

Descripteurs de forme : Etude comparée des approches 3D et 2D/3D

3D versus 2D/3D Shape Descriptors: A Comparative study

Titus Zaharia

Françoise Prêteux

Groupe des Ecoles des Télécommunications
Institut National des Télécommunications / Unité de Projets ARTEMIS
9, rue Charles Fourier 91011 Evry Cedex
email : titus.zaharia@int-evry.fr

Résumé

Cet article propose une étude comparée des approches 3D et 2D/3D pour l'indexation par descripteur de forme (DF) de maillages 3D, dans le cadre d'applications de requête par similarité.

Cinq DFs 3D et deux DFs 2D/3D ont été retenus : DH3DO - Descripteur de Hough 3D Optimisé, EGIs - Extended Gaussian Images, histogrammes des longueurs des cordes et des angles sphériques, histogramme des triangles aléatoires, descripteurs EEC - Espace Echelle de Contour et ART - Angular Radial Transform. Une version de basse complexité du DH3DO dérivée par quantification vectorielle est également développée et sélectionnée pour cette comparaison.

Les résultats expérimentaux, conduits sur la base de test catégorisée MPEG-7, établissent objectivement, par estimation du score Bull-Eye (SBE) et du First Tier, la supériorité du DH3DO et de sa version quantifiée (SBE respectivement de 81% et 79%) et un très bon comportement du DF EEC (SBE de 74%), qui surclasse les autres descripteurs qu'ils soient 3D ou 2D/3D.

Enfin, appliqué dans le cadre du projet RNRT SEMANTIC-3D, le DH3DO démontre sa pertinence en termes de discrimination, de scalabilité et de robustesse face à des versions fortement compressées de maillages.

Mots Clef

Indexation, maillages 3D, descripteurs de forme 2D et 3D, approches multi-vues, mesures de similarité, MPEG-7.

Abstract

This paper proposes a comparative study of 3D and 2D/3D shape descriptors (SDs) for indexing and retrieval of 3D mesh models.

Seven state of the art SDs are considered and compared, among which five are 3D (Optimized Hough 3D Descriptor-OH3DD, Extended Gaussian Images - EGIs, cords length and spherical angles histograms, random triangles histogram), and two 2D/3D, based on the MPEG-7 2D SDs (Contour Scale Space- CSS, and Angular Radial Transform - ART). A low complexity vector quantized (VQ) OH3DD is also proposed and considered for this comparison.

Experimental results were carried out upon the categorized MPEG-7 3D test database. By computing Bull-Eye Score (BES) and First Tier criteria, it is objectively established that OH3DD (even in its VQ version) outperforms (BES = 81% or 79%) all other SDs. The 2D/3D CSS-based descriptor exhibits a highly

discriminant behavior (BES = 74%) outperforming the other both 3D and 2D/3D approaches

Apply to the industrial framework of the RNRT SEMANTIC-3D Project, OH3DD demonstrated its relevance together with its scalability and robustness properties.

Keywords

Indexing and retrieval, 2D and 3D shape descriptors, multiview matching, similarity measures, MPEG-7 standard, 3D meshes.

1 Introduction

Aujourd'hui, les représentations 3D envahissent les applications multimédia destinées aussi bien au grand public, notamment à travers le monde des jeux, de la réalité virtuelle au service de la maison et de la domotique, ou encore le cinéma de synthèse, qu'aux professionnels avec la CAO et ses multiples déclinaisons spécifiques.

Le plus souvent, ces données 3D sont représentées sous forme de maillages 3D qui, contrairement à d'autres types de représentation (surfaces implicites, NURBS, données volumiques), permettent une visualisation simple et efficace, effectuée en temps réel par la plupart des cartes graphiques actuellement disponibles.

Le volume croissant de maillages 3D aujourd'hui disponibles sur Internet ou dans des bases de données spécialisées lance de nouveaux défis en termes d'indexation par le contenu et d'accès intelligent à ces contenus riches et protéiformes.

Cet article traite de l'indexation par descripteur de forme (DF) de maillages 3D, dans le cadre d'applications de requête par similarité, sous contrainte d'invariance géométrique aux similarités et de stabilité topologique aux représentations maillées multiples et/ou plus ou moins dégénérées [1].

L'objectif est d'analyser et de comparer les deux familles d'approches 3D et 2D/3D de la littérature pour en établir avantages, limitations et performances objectives en termes de résultats de requêtes par similarité sur un même corpus 3D avec vérité terrain. A notre connaissance, cette étude est la première de ce type proposée dans la littérature.

Nous présentons tout d'abord (Paragraphe 2) un état de l'art des approches 3D dans le contexte de la reconnaissance et de l'indexation d'objets 3D, puis enchaînons en détaillant le principe des approches 2D/3D. Cette synthèse nous conduit à retenir pour notre étude comparative un ensemble de cinq DF 3D ainsi que les deux DF 2D actuellement retenus dans le standard MPEG-7. Nous rappelons succinctement principe et

définition de ces différents DF (Paragraphe 3). Il ressort alors que le descripteur de Hough 3D optimisé (DH3DO) [1], [2], [3], offre la représentation la plus complète, au prix d'une capacité de stockage et d'une complexité de calcul de la mesure de similarité associée plus élevées que celles des autres DFs. Notre contribution est ici de montrer comment ce DF peut être rendu équivalent en terme de taille et de complexité aux autres DFs, via une quantification vectorielle.

Les huit DF retenus sont finalement évalués et comparés objectivement (Paragraphe 4). Leurs performances en terme de requêtes sont objectivement évaluées via le score *Bull-Eye* et le *First Tier*, sur la base catégorisée MPEG-7.

Enfin, dans la dernière partie, nous démontrons la pertinence du DH3DO pour des requêtes par similarité de forme dans le cadre d'une application de télédistribution de pièces automobiles, sur le corpus d'objets 3D du projet RNRT SEMANTIC 3D (Service d'Echange et de MANipulation (Tatouage, Indexation, Compression)) [4], et mettons notamment en évidence le comportement scalable et robuste du DH3DO pour des maillages dégradés par compression MPEG-4.

2 Etat de l'art

Commençons notre analyse de la littérature par les approches 3D.

2.1 Approches 3D

Les méthodes 3D décrites dans la littérature peuvent être répertoriées en approche statistique, structurale, par transformée et variationnelle. Etudions de façon synthétique les méthodes représentatives de chaque catégorie, et soulignons-en avantages et inconvénients.

2.1.2 Les approches statistiques

Dans le cadre spécifique de l'indexation, les DFs statistiques consistent en général soit à calculer divers moments statistiques, [5], [6], [7], soit à estimer la distribution de la mesure d'une primitive géométrique donnée (points, cordes, sécantes, triangles, tétraèdres), quelle soit déterministe [7], [8], ou aléatoire [6], [9].

Utiliser directement les moments statistiques pour la reconnaissance nécessite une normalisation en taille et position de l'objet, afin d'obtenir une certaine invariance géométrique *extrinsèque* de la représentation. Pour cette raison, nous recommandons d'utiliser des DFs satisfaisant *intrinsèquement* ces propriétés d'invariance, comme les invariants algébriques [10], qui fournissent des DFs globaux, s'exprimant en fonction des moments de différents ordres. Les mesures de similarité liées aux DFs définis comme des vecteurs d'invariants algébriques s'appuient en général sur la distance de Mahalanobis [11], pour prendre en compte les contributions spécifiques de chaque élément. La principale difficulté est ici de définir des invariants stables par rapport aux données.

Parmi les approches par distributions statistiques, mentionnons tout d'abord le spectre de forme 3D (SF3D) adopté dans MPEG-7¹ [12]. Intrinsèquement invariant aux transformations géométriques, le SF3D est en revanche très sensible aux représentations topologiques multiples [1], ce qui nécessite la mise en œuvre de procédures de filtrages lourdes et coûteuses en temps de calcul.

