

Recherche de modèles 3D de pièces mécaniques basée sur les moments de Zernike

Tarik Filali Ansary

Jean-Philippe Vandeborre

Mohamed Daoudi

MIIRE – GET / INT / LIFL UMR CNRS-USTL 8022

ENIC Télécom Lille 1

Cité Scientifique, rue G. Marconi

59658 Villeneuve d'Ascq cedex, France

{filali, vandeborre, daoudi}@enic.fr

Résumé

Les modèles tridimensionnels sont de plus en plus utilisés dans les applications où la nécessité de visualiser des objets réalistes se fait sentir (CAD/CAO, simulations médicales, jeux, réalité virtuelle, etc.). Par conséquent, la gestion de collections de données 3D de grande taille devient un domaine important. L'indexation de telles données permet, par exemple, à un concepteur de retrouver facilement des données similaires – au sens visuel ou sémantique – à un objet requête présenté. Pour cela, deux grandes approches existent : la recherche par une vue 2D de l'objet 3D recherché, et la recherche par l'exemple d'un modèle 3D directement. Dans cette contribution, nous nous intéressons à ce dernier point mais en nous appuyant sur une caractérisation des modèles 3D à partir d'une description des vues caractéristiques basée sur les moments de Zernike. De plus, le système présenté s'appuie sur une approche probabiliste bayésienne. Nous présentons les résultats obtenus sur une collection de 760 modèles 3D de pièces automobiles, et nous les comparons à ceux obtenus grâce à des descripteurs purement 3D classiques.

Mots clefs

Modèles 3D, indexation 3D, vues caractéristiques, moments de Zernike, probabilités.

1 Introduction

Les développements technologiques récents concernant l'imagerie tridimensionnelle (outils d'acquisition comme les scanners 3D, modeleurs, cartes graphiques dédiées, etc.) rendent possibles la création et le stockages de modèles tridimensionnels à grande échelle. L'utilisation des modèles tridimensionnels se retrouve dans de nombreuses applications telles que le patrimoine culturel, les simulations médicales, l'industrie mécanique, les jeux, la réalité virtuelle, et, de manière générale, tout ce qui touche au multimédia. De plus, avec l'expansion de l'Internet qui se poursuit et le développement des accès haut-débit, les modèles tridimensionnels feront de plus en plus partie inté-

grante des bases de données distribuées liées aux services multimédia.

De nombreux systèmes ont été et sont encore actuellement développés pour la recherche d'information dans des collections d'images ou de vidéos. Cependant, les solutions proposées dans le cas d'information purement bidimensionnelles, ne sont pas très efficaces ou difficilement adaptables à des données intrinsèquement tridimensionnelles.

Pour résoudre ce problème, nous proposons une méthode d'indexation de modèles 3D s'appuyant sur les vues caractéristiques 2D de ces derniers. Les vues caractéristiques sont décrites par un invariant mathématique : dans notre implémentation, les moments de Zernike 2D. De plus, pour renforcer l'efficacité de notre système, nous utilisons une approche probabiliste bayésienne pour le processus d'indexation.

Cet article est organisé de la manière suivante. Le paragraphe 2 présente quelques travaux récents concernant l'indexation de modèles 3D selon les deux grandes familles d'approches : 2D/3D et 3D/3D. Le paragraphe 3 détaille notre processus de sélection des vues caractéristiques d'un modèle 3D, avant de présenter, dans le paragraphe 4, notre approche probabiliste s'appuyant sur ces vues. Le paragraphe 5 détaille les moments de Zernike 2D que nous utilisons pour décrire les vues caractéristiques. Le corpus de test, les résultats obtenus, ainsi que des comparaisons avec des méthodes d'indexation 3D classiques sont présentés dans le paragraphe 6 avant de conclure.

2 Travaux récents

Pour l'indexation de modèles 3D, deux grandes approches s'imposent : la recherche par l'exemple d'un modèle 3D (approches 3D/3D), et la recherche par une vue 2D de l'objet 3D recherché (approches 2D/3D).

Par exemple, Vandeborre et al. [1] proposent d'utiliser des descripteurs statistiques de courbures et de distances sur des modèles maillés, afin de les combiner pour améliorer les résultats de chacun d'eux pris séparément. Cependant,

cette approche montre ses limitations face aux maillages non réguliers.

