle, uniformément répartis sur la surface, et un histogramme (appelé contexte de forme) des distances aux autres  $N - 1$  points leur est associé. Afin de conserver une notion de cohérence spatiale, les auteurs proposent de partitionner l'espace en  $S$  cellules selon trois types de décomposition classique (cf. Figure 1.11). Chaque point est localisé par la cellule dans laquelle il se trouve. Selon le type de représentation choisi (coquille, secteur ou combinée), il sera nécessaire d'appliquer ou non une normalisation de la pose de l'objet. Notons que la méthode s'inspire du modèle établi en 2D [Belongie et al., 2002], utilisé en reconnaissance de forme à partir d'images 2D. Chaque objet est décrit par  $N$  histogrammes de  $S$  cellules, avec  $N = 200$  et  $S = 288$  avec la représentation combinée ( $\theta * \varphi = 6 * 12$  secteurs<sup>3</sup> et 4 coquilles). Afin de réduire la grande taille des descripteurs les auteurs suggèrent d'utiliser une quantification vectorielle.

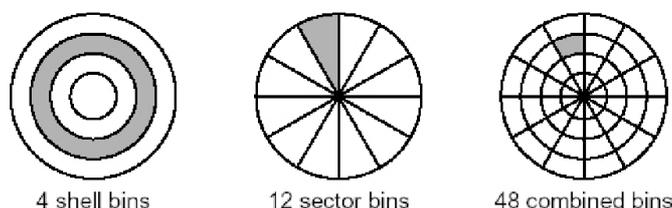


FIG. 1.11 – Décomposition classique de l'espace 2D en “coquilles”, “secteurs”, et “combinées” [Körtgen et al., 2004].

---

<sup>3</sup> $\theta$  et  $\varphi$  sont les angles en coordonnées sphériques.

### 1.6.3 Structures volumétriques élémentaires 3D

L'approche proposée par [Irani and Ware, 2000] consiste à décrire une forme en utilisant des structures volumétriques élémentaires 3D appelées géons (ou *geometric ions*). Les primitives volumétriques 3D utilisées sont alors des cylindres, cubes, parallélépipèdes, cônes (tronqués ou non) et ellipsoïdes. D'autres auteurs de [Zhou and Kambhamettu, 2002, Chevalier et al., 2003] utilisent quant à eux un jeu de superquadriques pour décrire la forme des objets 3D. Leur approche consiste dans un premier temps à segmenter l'objet en différentes régions, puis à associer à chaque région une forme élémentaire. Une structure similaire à celle d'un graphe permet ensuite de décrire la forme des objets de manière compacte.

Cependant, ces approches ne sont pas adaptées pour décrire de manière précise les objets de topologie complexe (par exemple une anse de vase ou une main). En effet, les différentes régions doivent avoir une forme proche des structures volumétriques élémentaires 3D pour que leur description soit pertinente.

## 1.7 Approches par graphes

Cette section présente les approches basées sur les représentations par graphe. Ces approches de haut-niveau permettent d'extraire des informations intrinsèques à la forme générale des objets 3D, et peuvent donc être de puissants outils de recherche ou comparaison de forme. On présentera rapidement la transformation par axe médian 3D, qui est une extension de la méthode la plus connue pour obtenir un squelette en 2D, puis des approches pour obtenir des graphes, et enfin les graphes de Reeb qui ont retenu notre attention par leur robustesse et efficacité, et dont nous allons présenter une étude détaillée au chapitre suivant.

### 1.7.1 Squelette par axe médian

L'axe médian d'un objet a été défini par Blum en 1967 comme étant l'ensemble des centres des disques maximaux inclus dans l'objet [Blum, 1967]. Un disque est maximal s'il n'est inclus dans aucun autre disque dans l'objet. On obtient donc un graphe linéaire en 2D (cf. Figure 1.12).

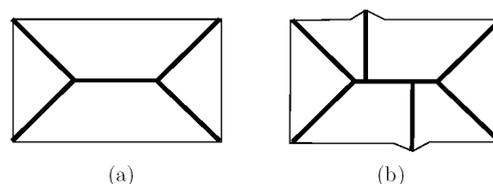


FIG. 1.12 – La transformée par axe médian produit un graphe linéaire en 2D. (a) et (b) montrent la sensibilité du graphe aux artefacts [Blum, 1967].

Chaque arc du graphe est médian, c'est-à-dire qu'il est situé au milieu d'une certaine partie de la forme et agit comme un axe de symétrie. Il existe différentes techniques pour obtenir un axe médian :

- la propagation du feu de prairie [M.Tanase and R.Veltkamp, 2003] qui produit un squelette analogue à celui par disques maximaux,
- l'amincissement homotopique [Lee et al., 1994],
- le diagramme de voronoï [Attali and Boissonnat, 2002],
- la carte des distances [L.Wade, 2000],
- transformation par axe médian 3D [Sherbrooke et al., 1996].

Bien que les squelettes soient une représentation compacte très riche, de haut-niveau d'abstraction et très utiles dans beaucoup d'applications, il reste que les principaux problèmes des approches par transformée par axe médian sont le manque de robustesse aux déformations locales, la complexité d'extraction du squelette, et la difficulté de coder des vecteurs caractéristiques. De plus en 3D, l'axe médian est plus complexe et peut contenir des surfaces. On l'appelle alors *surface médiane* [Sherbrooke et al., 1996].

### 1.7.2 Extraction de graphes

D'autres approches d'extraction de graphes permettent d'obtenir des représentations de haut niveau. Cependant elles restent coûteuses et très sensibles aux bruits de la surface des modèles. L'approche de [Verroust and Lazarus, 2000] permet d'extraire un graphe basé sur des ensembles de niveaux géodésiques. La représentation obtenue caractérise bien la forme des objets mais la construction du graphe est dépendante du choix d'un point source. Les graphes de shock de [Leymarie and Kimia, 2001] permettent d'obtenir une description assez riches en utilisant l'orientation des arêtes des graphes obtenus par la transformation par axe médian. Ils sont utilisés en reconnaissance de forme 2D, et leur extension en 3D consiste à utiliser des graphes de shock extraits de vues 2D de l'objet 3D à décrire. Leur approche repose donc sur des extractions d'axes médians qui peuvent être initialement très bruités. Ces approches ont été détaillées dans [Biasotti, 2004].

### 1.7.3 Graphe de Reeb

Les graphes de Reeb permettent d'obtenir une représentation de type squelette en conservant la structure topologique des objets [Reeb, 1946]. En effet, leurs constructions reposent sur la théorie de Morse [Shinagawa et al., 1991] qui permet de caractériser la topologie des surfaces fermées ; on retrouve donc autant de cycles dans les graphes que de trous dans les objets (cf. Figure 1.13).

Les graphes de Reeb sont utilisés dans divers domaines tels que la cartographie pour la modélisation de terrain [Takahashi et al., 1995], le médical pour la segmentation de scanner 3D de corps humain [Xiao et al., 2003], et la reconnaissance de formes 3D [Biasotti et al., 2003, Chen and OuhYoung, 2002].

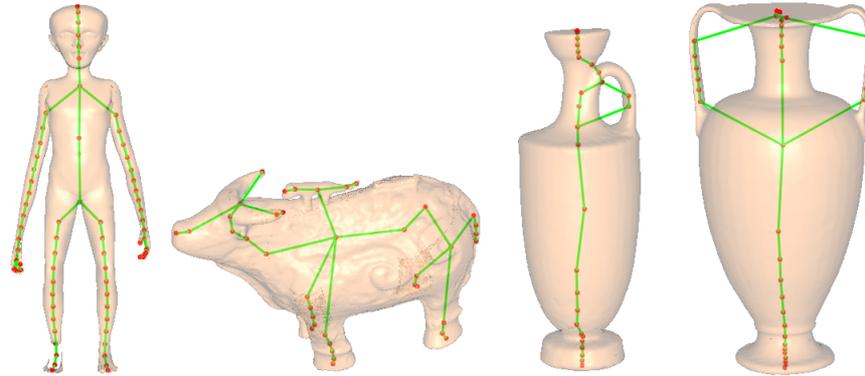


FIG. 1.13 – Modèles 3D et leur graphe de Reeb. Le genre de la surface de l'objet est retrouvé lorsque la profondeur du graphe devient suffisante.

Dans [Biasotti et al., 2003, Biasotti et al., 2000], une méthode d'appariement de graphe reposant sur la propagation de sous-graphes appariés est proposée. La méthode est de complexité quadratique et utilise des graphes de Reeb étendus (Extended Reeb Graph ou ERG) : les graphes de reeb sont orientés et possèdent des informations aux arêtes.

### Grphe de Reeb multirésolution

L'aspect multirésolution est intéressant car bien adapté aux recherches dans des grandes bases de données. En effet, on peut se limiter à des comparaisons à faible niveaux de résolutions afin d'obtenir des résultats rapides, ou utiliser la résolution maximale pour obtenir des résultats plus fins.

L'étude de la méthode d'appariement de graphes de Reeb multirésolution introduit par [Hilaga et al., 2001] nous a permis de conclure quant à la richesse et la pertinence des informations capturées par cette méthode sophistiquée que nous avons décidé d'étendre, et que nous détaillons dans le chapitre suivant.

Les travaux de [Chen and OuhYoung, 2002] sur une base de données de 445 objets 3D provenant d'Internet en utilisant la méthode de [Hilaga et al., 2001] permettent de juger les performances des MRG, mais aussi les limitations. En effet, les graphes ne sont pertinents que si les maillages sont "propres" et donc une phase de pré-traitement pour rééchantillonner les maillages peut être nécessaire.

## 1.8 Mesure de similarité

Les *descripteurs de forme* précédemment présentés permettent d'extraire des *vecteurs caractéristiques* qui sont introduits dans les bases de données en tant que clés pour représenter les modèles 3D et ainsi accélérer les recherches (car plus compacts que

les modèles 3D eux-mêmes). Afin de mesurer la similarité entre les modèles 3D, il est nécessaire de définir des fonctions de coût qui, appliquées à ces modèles retourneront un score quantifiant la similarité.

Certains vecteurs caractéristiques requièrent leur propre fonction de similarité, mais la plupart sont des histogrammes, ou des distributions de probabilité, et il est donc commode de leur appliquer des outils classiques de statistique pour les comparer.

Cette section présente les différentes techniques de calcul de distance pouvant être appliquées aux vecteurs caractéristiques des objets 3D suivant leur type.

### 1.8.1 Définitions

La mesure du degré de ressemblance de deux objets peut-être obtenue en évaluant avec une fonction de distance  $d$  la similarité entre leurs deux vecteurs caractéristiques. Soit  $S$  un ensemble de points et sa fonction de distance associée  $d$  (aussi appelée une *métrie*) telle que  $d : S \times S \rightarrow \mathbb{R}^+$ . Selon les applications,  $d$  peut vérifier les propriétés suivantes [Copson, 1968] :

1. *Identité* :  $\forall x \in S, d(x, x) = 0$ ,
2. *Positivité* :  $\forall (x, y) \in S \times S / x \neq y, d(x, y) > 0$ ,
3. *Symétrie* :  $\forall (x, y) \in S \times S, d(x, y) = d(y, x)$ ,
4. *Inégalité triangulaire* :  $\forall (x, y, z) \in S \times S \times S, d(x, z) \leq d(x, y) + d(y, z)$ ,

Par définition,  $d$  est une métrique si les propriétés 1 à 4 sont respectées,  $d$  est une semi-métrique lorsque seulement les propriétés 1 à 3 sont respectées, et  $d$  est une pseudo-métrique lorsque seulement les propriétés 1,3 et 4 sont respectées.

La distance entre deux points est donc positive et la distance de  $x$  à  $y$  est égale à la distance de  $y$  à  $x$ . L'inégalité triangulaire signifie que la distance à parcourir pour aller du point  $x$  au point  $z$  est inférieure à la distance de  $x$  à  $y$ , puis de  $y$  à  $z$ .

Selon [Barros et al., 1996] l'inégalité triangulaire est une propriété qui améliore les recherches basées sur des requêtes de similarité de forme dans les bases de données.

### 1.8.2 Normes de Minkowski

Elles sont aussi connues sous le nom de normes  $L_N$ . Ce sont des normes vectorielles, les plus répandues étant les normes  $L_1$ ,  $L_2$  et  $L_\infty$  :

$$\|q - v\|^N = \left( \sum_i |q_i - v_i|^N \right)^{\frac{1}{N}} \quad (1.5)$$

Ces normes sont très populaires car rapides à calculer et simples à implémenter. Elles sont donc utilisées dans presque toutes les approches précédemment présentées.

Les normes des fonctions de densité de probabilité (*FDP*) et les fonctions de densité cumulée (*FDC*) ont été comparées dans [Osada et al., 2002].

La norme  $L_N$  de la *FDP* est définie par

$$D(q, v) = \left( \int |q - v|^N \right)^{\frac{1}{N}}, \quad (1.6)$$

et la norme  $L_N$  de la *FDC* est définie par

$$D(q, v) = \left( \int |\hat{q} - \hat{v}|^N \right)^{\frac{1}{N}}, \quad (1.7)$$

où  $N$  est le degré de la norme,  $q$  et  $v$  sont les *FDP*, et  $\hat{q}$  et  $\hat{v}$  sont les *FDC* correspondantes, avec

$$\hat{q}(x) = \int_{-\infty}^x q. \quad (1.8)$$

On note que les normes de plus hauts degrés caractérisent seulement les grandes différences.

Les normes  $L_1$  et  $L_\infty$  de la *FDC* sont aussi appelées distance de Match et distance de Kolmogorov-Smirnov (KS) respectivement.

### 1.8.3 Distance de Bhattacharyya

La distance de Bhattacharyya  $D$  peut être utilisée pour comparée la similarité entre deux histogrammes  $Q$  et  $V$  [Thacker et al., 1997] :

$$D = 1 - \sum_i \sqrt{Q_i} \sqrt{V_i}. \quad (1.9)$$

Lorsque les données sont identiques et normalisées, on obtient  $D = 1 - \sum_i Q_i = 0$ .

### 1.8.4 Distance de Earth Mover

La distance de Earth Mover (Earth Mover's Distance ou EMD) est la quantité de travail minimale pour changer une signature en une autre. La notion de "travail" repose sur : 1) la quantité du contenu de la signature à transporter d'une composante à une autre, et 2) la distance Euclidienne choisie par l'utilisateur pour mesurer la distance entre deux composantes. Le nombre de composantes de deux signatures à comparer peut être différent, ainsi que la somme totale des composantes des deux

signatures.

Soient deux signatures à comparer  $P = \{(p_1, w_1), \dots, (p_m, w_m)\}$  et  $Q = \{(q_1, u_1), \dots, (q_n, u_n)\}$ ,  $P$  contenant  $m$  composantes  $p_i$  de poids  $w_i$ , et  $Q$  contenant  $n$  composantes  $q_j$  de poids  $u_j$ . Selon [Rubner et al., 1998], la première étape consiste à trouver l'ensemble des portions de contenu  $f_{ij}$  de la signature à transporter de la composante  $i$  vers la composante  $j$  qui minimisent le coût (travail) suivant :

$$\sum_{i=1}^m \sum_{j=1}^n d_{ij} f_{ij} \quad (1.10)$$

où  $d_{ij}$  est la distance entre les deux composantes  $i$  et  $j$  avec les contraintes suivantes :

$$\begin{aligned} f_{ij} &\geq 0 & 1 \leq i \leq m, 1 \leq j \leq n \\ \sum_{j=1}^n f_{ij} &\leq w_i & 1 \leq i \leq m \\ \sum_{i=1}^m f_{ij} &\leq u_j & 1 \leq j \leq n \\ \sum_{i=1}^m \sum_{j=1}^n f_{ij} &= \min(\sum_{i=1}^m w_i, \sum_{j=1}^n u_j) \end{aligned}$$

La première contrainte n'autorise que des mouvements des composantes de  $P$  vers  $Q$ . Les deux contraintes suivantes limitent la quantité de composantes déplacées de  $P$ , et la quantité de composantes reçu par  $Q$  à leur poids respectifs. La dernière contrainte implique le maximum de déplacement de composantes possible.

La distance EMD est alors définie comme

$$EMD(q, v) = \frac{\sum_{i=1}^m \sum_{j=1}^n d_{ij} f_{ij}}{\sum_{i=1}^m \sum_{j=1}^n f_{ij}}, \quad (1.11)$$

La distance EMD est normalisée au minimum des sommes des poids des signatures afin d'éviter de favoriser les signatures plus petites lors de comparaisons de signatures de tailles différentes.

### 1.8.5 La divergence de Kullback-Leibler

La divergence de Kullback-Leibler est une mesure souvent utilisée en théorie de l'information. Soit  $Q = \{q_i\}$  et  $V = \{v_i\}$  deux jeux de données, on a :

$$KL(Q|V) = \sum_i q_i \ln \frac{q_i}{v_i} \quad (1.12)$$

Dans sa forme originale,  $KL(Q|V)$  n'est pas égal à  $KL(V|Q)$ . Des versions symétriques ont été proposées par [Jeffreys, 1946] et [Hetzl et al., 2001].

Version symétrique de [Jeffreys, 1946] :

$$KL_j(Q|V) = \frac{KL(Q|V) + KL(V|Q)}{2}. \quad (1.13)$$

Version symétrique de [Hetzl et al., 2001] :

$$KL_h(Q|V) = \sum_i (q_i - v_i) \ln \frac{q_i}{v_i}. \quad (1.14)$$

### 1.8.6 Divergence du $\chi^2$

Le test du  $\chi^2$  est une méthode formelle statistique pour obtenir la dissimilarité entre deux distributions de probabilité. La divergence  $\chi^2$  est utilisée dans le test du  $\chi^2$  pour déterminer si les jeux de données proviennent de la même distribution ou non.

La divergence du  $\chi^2$  est employée pour comparer deux histogrammes  $Q$  et  $V$  [Hetzl et al., 2001] :

$$\chi^2(Q, V) = \sum_i \frac{(q_i - v_i)^2}{q_i + v_i} \quad (1.15)$$

### 1.8.7 Distance de Hausdorff

La distance de Hausdorff est la “distance maximale entre un ensemble de points et le point le plus proche d’un autre ensemble de points” :

$$h(Q, V) = \max_{q \in Q} \{ \min_{v \in V} \{ d(q, v) \} \} \quad (1.16)$$

où  $d(q, v)$  est la distance entre les points  $q$  et  $v$ .

### 1.8.8 Intersection d’histogrammes

L’intersection d’histogramme est une technique qui a été introduite pour l’indexation d’histogrammes couleur [Swain and Ballard, 1990]. Soient deux histogrammes couleur  $I = \{I_j\}$  (image) et  $M = \{M_j\}$  (modèle), de tailles  $N$ , l’intersection d’histogrammes  $H(I, M)$  est définie par :

$$H(I, M) = \sum_{j=1}^N \min(I_j, M_j) \quad (1.17)$$

Le résultat représente la quantité de pixels du modèle  $M$  ayant des pixels correspondants de la même couleur dans l’image  $I$ .

Afin d’obtenir une valeur normalisée entre 0 et 1, on norme l’histogramme par le nombre de pixels du modèle :

$$H_N(I, M) = \frac{\sum_{j=1}^N \min(I_j, M_j)}{\sum_{j=1}^N M_j} \quad (1.18)$$

En reconnaissance des formes, on a adapté l’approche par intersection d’histogrammes à d’autres types de signatures.

De plus, si les histogrammes sont de même dimension  $\sum_{i=j}^N M_j = \sum_{i=j}^N I_j = T$ , on obtient <sup>4</sup> :

$$1 - H_N(I, M) = 1 - \frac{\sum_{j=1}^N \min(I_j, M_j)}{T} = \frac{1}{2T} \sum_{j=1}^N |I_j - M_j|, \quad (1.19)$$

la fonction  $1 - H_N(I, M)$  est alors la distance métrique  $L_1$ .

## 1.9 Approches 2D

Dans certains cas de figures il peut s'avérer intéressant d'utiliser des approches spécifiques à la recherche par le contenu d'image 2D à des fins de recherche 3D. Par exemple :

1. Les modèles 3D peuvent être accompagnés d'images 2D dont on peut extraire des profils de l'objet, et dans ce cas les techniques d'appariement de forme 2D peuvent être utilisées.
2. Des cartes de texture 2D associées à la surface des objets peuvent être stockées avec les modèles 3D. Bien qu'il peut y avoir des problèmes de déformation géométrique au sein de l'image de texture, il est possible d'utiliser des techniques 2D de comparaison de texture pour retrouver des modèles 3D texturés.
3. En considérant les vidéos au format Quicktime <sup>5</sup> comme des représentations 3D, on peut appliquer des algorithmes de recherche d'image 2D sur ces vidéos. En prenant comme exemple des vidéos d'objets en rotation sur une table tournante, une image requête d'une simple vue 2D d'un objet pourrait être utilisée pour retrouver la vidéo Quicktime ayant l'image la plus proche de la requête.
4. Il est aussi possible de générer de multiples vues 2D d'un modèle 3D et d'effectuer des recherches sur ces images 2D pour retrouver l'objet 3D correspondant, de la même façon que pour l'approche avec les images 2D extraites des vidéos Quicktime [Chen et al., 2003, Ansary et al., 2004].

Les algorithmes 2D qui peuvent être utilisés pour ces applications de recherche par le contenu sont trop nombreux pour être exposés ici. On distingue cependant les approches basées sur la couleur, la texture, et la forme. On peut trouver dans la littérature des articles récents qui sont de bons états de l'art du domaine [Smeulders et al., 2000, Eakins and Graham, 1999].

Les cartes de texture des modèles 3D sont des images 2D représentant des informations d'apparence qui sont à rétroprojeter sur la surface des modèles 3D. Ces cartes étant définies dans un espace rectangulaire 2D, elles présentent de nombreuses distorsions géométriques locales dues aux projections 3D-2D.

---

<sup>4</sup>On se sert de la relation  $\min(a, b) = \frac{|a+b|}{2} - \frac{|a-b|}{2}$

<sup>5</sup><http://www.apple.com/quicktime>

En particulier les couleurs ou textures de modèles similaires peuvent être déformées de différentes manières dans l'espace 2D. Il en résulte que les algorithmes de comparaison basé sur la texture ne pourront pas fournir de très bons résultats, s'ils ne tiennent compte que des cartes de texture géométriquement déformées.

Cependant il est possible d'éviter ce problème en utilisant malgré tout des algorithmes 2D d'appariement de texture ou de cartes de texture, qui permettent d'obtenir des résultats intéressants. Les sections suivantes décrivent succinctement les algorithmes 2D implémentés dans SCULPTEUR [F.Schmitt and T.Tung, 2003] par l'équipe de recherche de Southampton et qui ont été développés dans le cadre du projet ARTISTE [Addis et al., 2002].

### 1.9.1 Histogramme de couleur

On utilise souvent des histogrammes de couleur (RVB, Lab, YDrDb, etc.) pour rechercher des images similaires en se basant sur la distribution générale des couleurs de l'image requête. L'histogramme contient les fréquences d'occurrence (c.-à-d. le nombre de pixels) des couleurs contenues dans l'image. Plus une couleur de l'image sera prédominante, plus sa fréquence sera forte dans l'histogramme (cf. Figure 1.14). Notons que l'utilisation d'un histogramme entraîne un compromis entre la vitesse et

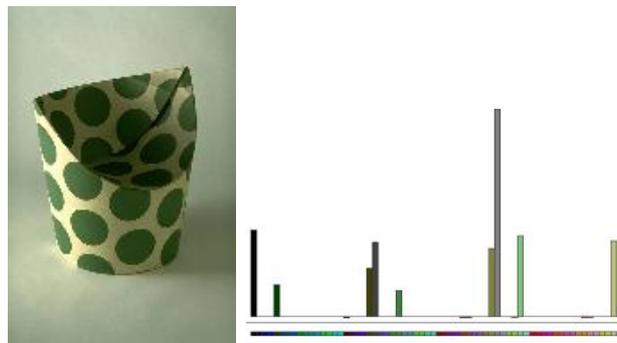


FIG. 1.14 – L'histogramme 3D dans un espace de couleur (RVB, Lab, YDrDb, etc.) une image par la distribution générale de ses couleurs [Addis et al., 2002].

la précision : plus l'histogramme contient d'intervalles, plus les comparaisons d'histogrammes prendront de temps, mais plus aussi les résultats deviendront discriminants.

Les histogrammes sont normalisés avant d'effectuer les comparaisons, tels que leur aire soit unitaire (division par le nombre de pixels de l'image). Ainsi, la taille de l'image requête n'affecte pas le processus de comparaison (si on néglige les problèmes de sous-échantillonnage). La méthode est bonne lorsque la requête est une image connue et que l'on souhaite retrouver les images ayant une distribution de couleur similaire. Elle n'est pas adaptée aux requêtes basées sur des extraits d'images. On note par ailleurs l'influence de la couleur du fond sur les histogrammes dans le cas particulier d'une

requête où l'image représente un objet 3D. C'est le principal inconvénient de cette méthode pour notre domaine d'application.

### 1.9.2 Histogramme monochrome

L'histogramme monochrome permet d'effectuer des recherches basées sur la distribution générale de l'intensité des pixels des images à niveaux de gris. L'histogramme contient les fréquences d'occurrence (i.e. le nombre de pixels) de chaque niveau d'intensité (cf. Figure 1.15). La méthode est adaptée aux images monochrome, comme les photographies en noir et blanc ou les radiographies.

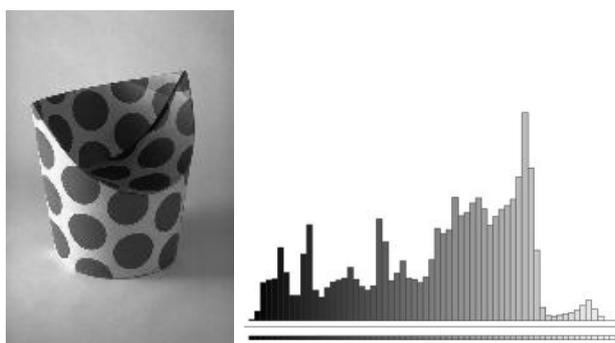


FIG. 1.15 – L'histogramme monochrome caractérise une image par la distribution générale de l'intensité des pixels.

Les histogrammes monochromes sont nécessaires lorsque la base de données contient des images monochromes car les histogrammes couleur ne sont pas assez discriminants avec ces images. Par exemple, un histogramme couleur à  $64 = 4^3$  intervalles de valeur ne contient que 4 intervalles dédiés aux valeurs de niveaux de gris (cf. Figure 1.16).

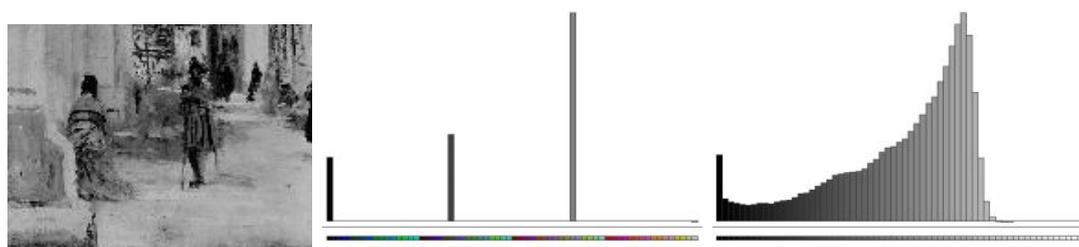


FIG. 1.16 – A nombre d'intervalles de valeur identique, l'histogramme monochrome (à droite) est plus discriminant que l'histogramme couleur (à gauche) pour les images monochromes. [Addis et al., 2002]

Les images en couleur sont converties en monochrome en convertissant les valeurs RVB en un intensité moyenne  $(R + V + B)/3$  ou en une valeur de luminance par une combinaison linéaire des composantes RVB dépendant de l'espace colorimétrique de référence.

### 1.9.3 Région de couleur CCV

La recherche d'images similaires par région de couleur (Colour Coherence Vector ou CCV) est basée sur la distribution générale des couleurs de l'image, avec une discrimination sur l'homogénéité des couleurs indépendamment de la taille variable des régions colorées. Une couleur dans une image est dite *cohérente* lorsqu'il existe dans l'image une région de cette couleur qui est plus grande qu'un certain seuil. On cherche ici les images qui ont la même distribution de couleurs cohérentes.

Pour cela on génère un histogramme pour les couleurs cohérentes, et un histogramme pour les couleurs incohérentes. Ces deux histogrammes sont utilisés séparément lors des recherches. La taille totale d'un vecteur caractéristique d'un CCV sera par exemple  $64 * 2$  nombres entiers. La cohérence et l'incohérence sont arbitrairement définis par rapport à une surface respectivement de plus ou moins 5% de la surface totale de l'image. Ainsi, une région de couleur dont la taille est inférieure à 5% de la taille totale de l'image contribuera à l'historgramme des couleurs incohérentes du CCV ; et une région de couleur dont la taille est supérieure à 5% de la taille totale de l'image contribuera à l'historgramme des couleurs cohérentes du CCV.

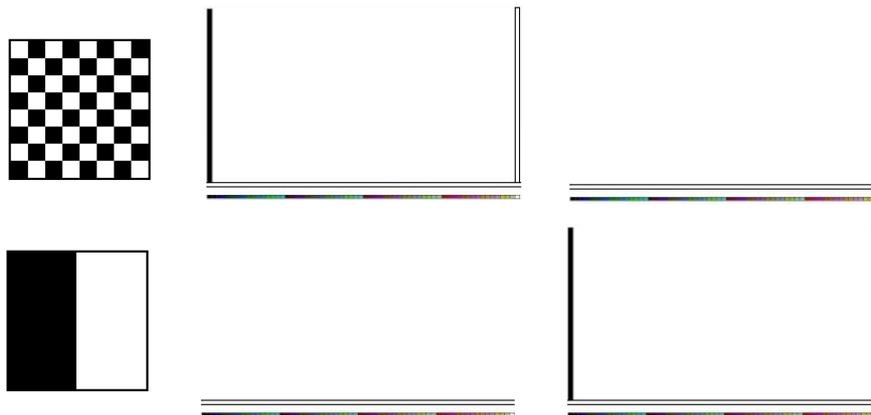


FIG. 1.17 – Histogrammes d'incohérence (centre) et de cohérence (droite) de deux images (gauche) ayant des régions de couleur de tailles différentes. En haut, les couleurs sont "incohérentes" car les régions associées sont de trop petite taille et l'historgramme de cohérence (à droite) est donc nul, alors qu'en bas les couleurs sont "cohérentes" car les régions associées étant de grande taille et l'historgramme d'incohérence est donc nul.

La figure 1.17 représente deux images ayant la même quantité de couleur : 50% de noir et 50% de blanc. L'image du haut est un échiquier contenant 32 cases blanches et 32 cases noires. Chaque case représente donc  $1/64^e$ , soit 1.5%, de la taille totale de l'image. Toutes les couleurs de cette image sont donc incohérentes et seul l'historgramme des couleurs incohérentes sert à la caractériser. Inversement, la seconde image contient deux régions couvrant chacune 50% de la taille totale de l'image. Ces régions sont donc cohérentes et seul l'historgramme des couleurs cohérentes sert à caractériser l'image.

Lorsque l'on compare les histogrammes d'incohérence et de cohérence des images, on obtient bien une grande distance entre les deux images, alors que des histogrammes classiques retourneraient une différence nulle comme si les deux images étaient similaires.

Comme précédemment la méthode convient pour des requêtes sur des images entières mais n'est pas adaptée aux requêtes basées sur des extraits d'images.

### 1.9.4 Histogramme monochrome multi-échelle

La méthode utilisant l'histogramme monochrome peut être étendue pour être adaptée aux requêtes basées sur des extraits d'images. Les images de la base de données sont divisées de manière à obtenir une hiérarchie pyramidale de sous-images et un histogramme monochrome est calculé pour chaque sous-image (cf. Figure 1.18). La requête est comparée avec toutes les parties d'images, toutes sous-images pouvant ainsi être localisées.



FIG. 1.18 – *Structure multi-échelle pyramidale.*

Comme pour l'histogramme monochrome, les images couleurs doivent être préalablement converties en niveaux de gris. La recherche de détail s'effectue en divisant de manière pyramidale la requête et les images de la base de données jusqu'à un certain niveau de résolution. Par exemple au niveau de résolution  $r = 6$ , l'image initiale est décomposée en  $64 * 64$  sous-images, et au niveau de résolution le plus faible ( $r = 0$ ) en une seule image. Pour chaque sous-image de la pyramide on dispose de son histogramme monochrome. Le vecteur caractéristique final est la concaténation de l'ensemble de ces histogrammes.

### 1.9.5 Région de couleur multi-échelle MCCV

Pour retrouver dans la base de données des détails d'une image en couleur, on peut utiliser le même procédé qu'avec les histogrammes monochromes multi-échelle. Les images sont divisées en une pyramide de sous-images et on calcule pour chaque sous-images les histogrammes d'incohérence et de cohérence des CCV (MCCV = Multi-scale CCV).

La division est appliquée à l'image requête et préalablement à toutes les images de la base de données et les comparaisons sont effectuées sur les CCV de chaque sous-images.

### 1.9.6 Transformée en ondelettes pyramidales (PWT)

La transformée en ondelettes pyramidales (Pyramidal Wavelet Transform ou PWT) permet de décomposer les images en composantes dans différentes bandes de fréquences (cf. Figure 1.19).

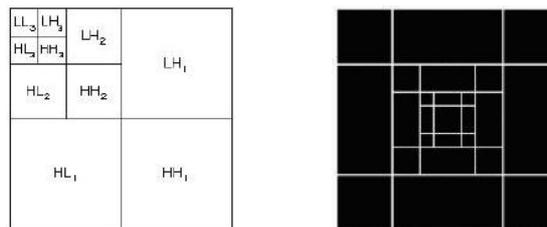


FIG. 1.19 – Gauche : décomposition dans l'espace image. Droite : décomposition dans l'espace des fréquences.

Après analyse, un nombre limite de descripteurs sont générés à partir des composantes fréquentielles. Par exemple, dans la base de données de SCULPTEUR, les images sont redimensionnées à 512\*512 pixels, et la décomposition retourne 22 descripteurs de fréquence pour une image. La recherche est donc très rapide, les comparaisons étant effectuées en utilisant une distance Euclidienne standard (L2) :

$$Dist(Q, M) = \sqrt{\sum_{j=1}^n (H_Q|j| - H_M|j|)^2}$$

où  $Dist(Q, M)$  représente la distance entre deux images  $Q$  et  $M$ , et  $H_Q$  et  $H_M$  les harmoniques composants les vecteurs caractéristiques de  $Q$  et  $M$ .

L'approche vise à retrouver les images ayant des textures similaires et elle requiert que les requêtes soient effectuées avec des images contenant suffisamment de texture.

### 1.9.7 Requête par fax (QBF)

Lorsque l'image requête est une image monochrome de mauvaise qualité, par exemple lorsqu'il s'agit d'un facsimilé d'une peinture à retrouver dans la base de données, on peut utiliser l'approche dénommée "requête par fax" (Query by Fax ou QBF). Elle repose sur des mesures de PWT d'images monochromes seuillées à différentes valeurs. Un vecteur caractéristique de QBF contient 99 vecteurs PWT correspondant aux différents niveaux de seuillage des images, entre 1% de noir et 99% de noir.

La procédure est la suivante :

1. binariser l'image requête,
2. déterminer le pourcentage de noir présent dans l'image,
3. extraire les vecteurs caractéristiques de la base de données d'images correspondant à la même proportion de noir
4. calculer les distances entre les vecteurs caractéristiques de PWT en utilisant la distance précédente.

La méthode est donc adaptée aux recherches d'images ayant les mêmes proportions de noir et de blanc.

## 1.10 Bilan

Les méthodes de requêtes par similarité de formes 3D sont beaucoup plus récentes que les méthodes 2D pour les images. Cependant la littérature scientifique est déjà très féconde dans ce domaine et des approches très différentes ont été proposées.

Les approches basées sur les histogrammes de courbure [Zaharia and Prêteux, 2002] ou les EGI (Extended Gaussian Image) [Horn, 1984, Kang and Ikeuchi, 1994] décrivent de manière très locale la géométrie des objets et de ce fait sont très sensibles aux modifications du maillage (bruit sur la position des sommets, décimation, etc.). De plus elles ne gardent aucune information topologique (connectivité du maillage). Inversement les approches globales comme les histogrammes de cordes [Paquet and Rioux, 1997], les distributions de forme [Osada et al., 2002], ou les ensembles de points pondérés [Tangelder and Velkamp, 2003] décrivent la forme des objets 3D de manière trop globale et imprécise pour être des outils de recherche pertinents, de même pour les descripteurs basés sur la transformée de Fourier ou les calculs de moments [Vranic et al., 2001, Kazhdan and Funkhouser, 2002].

Pour obtenir une description intuitive et capter des propriétés topologiques des formes, il s'avère intéressant d'employer une représentation par squelette ou graphe. Malheureusement, la faiblesse de ces dernières approches réside dans l'extraction du squelette qui est souvent coûteuse, trop sensible aux perturbations sur la surface de l'objet, ou nécessite un point source privilégiant une direction. Par ailleurs, la plupart n'offrent pas de procédure de comparaison [Goldak et al., 2001, Leymarie and Kimia, 2001, Verroust and Lazarus, 2000] ou de représentation multirésolution [Sundar et al., 2003, Biasotti et al., 2003] très adaptée à la recherche dans de grandes bases de données. L'utilisation des graphes de Reeb multirésolution (ou MRG) [Hilaga et al., 2001] nous a semblé être une solution très prometteuse, en particulier parce qu'elle constitue un moyen efficace d'exploiter avantageusement différentes techniques en les fusionnant, comme nous le proposons dans les chapitres suivants.

Le tableau 1.2 présente les méthodes d'indexation 3D que nous avons implémentées dans le cadre de nos recherches (cf. Chapitre 4). Les méthodes ont été choisies pour leur facilité d'implémentation, leur "notoriété" ou leur efficacité selon leurs auteurs. Le tableau reporte les paramètres proposés dans les références, ainsi que les principales caractéristiques des descripteurs.

De plus les distances suivantes ont été implémentées pour nos tests : les normes de Minkowski  $L_1$ ,  $L_2$ , et  $L_\infty$ , la distance de Bhattacharrya, la distance de Jeffreys, la divergence du  $\chi^2$ , l'intersection d'histogrammes, et la distance EMD.

Méthodes	Approche	Taille du descripteur	Distance	Normalisation requise (type ACP)	Sensible à la connectivité du maillage	Normales orientées requises
Rapport aire/vol 1.3.1	globale	1	$L_1$	non	oui	oui
Histogrammes de cordes 1.3.2	globale	192(64*3)	$L_2$	oui	non	non
Distribution de forme D2 1.3.3	globale	64	$L_1$	non	non	non
Descripteur de Hough 3D 1.3.4	globale	2560(20*128)	$L_1$ ou $L_2$	oui	non	oui
Spectre de forme 3D 1.4.1	locale	100	$L_1$ ou $L_2$	non	oui	oui
EGI complexes 1.4.2	locale	256(128*2)	$L_2$	oui	non	oui
Harmoniques sphériques 1.5.2	transf.	512(32*16)	$L_2$	non	non	non
MRG 1.7.3	graphe	variable	histo int	non	oui	non

TAB. 1.2 – Méthodes d'indexation 3D implémentées.



## CHAPITRE 2

# Graphe de Reeb multirésolution

D'après la théorie de Morse, une fonction continue définie sur une surface fermée caractérise la topologie de la surface en ses points critiques [Shinagawa et al., 1991]. Le graphe de Reeb d'un objet 3D se présente sous la forme d'un squelette construit à l'aide d'une telle fonction  $\mu$  basée sur la connectivité du maillage modélisant la surface de l'objet 3D. Les valeurs de la fonction  $\mu$  permettent de segmenter la surface de l'objet en régions auxquelles on associe à chacune d'entre elles un nœud du graphe. La structure du graphe est ensuite obtenue en reliant les nœuds dont les régions associées sont connexes. L'aspect multirésolution des graphes de Reeb multirésolution (MRG) provient de la discrétisation dichotomique des valeurs de la fonction et du regroupement hiérarchique des graphes de Reeb obtenus à chaque résolution. Nous présentons dans la première section 2.1 de ce chapitre la théorie de Morse. La section 2.2 décrit la construction du graphe de Reeb, et en particulier nous présentons les choix que nous avons faits concernant l'implémentation de la méthode. La section 2.3 traite de l'aspect multirésolution, et nous proposons en section 2.4 une discussion sur les propriétés des graphes de Reeb multirésolution.

## 2.1 Théorie de Morse

Dans le cadre de notre étude les objets 3D sont représentés par leurs surfaces. Notre approche vise à caractériser la topologie de ces surfaces. C'est pourquoi la théorie de Morse est bien adaptée car elle permet d'étudier la relation entre des fonctions définies dans un espace et la forme de cet espace. Dans notre contexte, les espaces considérés sont des 2-variétés<sup>1</sup>. La théorie de Morse s'applique aux espaces de dimension infinie. La théorie de Morse utilise des calculs variationnels pour établir une relation entre les points critiques d'une fonction différentiable définie sur une variété différentiable, et la topologie globale de la surface [Matsumoto, 1997, Milnor, 1963, Wood, 2003].

Pour illustrer la théorie de Morse, on peut par exemple étudier les points critiques de fonctions à une variable telles que  $\mu(x) = y$ . Ils permettent d'obtenir les maxima, minima, et points d'inflexion de la fonction. Par définition, les *points critiques* sont les points  $x_0$  satisfaisant  $\mu'(x_0) = 0$ . Un point critique est *non-dégénéré* si sa dérivée

---

<sup>1</sup>Une 2-variété est une surface où localement, autour de chaque point, la surface est Euclidienne, c'est-à-dire que la surface apparaît "plate" autour de chaque point.

## 2.1 Théorie de Morse

seconde est non-nulle ( $\mu''(x_0) \neq 0$ ). De même, pour les fonctions à deux variables telles que  $\mu(x, y) = z$ , les points critiques sont les points de gradient nul ( $\nabla\mu = 0$ , soit  $\partial\mu/\partial x = 0$  et  $\partial\mu/\partial y = 0$ ). Un point critique est dit non-dégénéré pour une fonction à deux variables si le déterminant de la Hessienne de  $\mu$  en ce point est non nul, la Hessienne étant la matrice des dérivées partielles du second ordre. L'expression du déterminant de la Hessienne est la suivante :

$$\det H\mu(x, y) = \partial^2\mu/\partial x^2 \times \partial^2\mu/\partial y^2 - (\partial^2\mu/\partial x\partial y)^2. \quad (2.1)$$

Ainsi, soit une variété différentiable  $S$ , et une fonction  $\mu : S \rightarrow \mathbb{R}$ , qui à tout point  $\mathbf{p} \in S$  associe un réel.  $\mu$  est une *fonction de Morse* si tous ses points critiques sont non-dégénérés et isolés<sup>2</sup>.

En reprenant l'exemple de la fonction à deux variables, la théorie de Morse démontre qu'au voisinage d'un point critique non-dégénéré, la fonction ne peut avoir qu'une allure très simple. En effet, en choisissant un système de coordonnées locales approprié, une fonction de deux variables aura l'une des trois formes suivantes au voisinage d'un point critique non-dégénéré (cf. Figure 2.1) :

- $\mu = x^2 + y^2 + c$
- $\mu = x^2 - y^2 + c$
- $\mu = -x^2 - y^2 + c$

où  $c$  est une constante.

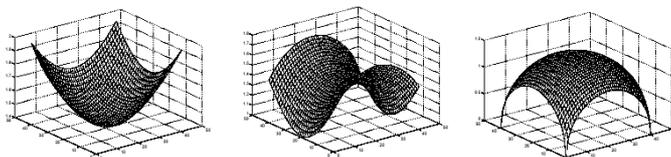


FIG. 2.1 – Allures des surfaces au voisinage des points critiques d'une fonction à deux variables : à gauche pour  $\mu = x^2 + y^2 + c$ , au milieu pour  $\mu = x^2 - y^2 + c$  et à droite pour  $\mu = -x^2 - y^2 + c$ .

La topologie et la forme globale de la variété différentiable peuvent donc être caractérisées en décomposant la surface en ces trois formes. La théorie de Morse permet alors d'obtenir des méthodes pour classer les points critiques selon leur type. En particulier, en étudiant le nombre de valeurs propres négatives de la Hessienne : un minimum n'a pas de valeurs propres négatives, un point selle en a une, et un maximum a deux valeurs propres négatives.

Un résultat classique de la théorie de Morse pour une fonction  $\mu : S \rightarrow \mathbb{R}$  définie sur une surface fermée  $S$ , est que si  $\mu$  a seulement deux points critiques non-dégénérés, alors  $S$  est topologiquement équivalent à une sphère. En effet, si on considère la fonction de hauteur, qui est la fonction de Morse classique, définie sur une sphère, les deux points critiques existants correspondent aux deux pôles Nord et Sud de la sphère.

<sup>2</sup>Les points critiques sont isolés si chaque point critique est associé à un unique réel.

Plus précisément, le nombre de points critiques est équivalent au nombre d'Euler caractérisant la surface :  $\chi = \#minima - \#selles + \#maxima$ . Un autre résultat de la théorie de Morse garantit qu'entre les points critiques, la topologie de la variété reste inchangée (lemme de la déformation).

La théorie de Morse est un outil puissant pour caractériser la topologie globale des 2-variétés différentiables. Dans le cadre de nos travaux, elle s'applique à un support discret, linéaire par morceaux. Des travaux récents ont visé à étendre la théorie de Morse pour les fonctions linéaires par morceaux sur des 2-variétés [Axen, 1999, Edelsbrunner et al., 2001].

Notre approche, présentée dans les sections suivantes, se base sur les travaux de [Hilaga et al., 2001] qui ont adapté la théorie de Morse pour des maillages 3D triangulés.

## 2.2 Graphe de Reeb

Le graphe de Reeb est une représentation graphique de la connectivité d'une surface entre ses points critiques. Les graphes de Reeb sont initialement utilisés pour représenter le squelette d'une variété topologique [Reeb, 1946]. Un graphe de Reeb devrait avoir un nœud associé à chaque point critique de la surface, et une arête entre les nœuds représentant les composantes connexes de la surface entre les points critiques. Cependant, la notion de graphe de Reeb peut être étendue de manière à représenter la relation entre les ensembles de niveaux pour une surface : on associe alors un nœud à chaque ensemble de niveau.

### 2.2.1 Définition du graphe de Reeb

Soit la fonction continue  $\mu : S \rightarrow \mathbb{R}$  définie sur la surface  $S$  d'un objet 3D. Le graphe de Reeb est par définition l'espace quotient du graphe de  $\mu$  dans  $S$ , définie par la relation d'équivalence suivante entre  $\mathbf{X} \in S$  et  $\mathbf{Y} \in S$  :

$$\mathbf{X} \sim \mathbf{Y} \iff \begin{cases} \mu(\mathbf{X}) = \mu(\mathbf{Y}) \\ \mathbf{X} \text{ et } \mathbf{Y} \text{ sont dans une même composante connexe de } \mu^{-1}(\mu(\mathbf{X})) \end{cases} \quad (2.2)$$

### 2.2.2 Exemple

Soit un ensemble  $S \subset \mathbb{R}^2$ ,  $\mathbf{v} \in S$ , et  $\mu$  tel que  $\mu(\mathbf{v}(x, y)) = y$ .

$\forall \mathbf{X}_1 = (x_1, y_1) \in S$  et  $\mathbf{X}_2 = (x_2, y_2) \in S$  :

- $\mu(\mathbf{X}_1) = \mu(\mathbf{X}_2)$  si et seulement si  $y_1 = y_2$ ,
- $\mu^{-1}(\mu(\mathbf{X}_1)) = \mu^{-1}(y_1)$  est la restriction de  $S$  à  $S \cap (\mathbb{R} \times y_1)$ , donc à la "section" d'ordonnée  $y_1$ .

Deux points appartiendront donc à la même classe d'équivalence si et seulement s'ils ont la même ordonnée, et s'ils sont dans la même composante connexe de la même

section (cf. Figure 2.2).

Le principal avantage du graphe de Reeb est qu'il permet de représenter de manière simple la topologie d'un ensemble. On note que l'aspect du graphe résultant est entièrement lié au choix de la fonction  $\mu$ .

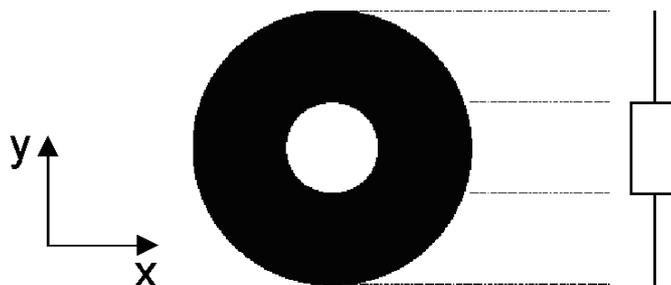


FIG. 2.2 – Exemple de graphe de Reeb en 2D.