Les différentes distributions de cordes [8], ou encore les distributions de longueurs, angles, aires ou volumes, associées à des ensembles de sécantes, triangles et tétraèdres, construits par un échantillonnage aléatoire du maillage considéré [9] contournent ces difficultés en faisant intervenir des représentations aléatoires, mais restent tout de même sensibles aux différentes triangulations qu'on peut associer à un même objet 3D.

Représentatives des DFs par approche surfacique, les EGI (*Extended Gaussian Images*) [13], [14], [15], définissent sur la sphère unité une fonction synthétisant l'information d'orientation, d'aire, de distance, et de courbure en tout point de la surface de l'objet 3D. L'inconvénient majeur de ces représentations vient de leur grande dépendance à l'information d'orientation. Ainsi, deux objets offrant le même aspect global, mais avec des facettes individuelles orientées différemment (comme, par exemple une pyramide en escalier et une pyramide lisse), conduiront à des EGI complètement différentes.

Toutes ces descriptions offrent l'avantage de la compacité et d'un faible coût de calcul. Toutefois, ces DF ne vérifient pas les contraintes d'invariance géométrique, à l'exception d'une des distributions proposées dans [9] (cf. Paragraphe 3.1.2) et du SF3D. Dans tous les autres cas, pour tenter d'y remédier de manière plus ou moins *ad-hoc*, les auteurs proposent d'introduire des mécanismes d'alignement spatial, fondés en général sur des mesures globales (comme par boîte englobante, analyse en composantes principales-ACP...), mais présentant en revanche un comportement assez instable, pouvant conduire à des faux alignements [1], [2].

2.1.2. Les approches structurales

Contrairement aux approches statistiques, les représentations structurales visent à décrire la notion de forme de manière plus complète et plus intuitive. Un premier type d'approche s'appuie sur une segmentation initiale de l'objet en sous-parties satisfaisant certains critères d'homogénéité par rapport à un attribut de forme préétabli, et représentée par des structures spécifiques comme des arbres ou des graphes. Si l'étape de segmentation vise à identifier les différentes structures élémentaires qui composent l'objet considéré, la deuxième phase permet de représenter les relations d'adjacence et éventuellement les positions relatives de ces différentes structures.

Dans une description surfacique d'objets obtenus à partir d'images de profondeur, Dorai et Jain [16] proposent un DF par graphe de *patches* maximaux au sens d'un index de forme fonction des courbures principales [17]. La connexité de ces *patches* maximaux est ensuite représentée par un graphe d'adjacence. Une technique d'appariement de graphes [18] permet alors de gérer la mise en correspondance des surfaces et d'en déduire un critère de similarité morphologique des objets.

Des approches représentatives de la segmentation d'objet 3D en *géons* et des représentations par graphe d'adjacence sont présentées dans [19], [20]. Les 10 *géons* utilisés ici sont des structures volumiques simples, correspondant à des formes géométriques élémentaires, comme des cylindres (déformés ou non), des cubes, des parallélépipèdes, des cônes et des ellipsoïdes (tronqués ou non). L'inconvénient majeur de cette approche réside dans le volume important des interactions requises avec l'utilisateur aussi bien pour la spécification des requêtes que pour l'annotation de la base.

Des approches similaires, mais beaucoup moins élaborées, sont décrites dans [21], où les *géons* sont des cubes, ellipsoïdes, ou cylindres ayant subi différentes déformations et modélisés

¹ Descripteur promu dans MPEG-7 par l'Institut National des Télécommunications, France.

soit par des surfaces superquadriques [22], soit par de simples primitives quadriques [23]. Cette dernière approche a été récemment considérée dans le cadre de la standardisation MPEG-7 et comparée en termes de performances de requêtes par similarité avec le SF3D MPEG-7. Les premiers résultats obtenus [24] montrent toutefois des performances nettement inférieures à celles du SF3D.

Récemment, Tangelder *et al.* [25], ont développé une approche fondée sur des représentations par points d'intérêt, définis en fonction de la courbure gaussienne. La mesure de similarité est fondée sur une distance de type *Earth Mover Distance* (EMD) qui permet d'effectuer l'appariement entre différentes représentations et de garantir l'invariance géométrique du descripteur proposé. Toutefois, les aspects d'invariance topologique, bien que cruciaux lorsqu'il s'agit d'estimer des caractéristiques différentielles de surface à partir de maillages 3D, ne sont pas abordés.

En général, les approches structurales par graphes de composantes conduisent à des représentations de haut niveau, complètes et élaborées, et en cohérence avec les conclusions des différentes théories de la perception. En outre, ces représentations *semi-globales*, sont utilisables avec des appariements aussi bien globaux que locaux. Toutefois, les algorithmes associés d'extraction/segmentation et d'appariement restent difficilement exploitables dans un contexte d'indexation, en raison de leur complexité élevée. Enfin, remarquons que l'analyse de connexité, indispensable lorsqu'il s'agit de déterminer les relations d'adjacence entre les différentes sous-composantes, requiert de disposer de données bien définies topologiquement. Notons pour conclure que ces bonnes propriétés topologiques ne sont en général pas vérifiées par des maillages quelconques.

2.1.3 Les approches par transformée

Caractéristiques du domaine de la reconnaissance, les représentations d'objets 3D fondées sur des transformées visent à déterminer des représentations de forme globales, définies en terme de transformation intégrale.

Dans ce contexte, mentionnons tout d'abord les approches réalisant une mise en correspondance d'un domaine 2D avec une surface 3D de topologie sphérique, comme celles par images d'attributs sphériques (IAS) [26], par applications harmoniques [27], ou encore par images de spin [28]. La surface 3D est alors décrite par une image de différents attributs, définie sur le domaine 2D considéré.

En pratique, les maillages ne satisfont jamais les contraintes de régularité géométrique et topologique requises par ces approches. La solution consiste alors en un remaillage préalable des maillages, mais ces procédures sont en général complexes en temps de calcul, impliquant des algorithmes d'optimisation laborieux. En outre, pour garantir la précision de la représentation, la discrétisation du domaine 2D associé doit être suffisamment dense, ce qui augmente considérablement la taille de ces représentations. L'appariement implique des algorithmes d'optimisation des fonctionnelles d'énergie ou des mesures de corrélation, complexes en temps de calcul.

Une littérature très riche souligne tout l'intérêt porté aux approches à base de transformée de Hough [29], [30], [31]. Ces différentes méthodes relèvent d'une approche purement surfacique, consistant à détecter différentes variétés de dimension $(n-1)$ plongées dans l'espace $\mathbb{R}^n$. Cependant, un autre point de vue peut être associé à la transformée de Hough, qui permet d'élargir son contexte de mise en œuvre notamment au cadre de l'indexation et des requêtes par similarité. C'est bien le

cas du descripteur de Hough 3D optimisé (DH3DO) proposé dans [3]. Intrinsèquement stable topologiquement, il est rendu invariant aux transformations de similarité via une procédure d'alignement qui conduit à une mesure de similarité naturellement symétrique.

Dans [32], les auteurs proposent une représentation par densité de nuages de points mesurée dans un ensemble de cellules volumiques entourant l'objet 3D. L'invariance aux rotations à 90° est ici assurée via une partition des cellules de l'espace en classes d'équivalence, le descripteur étant calculé par rapport à celles-ci. Toutefois, comme les auteurs le soulignent, cette fusion des cellules de l'espace en classes diminue significativement le pouvoir discriminant du descripteur. En outre, cette représentation, essentiellement volumique, est mal adaptée aux données surfaciques comme les maillages 3D, générant une complexité très élevée du descripteur proposé.

De cet ensemble méthodologique plus riche et plus élaboré, il ressort que ces DFs globaux fournissent des représentations de forme bien adaptées et quelquefois complètes (au sens où la forme initiale peut être reconstruite à partir de sa représentation). Toutefois, leur utilisation efficace dans le contexte de l'indexation nécessite : 1. d'assurer une certaine invariance aux déformations locales ; 2. d'introduire des procédures de normalisation/ alignement pour garantir leur invariance géométrique ; 3. de minimiser la taille des descripteurs et le coût de calcul de la mesure de similarité associée.

2.1.4 Les approches variationnelles

Les approches variationnelles s'appuient sur une modélisation physique des déformations que peut subir une surface 3D donnée. Elles recherchent la solution de l'équation d'équilibre dynamique entre tensions internes et champ de forces externes. La mesure de similarité entre les différentes surfaces est définie comme l'énergie globale de déformation nécessaire afin de mettre en correspondance les surfaces à comparer. Comme représentatifs de cette catégorie d'approche, mentionnons les travaux de [33], [34], où l'analyse des différents modes de déformation permet de définir une mesure de similarité, appelée *appariement modal*.

Ces approches sont particulièrement bien adaptées lorsqu'il s'agit de comparer des modèles appartenant à une même catégorie de formes, comme dans le cas des applications de reconnaissance de visage [35], pour lesquelles les différences entre les divers éléments sont efficacement modélisées par des déformations non-rigides relativement petites. En revanche, ces méthodes ne sont plus pertinentes dès lors qu'il s'agit de comparer des objets aussi différents qu'un chat et une bicyclette comme le dit avec humour Jitendra Malik lors de ses conférences [36].

Les difficultés spécifiques liées à la nature intrinsèquement surfacique des maillages 3D justifient la catégorie d'approche dite 2D/3D, dont le principe consiste à associer aux objets 3D un ensemble de projections 2D, correspondant à différents angles de vue.

2.2 Approches 2D/3D

La forme 3D est alors indirectement représentée par divers DF 2D associés à ces images de projection.

Notons en outre que ce principe permet également de réaliser des appariements entre modèles 3D et objets 2D extraits des bases d'images ou de vidéos.

Ce principe est actuellement pris en compte dans le standard MPEG-7, dans le cadre du schéma de description *MultiView DS*, qui définit un "conteneur" permettant d'intégrer l'ensemble des descripteurs 2D, qu'ils soient de forme, de couleur ou encore de texture.

En ce qui concerne la mesure de similarité, elle est dérivée à partir des mesures de similarité associées aux descripteurs 2D. Dans cet article, nous avons adopté l'approche suivante, recommandée par MPEG-7. Soit A et B deux objets 3D à comparer, chacun représenté par un nombre N de vues. On calcule tout d'abord une matrice d'erreur $E=(e_{ij})_{ij}$, où e_{ij} représente la mesure de similarité entre les descripteurs 2D associés respectivement à la vue i de l'objet A et à la vue j de l'objet B. La mesure de similarité $D_{3D}(A, B)$ entre les deux objets 3D est alors définie par l'équation (1) :

$$D_{3D}(A, B) = \min_{\pi \in \Pi_E} \{ \text{Trace} [\pi(E)] \} , \quad (1)$$

où π représente une permutation arbitraire de l'ensemble Π_E des permutations des colonnes de la matrice E.

Notons que cette approche assure la symétrie de la mesure de similarité $D_{3D}(A, B)$.

Pour comparer des objets 3D, ces approches 2D/3D présentent *a priori* les inconvénients suivants :

- perte d'information due à la projection,
- dépendance des résultats en fonction du nombre de vues,
- absence de méthode robuste pour sélectionner automatiquement le nombre de vues,
- complexité en terme de stockage proportionnelle au nombre de projections,
- temps de réponse aux requêtes quadratique avec le nombre de projections.

L'état de l'art actuel est bien représenté par les deux descripteurs de forme 2D MPEG-7 (DF-ART et DF-EEC) (cf. Paragraphe 3.2). Comparativement à de nombreux autres DFs 2D évalués dans le cadre des expérimentations techniques de MPEG-7, ils ont exhibé une meilleure capacité de discrimination des formes 2D et ont en outre été optimisés durant les procédures d'évaluation MPEG-7.

2.3 Conclusion

Le riche éventail de méthodes de représentation témoigne d'emblée du grand intérêt suscité par la problématique de la reconnaissance et de l'indexation d'objets 3D. Notre objectif maintenant est de sélectionner un ensemble d'approches, représentatives de chaque famille susmentionnée, pour réaliser une comparaison expérimentale des performances associées.

En ce qui concerne les représentations statistiques, nous avons préféré les approches par histogrammes à celles par moments statistiques et invariants algébriques, en raison de leur plus grande stabilité. Dans le cadre des DFs par transformées, nous avons opté pour la généralité. Puisque les représentations par IAS, par applications harmoniques ou encore par images de spin ne sont pas directement applicables à des objets de topologies arbitraires, nous avons porté notre choix sur le DH3DO, générique et déjà utilisé avec succès dans le cadre de requêtes par similarité.

Nous écartons de notre étude comparative d'une part les approches structurales, dont le calcul des mesures de similarité est bien trop complexe pour permettre des requêtes efficaces dans des grandes bases de données, et d'autre part les approches variationnelles, applicables uniquement dans le cas de bases

d'objets spécifiques appartenant à une même catégorie de formes.

Enfin, en ce qui concerne les approches 2D/3D, nous considérerons les deux descripteurs MPEG-7.

Présentons à présent les DFs 3D et 2D/3D retenus.

3 Les approches retenues

Six approches 3D ont été retenues. .

3.1 Approches 3D

Présentons succinctement leur principe et leur définition.

3.1.1 Histogrammes de cordes

Rappelons le principe de l'histogramme des longueurs des cordes (HLC), et de l'histogramme des angles sphériques des cordes (HASC), [8].

Dans le cas des descripteurs HLC et HASC, l'ensemble des cordes reliant le centre de gravité de l'objet aux centres de gravité de l'ensemble des facettes est tout d'abord construit. Le HLC représente la distribution de la longueur des cordes, tandis que le HASC donne la distribution 2D de leurs angles sphériques (azimut et élévation). Pour assurer un comportement invariant géométriquement, l'objet est tout d'abord normalisé par rapport à sa boîte englobante, déterminée après analyse en composantes principales (ACP) et alignement de l'objet par rapport aux axes d'inertie, comme décrit dans [8].

3.1.2 Histogrammes de triangles aléatoires (HTA)

Dans le cas du descripteur HTA un nombre prédéfini de triangles aléatoires est tout d'abord généré. Chaque triangle est construit en sélectionnant au hasard trois sommets du maillage, notés P_1 , P_2 et P_3 . Le descripteur HTA est la distribution de l'angle entre les segments P_1P_2 et P_2P_3 . Notons qu'il est intrinsèquement invariant géométriquement et qu'il ne nécessite donc pas l'inclusion de mécanismes d'alignement et de normalisation par ACP ou boîte englobante.

3.1.3 DF 3D par EGI

Ici, les EGI sont construites comme des cas particuliers du DH3DO (cf. § 3.1.4), ce qui permet d'améliorer leurs performances par rapport aux formulations et implantations algorithmiques traditionnelles [1].

3.1.4 Descripteur de Hough3D (DH3DO)

Rappelons tout d'abord la définition du DH3DO, dont le lecteur trouvera une description détaillée dans [1], [2].

Pour un repère $R = (Oxyz)$ donné, notons par C le groupe des changements de repère obtenus par permutations et inversions des sens des axes de R .

On appelle ensemble d'orientations O , un ensemble de vecteurs unitaires $\{u_{jk}\}$ de $\mathbb{R}^3$ respectivement spécifiés par leurs coordonnées sphériques (θ_k, φ_j) .