### 2.2.3 Choix de la fonction $\mu$

Dans le cadre de notre étude, les modèles 3D sont définis par leur surface à l'aide d'un maillage 3D triangulé dont les sommets sont localisés dans un repère cartésien  $\mathcal{R}(x, y, z)$ . Il existe différentes fonctions continues qui peuvent être utilisées pour la construction du graphe de Reeb [Biasotti et al., 2003] et qui ont des propriétés différentes :

- Soit un point  $\mathbf{v}(x, y, z)$  sur la surface  $S$  d'un objet, la fonction de hauteur définie par  $\mu(\mathbf{v}(x, y, z)) = z$  est bien adaptée pour les modèles privilégiant la verticale comme les représentations d'êtres humains. Cette fonction est élémentaire à implémenter et ne nécessite pas de calcul, mais par définition est totalement dépendante de l'orientation de l'objet. Par exemple, elle n'est pas adaptée aux objets dont les points sont majoritairement dispersés dans des directions perpendiculaires à l'axe  $z$ .
- La fonction définie comme la distance géodésique des extrema de courbure est basée sur des croissances de régions à partir de courbures Gaussienne locales sur des sommets (germes). Les résultats dépendent de la position des germes et nécessitent le calcul des courbures locales ce qui n'est pas toujours réaliste, spécialement pour les objets provenant d'Internet dont les maillages sont très souvent de mauvaise qualité géométrique [Shilane et al., 2004, 3dcafe, 2005].
- La fonction centroïde définie comme la distance d'un point de la surface  $\mathbf{v}$  au centre de masse  $\mathbf{G}$  de l'objet  $\mu(\mathbf{v}) = d(\mathbf{G}, \mathbf{v})$ , où  $d$  est la distance Euclidienne, est simple à calculer et stable aux petites perturbations. Par contre, cette fonction donne un graphe de Reeb qui n'est pas suffisamment flexible dans le cadre de notre étude, comme expliqué ci-après.

- La fonction  $\mu$  proposée dans [Hilaga et al., 2001] et que l'on a retenue, est définie comme l'intégrale des distances géodésiques  $g(\mathbf{v}, \mathbf{p})$  de  $\mathbf{v}$  aux autres points  $\mathbf{p}$  de la surface (cf. Figure 2.3) :

$$\mu(\mathbf{v}) = \int_{\mathbf{p} \in S} g(\mathbf{v}, \mathbf{p}) dS. \quad (2.3)$$

La fonction  $\mu$  choisie a la propriété d'invariance à la rotation. Elle est stable en cas de modifications locales du maillage car de forme intégrale sur l'ensemble de l'objet. L'intégrale de la distance géodésique sur toute la surface de l'objet permet de mesurer l'excentricité des points de la surface de l'objet dans le sens où plus la valeur de  $\mu$  en un point est élevée, plus ce point est excentré, et inversement le point ayant une valeur minimale de  $\mu$  occupe une partie centrale de l'objet pour en constituer une sorte de point nombrilique. En effet, lorsqu'un point est excentré, sa distance aux autres points est grande et donc l'intégrale des distances géodésiques est plus élevée. Ainsi, cette fonction caractérise de manière plus intuitive la topologie des formes comparé à la fonction centroïde. Les mains d'un modèle humain conserveront toujours la même valeur de  $\mu$  indépendamment de la position du bras, ce qui n'est pas le cas avec la fonction centroïde (cf. Figure 2.3).

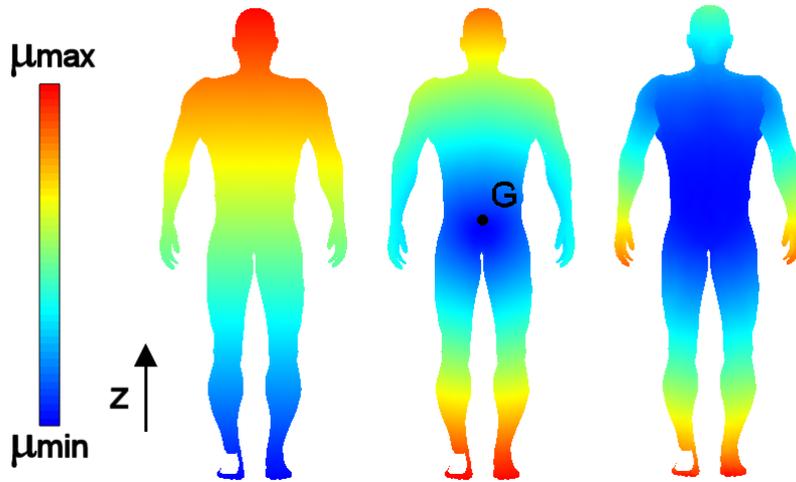


FIG. 2.3 – Exemple de fonctions  $\mu$  : hauteur (gauche), distance au centre de masse (milieu), intégrale de distance géodésique (droite).

Lazarus *et al.* ont proposé un diagramme d'ensemble de niveau (Level Set Diagram ou LSD) [Lazarus and Verroust, 1999] dont la construction repose sur une fonction  $\mu$  représentant la distance géodésique à un point source. Leur approche permet d'obtenir un découpage régulier des objets 3D en tranche que l'on peut relier pour obtenir un squelette, mais la détermination du point source, dans le cadre de notre application, devrait être effectuée de manière automatique et robuste, ce qui reste un problème difficile.

### 2.2.4 Discrétisation de la fonction $\mu$

Il y a un nombre infini de points  $\mathbf{v}$  sur la surface de l'objet, et il est donc impossible de calculer  $\mu$  en tout point. On se limite donc au calcul de  $\mu$  en chaque sommet du maillage, et la valeur en un autre point quelconque est obtenue par interpolation barycentrique à partir des valeurs de  $\mu$  calculées aux sommets du triangle contenant le point. Si le point est situé sur une arête sa valeur  $\mu$  se déduit par interpolation linéaire des valeurs de  $\mu$  aux deux sommets de l'arête.

Dans notre étude nous ne considérons que des modèles 3D représentés par des maillages, et donc le calcul de  $\mu$  sur la surface s'effectue de manière discrète :

$$\mu(\mathbf{v}) = \sum_{\mathbf{p} \in V} g(\mathbf{v}, \mathbf{p}) \cdot aire(S_p)$$

où  $V$  est l'ensemble des sommets de l'objet et  $\{S_p\}_{\mathbf{p} \in V}$  une partition de la surface de l'objet associée aux sommets  $p$ .

Cependant le coût de calcul reste extrêmement élevé en prenant tous les sommets du maillage. On le réduit par les deux approches suivantes :

- La distance géodésique entre deux sommets du maillage est approximée par la plus courte distance entre ces deux sommets dans le parcours du graphe des arêtes. Théoriquement, l'erreur introduite n'est pas bornée mais [Hilaga et al., 2001] proposent de définir un seuil et de subdiviser les triangles du maillage jusqu'à ce que par exemple, la longueur de chaque arête soit inférieure à ce seuil, et ceci dans le but de minimiser l'erreur. Le test du seuil n'a pas été implémenté ici car presque tous les modèles utilisés pour nos travaux de recherche sont très résolus.
- La somme sur les sommets du maillage n'est effectuée qu'en un nombre restreint de sommets appelés *bases* ce qui va permettre de réduire considérablement le temps de calcul. On définit  $r$  comme étant une distance seuil telle que tous les sommets se trouvant à une distance inférieure à  $r$  d'une base soient associés à cette base. Le nombre de bases est jugé suffisant lorsque tous les sommets du maillage appartiennent à une base.

Ainsi,  $\mu$  est donc approximé par :

$$\mu(v) = \sum_{i=1}^N d(v, b_i) \cdot aire(b_i) \quad (2.4)$$

où  $\{b_i\}_{i \in [1;n]}$  représente l'ensemble des  $N$  bases,  $d(.,.)$  est la distance la plus courte entre deux sommets du maillage dans le graphe des arêtes, et  $aire(b_i)$  est l'aire de la surface associée à la base  $b_i$  (cf. Figure 2.4). Evidemment si le maillage de la surface n'est pas suffisamment résolu, les calculs de distance géodésique ne seront pas assez précis, et il sera nécessaire de subdiviser les triangles jusqu'à obtenir un niveau de résolution satisfaisant.

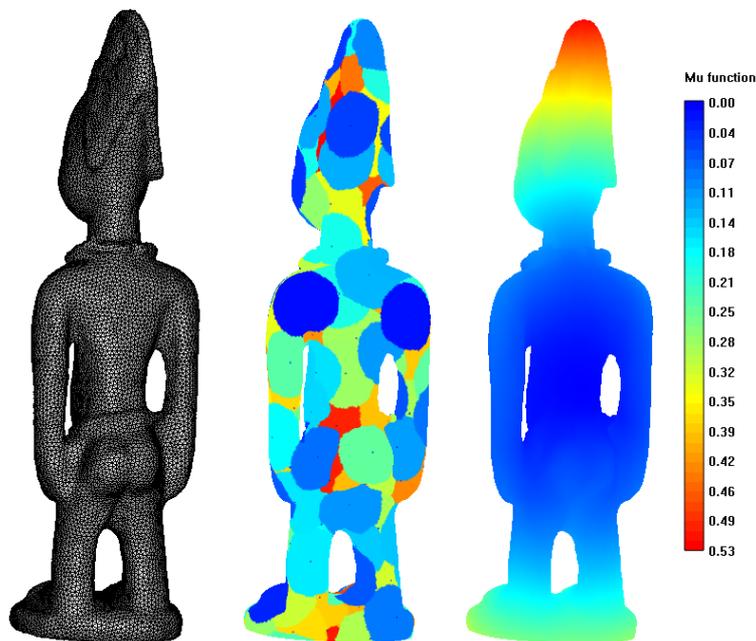


FIG. 2.4 – Gauche : maillage initial. Milieu : les bases en couleur représentent les points de calcul. Droite : fonction  $\mu$  obtenue.

### 2.2.5 Choix des bases

La technique utilisée pour le choix des bases n'a pas été détaillée dans [Hilaga et al., 2001]. La méthode suivante a été implémentée :

- La première base est choisie au hasard : par exemple, le dernier sommet du maillage de l'objet.
- Quand  $n$  bases ont été choisies et s'il reste des sommets qui ne sont dans aucune région d'influence délimitée par le seuil  $r$  : parmi ces sommets, la base  $n + 1$  est le sommet le plus éloigné en moyenne des  $n$  premières bases.

L'avantage de cette méthode par rapport à un choix totalement aléatoire, est que si la répartition des sommets sur le maillage de l'objet est inhomogène, cela empêche les bases de s'accumuler dans les zones où la densité de sommets est plus élevée.

Par ailleurs, le seuil  $r$  doit également être choisi avec attention pour permettre un échantillonnage des sommets qui soit un bon compromis entre la rapidité de calcul, et la minimisation de l'erreur induite dans le calcul de  $\mu$ . Les auteurs de [Hilaga et al., 2001] proposent :

$$r = \sqrt{5 \cdot 10^{-3} \cdot aire(S)}.$$

où  $aire(S)$  est l'aire totale de la surface de l'objet. Cette définition permet d'obtenir généralement 150 à 200 bases, ce qui réalise un compromis satisfaisant. Elle n'a donc pas été remise en cause.

Pour déterminer le terme  $aire(b_i)$  de 2.4, on considère que chaque sommet apporte à la base à laquelle il est associé (s'il fait partie de plusieurs régions, on n'en considère qu'une pour ce calcul) une contribution égale à la somme du tiers de l'aire de chaque triangle dont il est un sommet. L'important étant de vérifier :

$$\sum_{i=1}^N aire(b_i) = aire(S) \quad (2.5)$$

ce qui est réalisé ici en divisant chaque triangle en trois parties égales rattachées chacune à un des trois sommets.

Les centres des  $N$  bases obtenues servent donc de points de calcul de distances géodésiques pour la fonction  $\mu$ . Cependant malgré l'approximation proposée il reste coûteux de calculer les  $N$  sommes de distances géodésiques d'une base aux  $N - 1$  autres bases, soit  $N(N - 1)$  opérations. Pour pallier ce problème, [Hilaga et al., 2001] proposent d'appliquer l'algorithme de calcul des plus courts chemins de Dijkstra.

### 2.2.6 Algorithme de Dijkstra

Les méthodes pour calculer précisément une distance géodésique ont été bien étudiées dans la littérature [Sharir and Schorr, 1986, Mitchell et al., 1987, Chen and Han, 1990, Lanthier et al., 1999, Kanai and Suzuki, 2000, Peyre and Cohen, 2003]. Cependant, il reste très coûteux de calculer l'intégrale des distances géodésiques avec ces méthodes. Pour obtenir un compromis entre précision et temps de calcul, l'algorithme de Dijkstra (pour le calcul du plus court chemin sur un graphe pour aller d'un sommet à un autre) a été choisi pour approximer les distances géodésiques. L'intérêt de cet algorithme est qu'il réalise le calcul avec une complexité en temps de  $\mathcal{O}(N \cdot \log(N))$ ,  $N$  étant le nombre de sommets utilisés pour le calcul.

Soit  $d_i(\cdot)$ , la distance de la base  $b_i$  aux autres sommets du maillage, l'algorithme est le suivant :

1. Poser  $d_i(v) = +\infty$  pour tous les sommets du maillage,  $d_i(b_i) = 0$ , et créer une liste de points  $L$  contenant initialement l'ensemble des  $\{b_i\}$  ( $L$  est en fait un arbre binaire de recherche [Hilaga et al., 2001]).
2. Trouver dans  $L$  le sommet  $v$  qui minimise  $d_i(v)$  et l'extraire de  $L$ .
3. Pour tous les sommets  $v_a$  adjacents à  $v$  :  
 si  $d_i(v) + |v - v_a| < d_i(v_a)$ , faire  $d_i(v_a) = d_i(v) + |v - v_a|$   
 et insérer  $v_a$  dans  $L$   
 (deux points sont adjacents s'ils sont reliés par une arête;  $|v - v_a|$  est égal à la longueur de cette arête).
4. Tant que  $L$  est non vide, aller en 2.

## 2.2.7 Normalisation de $\mu$

La fonction  $\mu$  précédemment définie n'est pas invariante en cas de changement d'échelle dans toutes les directions, et il est donc nécessaire de la normaliser. Notons  $\mu_{min} = \min_{\mathbf{p} \in S} \mu(\mathbf{p})$  et  $\mu_{max} = \max_{\mathbf{p} \in S} \mu(\mathbf{p})$ . Les auteurs [Hilaga et al., 2001] ont choisi de normaliser la fonction avec au dénominateur  $\mu_{max}$  :

$$\mu_N(\mathbf{v}) = \frac{\mu(\mathbf{v}) - \mu_{min}}{\mu_{max}}. \quad (2.6)$$

Les valeurs de  $\mu(\mathbf{v})$  sont donc comprises entre 0 et  $\frac{\mu_{max} - \mu_{min}}{\mu_{max}} \leq 1$ .

Afin de profiter d'une dynamique de  $\mu(\mathbf{v})$  sur  $[0, 1]$ , on peut aussi normaliser la fonction avec ses valeurs *min* et *max* :

$$\mu_N(\mathbf{v}) = \frac{\mu(\mathbf{v}) - \mu_{min}}{\mu_{max} - \mu_{min}}. \quad (2.7)$$

Cette normalisation est évidemment valable uniquement si le dénominateur est non nul. En particulier, la fonction  $\mu$  est constante sur toute la surface d'une sphère et dans ce cas on ne peut donc définir de fonction  $\mu_N$  avec cette formulation.

Il est important de noter ici que la normalisation est directement liée à la dynamique de  $\mu$ , et l'allure du graphe de Reeb dépendra de son choix. En effet, dans la suite du procédé de construction du graphe de Reeb, nous effectuons un découpage de la surface en intervalle de taille  $1/2^R$ , suivant les valeurs de  $\mu$  calculée en chaque point,  $R$  étant le niveau de résolution maximal. Le découpage de la surface est donc totalement lié à la résolution du graphe comme nous le verrons en détail dans les sections suivantes.

Si nous considérons le rapport  $\frac{\mu_{min}}{\mu_{max}}$  on observe que :

- si  $\frac{\mu_{min}}{\mu_{max}} \rightarrow 1 \Rightarrow \mu_{min} \sim \mu_{max}$  : la dynamique est faible, et la normalisation de l'Eq. 2.6 avec le *max* pourrait s'avérer insuffisante pour capturer les détails de la surface de l'objet si la résolution du découpage de la surface n'est pas assez fine. La formulation de l'Eq. 2.7 risque au contraire de produire un graphe peu robuste et très bruité si la surface est bruitée.
- si  $\frac{\mu_{min}}{\mu_{max}} \rightarrow 0 \Rightarrow \mu_{min} \ll \mu_{max}$  : plus  $\mu_{max}$  s'éloigne de  $\mu_{min}$ , plus les deux normalisations de l'Eq. 2.6 et de l'Eq. 2.7 mèneront au même découpage de la surface, et donc au même graphe de Reeb. On notera que si la taille de l'objet est normalisée il est impossible que  $\mu_{min} \rightarrow 0$  avec la fonction  $\mu(\mathbf{v}) = \int_{\mathbf{p} \in S} g(\mathbf{v}, \mathbf{p}) dS$ .

Ce que l'on peut conclure est que quelque soit la dynamique de  $\mu$ , la normalisation de l'Eq. 2.7 avec *max* – *min* permettrait toujours d'avoir une représentation de la surface sans perte de détail a priori, car tout l'intervalle  $[0, 1]$  est couvert par cette normalisation. Malheureusement cela implique que l'Eq. 2.7 est beaucoup plus sensible aux bruits. Dans l'implémentation de la méthode, nous avons donc choisi de normaliser la fonction avec sa valeur maximale  $\mu_{max}$ .

## 2.2.8 Construction du graphe de Reeb

### Discrétisation du graphe de Reeb

Pour une implémentation algorithmique, le graphe de Reeb n'est pas construit de manière continue, comme le suggère la définition en 2.2.6, mais de manière discrète. L'intervalle de définition de  $\mu$  est divisé en  $k$  intervalles disjoints  $I_1, \dots, I_k$ , et les conditions en 2.2 deviennent :

$$\mathbf{X} \sim \mathbf{Y} \iff \begin{cases} \exists k_0 \in [1; k] : & \mu(X) \in I_{k_0} \text{ et } \mu(Y) \in I_{k_0} \\ \mathbf{X} \text{ et } \mathbf{Y} \text{ sont dans une même composante connexe de } & \mu^{-1}(I_{k_0}) \end{cases} \quad (2.8)$$

Ainsi, cela signifie que pour un intervalle donné  $I_k$ , si on considère seulement la partie de l'objet dont l'image par  $\mu$  est dans  $I_k$  (c'est-à-dire  $\mu^{-1}(I_k)$ ), il y a dans cette restriction autant de classes d'équivalence que de composantes connexes.

L'intervalle de définition de  $\mu_N(\mathbf{v})$  est inclus dans  $[0, 1]$ . On découpe l'intervalle  $[0, 1]$  de façon régulière en un nombre de sous-intervalles égal à une puissance entière de 2. Si on divise  $[0, 1]$  en  $2^r$  intervalles ( $r \geq 0$ ), on appelle  $r$  le *niveau de résolution* du graphe de Reeb ainsi créé. Un des intérêts de cette subdivision est bien sûr que chaque intervalle du niveau de résolution  $r$  contient exactement deux intervalles du niveau  $r + 1$ .

A un niveau de résolution fixé, les nœuds et les arêtes du graphe de Reeb discrétisé sont définis de la manière suivante :

**Définition 1** *Chaque classe d'équivalence de la relation définie par les conditions 2.8 définit un **nœud** du graphe de Reeb discrétisé. Un nœud du graphe de Reeb correspond donc à un intervalle<sup>3</sup> de  $\mu$  et à un sous-ensemble connexe de l'objet.*

Pour le sous-ensemble de l'objet qui est l'image inverse d'un sous-intervalle de l'intervalle de définition de  $\mu$ , il y a autant de nœuds que de composantes connexes. Pour l'objet entier, on peut construire une bijection entre les nœuds du graphe à une résolution  $r$  fixé et les éléments d'une partition de l'objet. Deux éléments de la partition sont connexes si et seulement si leurs nœuds associés sont reliés par une arête du graphe. On fera souvent référence, en parlant d'un nœud du graphe, au sous-ensemble ce l'objet qui lui correspond.

**Définition 2** *Une **arête** relie deux nœuds du graphe de Reeb discrétisé si et seulement si les deux sous-ensembles de l'objet qui correspondent à ces deux nœuds sont connexes.*

Comme chacun des deux sous-ensembles considérés est par construction connexe, cela signifie qu'ils doivent être "en contact". Notamment, une condition nécessaire non suffisante pour que deux nœuds soient reliés est que les deux sous-intervalles de l'intervalle de définition qui ont servi à leur construction soient eux-mêmes connexes.

---

<sup>3</sup>L'amplitude de cet intervalle permet de retrouver la niveau de résolution auquel est défini le nœud.

## Représentation

Concrètement, et en se basant sur ce qui a été présenté précédemment, le graphe de Reeb est obtenu en partitionnant itérativement l'objet suivant des intervalles réguliers de valeurs de  $\mu$ , et en affectant un nœud à chaque ensemble connexe de triangles. Les régions connexes (les nœuds) sont reliées entre-elles par des arêtes (cf. Figure 2.5).

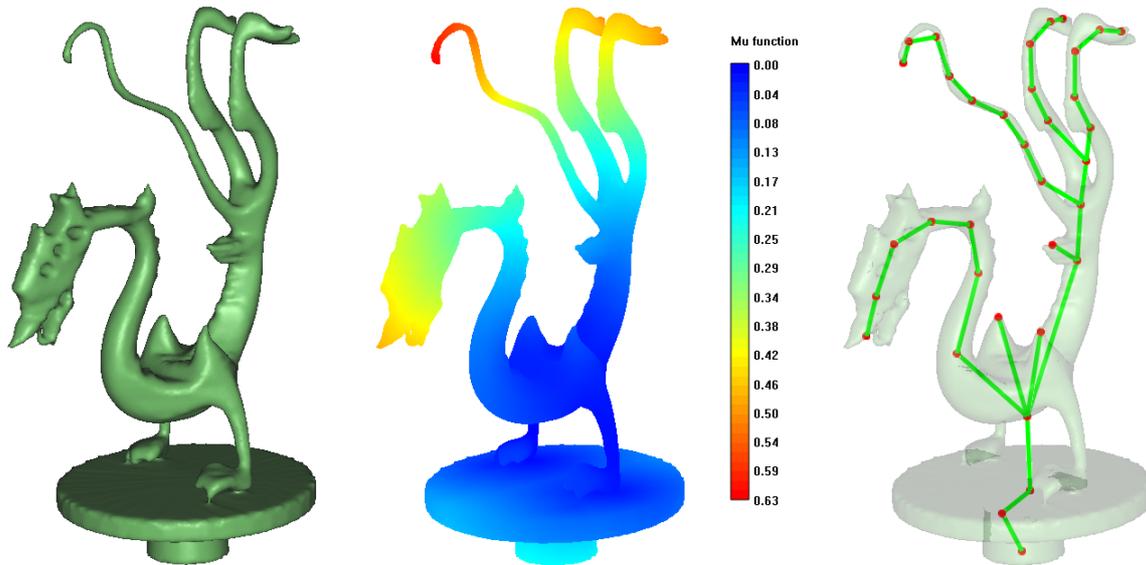


FIG. 2.5 – Graphe de Reeb : modèle 3D (gauche), fonction  $\mu_N$  (centre), graphe de Reeb; arêtes en vert; nœuds en rouge (droite).

## 2.3 Graphe de Reeb Multirésolution

### 2.3.1 Définition

Cette section présente le graphe de Reeb multirésolution (Multirésolution Reeb Graph ou MRG) [Hilaga et al., 2001]. L'idée est de construire des graphes de Reeb d'un objet 3D à plusieurs niveaux de résolution. Comme présentée précédemment, pour construire un graphe de Reeb à une résolution donnée  $r$ , l'objet est partitionné en régions basées sur les valeurs de la fonction  $\mu$ . Un nœud du graphe de Reeb représente une composante connexe dans une région particulière, et les nœuds adjacents sont liés par une arête si leurs composantes connexes associées sont en contact avec elle. Le graphe de Reeb au niveau immédiatement supérieur  $r+1$  est obtenu en re-partitionnant chaque région (en l'occurrence, chaque région est divisée en deux).

Soient  $G_r$ ,  $r \in [0; R]$ , les  $R + 1$  graphes obtenus en faisant varier le niveau de résolution entre 0 et  $R$ . Pour passer d'un niveau de résolution  $r$  ( $0 \leq r \leq R - 1$ ) au niveau  $r + 1$ , on re-partitionne chaque intervalle :

$$I_{r,k} = \left[ \frac{k}{2^r}; \frac{k+1}{2^r} \right], k \in [0; 2^r - 1] \quad (2.9)$$

du niveau  $r$  en deux sous-intervalles égaux :

$$I_{r+1,2k} = \left[ \frac{2k}{2^{r+1}}; \frac{2k+1}{2^{r+1}} \right] \text{ et } I_{r+1,2k+1} = \left[ \frac{2k+1}{2^{r+1}}; \frac{2k+2}{2^{r+1}} \right]. \quad (2.10)$$

Si on considère un nœud du niveau  $r$ , l'image du sous-ensemble de l'objet auquel il correspond par  $\mu$  est comprise dans  $I_{r,k}$ , donc quand  $I_{r,k}$  est repartitionné en  $I_{r+1,2k}$  et  $I_{r+1,2k+1}$ , les nouveaux sous-ensembles formés par les classes d'équivalence forment une partition du sous-ensemble d'origine. Ainsi, il y a une relation naturelle parent-fils entre le nœud du niveau  $r$  et les nœuds du niveau  $r+1$  (cf. Figure 2.6).

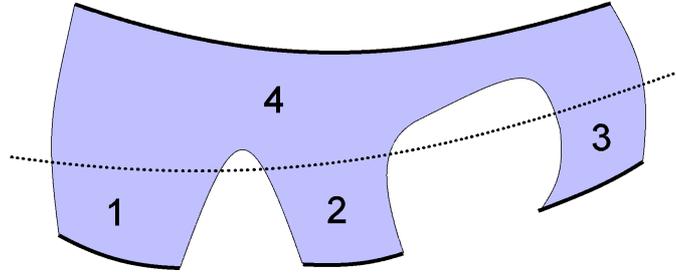


FIG. 2.6 – *Passage à un niveau de résolution supérieur : les traits épais continus représentent les frontières de l'intervalle  $I_{r,k}$  reportées sur la surface de l'objet. La surface fermée par ces traits et les traits fins continus est un sous-ensemble connexe de l'objet compris entièrement dans  $I_{r,k}$  : il correspond donc exactement à un nœud du niveau de résolution  $r$ . Le trait pointillé représente la nouvelle frontière de  $\mu$  qu'il faut considérer au niveau  $r+1$  : elle partage  $I_{r,k}$  en  $I_{r+1,2k}$  et  $I_{r+1,2k+1}$ . Ce trait partitionne le sous-ensemble initial en 4 sous-ensembles connexes : le nœud du niveau  $r$  aura donc 4 nœuds au niveau  $r+1$ .*

Le graphe de Reeb multirésolution à la résolution  $R$  (niveau de résolution maximal pris en compte pour le calcul des graphes discrétisés) est constitué de l'“empilement” des graphes de Reeb discrétisés des niveaux de résolution compris entre 0 et  $R$ . Grâce aux relations parent-fils qu'on a mises en évidence entre les niveaux de résolution voisins, cette construction est cohérente. Ainsi, un nœud du graphe de Reeb multirésolution a exactement un nœud-parent (sauf si c'est un nœud du niveau de résolution 0, auquel cas il n'a pas de parent) ; et il peut avoir un nombre quelconque de fils  $\geq 1$ .

En résumé, le graphe de Reeb multirésolution à la résolution  $R$  est un graphe dont l'ensemble des nœuds est l'union des nœuds des graphes de Reeb discrétisés des niveaux de résolutions compris dans  $[0; R]$ , et l'ensemble des arêtes est l'union des arêtes de ces graphes et des arêtes parents-fils entre deux graphes consécutifs.

### 2.3.2 Exemple

Nous choisissons un exemple simple où la fonction de hauteur est choisie comme fonction de Morse  $\mu_N$ . Pour construire un graphe de Reeb à  $R$  niveaux de résolution,  $\mu_N$  est subdivisé en  $2^R$  intervalles à partir desquels la surface de l'objet est partitionnée au niveau de résolution le plus grand. La construction est ensuite effectuée de manière hiérarchique, le passage à un graphe de résolution inférieur se faisant par fusion deux par deux des intervalles voisins définis en à l'Eq. 2.10 [Hilaga et al., 2001]. Un nœud parent est associé à chaque ensemble connexe de chaque nouvel intervalle obtenu et est relié à ces  $n$  nœuds fils ( $n \geq 1$ ) appartenant au niveau de résolution immédiatement supérieur et dont il est la fusion (cf. Figure 2.7).

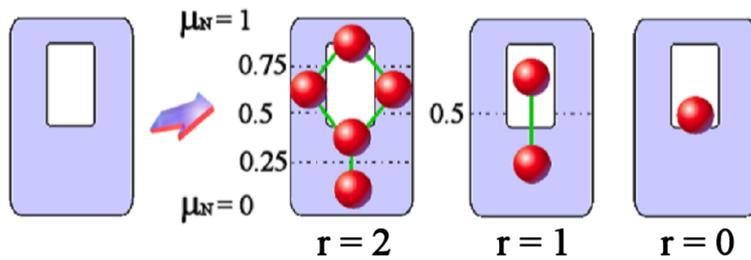


FIG. 2.7 – Construction d'un MRG par fusion des intervalles de  $\mu_N$  partitionnant la surface de l'objet. Les nœuds d'un niveau de résolution  $r + 1$  fusionnent au niveau inférieur  $r$  en des nœuds parents (de gauche à droite : résolution  $r = 2, 1$  et  $0$ ).

### 2.3.3 Propriétés

Le graphe de Reeb multirésolution a les propriétés suivantes :

- Propriété 1 :** Il existe des relations parents-fils entre les nœuds des niveaux de résolution adjacents.
- Propriété 2 :** En réitérant la subdivision, le MRG converge vers le graphe de Reeb original comme défini par Reeb [Reeb, 1946]. Les niveaux de résolution plus fins approximent l'objet original avec plus de précision.
- Propriété 3 :** Un graphe de Reeb d'un certain niveau de résolution contient implicitement toutes les informations des niveaux plus grossiers. Lorsque le niveau de résolution le plus fin a été choisi, un graphe de Reeb d'un niveau de résolution inférieur peut être construit par fusion de nœuds adjacents. La fusion des nœuds conduit à la relation parent-enfant.
- Propriété 4 :** Pour comparer deux graphes de Reeb multirésolutions, il est naturel de comparer d'abord les deux graphes à des niveaux de résolution grossiers, puis de descendre à des niveaux plus fins si cela a été concluant. Ajouté au fait qu'à un niveau de résolution suffisamment élevé, le graphe de Reeb discrétisé est une simplification fidèle de la forme de l'objet, cela explique l'intérêt de cette méthode

dans le cadre de l'indexation (pour plus de détails sur le calcul de similarité, voir le chapitre suivant).

## 2.4 Discussion sur les graphes MRG

Les sections suivantes présentent quelques cas d'études mettant en évidence les points sensibles des graphes de Reeb multirésolution. En particulier, le choix de la fonction  $\mu$ , du niveau de résolution maximal, ainsi que la qualité des maillages sont des paramètres d'entrée primordiaux qui affectent les détails que peuvent capturer les graphes.

Les résultats de cette étude nous ont orientés dans nos différents choix pour les expériences présentées au chapitre 4.

### 2.4.1 Choix de la fonction $\mu$

Les figures 2.8 et 2.9 illustrent l'influence du choix de la fonction  $\mu$ . La figure 2.8 présente les MRG obtenus aux résolutions  $r = 1$  à 5 avec la fonction de hauteur  $\mu_h(\mathbf{P}) = z$  pour tout point  $\mathbf{P} = (x, y, z)$  de la surface  $S$  de l'objet. On constate que les bras du modèle *baby* [3dcafe, 2005], perpendiculaires à l'axe  $z$ , ont du mal à être capturés par la fonction de hauteur. Tandis qu'avec la fonction géodésique intégrale  $\mu_g(\mathbf{P}) = \int_{\mathbf{v} \in S} g(\mathbf{v}, \mathbf{P}) dS$  (cf. Eq 2.3), tous les membres sont capturés dès le niveau de résolution  $r = 2$ .

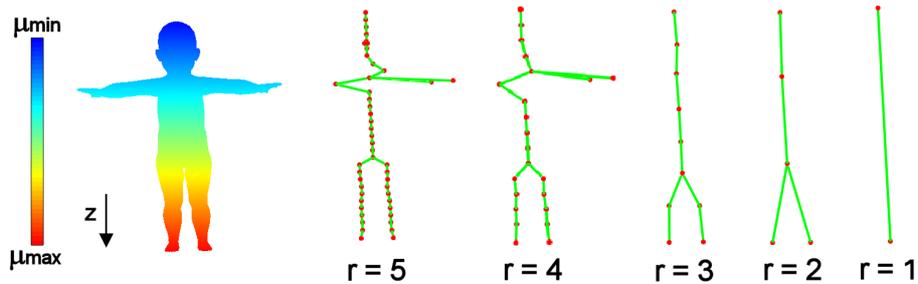


FIG. 2.8 – MRG d'un modèle humain aux résolutions  $r = 1$  à 5 obtenu avec la fonction de hauteur.

### 2.4.2 Résolution du graphe

La précision du graphe dépend de sa résolution. On observe par exemple que certains détails n'apparaissent qu'à partir de certains niveaux de résolutions (cf. Figures 2.10 et 2.11).

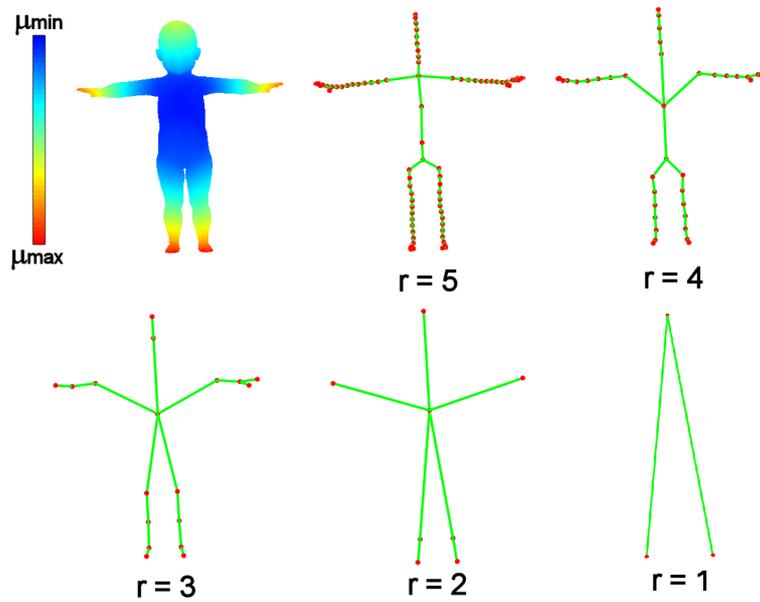


FIG. 2.9 – MRG d'un modèle humain aux résolutions  $r = 1$  à 5 obtenu avec la fonction géodésique intégrale.

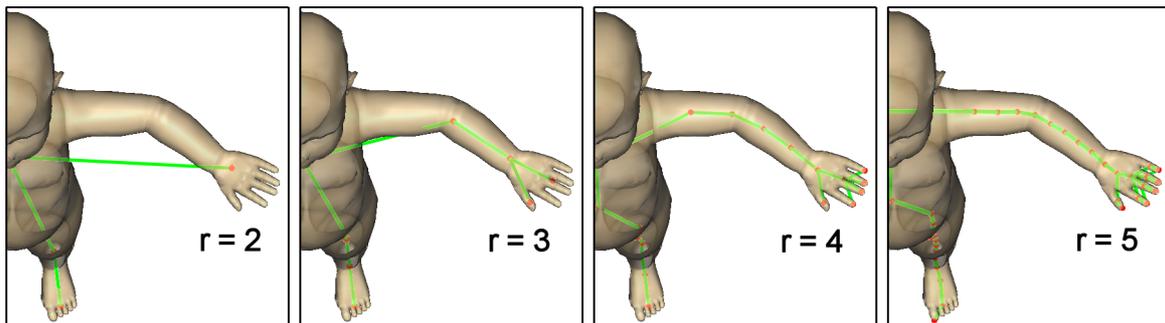


FIG. 2.10 – Détail du modèle baby. Les cinq doigts apparaissent à la résolution  $r = 4$ .

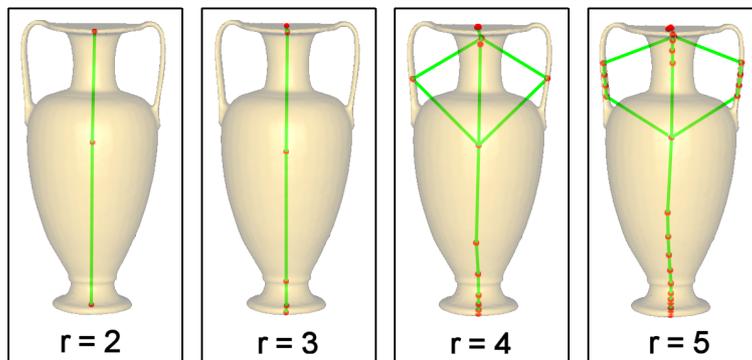


FIG. 2.11 – Modèle vase. Les anses apparaissent à la résolution  $r = 4$ .

### 2.4.3 Maillage

On illustre ici la dépendance du graphe à la résolution et à la connectivité du maillage. La figure 2.12 représente deux maillages différents d'un cône, l'un avec 33 triangles (cône a)) et l'autre avec 242 triangles (cône b)), ainsi que le maillage de leur base qui est excentré. La surface du cône a) est constituée de triangles générateurs (dont un sommet est le sommet du cône, et les deux autres appartiennent à la base du cône). La surface du cône b) est constituée de triangles générateurs subdivisés.

La figure 2.13 représente les valeurs de la fonction  $\mu$  géodésique intégrale calculée en chaque sommet des deux cônes avec l'algorithme discret décrit dans les sections 2.2.4 à 2.2.7 (vues de dessus, et vues de la base). Les cônes étant parfaitement à symétrie de révolution, les valeurs de  $\mu$  minimales  $\mu_{min}$  devraient se situer dans les zones centrales, c'est-à-dire entre la base et le sommet (zones de  $\mu$  maximal). On constate que le minimum  $\mu_{min}$  calculé sur le cône a), faiblement résolu, est complètement excentré sur le bord de la base. Ceci est dû au maillage asymétrique et de faible résolution de la base du cône. Les valeurs n'étant attribuées qu'aux sommets du maillage, un objet faiblement résolu ne possèdera que peu de points de calcul de  $\mu$ .

Le minimum  $\mu_{min}$  calculé sur le cône b), fortement résolu, se situe vers la zone centrale du cône, ce qui est plus correct. On a donc ici atténué les effets du maillage trop grossier sur la surface du cône, mais il reste que la forte asymétrie du maillage de la base produit des valeurs inhomogènes dans cette région. En effet, la fonction  $\mu$  discrétisée doit être appliquée sur des maillages assez réguliers avec des longueurs d'arêtes relativement homogènes de sorte que les voisinages de chaque sommet soient proches pour que le calcul de  $\mu$  soit correct.

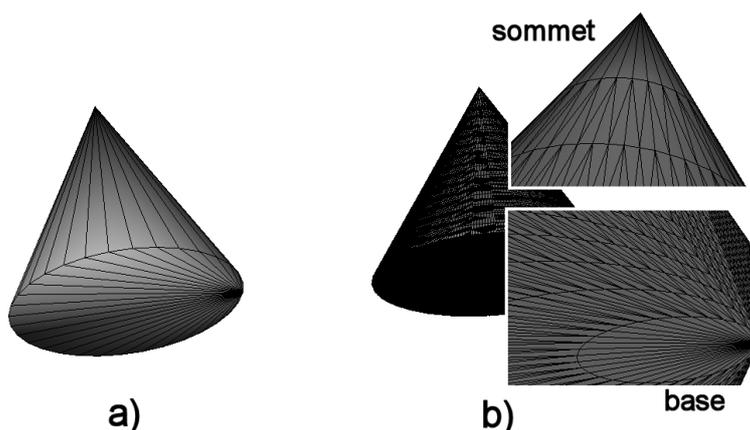


FIG. 2.12 – Cônes avec maillages différents : a) 33 triangles. b) 242 triangles.

La figure 2.14 présente les graphes MRG obtenus pour ces deux maillages très différents du cône. Les nœuds sont positionnés au barycentre des sections auxquelles

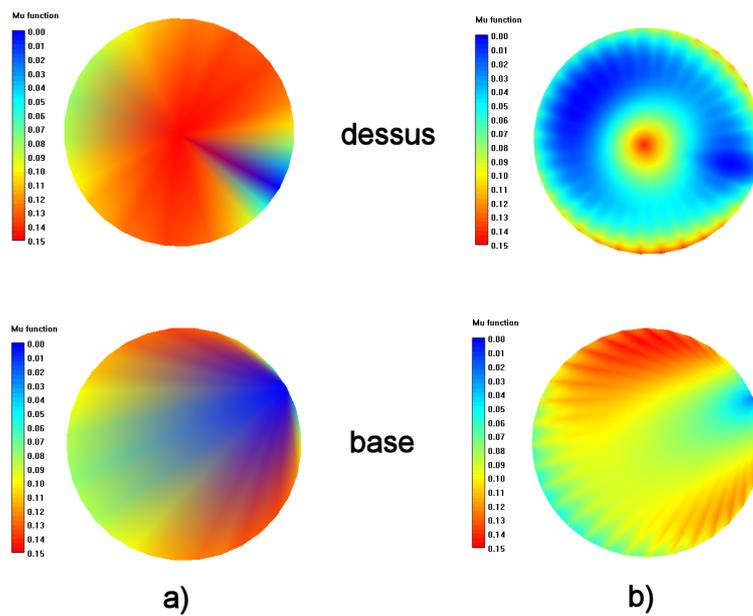


FIG. 2.13 – Valeurs de la fonction intégrale géodésique calculée sur les deux cônes de maillage différent (vues de dessus et vues de la base). a) les triangles générateurs ne sont pas subdivisés (le maillage est très irrégulier et anisotrope sur la base), le calcul de  $\mu$  est inexact et le minimum est excentré. b) les triangles générateurs sont suffisamment subdivisés, le calcul approche une valeur exacte sur la surface et le minimum se trouve bien dans la zone centrale du dessus du cône où on retrouve approximativement la symétrie de révolution de la surface conique. Les valeurs sur la base restent cependant inhomogènes à cause du maillage irrégulier trop asymétrique.

ils sont associés. La présence de la déviation des nœuds du graphe du cône a permis de constater que les valeurs de  $\mu$  calculées ne sont pas homogènes sur la surface et à la base.

On constate aussi que lorsque la résolution augmente le graphe du cône b tend vers un graphe plus conforme à ce que l'on souhaiterait obtenir (un graphe possédant un axe central de symétrie de révolution).

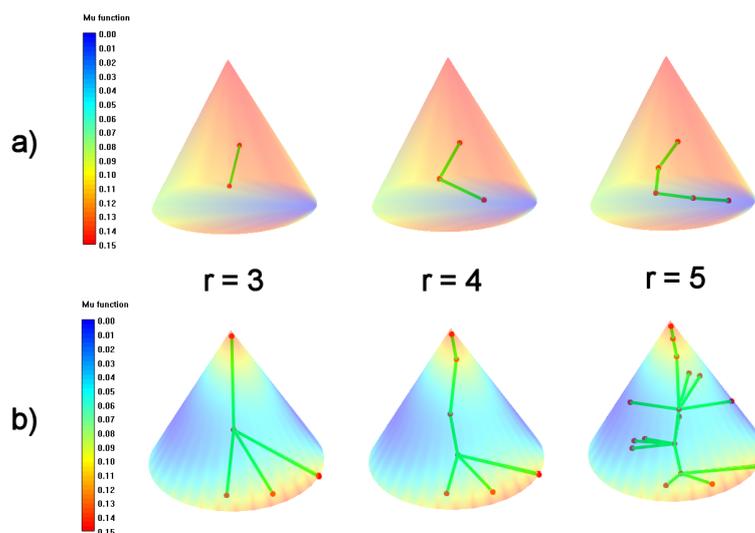


FIG. 2.14 – a) Graphes MRG ( $r = 3$  à  $5$ ) du cône à maillage très irrégulier de la figure 2.12 a). La déviation de la position des nœuds met en évidence l’inhomogénéité des valeurs de  $\mu$  sur la surface, liée à l’asymétrie du maillage de la base. b) Graphes MRG ( $r = 3$  à  $5$ ) du cône de la figure 2.12 b) à maillage subdivisé et 2-variété. On constate que le graphe possède un axe central de symétrie dès la résolution  $r = 3$  qui est plus proche de l’axe de symétrie de révolution du cône.

## 2.5 Bilan

Le graphe de Reeb multirésolution (MRG) permet d’obtenir une représentation de haut-niveau de type squelette des objets 3D. De plus l’aspect multirésolution permet de capturer les détails des modèles si la résolution des graphes est suffisamment élevée. Reposant sur la théorie de Morse, les graphes de Reeb sont une représentation de la topologie de la surface des objets 3D. Il est donc possible de retrouver le genre d’un modèle en étudiant les cycles de son graphe. Cependant la théorie assume que les surfaces sont des 2-variétés afin de pouvoir y calculer la fonction continue  $\mu$  qui permet de construire le graphe MRG. Les conditions de continuité impliquent que les maillages doivent être de bonne qualité et suffisamment résolu pour que les calculs de la fonction  $\mu$  soient exacts.

En s’appuyant sur la structure des MRG, les auteurs de [Hilaga et al., 2001] ont proposé un algorithme d’appariements de graphes dans le but d’estimer leur similarité. Nous présentons leur approche en détail dans le chapitre suivant, ainsi que nos contributions qui ont permis d’améliorer leurs résultats.

## CHAPITRE 3

# Indexation 3D avec les graphes de Reeb multirésolution augmentés

Profitant des avantages de la représentation multirésolution que procure la structure des MRG, nous avons injecté des attributs topologiques, géométriques et visuelles (couleur, texture) dans les nœuds des graphes, afin d'obtenir une description hiérarchique des modèles, tant locale que globale, et de parvenir à des calculs de similarités plus fins. Tous ces attributs sont librement pondérables afin de s'adapter au mieux aux requêtes d'un utilisateur. L'indexation des objets 3D avec notre approche consistera donc à insérer dans les bases de données des vecteurs caractéristiques décrivant l'ensemble des informations contenues dans les nœuds des graphes de Reeb à toutes les résolutions. Nous obtenons une représentation souple, multicritère et multirésolution que nous nommons graphe de Reeb multirésolution augmenté (aMRG).

Nous présentons dans la section 3.1 de ce chapitre la méthode d'appariement de graphes MRG proposée par [Hilaga et al., 2001], ainsi que l'extension des critères de cohérence topologique qui pallie certains problèmes d'appariements subsistants, et les attributs topologiques et géométriques supplémentaires pour améliorer l'appariement des nœuds des aMRG. Nous décrivons dans la section 3.2 le calcul de similarité entre les nœuds, et dans la section 3.3 les attributs géométriques et visuels que nous avons ajouté. Le calcul final de similarité entre deux objets est présenté à la section 3.4, et le principe de mise en correspondance d'ensemble de nœuds à la section 3.5. La section 3.6 précise le format des vecteurs caractéristiques contenant l'information des nœuds en binaire. Enfin nous présentons en section 3.7 un récapitulatif de nos contributions.

## 3.1 Appariements de graphes aMRG

L'approche repose sur les appariements de graphes de Reeb multirésolution proposée par [Hilaga et al., 2001]. La stratégie de comparaison de deux MRG consiste à appairer entre deux graphes, tous les nœuds ayant une configuration topologique similaire en terme de connexion avec les voisins à la même résolution et de relation parents/fils entre les niveaux de résolutions successifs. Dans un premier temps, on détecte les paires de nœuds respectant un ensemble de critères de cohérence topologique prédéfinis. Puis, à l'aide d'une fonction de discrimination, on retient parmi

ces paires de nœuds ceux qui serviront à estimer la similarité des deux MRG. L'approche est multirésolution. L'appariement débute au niveau de résolution  $r = 0$ , puis la procédure d'appariement est appliquée aux nœuds fils des nœuds appariés, et ainsi de suite jusqu'au niveau  $R$  ou jusqu'à ce qu'il n'y ait plus d'appariement de nœuds possible. Cette approche est une bonne alternative pour éviter le problème de NP-complétude inhérent à la recherche d'isomorphisme de sous-graphes [Foulds, 1992].