O est dit *C-invariant* si et seulement si $C(O) = O$, ce qui est noté et exprimé par :

$$\forall c \in$$

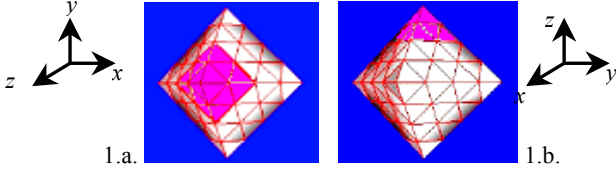


Figure 1. Partition C-invariante.

Le maillage de la Figure 1 est obtenu en appliquant deux étapes successives de subdivision à un octaèdre régulier, centré à l'origine du repère. Les sommets normalisés de l'octaèdre subdivisé constituent un ensemble de vecteurs unitaires, qui définissent l'ensemble C-invariant d'orientations. La propriété d'invariance est illustrée pour les domaines en gris foncé. Pour le changement de repère considéré, le domaine de la Figure 1.a se recalc exactement sur le domaine de la Figure 1.b.

Notons que d'autres constructions d'ensembles invariants d'orientations sont également possibles. Dans cet article, nous nous limiterons uniquement aux ensembles d'orientations obtenus par subdivision récursive de l'octaèdre régulier.

Le DH3DO est défini comme une transformation de Hough 3D unique, associée à un repère arbitraire de l'ensemble des 48 repères ACP possibles de l'objet et à un ensemble C-invariant d'orientations O . Il est défini comme une fonction de trois variables, notée $\{h(s_i, \theta_k, \varphi_j)\}_{s_i \in \{0, \Delta, \dots, (N_s-1)\Delta\}, (\theta_k, \varphi_j) \in O}$, où

$$\Delta_s = S_{\max} / N_s, \text{ et } S_{\max} \text{ est la taille du maillage.}$$

La valeur $h(s_i, \theta_k, \varphi_j)$ cumule les contributions (exprimées en termes d'aire et/ou d'orientation) des facettes du maillage dont les centres de gravité sont traversés par au moins un plan de vecteur normal n_{jk} , situé à une distance de l'origine du repère dans l'intervalle $[s_i, s_{i+1})$.

Remarquons que pour $N_s = 1$, le DH3DO devient une variante d'EGI [1].

La mesure de similarité entre deux DH3DO, notés h_1 et h_2 , est définie par :

$$d(h_1, h_2) = \min_{c \in C} \left\{ \|h_1 - h_2^c\| \right\}, \quad (3)$$

où h_2^c représente la transformée de Hough 3D associée au repère $c(R)$ de C .

Grâce à la propriété de C-invariance de l'ensemble d'orientations considéré, h_2^c peut être entièrement déduite à partir de h_2 pour tout c de C , et la mesure de similarité ainsi définie est naturellement symétrique [1], [2], [37].

Pour une subdivision de l'octaèdre à deux niveaux de résolution et pour $N_s = 10$ intervalles de quantification des distances, le DH3DO est représenté par un vecteur de 1280 nombre flottants. Afin de réduire la taille du DH3DO et de la rendre comparable avec celle des autres descripteurs considérés ici, nous proposons d'effectuer une quantification vectorielle (QV) des valeurs du DH3DO.

3.1.5 Descripteur DH3DO avec QV (DH3DO-QV)

Pour chaque orientation (θ_k, φ_j) , nous construisons le vecteur 1D suivant :

$$\bar{h}(\theta_k, \varphi_j) = (h(0, \theta_k, \varphi_j), \dots, h((N_s - 1)\Delta, \theta_k, \varphi_j))^t \quad (4)$$

et le quantifions par rapport à un dictionnaire de prototypes fixé, en utilisant un critère de distance euclidienne minimale.

Le DH3DO s'exprime alors comme une fonction de deux variables, notée $h^q(\theta_k, \varphi_j)$. Pour chaque orientation (θ_k, φ_j) , $h^q(\theta_k, \varphi_j)$ représente l'index du vecteur prototype le plus proche de $\bar{h}(\theta_k, \varphi_j)$.

Dans l'exemple considéré, le DH3DO quantifié sera ainsi représenté par un vecteur de 128 indices entiers !

Cette approche par QV permet également de réduire drastiquement (d'un facteur N_s) le temps de calcul de la mesure de similarité, puisque les distances mutuelles entre les vecteurs prototypes du dictionnaire peuvent être pré-calculées au niveau du moteur de recherche.

Pour déterminer le dictionnaire des prototypes nous avons appliqué un algorithme LBG [38] sur un ensemble d'apprentissage constitué de la totalité des vecteurs $\bar{h}(\theta_k, \varphi_j)$ associés aux DH3DOs d'un sous-ensemble de 32 objets 3D tirés au hasard de la base MPEG-7 catégorisée de 363 modèles (cf. § 4). La taille du dictionnaire varie entre 512 et 1024 vecteurs prototypes.

3.1.6 Le spectre de forme MPEG-7

Le spectre de forme 3D (SF3D) [12] est défini comme la distribution de l'index de forme (IF) [17] sur l'ensemble des facettes du maillage. L'IF est un critère géométrique différentiel bien connu, défini pour tout point d'une surface 3D en fonction du rapport des deux courbures principales, qui permet de classer différentes formes géométriques locales en familles élémentaires comme cylindres, sphères, selles...

3.2 Approches 2D/3D

Les approches 2D/3D retenues sont fondées sur les deux DF 2D MPEG-7, que nous décrivons ci-dessous.

3.2.1 Descripteur ART (Angular Radial Transform)

Le descripteur ART² est défini comme l'ensemble des coefficients $c_{m,n}$ du développement de la fonction support de l'objet, notée $r(\rho, \theta)$, dans une base de fonctions $\{f_{mn}\}_{mn}$ définies sur le disque unitaire de façon séparable et données par :

$$f_{m,n}(\rho, \theta) = \frac{1}{\pi} \cos(\pi n \rho) e^{jm\theta}, \quad (5)$$

où $m \in \{0, \dots, M-1\}$ et $n \in \{0, \dots, N-1\}$. Les valeurs de M et de N recommandées par MPEG-7 sont respectivement

obtenue de manière intrinsèque, en considérant uniquement les valeurs absolues des coefficients de la décomposition.

Enfin, les mesures de similarité sont fondées sur la norme L_1 entre les vecteurs de coefficients.

Le descripteur ART présente l'avantage de la généralité, étant applicable à toute forme. Toutefois, dans des cas spécifiques où les formes considérées peuvent être représentées par un unique contour, il est plus approprié de considérer une représentation fondée sur l'information de contour, comme dans le cas du deuxième descripteur adopté par MPEG-7, décrit ci-dessous.

3.2.2 Descripteur par espace échelle de contour (EEC)

Ce descripteur³ propose de décrire une forme en en caractérisant le contour, en exploitant les maxima de courbure [39], détectés à travers une analyse multi-échelle dans un espace échelle gaussien. Cette analyse multi-échelle garantit la robustesse de la représentation au bruit et assure la sélection des maxima de courbure prédominants, *i.e.* persistants à plusieurs échelles. Elle implique un lissage linéaire récursif de la courbe, qui s'arrête au moment où l'on obtient une courbe lissée convexe.

Un nombre maximal de 64 maxima de courbure, déterminés de manière robuste dans le cadre d'une analyse multi-échelle gaussienne, est ensuite retenu. Le descripteur inclut leur coordonnées, leur valeurs et deux paramètres globaux de forme, la circularité (C) et l'excentricité (E).

L'invariance du descripteur par rapport aux transformations de similarité est réalisée à l'aide de mécanismes classiques d'alignement et normalisation. Ainsi, le choix de l'origine du repère au centre de gravité de la courbe garantit-il l'invariance aux translations. Pour assurer un comportement invariant aux rotations, les pics de courbure sont stockés à partir de celui de plus grande amplitude. Quant à l'invariance à l'échelle, elle est réalisée par une normalisation de type boîte englobante.

La mesure de similarité associée est fondée sur une fonction de coût d'appariement entre les maxima de courbure détectés. Les paramètres globaux de forme sont également utilisés dans l'étape d'appariement pour éliminer d'emblée les formes peu susceptibles d'être similaires à l'objet requête.

4 Analyse comparée et critères

La comparaison des différents DFs repose sur un ensemble de critères, sélectionnés en raison de leur pertinence au regard d'applications de requête par similarité dans des bases 3D génériques :

- coût de stockage, noté T , représentant la taille du descripteur exprimée en nombre de composantes du vecteur de représentation,
- complexité de l'algorithme d'extraction du DF à partir des données 3D, notée C_E
- complexité de calcul de la mesure de similarité associée, notée C_{MS} ,
- généralité (si le DF s'applique ou non aux maillages de géométrie et topologie arbitraires), notée G ,
- invariance géométrique (I_G) aux transformées de similarité et topologique (I_T) aux représentations maillées multiples,

Le Tableau 1 synthétise la façon dont les différents DFs 3D et 2D/3D considérés répondent à ces divers critères.

Chaque DF est représenté comme un vecteur 1D, 2D ou 3D, composé respectivement d'un nombre N_{1D} , N_{2D} et N_{3D}

d'éléments. Les valeurs de N_{1D} , N_{2D} et N_{3D} classiquement associées (notées V_{type}) à chaque descripteur sont également indiquées dans le tableau 1. Pour les valeurs types spécifiées, elles conduisent à des coûts de stockage tout à fait comparables pour l'ensemble des DF considérés, à l'exception du DH3DO, d'un ordre de grandeur plus complexe, mais rendu compact via le DH3DO-VQ.

Concernant le coût d'extraction, il est difficile à estimer en raison des multiples procédures d'optimisation, envisageables. De plus, il n'est pas déterminant pour le temps de réponse aux requêtes. Pour ces raisons, nous n'avons présenté qu'une évaluation qualitative, par des notes "+" et "++".

En ce qui concerne les temps de calcul des mesures de similarité, le DH3DO présente une complexité plus élevée que celle des autres DFs, en raison de l'intégration de la procédure d'alignement qui lui assure un *comportement invariant* (noté par CI dans le Tableau 1) géométriquement, qui évite les faux alignements. Ce handicap est levé par le DH3DO-VQ, dont la complexité est constante ($O(1)$), grâce aux distances pré-calculées entre vecteurs prototypes.

Enfin, chacun des DFs est bien générique, à l'exception du DF-EEC, applicable uniquement à des modèles dont les projections 2D sont représentables par des contours fermés.

Pour établir une comparaison objective des performances des requêtes par similarité des différents descripteurs considérés, les résultats expérimentaux ont été tout d'abord conduits sur la base MPEG-7 de 1300 modèles au format VRML.

4.1 Evaluation comparative objective

Cette base contient un sous-ensemble de 363 modèles catégorisés en 23 catégories, selon des critères géométriques et sémantiques. Cela conduit à une variabilité assez importante en terme de perception intuitive de forme : des avions à une ou plusieurs ailes, simples ou en delta, des humanoïdes avec postures et attitudes très différentes, des voitures anciennes ou modernes, décapotables ou non, sont respectivement regroupés dans la classe générique "Avions", "Humanoïdes" et "Voitures". Cela rend la base catégorisée suffisamment complexe pour tester effectivement les performances des DF 3D et 2D/3D. Pour plus de détails sur la base MPEG-7, le lecteur est invité à se reporter à [1], [37].

Les résultats expérimentaux présentés ici sont obtenus avec les paramètres suivants : histogrammes à 256 éléments pour les descripteurs HLC, HASC et HTA ; histogramme à 100 éléments pour le SF3D ; $N_s = 20$ et deux niveaux de décomposition de l'octaèdre pour le DH3DO ; 34 coefficients ART ; 8 maxima de courbure pour le EEC. Pour les DFs 2D/3D, nous avons utilisé un nombre de 7 vues, déterminées à partir des modèles 3D, comme recommandé dans MPEG-7 : 3 vues principales, correspondant aux directions des axes d'inertie de l'objet, et 4 vues secondaires, correspondant aux directions diagonales des quatre différents octants du repère ACP (Figure 2).



Figure 2. Les sept vues de projection, principales (les trois premières) et secondaires (les quatre dernières), pour un modèle d'humanoïde.

³ Descripteur promu

Pour évaluer quantitativement et objectivement les performances des DFs proposés, nous avons calculé les deux paramètres suivants :

- le score *Bull-Eye* (SBE), retenu comme critère d'évaluation de tous les descripteurs de forme considérés dans MPEG-7,
- le score *First Tier* (FT), utilisé dans [9].

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

$$SBE = \frac{N_{cor}(2Q)}{Q}, FT = \frac{N_{cor}(Q)}{Q} \quad (7)$$

où $N_{cor}(x)$ est le nombre de résultats corrects (*i.e.* maillage appartenant à la même catégorie que celle de la requête) parmi

les x premiers résultats retrouvés, et Q est le nombre d'éléments de la catégorie correspondant à la requête.

Si le SBE est relativement permissif, admettant une marge d'erreur de 50% pour retrouver l'ensemble des modèles recherchés, le FT est beaucoup plus drastique, un FT de 100% n'autorisant aucune erreur.

Afin d'étudier la scalabilité (*i.e.* persistance de bonnes performances avec l'augmentation du volume de données) des DF analysés, nous avons plongé l'ensemble des modèles catégorisés dans la base des 1300 modèles 3D et calculé le nouveau SBE, noté SBE_{tot} .

Les valeurs obtenues pour ces trois scores sont synthétisées dans le Tableau 2.

DF	Représentation	T	V _{type}	C _E	C _{MS}	I _G	I _T	G
SF3D	Vecteur 1D	N _{1D}	64-128	+	O(N _{1D})	Oui	Non	Oui
HLC	Vecteur 1D	N _{1D}	64-256	+	O(N _{1D})	Non	Non	Oui
HLAS	Vecteur 2D	N _{2D}	128-256	++	O(N _{2D})	Non	Non	Oui
HTA	Vecteur 1D	N _{1D}	64-256	+	O(N _{1D})	Oui	Non	Oui
EGI	Vecteur 2D	N _{2D}	128-256	++	O(N _{2D})	Non	Non	Oui
DH3DO	Vecteur 3D	N _{3D}	1280-2560	++	O(48 x N _{3D})	CI	Oui	Oui
DH3DO-QV	Vecteur 2D	N _{2D}	128-256	++	O(1)	CI	Oui	Oui
2D/3D ART	Vecteur 2D	V x N _{2D}	34	++	O(V ² x N _{2D})	Non	Oui	Oui
3D/3D EEC	Vecteur 1D	V x N _{1D}	8-64	++	O(V ² x N _{2D})	Non	Oui	Non

Tableau 1. Comparaison des DFs et performances associées. N_{iD}, est le nombre d'éléments d'un vecteur iD (i=1, 2, 3), V_{type} désigne les valeurs classiquement associées aux paramètres N_{iD} et CI comportement invariant.

Approche	SF3D	HLC	HASC	HTA	EGI	DH3DO	DH3DO QV	2D/3D ART	2D/3D EEC
SBE (%)	53	56	47	58	66	81	79	66	74
FT (%)	44	48	44	43	57	76	72	58	64
SBE _{tot} (%)	41	44	38	45	54	76	73	53	65

Tableau 2. Performances des différents DF selon les critères : SBE, FT et SBE_{tot}.

Ces résultats mettent en évidence la nette supériorité du DH3DO, avec un SBE et un FT de 81% et 76%, respectivement. La quantification vectorielle (réalisée ici avec un dictionnaire de 512 prototypes), conduit à une légère dégradation des performances, de 2% en terme de SBE, performances qui restent toutefois supérieures à celles des autres DFs considérés. Cela valide notre construction par quantification vectorielle d'un DH3DO de faible complexité, mais avec des performances pratiquement équivalentes.