#### 3.1.1 Attributs des nœuds

Le graphe de Reeb multirésolution tel qu'il est construit ne contient aucune information géométrique (d'où le nom de "Topology Matching" donné à cette méthode par [Hilaga et al., 2001]). Cependant, pour le calcul de similarité, des informations autres que la simple topologie du graphe sont utiles. Ils se présentent sous la forme d'"attributs" associés à chaque nœud, c'est-à-dire un certain nombre de fonctions à valeurs réelles définies sur l'ensemble des nœuds. L'attribut d'un nœud  $n$  est un élément de  $\mathbb{R}^2$  et est noté  $\bar{n}$ . Dans [Hilaga et al., 2001], deux fonctions sont utilisées (mais on peut en envisager d'autres comme on le verra par la suite). Les deux fonctions utilisées sont notées  $a$  et  $l$  (donc  $\bar{n} = (a(n), l(n))$ ). Au niveau de résolution le plus élevé du graphe de Reeb multirésolution, elles sont définies ainsi :

**Définition 3** ( $a(n)$ ) Soient  $aire(n)$  l'aire de la surface du sous-ensemble de l'objet correspondant à  $n$ ,  $aire(S)$  l'aire totale de l'objet et  $R$  la résolution maximale du graphe de Reeb multirésolution, alors :

$$a(n) = \frac{aire(n)}{(R + 1) \cdot aire(S)} \quad (3.1)$$

**Définition 4** ( $l(n)$ ) Soit  $amp(m) = \max\mu(m) - \min\mu(m)$  l'amplitude d'un nœud  $m$ , où  $\max\mu(m)$  et  $\min\mu(m)$  sont le maximum et le minimum de  $\mu$  respectivement, sur le sous-ensemble de l'objet correspondant au nœud  $m$ , et  $R$  la résolution du graphe de Reeb multirésolution, alors :

$$l(n) = \frac{amp(n)}{(R + 1) \cdot \sum_{m \in F} amp(m)} \quad (3.2)$$

où  $F$  est l'ensemble des nœuds au niveau de résolution le plus élevé du graphe. Il faut noter qu'on n'a pas forcément  $amp(m) = \frac{1}{2R}$  mais plutôt  $amp(m) \leq \frac{1}{2R}$ .

En remarquant que :

$$aire(S) = \sum_{m \in F} aire(m) \quad (3.3)$$

on voit que les deux définitions sont très similaires. Ainsi, la somme des fonctions  $a$  ou  $l$  pour les nœuds d'un niveau de résolution fixé est égale à  $\frac{1}{R+1}$ , donc elle est égale à 1 sur un graphe entier.

Pour les autres niveaux de résolution, on définit  $\bar{n}$  itérativement en utilisant les nœuds-fils de  $n$  :

$$\begin{aligned}\bar{n} &= \sum_{m \in \{\text{fils de } n\}} \bar{m} \\ &= \left( \sum_{m \in \{\text{fils de } n\}} a(m), \sum_{m \in \{\text{fils de } n\}} l(m) \right)\end{aligned}\tag{3.4}$$

La signification de  $a(\cdot)$  est claire : il s’agit de la proportion de la surface totale “capturée” par le nœud. Celle de  $l(\cdot)$  l’est moins à première vue. D’après l’équation 3.2, pour un nœud à un niveau de résolution autre que le plus élevé, la valeur  $amp(n)$  de l’amplitude de  $\mu$  pour le nœud  $n$  est pondérée par la somme *totale* des valeurs de l’amplitude pour les nœuds qui sont ses descendants au niveau de résolution le plus élevé. Comme la valeur de  $l(\cdot)$  est comprise entre 0 et  $\frac{1}{2^R}$ , pour chacun des nœuds de niveau de résolution  $R$ ,  $l(\cdot)$  est en quelque sorte un indicateur de la “complexité” du graphe qui est la descendance du nœud considéré. Par “complexité” on entend par exemple le nombre de ramifications.

Nous proposons dans les sections suivantes de nouveaux attributs reposant sur la topologie du graphe permettant de mieux caractériser la complexité des graphes descendants de nœuds.

### 3.1.2 Stratégie

On considère que l’on a deux graphes de Reeb multirésolution  $M$  et  $N$  de même résolution, dont on veut calculer la similarité  $SIM(M, N)$ . S’ils n’ont pas la même résolution, on peut envisager de tronquer celui de résolution la plus élevée au niveau de résolution de l’autre. Mais dans ce cas, comme les attributs des nœuds n’ont pas été normalisés de la même façon (voir définitions 4 et 5) ; il faudrait donc les renormaliser.

La stratégie de comparaison de deux MRG consiste à constituer des paires de nœuds (un sur chaque graphe) qui “se ressemblent”. Pour déterminer quelle paire choisir, on cherche des paires de nœuds qui respectent des conditions de cohérence topologique. Une fois qu’on a constitué toutes les paires possibles qui respectent ces conditions, on calcule un indice de similarité entre les deux nœuds ainsi mis en correspondance qui est fonction de leur “ressemblance” et de leur “importance”. Cet indice sert également à départager les candidats à la mise en correspondance, si jamais plusieurs nœuds d’un graphe respectent les conditions de cohérence topologique avec un même nœud de l’autre modèle. Le résultat de ce calcul est un nombre réel compris entre 0 et 1 (un résultat égal à 1 indiquant que les deux modèles sont identiques).

### 3.1.3 Critères de cohérence topologique

Soient deux objets  $M$  et  $N$  dont on souhaite comparer les graphes de Reeb multirésolution. Les critères de cohérence topologique permettent de sélectionner les paires

### 3.1 Appariements de graphes aMRG

---

de nœuds  $(m, n)$ , avec  $m \in \text{MRG}(M)$  et  $n \in \text{MRG}(N)$ , qui seront éventuellement appariées dans la suite du processus. Les nœuds  $m$  et  $n$  sont cohérents topologiquement si :

**Condition 1 :** les parents de  $m$  et  $n$  ont été appariés ensemble au niveau de résolution précédent (sauf s'ils n'ont pas de parent, c'est-à-dire s'ils correspondent au nœud racine du niveau de résolution 0). On comprend avec cette condition que, conformément à la remarque faite précédemment, le calcul de similarité entre  $M$  et  $N$  se fera bien en commençant par le niveau le plus grossier puis passera successivement aux niveaux plus fins si à chaque fois la comparaison au niveau précédent a été concluante.

**Condition 2 :**  $m$  et  $n$  correspondent au même intervalle de fonction  $\mu_N$  (ce qui implique que l'on compare les nœuds de deux graphes de même résolution).

**Condition 3 :** Lorsque deux nœuds sont appariés, une même marque leur est attribuée (un numéro unique) et est propagée parallèlement dans les deux graphes à leurs voisins connexes en suivant respectivement le sens croissant et le sens décroissant des valeurs de  $\mu_N$  [Hilaga et al., 2001]. Les nœuds qui reçoivent cette marque sont au même niveau de résolution que les nœuds mis en correspondance. Chaque nœud va donc conserver une liste (initialement nulle) des marques qui lui parviennent.

Ainsi, si les nœuds  $m$  et  $n$  font chacun partie d'une "branche"<sup>1</sup> de graphe dont certains des nœuds ont déjà été appariés, alors ils doivent avoir reçu exactement les mêmes marques pour être appariés à leur tour (cf. Figure 3.1).

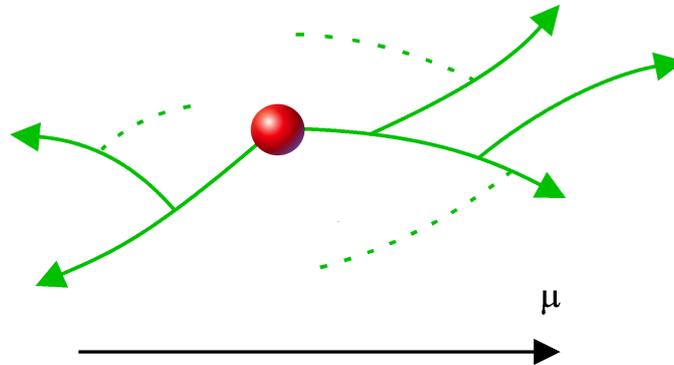


FIG. 3.1 – Propagation de la marque : la marque est propagée à partir du nœud apparié (en rouge). Les nœuds situés sur les branches en trait plein vont le recevoir, au contraire de ceux qui sont sur des branches en traits pointillés. Dans ce schéma la droite et la gauche correspondent au sens de croissance et décroissance de  $\mu$  respectivement.

---

<sup>1</sup>On appelle branche de graphe un ensemble de nœuds successivement reliés par seulement une arête entre eux. Deux branches sont appariées entre elles lorsque les nœuds leurs appartenant sont appariés entre eux.

### 3.1.4 Critère de cohérence topologique supplémentaire

Le simple énoncé des critères précédemment présentés est insuffisant pour assurer que deux branches restent en correspondance lorsqu'on passe à un niveau de résolution plus fin. En effet, l'appariement des branches n'est pas une information transmise aux fils, et il se peut donc que deux branches qui avaient été mises en correspondance ensemble ne le soient plus lorsque l'on passe au niveau de résolution supérieur suivant. Il apparaît nécessaire d'ajouter un critère supplémentaire d'héritage aux trois critères de cohérence topologique de [Hilaga et al., 2001] :

**Condition 4 :**  $m$  et  $n$  sont cohérents topologiquement si les parents de leurs voisins (lorsque  $m$  et  $n$  ont des voisins) ont été appariés ensemble au niveau de résolution précédent.

On appariera donc deux nœuds  $m$  et  $n$  du niveau de résolution  $r + 1$  s'ils ont deux voisins  $a$  et  $b$  dont les parents  $a'$  et  $b'$ , voisins directs des parents  $m'$  et  $n'$  de  $m$  et  $n$ , ont déjà été appariés ensemble au niveau de résolution  $r$ .

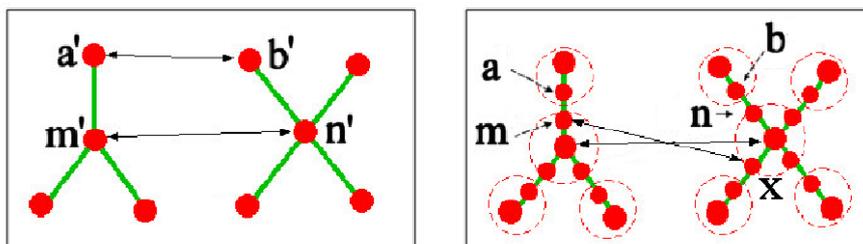


FIG. 3.2 – Gauche : deux graphes de Reeb au niveau de résolution  $r$  avec deux branches initialement appariées ( $a' - m'$  avec  $b' - n'$ ). Droite : les mêmes graphes au niveau de résolution  $r+1$ . Si deux nœuds fils sur des branches différentes sont appariés par erreur, par exemple  $m$  et  $x$ , une marque étant propagée le long des branches auxquelles ils appartiennent, les branches  $a - m$  et  $b - n$  ne pourront plus être correctement appariées car leurs nœuds ne pourront pas vérifier la condition 3 de cohérence topologique. Il est donc nécessaire de vérifier l'appariement des parents  $a'$  et  $b'$  des voisins  $a$  et  $b$  de deux nœuds  $m$  et  $n$  candidats à l'appariement afin de préserver le bon appariement des branches précédemment appariées au niveau de résolution inférieur.

La Figure 3.2 prend pour exemple deux graphes aux niveaux de résolution  $r$  et  $r + 1$ , à gauche et à droite respectivement. Au niveau de résolution  $r$ , les branches  $a' - m'$  et  $b' - n'$  comprenant les nœuds  $a'$  et  $m'$ , et  $b'$  et  $n'$  respectivement, sont appariées correctement. Si on s'en tient juste aux conditions de cohérence topologique de [Hilaga et al., 2001], on constate que n'importe quel nœud fils à la résolution  $r + 1$  du nœud  $m'$  peut être apparié à n'importe quel nœud fils à la résolution  $r + 1$  du nœud  $n'$  lorsqu'ils appartiennent au même intervalle de valeur de  $\mu$  : sur la figure 3.2 à droite, à la résolution  $r + 1$ ,  $m$  pourrait donc ne pas être apparié avec  $n$ , créant

ainsi une erreur d'appariement de branche, une marque étant propagée le long des branches auxquelles appartiennent les deux nœuds. Il est donc nécessaire de vérifier l'appariement des parents  $a'$  et  $b'$  des voisins  $a$  et  $b$  de deux nœuds  $m$  et  $n$  candidats à l'appariement afin de préserver le bon appariement des branches précédemment appariées au niveau inférieur.

L'ajout de cette condition supplémentaire assure donc qu'il n'y a pas de perte ou de mélange d'appariement de branches lorsque l'on passe à un niveau de résolution supérieur.

#### 3.1.5 Discrimination analytique pour l'appariement des nœuds

Les critères topologiques ne suffisent pas pour discriminer tous les nœuds candidats à l'appariement. Pour cela on définit une fonction *loss* qui permet de quantifier la différence de similarité entre deux nœuds  $m$  et  $n$ .

En se basant sur [Hilaga et al., 2001], on choisit la constante  $w$ , qui est le poids accordé à la fonction  $a(\cdot)$  dans l'attribut. La fonction  $imp(m)$  (“*imp*” pour importance) d'un nœud  $m$  est alors :

$$imp(m) = w \cdot a(m) + (1 - w) \cdot l(m). \quad (3.5)$$

La fonction *sim* renvoie l'indice de similarité entre deux nœuds  $m$  et  $n$  :

$$sim(m, n) = w \cdot \min(a(m), a(n)) + (1 - w) \cdot \min(l(m), l(n)). \quad (3.6)$$

On remarque que  $imp(m) = sim(m, m)$ , et  $sim(m, n) = sim(n, m)$ .

Pour former une paire  $(m, n)$  lorsqu'il y a plusieurs candidats (tous cohérents topologiquement), on utilise les deux indices précédemment présentés. Soient  $m_1, \dots, m_i$ ,  $i$  nœuds du graphe  $M$  cohérents topologiquement avec le nœud  $n$  du graphe  $N$ . On introduit la fonction de discrimination  $loss(m_i, n)$  (“perte”) qui mesure la décote moyenne de la similarité globale par rapport à une situation où  $m_i$  et  $n$  seraient mis en correspondance avec eux-mêmes :

$$\begin{aligned} loss(m, n) &= \frac{imp(m) + imp(n)}{2} - sim(m, n) \\ &= w \cdot \left( \frac{a(m) + a(n)}{2} - \min(a(m), a(n)) \right) \\ &\quad + (1 - w) \cdot \left( \frac{l(m) + l(n)}{2} - \min(l(m), l(n)) \right) \\ &= w \cdot \frac{|a(m) - a(n)|}{2} + (1 - w) \cdot \frac{|l(m) - l(n)|}{2}. \end{aligned} \quad (3.7)$$

Pour augmenter la robustesse des appariements on prend également en compte le voisinage immédiat. On note  $adj(I, n)$ , l'ensemble des nœuds reliés à  $n$  par une arête et qui correspondent à un intervalle de  $\mu$  égal à  $I$ .

Lorsqu'un nœud  $n$  est en présence de plusieurs voisins, on pose les relations :

$$a(adj(I, n)) = \sum_{m \in adj(I, n)} a(m) \quad \text{et} \quad l(adj(I, n)) = \sum_{m \in adj(I, n)} l(m). \quad (3.8)$$

Si l'intervalle de  $\mu$  correspondant à  $n$  (et donc aussi à  $m_1, \dots, m_i$ ) est  $I_{r,k}$ , on définit  $mat(m, n)$  (“*mat*” pour “*matching*”) :

$$\begin{aligned} mat(m, n) = & -loss(m, n) \\ & -loss(adj(I_{r,k-1}, n), adj(I_{r,k-1}, m)) \\ & -loss(adj(I_{r,k+1}, n), adj(I_{r,k+1}, m)) \end{aligned} \quad (3.9)$$

bien sûr, si  $k = 0$ , on ne prend pas en compte le deuxième terme, et si  $k = 2^r - 1$ , on ne prend pas en compte le troisième.

Finalement, on choisit parmi les nœuds  $m_1, \dots, m_i$  celui qui maximise  $mat(\cdot, n)$ .

Afin d'améliorer cette étape, nous proposons d'introduire des attributs discriminants supplémentaires et d'écrire la fonction  $loss$  de la manière suivante :

$$loss(m, n) = \sum_{i=0}^{N_f} \alpha_i d_i(f_i(m), f_i(n)), \quad (3.10)$$

où  $N_f$  est le nombre d'attributs  $f_i$ ,  $d_i$  une distance associée à  $f_i$ , et  $\alpha_i$  sont des constantes de pondération des attributs  $f_i$  tels que  $\sum_i \alpha_i = 1$ . Cette formule procure beaucoup de liberté dans l'appariement des nœuds, et en particulier les constantes  $\alpha_i$  permettent de pondérer l'influence de chaque attribut  $f_i$  selon le type de requête désirée. Le choix des fonctions  $f_i$  est explicité dans les sections suivantes. Le choix des constantes  $\alpha_i$  et des distances  $d_i$  peut être effectué de manière empirique comme décrit au chapitre 5.

### 3.1.6 Ajout d'attributs topologiques

Pour améliorer l'appariement des nœuds entre les graphes aMRG nous proposons d'utiliser la topologie du graphe, d'exploiter l'aspect multirésolution des aMRG, et de profiter de l'orientation des arêtes (induites par les valeurs de  $\mu$ ).

Pour cela, nous introduisons au nœud  $n$  à la résolution maximale  $r = R$  quatre paramètres entiers suivants :

**Up<sub>N</sub>(n)** : le nombre de nœuds voisins reliés à  $n$  et appartenant à l'intervalle de  $\mu$  directement supérieur,

### 3.1 Appariements de graphes aMRG

---

**Down<sub>N</sub>(n)** : le nombre de nœuds voisins reliés à  $n$  et appartenant à l'intervalle de  $\mu$  directement inférieur,

**Up<sub>E</sub>(n) ∈ {0, 1}** : balise indiquant si  $n$  est un nœud terminal “supérieur”  
( $Up_E(n) = 1$ ) ou non ( $Up_E(n) = 0$ ),

**Down<sub>E</sub>(n) ∈ {0, 1}** : balise indiquant si  $n$  est un nœud terminal “inférieur”  
( $Down_E(n) = 1$ ) ou non ( $Down_E(n) = 0$ ).

Par définition nous avons les relations suivantes au niveau  $R$  :

- si  $Up_N(n) = 0$  alors  $n$  est un nœud terminal “supérieur”, et dans ce cas  $Up_E(n) = 1$  et  $Down_E(n) = 0$ ,
- si  $Down_N(n) = 0$  alors  $n$  est un nœud terminal “inférieur”, et dans ce cas  $Up_E(n) = 0$  et  $Down_E(n) = 1$ ,
- si  $Up_N(n) = Down_N(n) = 0$  alors le graphe de résolution maximale  $R$  est représenté par un seul nœud racine et  $Up_E(n) = Down_E(n) = 1$ .

On possède ainsi une information topologique locale en chaque nœud au niveau de résolution le plus fin  $R$ .

Au niveau de résolution  $r < R$ , nous cumulons les informations topologiques (cf. Figure 3.3). Soit le nœud  $m$  à la résolution  $r < R$ , on a :

- $Up_N(m) = \sum_{n \in \{\text{fils de } m\}} Up_N(n)$ ,
- $Down_N(m) = \sum_{n \in \{\text{fils de } m\}} Down_N(n)$ ,
- $Up_E(m) = \sum_{n \in \{\text{fils de } m\}} Up_E(n)$ ,
- $Down_E(m) = \sum_{n \in \{\text{fils de } m\}} Down_E(n)$ .

Avec ces définitions, à la racine  $r = 0$  de l'aMRG on obtient  $Up_N = Down_N = \#$  d'arêtes à la résolution maximale  $r = R$ . Les informations contenues dans  $Up_N$  et  $Down_N$  sont donc redondantes et manquent de pertinences. Pour cela, on propose la règle supplémentaire suivante à la résolution  $R$  :

- si  $n$  est un nœud terminal alors on pose  $Up_N(n) = 0$  et  $Down_N(n) = 0$ ,
- et on obtient la relation  $\#$  d'arêtes  $\geq \max(Up_N + Down_E, Down_N + Up_E)$ .

Le fait de cumuler l'information topologique permet de caractériser la complexité des graphes descendants de chaque nœud. En effet, appelons  $\sigma = Up_N(m) + Down_N(m) + Up_E(m) + Down_E(m)$ . Si à la résolution  $r < R$ ,  $\sigma \gg 1$  alors  $m$  a beaucoup de nœuds descendants (fils et petits-fils), et si  $\sigma \rightarrow 1$  alors  $m$  a peu de descendants et devient terminal si  $\sigma = 1$ .

Notre étude a révélé que ces paramètres sont des indicateurs pertinent de la topologie des graphes. Par exemple, l'étude de l'aMRG à la résolution  $r = 5$  d'un cube à 5158 triangles donne (cf. cube déformé Figure 3.4) :

- $Up_N(m) = 22$ ,
- $Down_N(m) = 20$ ,
- $Up_E(m) = 8$  correspondant aux 8 sommets du cube,
- $Down_E(m) = 6$  correspondant aux 6 faces du cube.

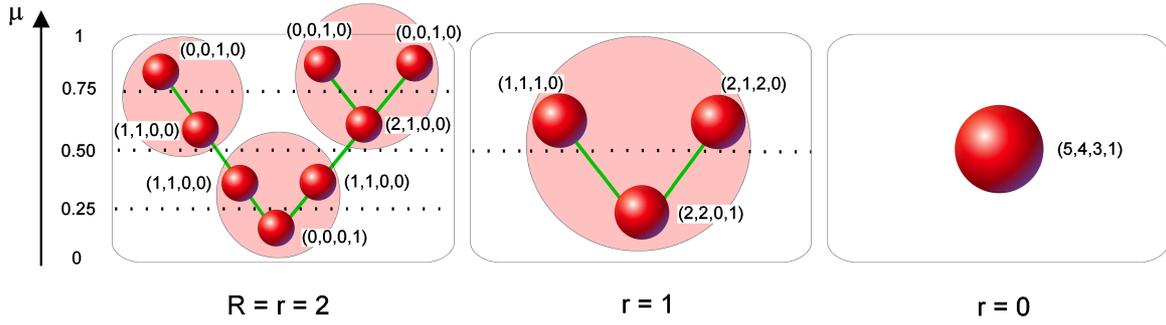


FIG. 3.3 – Gauche : au niveau de résolution  $r = R = 2$ , on introduit le vecteur  $(Up_N, Down_N, Up_E, Down_E)$  relatif à la topologie locale de chaque nœud. Milieu : au niveau de résolution inférieur  $r = R - 1 = 1$ , l'information topologique est cumulée. Droite : à la racine ( $r = 0 =$  tous les attributs topologiques sont cumulés).

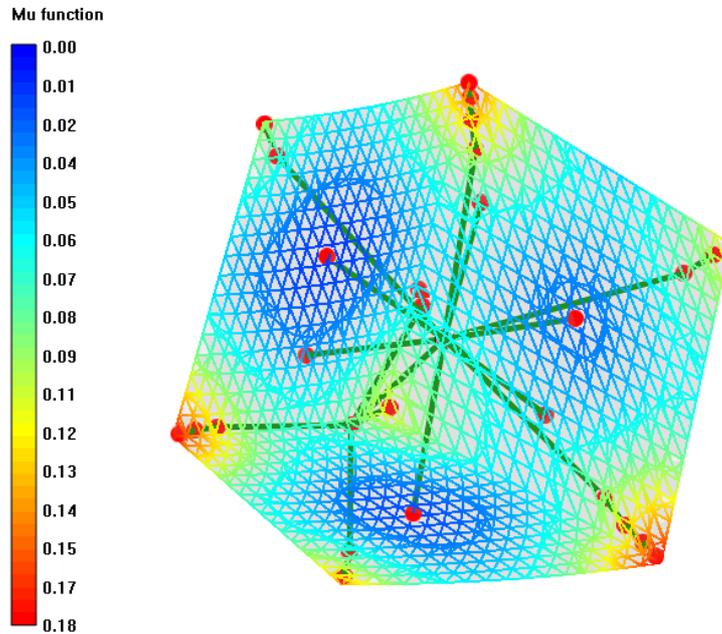


FIG. 3.4 – aMRG d'un cube déformé à  $R = 5$ .  $\mu$  est minimal aux centres des faces et maximal aux sommets.

Les paramètres topologiques permettent de dénombrer les arêtes du graphe, le nombre total d'arêtes étant lié au nombre d'intervalles de  $\mu$  qui segmentent la surface de l'objet. Le choix de la fonction  $\mu$  et particulièrement sa normalisation jouent donc ici un rôle important. En effet, la normalisation de  $\mu$  par  $\mu_{max}$  (cf. Eq 2.6) réduit la dynamique de  $\mu_N \subset [0, 1]$  et donc le nombre d'intervalles de  $\mu$  possible sur l'objet, tandis que la normalisation de  $\mu$  par  $\mu_{max} - \mu_{min}$  procure une dynamique complète sur  $[0, 1]$  et donc d'atteindre le nombre maximal  $2^R$  d'intervalles de  $\mu$  lors du partitionnement dichotomique (pour la construction du graphe multirésolution au niveau le plus

fin  $r = R$ ).

Ainsi, par exemple en utilisant la fonction  $\mu$  géodésique normalisée par  $\mu_{max}$ , il est possible de se servir des paramètres topologiques pour effectuer la classification suivante :

**Objets étoilés (main, modèle humain) :** plus l’objet est étoilé, plus il a de nœuds terminaux “supérieurs” et plus  $Up_E$  est grand. On a alors la relation suivante entre nœud terminaux supérieurs et inférieurs :  $Up_N > Down_N$ .

**Objets allongés (cylindre) :**  $\mu_{max}$  se trouve aux deux extrémités et  $\mu_{min}$  est au voisinage du barycentre. Le graphe est filiforme et on a  $Up_N = Down_N$ ,  $Up_E = 2$  et  $Down_E = 1$ . De plus, plus l’objet est allongé, plus  $Up_N = Down_N \rightarrow 2^R$ .

**Objets ramassés (sphère) :**  $\mu_{max} \sim \mu_{min}$  la dynamique de  $\mu_N$  normalisé par  $\mu_{max}$  est très faible. Le graphe ne contient donc que très peu de nœud, même à la résolution  $R = 5$ .  $Up_N$ ,  $Down_N$ ,  $Up_E$  et  $Down_E$  sont petits.

**Objets de type polyédrique convexe :** les centres des facettes planes et les sommets du polyèdre correspondent à des valeurs de  $\mu$  minimales et maximales respectivement. Si la résolution maximale  $R$  est suffisante,  $Up_E$  et  $Down_E$  vont tendrent vers le nombre de sommets et de facettes planes respectivement, comme dans l’exemple précédent du cube déformé (cf. Figure 3.4).  $Up_E \sim \#$  sommets et  $Down_E \sim \#$  facettes.

Le pouvoir discriminant des attributs topologiques  $Up_N$ ,  $Down_N$ ,  $Up_E$ , et  $Down_E$  comme fonction discriminante  $f_i$  dans la fonction *loss* 3.10 va donc nous permettre de terminer le processus d’appariements des nœuds de manière intuitive et topologique.

L’efficacité de ces attributs topologiques fait l’objet d’une étude expérimentale au chapitre 5.

#### 3.1.7 Ajout d’attributs géométriques

Les appariements de graphes topologiques comme en [Hilaga et al., 2001] ne sont pas toujours satisfaisants. Par exemple sur un modèle humain un bras équivaut topologiquement à une jambe (ce sont des membres reliés au tronc), alors que nous souhaiterions obtenir des appariements naturellement corrects dans le cadre de notre étude. De plus les critères de cohérence topologique sont tels que si deux nœuds ne sont pas mis en correspondance à un certain niveau de résolution alors qu’ils auraient dû l’être, il n’y a aucune possibilité pour que les fils le soient à un niveau plus fin. Il est donc crucial que les nœuds soient appariés au mieux dès les premiers tests de cohérence. Pour cela nous proposons d’apporter des informations de localisation aux nœuds. Un nœud  $m$  correspondant à un ensemble de triangles connexes peut être localisé au barycentre  $m(x, y, z)$  de ces triangles (cf. Figures 2.5 et 2.7). Nous choisissons une représentation en coordonnées sphérique  $(\mathbf{r}, \theta, \varphi)$  :

- $\mathbf{r}(m) \geq 0$  est la distance du nœud au centre du repère,
- $\theta(m) = \arctan\left(\frac{y}{x}\right) \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$  l’azimut,
- $\varphi(m) = \arccos\left(\frac{z}{r}\right) \in [0, \pi]$  l’élévation.

En prenant l'exemple d'un modèle humain (cf. Figure 3.5) dont l'axe  $x$  traverserait le corps de bas en haut, l'axe  $y$  de gauche à droite et l'axe  $z$  d'arrière en avant, on observe que  $\theta$  permet d'identifier la gauche de la droite,  $\varphi$  de distinguer les bras des jambes et  $r$  de trancher lorsqu'il y a ambiguïté afin par exemple de ne pas confondre un bras levé avec une tête.

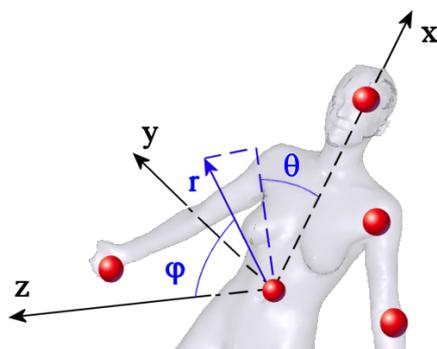


FIG. 3.5 – Localisation des nœuds en coordonnées sphériques.

Il devient donc impératif que les objets soient bien orientés par rapport aux directions principales ( $x, y, z$ ). Ainsi on peut utiliser les axes naturels des objets rapportés au centre de gravité avec des coordonnées normalisées, ou une technique classique d'alignement spatial comme l'ACP classiquement utilisé pour les modèles 3D [Vranic et al., 2001].

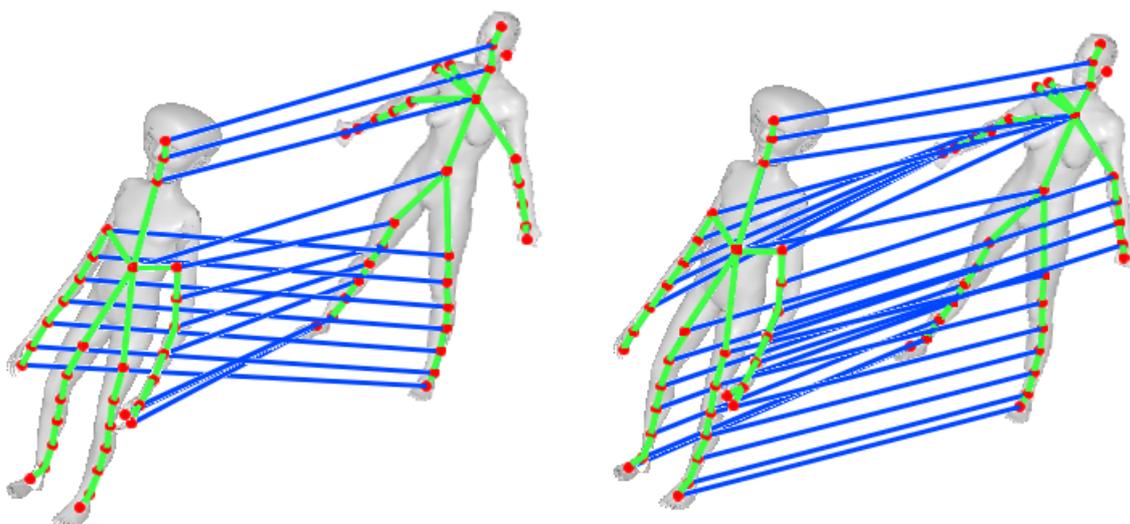


FIG. 3.6 – Gauche : sans l'information géométrique, les jambes peuvent être appariées aux bras car ils sont topologiquement équivalents. Droite : en ajoutant l'information géométrique, tous les appendices sont bien appariés.

La Figure 3.6 illustre l'appariement de deux modèles humanoïdes et le gain obtenu en introduisant ces attributs géométriques de localisation de nœuds dans la formulation de la fonction *loss* précédemment définie à l'Eq. 3.10.

Les informations relatives à la position des nœuds ont permis d'améliorer considérablement l'appariement des nœuds des graphes. Les résultats illustrant les performances obtenues sont présentés au chapitre 4 (Résultats Expérimentaux).

## 3.2 Procédure d'appariements des nœuds

Soient deux objets  $M$  et  $N$ , l'algorithme d'appariement des nœuds de leur graphe est le suivant :

### Initialisation :

Créer :

1. une liste de nœuds  $Objet_M$ , initialement remplie avec le nœud du niveau de résolution 0 de l'objet  $M$ ,
2. une liste de nœuds  $Objet_N$ , initialement remplie avec le nœud du niveau de résolution 0 de l'objet  $N$ ,
3. une liste de nœuds  $Candidats$ , initialement vide,
4. une liste des paires de nœuds appareillées  $PairesOK$ , initialement vide.

**Boucle :** Tant que  $Objet_M$  ou  $Objet_N$  ne sont pas vides :

{

**Etape 1 :** Choisir dans  $Objet_M$  le nœud  $m$  qui maximise la fonction  $imp(\cdot)$  (cf. Eq. 3.5) et l'enlever de  $Objet_M$ .

**Candidats :**

1. Initialiser  $Candidats$  à liste vide,
2. insérer dans  $Candidats$  tous les nœuds de  $Objet_N$  (sans les enlever) qui sont cohérents topologiquement avec  $m$ .

**Condition :** Si  $Candidats$  n'est pas vide :

{

**Etape 2 :**

1. Trouver dans  $Candidats$  le nœud  $n$  qui maximise  $mat(m, \cdot)$  (cf Eq. 3.9), l'enlever de  $Objet_N$ ,
2. ajouter la paire  $(m, n)$  à  $PairesOK$ ,
3. ajouter à  $Objet_M$  tous les fils de  $m$ , et
4. ajouter à  $Objet_N$  tous les fils de  $n$ .

}

}

**Fin :** Le processus s'arrête lorsqu'il n'y a plus de nœuds à appairer, c.-à-d. lorsque  $Objet_M$  ou  $Objet_N$  sont vides. A la fin de l'appariement,  $PairesOK$  contient la liste

des nœuds appariés.

Dans [Hilaga et al., 2001] les auteurs proposent d'utiliser  $w = 0.5$  avec les attributs  $a(\cdot)$  et  $l(\cdot)$  dans les fonctions  $imp(\cdot)$  et  $loss(\cdot)$  définies aux Eq. 3.5 et 3.7.

Nos études nous ont montré que l'utilisation des seuls attributs topologiques dans la fonction  $loss(\cdot)$  donne de meilleurs appariements que la formulation originale de l'Eq. 3.7. L'appariement des nœuds est alors totalement topologique et indépendant de l'orientation des objets.

Ainsi en pratique, la formulation de  $imp(\cdot)$  est inchangée, mais pour la fonction  $loss$  on utilise, pour deux nœuds  $m$  et  $n$  :

$$loss(m, n) = \alpha_1 |Up_N(m) - Up_N(n)| + \alpha_2 |Down_N(m) - Down_N(n)| \\ + \alpha_3 |Up_E(m) - Up_E(n)| + \alpha_4 |Down_E(m) - Down_E(n)|. \quad (3.11)$$

avec  $\sum_{i=1}^4 \alpha_i = 1$ .

De plus si tous les objets ont une pose normalisée, alors on peut encore améliorer l'appariement des nœuds en introduisant les attributs géométriques de localisation (cf. Section 3.1.7), la fonction  $loss$  peut s'écrivant alors :

$$loss(m, n) = \alpha_1 |Up_N(m) - Up_N(n)| + \alpha_2 |Down_N(m) - Down_N(n)| \\ + \alpha_3 |Up_E(m) - Up_E(n)| + \alpha_4 |Down_E(m) - Down_E(n)| \\ + \alpha_5 |\mathbf{r}(m) - \mathbf{r}(n)| + \alpha_6 \frac{|\theta(m) - \theta(n)|}{2\pi} + \alpha_7 \frac{|\varphi(m) - \varphi(n)|}{2\pi}, \quad (3.12)$$

avec  $\sum_{i=1}^7 \alpha_i = 1$ .

### 3.3 Calcul de similarité de nœuds

Après avoir obtenu la liste des nœuds appariés *PairesOK*, des fonctions de similarités vont permettre de préciser le processus de reconnaissance des formes entre graphes en quantifiant la similarité des nœuds. On calcule dans un premier temps une fonction de similarité sur chaque paire de nœuds appariés. Cette fonction a été fortement enrichie par un ensemble de mesures comparée à celle proposée par [Hilaga et al., 2001] afin d'affiner de manière significative le calcul de la similarité entre deux objets. Ces mesures peuvent être pondérées au goût des utilisateurs selon qu'ils souhaitent accorder ou non plus d'impact à telle ou telle mesure.

Le calcul de la nouvelle fonction de similarité entre deux nœuds  $sim(m, n)$  a la

même formulation que la fonction *loss* 3.10 :

$$sim(m, n) = \sum_{i=0}^{N_{f'}} \lambda_i d'_i(f'_i(m), f'_i(n)), \quad (3.13)$$

où  $N_{f'}$  est le nombre d'attributs  $f'_i$ ,  $d'_i$  est une distance associée à  $f'_i$  et  $\lambda_i$  sont des constantes de pondération des attributs  $f'_i$  telles que  $\sum_i \lambda_i = 1$ .

Afin d'améliorer les résultats de requêtes par similarité de formes obtenus avec les MRG classiques, et d'élargir le contexte à des requêtes privilégiant d'autres caractéristiques (géométriques, colorimétriques, etc.), nous avons fusionné des propriétés géométriques, globales et locales, et des propriétés visuelles de la section de l'objet  $S(m)$  associée en chaque nœud  $m$ . Ces propriétés sont les suivantes :

- $(\mathbf{r}(m), \theta(m), \varphi(m))$  déjà défini en section 3.1.7,
- $v(m)$ , volume relatif de  $S(m)$ ,
- $cord(m)$ , mesure statistique de l'étendue de la section  $S(m)$ ,
- $courb(m)$ , mesure statistique de l'indice de forme de Koenderink [Zaharia and Prêteux, 2002] calculé localement sur la surface  $S(m)$ ,
- $hough(m)$ , mesure statistique caractérisant l'orientation des normales des triangles du maillage associé à  $S(m)$ ,
- $color(m)$ , mesure statistique de la couleur associée à  $S(m)$  lorsqu'une carte de texture est disponible.

Le choix de ces attributs a été guidé par la littérature sur la recherche par le contenu 3D. Le choix des distances  $d'_i$  et des valeurs des paramètres  $\lambda_i$  a été établi de manière empirique et est discuté au chapitre 5. Nous précisons dans la section suivante ce qui a motivé l'ajout de ces attributs géométriques et visuels, puis nous les définissons successivement.

## 3.4 Ajout d'attributs géométriques et visuels

### 3.4.1 Motivations

Lorsque la topologie des objets n'est pas suffisante pour caractériser la forme des objets, il est alors nécessaire d'extraire de l'information "ailleurs". La figure 3.7 présente en a) les graphes MRG de deux objets dont les graphes sont similaires (quasi linéaires). Il est alors difficile de bien différencier les deux modèles avec la seule approche par MRG. Cependant, l'ajout d'informations géométriques locales en chaque nœud, par exemple en b) les indices de forme [Koenderink, 1990], permet d'avoir une caractérisation beaucoup plus fine des nœuds du graphe et donc de discriminer les modèles 3D.

L'enrichissement des nœuds des graphes par de l'information géométrique paraît d'autant plus nécessaire lorsque les graphes MRG sont peu significatifs, en particulier pour les objets compacts où la dynamique des valeurs de  $\mu$  est faible (cf. Figure 3.8 : buste, pavé, cube, sphère, à la résolution  $r = 5$ ). Notre approche vise donc à déterminer les attributs de nœuds les plus pertinents pour décrire la géométrie des objets.

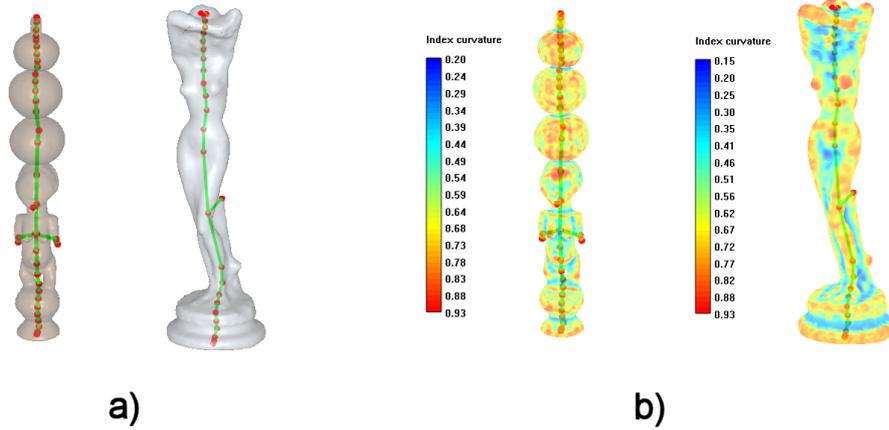


FIG. 3.7 – a) Les deux modèles semblent avoir des graphes MRG similaires. b) L'information introduite dans chaque nœud (ici l'indice de forme) permet de mieux quantifier les différences entre les nœuds des graphes.)

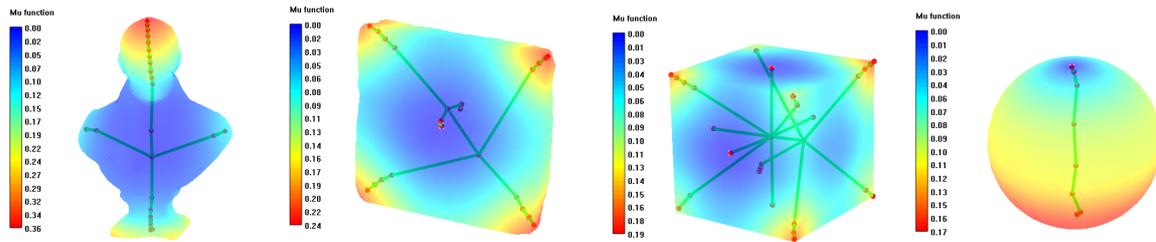


FIG. 3.8 – Exemples de graphes MRG à la résolution  $r = 5$  d'objets à géométrie compacte. Les graphes en soit sont assez peu discriminants.

### 3.4.2 Calcul du volume

Soit  $[\mu_1, \mu_2]$  l'intervalle de la fonction  $\mu$  associé au nœud  $m$ . On définit le volume relatif par :

$$v(m) = \frac{1}{R + 1} \cdot \frac{vol(m)}{vol(S)} \quad (3.14)$$

où  $vol(S)$  représente le volume total du modèle délimité par la surface  $S$ , et  $vol(m)$  le volume correspondant à la section  $S(m)$  associée au nœud  $m$ .  $vol(S)$  est calculé en additionnant les volumes signés des tétraèdres formés par les triangles orientés du maillage et le centre de gravité de l'objet [Zhang and Chen, 2001].  $S(m)$  étant la section de la surface de l'objet associé au nœud  $m$ ,  $vol(m)$  est calculé en additionnant :

1. le volume des tétraèdres formés par les triangles de la section  $S(m)$  associés à  $m$  et le barycentre de ces triangles,
2. le volume des tétraèdres formés par ce barycentre, une arête du maillage correspondant au bord  $\mu = \mu_1$  de la surface associée à  $m$ , et le barycentre des points

associés à ces arêtes,

3. le volume symétrique au précédent pour le bord  $\mu = \mu_2$  (cf. Figure 3.9).

Le facteur  $\frac{1}{R+1}$  dans l'Eq. 3.14 correspond aux niveaux de résolution  $r = 0, \dots, R$  du graphe aMRG( $M$ ), ce qui donne la relation  $\sum_{m \in \text{aMRG}(M)} v(m) = 1$ . On remarque que cette méthode de calcul de volume n'est valide que si les triangles sont orientés, et les maillages fermés, ce qui est le cas pour la plupart des modèles formant notre base de données.

En couplant cet attribut avec l'aire relative du nœud  $a(m)$ , on obtient une caractérisation supplémentaire de la section de l'objet associée au nœud  $m$  : nous pouvons introduire le rapport sans dimension  $ratio(m) = 36.\pi.vol(m)^2/aire(m)^3$ , où  $aire(m)$  est l'aire de la section volumique du modèle correspondant à  $S(m)$ . Ce rapport est particulièrement discriminant au niveau de résolution  $r = 0$  pour certaines formes élémentaires (cf. Chapitre 5). Par exemple nous avons pour la sphère  $ratio = 1$  car c'est une surface minimale, pour le cube  $ratio = \pi/6 \sim 0.524$ . Pour les formes filiformes comme celles de la figure 3.7 nous avons  $ratio = 0.126$  et  $ratio = 0.131$  pour l'objet africain et la déesse respectivement.

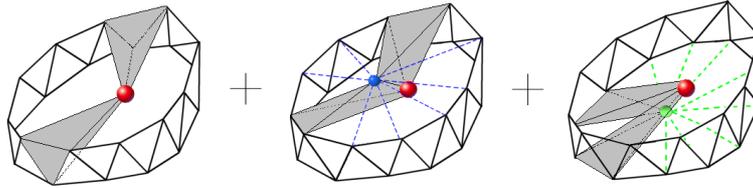


FIG. 3.9 – Volume associé à un nœud obtenu en sommant les volumes des trois types de tétraèdres contenus dans la section.

#### 3.4.3 Mesures des cordes

La mesure des cordes introduite dans [Paquet and Rioux, 1997] correspond à l'histogramme normalisé des distances des barycentres des triangles au centre du repère. Appliqué à un nœud  $m$  du niveau de résolution le plus fin, l'histogramme  $cord_L(m)$  est constitué des distances entre le barycentre de la section de l'objet  $S(m)$  associée au nœud  $m$ , et les barycentres des triangles appartenant à  $S(m)$ . Cette mesure statistique permet d'obtenir des informations sur l'étendue d'un ensemble de triangles et est un indicateur sur la géométrie du maillage associé au nœud (cf. Figure 3.10). Chaque distance calculée est pondérée par l'aire relative du triangle (aire normalisée par l'aire totale de l'objet). Les longueurs des cordes de  $m$  sont aussi normalisées par rapport à leur longueur maximale. On adapte ainsi la dynamique de l'histogramme pour utiliser la totalité des intervalles pour chaque nœud.

A l'histogramme des longueurs des cordes nous ajoutons comme dans [Paquet and Rioux, 1997] les deux histogrammes des angles des cordes avec les deux premiers axes

de l'analyse en composante principale,  $cord_1(n)$  et  $cord_2(n)$  respectivement.

Au niveau de résolution  $r = 0$  le barycentre du nœud unique, racine du graphe aMRG est le centre de gravité de l'objet et les histogrammes obtenus sont égaux à ceux obtenus par l'approche globale sur l'objet entier [Paquet and Rioux, 1997].

L'utilisation de méthodes statistiques utilisant des histogrammes pour décrire un objet, comme les cordes de [Paquet and Rioux, 1997] ou les distributions de formes de [Osada et al., 2002], sont réputées peu efficaces lorsqu'il s'agit de décrire des formes complexes, comme nous pouvons le constater au Chapitre 4. En revanche, dans notre contexte et à partir d'un certain niveau de résolution  $r$ , les nœuds sont en majorité associés localement à des formes simples de types couronnes ou paraboloides, pour lesquelles ces mesures géométriques aisées à calculer peuvent être suffisamment discriminantes.