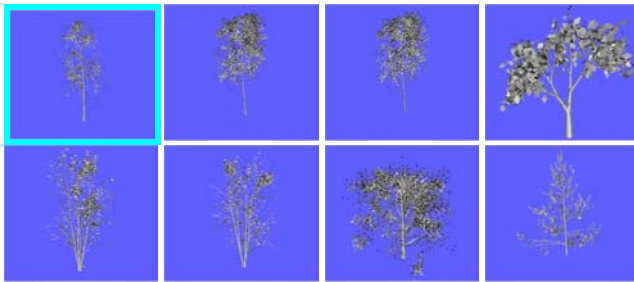
Pour illustrer les performances des requêtes par similarité du DH3DO, présentons trois exemples de requêtes par similarité, empruntées aux catégories "Arbres avec feuilles", "Humanoïdes" et "Avions" (Figure 3). Les modèles retrouvés sont triés par ordre décroissant de similarité par rapport au modèle requête, de gauche à droite et de haut en bas.

Dans tous les cas, les modèles retrouvés aux huit premières positions correspondent à la même catégorie que la requête, malgré une variabilité relativement importante de leur forme.

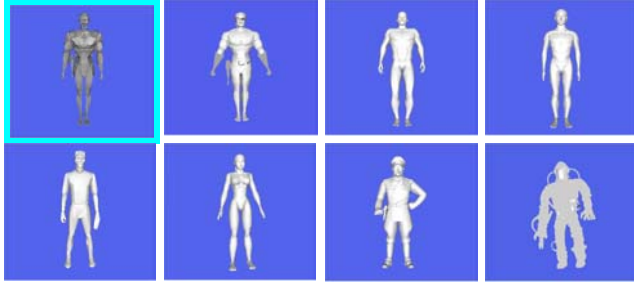
Le deuxième élément à retenir de cette étude est lié aux très bonnes performances des approches 2D/3D, et en particulier à celles du DF-EEC, qui se place en deuxième position avec un bon SBE de 74%. Même si le FT descend à 64%, ce score dépasse largement ceux des DFs 3D par EGI ou par histogrammes (SF3D, HLC, HAS, et HTA), malgré le fait que ces derniers exploitent une information complète 3D. Cela

démontre la pertinence élevée de cette approche qui, en exploitant une information de projection 2D, beaucoup moins complète mais bien plus maniable, bien définie topologiquement et admettant une paramétrisation naturelle par abscisse curviligne, permet de discriminer les objets 3D mieux que les approches purement 3D susmentionnées.

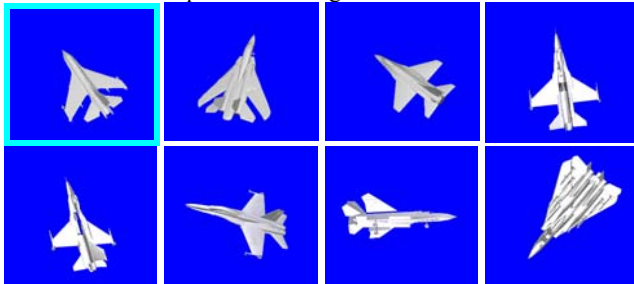
Quant au DF-ART, il conduit à des résultats inférieurs à ceux du DF-EEC, avec un SBE de seulement 66%. Cela n'est pas surprenant, la diminution des performances étant le prix à payer pour une plus grande généralité. Toutefois, les



3.a. Requête de la catégorie "Arbres avec feuilles".



3.b. Requête de la catégorie "Humanoïdes".



3.c. Requête de la catégorie "Avions".

Figure 3. Modèles retrouvés pour des requêtes par similarité sur la base MPEG-7 avec le DH3DO.

Afin d'élargir la portée de l'étude comparative proposée, nous avons également considéré les bases de modèles 3D de Princeton [9] (dont la quasi-totalité des 133 modèles fait partie de la base MPEG-7 de 1300 maillages) et la base Utrecht [25], composée essentiellement de modèles d'avions (512 modèles catégorisés en cinq classes : conventionnels, biplans, hélicoptères, deltas et multifuselages).

Sur la base restreinte de Princeton, l'ensemble des différentes méthodes proposées dans [9] conduisent à un score SBE entre 50 et 60%, tandis que l'approche de [25] obtient un score compris entre 66 et 71%. Les scores SBE de 81%, 79% et 74% respectivement des DH3DO, DH3DO-VQ et EEC, sur la base MPEG-7 dépassent donc largement ceux rapportés dans [9] et [25] sur un sous-ensemble de cette base de cardinalité plus inférieure (133 modèles).

En ce qui concerne la base Utrecht, elle est mal adaptée aux descripteurs considérés dans cet article, qui ne visent pas une discrimination si fine. Bien au contraire, l'objectif ici est de retrouver, dans une base de modèles variés, des avions de tout type, malgré leur diversité. Ainsi, la catégorie "Avions" que nous avons définie dans la base MPEG-7, regroupe-t-elle aussi bien des avions conventionnels que des multifuselages, des biplans ou des deltas. L'excellent score SBE obtenu par le DH3DO (94%) pour cette catégorie démontre que cet objectif est bien atteint !

Nous retenons donc de cette étude la supériorité du DH3DO et du DF-EEC, seuls DFs à offrir une description de forme compacte et suffisamment complète pour rendre les résultats des requêtes par similarité à la fois pertinentes et

efficaces. Au prix d'une légère baisse en termes de performances des requêtes, le DF EEC offre en plus la possibilité d'effectuer des appariements entre images 2D (pouvant provenir d'esquisses manuelles ou d'acquisitions numériques 2D en imagerie multicaméra) et modèles 3D. Cette fonctionnalité peut être un atout important dans le cadre de certaines applications (cf. § 5.2). Il serait alors intéressant de savoir si le DH3DO ne pourrait être également doté d'un mécanisme de requête 2D/3D.

La réponse à cette question est positive, à condition que les différentes images 2D correspondent à des angles de vues prédéterminés et supposés connus. Cette hypothèse permet alors de reconstruire, à partir des projections un modèle 3D plus ou moins approximatif. La Figure 4 présente un exemple de reconstruction 3D à partir des sept vues présentées Figure 2, et obtenue avec un algorithme classique d'intersection de volumes [40].

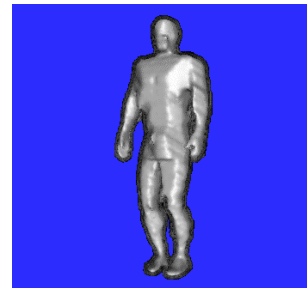


Figure 4. Objet 3D volumique reconstruit par intersection de volumes à parties des sept vues 2D de la Figure 2.

Pour tester la viabilité de cette solution, nous avons tout d'abord créé pour chaque modèle de la base MPEG-7 une version volumique par intersection de volumes, que nous avons ensuite indexée par le DH3DO. Les résultats des requêtes par similarité ont conduit alors à un score SBE de 70%, légèrement inférieur à celui du DF-EEC, mais supérieur à celui du descripteur ART. Cela prouve la pertinence de la solution proposée, d'autant plus que l'approche est plus générique que celle par EEC, pouvant s'appliquer à des formes arbitraires (en topologie notamment).

En raison de ses excellentes performances en termes de requêtes par similarité et des fonctionnalités offertes, à la fois 3D et 2D/3D, nous avons considéré le DH3DO dans un contexte applicatif nouveau et spécifique, qui est celui défini par le projet RNRT SEMANTIC-3D : Service d'Echange et de MANipulation (Tatouage, Indexation, Compression) de pièces automobile [4].

4.2 Application SEMANTIC-3D

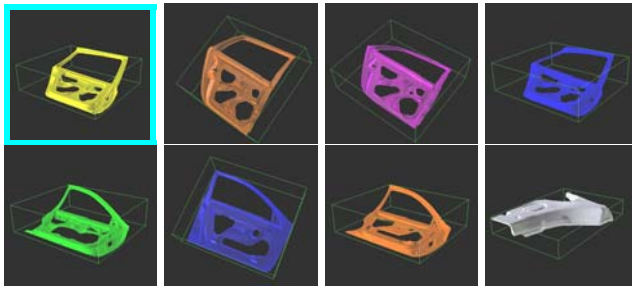
Ce projet vise à élaborer un ensemble méthodologique unifié pour la compression, l'indexation et le tatouage d'objets 3D, dans le cadre d'une application de télédistribution de pièces automobiles. Dans sa composante indexation, l'objectif du projet est d'élaborer et de mettre au point des descripteurs aussi bien 3D que 2D/3D, adaptés aux différents scénarios d'interrogation à distance des bases de pièces, établis par le partenaire industriel qui est RENAULT. Ces scénarios incluent des requêtes à partir des esquisses réalisées à la main, à partir d'un ensemble de vues des pièces 3D acquises avec une ou plusieurs caméra numérique, à partir d'un modèle acquis avec un scanner 3D, ou encore, pour des raisons de bande passante, à partir de versions compressées et donc dégradées d'un objet 3D.

Aux contraintes habituelles d'invariance géométrico-topologique s'ajoute donc ici celle de robustesse des requêtes par rapport aux versions compressées d'un même objet 3D.

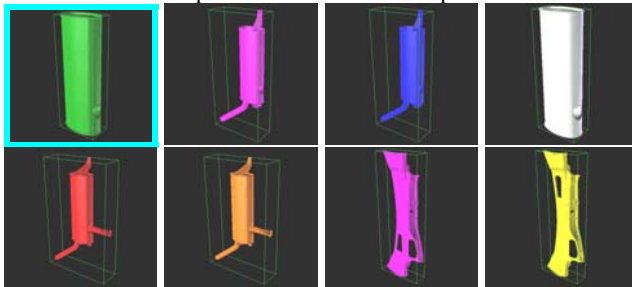
En outre, notons que dans le cadre de cette application spécifique, la contrainte de connaissance *a priori* des angles de vue qui permet au DH3DO d'effectuer des requêtes 2D/3D peut être satisfaite dans des conditions complètement réalistes.

Le corpus de données constitué et mis à la disposition des partenaires du projet par RENAULT, inclut environ 2200 modèles 3D au format VRML, représentant l'ensemble des pièces, pour les différentes versions de la partie avant d'une automobile RENAULT de type Laguna et correspondant aussi bien à des pièces de tolérances que de mécanique, de sellerie ou de câbles..

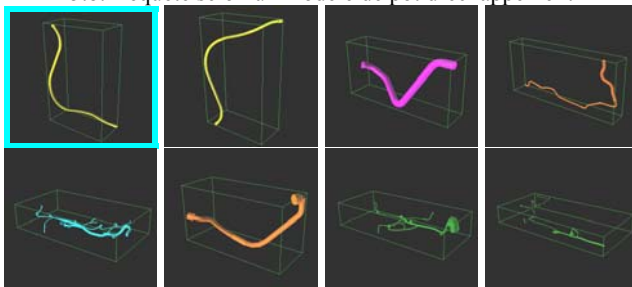
La Figure 5 présente quelques résultats de requêtes par similarité pour trois modèles requête correspondant à : une portière, un pot d'échappement et un câble.



5.a. Requête selon un modèle de portière.



5.b. Requête selon un modèle de pot d'échappement



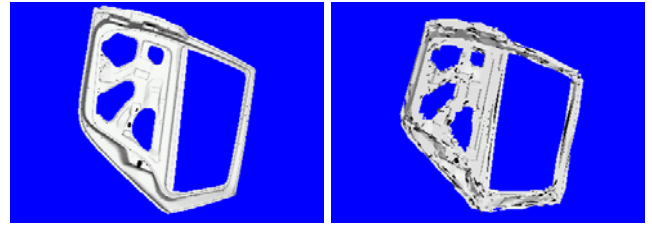
5.c. Requête selon un modèle de pot d'échappement.

Figure 5. Modèles retrouvés pour des requêtes par similarité sur le corpus SEMANTIC-3D (DH3DO-QV).

Dans tous les cas, les résultats retrouvés en premières positions correspondent bien à des modèles similaires à celui de la requête. Remarquons que pour la requête "portière", des portières aussi bien avant que arrière sont retrouvées. Pour la requête "pot d'échappement" les différents autres pots, avec ou sans les tuyaux adjacents, sont également trouvés en premières positions. Enfin, dans le dernier exemple sont retrouvées uniquement des structures filiformes.

Pour déterminer la robustesse du DH3DO-QV par rapport à la compression, nous avons reproduit les requêtes, mais cette fois avec des versions fortement compressées (à 6 bits par

sommet) des modèles donnés comme exemple, obtenues avec le logiciel de référence MPEG-4. Le modèle requête de la Figure 5.a dans sa version originale (fichier VRML de 850 koctets) et compressé (fichier binaire MPEG-4 de 18 koctets) est illustré Figure 6.



6.a. Modèle initial.

6.b. Version compressée.

Figure 6. Le modèle requête de la figure 5.a en version originale (6.a) et compressée à 6 bits/sommet (6.b).

Pour l'ensemble des trois requêtes illustrées Figure 6, l'ordre des résultats retrouvés reste le même. Cette robustesse du DH3DO à la compression s'explique par le caractère global de la description qui le rend naturellement peu sensible aux déformations locales de faibles amplitudes, caractéristiques des maillages compressés.

5 Conclusion et perspectives

Dans cet article, nous avons présenté une étude comparée des approches aussi bien 3D que 2D/3D pour la représentation de modèles polygonaux 3D dans le cadre d'applications de requête par similarité de forme.

Les critères utilisés renvoient aussi bien à la capacité de la représentation des DF à vérifier des propriétés d'invariance géométrique et/ou topologique qu'à leur caractère général permettant de les appliquer ou non à tout type d'objet 3D maillé, ou encore qu'à leur complexité de calcul ou la complexité de calcul de la mesure de similarité mise en œuvre.

Des sept DF différents et dont les performances ont été comparées objectivement par évaluation des SBE et FT sur la base catégorisée MPEG-7, le DH3DO surclasse nettement tous les autres, quelle que soit leur nature, 3D ou 2D/3D, avec un SBE de 81%. De plus, la nouvelle version du DH3DO, obtenue via une quantification vectorielle permet de réduire d'un ordre de grandeur la taille du DH3DO et la complexité de calcul de la mesure de similarité associée, tout en conservant d'excellentes performances en terme de requêtes par similarité (SBE de 79%).

Retenons également les très bonnes performances des approches 2D/3D et en particulier au travers du DF EEC, qui, avec un SBE de 74%, devance significativement les autres descripteurs 3D considérés ici (EGI, HLC, HASC, HTA et SF3D).

Enfin, appliqué dans le contexte spécifique de télédistribution de pièces automobiles, définie dans le cadre du projet RNRT SEMANTIC-3D, le DH3DO a démontré ses très bonnes performances en termes de discrimination entre les différents types de pièces considérés, de scalabilité et de robustesse face à des versions fortement compressées de maillages (6 bits par sommet).

A court terme, les perspectives de notre recherche et développement concernent la mise en œuvre complète du DH3DO-VQ dans le cadre du projet SEMANTIC-3D, et à moyen terme la prise en compte des différentes composantes sémantiques auxquelles un DF doit renvoyer au travers des schémas de description conceptuelle aux structures adaptatives.