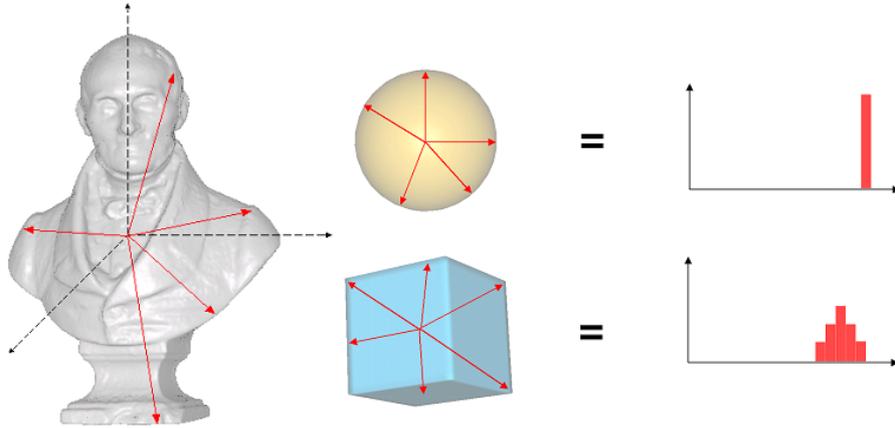


FIG. 3.10 – *La mesure des cordes est une caractérisation de la géométrie des objets.*

### 3.4.4 Mesure des courbures

Les propriétés géométriques de la surface peuvent être caractérisées localement par une mesure de la courbure locale. On utilise pour cela l'indice de forme de Koenderink [Koenderink, 1990, Zaharia and Prêteux, 2002] (cf. Figure 3.11). Au point  $\mathbf{P}$ , l'indice de forme  $I(\mathbf{P})$  s'écrit :

$$I(\mathbf{P}) = \frac{1}{2} - \frac{1}{\pi} \operatorname{atan}\left(\frac{k_2(\mathbf{P})}{\sqrt{|k_2^2(\mathbf{P}) - k_1(\mathbf{P})|}}\right), \quad (3.15)$$

$k_1(\mathbf{P})$  étant la courbure Gaussienne et  $k_2(\mathbf{P})$  la courbure moyenne.

La courbure Gaussienne en un sommet  $\mathbf{P}$  du maillage est obtenue par le calcul du

déficit angulaire [Schroeder et al., 1997] :

$$k_1(\mathbf{P}) = \frac{2\pi - \sum_j \theta_j}{\frac{1}{3} \sum_j A_j}, \quad (3.16)$$

$\theta_j$  étant l'angle au sommet  $\mathbf{P}$  du  $j^{\text{e}}$  triangle du 1<sup>er</sup> anneau de  $\mathbf{P}$ , et  $A_j$  est l'aire de ce  $j^{\text{e}}$  triangle.  $k_1(\mathbf{P})$  est homogène à l'inverse d'une surface.

La courbure moyenne  $k_2(\mathbf{P})$  au sommet  $\mathbf{P}$  est calculée selon la formule classique [Schroeder et al., 1997], en sommant la moitié des angles  $\phi_i$ ,  $i = 1, \dots, N_e$  des dièdres formés par les normales des paires de triangles consécutifs dans le 1<sup>er</sup> anneau de  $\mathbf{P}$ ,  $N_e$  étant le nombre de triangles adjacents à  $\mathbf{P}$ . Les angles sont pondérés avec le tiers de l'aire moyenne  $\bar{A}_i$  des deux faces correspondantes.  $k_2(\mathbf{P})$  est homogène à l'inverse d'une distance.

$$k_2(\mathbf{P}) = \frac{1}{N_e} \sum_i \frac{\frac{1}{2}\phi_i}{\frac{1}{3}\bar{A}_i}. \quad (3.17)$$

L'indice de forme  $I(\mathbf{P})$  est compris entre 0 et 1, et est donc bien adapté pour des statistiques normalisées. L'histogramme des indices de forme  $curv(m)$  est calculé en chaque nœud  $m$  au niveau de résolution le plus fin  $R$ . Il est calculé pour tous les sommets du maillage associé à  $m$ , chaque contribution d'un point  $\mathbf{P}$  étant pondérée par le tiers des aires relatives (normalisées par rapport à l'aire totale de l'objet) des triangles adjacents à  $\mathbf{P}$  et appartenant à la section associée à  $m$ .

Les histogrammes des nœuds parents sont obtenus en sommant les histogrammes des nœuds fils. Au niveau de résolution  $r = 0$ , l'histogramme est équivalent au descripteur de spectre de forme 3D calculé sur l'objet entier [Zaharia and Prêteux, 2002].

Cependant, un des inconvénients majeurs de cette approche est qu'elle est inapplicable directement sur des modèles dont le maillage est trop bruité. En effet, l'estimation locale de la courbure requiert une bonne continuité de la surface et donc un maillage de bonne qualité. Il peut alors être nécessaire d'appliquer un filtrage des données avant d'appliquer les calculs de courbure en cas de données bruitées.

Dans le cadre de notre étude nous avons régularisé les maillages problématiques avec une combinaison de filtrage Laplacien (pour le lissage) et bi-harmonique (pour la préservation de la courbure) [Esteban, 2004] (cf. Figure 3.12). Au sommet  $\mathbf{v}$  de voisinage de 1<sup>er</sup> anneau  $\mathcal{N}_1(\mathbf{v})$  les opérateurs sont :

**Filtre Laplacien  $\Delta$  :**

$$\Delta \mathbf{v} = \left( \sum_{i \in \mathcal{N}_1(\mathbf{v})} \frac{\mathbf{v}_i}{m} \right) - \mathbf{v}, \quad (3.18)$$

où  $\mathbf{v}_i$  sont les voisins de  $\mathbf{v}$  et  $m$  le nombre total des voisins (valence).

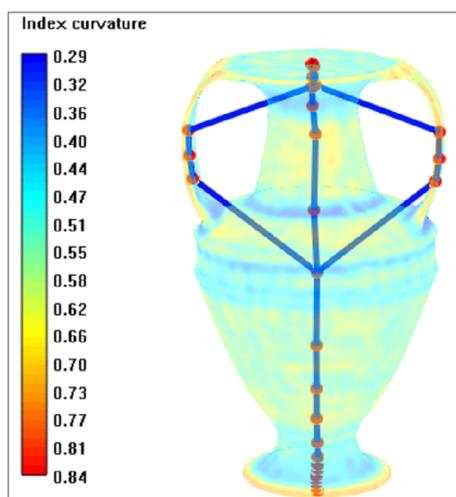


FIG. 3.11 – *Indice de forme de Koenderink calculé en chaque sommet du maillage et utilisé comme attribut de nœud.*

**Filtre bi-harmonique  $\Delta^2$  :**

$$\Delta^2 \mathbf{v} = \Delta(\Delta \mathbf{v}) = \left( \sum_{i \in \mathcal{N}_1(\mathbf{v})} \frac{\Delta \mathbf{v}_i}{m} \right) - \Delta \mathbf{v}. \quad (3.19)$$

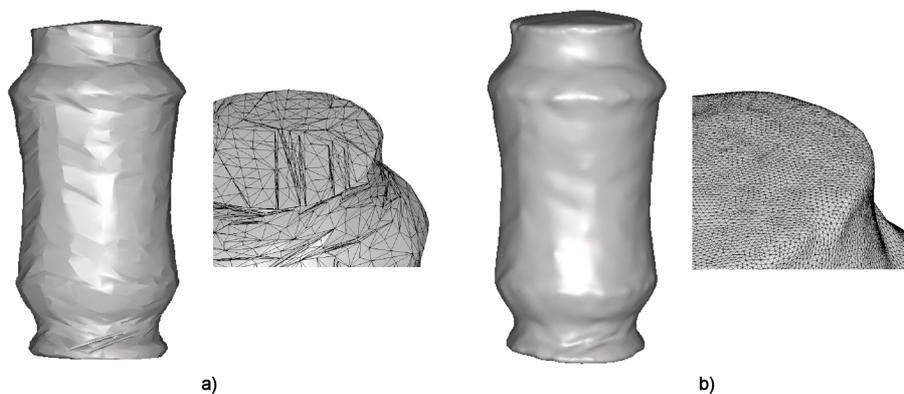


FIG. 3.12 – *a) Modèle ayant un maillage trop bruité pour l'estimation locale de la courbure. b) Modèle filtré pour le calcul des courbures.*

Ils peuvent être itérés plusieurs fois en cas de maillages fortement bruités.

### 3.4.5 Descripteur de Hough 3D

Le descripteur de Hough 3D a été proposé dans [Zaharia and Prêteux, 2002] pour succéder au descripteur de Spectre de forme 3D. Un des avantages de cette approche

est sa robustesse face aux déformations locales du maillage (connectivité du maillage, décimation).

Le descripteur de Hough 3D  $hough(m)$  d'un nœud  $m$  caractérise la distribution spatiale des plans contenant les triangles du maillage associé à  $m$ . Le plan  $\Pi$  de chaque triangle est paramétré en coordonnées sphériques  $(r, \theta, \phi)$ ,  $r$  représentant la distance du centre du nœud  $m$  à  $\Pi$ , et  $(\theta, \phi)$  les angles d'azimut et d'élévation exprimant la direction de la normale du plan  $\Pi$  dans le repère de l'objet, la normale étant orientée vers le demi-espace ne contenant pas l'origine :  $r \geq 0, \theta \in [-\pi, \pi]$ , et  $\phi \in [0, \pi]$ .

Afin d'obtenir une partition homogène de l'orientation des normales on utilise les faces d'un octaèdre subdivisé projetées sur la sphère unité. On élimine ainsi les problèmes d'inhomogénéité aux pôles liés au paramétrage en coordonnées sphériques classiquement obtenu par un échantillonnage uniforme des axes  $\theta$  et  $\phi$ . L'attribut  $hough(m)$  du nœud  $m$  est un histogramme 3D accumulant les triplets  $(r, \theta, \phi)$  caractérisant les plans des triangles de  $m$ . La contribution apportée par chaque triangle est pondérée par l'aire relative du triangle. Au niveau de résolution  $r = 0$ , l'histogramme est équivalent au descripteur de Hough 3D calculé sur l'objet entier [Zaharia and Prêteux, 2002].

#### 3.4.6 Quantification des couleurs

Lors d'une recherche d'objets 3D dans une base de donnée, il est aussi intéressant de pouvoir faire des requêtes sur la couleur et la texture des modèles si ceux-ci en disposent. L'alliance de critères de forme et de couleur peut s'avérer hautement discriminant. On peut regretter d'ailleurs que l'information couleur soit encore peu présente dans les bases de données d'objets 3D.

Nous utilisons pour la couleur les informations contenues dans les cartes de textures associées aux modèles 3D. Ces informations 2D viennent naturellement s'ajouter aux attributs 3D précédemment présentés (cf. Figure 3.13). Notre caractérisation de la couleur est effectuée simplement au moyen des trois histogrammes normalisés dans l'espace  $R, V$  et  $B$ . D'autres espaces colorimétriques du type CIE-Lab pourraient être préférés si on dispose de carte de texture de bonne qualité. Il est évidemment possible de choisir d'autres approches pour caractériser les couleurs et la texture comme décrit en section 1.9. Ces histogrammes sont initialement déterminés pour chacun des nœuds du aMRG au niveau de résolution le plus fin  $r = R$ , pour chaque triangle appartenant à la section  $S(m)$  associée au nœud  $m$ , on collecte la contribution des texels<sup>2</sup> sur lesquels se projette le triangle dans la carte de texture. La contribution aux histogrammes de chaque texel est proportionnelle à la taille relative du texel par rapport à la taille du triangle projeté et la somme de ces contributions est égale à la surface relative du triangle. Puis de manière itérative et hiérarchique pour tous les niveaux de résolution

---

<sup>2</sup>un texel (*texture element*) est un pixel élémentaire de la carte de texture.

inférieurs :

$$color(m) = \frac{1}{R+1} \sum_{n \in \{fils\ de\ m\}} color(n), \quad (3.20)$$

avec à chaque niveau de résolution  $r \in [0, R]$ ,  $\sum_{m \in res(r)} color(m) = 1$ .



FIG. 3.13 – Exemple de maillage à droite avec la carte de texture associée à gauche.

### 3.5 Similarité d'objets

Soient deux objets  $M$  et  $N$  dont on désire calculer la similarité. On pose :

$$SIM(M, N) = 1 - \sum_{(m,n) \in PairesOK} sim(m, n), \quad (3.21)$$

où *PaireOK* est la liste des couples de nœuds appariés définie en section 3.2,  $sim(m, n)$  est la fonction de similarité entre les nœuds  $m$  et  $n$  définie à l'Eq. 3.13.  $SIM(M, N)$  est contenu dans  $[0, 1]$  car défini comme une combinaison linéaire d'histogrammes normalisés de valeurs relatives (aire, volume, ratio) avec des termes de pondération barycentriques. La fonction  $SIM$  a les propriétés suivantes :

- Identité :  $SIM(M, M) = 0$ ,
- Positivité :  $SIM(M, N) \geq 0$ ,
- Symétrie :  $SIM(M, N) = SIM(N, M)$ .

$SIM$  est une semi-métrique [Barros et al., 1996]. L'inégalité triangulaire ne peut être vérifiée car le calcul de similarité de deux graphes aMRG dépend du nombre de paires de nœuds appariés, et ce nombre peut varier lorsque l'on compare différents objets à un même objet.

### 3.6 Mise en correspondance d'ensemble de nœuds

Un problème délicat de mise en correspondance d'ensemble de nœuds a été soulevé par Hilaga *et al.* dans [Hilaga et al., 2001]. Il se produit quand deux modèles comportent des ramifications fortement similaires (et qu'on aimerait donc logiquement mettre en correspondance) mais qui apparaissent pour un certain niveau de résolution de part et d'autre d'une transition de fonction  $\mu$  passant d'un intervalle  $I_k$  à l'intervalle suivant  $I_{k+1}$  (cf. Figure 3.14).

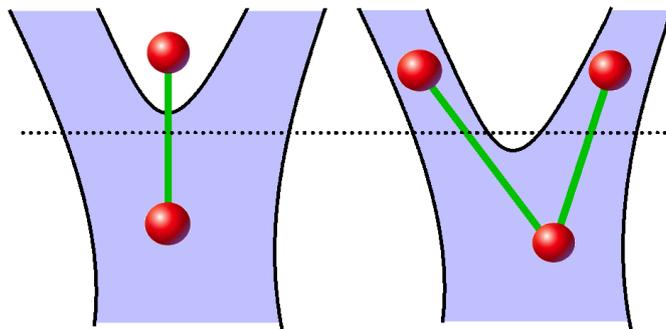


FIG. 3.14 – *Problème de transition d'intervalle de  $\mu$  : la ligne de tirets représente une telle transition. Deux paires de modèles sont représentées, ainsi que leurs graphes de Reeb, à un niveau de résolution correspondant à cette transition d'intervalle de  $\mu$ . Les modèles 3D sont très similaires. On aimerait pouvoir mettre en correspondance le nœud supérieur du graphe de gauche avec les deux nœuds supérieurs du graphe de droite.*

Pour pallier ce problème, il serait intéressant de pouvoir mettre en correspondance non pas une paire de nœuds mais plutôt une paire d'ensemble de nœuds. En fait, en modifiant l'algorithme d'appariement de nœuds présenté en section 3.2, on peut mettre en correspondance un nœud avec un ensemble de nœuds de la manière suivante :

- le premier élément  $m$  de la paire mise en correspondance est toujours un nœud (cf. section 3.2),
- le second élément de la paire peut-être un ensemble de nœuds : l'étape 2 de 3.2 est modifiée. Ainsi, si plusieurs nœuds sont candidats à la mise en correspondance, on cherche à constituer des ensembles avec la condition suivante : *pour former un ensemble, les nœuds doivent être voisins d'un même nœud qui se situe à leur résolution* (par exemple, dans la Figure 3.14, les deux nœuds supérieurs à droite sont bien voisins du nœud inférieur : la paire qu'ils forment peut donc être candidate à la mise en correspondance avec le nœud supérieur à gauche). L'attribut d'un type donné d'un ensemble de nœuds est calculé par simple cumul des attributs de même type de ces nœuds, comme en section 3.2. Les ensembles de nœuds formés ainsi sont alors candidats au même titre que les nœuds seuls, et on choisit parmi tous ces candidats celui qui maximise la fonction  $mat(m, \cdot)$  (cf. Section 3.2). Il est donc toujours possible de choisir un seul nœud comme

second élément.

Cet aspect permet de donner plus de souplesse au 2<sup>e</sup> critère de cohérence topologique qui n'autorise que des appariements de nœuds appartenant au même intervalle de valeurs de  $\mu$ . En effet, il permet de pallier les éventuels problèmes de bords (pouvant apparaître aux embranchements) qui sont liés au découpage de la surface par la fonction  $\mu_N$ .

### 3.7 Indexation des graphes

L'indexation des graphes peut être effectuée en sauvegardant la structure du graphe aMRG ainsi que les attributs des nœuds. Les informations peuvent être stockées dans des fichiers en binaire par soucis de gain de place, la taille de ces fichiers variant linéairement en fonction du nombre de nœuds du graphe. Nous avons implémenté la structure de fichier suivante :

Taille (en octets)	Nom	Description
$2 * \text{sizeof}(\text{char}) = 2$		Chaîne de caractères égale à "BE" ou "LE", selon que le système générant les données est <i>big-endian</i> ou <i>little-endian</i> .
$\text{sizeof}(\text{char}) = 1$	res	Résolution maximale $R$ du graphe de Reeb Multirésolution Augmenté.
$\text{sizeof}(\text{char}) = 1$	nbAttr	Nombre de mesures associées à chaque nœud.
$(\text{res} + 1) * \text{sizeof}(\text{int}) = (\text{res} + 1) \cdot 4$	nbNodes	Tableau contenant le nombre de nœuds pour chaque niveau de résolution.
$\text{sizeof}(\text{int}) = 4$	NbEdges	Nombre d'arêtes.
$\text{nbNodes} * \text{nbAttr} * \text{sizeof}(\text{float}) = \text{nbNodes} * \text{nbAttr} * 4$		Tous les nœuds, décrits par leurs attributs.
$2 * \text{nbEdges} * \text{sizeof}(\text{int}) = 2 * \text{nbEdges} * 4$		Toutes les arêtes, décrites par les indices des deux nœuds extrémités.

La taille du fichier est donc :  $4 * [3 + \text{res} + \text{nbNodes} * \text{nbAttr} + \text{nbEdges} * 2]$  octets. Notons que cette formulation ne considère que le cas simple d'attributs de taille 1 flottant. Dans la pratique on a inclus les attributs suivants pour chaque nœud :

Taille (en octets)	Nom	Description
sizeof(float)=4	$a$	Aire relative.
sizeof(float)=4	$l$	Longueur d'intervalle de $\mu$ .
sizeof(int)=4	$i$	Indice de l'intervalle de $\mu$ .
sizeof(float)=4	$r$	Distance du nœud au centre du repère.
sizeof(float)=4	$\theta$	Azimut.
sizeof(float)=4	$\varphi$	Élévation.
sizeof(float)=4	$volume$	Volume relatif.
sizeof(float)=4	$ratio$	$ratio = 36.\pi.vol^2/aire^3$ .
sizeof(int)=4	$Up_N$	Nombre de voisins supérieurs.
sizeof(int)=4	$Down_N$	Nombre de voisins inférieurs.
16*sizeof(float)=64	$cordL$	Histogramme de longueurs de cordes.
16*sizeof(float)=64	$cord1$	Histogramme d'angles des cordes avec le 1 <sup>er</sup> axe.
16*sizeof(float)=64	$cord2$	Histogramme d'angles des cordes avec le 2 <sup>e</sup> axe.
16*sizeof(float)=64	$curv$	Histogramme d'indices de forme.
4*32*sizeof(float)=512	$curv$	Descripteur de Hough 3D.

Ainsi pour chaque nœuds, les descripteurs sont codés sur 808 octets, ou  $808 + 192 = 1000$  octets si on ajoute 3 histogrammes couleurs *RVB* de 16 flottants.

Par exemple, le vecteur caractéristique de l'objet *botijo*<sup>3</sup> dont le format binaire *.tri* faile une taille de  $2252Ko$  (cf. Figure 3.15), a été calculé avec une résolution maximale de graphe aMRG de  $R = 5$ . Avec tous les attributs précédemment cités ce vecteur caractéristique fait  $90Ko$ . Il est clair que la taille du vecteur caractéristique n'est pas optimale. Il peut être compressé en zip sur  $44Ko$ . Mais il est préférable de conserver la structure initiale du fichier qui est claire et flexible et qui permet facilement d'enrichir les nœuds avec de nouveaux attributs.

De plus, bien que la taille du descripteur soit conséquente, les tests reportés au chapitre 4 ont démontré que les temps de recherches dans des base de données n'étaient pas excessifs. La procédure d'appariement de nœuds multirésolution, par son approche de haut en bas, permet de rejeter très rapidement les appariements de modèles qui ne comportent pas de structure de graphe similaire. En effet, la discrimination est effectuée à l'étape d'appariement topologique des nœuds avant l'étape de calcul de similarité.

<sup>3</sup>Le modèles *botijo* est composé de 82340 triangles et 41160 sommets.

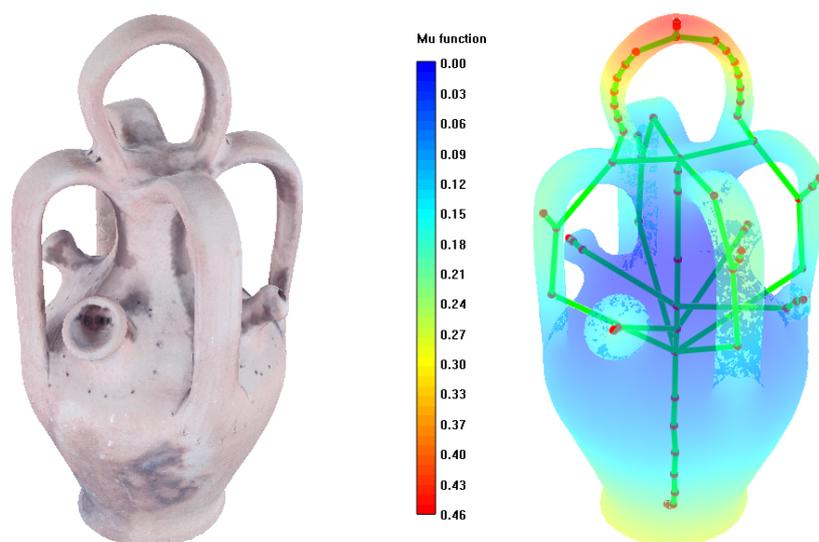


FIG. 3.15 – *Objet botijo et son graphe de Reeb à la résolution 5.*

### 3.8 Bilan

Nous avons proposé une extension de la méthode de calcul de similarité entre modèles 3D présentée dans [Hilaga et al., 2001]. En particulier, nous avons augmenté la robustesse de l'appariement des graphes de Reeb multirésolution en étendant les conditions de cohérence topologique. Les attributs qui décrivent localement la topologie des graphes et que nous proposons en 3.1.6 pour discriminer les nœuds se révèlent très pertinents et riches d'informations. Ils permettent notamment de discriminer facilement les structures de graphes complexes des structures simples. Nous pensons cependant que leurs propriétés ne sont pas encore optimalement exploités et peuvent faire encore l'objet de travaux de recherche. D'autre part, nous avons enrichi les nœuds de divers types d'attributs : histogrammes de cordes, histogramme d'indice de forme, descripteur de Hough 3D, et histogrammes couleur. Ces informations sont fusionnées dans chaque nœud pour le calcul de similarité avec des termes de pondération pour chaque type d'attributs. L'utilisateur est donc libre de choisir quel jeu d'attributs il souhaite privilégier. L'encodage des attributs contenus dans les nœuds dans les bases de données nous permet d'aboutir à un schéma d'indexation 3D multirésolution contenant simultanément des informations globales et locales, et de nature très variée (topologique, géométrique, et couleur). Les descripteurs ont des grandes tailles par rapport aux approches classiques (808 octets pour les aMRG contre 64 octets pour un histogramme de 16 flottants) mais les performances en terme de temps de requête ne sont pas pénalisées grâce à l'approche multi-résolution.

Dans le chapitre suivant, nous allons étudier les performances de l'aMRG et les comparer à celles des autres méthodes à l'aide d'outils visuels et quantitatifs qui sont classiquement utilisés par la recherche de classe en reconnaissance des formes sur des bases de données.



## CHAPITRE 4

# Résultats expérimentaux

Dans ce chapitre nous comparons notre approche par graphe de Reeb multirésolution augmenté avec différentes méthodes d'indexation 3D afin de mettre en évidence son efficacité. Nous avons pour cela constitué trois bases de données de taille et de complexité différentes, et implémenté plusieurs outils statistiques d'évaluation de performance (images, scores, etc.). Il ressort de nos tests que l'approche que nous proposons est la plus efficace pour retrouver les classes d'objets 3D similaires, et retourne les résultats les plus intuitivement homogènes. La section 4.1 présente les trois bases de données de 116, 266 et 567 objets que nous avons utilisées pour nos tests. Les deux premières classes comportent 16 et 27 classes respectivement. La troisième comporte 301 objets non classés, ajoutés aux 266 objets de la deuxième classe. La section 4.2 présente les différentes techniques que nous avons employées pour évaluer les différentes méthodes d'indexation 3D testées. Nous présentons en section 4.3 les résultats obtenus sur la première base de données de chaque méthode avec chaque distance implémentée, afin d'associer aux descripteurs les distances qui donnent les meilleurs résultats. La section 4.4 présente les résultats obtenus avec les aMRG en comparaison avec les autres méthodes. Nous terminons avec un bilan des expériences menées à la section 4.5.

### 4.1 Bases de données

Afin d'évaluer les performances de l'approche des aMRG pour retrouver les objets et classes d'objets de formes similaires, et de comparé les aMRG aux méthodes existantes, nous avons construit trois bases de données de taille et complexité différentes.

La première base de données est composée de 116 modèles classés en 16 classes (cf. Figure 4.1 et Figure A.1 en Annexe A). Les objets sont tous différents, en forme, en topologie ou en résolution de maillage. La base de données comporte des classes composées de formes primitives comme des cubes, des sphères, des cônes, des tores ou double tores qui ont été conçues interactivement avec un outil de modélisation 3D. D'autres classes sont composées de modèles reconstruits à partir d'objets réels [Esteban and Schmitt, 2004, Goodall et al., 2004], et d'autres ont été téléchargés sur l'Internet [3dcafe, 2005].

## 4.1 Bases de données

Les maillages sont de différentes résolutions. Les maillages de plus haute résolution comptent  $\sim 150\,000$  triangles (statues), et les maillages de faible résolution  $\sim 500$  triangles (pièce d'échecs). Quelques maillages sont de qualité médiocre, contiennent des trous ou même ne sont pas manifolds <sup>1</sup>. Tous les objets de chaque classe ont des formes similaires, excepté ceux de la classe n° 16 qui est composée de 9 objets d'aspect très différent dont 7 sont des statues de musées.

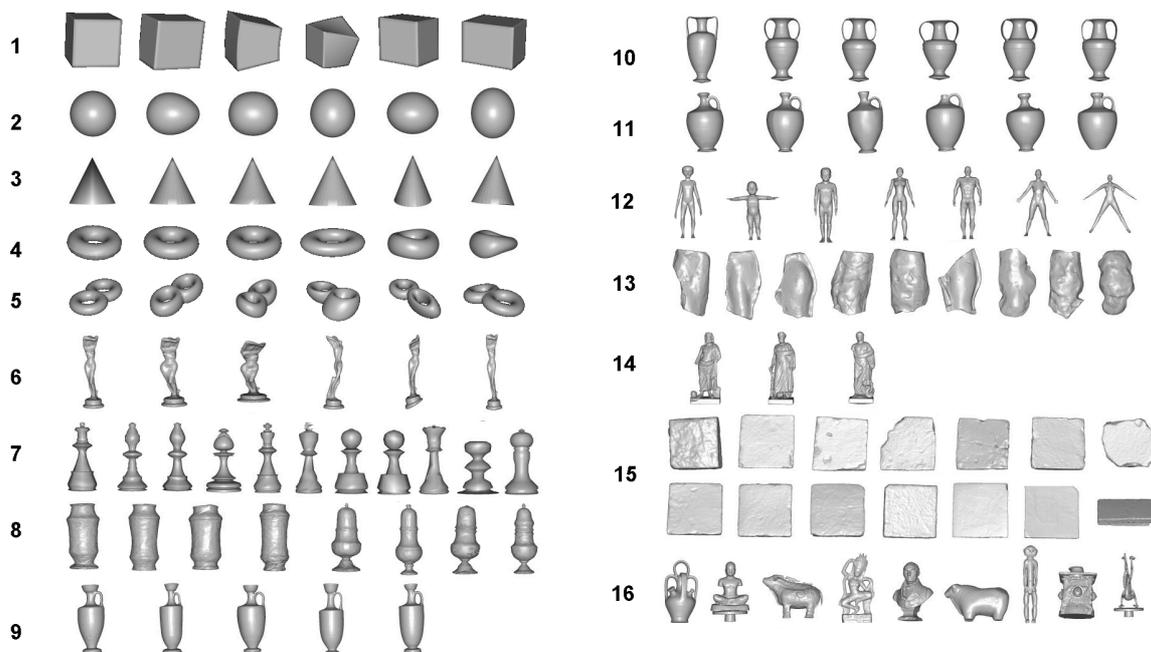


FIG. 4.1 – Base de données de test composée de 116 objets répartis en 16 classes.

Certaines classes sont très éloignées de toutes les autres mais restent assez proches entre elles et peuvent être facilement confondues par un mauvais descripteur. Par exemple les trois classes de vases n° 9, 10 et 11 peuvent être confondues, ou bien la classe de modèle humain n° 12 avec les classes de statues n° 6 et n° 14. La classe n° 6 est constituée du même modèle initial ayant subi différentes déformations 3D. La première base de données est utilisée comme base d'entraînement pour optimiser le jeu de facteurs de pondération des attributs des aMRG, et pour choisir la distance qui produit les meilleurs résultats de classification pour chaque méthode testée.

La deuxième base de données est composée de 266 objets classés en 27 classes (cf. Figure 4.2 et Figure 4.2 en Annexe A). Elle inclut la base d'entraînement précédente sans la classe de statues disparates n° 16. Mais en plus, des objets 3D de musées de très haute résolution, avec plus de 400 000 triangles, ont été ajoutés. L'intérêt de cette base de données réside dans la complexité de sa classification.

<sup>1</sup>Si le modèle n'est que très localement non manifold, le calcul des descripteurs peut tout de même s'effectuer sur l'ensemble du maillage. Dans ce cas l'objet a été conservé dans la base de données.



FIG. 4.2 – Base de données de 266 objets répartis en 27 classes.

## 4.1 Bases de données

---

En effet, certaines des 12 classes additionnelles sont très proches. Il y a par exemple différentes classes de fragments de moules. Les expériences menées sur cette base de données vont nous permettre de tester la précision des descripteurs 3D.

Cette deuxième base de données a ensuite été complétée avec un ensemble de 301 objets, tous différents, pour former la troisième base de données de 567 objets (cf. Figure 4.3 et Figure A.3 en Annexe A). Les objets 3D qui ont été ajoutés proviennent encore de numérisation de modèles de musées. Certains objets ont une forme assez proche d'objets classés, mais restent tout de même suffisamment intuitivement différents pour ne pas être classés avec les autres. Cette base de données peut être interprétée comme étant une base de données très bruitée et devrait permettre de mettre en évidence l'habileté de chaque descripteur 3D à retrouver les classes d'objets similaires.

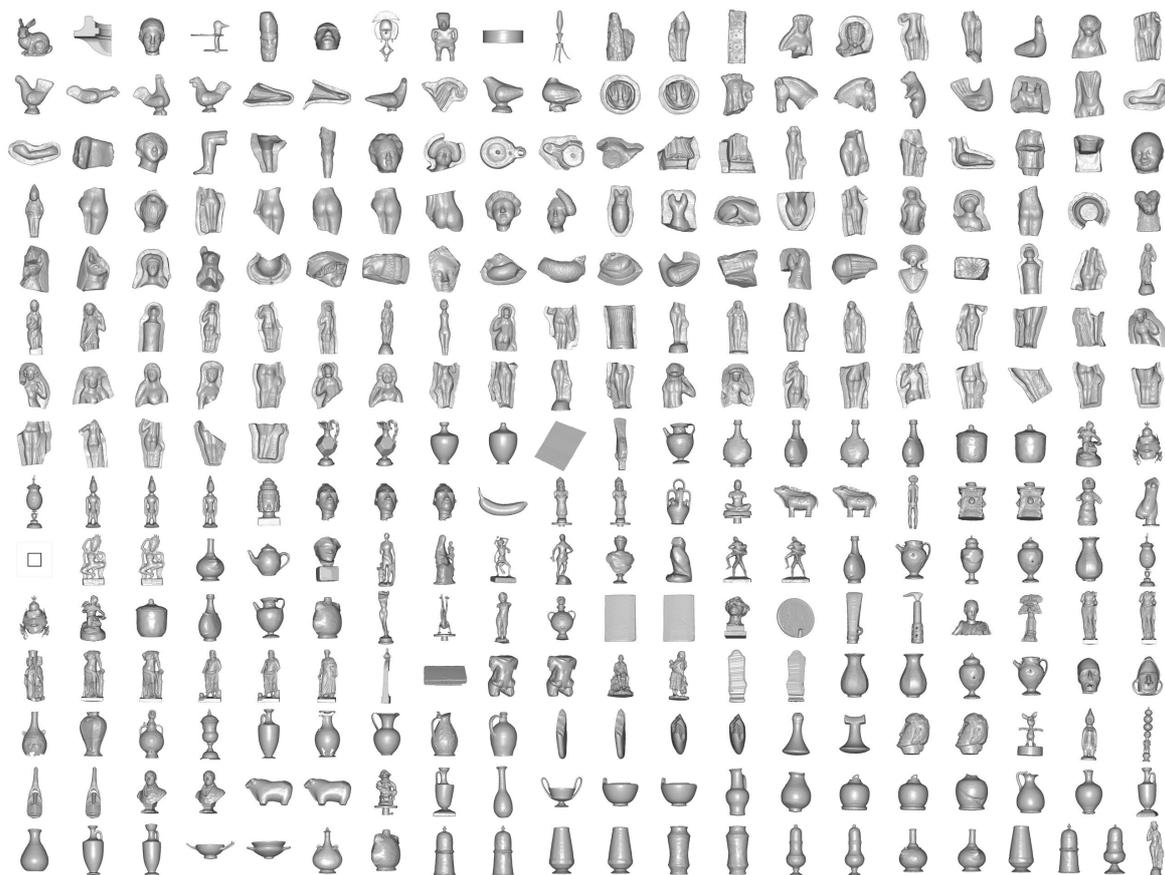


FIG. 4.3 – Base de données de 301 objets non classés.

## 4.2 Techniques d'évaluation

Pour chaque objet, la structure du graphe aMRG ainsi que ses attributs sont codés dans un fichier binaire (cf. Section 3.7). La taille du fichier dépend de la complexité de l'objet et croît linéairement en fonction du nombre de nœuds. Ces fichiers sont sauvegardés dans une base de données et représentent les *vecteurs caractéristiques* des objets.

Les évaluations pour les requêtes ont été effectuées sur un PC Pentium 4 à 2 GHz avec 1 Go de RAM. Avec un niveau de résolution maximal de  $R = 5$ , le temps de réponse d'une requête est  $\sim 5s$  sur la base de données de 116 objets,  $\sim 10s$  sur la base de données de 260 objets, et  $\sim 30s$  sur la base de 560 objets.

À la résolution  $r = 3$ , les temps de réponse sont réduits à  $\sim 4s$ ,  $\sim 8s$ , et  $\sim 19s$  respectivement. En comparaison, le temps de réponse d'une requête avec une méthode statistique utilisant des vecteurs caractéristiques constitués de 128 flottants (par exemple un histogramme de 128 valeurs) sont  $\sim 3s$ ,  $\sim 6s$ , et  $\sim 14s$  respectivement.

Le niveau de résolution des aMRG a un fort impact sur la rapidité des calculs. On observe que le temps de génération et de comparaison est plus coûteux pour les graphes de formes complexes que pour ceux de forme simple. En effet, les objets "compliqués" comportent plus de nœuds pour leur représentation à chaque niveau de résolution, ce qui implique plus de combinaisons d'appariements à tester lors de comparaisons avec d'autres objets. La génération d'un vecteur caractéristique à la résolution  $R = 5$  requiert  $\sim 30s$  pour un objet de  $\sim 50\,000$  triangles. Le calcul de la fonction géodésique intégrale  $\mu$  (cf. Section 2.2.3) reste l'étape la plus coûteuse. La complexité de l'algorithme est en  $O(N \log(N))$ ,  $N$  étant le nombre de points de calcul (cf. Section 2.2.6).

Afin de mettre en évidence les propriétés discriminantes de l'approche par aMRG, nous avons effectué des comparaisons avec les différentes méthodes d'indexation 3D suivantes déjà décrites au chapitre 1 :

- histogrammes de cordes [Paquet and Rioux, 1997] (cf. Section 1.3.2),
- distribution de forme D2 [Osada et al., 2002] (cf. Section 1.3.3),
- harmoniques sphériques [Kazhdan and Funkhouser, 2002] (cf. Section 1.5.2),
- descripteur de spectre de forme 3D [Zaharia and Prêteux, 2002] (cf. Section 1.4.1),
- descripteur de Hough 3D [Zaharia and Prêteux, 2002] (cf. Section 1.3.4),
- EGI complexes [Kang and Ikeuchi, 1994] (cf. Section 1.4.2),
- rapport aire/volume (cf. 3.3.2 et [Zhang and Chen, 2001]) (cf. Section 1.3.1).

Chacune de ces méthodes décrivent une caractérisation géométrique des modèles 3D sous la forme d'un vecteur caractéristique. La similarité entre deux modèles est mesurée au moyen d'une distance entre deux vecteurs caractéristiques [Bimbo, 2001]. Pour trouver la meilleure distance pour chaque méthode (cf. Section 1.8), nous avons étudié et comparé les performances de chacune des distances suivantes (leurs définitions ont été rappelées à la section 1.8) sur les trois bases de données :

- distance  $L_1$  de Minkowski (ou distance city-block, ou Manhattan),

- distance  $L_2$  de Minkowski (distance Euclidienne),
- distance  $L_\infty$  de Minkowski,
- distance de Bhattacharyya,
- distance de Jeffreys,
- divergence  $\chi^2$ ,
- intersection d'histogramme,
- distance Earth Mover [Tangelder and Velkamp, 2003, Rubner et al., 1998].

La distance la plus coûteuse est la Earth Mover, de complexité  $O(N^4 \log N)$ ,  $N$  étant le nombre de valeurs. Pour cette distance, nous limitons les calculs à  $N < 512$  en utilisant l'implémentation proposée par [Rubner et al., 1998]. Les autres distances sont en  $O(N)$ .

La dimension des vecteurs caractéristiques varie en fonction du nombre  $N$  de niveaux de quantification choisis pour obtenir les histogrammes. Nous avons évalué l'influence de ce choix pour chaque méthode avec les valeurs suivantes :

- 16 et 128 valeurs pour les histogrammes de cordes et le descripteur de spectre de forme 3D.
- 16, 128 valeurs pour la distribution de forme D2. Bien que les auteurs [Osada et al., 2002] préconisent 1024 valeurs, nous avons limité nos tests à 16 et 128 valeurs car nos essais effectués avec 1024 valeurs ne se sont pas avérés plus concluants sur nos bases de données.
- 128 ( $4 * 32$ ) et 512 ( $4 * 128$ ) valeurs pour le descripteur de Hough 3D projeté sur un octaèdre ayant ses 8 faces subdivisées en 4, résultant en 32 orientations discrètes des normales, et en 16 (par itération de 2 subdivisions en 4), résultant en 128 orientations, respectivement.
- 50 ( $2 * 5 * 5$ ) et 384 ( $3 * 128$ ) valeurs pour des EGI complexes que nous avons adaptées. Le descripteur de 50 valeurs caractérise deux grandeurs correspondant à l'ensemble des triangles associés en chaque cellule : la somme des aires relatives des triangles, et la moyenne des longueurs des cordes (distance entre le centre de gravité du modèle et les barycentres des triangles). Les cellules correspondant à la discrétisation des orientations des normales sont obtenues par partition de la sphère de Gauss en coordonnées sphériques selon les latitudes et les longitudes. On effectue 5 partitions suivant les deux coordonnées.

Le descripteur de 384 valeurs caractérise trois grandeurs dans chaque cellule : la somme des aires relatives des triangles, la moyenne des longueurs des cordes, et la moyenne des angles entre les normales et les cordes. Les 128 cellules correspondent aux 128 faces obtenues par double subdivision d'un octaèdre.

Pour évaluer l'habileté d'une méthode à retrouver des objets similaires dans une base de données, nous utilisons les outils de classification classiques suivants [Shilane et al., 2004, Goodall et al., 2005] :

**Les k-meilleurs scores :** pour une requête donnée on calcule les  $k$  premiers objets choisis par la méthode comme ayant les formes les plus similaires et on en forme une image composée. Il s'agit d'un outil visuel qui permet d'apporter une évaluation qualitative subjective (cf. Figure 4.4).

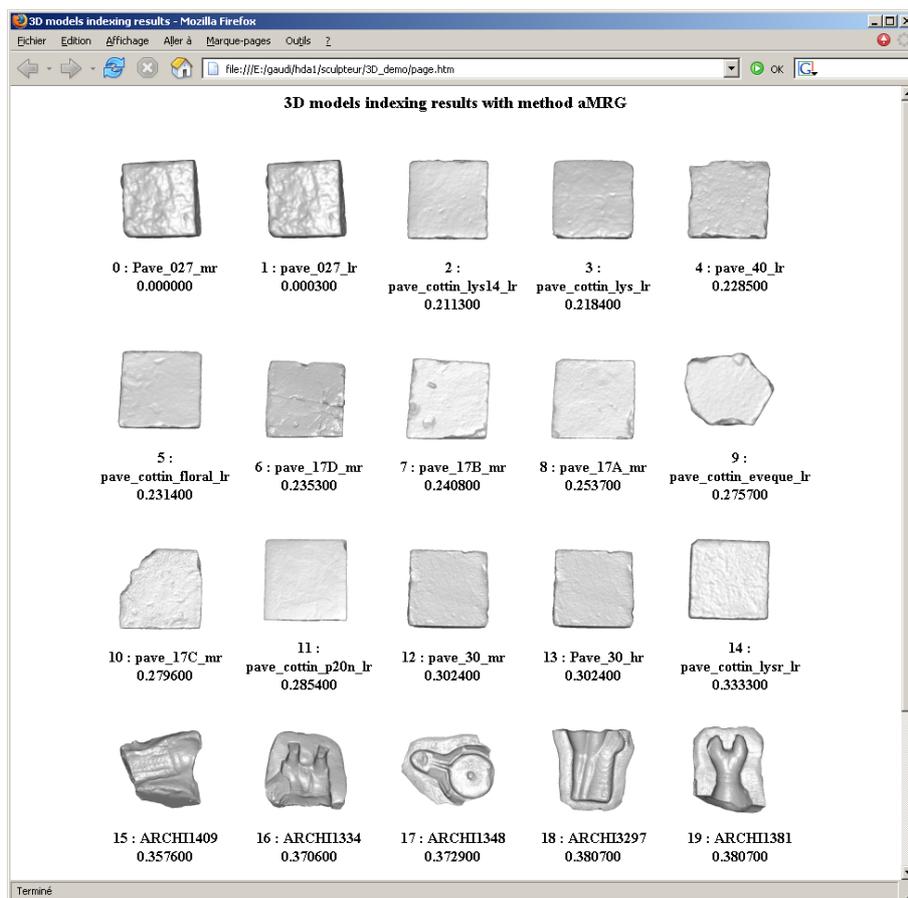


FIG. 4.4 – Exemple d'image de meilleurs scores. L'image présente un classement des objets les plus similaires à l'objet requête situé en haut à gauche (à la position 0). Le rang et le score relatifs à la requête sont indiqués sous l'image des objets sélectionnés pour la requête.

**La matrice des distances :** il s'agit de la matrice symétrique dont chaque élément  $(i, j)$  représente la valeur de similarité entre les objets  $i$  et  $j$ , les objets étant groupés par classe selon les deux axes. Cette représentation permet d'avoir rapidement un aperçu du pouvoir classificateur de la méthode étudiée. En effet, en mettant par exemple une couleur sombre pour les objets les plus similaires, la matrice des distances d'une méthode discriminante devrait comporter une forte prédominance sombre à l'intérieur des carrés alignés sur la diagonale correspondant à des objets d'une même classe. La matrice des distances permet notamment de détecter les classes d'objets les plus homogènes pour la mesure de similarité correspondante (cf. Figure 4.5).

**L'image tier :** elle représente pour chaque objet d'une classe de taille  $C$  :

1. les objets les plus proches de la requête et appartenant à la même classe parmi les  $C - 1$  premiers objets retournés (pixels en noir),
2. les objets parmi les  $C - 1$  premiers objets retournés qui ne sont pas de la

même classe que la requête (pixels rouge),

- les  $C - 1$  objets suivants, retournés entre les rangs  $C$  et  $2(C - 1)$  inclus (pixels bleus) (cf. Figure 4.6).

En associant une couleur pour chaque catégorie, cette représentation permet, comme la précédente, d'obtenir un rapide aperçu du pouvoir classificateur d'un descripteur. Les carrés sur la diagonale de l'image, correspondant aux classes, sont complètement noirs pour une méthode très discriminante, et noirs et bleus pour une moins discriminante. La présence de pixels blancs  $(i, j)$  dans les carrés signifie que pour la requête correspondant à l'objet  $i$ , l'objet  $j$  appartenant à la même classe que  $i$  n'a pas été retrouvé par la méthode parmi les  $2(C - 1)$  premiers objets. Cette représentation n'est plus symétrique et doit se lire ligne par ligne.

**La courbe précision-rappel :** il s'agit d'un calcul statistique sur la base de données mettant aussi en évidence la capacité de la méthode étudiée à retrouver les classes d'objets. Il ne faut prendre en compte que les objets classables. Soit un objet  $i$  d'une classe de taille  $C$ ,  $R$  le nombre choisi d'objets retournés par la requête, et  $Rr$  le nombre d'objets parmi les  $R$  appartenant à la même classe que  $i$ , la courbe est composée en ordonnée par la précision, définie comme le rapport  $Precision = Rr/R$ , et en abscisse par le rappel, défini comme le rapport  $Rappel = Rr/C$ . Une méthode discriminant parfaitement bien les classes produirait une ligne horizontale à  $Precision = 1.0$  (cf. Figure 4.7). L'approche étant statistique, on compare systématiquement les résultats obtenus avec des tirages aléatoires afin d'obtenir une évaluation du bruit.

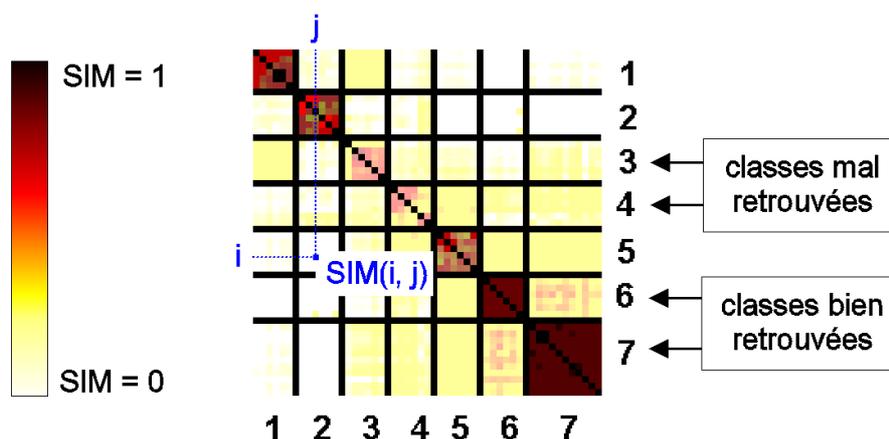


FIG. 4.5 – Exemple de matrice de distance. Les classes d'objets sont numérotées et réparties le long des deux dimensions. L'élément  $(i, j)$  représente la similarité entre les objets  $i$  et  $j$ . Les pixels les plus sombres représentent les objets les plus similaires et inversement les pixels clairs représentent les objets les moins similaires. La présence de grands carrés sombres sur la diagonale signifie une bonne discrimination des classes.

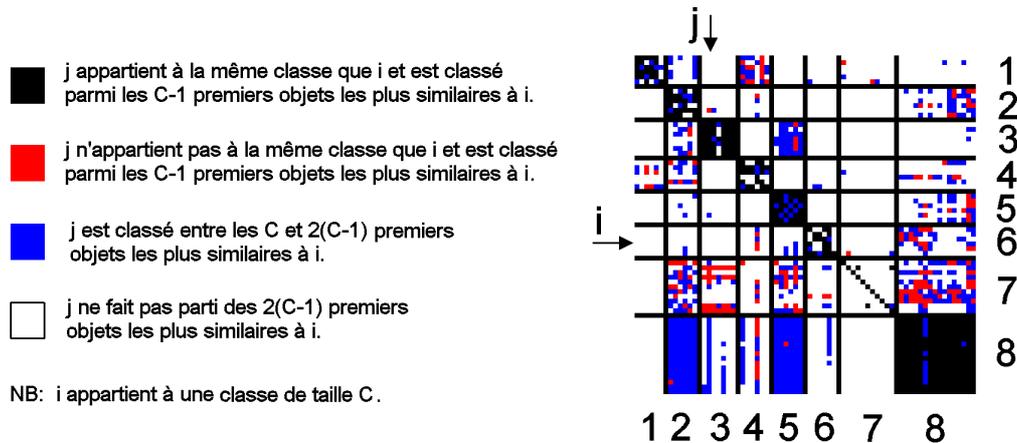


FIG. 4.6 – Exemple d’image tier. Les pixels  $(i, j)$  en noirs indiquent que l’objet  $j$  appartient à la même classe de taille  $C$  que l’objet  $i$ , et se trouve parmi les  $C - 1$  objets les plus similaires à  $i$ . Les pixels rouges indiquent que l’objet  $j$  se trouve parmi les  $C - 1$  objets plus proches de  $i$  mais n’appartient pas à la même classe que  $i$ . Les pixels bleus représentent les objets  $j$  retournés entre les rangs  $C$  et  $2(C - 1)$  inclus.

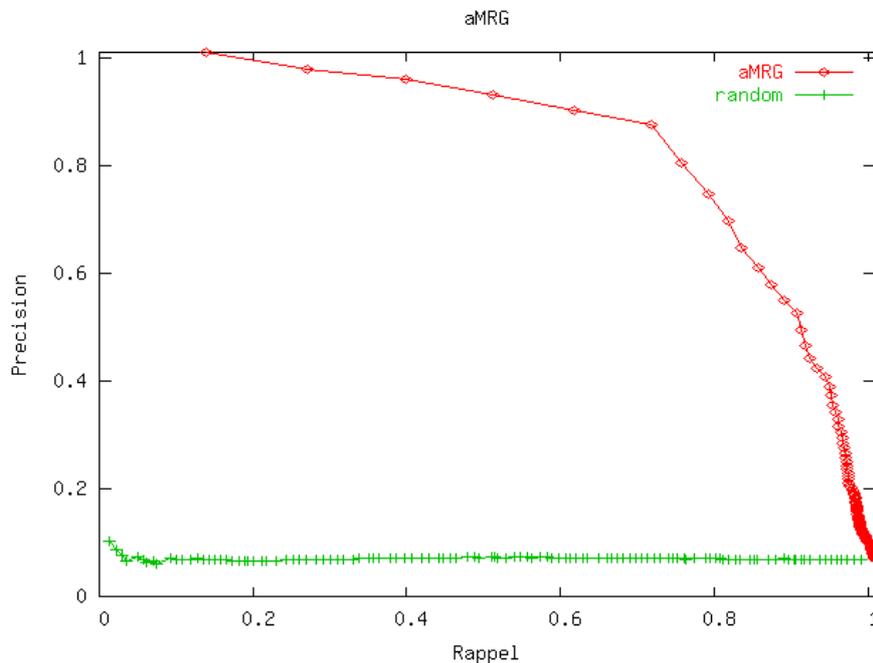


FIG. 4.7 – Exemple de courbe précision-rappel. La représentation caractérise l’efficacité de la méthode à retrouver les classes. Une méthode discriminante parfaite produirait une ligne horizontale à  $Precision = 1.0$ .

L'inconvénient avec ces représentations, et particulièrement avec l'image tier, est qu'elles deviennent difficiles à interpréter lorsque les bases de données sont de taille très importante et comportent un très grand nombre de classes. De plus, il devient difficile d'évaluer les résultats lorsque ces représentations sont visuellement similaires. C'est pourquoi il est aussi utile de calculer des scores, plus synthétiques, pour évaluer les performances des différentes méthodes [Shilane et al., 2004] :

**Le score first tier  $T_1$**  : il est obtenu en calculant la macro moyenne des pourcentages des objets bien retrouvés (appartenant à la même classe de taille  $C$  que l'objet requête) parmi les  $C - 1$  premiers objets retournés, c'est-à-dire la moyenne des moyennes des pourcentages de chaque classe :

$$T_1 = \frac{1}{N_c} \sum_{i=1}^{N_c} T_{C_i} \quad \text{et} \quad T_{C_i} = \frac{1}{C_i} \sum_{j=1}^{C_i} T_j^i, \quad (4.1)$$

où  $N_c$  est le nombre de classes dans la base de données,  $T_{C_i}$  le score first tier de la  $i^e$  classe,  $C_i$  la taille de la  $i^e$  classe et  $T_j^i$  le score first tier du  $j^e$  objet de la  $i^e$  classe (c'est-à-dire le pourcentage d'objets de la même classe  $C_i$  retournés parmi les  $C - 1$  premiers objets).

**Le score second tier** : il est obtenu de façon similaire au score first tier, mais avec les pourcentages calculés sur les  $2(C - 1)$  premiers objets retrouvés. On obtient aussi de façon similaire le score bull-eye (SBE), calculé sur les  $2C$  premiers objets retrouvés. Le SBE a été officiellement choisi pour la norme MPEG-7 [Zaharia and Prêteux, 2002].

**La mesure E** : elle caractérise l'efficacité d'une classification, basée sur seulement les 32 premiers objets retrouvés [van Rijsbergen, 1975, Shilane et al., 2004]. Le nombre est fixe et représente le nombre de résultats auxquels s'intéressent les utilisateurs d'un moteur de recherche. On estime qu'intuitivement les utilisateurs sont plus intéressés par les premiers objets retrouvés que par les suivants. Par exemple dans [Shilane et al., 2004], les auteurs proposent de présenter les 32 objets les plus similaires à une requête sur une page de résultats. Le calcul repose sur les valeurs de précision et de rappel :

$$E = \frac{2}{\frac{1}{P} + \frac{1}{R}}. \quad (4.2)$$

Des valeurs hautes de E signifient une bonne discrimination sur les 32 premiers résultats, avec un score maximum de 1. Ce maximum ne peut être théoriquement atteint que si toutes les classes ont un nombre d'objets  $\geq 32$  car sinon la valeur de précision est majorée par  $C/32$ .

**Le décompte de gain cumulé (ou Discounted Cumulative Gain, DCG)** : il s'agit d'une mesure statistique dont le calcul consiste à cumuler les contributions de tous les objets de la base de données, avec des poids dépendant du classement des objets retournés par la méthode évaluée [Leifman et al., 2003, Shilane et al., 2004]. Soit  $N$  le nombre d'objets dans la base de données, et  $G_r$  ( $r \in [1, N]$ ) la contribution du  $r^e$  objet retrouvé telle que :

- $G_r = 0$  si le  $r^e$  objet retrouvé n'appartient pas à la même classe que l'objet requête,
- $G_r = \frac{1}{\log_2(r)}$  si le  $r^e$  objet retrouvé appartient à la même classe que l'objet requête.

Le DCG de l'objet requête est défini par :

$$DCG = \begin{cases} 1 & \text{si } N = 1 \\ 1 + \sum_{r=2}^N G_r & \text{sinon} \end{cases} \quad (4.3)$$

On normalise ensuite le DCG par la valeur maximale possible du score DCG, correspondant à une classification parfaite :

$$DCG_N = \frac{DCG}{1 + \sum_{j=2}^C \frac{1}{\log_2(j)}}, \quad (4.4)$$

$C$  étant la taille de la classe de l'objet requête. La valeur maximale de  $DCG_N$  est donc 1. Une valeur haute du DCG calculé pour une requête signifie que tous les objets de la même classe que l'objet requête sont retrouvés proches.

### 4.3 Etude des méthodes implémentées

Cette partie présente les performances des méthodes d'indexation 3D précédemment listées obtenues sur la base de test de 116 objets. Afin d'obtenir les meilleurs résultats pour chaque méthode, nous avons étudié le comportement de chacune d'entre elles avec chaque distance et avec les différentes dimensions des vecteurs caractéristiques sélectionnés.

L'objectif de cette étude est double :

1. trouver les meilleurs candidats à comparer à notre approche afin d'en évaluer l'efficacité de manière objective,
2. déterminer les approches les plus pertinentes afin de les fusionner en tant qu'attributs de nœuds des graphes de Reeb multirésolution augmentés.

L'évaluation de chaque combinaison de méthode et distance est effectuée au moyen des techniques visuelles et qualitatives présentées dans la partie précédente.

### 4.3.1 Le rapport aire/volume

Le calcul du rapport aire/volume<sup>2</sup> présenté en section 3.4.2 permet de caractériser la forme des objets 3D par une valeur réelle comprise entre 0 et 1. Le descripteur étant scalaire les distances  $L_1$ ,  $L_2$ ,  $L_\infty$  et EMD sont équivalentes. Les distances de Bhattacharyya et d'intersection d'histogrammes ne s'appliquent pas car ils requièrent que les données soient normalisées à 1.

La Figure 4.8 présente les courbes de précision-rappel obtenues sur la base de données de 116 objets en effectuant des calculs de similarité avec le rapport et toutes les distances implémentées. Les résultats sont équivalents pour les trois distances  $L_1$ , Jeffreys et  $\chi^2$ , ce qu'indiquent aussi les scores (cf. Table 4.1).

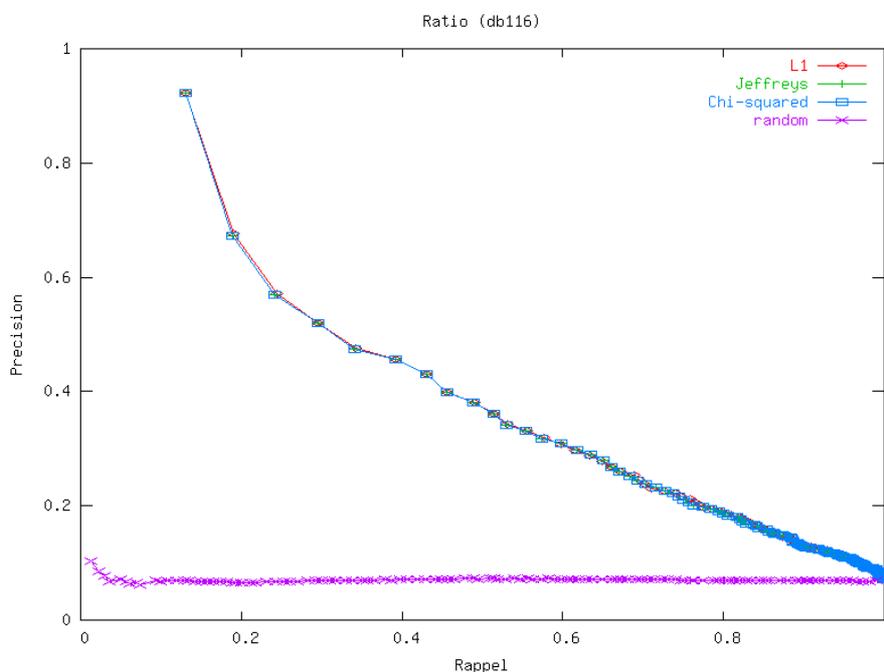


FIG. 4.8 – Courbes de précision-rappel des distances avec le calcul du rapport aire/volume. Toutes les distances donnent sensiblement les mêmes résultats.

Les performances sont globalement faibles. En moyenne, pour une requête d'un objet appartenant à une classe de taille  $C$ , seulement la moitié des objets de la classe sont correctement retrouvés parmi les  $C$  premiers résultats retournés par le rapport :  $T_1 = 49.4\%$ , et  $SBE = 68.3\%$ . La mesure E faible indique que les objets appartenant à une même classe ne sont pas retrouvés proches.

La figure 4.9 présente la matrice de distances et l'image de tier obtenues avec le rapport et la distance  $L_1$ . On constate que le rapport est pertinent pour les classes

<sup>2</sup>On fait ici référence au  $Ratio = 36.\pi.\text{volume}^2/\text{aire}^3$  défini en section 3.4.2.

	$L_1, L_2, L_\infty, EMD$	Jeffreys	$\chi^2$
$T_1$	0.491	0.494	0.494
$SBE$	0.681	0.683	0.683
$E$	0.296	0.295	0.295
$DCG$	0.726	0.724	0.724

TAB. 4.1 – Scores ( $T_1$ ,  $SBE$ , mesure  $E$  et  $DCG$ ) sur la base de données de 116 modèles du rapport aire/volume pour les diverses distances.

de sphère n° 2, de cubes n° 1, de cônes n° 3, et des vases “ronds” n° 11. Par contre les résultats sont faibles pour les autres classes, et particulièrement pour les classes de modèles humains n° 12, les statues n° 1, les pièces d’échecs n° 7, les moules n° 13, et les pavés n° 15. On conclut que le calcul du rapport permet de distinguer les objets de formes sphériques (où le rapport est maximal) des objets filiformes (ou allongés) et des objets plats.

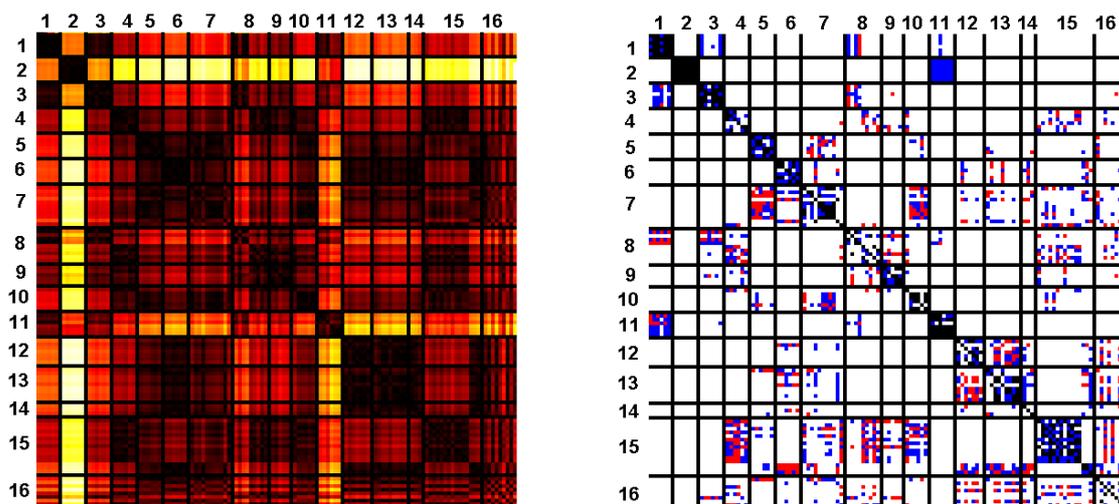


FIG. 4.9 – Matrice de distances et image tier du rapport aire/volume avec la distance  $L_1$ . Le rapport n’est pas suffisamment pertinent pour retrouver les classes.

### 4.3.2 Histogrammes de cordes

#### Histogramme des longueurs de cordes

Les figures 4.10 et 4.11 présentent les courbes de précision-rappel obtenues avec les histogrammes des longueurs de cordes pour toutes les distances, les histogrammes contenant 16 et 128 valeurs respectivement.

Les résultats obtenus avec les histogrammes de 16 valeurs sont globalement meilleurs que ceux de 128 valeurs et présentent moins d'écart entre les différentes distances. Cependant les meilleurs résultats sont obtenus avec les histogrammes de 128 valeurs avec la distance Earth Mover. La distance Earth Mover donne aussi les meilleurs résultats avec les histogrammes de 16 valeurs, bien que la différence avec les autres distances soit moins notable qu'avec les histogrammes de 128 valeurs : l'EMD est  $\sim 1\%$  supérieure aux autres distances avec les histogrammes de 16 valeurs, et  $\sim 10\%$  supérieure avec les histogrammes de 128 valeurs.

Ces résultats permettent d'illustrer la stabilité de l'EMD face à la quantification des données. Les scores des autres distances ont chuté de 5 à 10% lors du passage de 16 à 128 valeurs, tandis que l'EMD a gagné de  $\sim 1\%$ .

Les performances restent moyennes, avec  $T_1 = 69.1\%$  et  $70.4\%$  pour 16 et 128 valeurs respectivement.

	$L1$	$L2$	$L_\infty$	Bhatta	Jeffreys	$\chi^2$	Histo int	EMD
$T_1$	0.681	0.592	0.557	0.681	0.619	0.669	<b>0.682</b>	<b>0.691</b>
$SBE$	0.787	0.682	0.634	<b>0.801</b>	0.736	0.794	0.789	<b>0.808</b>
$E$	0.303	0.275	0.261	0.302	0.288	<b>0.304</b>	0.303	<b>0.308</b>
$DCG$	<b>0.838</b>	0.788	0.752	<b>0.838</b>	0.805	0.836	<b>0.838</b>	<b>0.850</b>

TAB. 4.2 – Scores ( $T_1$ ,  $SBE$ , mesure  $E$  et  $DCG$ ) sur la base de données de 116 modèles avec les histogrammes de cordes de 16 valeurs pour les diverses distances.

	$L1$	$L2$	$L_\infty$	Bhatta	Jeffreys	$\chi^2$	Histo Int	EMD
$T_1$	<b>0.629</b>	0.527	0.467	0.613	0.527	0.627	<b>0.629</b>	<b>0.704</b>
$SBE$	<b>0.729</b>	0.596	0.525	0.726	0.633	0.724	0.727	<b>0.821</b>
$E$	<b>0.279</b>	0.234	0.223	0.272	0.272	0.275	<b>0.279</b>	<b>0.310</b>
$DCG$	<b>0.804</b>	0.728	0.673	0.791	0.733	0.797	0.803	<b>0.855</b>

TAB. 4.3 – Scores ( $T_1$ ,  $SBE$ , mesure  $E$  et  $DCG$ ) sur la base de données de 116 modèles avec les histogrammes de longueurs de cordes de 128 valeurs pour les diverses distances.

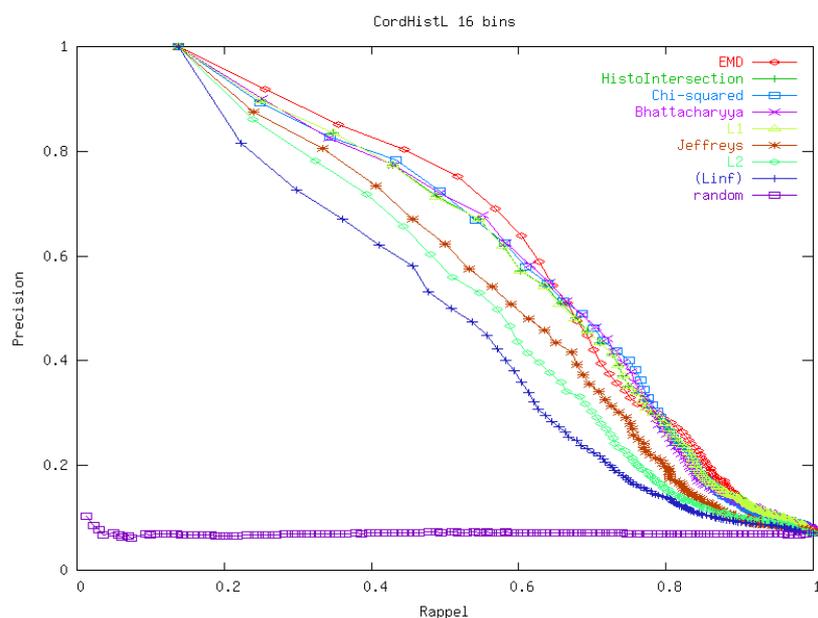


FIG. 4.10 – Courbes de précision-rappel des distances pour les histogrammes de longueurs de cordes de 16 valeurs. La meilleure distance est la Earth Mover.

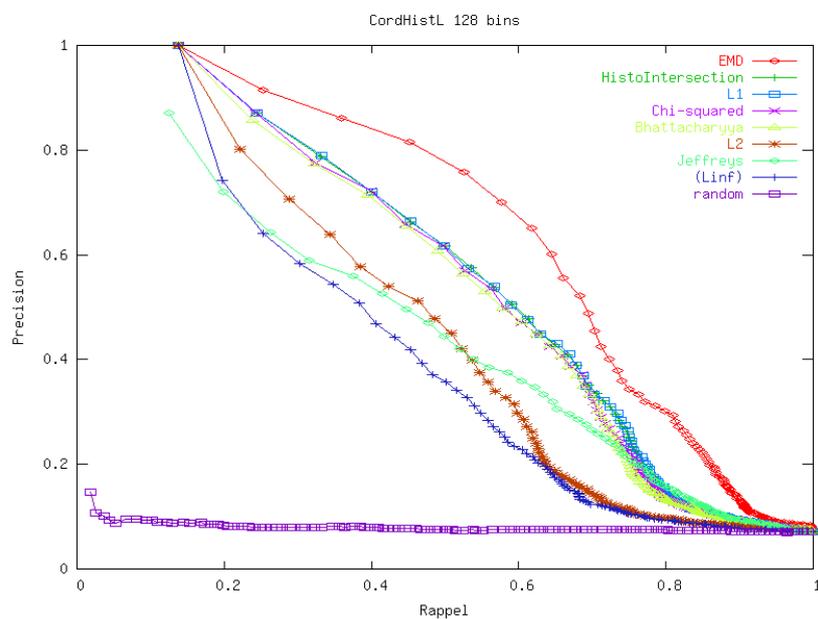


FIG. 4.11 – Courbes de précision-rappel des distances pour les histogrammes de longueurs de cordes de 128 valeurs. La meilleure distance est la EarthMover.

### Histogramme des angles des cordes avec le 1<sup>er</sup> axe

Les figures 4.12 et 4.13 présentent les courbes de précision-rappel obtenues avec les histogrammes d'angles des cordes par rapport au 1<sup>er</sup> axe principal de l'ACP, avec toutes les distances, et avec des tailles de 16 et 128 valeurs respectivement.

Les meilleurs résultats sont obtenus avec la divergence du  $\chi^2$  pour les deux tailles d'histogrammes, bien que sa supériorité ne soit pas significative. Les distances Earth Mover et Bhattacharyya donnent les moins bons résultats pour les histogrammes de 16 valeurs, et les distances  $L_\infty$  et Bhattacharyya donnent les moins bons résultats pour les histogrammes de 128 valeurs. Les autres distances retournent des performances équivalentes à 1% près.

Les histogrammes de 16 valeurs donnent des résultats légèrement supérieurs à ceux de 128 valeurs. On note une fois de plus que l'EMD reste stable lors du passage de 16 à 128 valeurs, et retourne des scores de SBE et DCG légèrement supérieurs alors que les performances de toutes autres distances chutent.

Les performances sont globalement faibles, avec  $T_1 = 55.1\%$  et  $54.7\%$  pour 16 et 128 valeurs respectivement.

	$L1$	$L2$	$L_\infty$	Bhatta	Jeffreys	$\chi^2$	Histo int	EMD
$T_1$	0.542	<b>0.561</b>	0.552	0.536	0.547	<b>0.551</b>	0.544	0.534
$SBE$	0.694	0.697	0.700	0.710	0.704	<b>0.713</b>	0.691	0.657
$E$	0.296	0.294	0.293	0.299	0.300	<b>0.301</b>	0.296	0.290
$DCG$	0.777	0.779	0.773	<b>0.789</b>	0.775	0.781	0.777	0.749

TAB. 4.4 – Scores ( $T_1$ ,  $SBE$ , mesure  $E$  et  $DCG$ ) sur la base de données de 116 modèles avec les histogrammes d'angles des cordes avec le 1<sup>er</sup> axe de 16 valeurs pour les diverses distances.

	$L1$	$L2$	$L_\infty$	Bhatta	Jeffreys	$\chi^2$	Histo int	EMD
$T_1$	0.530	0.532	0.477	0.516	0.533	<b>0.547</b>	0.526	0.529
$SBE$	0.679	0.661	0.603	0.663	0.655	<b>0.688</b>	0.682	0.668
$E$	0.278	0.270	0.252	0.280	0.281	0.282	0.279	<b>0.290</b>
$DCG$	0.762	0.758	0.709	0.746	0.755	<b>0.767</b>	0.761	0.759

TAB. 4.5 – Scores ( $T_1$ ,  $SBE$ , mesure  $E$  et  $DCG$ ) sur la base de données de 116 modèles avec les histogrammes d'angles des cordes avec le 1<sup>er</sup> axe de 128 valeurs pour les diverses distances.

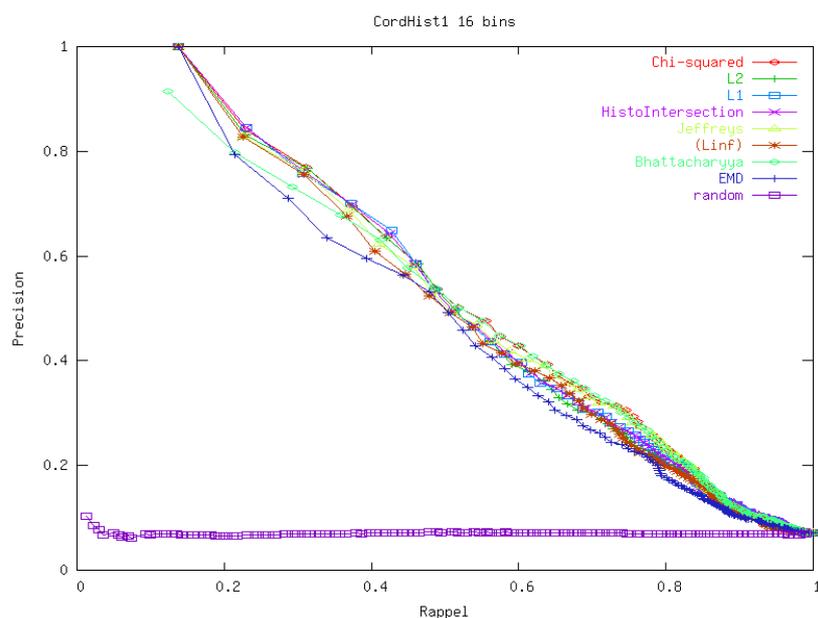


FIG. 4.12 – Courbes de précision-rappel pour les histogrammes d’angles des cordes avec le 1<sup>er</sup> axe de 16 valeurs. Les performances sont équivalentes pour les 6 premières distances.

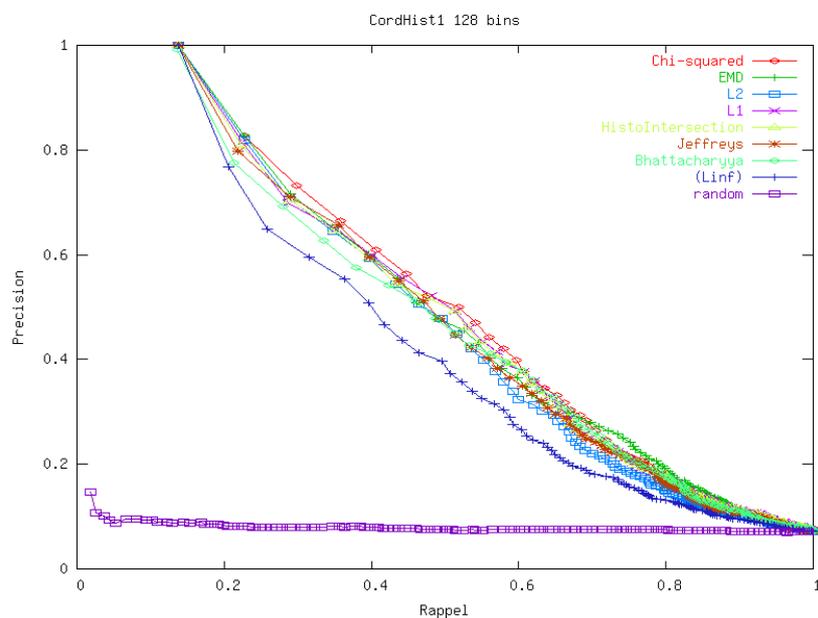


FIG. 4.13 – Courbes de précision-rappel pour les histogrammes d’angles des cordes avec le 1<sup>er</sup> axe de 128 valeurs. La meilleure distance est la divergence du  $\chi^2$ .

### Histogrammes des angles des cordes avec le 2<sup>e</sup> axe

Les figures 4.14 et 4.15 présentent les courbes de précision-rappel obtenues avec les histogrammes des angles des cordes par rapport au 2<sup>e</sup> axe principal de l'ACP, avec toutes les distances, et avec des tailles de 16 et 128 valeurs respectivement.

Les performances de toutes les distances sont équivalentes pour les histogrammes de 16 valeurs. Les distances de Bhattacharyya et  $\chi^2$  sont les plus performantes et  $L_\infty$  donne les moins bons résultats.

Les résultats sont moins homogènes avec les histogrammes de 128 valeurs. L'EMD donne les meilleurs résultats et la distance  $L_\infty$  les moins bons. Les histogrammes de 128 valeurs donnent des résultats globalement équivalents à ceux de 16 valeurs, les meilleurs résultats étant obtenus avec l'EMD pour les histogrammes de 128 valeurs.

Les performances sont globalement moyennes, avec  $T_1 = 71.9\%$  et  $69.8\%$  pour 16 et 128 valeurs.

	$L1$	$L2$	$L_\infty$	Bhatta	Jeffreys	$\chi^2$	Histo int	EMD
$T_1$	0.687	0.667	0.645	<b>0.719</b>	0.677	0.711	0.687	0.678
$SBE$	0.821	0.800	0.787	<b>0.831</b>	0.819	<b>0.831</b>	0.823	0.815
$E$	<b>0.310</b>	0.304	0.297	0.309	0.308	<b>0.310</b>	<b>0.310</b>	0.307
$DCG$	0.834	0.823	0.813	0.846	0.826	<b>0.847</b>	0.833	0.833

TAB. 4.6 – Scores ( $T_1$ ,  $SBE$ , mesure  $E$  et  $DCG$ ) sur la base de données de 116 modèles avec les histogrammes d'angles des cordes avec le 2<sup>e</sup> axe de 16 valeurs pour les diverses distances.

	$L1$	$L2$	$L_\infty$	Bhatta	Jeffreys	$\chi^2$	Histo int	EMD
$T_1$	<b>0.686</b>	0.627	0.535	0.667	0.628	0.673	<b>0.688</b>	<b>0.698</b>
$SBE$	<b>0.800</b>	0.712	0.670	0.774	0.772	0.789	<b>0.800</b>	<b>0.832</b>
$E$	<b>0.296</b>	0.269	0.264	0.289	0.301	0.294	<b>0.296</b>	<b>0.311</b>
$DCG$	<b>0.839</b>	0.800	0.750	0.826	0.801	0.835	<b>0.838</b>	<b>0.849</b>

TAB. 4.7 – Scores ( $T_1$ ,  $SBE$ , mesure  $E$  et  $DCG$ ) sur la base de données de 116 modèles avec des histogrammes d'angles des cordes avec le 2<sup>e</sup> axe de 128 valeurs pour les diverses distances.

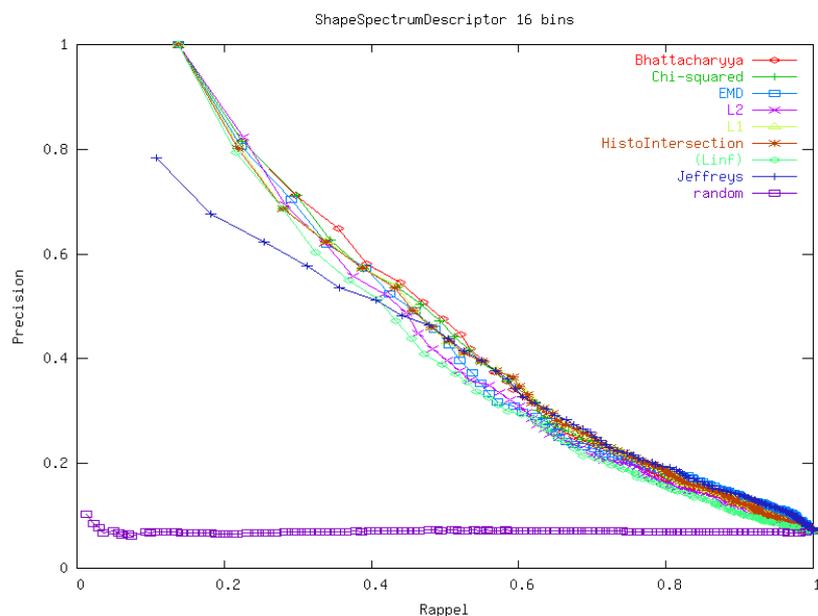


FIG. 4.14 – Courbes de précision-rappel pour les histogrammes d'angles des cordes avec le 2<sup>e</sup> axe de 16 valeurs. Les meilleures distances sont les distances de Bhattacharyya et  $\chi^2$ .

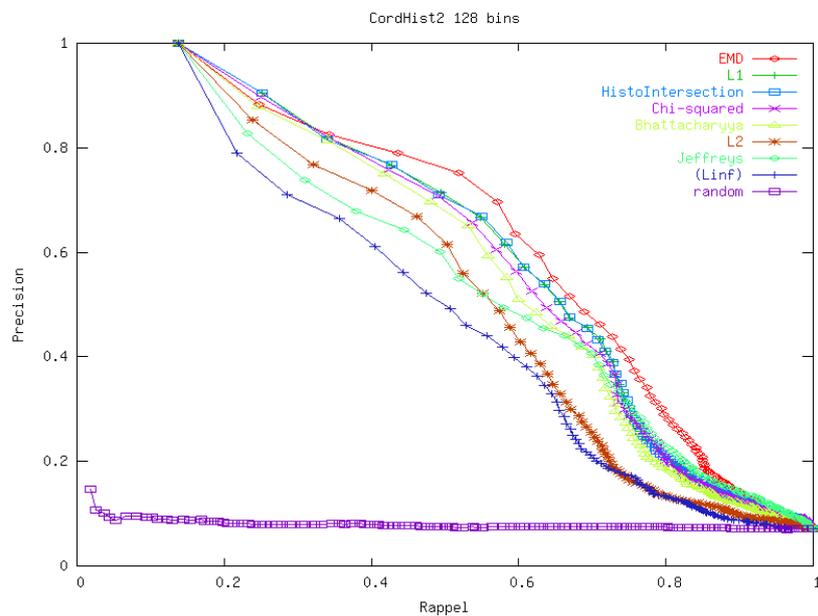


FIG. 4.15 – Courbes de précision-rappel pour les histogramme d'angles des cordes avec le 2<sup>e</sup> axe de 128 valeurs. La meilleure distance est la distance EMD.

### 4.3 Etude des méthodes implémentées

---

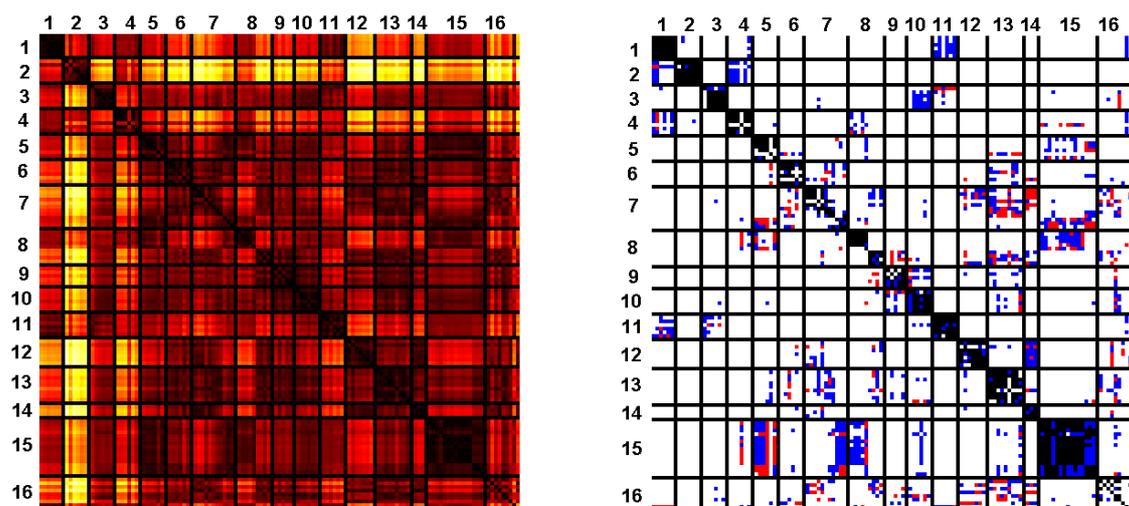


FIG. 4.16 – *Matrice de distances et image tier avec les histogrammes de longueurs des cordes de 16 valeurs et la distance EMD.*

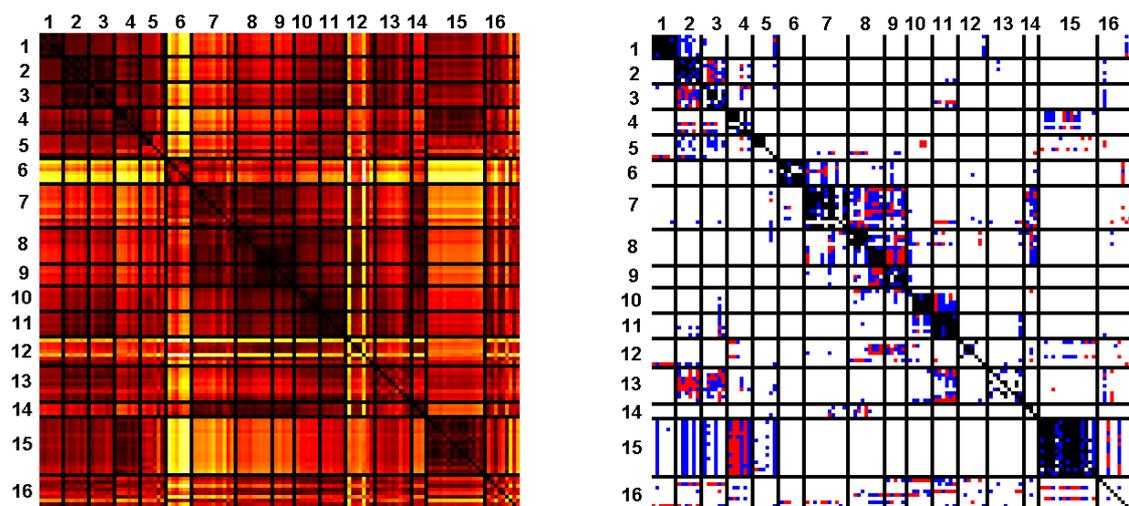


FIG. 4.17 – *Matrice de distances et image tier avec les histogrammes d'angles des cordes de 16 valeurs avec le 1<sup>er</sup> axe et la distance  $L_2$ .*

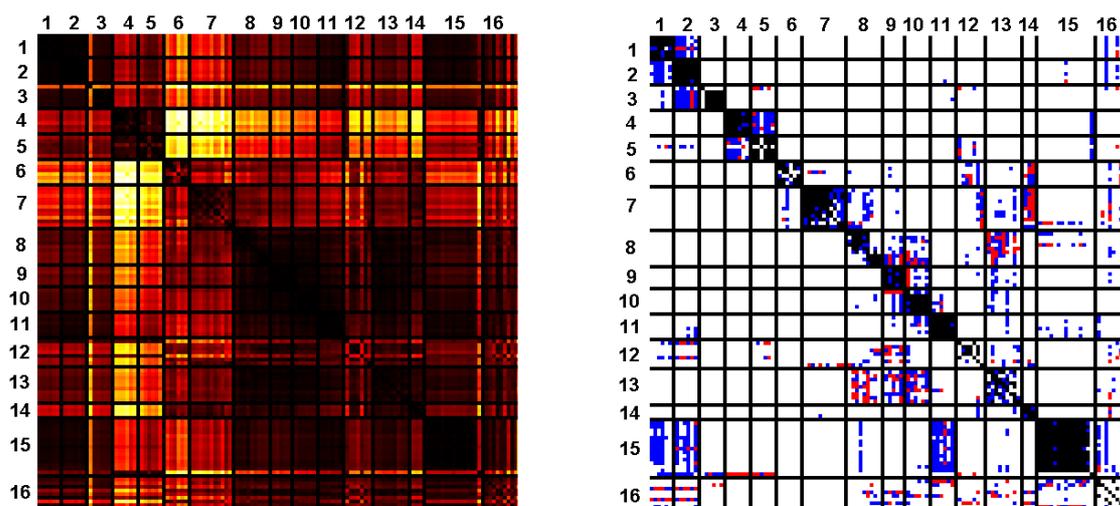


FIG. 4.18 – *Matrice de distances et image tier avec les histogrammes d'angles des cordes de 16 valeurs avec le 2<sup>e</sup> axe et la distance de Bhattacharyya.*

Les histogrammes de longueurs de cordes permettent de retrouver partiellement la plupart des classes (cf. Figure 4.16). Le principal problème de cette approche est la précision des résultats, les classes retrouvées étant pour la plupart incomplètes. Les histogrammes des angles des cordes avec le 2<sup>e</sup> axe de l'ACP donnent de meilleurs résultats que ceux avec le 1<sup>er</sup> axe, et se sont révélés efficaces pour retrouver les formes primitives (n° 1 à 5), certains vases (n° 9, 10 et 11) et les pavés (n° 15), mais très peu performantes pour les modèles humains (n° 12) et les moules (n° 13) (cf. Figures 4.17 et 4.18).

### 4.3.3 Distribution de forme D2

Les performances de la distribution de forme D2 sont très mauvaises sur notre base de données quelque soit la distance employée, et pour les deux tailles d'histogrammes de 16 et 128 valeurs (cf. Figures 4.19 et 4.20). Les scores sont assez homogènes et équivalents pour les deux tailles d'histogrammes ( $\sim 1\%$  d'écart).

Les tests ont été effectués avec un grand nombre de tirages aléatoires (1024 et  $1024 \times 1024$ ) mais sans grandes conséquences sur les performances. On obtient en moyenne les faibles scores de  $T_1 = 33.7\%$  et  $34.2\%$  pour 16 et 128 valeurs respectivement. On observe de plus que les matrices de distances et image tier sont très bruitées (cf. Figure 4.21).

	$L1$	$L2$	$L_\infty$	Bhatta	Jeffreys	$\chi^2$	Histo int	EMD
$T_1$	<b>0.337</b>	0.306	0.281	0.313	0.270	0.320	0.336	0.330
$SBE$	<b>0.458</b>	0.427	0.383	0.435	0.377	0.448	0.455	0.447
$E$	0.219	0.209	0.196	0.199	0.203	0.203	<b>0.220</b>	0.216
$DCG$	0.625	0.612	0.596	0.622	0.589	0.622	0.625	<b>0.633</b>

TAB. 4.8 – Scores ( $T_1$ ,  $SBE$ , mesure  $E$  et  $DCG$ ) sur la base de données de 116 modèles avec la distribution de forme D2 sur 16 valeurs pour les diverses distances.

	$L1$	$L2$	$L_\infty$	Bhatta	Jeffreys	$\chi^2$	Histo int	EMD
$T_1$	0.327	0.311	0.290	0.286	<b>0.342</b>	0.307	0.336	0.314
$SBE$	0.418	0.395	0.390	0.353	<b>0.474</b>	0.371	0.415	0.421
$E$	0.208	0.205	0.207	0.158	<b>0.244</b>	0.172	0.208	0.208
$DCG$	0.624	0.615	0.600	0.582	<b>0.644</b>	0.599	0.629	0.624

TAB. 4.9 – Scores ( $T_1$ ,  $SBE$ , mesure  $E$  et  $DCG$ ) sur la base de données de 116 modèles avec la distribution de forme D2 sur 128 valeurs pour les diverses distances.

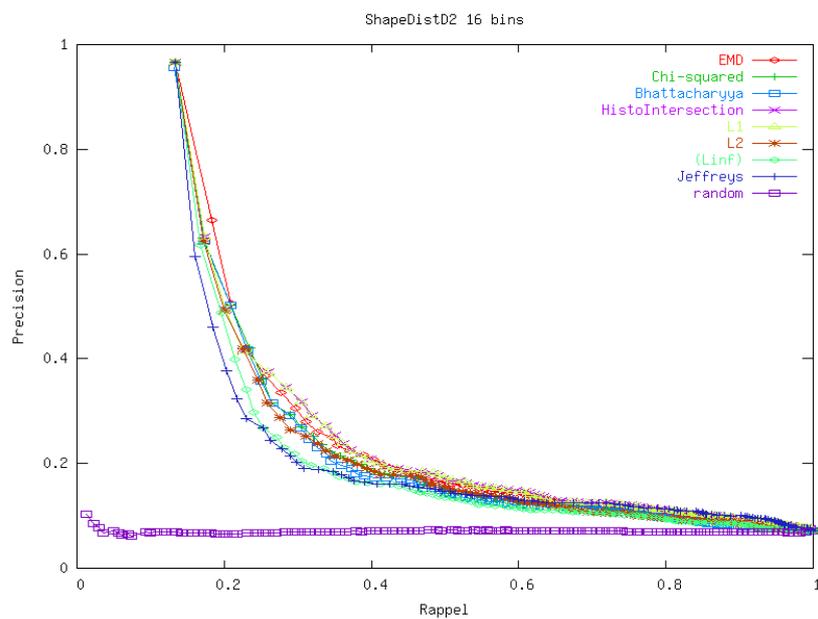


FIG. 4.19 – Courbes de précision-rappel pour la distribution de forme  $D2$  sur 16 valeurs. La meilleure distance est la  $L_1$ .

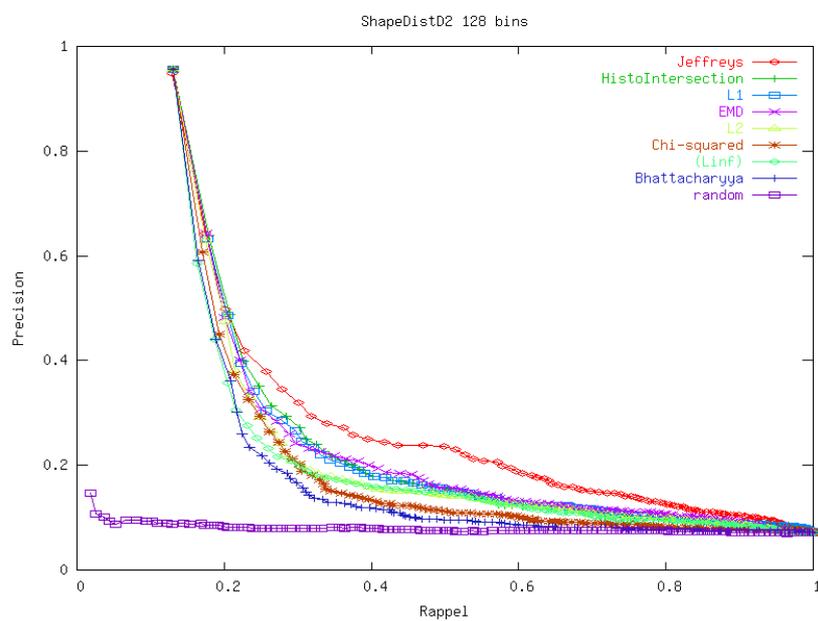


FIG. 4.20 – Courbes de précision-rappel pour la distribution de forme  $D2$  sur 128 valeurs. La meilleure distance est celle de Jeffreys.

### 4.3 Etude des méthodes implémentées

---

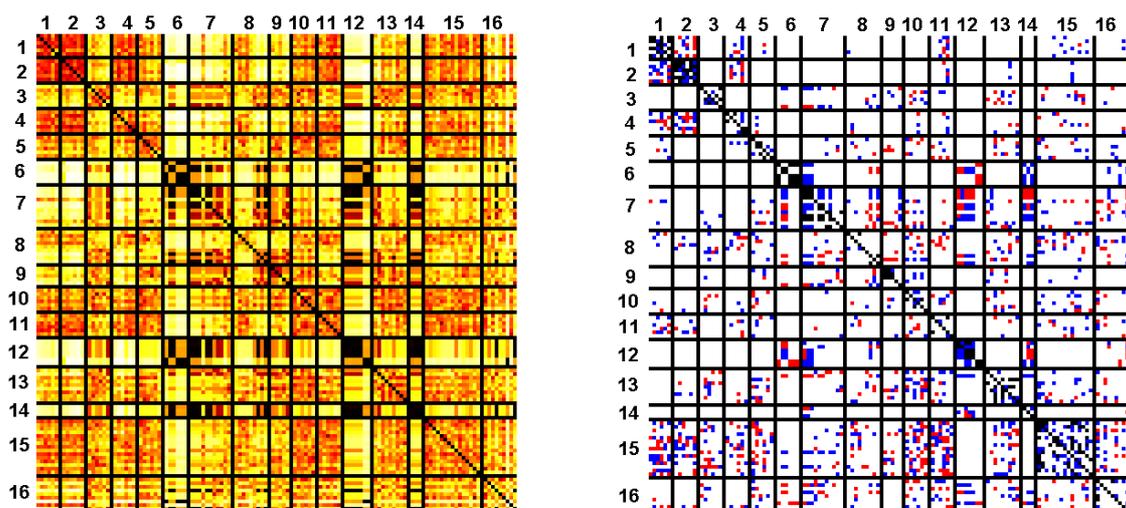


FIG. 4.21 – *Matrice de distances et image tier avec la distribution de forme D2 de 16 valeurs et l'intersection d'histogrammes.*

### 4.3.4 Les harmoniques sphériques

La figure 4.3.4 présente les courbes obtenues avec les harmoniques sphériques. A l'instar des résultats obtenus avec les distributions de forme D2, les résultats sont très faibles. La divergence du  $\chi^2$  retourne les meilleurs résultats, tandis que les autres distances retournent des résultats plus faibles similaires (cf. Table 4.10). Il n'est pas possible de calculer les distances de Bhattacharyya et d'intersection d'histogrammes car les coefficients de la décomposition en harmoniques sphériques ne sont pas normalisés à 1.