6 Bibliographie

- [1] T. Zaharia, F. Prêteux, "Indexation de maillages 3D par descripteurs de forme", *13ème Congrès Francophone de Reconnaissance de Forme et Intelligence Artificielle* (RFIA2002), vol. pp. Angers, France, Janvier 2002.
- [2] T. Zaharia, F. Prêteux, "Shape-based retrieval of 3D mesh models", *Proc. IEEE International Conference on Multimedia and Expo* (ICME 2002), Suisse, Août 2002.
- [3] T. Zaharia, F. Prêteux, "Hough transform-based 3D mesh retrieval", *Proc. SPIE Conference on Vision Geometry*, San Diego, Etats-Unis, Août 2001.
- [4] <http://www.telecom.gouv.fr/RNRT>.
- [5] T. Mura, "Descriptors of polyhedral data for 3D-shape similarity search", *Proposal P177, MPEG-7 Proposal Evaluation Meeting*, UK, Fév. 1999.
- [6] M. Elad, A. Tal, S. Ar, "Directed search in a 3D objects database using SVM", Hewlett-Packard Research Report HPL-2000-20R1, 2000.
- [7] C. Zhang, T. Chen, "Efficient feature extraction for 2D/3D objects in mesh representation", *Proc. of the International Conference on Image Processing* (ICIP 2001), 2001.
- [8] E. Paquet, M. Rioux, A. Murching, T. Naveen., A. Tabatabai, "Description of shape information for 2-D and 3-D objects", *Signal Processing: Image Communications*, Vol. 16, pp. 103-122, Sept. 2000.
- [9] R. Osada, T. Funkhouser., B. Chazelle., D.P. Dobkin., "Matching 3D models with shape distributions", *Proc. International Conference on Shape Modeling and Applications*, Mai 2001.
- [10] G. Taubin et D.B. Cooper, "Object recognition based on moment (or algebraic) invariants", *J.L. Mundy and A. Zisserman, editors, Geometric Invariants in Computer Vision*, pp. 375-397, MIT Press, 1992.
- [11] J. Lefebvre, "Introduction aux analyses statistiques multidimensionnelles", Masson, Paris, 1983.
- [12] T. Zaharia, F. Prêteux, "3D-shape-based retrieval within the MPEG-7 framework", *Proc. SPIE Conf. on Nonlinear Image Processing and Pattern Analysis XII*, Vol. 4304, pp.133-145, San Jose, Etats-Unis, Janv. 2001.
- [13] B.K.P Horn, "Extended Gaussian Image", *Proc. of the IEEE*, Vol. 72, pp. 1671-1686, 1984.
- [14] S.B. Kang, K. Ikeuchi., "The complex EGI: a new representation for 3D pose determination", *IEEE Trans. on PAMI*, Vol. 16 (3), pp. 249-58, Mars 1994.
- [15] H. Matsuo, A. Iwata, "3D Object Recognition using MEGI model from range data", *Proc. 12th Int. Conf. on Pattern Recognition*, pp. 843-846, Oct. 1994.
- [16] C. Dorai, A.K. Jain, "Shape Spectrum-Based View Grouping and matching of 3D Free-Form Objects", *IEEE Trans. on PAMI*, 19(10), pp. 1139-1145, Oct. 1997.
- [17] J. Koenderink, "Solid shape", The MIT Press, Cambridge, Massachusetts, 1990.
- [18] J. R. Ullmann. "An algorithm for subgraph isomorphism", *Journal of the ACM*, Vol. 1, No. 23, pp. 31-42, Janv. 1976.
- [19] S. J. Dickinson, D. Metaxas, A. Pentland, "The role of model-based segmentation in the recovery of volumetric parts from range data", *IEEE Trans. on PAMI*, Vol. 19, No. 3, pp. 259-267, Mars 1997.
- [20] S. Dickinson, A. Pentland, S. Stevenson, "Viewpoint-invariant indexing for content-based image retrieval", *Proc. IEEE Int. Workshop on Content-Based Access of Image and Video Database*, pp. 20-30, Janv. 1998.
- [21] K. Wu, M.D. Levine, "Recovering of parametric geons from multiview range data", *Proc. IEEE Conf. on CVPR.*, pp. 159-166, Juin 1994.
- [22] D. Terzopoulos, D. Metaxas, "Dynamic 3D models with local and global deformations: Deformable superquadrics", *IEEE Trans. on PAMI*, 13 (7), pp. 703-714, Juillet 1991.
- [23] Park I.K., Kim J.D., Kim D.H., Yun I.D., Lee S.U., "3D perceptual shape descriptor: Result of exploration experiments and proposal for core experiments", *ISO/IEC JTC1/SC29/WG11*, MPEG02/M9210, Japon, Déc. 2002.
- [24] Park I.K., Kim J.D., Kim D.H., Yun I.D., Lee S.U., "Perceptual 3D shape descriptor: Result of core experiment", *ISO/IEC JTC1/SC29/WG11*, MPEG03/M9481, Thaïlande, Mars 2003.
- [25] J.W.H. Tangelder, R.C. Veltkamp, "Polyhedral model retrieval using weighted point sets", Rapport technique no UU-CS-2002-019, Université de Utrecht, Pays-Bas, 2002.
- [26] M. Hebert, K. Ikeuchi, H. Delingette, "A spherical representation for recognition of free-form surfaces", *IEEE Trans. on PAMI*, 17(7), pp. 681-90, Juil. 1995.
- [27] D. Zhang and M. Hebert, "Harmonic maps and their applications in surface matching" *Proceedings of the Int. Conf. on CVPR*, pp. 524-530, Juin 1999.
- [28] A. Johnson et M. Hebert, "Using spin-images for efficient multiple model recognition in cluttered 3-d scenes", *IEEE Trans. on PAMI*, Vol. 21, No. 5, pp. 433-449, Mai 1999.
- [29] P. V. C. Hough, "Method and means for recognizing complex patterns", U.S. Patent 3 069 654, 1962.
- [30] D. H. Ballard, "Generalizing the Hough transform to detect arbitrary shapes", *Pattern Recognition*, Vol. 13, No. 2, pp. 111-122, 1981.
- [31] J. Illingworth et J. Kittler, "A survey of the Hough transform", *Computer Vision, Graphics and Image Processing*, Vol. 44, pp. 87-116, Oct. 1988.
- [32] M.T Suzuki, T. Kato, N. Otsu., "A similarity retrieval of 3D polygonal models using rotation invariant shape descriptors", *IEEE Int. Conf. on Systems, Man, and Cybernetics*, pp. 2946-2952, Etats-Unis, Oct. 2000.
- [33] A. Pentland, S. Sclaroff, "Closed form solutions for physically-based shape modelling and recognition", *IEEE Trans. on PAMI*, Vol. 13, No. 7, pp. 715-729, Juillet 1991.
- [34] S. Sclaroff, "Deformable prototypes for encoding shape categories in image databases", *Pattern Recognition*, Vol. 30, No. 4, pp. 627-641, Avril 1997.
- [35] C. Nastar, "Generalized Image matching : Statistical learning of physically-based deformations", *Rapport de recherche MIT Media Lab.*, No. 368, Avr. 1996.
- [36] J. Malik, "Shape Matching for Content based Image Retrieval", *Int. Workshop on Multimedia Content-based Indexing and Retrieval*, INRIA Rocquencourt, France, Sept. 2001.
- [37] T. Zaharia, "Indexation de vidéos et de maillages 3D dans le contexte MPEG-7", Thèse de Doctorat, Univ. Paris V René Descartes, Paris, France, Déc. 2001.
- [38] Y. Linde, A. Buzo, R. M. Gray, "An Algorithm for Vector Quantizer Design", *IEEE Trans. on Communications*, Vol. COM-28, No.1, pp. 84-95, Janv. 1980.
- [39] F. Mokhtarian, A.K. Mackworth, "A theory of multiscale, curvature-based shape representation for planar curves", *IEEE Trans. on PAMI*, 14 (