Les performances sont très faibles, avec en moyenne seulement un tiers des objets retrouvés par classe ( $T_1 = 38.3\%$ ).

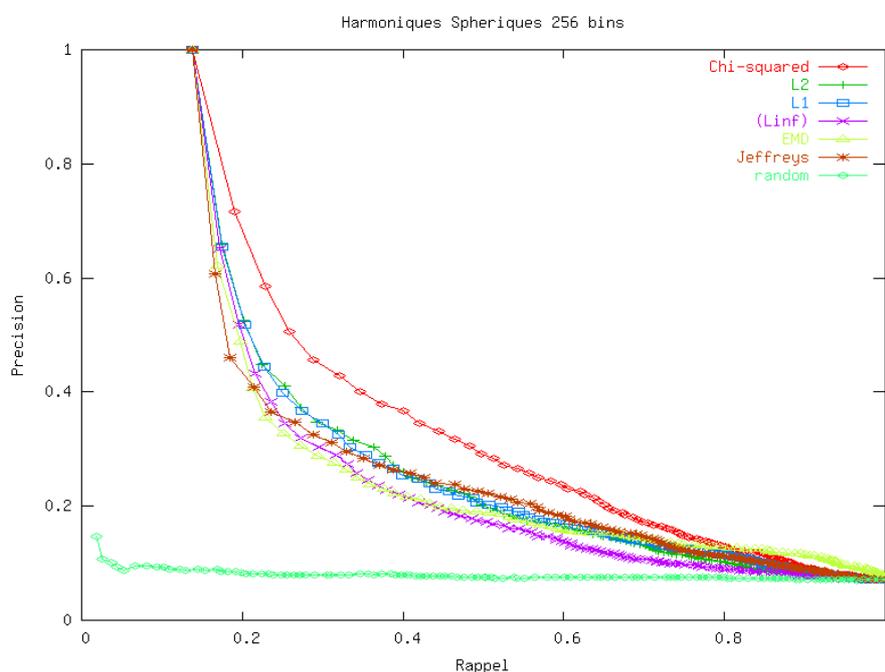


FIG. 4.22 – Courbes de précision-rappel pour les harmoniques sphériques. La meilleure distance est la divergence du  $\chi^2$ .

	$L1$	$L2$	$L_{inf}$	Jeffreys	$\chi^2$	EMD
$T_1$	0.322	0.324	0.297	0.317	<b>0.383</b>	0.301
$SBE$	0.441	0.431	0.397	0.453	<b>0.529</b>	0.432
$E$ measure	0.224	0.216	0.197	0.226	<b>0.247</b>	0.219
$DCG$	0.612	0.606	0.587	0.603	<b>0.657</b>	0.606

TAB. 4.10 – Scores ( $T_1$ ,  $SBE$ , mesure  $E$  et  $DCG$ ) sur la base de données de 116 modèles des harmoniques sphériques pour les diverses distances.

### 4.3.5 Les EGI complexes

Les courbes de précision-rappel obtenues avec les CEGI sur 50 valeurs sont homogènes, avec une légère supériorité pour la distance de Jeffreys et les moins bonnes performances pour la distance  $L_\infty$  (cf. Figure 4.23). Les autres distances donnent des performances équivalentes à  $\sim 1\%$  près selon les scores (cf. Table 4.11).

Les courbes de précision-rappel obtenues avec les CEGI sur 384 valeurs sont moins homogènes que les précédentes. La distance  $L_1$  donne les meilleures performances, et les moins bonnes performances sont encore obtenues par la distance  $L_\infty$  (cf. Figure 4.24). Les autres distances donnent des performances équivalentes à  $\sim 1\%$  près selon les scores (cf. Table 4.12).

Les performances globales des CEGI sont moyennes avec 50 valeurs et faibles avec 384 valeurs, avec  $T_1 = 73.5\%$  et  $56.5\%$  respectivement. Les CEGI avec 50 valeurs sont efficaces pour retrouver les formes simples comme les primitives (n° 1 à 5), les pièces d'échecs (n° 7), les vases (n° 8 à 11) et les pavés (n° 15), mais ont des difficultés pour les modèles humains (n° 12), les moules (n° 13) et les statues (n° 14) (cf. Figure 4.25). On remarque que les performances sont assez similaires aux histogrammes d'angles des cordes avec le 2<sup>e</sup> axe. Les formes plus complexes comme les modèles humains et les moules sont difficilement caractérisables par ces descripteurs "angulaires", quand bien même en augmentant le pas de quantification, trop sensibles aux variations de la surface des objets.

	$L_1$	$L_2$	$L_\infty$	Jeffreys	$\chi^2$
$T_1$	0.701	0.703	0.684	<b>0.735</b>	0.712
$SBE$	0.844	0.822	0.778	<b>0.864</b>	0.826
$E$	0.322	0.313	0.302	<b>0.331</b>	0.309
$DCG$	0.859	0.850	0.833	<b>0.869</b>	0.856

TAB. 4.11 – Scores ( $T_1$ ,  $SBE$ , mesure  $E$  et  $DCG$ ) sur la base de données de 116 modèles avec les CEGI sur 50 valeurs pour les diverses distances.

	$L_1$	$L_2$	$L_\infty$	Jeffreys	$\chi^2$
$T_1$	0.612	<b>0.565</b>	0.473	0.517	0.585
$SBE$	0.731	<b>0.690</b>	0.561	0.628	0.708
$E$	0.284	<b>0.276</b>	0.223	0.257	0.280
$DCG$	0.801	<b>0.775</b>	0.715	0.750	0.792

TAB. 4.12 – Scores ( $T_1$ ,  $SBE$ , mesure  $E$  et  $DCG$ ) sur la base de données de 116 modèles avec les CEGI sur 384 valeurs pour les diverses distances.

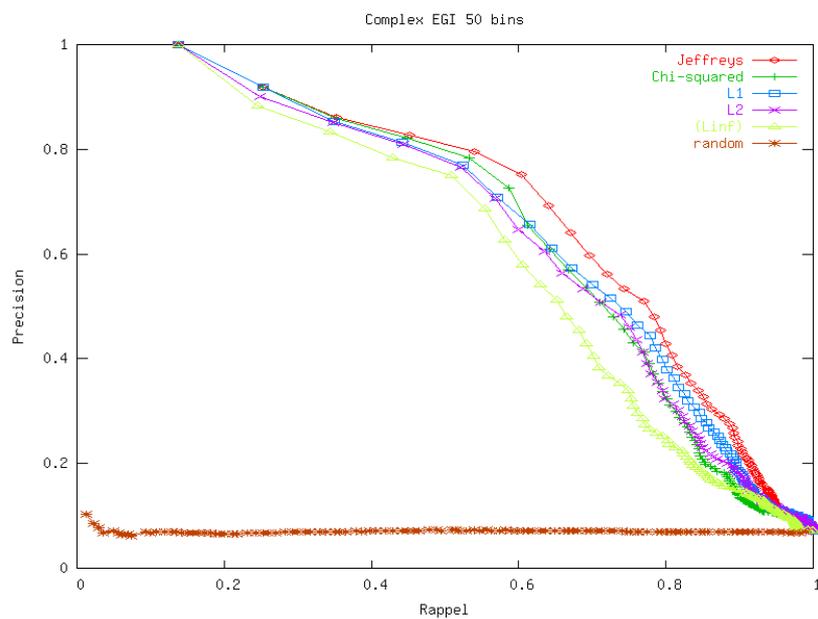


FIG. 4.23 – Courbes de précision-rappel pour les CEGI avec 50 valeurs. La meilleure distance est celle de Jeffreys.

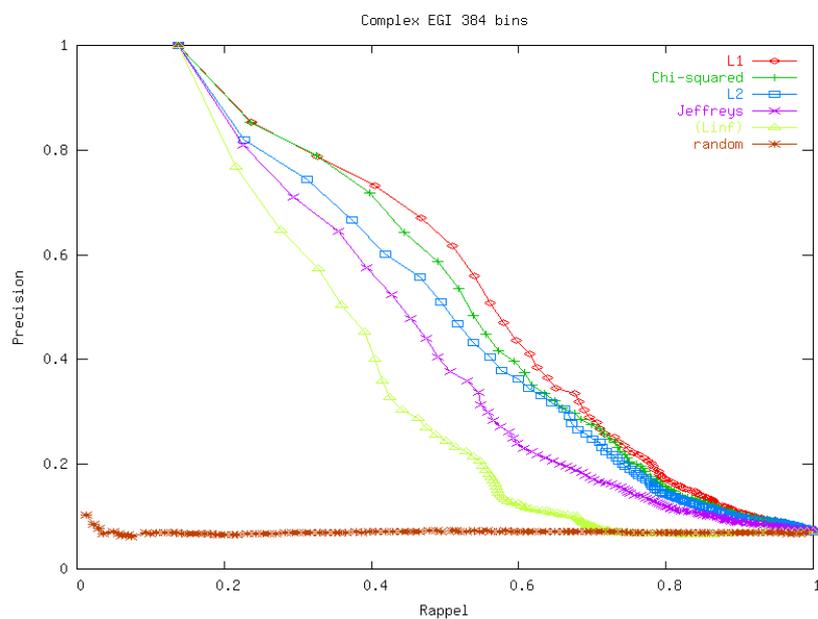


FIG. 4.24 – Courbes de précision-rappel pour les CEGI sur 384 valeurs. La meilleure distance est la  $L_1$ .

### 4.3 Etude des méthodes implémentées

---

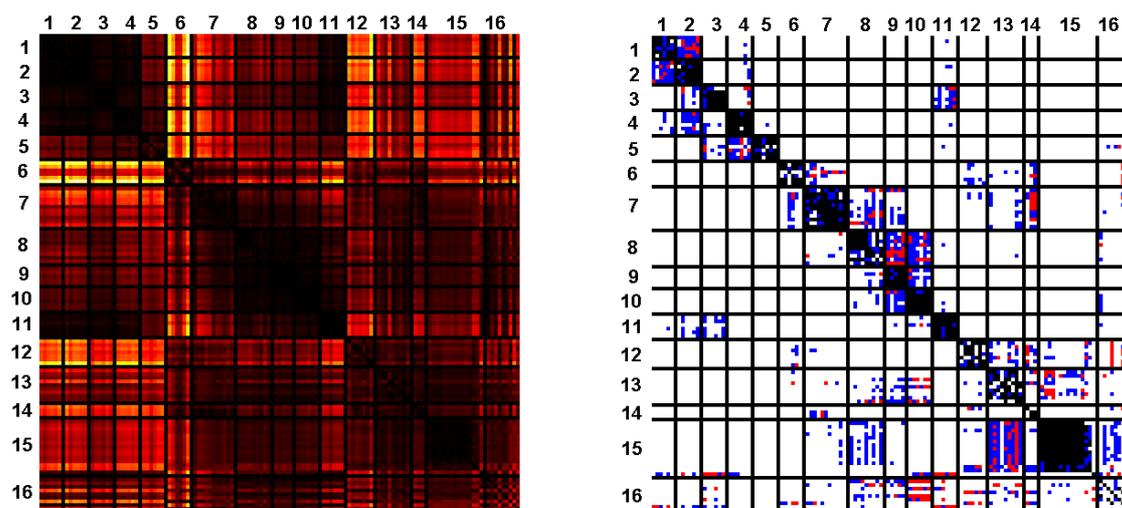


FIG. 4.25 – *Matrice de distances et image tier des CEGI avec 50 valeurs et la distance de Jeffreys.*

### 4.3.6 Le descripteur de spectre de forme 3D

Les courbes de précision-rappel obtenues avec les descripteurs de spectre de forme 3D de 16 et 128 valeurs montrent des résultats équivalents pour la majorité des distances (cf. Figures 4.26 et 4.27 respectivement). On observe que la distance de Jeffreys a donné les plus faibles résultats pour ce descripteur comparé aux autres distances.

Les scores montrent des performances moyennes sur la base de données de 116 objets avec seulement  $T_1 = 54.1\%$  et  $53.9\%$  pour les deux histogrammes de 16 et 128 valeurs (cf. Tables 4.13 et 4.14). Toutefois le descripteur de spectre de forme 3D, caractérisant les courbures locales, a été le seul à pouvoir retrouver parfaitement la classe de déesses déformées (n° 6) (cf. Figure 4.28). La classe de statues (n° 14) et les pavés (n° 15) ont bien été retrouvés aussi. Cependant les résultats sont très peu performants pour toutes les autres classes.

	$L_1$	$L_2$	$L_\infty$	Bhatta	Jeffreys	$\chi^2$	Histo int	EMD
$T_1$	0.526	0.508	0.494	0.539	0.494	<b>0.541</b>	0.527	0.515
$SBE$	0.661	0.615	0.619	0.659	0.652	<b>0.663</b>	0.660	0.654
$E$	<b>0.276</b>	0.268	0.259	0.270	0.278	0.273	0.276	<b>0.276</b>
$DCG$	0.743	0.736	0.723	0.751	<b>0.756</b>	0.751	0.741	0.749

TAB. 4.13 – Scores ( $T_1$ ,  $SBE$ , mesure  $E$  et  $DCG$ ) sur la base de données de 116 modèles avec le descripteur de spectre de forme 3D de 16 valeurs pour les diverses distances.

	$L_1$	$L_2$	$L_\infty$	Bhatta	Jeffreys	$\chi^2$	Histo int	EMD
$T_1$	0.531	0.494	0.453	0.506	0.489	<b>0.539</b>	0.524	0.517
$SBE$	0.651	0.613	0.592	0.629	0.619	0.647	0.647	<b>0.678</b>
$E$	0.265	0.250	0.245	0.257	0.266	0.262	0.266	<b>0.278</b>
$DCG$	0.737	0.716	0.700	0.729	<b>0.753</b>	0.738	0.737	0.748

TAB. 4.14 – Scores ( $T_1$ ,  $SBE$ , mesure  $E$  et  $DCG$ ) sur la base de données de 116 modèles avec le descripteur de spectre de forme 3D de 128 valeurs pour les diverses distances.

### 4.3 Etude des méthodes implémentées

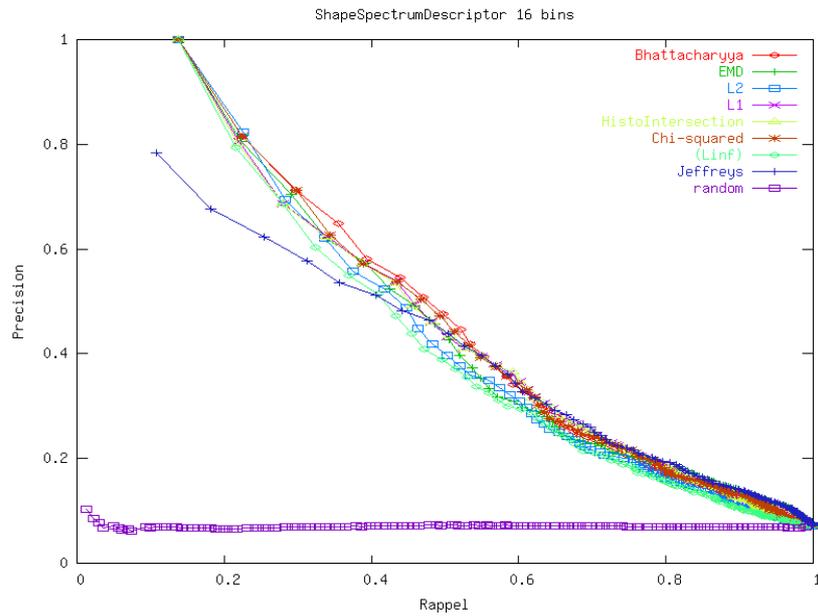


FIG. 4.26 – Courbes de précision-rappel pour le descripteur de spectre de forme 3D de 16 valeurs. Les résultats sont homogènes.

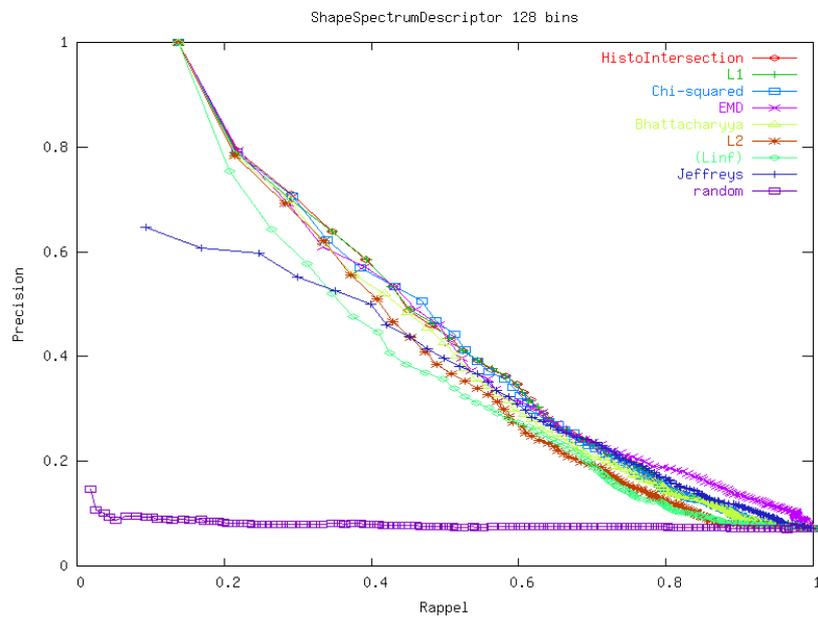


FIG. 4.27 – Courbes de précision-rappel pour le descripteur de spectre de forme 3D de 128 valeurs. Les résultats sont homogènes.

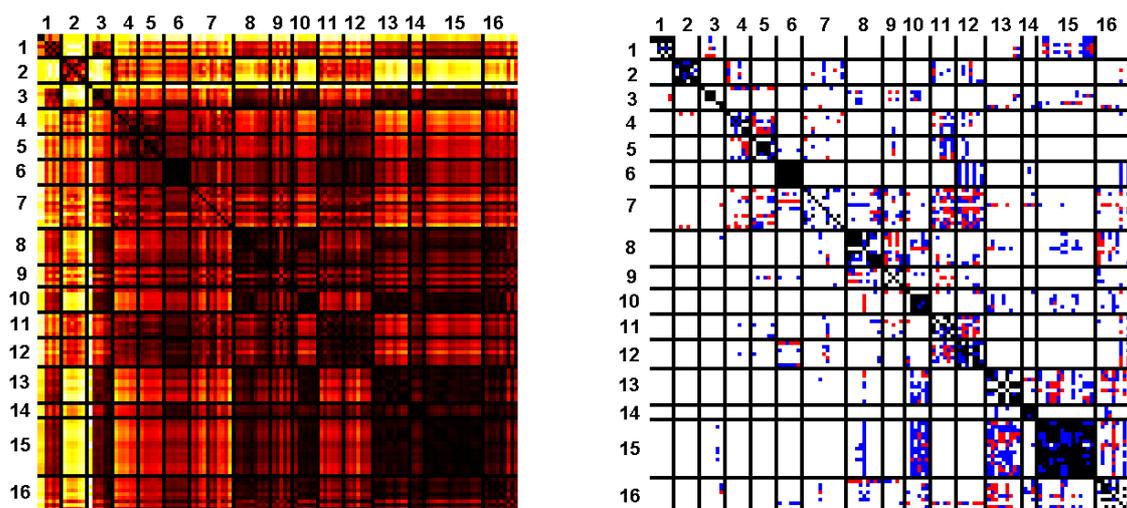


FIG. 4.28 – Matrice de distances et image tier du descripteur de spectre de forme 3D avec 16 valeurs et la divergence du  $\chi^2$ .

### 4.3.7 Le descripteur de Hough 3D

Les courbes de precision-rappel obtenues avec le descripteur de Hough 3D donnent les meilleurs résultats avec la distance Earth Mover comparées aux autres distances testées (cf. Figure 4.29). Cependant, il n'y a pas de différences significatives avec trois autres distances moins coûteuses ( $\chi^2$ , Bhattacharyya et l'intersection d'histogrammes).

Les résultats des tests correspondant à une simple subdivision de l'octaèdre sont meilleurs sur la base de données d'entraînement que ceux correspondant à deux subdivisions (cf. Figures 4.29 et 4.30). La distance EMD a donné les meilleurs résultats sur la première base de données mais n'a pu être appliquée sur la deuxième. En fait l'implémentation de la méthode de [Rubner et al., 1998] ne nous a pas permis d'utiliser le descripteur de  $4 * 128 = 512$  flottants. La taille du vecteur caractéristique que nous avons retenu est alors de  $4 * 32 = 128$  flottants. Les meilleures distances donnent  $T_1 = 79.2\%$  et  $76.5\%$  avec 128 et 512 valeurs respectivement, soit plus de 3/4 des objets retrouvés par classe (cf. Tables 4.15 et 4.16).

	$L_1$	$L_2$	$L_\infty$	Bhatta	Jeffreys	$\chi^2$	Histo int	EMD
$T_1$	0.758	0.681	0.600	0.761	0.487	0.773	0.759	<b>0.792</b>
$SBE$	0.854	0.783	0.678	0.860	0.609	0.866	0.852	<b>0.900</b>
$E$	<b>0.330</b>	0.297	0.273	0.328	0.261	<b>0.330</b>	<b>0.330</b>	0.329
$DCG$	0.881	0.836	0.787	0.885	0.725	<b>0.885</b>	0.881	<b>0.895</b>

TAB. 4.15 – Scores ( $T_1$ ,  $SBE$ , mesure  $E$  et  $DCG$ ) sur la base de données de 116 modèles avec le descripteur de Hough 3D de 128 valeurs pour les diverses distances.

	$L_1$	$L_2$	$L_\infty$	Bhatta	Jeffreys	$\chi^2$	Histo int
$T_1$	0.744	0.673	0.577	<b>0.765</b>	0.202	0.758	0.745
$SBE$	0.875	0.755	0.670	<b>0.886</b>	0.291	0.874	0.880
$E$	0.330	0.277	0.266	<b>0.332</b>	0.162	0.333	0.332
$DCG$	0.882	0.822	0.773	<b>0.884</b>	0.580	0.884	0.882

TAB. 4.16 – Scores ( $T_1$ ,  $SBE$ , mesure  $E$  et  $DCG$ ) sur la base de données de 116 modèles avec le descripteur de Hough 3D de 512 valeurs pour les diverses distances.

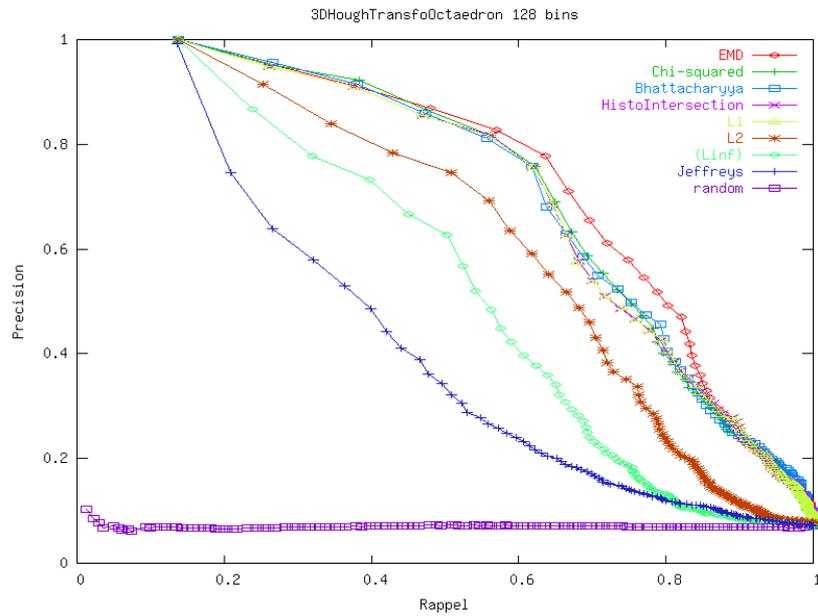


FIG. 4.29 – Courbes de précision-rappel pour le descripteur de Hough 3D avec 128 valeurs. La meilleure distance est l'EMD.

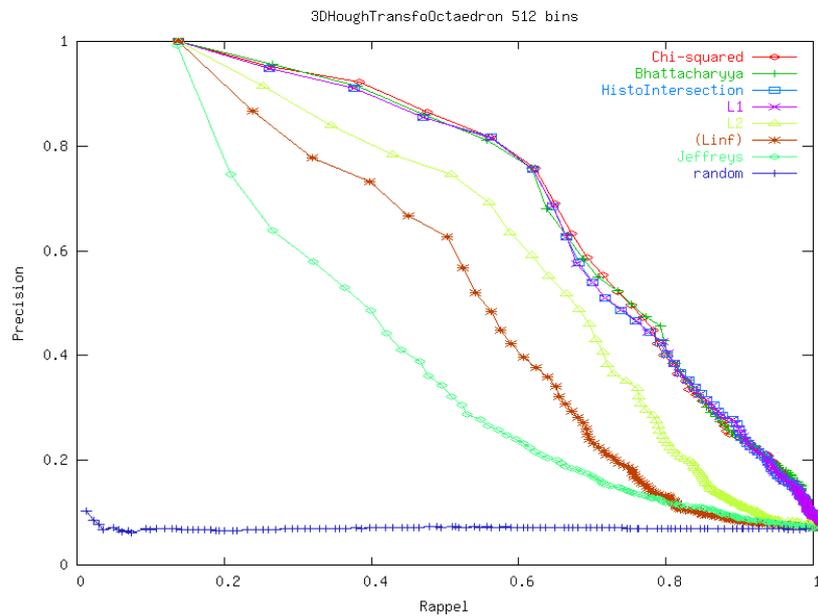


FIG. 4.30 – Courbes de précision-rappel pour le descripteur de Hough 3D avec 512 valeurs. Quatre distances retournent des scores équivalents.

### 4.3 Etude des méthodes implémentées

Les calculs de la matrice de distances et de l'image tier révèlent l'habileté du descripteur de Hough 3D pour retrouver 13 classes parmi 16. La classe inhomogène (n° 16) a correctement non été retrouvée, mais les classes de modèles humains et de moules (n° 12 et 13) sont seulement partiellement retrouvées (cf. Figure 4.31). Le descripteur de Hough 3D semble présenter des difficultés pour retrouver les classes de modèles humains et de moules (n° 12 et 13), alors qu'il produit de bons résultats avec les classes d'objets de formes primitives (n° 1 à 5), les vases (n° 8 à 11) et les pavés (n° 15). On constate également que les objets correspondant à des variantes déformées d'un même modèle ne sont pas bien retrouvés. Ceci peut être observé avec les tores (classe n° 4) et les statues de déesses (classe n° 6). Ce comportement du descripteur de Hough 3D est intrinsèquement dû au calcul de l'orientation des normales, tout comme pour les CEGI. Des variations géométriques importantes de la surface affecte de manière conséquente la représentation des vecteurs caractéristiques. De plus, l'approche originale ne prend en compte que les distances des plans à l'origine supérieures à 70% de la distance maximale. Les moules étant globalement plats et convexes, et les modèles humains filiformes, par conséquent ils sont difficilement caractérisables par cette approche.

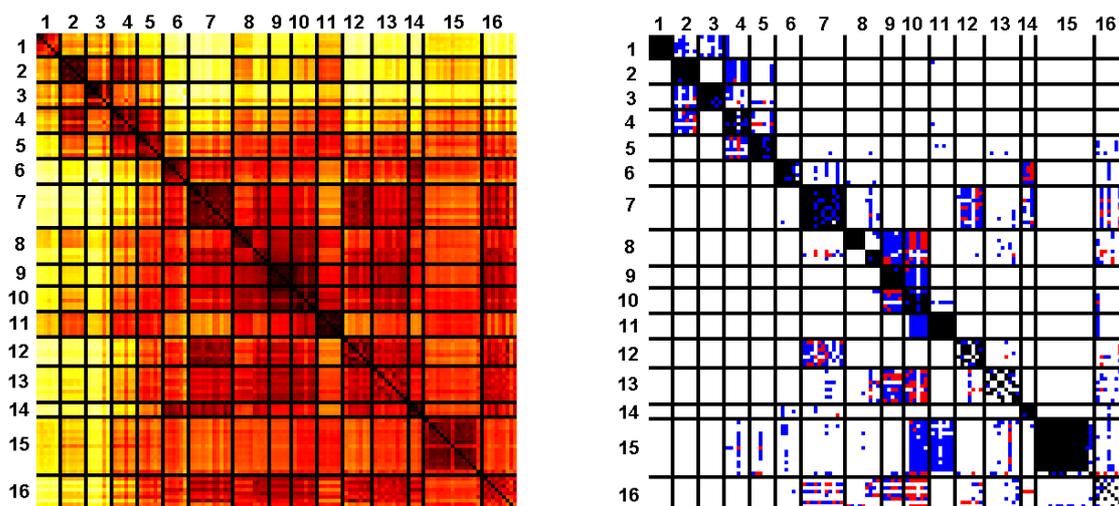


FIG. 4.31 – Matrice de distances et image tier du descripteur de Hough 3D sur 128 valeurs avec la distance EMD.

### 4.3.8 Bilan

Parmi toutes les méthodes que nous avons implémentées comme comparatifs à l'approche des aMRG que nous proposons, le descripteur de Hough 3D [Zaharia and Prêteux, 2002] a montré les meilleures performances sur la base d'entraînement de 116 objets avec la distance de Earth Mover, et avec des vecteurs caractéristiques de taille 128 ( $4 * 32$ ).

Les autres méthodes qui retournent des résultats corrects sont les CEGI et les histogrammes de cordes. La distribution de forme D2 et les harmoniques sphériques se sont révélées très faibles.

Les performances varient selon les descripteurs de même que l'habileté à retrouver certains types d'objets. Aucun descripteur n'a retrouvé parfaitement toutes les classes, et en particulier les classes de modèles humains et de moules (n° 12 et 13) ont posé des difficultés. Cela est d'autant plus gênant que les deux bases de données, plus grandes, contiennent beaucoup de fragments de moules.

Une solution plausible pour améliorer les résultats est de combiner les différents descripteurs, ce que nous proposons intrinsèquement dans notre approche par aMRG.

Ainsi, les résultats nous ont permis de choisir les descripteurs suivants comme attributs des nœuds :

- le ratio,
- les trois histogrammes de cordes de 16 valeurs,
- le descripteur de spectre de forme 3D de 16 valeurs,
- et le descripteur de Hough 3D de 128 valeurs.

Le descripteur de Hough 3D étant une variante améliorée des CEGI, ces dernières n'ont donc pas été incluses comme attributs.

## 4.4 Résultats des aMRG

En s'appuyant sur les résultats expérimentaux obtenus, nous pouvons à présent comparer notre approche par aMRG avec les diverses méthodes implémentées. Pour cela, nous avons choisit pour chacune d'entre elles la distance qui a fourni les meilleurs résultats aux tests précédents. On note qu'à  $r = 0$  le calcul des aMRG est équivalent à une combinaison des méthodes classiques, et que les résultats sont inférieurs à ceux obtenus aux résolutions supérieures. Les aMRG ont été calculés à la résolution  $r = 4$ , avec les attributs topologiques présentés en section 3.1.6 pour l'appariement des nœuds, et avec les intersections d'histogrammes comme distance (introduite dans les fonctions de similarité *sim* en section 3.3). Des tests aux différents niveaux de résolution  $r = 0$  à  $r = R = 5$ , et avec toutes les distances nous ont permis de faire ces choix (cf. Tableau 4.17). Les résultats à  $r = 4$  et  $r = 5$  étant très proches, nous avons opté pour la plus petite résolution car les temps de calculs sont plus courts.

$r$	0	1	2	3	4	5
$T_1$	0.867	0.868	0.882	0.905	<b>0.912</b>	0.914
$SBE$	0.962	0.964	0.976	0.983	<b>0.985</b>	0.984
$E$	0.346	0.346	0.346	0.348	<b>0.348</b>	0.347
$DCG$	0.921	0.922	0.928	0.932	<b>0.935</b>	0.935

TAB. 4.17 – Scores ( $T_1$ ,  $SBE$ , mesure  $E$  et  $DCG$ ) sur la base de données de 116 modèles pour les aMRG aux résolutions  $r = 0$  à 5, et avec la distance intersection d'histogrammes.

Le tableau 4.18 présente les résultats obtenus sur la base de données de 116 objets.

	Shape dist D2	Sph harmo	Vol area ratio	Shape spec	Cord length	CEGI	3D Hough desc	aMRG $r = 4$
	L1	$\chi^2$	L1	$\chi^2$	EMD	Jeffreys	EMD	Histo int
$T_1$	0.337	0.383	0.491	0.541	0.691	0.735	0.792	<b>0.912</b>
$SBE$	0.458	0.529	0.681	0.663	0.808	0.864	0.900	<b>0.985</b>
$E$	0.219	0.247	0.296	0.273	0.308	0.331	0.329	<b>0.348</b>
$DCG$	0.625	0.657	0.726	0.751	0.850	0.869	0.895	<b>0.935</b>

TAB. 4.18 – Scores ( $T_1$ ,  $SBE$ , mesure  $E$  et  $DCG$ ) sur la base de données de 116 modèles pour les meilleures combinaisons de distances avec les méthodes implémentées.

Sur cette base de données, les aMRG obtiennent largement les meilleurs scores puisqu'ils retrouvent en moyenne avec le score first tier  $T_1 = 91.2\%$  des objets de chaque classe. Lorsque l'on considère deux fois la taille de chaque classe ( $SBE$ ), les aMRG retrouvent 98.5% des objets. Les courbes de précision-rappel démontrent aussi la supériorité des aMRG comparée aux autres méthodes sur cette première base de

données (cf. Figure 4.32). La matrice de distances et l'image tier montrent l'habileté des aMRG à retrouver les classes (cf. Figure 4.33), excepté pour la classe hétérogène n° 16 composée de statues de formes différentes. Les formes primitives (n° 1 à 5) sont bien classifiées, ainsi que les différents vases dont les classes (n° 8 à 11) restent bien groupées. La classe de moule n° 13 a aussi bien été retrouvée et est proche de la classe des pavés n° 15. En comparaison avec les autres méthodes, les aMRG ont ainsi l'avantage de produire une classification plus homogène (cf. Figure 4.33 droite).

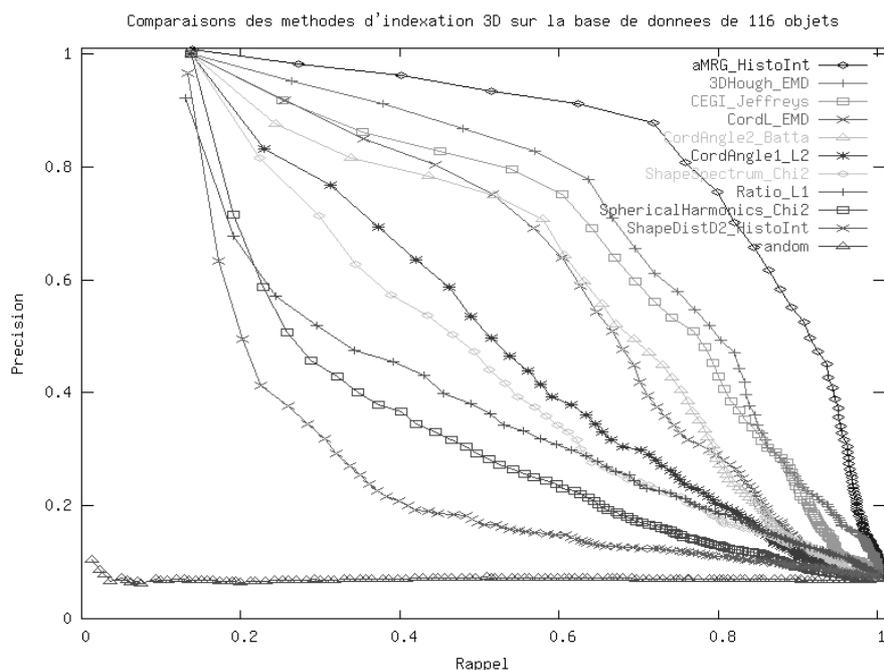


FIG. 4.32 – Courbes de précision-rappel pour les méthodes d'indexation 3D testées sur la première base de données de 116 objets. Les aMRG donnent les meilleures performances.

Pour ces tests, les attributs et poids des aMRG ont été choisis empiriquement comme suit :

- $f_0 = a$ ,  $\lambda_0 = 0.2$ ,
- $f_1 = cordL$ ,  $\lambda_1 = 0.2$ ,
- $f_2 = cord1$ ,  $\lambda_2 = 0.1$ ,
- $f_3 = cord2$ ,  $\lambda_3 = 0.1$ ,
- $f_4 = curv$ ,  $\lambda_4 = 0.1$ ,
- $f_5 = hough$ ,  $\lambda_5 = 0.2$ ,
- $f_6 = ratio$ ,  $\lambda_6 = 0.1$ .

Les résultats présentés à la Figure 4.34 correspondent à la matrice de distances et l'image tier obtenues avec l'approche classique des MRG [Hilaga et al., 2001]. On observe une beaucoup plus grande dispersion des résultats de la classification et une bien moins grande qualité des résultats. On peut donc conclure que l'ajout de critères et d'attributs supplémentaires ayant conduit aux aMRG est tout à fait pertinent.

#### 4.4 Résultats des aMRG

L'approche classique des MRG donne les scores suivants :  $T_1 = 0.661$ ,  $SBE = 0.790$ ,  $E = 0.320$  et  $DCG = 0.830$ , à comparer aux scores obtenus avec les aMRG dans le tableau 4.17 ( $T_1 = 0.912$ ,  $SBE = 0.985$ ,  $E = 0.348$  et  $DCG = 0.935$ ).

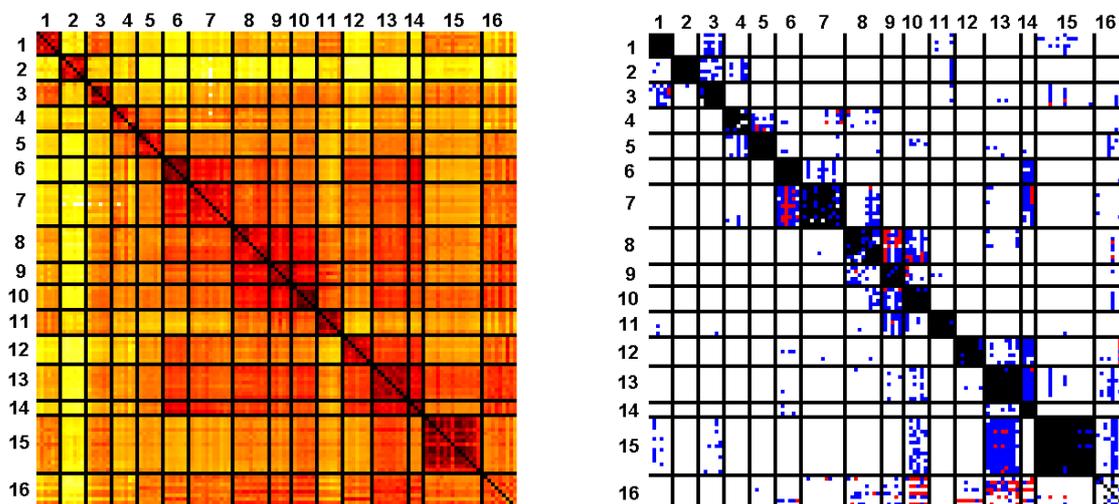


FIG. 4.33 – Gauche : matrice de distances des aMRG. Droite : image tier des aMRG. Les aMRG retrouvent bien les classes (exceptés la n° 16).

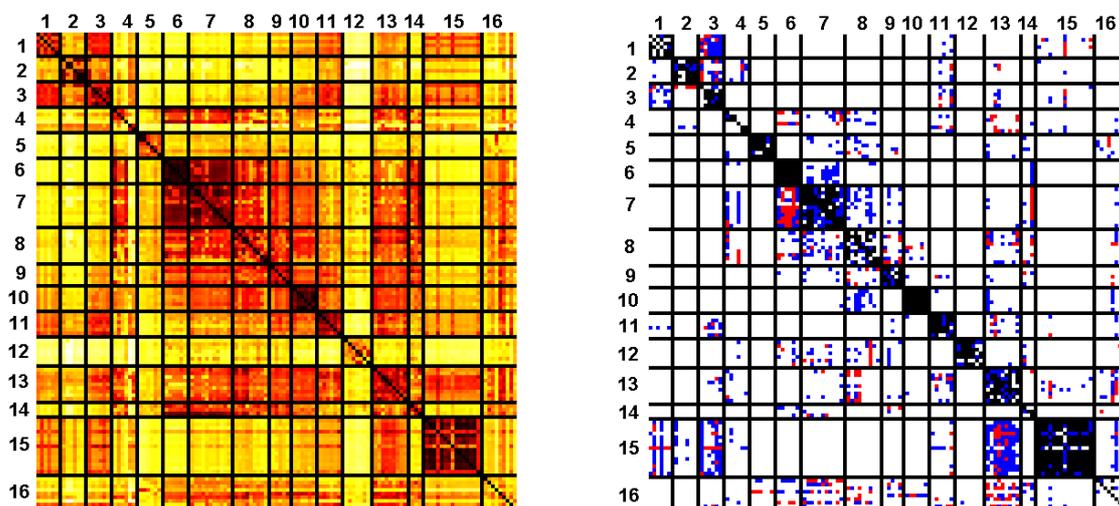


FIG. 4.34 – Gauche : matrice de distances des MRG .Droite : image tier des MRG. L'approche originale des MRG donne de moins bonnes performances que l'approche des aMRG.

Les scores suivants sont obtenus lorsque l'on utilise uniquement dans l'approche classique des MRG les attributs topologiques pour l'appariement des graphes (cf. sec-

tion 3.1.6) à la place des attributs  $a$  et  $\mu$  originellement utilisés dans [Hilaga et al., 2001] :  $T_1 = 0.712$ ,  $SBE = 0.831$ ,  $E = 0.320$  et  $DCG = 0.853$ . L'introduction de ces attributs topologiques que nous avons proposés pour la mise en correspondance de nœuds entre deux graphes permet donc d'obtenir de réels gains.

Les tests effectués sur la deuxième base de données des 266 objets classés montrent aussi que l'approche des aMRG donne les meilleurs résultats pour retrouver les objets et classes d'objets similaires (cf. Figure 4.36).

L'étude des aMRG à différents niveaux de résolution avec les mêmes paramètres que le test précédent révèle que, sur cette base de données, les aMRG donnent de meilleurs résultats avec des calculs à la résolution  $r = 3$  et avec la distance d'intersection d'histogrammes (cf. Tableau 4.19 et Figure 4.35).

$r$	0	1	2	3	4	5
$T_1$	0.749	0.750	0.762	<b>0.770</b>	0.747	0.733
$SBE$	0.858	0.859	0.871	<b>0.874</b>	0.852	0.833
$E$	0.391	0.391	0.394	<b>0.394</b>	0.387	0.380
$DCG$	0.871	0.872	0.875	<b>0.880</b>	0.867	0.859

TAB. 4.19 – Scores ( $T_1$ ,  $SBE$ , mesure  $E$  et  $DCG$ ) sur la base de données de 266 modèles pour les aMRG aux résolutions  $r = 0$  à 5, et avec la distance intersection d'histogrammes.

Les méthodes ont été de nouveau testées avec toutes les distances, et on a constaté que les scores étaient cohérents avec ceux obtenus sur la base de 116 objets. Les performances obtenues par les diverses distances sont globalement les mêmes sur les deux bases : les distances qui donnent les meilleurs résultats sur la première base restent parmi celles qui donnent les meilleurs résultats sur la deuxième base, et des distances donnant des performances proches entre elles sur la première base donnent aussi des performances proches entre elles sur la deuxième base. Cependant nos tests révèlent que certaines distances qui n'étaient pas les plus performantes sur la première base de données (malgré de bons scores) ont donné de meilleurs résultats sur cette deuxième base. En particulier on obtient avec le descripteur de Hough 3D associé à la divergence du  $\chi^2$  :  $T_1 = 0.703$ ,  $SBE = 0.815$ ,  $E = 0.378$  et  $DCG = 0.849$ . Le tableau 4.20 reporte les résultats obtenus sur la base de données de 266 objets. Nous présentons les résultats des distances précédemment établies afin de rester homogène avec le tableau précédent. On observe que le descripteur de Hough 3D retourne toujours de meilleurs résultats que les autres approches testées.

Pour conclure notre évaluation nous avons effectué des tests sur la base complète de 567 objets, et les aMRG se sont encore révélés plus performants que les autres méthodes. Les aMRG ont été calculés au niveau de résolution  $r = 3$ , et avec les mêmes paramètres que le test précédent. Les autres méthodes ont été testées avec les distances

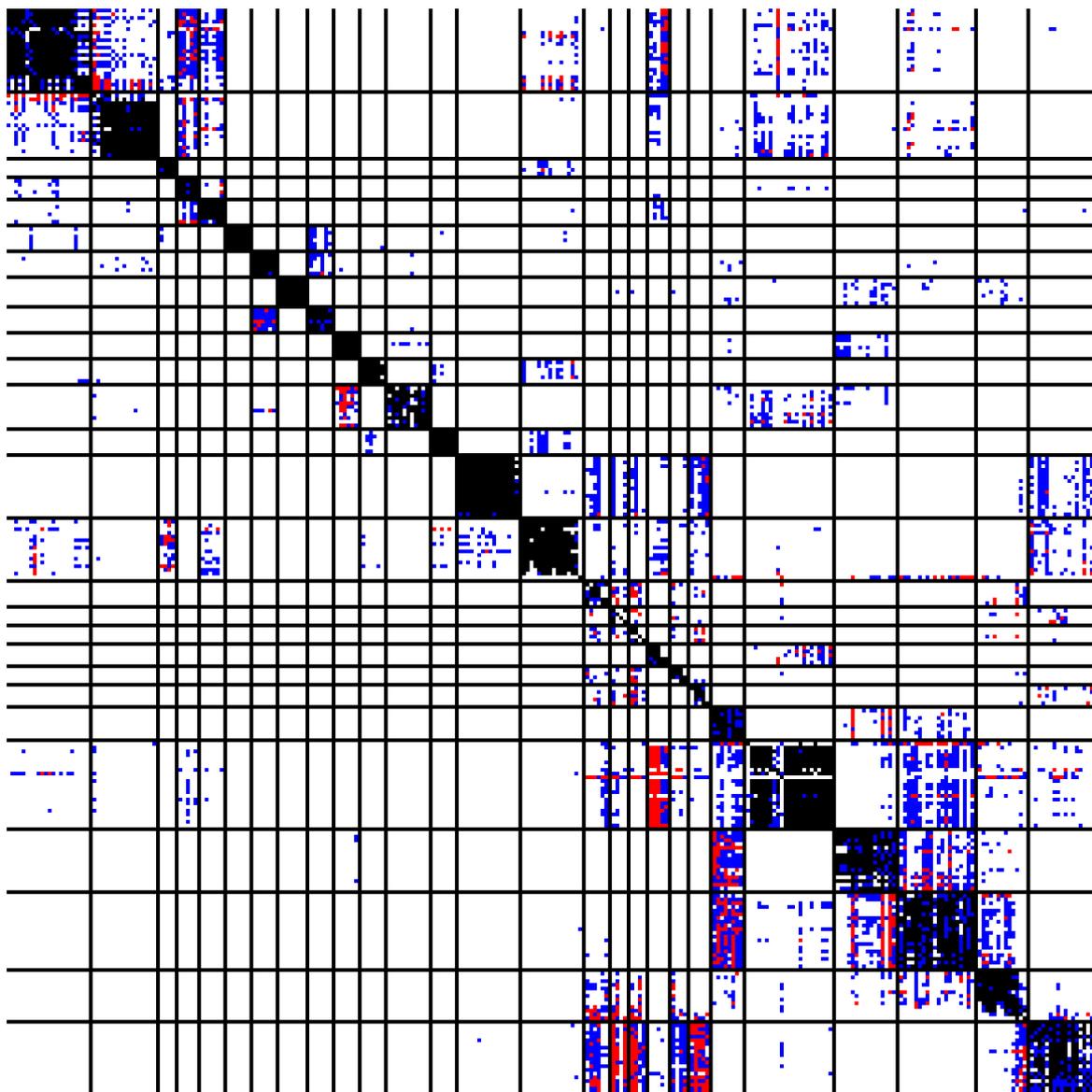


FIG. 4.35 – Image tier des aMRG à  $r = 3$  sur la base de données de 266 objets. La plupart des classes sont bien retrouvées.

précédemment établies.

On constate que les performances de toutes les méthodes ont diminué (cf. Figure 4.37 et Table 4.21). La grande quantité d'objets non classés ajoutés à la troisième base perturbe fortement les recherches, en particulier la mesure E qui dépend des 32 premiers objets retournés.

On observe toutefois que le score DCG des aMRG reste haut, ce qui signifie que tous

les objets d'une même classe sont retrouvés proches. Par ailleurs on observe encore une fois que le descripteur de Hough 3D donne de bien meilleurs résultats que les autres méthodes, bien qu'ils restent inférieurs en performances aux aMRG.

	Shape dist D2	Sph harmo	Vol area ratio	Shape spec	Cord length	CEGI	3D Hough desc	aMRG $r = 3$
	L1	$\chi^2$	L1	$\chi^2$	EMD	Jeffreys	EMD	Histo int
$T_1$	0.250	0.294	0.304	0.404	0.551	0.614	0.667	<b>0.770</b>
$SBE$	0.326	0.416	0.468	0.499	0.656	0.751	0.790	<b>0.874</b>
$E$	0.160	0.205	0.246	0.250	0.311	0.354	0.360	<b>0.394</b>
$DCG$	0.555	0.595	0.637	0.672	0.767	0.804	0.835	<b>0.880</b>

TAB. 4.20 – Scores ( $T_1$ ,  $SBE$ , mesure  $E$  et  $DCG$ ) sur la base de données de 266 modèles pour les meilleures combinaisons de distances avec les méthodes implémentées.

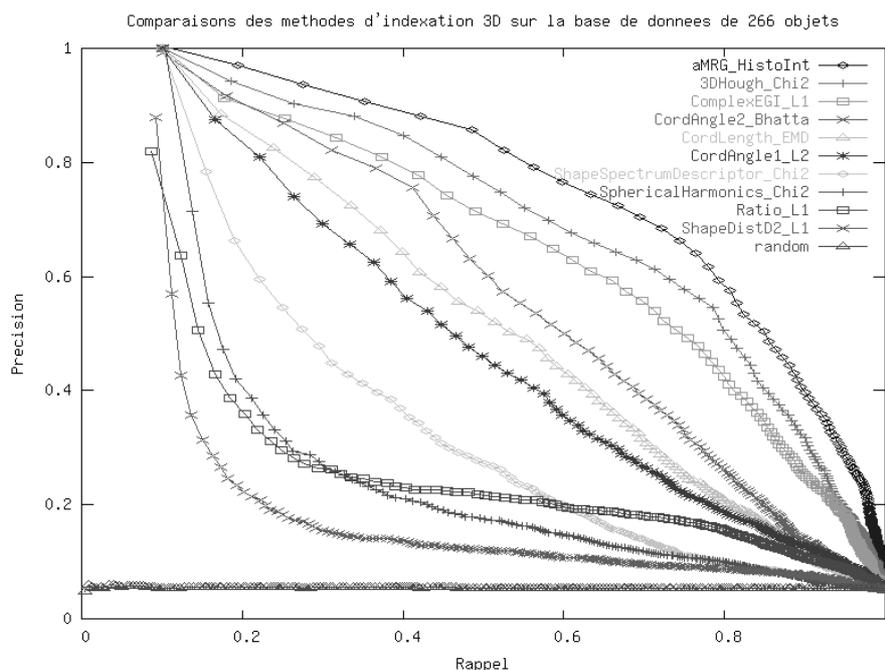


FIG. 4.36 – Courbes de précision-rappel pour les méthodes d'indexation 3D testées sur la deuxième base de données de 266 objets. Les aMRG donnent les meilleures performances.

#### 4.4 Résultats des aMRG

	Shape dist D2	Sph harmonic	Vol area ratio	Shape spec	Cord length	CEGI	3D Hough desc	aMRG $r = 3$
	L1	$\chi^2$	L1	$\chi^2$	EMD	Jeffreys	$\chi^2$	Histo int
$T_1$	0.201	0.249	0.243	0.332	0.467	0.526	0.640	<b>0.699</b>
$SBE$	0.248	0.314	0.324	0.415	0.568	0.657	0.752	<b>0.820</b>
$E$	0.114	0.155	0.158	0.194	0.268	0.318	0.352	<b>0.367</b>
$DCG$	0.500	0.538	0.574	0.599	0.702	0.746	0.813	<b>0.841</b>

TAB. 4.21 – Scores ( $T_1$ ,  $SBE$ , mesure  $E$  et  $DCG$ ) sur la base de données de 567 modèles pour les meilleures combinaisons de distances avec les méthodes implémentées.

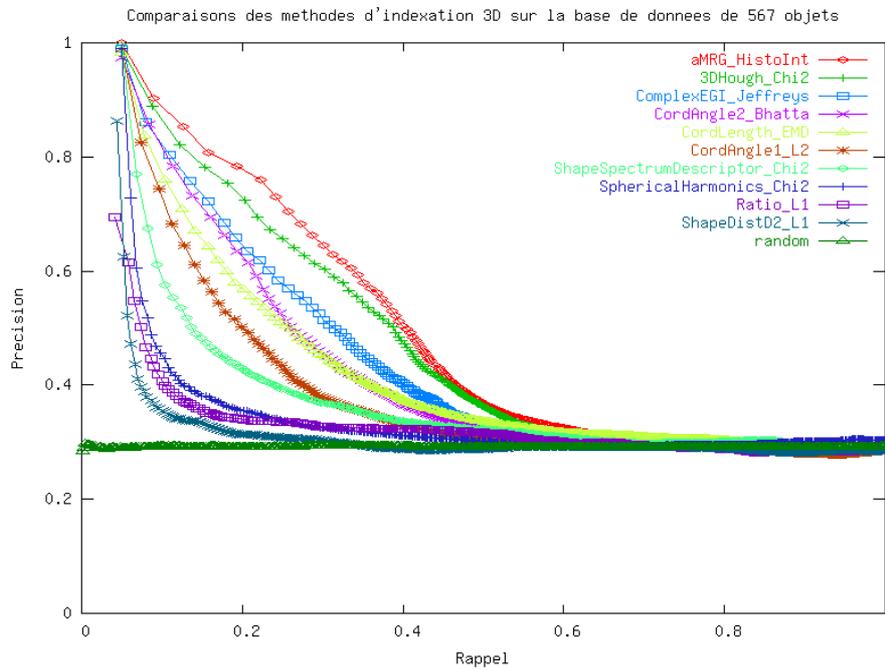


FIG. 4.37 – Courbes de precision-rappel des methodes d'indexation 3D testees sur la troisieme base de donnees de 567 objets. Les aMRG donnent les meilleures performances.

## 4.5 Bilan

Nos résultats sont encourageants car les paramètres  $\lambda_i$  ne sont pas optimaux. Un meilleur jeu de paramètres devrait davantage améliorer les résultats. Nous avons effectué plusieurs tests inspirés de la méthode de minimisation de paramètres de Powell (la fonction à minimiser étant  $1-T_1$ ) afin de trouver le meilleur jeu de paramètres, et les valeurs que nous proposons correspondent à un minimum local. Nous n'avons pas poursuivi la recherche de paramètres optimaux car l'évaluation des résultats reste subjective, dépendant de la base de données et des désirs de l'utilisateur. Intuitivement on peut sélectionner "à la main" des paramètres pour une requête spécifique où on est plus intéressé par soit les propriétés topologiques de l'objet, soit ses courbures, soit ses dimensions, etc.

On peut également faire remarquer que les approches par appariement de graphes semblent a priori mal adaptées pour les objets compacts comme les sphères où les moules, les graphes des objets compacts étant assez peu significatifs. Mais les expérimentations effectuées sur notre base de données nous montrent cependant que les aMRG parviennent très bien à caractériser ces formes. Cela est également dû aux attributs ajoutés aux MRG qui permettent d'étendre les approches par graphes en y fusionnant les informations géométriques spécifiques aux autres méthodes.

Les figures 4.38, 4.39 et 4.40 présentent des exemples de requêtes effectuées sur la base de données de 567 objets avec les graphes de Reeb multirésolution augmentés et avec le descripteur de Hough 3D en comparaison (car il obtient aussi des performances intéressantes). Notre implémentation repose en partie sur les théories présentées au chapitre 2, et nécessite donc que les objets soient bien modélisés avec en particulier un maillage de leur surface manifold et fait d'une seule composante connexe. Cette condition s'est révélée assez restrictive quant au choix des modèles 3D provenant d'Internet, et c'est pourquoi nous n'avons pu effectuer de tests sur des bases de données plus grandes (comme par exemple la base de Princeton [Shilane et al., 2004]). Afin de pouvoir traiter le maximum de données, il est donc nécessaire de soit filtrer les modèles pour obtenir un maillage propre, soit modifier notre implémentation pour rendre la construction des aMRG plus permissive. Actuellement, nous pouvons estimer que notre approche se destine davantage à des utilisations spécialisées avec des bases de données de modèles 3D à haute résolution, et que pour des applications moins fines le descripteur de Hough 3D de [Zaharia and Prêteux, 2002] reste un très bon candidat (cf. comparaison de requêtes aux figures 4.40 et 4.41).

## 4.5 Bilan

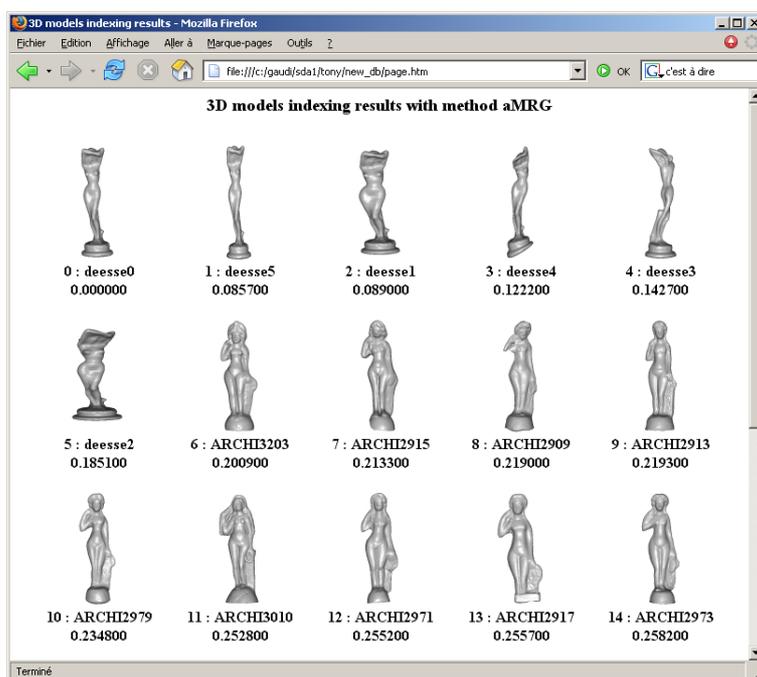


FIG. 4.38 – Exemple de requête avec les aMRG.

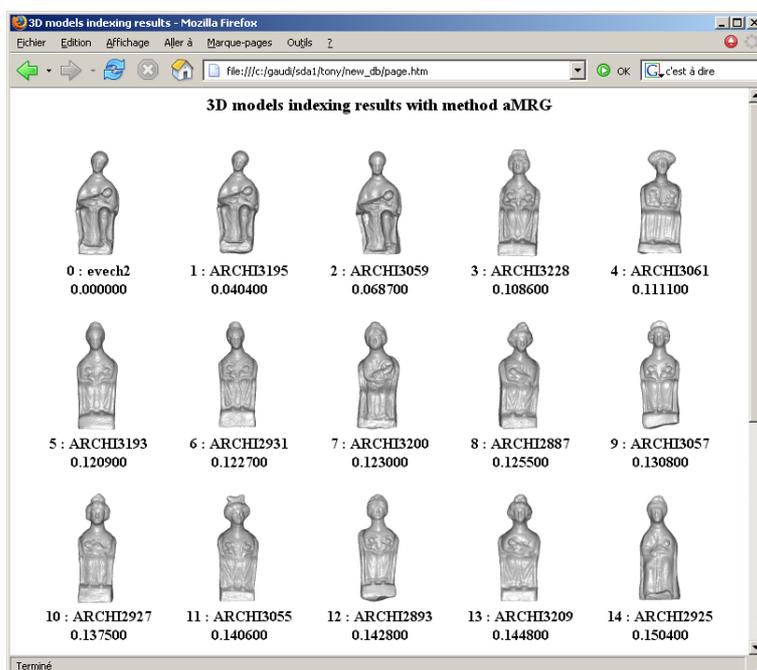


FIG. 4.39 – Exemple de requête avec les aMRG.

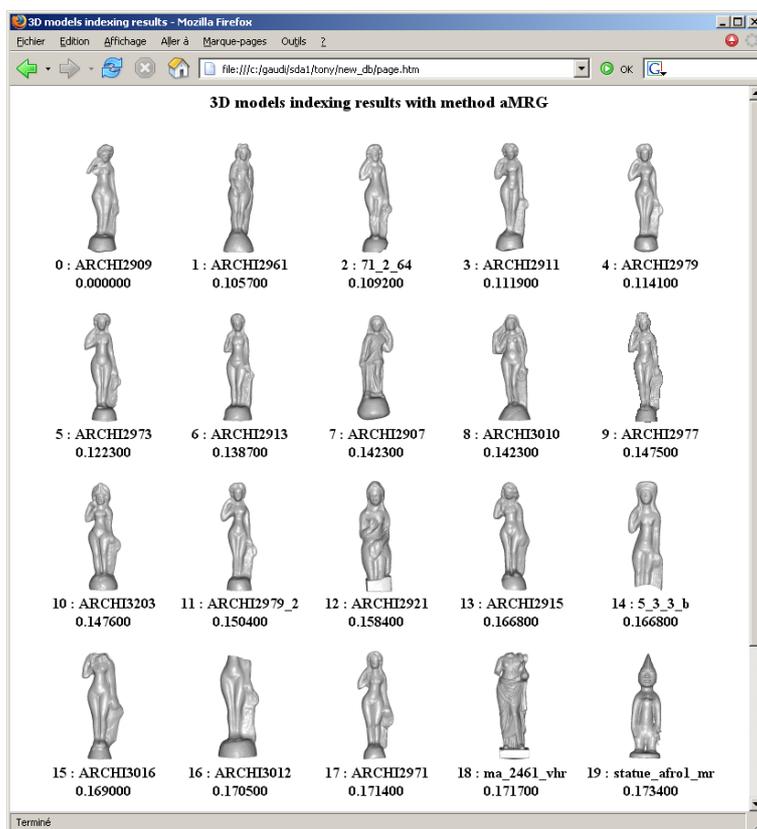


FIG. 4.40 – Exemple de requête avec les aMRG.

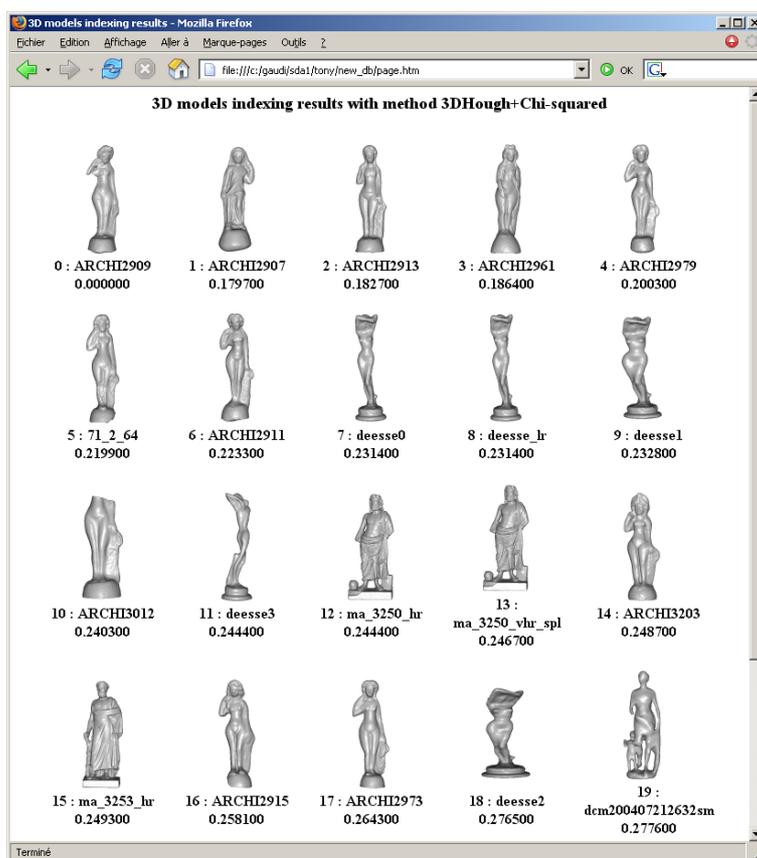


FIG. 4.41 – Exemple de requête avec le descripteur de Hough 3D.

## CHAPITRE 5

# Applications

Nous présentons dans ce chapitre trois applications mettant en œuvre notre approche par graphes de Reeb multirésolution augmentés. La première application présente le cas d'une recherche basée sur la forme et les couleurs. La structure des aMRG permet de prendre en compte la texture des objets (cf. Section 3.4.6). Nous présentons les résultats obtenus sur une petite base de données d'objets 3D texturés que nous avons constituée pour effectuer nos tests. La deuxième application utilise directement les algorithmes d'appariements décrits dans ce manuscrit pour effectuer des recherches de sous-graphes. On se sert alors de l'aspect multirésolution et de la pertinence des informations que nous avons ajoutées pour retrouver dans une base de données de modèles 3D des formes similaires à une partie d'un objet 3D (par exemple une main). La troisième application met en œuvre une approche de requête dessinée (dans l'esprit de celle développée à Princeton [Funkhouser et al., 2003]). Nous proposons un système de reconnaissance de formes 2D/3D : l'utilisateur dessine une forme 2D ; l'approche des aMRG est appliquée afin d'extraire le graphe et les attributs correspondants et retrouver les objets 3D similaires dans une base de données.

### 5.1 Recherche d'objets 3D texturés

Les résultats présentés au chapitre précédent ont permis de mettre en évidence les performances des aMRG pour retrouver les classes d'objets similaires par la forme. Cependant, les utilisateurs peuvent être aussi intéressés par des requêtes sur les couleurs des modèles. Par exemple, on pourrait souhaiter retrouver tous les objets ayant une forme et une couleur similaires. Dans le cadre de notre étude, ceci est réalisable car certains objets 3D possèdent des cartes de texture d'où il est possible d'extraire l'ensemble des texels couleur associés à chaque triangle du maillage de surface.

#### Exemple :

Nous étudions la faisabilité de cette application avec des requêtes sur la classe des 15 pavés. Les résultats présentés à la section 4.4 montrent que les aMRG n'ont aucune difficulté à retrouver tous les objets de cette classe à partir de requête sur la forme. Nous proposons d'affiner les résultats en utilisant la couleur des modèles.

Pour caractériser la couleur des modèles, nous avons utilisé trois histogrammes de

## 5.1 Recherche d'objets 3D texturés

couleur RVB calculés à partir des cartes de texture associées aux objets 3D. Il est évidemment possible d'utiliser d'autres types de caractérisation plus sophistiqués (ex : histogrammes Lab) comme ceux présentés en section 19. Les histogrammes RVB ont été calculés au niveau de résolution le plus fin, et introduits dans les nœuds comme nous le proposons en section 3.4.6.

Les premiers tests consistent à retrouver les pavés ayant les critères colorimétriques les plus proches de ceux des requêtes. Les aMRG ont été calculés à la résolution  $R = 5$  et seul les trois histogrammes RVB ont été utilisés dans un premier temps pour calculer la similarité des objets. Les figures 5.1 et 5.2 présentent les résultats des recherches des pavés de couleur "rouge brique" et de couleur "vert clair". On observe que les objets retournés sont cohérents avec la requête <sup>1</sup>.

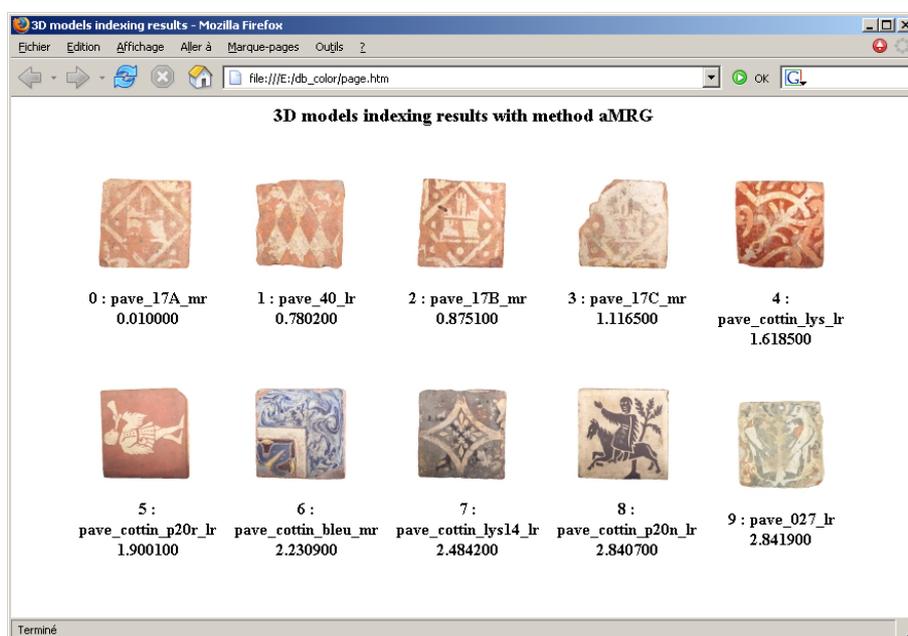


FIG. 5.1 – Résultats de la recherche d'un pavé couleur "rouge brique".

<sup>1</sup>Ici il faut faire attention lors de l'interprétation des résultats à bien tenir compte de la couleur du dos des pavés qui est entièrement rouge brique.



FIG. 5.2 – Résultats de la recherche d’un pavé couleur “vert clair”.

La figure 5.3 présente les résultats prometteurs d’une requête sur la forme et la couleur. Nous avons constitué une base de données de tests comprenant divers objets 3D texturés (vases, statues, pavés, etc.). La difficulté, intrinsèque à la méthode des aMRG, consiste à trouver les bons paramètres de pondération pour les différents attributs géométriques et colorimétriques. Pour ces tests, nous nous sommes basés sur les valeurs des précédents tests (cf. Section 4.4) et nous avons choisi empiriquement :

- $f_0 = a$ ,  $\lambda_0 = 0.2$ ,
- $f_1 = cordL$ ,  $\lambda_1 = 0.1$ ,
- $f_2 = cord1$ ,  $\lambda_2 = 0.1$ ,
- $f_3 = cord2$ ,  $\lambda_3 = 0.1$ ,
- $f_4 = curv$ ,  $\lambda_4 = 0.1$ ,
- $f_5 = hough$ ,  $\lambda_5 = 0.2$ ,
- $f_6 = histo_{RVB}$ ,  $\lambda_6 = 0.2$ , les 3 histogrammes étant équirépartis.

Les résultats obtenus sur la base de tests sont cohérents compte tenu des valeurs des paramètres. On constate que plus la couleur est privilégiée par rapport à la géométrie, et plus les premiers objets retrouvés ont bien une forme et une couleur similaires.

Le paramètre  $\lambda_6$  permet aux utilisateurs de choisir l’importance à accorder à la couleur. Pour déterminer le jeu de paramètres le plus adapté à une requête de similarité conjointe sur la forme et la couleur, il est indispensable de disposer pour l’apprentissage une base de données d’objets 3D texturés.

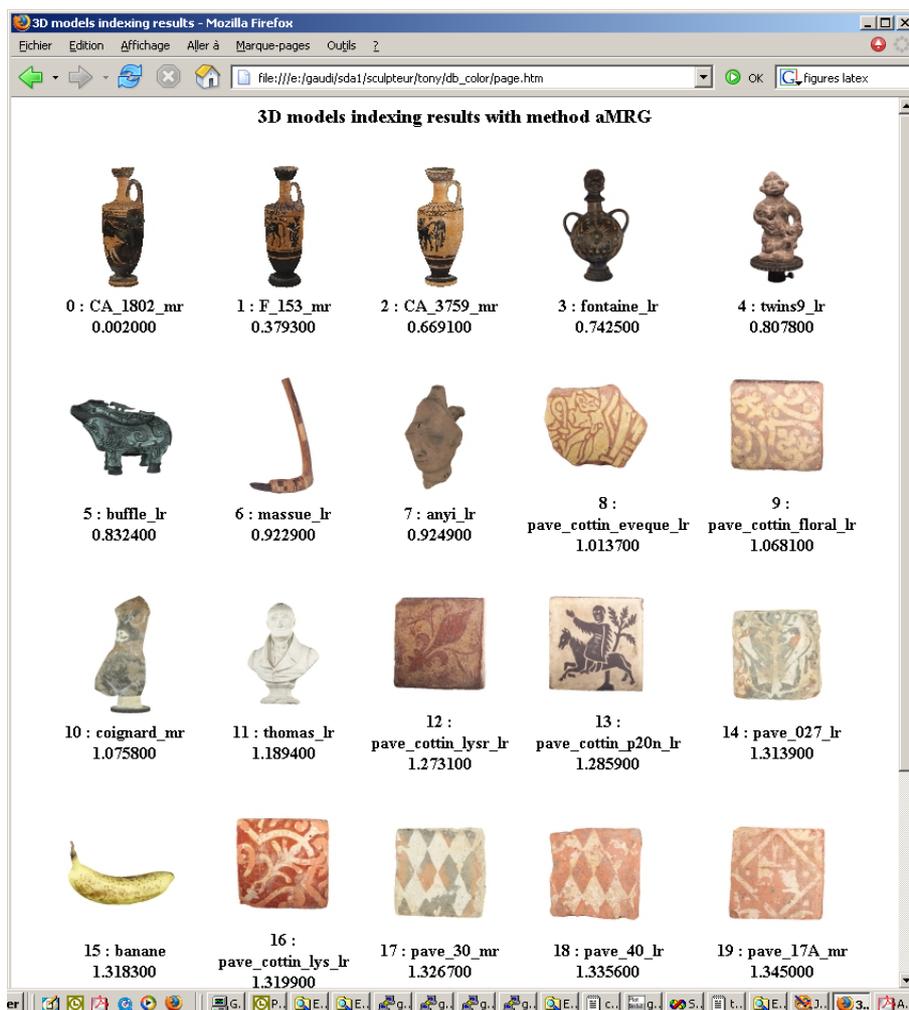


FIG. 5.3 – Résultat d'une recherche sur la forme et la couleur. Les premiers objets retrouvés ont une forme et des couleurs similaires à la requête.

## 5.2 Recherche de sous-graphes

Les aMRG sont à la fois des descripteurs de haut-niveau grâce à la représentation globale qu'apporte le graphe de Reeb multirésolution, et des descripteurs de bas-niveau avec les caractéristiques locales 2D et 3D extraites pour chaque nœud de chaque niveau de résolution, spécialement au niveau le plus fin.

L'aspect multirésolution peut être utilisé dans de nombreux contextes en particulier dans le cadre d'une application très prometteuse qui consiste à exploiter les nœuds enrichis des aMRG pour effectuer un appariement entre parties d'objets.

En effet, les informations qui ont été ajoutées dans les nœuds des graphes sont suffisamment riches d'information pour être discriminantes et permettre de retrouver dans une base de données des objets ayant une partie spécifique de leur forme similaire à celle d'un objet requête. Ce sont les différents sous-graphes aux différents niveaux de résolution, incluant aussi leurs nœuds fils qui assurent la précision de l'appariement.

### Schéma de l'appariement par parties

*Initialisation :*

1. Choix d'un objet  $M$ .
2. Choix d'un niveau de résolution.
3. Choix d'un nœud  $m$  requête à ce niveau de résolution, correspondant à la partie à appairier avec les autres parties de la base de données.

*Appariement :*

4. Choix des nœuds  $n$  parmi tous les objets  $N$  de la base de données topologiquement similaires à  $m$  : les nœuds  $n$  sont au même niveau de résolution que  $m$  (ou éventuellement à des niveaux de résolution proches), et on retient les nœuds  $n$  tels que  $loss(m, n) < s$ ,  $s$  étant un seuil prédéfini.
5. Nous considérons maintenant le sous-graphe de  $m$  et de ses nœuds fils comme un nouvel aMRG, qui est en fait l'aMRG de la région de l'objet tronqué associée au nœud  $m$ . On considère de même les sous-graphes de tous les nœuds cohérents  $n$  comme étant les aMRG des parties associées aux nœuds  $n$ . On applique ensuite l'algorithme général d'appariement de nœuds décrit dans le chapitre précédent à ces sous-graphes pour déterminer la partie la plus similaire à celle choisie dans l'objet  $M$  en tant que requête.

Une approche d'appariement de parties a été proposée dans [Sundar et al., 2003], mais la représentation du squelette des objets n'était pas bien adaptée à cette application. Le manque d'une stratégie d'appariement multirésolution est évident car par exemple, une partie de squelette pourrait être une description trop grossière pour être retrouvée ailleurs (un segment peut être apparié n'importe où).

Nous présentons un nouveau schéma d'appariement de parties reposant sur la procédure exposée au chapitre 3. Pour retrouver les nœuds les plus similaires nous exploitons l'aspect multirésolution des aMRG, les conditions de cohérences topologiques adaptées à l'appariement de sous-graphes, et l'extraction des diverses caractéristiques.

#### Exemple :

Nous illustrons ici un cas simple de recherche de partie d'objet en proposant la recherche d'une main d'un modèle humain dans une base de données de modèles 3D. Soit le modèle humain 3D *hunk* [3dcafe, 2005]. La construction du graphe aMRG à la résolution maximale  $R = 5$  nous permet d'observer qu'à la résolution  $r = 1$ , les sous-graphes des quatre nœuds terminaux positifs correspondent aux quatre membres du modèle (cf. figure 5.4).

Pour notre test, nous avons constitué une base de données d'environ 200 objets pris

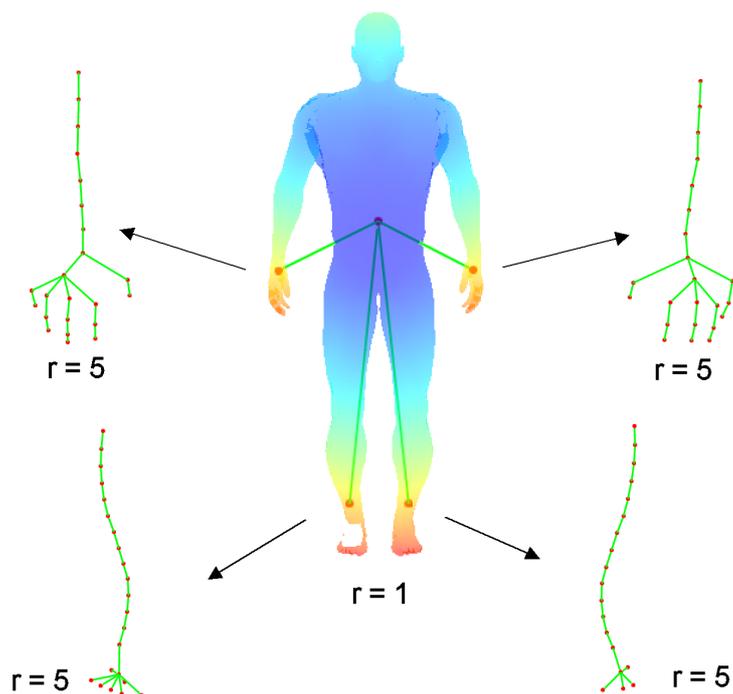


FIG. 5.4 – *Modèle hunk*. Les sous-graphes à  $r = 5$  des quatre nœuds terminaux à  $r = 1$  correspondent aux quatre membres du modèle.

aléatoirement parmi les objets de la base de données de 567 objets (cf. Section 4.1). La base est composée de modèles humains, de vases, de formes primitives et de statues (cf. Figure 5.5). Nous avons posé comme requête le nœud correspondant à la main droite du modèle *hunk* à  $r = 1$  (cf. Figure 5.4). Les quatre modèles humains de la base de données possèdent des mains. Les autres objets ne possèdent pas de structure topologique équivalente. Pour chaque objet comparé, l'algorithme recherche les sous-graphes les plus similaires en parcourant les nœuds à la résolution fixée pour la recherche.

Pour l'appariement des sous-graphes, la discrimination des nœuds (cf. Table 5.2 étape 4) est effectuée uniquement en utilisant les attributs topologiques proposés en 3.1.6. Soit  $m$  et  $n$ , deux nœuds à comparer, on choisit empiriquement  $s = 5$  et

$$\begin{aligned} \text{loss}(m, n) = & |Up_N(m) - Up_N(n)| + |Down_N(m) - Down_N(n)| \\ & + |Up_E(m) - Up_E(n)| + |Down_E(m) - Down_E(n)|. \end{aligned} \quad (5.1)$$

Le choix assez restrictif du seuil  $s = 5$  a permis de retrouver les nœuds ayant une structure topologique très proche de la requête, en l'occurrence des mains. En effet, on observe que les parties d'objets détectées correspondent aux mains droites et gauches des trois modèles humains de la base de données (*child*, *baby* et *alien*).

Pour ce test, les attributs topologiques ont été très discriminants mais étant indépendants de l'orientation, ils ne permettent pas de différencier les directions (la droite de la gauche). Ainsi, en supposant que les modèles ont été préalablement orientés, nous pouvons introduire l'information géométrique de localisation des nœuds  $(\mathbf{r}, \theta, \varphi)$  (cf. 3.1.7) dans le calcul de la similarité (cf. 5.2 étape 5) pour pallier les problèmes de miroir. Le choix des autres attributs des aMRG est moins évident et significatif pour ce test. Cependant, on remarque qu'en faisant varier la valeur des poids l'utilisateur peut retrouver les mains d'autres modèles humains.

La figure 5.6 présente les sous-graphes correspondant à la partie de l'objet requête et aux parties des objets retrouvées en premiers parmi les objets de la base de données. La partie de l'objet requête correspond à la main droite du modèle *hunk*, et la partie de l'objet retrouvée en première position correspond à la main droite du modèle *alien*. Les parties des objets retrouvées en deuxième et troisième positions correspondent respectivement aux mains droites des modèles *child* et *baby*. La méthode proposée retourne donc des résultats cohérents.

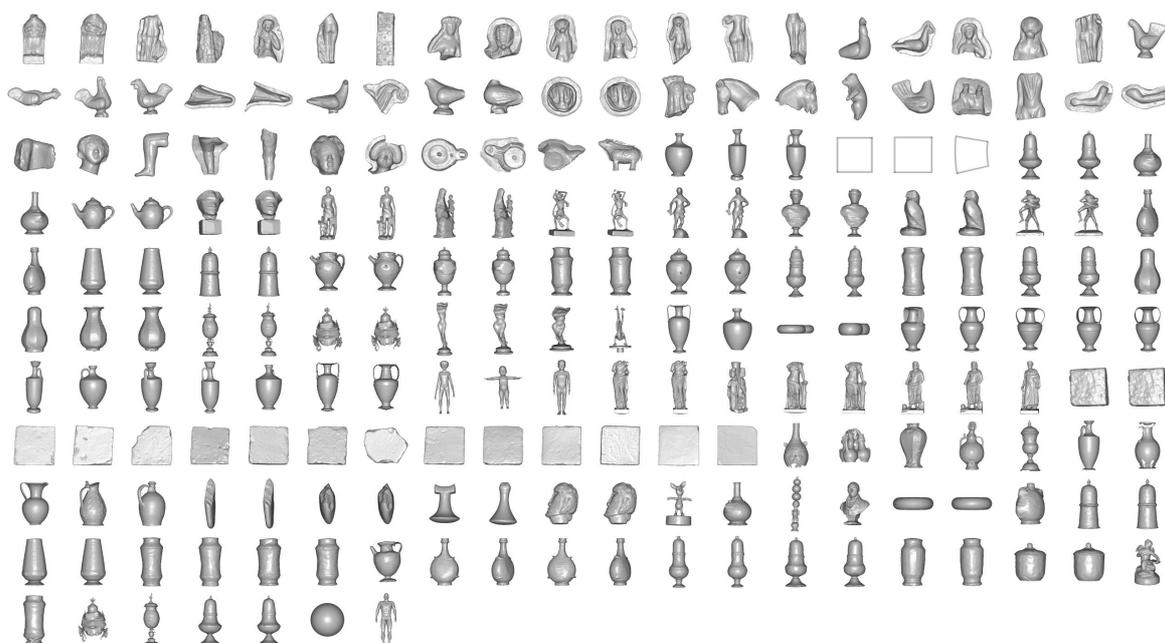


FIG. 5.5 – Base de données de 207 objets pour les tests d'appariement de parties d'objets.

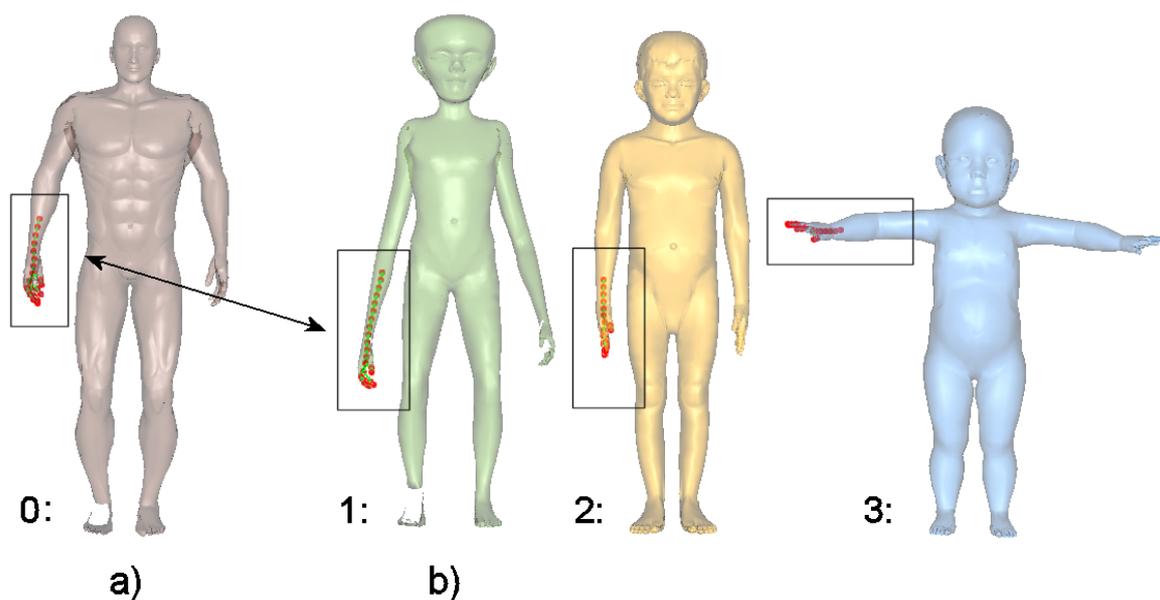


FIG. 5.6 – a) Sous-graphe aMRG à  $r = 5$  de la partie de l'objet requête. b) Sous-graphes aMRG à  $r = 5$  des parties retrouvées.

## 5.3 Recherche 2D/3D

Au même titre que la recherche de forme 3D la recherche d'objets 3D à partir de formes 2D est un domaine en pleine expansion [Funkhouser et al., 2003].

Nous proposons d'appliquer notre approche par aMRG à la recherche de forme 3D à partir d'une requête 2D. Un utilisateur peut effectuer une requête en dessinant une forme à l'aide d'une interface graphique. Le système extrait alors de cette forme 2D le graphe de Reeb multirésolution augmenté correspondant.

Ceci part du constat que pour certains types d'objets, et plus particulièrement les modèles étoilés comme les humains, les mains, etc., les graphes de Reeb sont “presque toujours” plans. Il apparaît donc possible pour ces types d'objets de comparer le graphe de Reeb extrait d'un dessin 2D au graphe extrait d'un modèle 3D.

Afin de pouvoir appliquer directement nos algorithmes, nous proposons de construire un maillage de la requête 2D. Cette application requiert l'implémentation d'une routine pour fermer les contours et pour fabriquer le maillage associé.

La figure 5.7 présente un exemple de silhouette et de maillage 2D. Pour ce test de recherche 2D/3D, nous avons créé la requête 2D à partir du modèles 3D *hunk* aplati en mettant la coordonnée  $z$  à 0. La figure 5.8 présente le graphe aMRG obtenu à la résolution  $r = 5$  de la requête 2D.

Pour pouvoir effectuer une recherche dans une base de données 3D, il est nécessaire de choisir des attributs adaptés aussi bien pour les formes 2D que 3D ce qui n'est pas

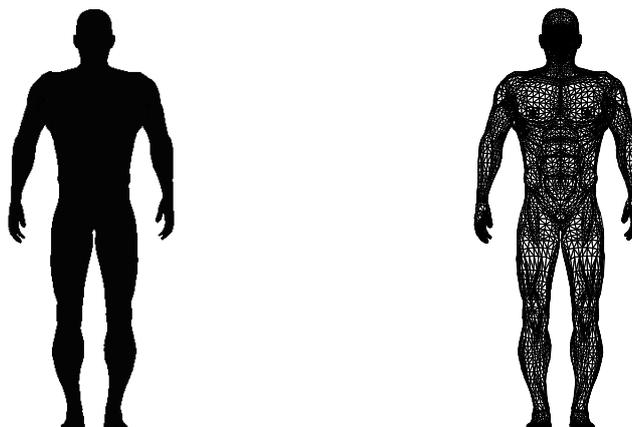


FIG. 5.7 – Exemple de silhouette et son maillage associé.

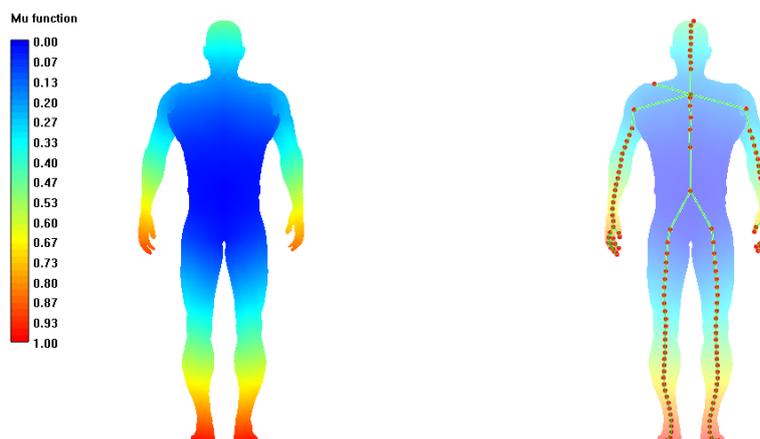


FIG. 5.8 – Fonction  $\mu$  calculée sur la surface de la forme et son graphe aMRG associé ( $r = 5$ ).

évident à moins de calculer des caractéristiques 2D pour chaque objet à partir de vues 2D extraites (cf. Section 1.9).

L'approche des aMRG permet de pallier ce problème car les appariements reposent sur la topologie des graphes et sur l'utilisation des attributs topologiques que nous avons proposés (cf. Section 3.1.6). Ces attributs permettent de caractériser la complexité locale des nœuds des graphes, et nous permettent d'élaguer efficacement la base de données. La figure 5.9 présente des résultats préliminaires de cette application. Nous avons effectué un test sur la base de données de 266 objets (cf. Section 4.1) avec des graphes aMRG à la résolution maximale de  $R = 5$ .

Pour ces tests, les attributs et poids des aMRG ont été choisis empiriquement comme suit :

### 5.3 Recherche 2D/3D

- $f_0 = a$ ,  $\lambda_0 = 0.33$ ,
- $f_1 = \text{cord}L$ ,  $\lambda_1 = 0.33$ ,
- $f_2 = \text{cord}2$ ,  $\lambda = 0.34$ ,

On observe que les résultats sont cohérents. En effet, les premiers objets retrouvés sont des modèles humains, et les objets suivants présentent des formes similaires à celles de la requête. On note qu'avec les paramètres choisis les modèles *alien* et *baby* sont malencontreusement précédés de deux statuettes. Cependant, les attributs et les poids peuvent encore être optimisés.

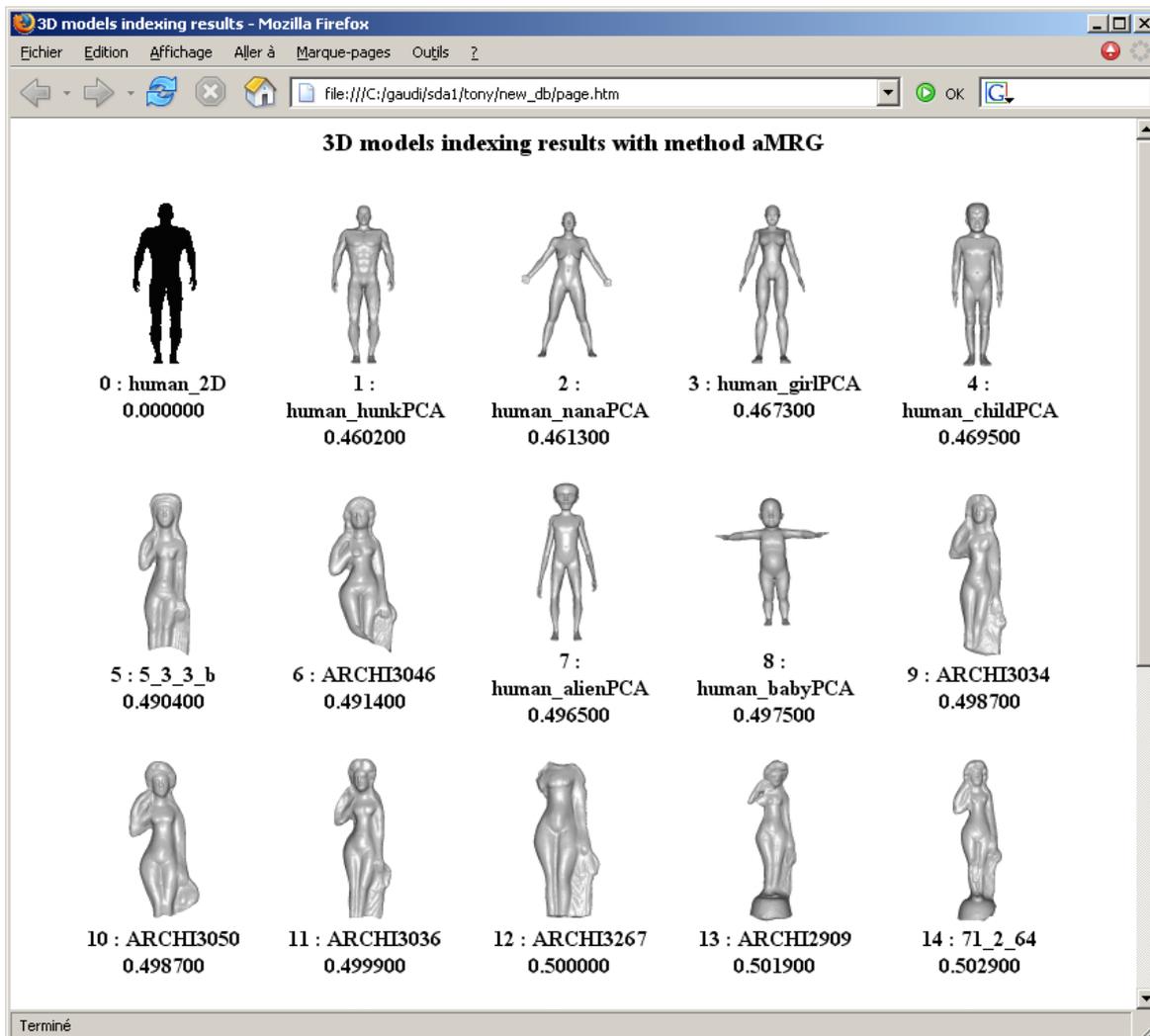


FIG. 5.9 – Résultat de recherche 2D/3D avec les aMRG. Les objets retrouvés sont cohérents avec la requête.

## CHAPITRE 6

# Conclusion et perspectives

## 6.1 Conclusion

Nous avons enrichi l’approche originale de [Hilaga et al., 2001] afin de répondre à notre problématique d’indexation et de recherche par le contenu d’objets 3D texturés. Les critères topologiques ont été étendus, et des attributs géométriques et colorimétriques ont été ajoutés. La procédure d’appariement des nœuds s’avère très puissante et robuste, mais son implémentation a cependant été assez délicate à mettre au point. Les calculs de similarité ont été améliorés en fusionnant différentes caractéristiques géométriques et visuelles, et en proposant diverses pondérations et mesures de distances. On est donc finalement parvenu à un descripteur de forme multirésolution multicritère.

Les résultats expérimentaux ont permis de mettre en évidence l’efficacité des aMRG, non seulement pour discerner les différentes classes d’objets, mais aussi pour regrouper de manière intuitive les classes d’objets qui sont proches entre-elles.

Les tests comparatifs effectués sur notre base de données contenant divers types de modèles 3D à différentes résolutions nous permettent de conclure sur la robustesse de l’approche des aMRG et sur sa supériorité par rapport aux autres méthodes testées. Ainsi la méthode a été implémentée dans le système qu’a produit le projet SCULPTEUR pour les musées européens partenaires du projet.

L’approche des aMRG apparaît riche et prometteuse. Cependant, la fonction  $\mu$  n’étant applicable que sur des maillages à une seule composante connexe, les objets doivent être modélisés correctement. La caractérisation par les informations de couleur et de textures reste à être approfondie. La difficulté ici est de pouvoir disposer d’une base de données d’objets 3D texturés suffisamment importante pour pouvoir effectuer des tests statistiques significatifs.

Un des intérêts de la méthode des aMRG est aussi d’offrir beaucoup de souplesse dans la recherche par le contenu, l’utilisateur pouvant sélectionner et pondérer les attributs qui correspondent le plus à sa recherche. Mais cet avantage se transforme aussi en inconvénient, celui de déterminer une valeur à l’ensemble de ces paramètres. Il serait donc intéressant d’appliquer des techniques d’optimisation de paramètres afin

de proposer à l'utilisateur un ensemble de valeurs par défaut.

## 6.2 Perspectives de recherche

Les résultats expérimentaux que nous avons obtenus ainsi que les premiers résultats des applications nous encouragent à continuer à prospecter dans la même direction. Les graphes de Reeb multirésolution augmentés n'ont pas une structure figée. Il est toujours possible d'améliorer le choix des attributs à inclure dans les nœuds, optimiser les poids, les distances, l'appariement des nœuds, etc.

Les premières améliorations pourraient porter sur le choix des distances et des paramètres de pondération des attributs. Il serait intéressant de définir une méthode d'optimisation des paramètres. On pourrait appliquer des algorithmes génétiques pour minimiser une fonction (par exemple :  $1 - T_1$ ) en travaillant avec une base de données d'entraînement comme nous l'avons proposé. Aussi, nous pensons qu'il peut être intéressant d'étudier les diverses normalisations des attributs pour pouvoir utiliser des distances différentes pour chaque attribut dans le calcul de similarité. Parallèlement à ces travaux, il paraît indispensable de développer un outil de "nettoyage de maillage". Il pourrait être appliqué en prétraitement sur les maillages de mauvaise qualité provenant d'internet. Les opérations nécessaires seraient : fusion des différentes composantes connexes, remaillage de la surface (rééchantillonnage), uniformisation des normales, et régularisation par filtrage pour l'estimation des courbures. Cependant les modèles de piètre qualité ("soupe" de triangles) ne pourront jamais être reconvertis automatiquement en de bons modèles 3D avec de belles propriétés géométriques.

On a pu observer lors de nos travaux que les attributs caractérisant localement la topologie que nous avons introduits sont très discriminants. Il serait intéressant de continuer à chercher un moyen de bien exploiter ces informations. A court terme, il s'agirait de définir une normalisation et une distance adaptée à ces attributs. De plus, il serait aussi intéressant de pouvoir continuer à améliorer les applications que nous avons proposées. Nous pensons par exemple utiliser d'autres descripteurs colorimétriques pour les requêtes basées sur les formes et les couleurs (par exemple : des histogrammes en CIE-Lab) ou utiliser des méthodes de caractérisation de texture. Aussi, les attributs et les poids les mieux adaptés pour les recherches de parties d'objets et les requêtes 2D/3D restent à être trouvés. Une interface graphique adaptée pourrait être développée, permettant la soumission de différents formats de requêtes : objet 3D, image, ou esquisse.

A moyen terme, il serait intéressant de pouvoir intégrer la méthode dans un système de classification non supervisée qui pourrait optimiser automatiquement les paramètres des attributs en fonction de la base de données. Les attributs et les poids devraient pouvoir s'adapter automatiquement au type de requête désiré (recherche de forme globale, de couleur, de parties, ou 2D/3D). On pourrait aboutir à un outil de recherche efficace et multi-usage. Lors des diverses discussions réunissant les personnels des musées par-

tenaires du projet européen SCULPTEUR, nous avons pu prendre connaissance des besoins et problèmes auxquels l'indexation et la reconnaissance de formes 3D pourraient répondre. Par exemple, les historiens sont intéressés par un outil permettant de retrouver précisément les objets de formes similaires et permettant d'effectuer diverses mesures pour pouvoir par exemple étudier l'usure des objets dans le temps et retrouver la chronologie : statue originale, moule, statue moulée, moule de deuxième génération,...

A long terme, il est clair que la reconnaissance des formes 3D pourrait être appliquée à beaucoup d'autres domaines : par exemple les recherches 3D/2D, les recherches d'objets dans les vidéos, la recherche de données sur Internet, etc. En effet, les contraintes de temps de calculs que nous connaissons actuellement n'auront plus raison d'être dans un futur proche tant les performances des ordinateurs croissent rapidement. On pourra donc réfléchir à diverses applications en temps-réel comme faire l'inventaire automatique d'un chariot de supermarché pour calculer le montant des courses : on peut imaginer un système de reconnaissance de forme et de couleur pour reconnaître les produits et rechercher leur prix dans une base de données.



## ANNEXE A

# Bases de données

Les figures [A.1](#), [A.2](#) et [A.3](#) représentent respectivement la base de données de 116 objets répartis en 16 classes, la base de données de 266 objets répartis en 27 classes, et les 301 objets non classés qui ont été ajoutés à la base de données de 266 objets pour former la base de données de 567 objets.

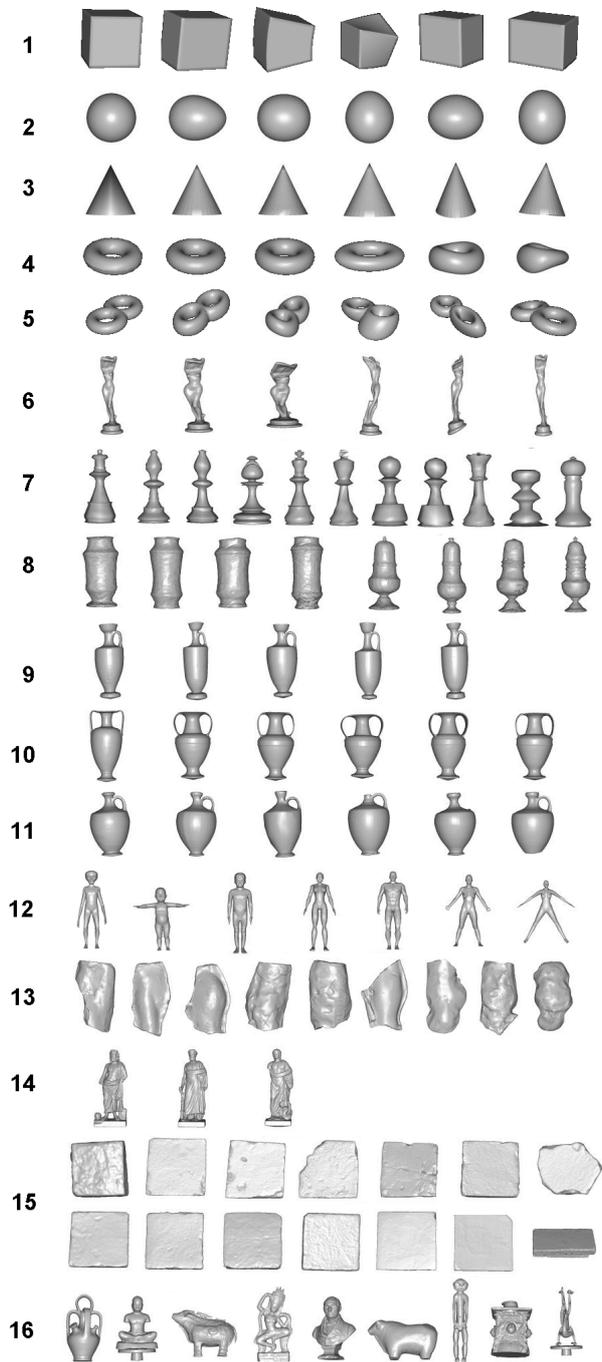


FIG. A.1 – Base de données de test composée de 116 objets répartis en 16 classes.

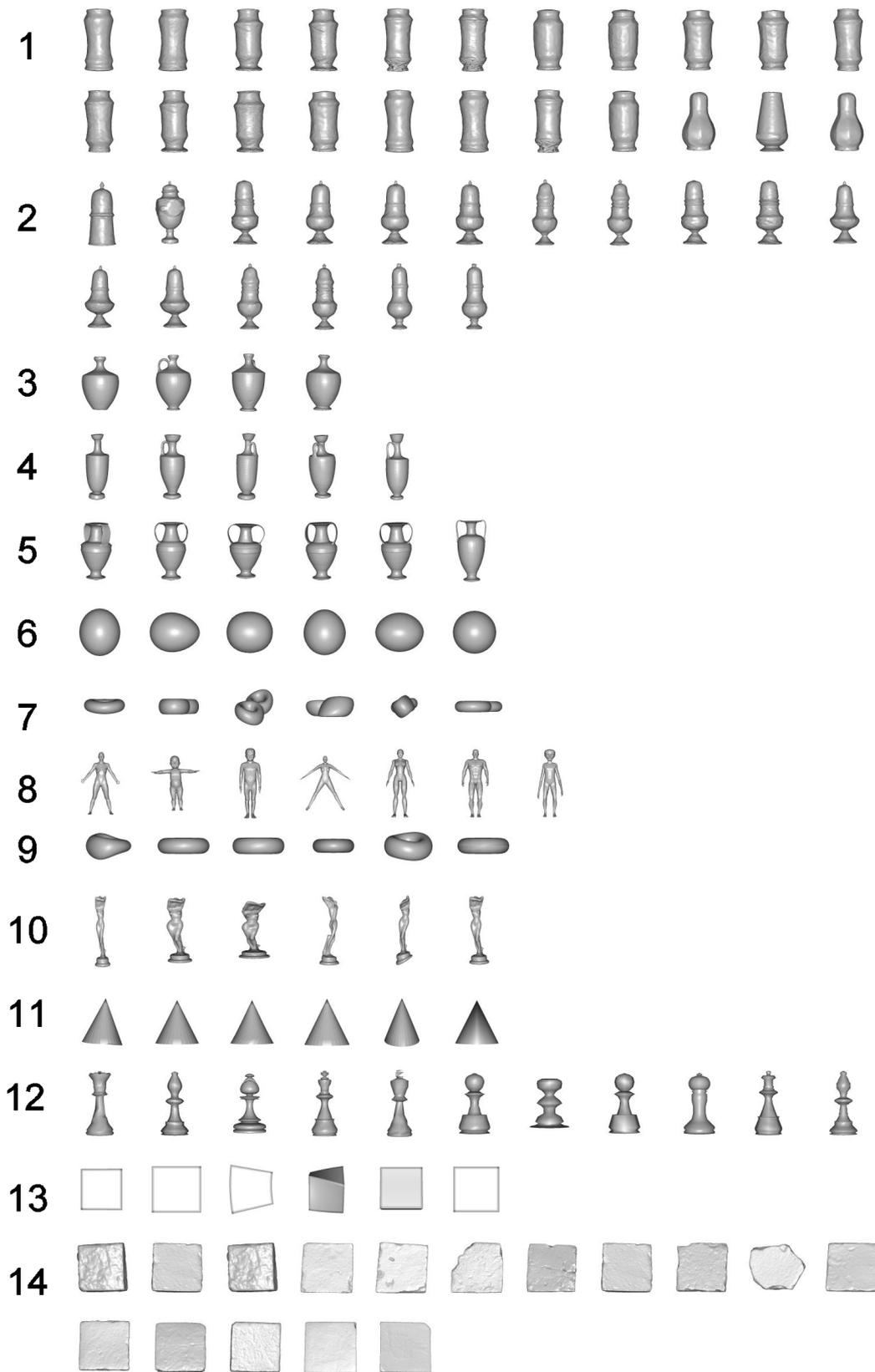
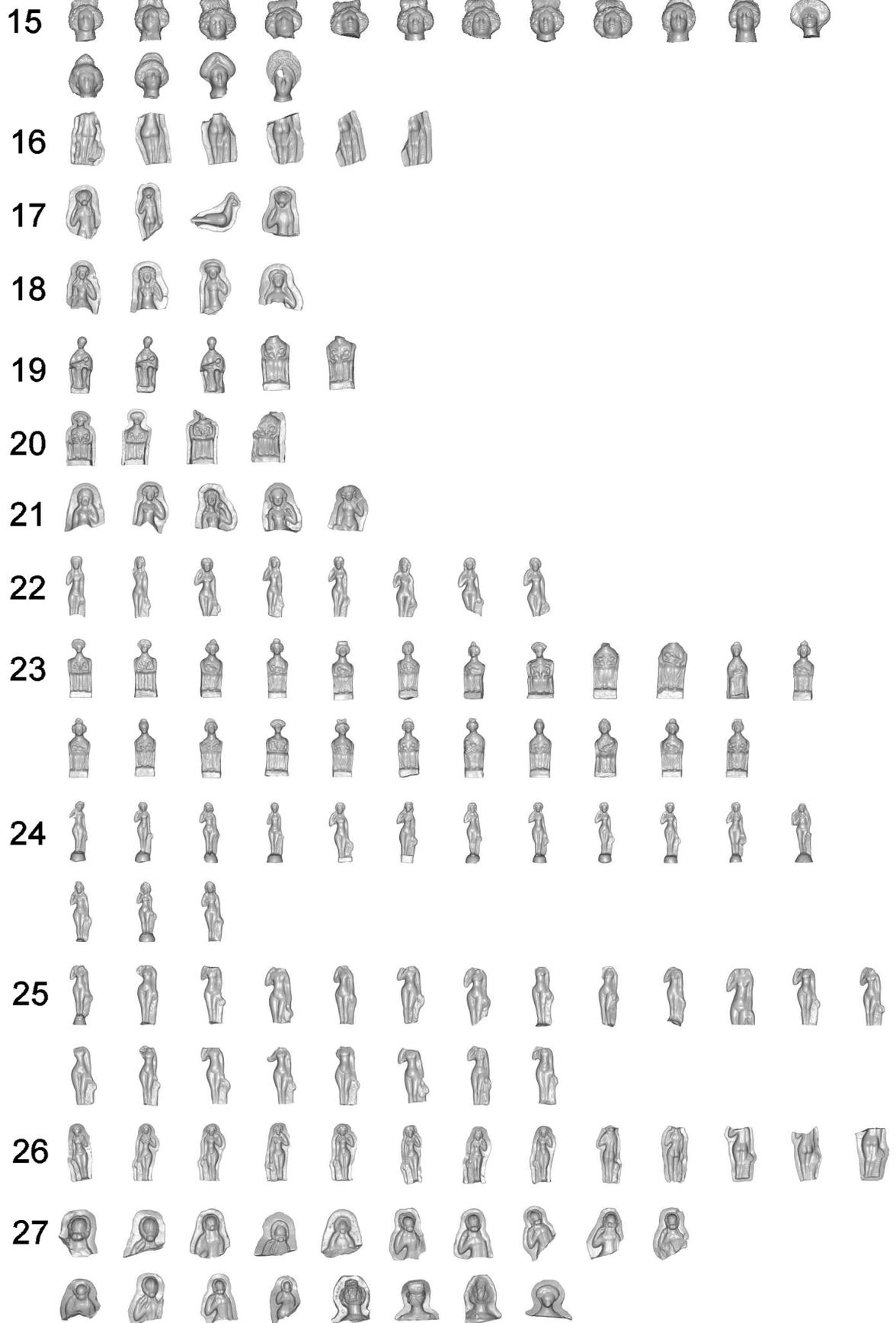


FIG. A.2 – Base de données de 266 objets répartis en 27 classes.



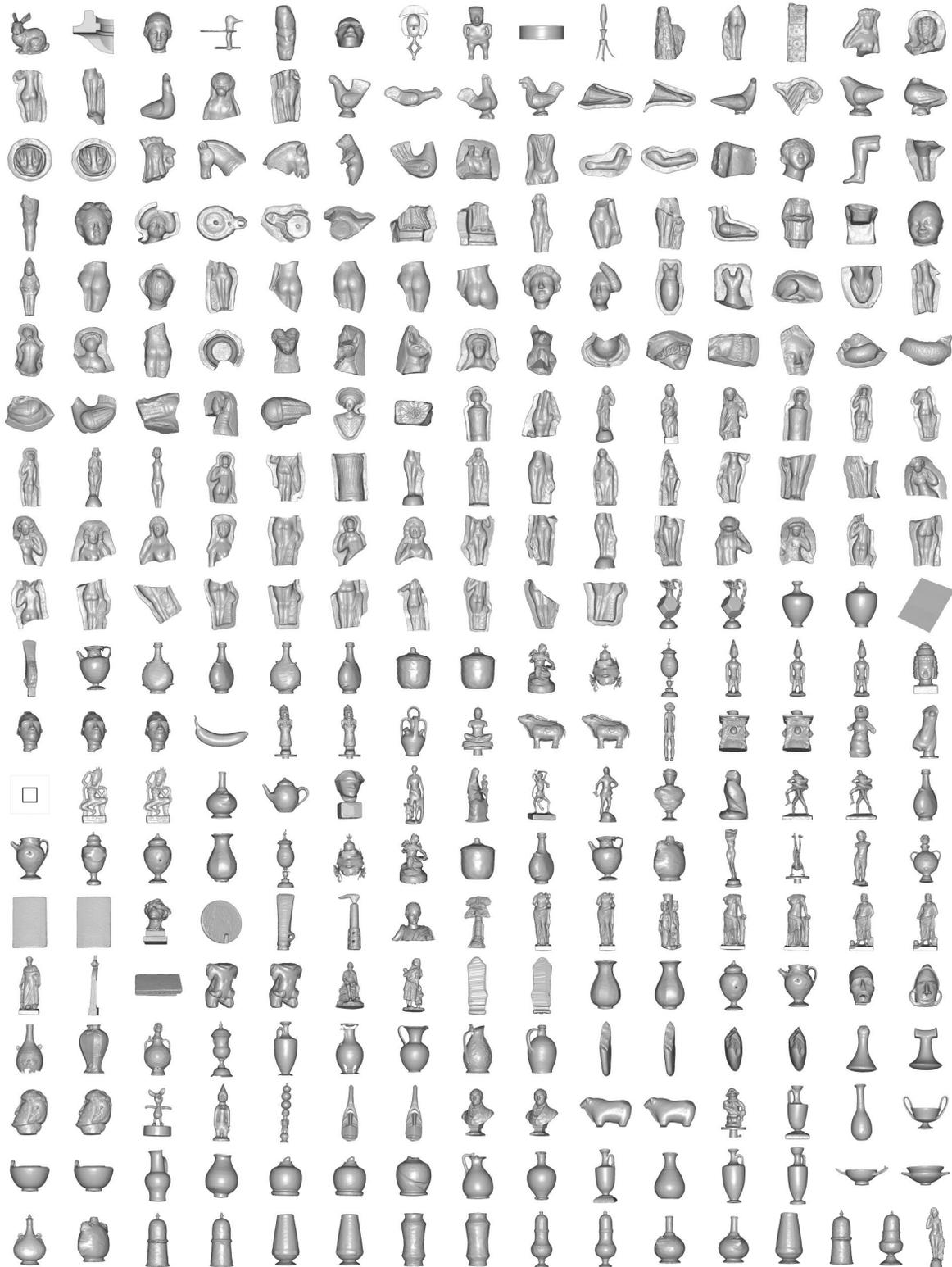


FIG. A.3 – Base de données de 301 objets non classés.



## A N N E X E B

# Interface développée

La figure B.1 représente l'interface développée pour les différents tests. Elle permet de fabriquer les vecteurs caractéristiques correspondant aux méthodes présentées dans le manuscrit (cf. Chapitre 4) :

- histogrammes des longueurs et angles des cordes,
- distribution de forme D2,
- aire, volume et ratio,
- EGI complexes sur la sphère et l'octaèdre subdivisé 1 et 2 fois,
- descripteur de Hough 3D sur la sphère et l'octaèdre subdivisé 1 et 2 fois,
- descripteur de spectre de forme 3D,
- harmoniques sphériques,
- graphe de Reeb multirésolution augmentés.

De plus, elle permet de constituer une base de données d'objets 3D, et effectuer des recherches en utilisant une méthode ou en moyennant deux méthodes. Les distances implémentées sont celles présentées au chapitre 4 :

- distance de Minkowski  $L_1$ ,  $L_2$  et  $L_\infty$ ,
- distance de Bhattacharyya,
- distance de Jeffreys,
- divergence du  $\chi^2$ ,
- intersection d'histogrammes,
- distance EMD.

De plus l'outil réalisé permet, pour les aMRG, de choisir les poids des attributs des nœuds à utiliser pour les calculs de similarité :

- Area = aire relative  $a$ ,
- Mu length = longueur relative de l'intervalle de  $\mu$ ,
- Volume = volume relatif,
- Cords L = histogramme des longueurs des cordes,
- Cords 1 = histogramme des angles des cordes avec le 1<sup>er</sup> axe principal,
- Cords 2 = histogramme des angles des cordes avec le 2<sup>e</sup> axe principal,
- Curv Index = histogramme des indices de forme,
- Hough 3D = descripteur de Hough 3D,
- Ratio(0) = ratio aire/volume au niveau de résolution  $r = 0$ .

Tous les poids sont paramétrables sur  $[0, 1]$ .

Les images de “meilleurs scores”, matrices de distances, et images tier présentés dans le manuscrit ont été aussi générées par l’interface. De même que les calculs des valeurs de précision-rappel et des différents scores  $T_1$ ,  $SBE$ ,  $E$  et  $DCG$ .

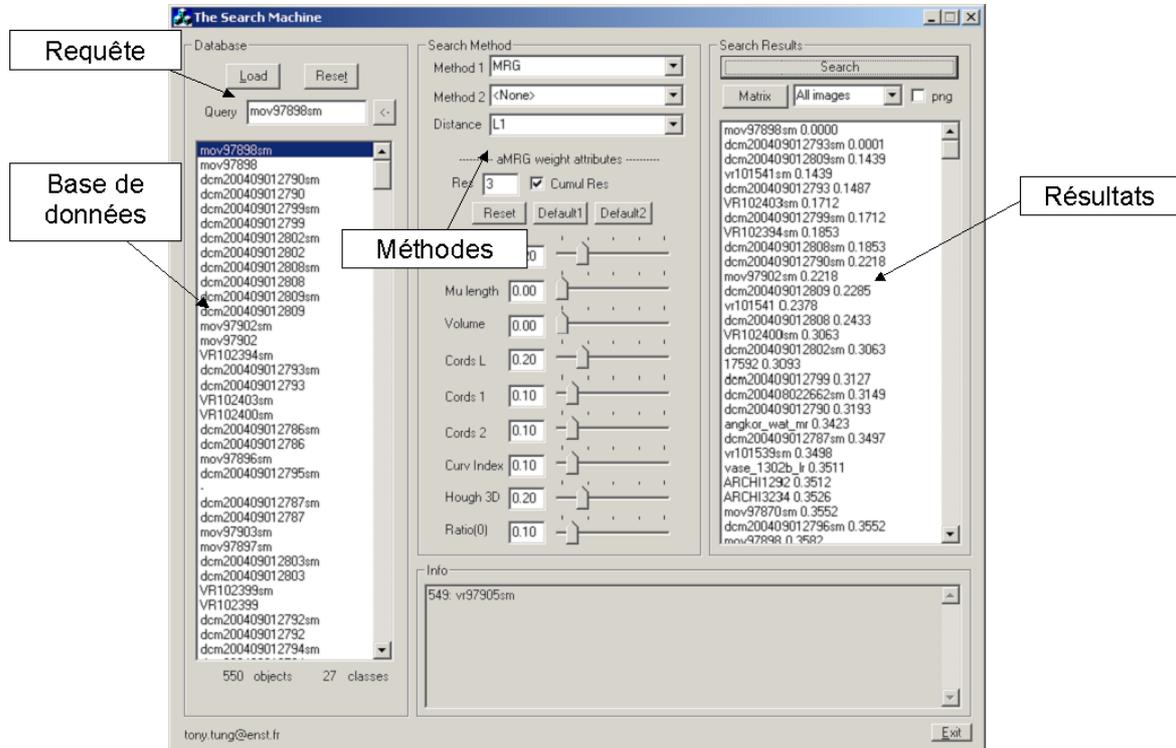


FIG. B.1 – Interface développée pour les tests des méthodes implémentées.

# Bibliographie

- [3dcafe, 2005] 3dcafe (2005). Site internet 3d café : <http://www.3dcafe.com>. 1.1, 2.2.3, 2.4.1, 4.1, 5.2
- [Adán et al., 2003] Adán, M., Adán, A., C.Cerrada, Merchán, P., and Salamanca, S. (2003). Weighted cone-curvature :applications for 3d shapes similarity. In *3DIM, 4th Int. Conf. on 3D Digital Imaging and Modeling*, pages 458–465, Banff, Canada. 1.4.1
- [Addis et al., 2002] Addis, M., Lewis, P., and Martinez, K. (2002). Artiste image retrieval system puts european galleries in the picture (<http://www.artisteweb.org>). In *Cultivate Interactive*. 1.9, 1.14, 1.16
- [Ansary et al., 2004] Ansary, T., Vandeborre, J., and M.Daoudi (2004). Bayesian approach for 3d models retrieval based on characteristic views. In *17th IEEE International Conference On Pattern Recognition*, pages 898–901. 4
- [Assfalg et al., 2003] Assfalg, J., Bimbo, A. D., and Pala, P. (2003). Retrieval of 3d objects using curvature maps and weighted walkthroughs. In *12th International Conference on Image Analysis and Processing (ICIAP)*, pages 348–353. 1.4.1
- [Attali and Boissonnat, 2002] Attali, D. and Boissonnat, J.-D. (2002). Approximation of the medial axis. In *Technical report ECG-TR-12403*, INRIA Sophia-Antipolis. 1.7.1
- [Axen, 1999] Axen, U. (1999). Computing morse functions on triangulated manifolds. In *In Proc. of the SIAM Symposium on Discrete Algorithms (SODA)*. 2.1
- [Barros et al., 1996] Barros, J., French, J., Martin, W., Kelly, P., and Cannon, M. (1996). Using the triangle inequality to reduce the number of comparisons required for similarity-based retrieval. In *Proc. of SPIE, Vol. 2670*, pages 392–403. 1.8.1, 3.5
- [Belongie et al., 2002] Belongie, S., Malik, J., and Puzicha, J. (2002). Shape matching and object recognition using shape contexts. *IEEE Trans. PAMI*, 24(24) :509–522. 1.6.2
- [Biasotti, 2004] Biasotti, S. (2004). Computational topology methods for shape modelling applications. In *PhD thesis, Università di Genova, Italy*. 1.7.2
- [Biasotti et al., 2000] Biasotti, S., Falcidieno, B., and Spagnuolo, M. (2000). Extended reeb graphs for shape analysis and model compression. *Lecture Notes in Computer Science : Proc. of the 9th International Conference on Discrete Geometry for Computer Imagery*, 1953 :185–197. 1.7.3

- [Biasotti et al., 2003] Biasotti, S., Marini, S., Mortara, M., Patanè, G., Spagnuolo, M., and Falcidieno, B. (2003). 3d shape matching through topological structures. *11th Discrete Geometry for Computer Imagery conference, Springer-Verlag, Lecture Notes in Computer Science*, 2886. [1.7.3](#), [1.10](#), [2.2.3](#)
- [Bimbo, 2001] Bimbo, A. D. (2001). Visual information retrieval. In *Morgan Kaufman Publishers, Inc.* [4.2](#)
- [Blum, 1967] Blum, H. (1967). A transformation for extracting new descriptors of shape. In *Proc. Symp. Models for the perception of Speech and Visual Form, MIT Press*, pages 362–380, Cambridge, MA. [1.7.1](#), [1.12](#)
- [Canterakis, 1999] Canterakis, N. (1999). 3d zernike moments and zernike affine invariants for 3d image analysis and recognition. In *In 11th Scandinavian Conf. on Image Analysis*, Greenland. [1.5.3](#)
- [Chen et al., 2003] Chen, D., Ouhyoung, M., Tian, X., and Shen, Y. (2003). On visual similarity based 3d model retrieval. In *Computer Graphics Forum*, pages 223–232. [4](#)
- [Chen and OuhYoung, 2002] Chen, D.-Y. and OuhYoung, M. (2002). A 3d object retrieval system based on multi-resolution reeb graph. In *Proc. of Computer Graphics Workshop*, page 16, Tainan, Taiwan. [1.7.3](#), [1.7.3](#)
- [Chen and Han, 1990] Chen, J. and Han, Y. (1990). Shortest paths on polyhedron. In *Proc. Symp. Theory of Computing*, pages 360–369. [2.2.6](#)
- [Chevalier et al., 2003] Chevalier, L., Jaillet, F., and Baskurt, A. (2003). Segmentation and superquadrics modeling of 3d objects. *Journal of Winter School of Computer Graphics*, 11(2) :232–239. [1.6.3](#)
- [Copson, 1968] Copson, E. (1968). Metric spaces. In *Cambridge University Press.* [1.8.1](#)
- [Dutagaci and Sankur, 2005] Dutagaci, H. and Sankur, B. (2005). Transform-based methods for indexing and retrieval of 3d objects. In *The 5th International Conference on 3-D Digital Imaging and Modeling*, Ottawa, Ontario. [1.5.1](#)
- [Eakins and Graham, 1999] Eakins, J. and Graham, M. (1999). Content based image retrieval. In *Technical Report 39, U.K. JISC Technology Application Programme (<http://www.jtap.ac.uk>)*. [1.9](#)
- [Edelsbrunner et al., 2001] Edelsbrunner, H., Harer, J., and Zomorodian, A. (2001). Hierarchical morse complexes for piecewise linear 2-manifolds. In *Proc. 17th Ann. ACM Symposium on Computational Geometry*, pages 70–79. [2.1](#)
- [Esteban, 2004] Esteban, C. H. (2004). Stereo and silhouette fusion for 3d object modeling from uncalibrated images under circular motion. In *Ph.D. dissertation, Ecole Nationale Supérieure des Télécommunications.* [3.4.4](#)
- [Esteban and Schmitt, 2004] Esteban, C. H. and Schmitt, F. (2004). Silhouette and stereo fusion for 3d object modeling. *Computer Vision and Image Understanding, Special issue on “Model-based and image-based 3D Scene Representation for Interactive Visualization” (<http://www.tsi.enst.fr/3dmodels>)*, 93(3) :367–392. [1.1](#), [4.1](#)

- 
- [Foulds, 1992] Foulds, L. R. (1992). Graph theory applications. In *Springer Verlag, New York*. 3.1
- [F.Schmitt and T.Tung, 2003] F.Schmitt and T.Tung (2003). Sculpteur report on survey on content-based retrieval algorithms (<http://www.sculpteurweb.org>). In *Project report ENST Paris, IST-2001-35372*. 1.9
- [Funkhouser et al., 2003] Funkhouser, T., Min, P., Kazhdan, M., Chen, J., Helderman, A., Dobkin, D., and Jacobs, D. (2003). A search engine for 3d models. *ACM Trans on Graphics*, 22(1) :83–105. 5, 5.3
- [Goldak et al., 2001] Goldak, J., Yu, X., Knight, A., and Dong, L. (2001). Constructing discrete medial axis of 3-d objects. In *International Journal of Computational Geometry and Applications*, 1(3), pages 327–339. 1.10
- [Goodall et al., 2005] Goodall, S., Lewis, P., and Martinez, K. (2005). 3-d shape descriptors and distance metrics for content-based artefact retrieval. In *Proceedings of Storage and Retrieval Methods and Applications for Multimedia, San Jose, California, USA*, pages 87–97. 4.2
- [Goodall et al., 2004] Goodall, S., Lewis, P., Martinez, K., Sinclair, P. A., Giorgini, F., Addis, M. J., Boniface, M. J., Lahanier, C., and Stevenson, J. (2004). Sculpteur : Multimedia retrieval for museums, image and video retrieval (<http://www.sculpteurweb.org>). In *Third International Conference, CIVR, Dublin, Ireland*, pages 638–646. 1.1, 4.1
- [Healy et al., 1998] Healy, D., D.Rockmore, Kostelec, P., and Moore, S. (1998). Ffts for the 2-sphere - improvements and variations. In *TR2002-419, Departement of CS, Darmouth College*. 1.5.2
- [Hetzl et al., 2001] Hetzel, G., Leibe, B., Levi, P., and Schiele, B. (2001). 3d object recognition from range images using local feature histograms. In *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR'01)*, volume 2, pages 394–399. 1.8.5, 1.8.5, 1.8.6
- [Hilaga et al., 2001] Hilaga, M., Shinagawa, Y., Kohmura, T., and Kunii, T. (2001). Topology matching for fully automatic similarity estimation of 3d shapes. In *ACM SIGGRAPH*, pages 203–212, Los Angeles, CA, USA. (document), 1.1, 1.7.3, 1.10, 2.1, 2.2.3, 2.2.4, 2.2.5, 2.2.5, 1, 2.2.7, 2.3.1, 2.3.2, 2.5, 3, 3.1, 3.1.1, 3.1.3, 3.1.4, 3.1.4, 3.1.5, 3.1.7, 3.2, 3.3, 3.6, 3.8, 4.4, 4.4, 6.1
- [Horn, 1984] Horn, B. (1984). Extended gaussian image. In *Proc. of the IEEE, Vol.72*, pages 1671–1686. 1.4.2, 1.10
- [Irani and Ware, 2000] Irani, P. and Ware, C. (2000). Diagrams based on structural object perception. In *Proceedings of the working conference on Advanced visual interfaces. ACM Press.*, pages 61–67. 1.6.3
- [Jeffreys, 1946] Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. *Proc. Roy. Soc. A*, 186 :186–461. 1.8.5
- [Kanai and Suzuki, 2000] Kanai, T. and Suzuki, H. (2000). Approximate shortest path on polyhedral surface based on selective refinement of the discrete graph and
-

- its application. In *IEEE Proc. Geometric Modeling and Processing*, pages 241–250. [2.2.6](#)
- [Kang and Ikeuchi, 1994] Kang, S. B. and Ikeuchi, K. (1994). The complex egi : a new representation for 3d pose determination. In *IEEE Trans on PAMI, Vol.16, No.3*, pages 249–258. [1.4.2](#), [1.10](#), [4.2](#)
- [Kazhdan and Funkhouser, 2002] Kazhdan, K. and Funkhouser, T. (2002). Harmonic 3d shape matching. In *SIGGRAPH Technical Sketches*, page 191. [1.5.2](#), [1.10](#), [1.10](#), [4.2](#)
- [Kazhdan et al., 2004] Kazhdan, M., Funkhouser, T., and Rusinkiewicz, S. (2004). Shape matching and anisotropy. *ACM Transactions on Graphics (SIGGRAPH)*, 23(3) :623–629. [1.2](#)
- [Koenderink, 1990] Koenderink, J. (1990). Solid shape. In *The MIT Press*. [1.4.1](#), [3.4.1](#), [3.4.4](#)
- [Körtgen et al., 2004] Körtgen, M., Park, G. J., Novotni, M., and Klein, R. (2004). 3d shape matching with 3d shape contexts. In *Central European Seminar on Computer Graphics, CESC G*, Slovakia. [1.6.2](#), [1.11](#)
- [Lanthier et al., 1999] Lanthier, M., Maheshwari, A., and Sack, J. (1999). Approximating weighted shortest paths on polyhedral surfaces. In *Proc. Symp. Computational Geometry*, pages 274–283. [2.2.6](#)
- [Lazarus and Verroust, 1999] Lazarus, F. and Verroust, A. (1999). Level set diagrams of polyhedral objects. In *Proc. Symp. Solid Modeling*, pages 130–140. [2.2.3](#)
- [Lee et al., 1994] Lee, T., Kashyap, L., and Chu, C. (1994). Building skeleton models via 3d medial surface/axis thinning algorithms. In *Graphical models and image processing Vol. 56, No. 6*, pages 462–478. [1.7.1](#)
- [Leifman et al., 2003] Leifman, G., Katz, S., Tal, A., and Meir, R. (2003). Signatures of 3d models for retrieval. In *The 4th Israel-Korea Bi-National Conference on Geometric Modeling and Computer Graphics*, pages 159–163. [4.2](#)
- [Leymarie and Kimia, 2001] Leymarie, F. and Kimia, B. (2001). The shock scaffold for representing 3d shape. *Visual Form, Springer-Verlag, Lecture Notes in Computer Science*, 2059 :216–229. [1.7.2](#), [1.10](#)
- [L.Wade, 2000] L.Wade (2000). Automated generation of control skeleton for use in animation. In *PhD thesis*, Ohio State university. [1.7.1](#)
- [Matsumoto, 1997] Matsumoto, Y. (1997). An introduction to morse theory. In *American Mathematical Society*. [2.1](#)
- [Milnor, 1963] Milnor, J. (1963). Morse theory. In *Princeton University Press*. [2.1](#)
- [Mitchell et al., 1987] Mitchell, J., Mount, D., and Papadimitriou, C. (1987). The discrete geodesic problem. *SIAM J. Computing*, 16 :647–667. [2.2.6](#)
- [M.Tanase and R.Veltkamp, 2003] M.Tanase and R.Veltkamp (2003). Polygon decomposition based on the straight line skeleton. In *Lecture Notes in Computer Science, vol. 2616-2003, Springer-Verlag*, pages 247–267, Heidelberg. [1.7.1](#)

- [Novotni and Klein, 2003] Novotni, M. and Klein, R. (2003). 3d zernike descriptors for content based shape retrieval. In *The 8th ACM Symposium on Solid Modeling and Applications*, Seattle, USA. [1.5.3](#)
- [Osada et al., 2002] Osada, R., Funkhouser, T., Chazelle, B., and Dobkin, D. (2002). Shape distributions. In *ACM Trans. on Graphics*, 21(4), pages 807–832. [1.3.3](#), [1.5](#), [1.6.2](#), [1.8.2](#), [1.10](#), [3.4.3](#), [4.2](#)
- [Paquet and Rioux, 1997] Paquet, E. and Rioux, M. (1997). Nefertiti : a query by content software for three-dimensional models databases management. In *Proc. Int. Conf. on Recent Advances, 3-D Digital Imaginig and Modeling*, pages 345–352. [1.3.2](#), [1.10](#), [3.4.3](#), [4.2](#)
- [Paquet and Rioux, 1999] Paquet, E. and Rioux, M. (1999). The mpeg-7 standard and the content-based management of three-dimensional data : A case study. In *ICMCS, Vol. 1*, pages 375–380. [1.5.3](#)
- [Peyre and Cohen, 2003] Peyre, G. and Cohen, L. D. (2003). Geodesic remeshing and parameterization using front propagation. In *In Proc. of 2nd IEEE Workshop on Variational, Geometric and Level Set Methods in Computer Vision (VLSM'03)*, Nice. [2.2.6](#)
- [Reeb, 1946] Reeb, G. (1946). Sur les points singuliers d'une forme de pfaff complètement intégrable ou d'une fonction numérique. In *Comptes Rendus, Académie des Sciences de Paris, Vol. 222*, pages 847–849. [7](#), [1.7.3](#), [2.2](#), [2.3.3](#)
- [Ricard et al., 2004] Ricard, J., Coeurjolly, D., and Baskurt, A. (2004). Art extension for description, indexing and retrieval of 3d objects. In *17th International Conference on Pattern Recognition, ICPR*, Cambridge, UK. [1.5.4](#), [1.5.4](#)
- [Rubner et al., 1998] Rubner, Y., Tomasi, C., and Guibas, L. J. (1998). A metric for distributions with applications to imge databases. In *Proceedins of the 1998 IEEE International Conference on Computer Vision*, Bombay, India. [1.8.4](#), [4.2](#), [4.3.7](#)
- [Saupe and Vranic, 2001] Saupe, D. and Vranic, D. (2001). 3d model retrieval with spherical harmonics and moments. In *Proceedings of the DAGM*, pages 392–397, Munich, Germany. [1.5.2](#), [1.5.3](#)
- [Schroeder et al., 1997] Schroeder, W., Martin, K., and Lorensen, B. (1997). The visualization toolkit, 2nd edition. In *Prentice Hall*. [3.4.4](#), [3.4.4](#)
- [Sharir and Schorr, 1986] Sharir, M. and Schorr, A. (1986). On shortest paths in polyhedral spaces. *SIAM J. Computing*, 15(1) :193–215. [2.2.6](#)
- [Sherbrooke et al., 1996] Sherbrooke, E., Patrikakalis, N., and Brisson, E. (1996). An algorithm for the medial axis transform of 3d polyhedral solids. In *IEEE Transactions on Visualization and Computer Graphics, Vol. 2, No. 1*. [1.7.1](#)
- [Shilane et al., 2004] Shilane, P., Min, P., Kazhdan, M., and Funkhouser, T. (2004). The princeton shape benchmark. In *Shape Modeling International, IEEE Computer Society Press, Genova, Italy*, pages 167–178. [1.1](#), [1.1](#), [2.2.3](#), [4.2](#), [4.2](#), [4.2](#), [4.2](#), [4.5](#)
- [Shinagawa et al., 1991] Shinagawa, Y., Kunii, T., and Kergosien, Y. (1991). Surface coding based on morse theory. *IEEE Computer Graphics and Applications*, 11(5) :66–78. [1.7.3](#), [2](#)

- [Smeulders et al., 2000] Smeulders, A. W. M., Worring, M., Santini, S., Gupta, A., and Jain, R. (2000). Content-based image retrieval at the end of the early years. In *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 22(12) :1349–1380. [1.9](#)
- [Sundar et al., 2003] Sundar, H., Silver, D., Gagvani, N., and Dickinson, S. (2003). Skeleton based shape matching and retrieval. In *Shape Modeling International, SMI*, pages 130–142, Seoul, Korea. [1.10](#), [5.2](#)
- [Swain and Ballard, 1990] Swain, M. J. and Ballard, D. H. (1990). Indexing via color histograms. In *ICCV*, pages 390–393. [1.8.8](#)
- [Takahashi et al., 1995] Takahashi, S., Ikeda, T., Shinagawa, Y., Kunii, T. L., and Ueda, M. (1995). Algorithms for extracting correct critical points and constructing topological graphs from discrete geographical elevation data. *Eurographics*, 14 :181–192. [1.7.3](#)
- [Tangelder and Veltkamp, 2003] Tangelder, J. and Veltkamp, R. (2003). Polyhedral model retrieval using weighted point sets. In *Shape Modeling International, SMI*, Seoul, Korea. [1.1](#), [1.6.1](#), [1.10](#), [4.2](#)
- [Tangelder and Veltkamp, 2004] Tangelder, J. W. and Veltkamp, R. C. (2004). A survey of content based 3d shape retrieval methods. In *Shape Modeling International, IEEE Computer Society Press, Genova, Italy*, pages 145–156. [1.1](#)
- [Thacker et al., 1997] Thacker, N. A., F.Ahearne, and Rockett, P. I. (1997). The bhattacharyya metric as an absolute similarity measure for frequency coded data. *Kybernetika*, 32(4) :363–368. [1.8.3](#)
- [van Rijsbergen, 1975] van Rijsbergen, C. K. (1975). Information retrieval. In *Butterworths*. [4.2](#)
- [Verroust and Lazarus, 2000] Verroust, A. and Lazarus, F. (2000). Extracting skeletal curves from 3d scattered data. *The Visual Computer, Springer*, 16(1) :15–25. [1.7.2](#), [1.10](#)
- [Vranic, 2003] Vranic, D. (2003). An improvement of rotation invariant 3d-shape descriptor based on functions on concentric spheres. In *IEEE International Conference on Image Processing, ICIP*, pages 757–760, Barcelona, Spain. [1.1](#), [1.5.2](#)
- [Vranic and Saupe, 2001] Vranic, D. and Saupe, D. (2001). 3d shape descriptor based on 3d fourier transform. In *Proc. of the EURASIP Conference on Digital Signal Processing for Multimedia Communications and Services, ECMCS*, pages 271–274, Budapest, Hungary. [1.5.1](#)
- [Vranic et al., 2001] Vranic, D., Saupe, D., and J.Richter (2001). Tools for 3d-object retrieval : Kahrune-loeve transform and spherical harmonics. In *IEEE Workshop Multimedia Signal Processing*, pages 293–298, Cannes, France. [1.2](#), [1.10](#), [3.1.7](#)
- [Wood, 2003] Wood, Z. J. (2003). Computational topology algorithms for discrete 2-manifolds. In *PhD thesis, California Institute of Technology, Pasadena, California, USA*. [2.1](#)

- [Xiao et al., 2003] Xiao, Y., Siebert, P., and Werghe, N. (2003). A discrete reeb graph approach for the segmentation of human body scans. In *3DIM, 4th Int. Conf. on 3D Digital Imaging and Modeling*, pages 378–385, Banff, Alberta, Canada. [1.7.3](#)
- [Zaharia and Prêteux, 2002] Zaharia, T. and Prêteux, F. (2002). Indexation de maillages 3d par descripteurs de forme. In *RFIA*, pages 48–57, Angers, France. [1.1](#), [1.2](#), [1.3.4](#), [1.4.1](#), [1.8](#), [1.4.1](#), [1.10](#), [3.3](#), [3.4.4](#), [3.4.4](#), [3.4.5](#), [4.2](#), [4.2](#), [4.3.8](#), [4.5](#)
- [Zaharia and Prêteux, 2004] Zaharia, T. and Prêteux, F. (2004). Descripteurs de forme : Etude comparée des approches 3d et 2d/3d. In *RFIA*, Toulouse, France. [1.3.4](#)
- [Zhang and Chen, 2001] Zhang, C. and Chen, T. (2001). Efficient feature extraction for 2d/3d objects in mesh representation. In *IEEE International Conference on Image Processing, ICIP*, Thessaloniki, Greece. [1.3.1](#), [3.4.2](#), [4.2](#)
- [Zhou and Kambhamettu, 2002] Zhou, L. and Kambhamettu, C. (2002). Representing and recognizing complete set of geons using extended superquadrics. In *ICPR*, pages 713–718. [1.6.3](#)



# Publications

## Journaux

T.Tung and F.Schmitt.

**The augmented multiresolution Reeb graph approach for content-based retrieval of 3D shapes,**

*International Journal of Shape Modeling (IJSM), SMI04 Special issue*, June 2005.

(accepté, à paraître)

T.Tung and F.Schmitt.

**Indexation de modèles 3D par graphe de Reeb multirésolution augmenté,**

*Annales des télécommunications - Annals of telecommunications, Special issue on 3D*, Sep. 2005.

(sélectionné, en révision)

## Conférences & Workshops avec actes et comité de lecture

T.Tung and F.Schmitt.

**Augmented Reeb Graphs for Content-based Retrieval of 3D Mesh Models,**

*International Conference on Shape Modeling and Applications (SMI'04), IEEE Computer Society Press*, pp. 157-166, Genova, Italy, June 2004.

T.Tung et F.Schmitt.

**Indexation d'objets 3D par Graphe de Reeb Multirésolution Augmenté,**

*9èmes journées d'études et d'échanges "COmpression et REprésentation des Signaux Audiovisuels" (CORESA'04)*, pp. 133-136, Lille, France, Mai 2004.

T.Tung et F.Schmitt.

**Indexation d'objets 3D par Graphe de Reeb Multirésolution Augmenté,**

*14ème Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle (RFIA'04)*,

*Atelier Analyse de données, Statistique et Apprentissage pour la Fouille d'Images*, pp. 49-54, Toulouse, France, Jan. 2004.

---

## Communications

SCULPTEUR team.

**New ways to produce, visualise, and navigate multimedia digital collections of 2D & 3D objects,**

*EVA Florence 2005 workshop, Electronic Imaging & the Visual Arts, Florence, Italy, Mar. 2005.*

T.Tung and F.Schmitt.

**3D object indexing : overview and MRG approach,**

*TELEGEO'03, Workshop on Geometry compression, Sophia-Antipolis, France, Nov. 2003.*

## Rapports internes

F.Schmitt, T.Tung, S.Goodall, P.Lewis and K.Martinez.

**SCULPTEUR Report : Documentation on 3D object content-based analysis,**

*Project report ENST Paris, IST-2001-35372, May. 2005.*

T.Tung and F.Schmitt.

**Augmented Reeb Graphs for Content-based Retrieval of 3D Mesh Models,**

*Research report ENST Paris, No 2003D010, Nov. 2003.*

F.Schmitt and T.Tung.

**SCULPTEUR Report on Survey on Content-based Retrieval Algorithms,**

*Project report ENST Paris, IST-2001-35372, Mar. 2003.*