Sundar et al. [2] représentent le modèle 3D maillé sous la forme d'un squelette. Ils utilisent des techniques de mise en correspondance de graphes pour comparer les squelettes et, par conséquent, les objets 3D dont ils sont issus. Ces auteurs suggèrent également que leur approche par squelette a la particularité de permettre la recherche par partie, ainsi que d'aider à définir les requêtes de manière intuitive.

Kazhdan et al. [3] décrivent une approche basée sur les harmoniques sphériques. A partir d'un ensemble de fonctions sphériques calculées sur une grille de voxels de l'objet 3D, un descripteur invariant à la rotation est calculé en décomposant la fonction par ses harmoniques sphériques et en sommant ces harmoniques pour chaque fréquence. La norme L_2 est ensuite calculée pour chaque composante. Le résultat est obtenu sous la forme d'un histogramme 2D indexé par le rayon et la fréquence.

Enfin, pour les approches purement tridimensionnelles, Tung et Schmitt [4] utilisent des graphes de Reeb mutirésolution pour indexer les modèles grâce à leurs topologies. Concernant les approches 2D/3D, les deux principaux problèmes sont les suivants : comment caractériser les modèles 3D avec des vues 2D, et comment utiliser ces vues pour retrouver le modèle dans une collection de modèles 3D.

Ohbuchi et al. [5] proposent une méthode assez *brute* basée sur l'apparence des objets. En effet, la méthode consiste à comparer directement quarante-deux images de profondeur créées pour chaque modèle.

Abbasi et Mokhtarian [6] présentent une méthode qui consiste à éliminer les vues similaires au sens de la distance entre les CSS (Curvature Scale Space) des contours des vues. Le nombre minimal de vues est enfin déterminé par un algorithme d'optimisation.

Mahmoudi et Daoudi [7] suggèrent également d'utiliser le CSS des contours des vues extraites des modèles 3D. Les CSS sont rangés dans une structure appelée *M-Tree* pour accélérer les accès et les comparaisons.

Dorai et Jain [8] utilisent, pour chaque modèle 3D dans leur collection, un algorithme de génération de trois cent-vingt vues. Ensuite, une classification hiérarchique, utilisant la mesure d'une distance entre des histogrammes de courbures, permet de réduire le nombre de vues.

Chen et Stockman [9], ainsi que Yi et Chelberg [10], proposent une méthode probabiliste bayésienne. Une probabilité *a posteriori* de reconnaître un modèle 3D, sachant que certaines caractéristiques sont observées, est calculée. D'après ces auteurs, cette méthode donne de bons résultats mais n'a été testée que sur une collection de petite taille, c'est-à-dire vingt modèles 3D.

Chen et al. [11][12] défendent l'idée intuitive que deux modèles 3D sont similaires si, vus sous différents angles, ils semblent également similaires. Par conséquent, ils utilisent cent projections orthogonales par modèle 3D, et les décrivent par les moments de Zernike et des descripteurs de Fourier.

Enfin, Tangelder et Velkamp [13] présentent un état de l'art de l'indexation et de la reconnaissance de modèles 3D très complet. Zaharia et Prêteux [14] proposent également un tel état de l'art tout en comparant les performances de descripteurs connus 2D/3D et 3D/3D.

C'est en se basant sur ces considérations que nous présentons, dans le paragraphe suivant, notre méthode de sélection des vues caractéristiques d'un modèle 3D, puis notre méthode probabiliste bayésienne d'indexation.

3 Sélection des vues caractéristiques

Puisque nous nous appuyons sur les vues caractéristiques d'un modèle 3D pour l'indexation, il convient de commencer par la méthode de sélection des vues caractéristiques d'un modèle 3D.

Soit $D_b = \{M_1, M_2, \dots, M_N\}$ une collection composée de N modèles 3D. Chaque modèle 3D de cette collection doit être représenté par un ensemble de vues. Dans ce but, pour chaque modèle 3D de la collection, le processus consiste à générer un ensemble de vues initiales, puis à réduire cet ensemble de vues à celles, et seulement celles, qui caractérisent au mieux le modèle 3D en question en fonction d'une distance entre les descripteurs de vues (dans notre implémentation, les moments de Zernike, voir paragraphe 5). Finalement, afin d'être certains d'obtenir l'ensemble minimal de vues caractéristiques, mais également de ne pas oublier d'information en abandonnant une vue, nous nous assurons que les deux critères suivants sont vérifiés : toute vue initiale du modèle doit avoir une vue caractéristique qui la représente – pour éviter le problème de sous-représentation –, et chaque vue caractéristique représente au moins une vue qui n'est pas représentée par une autre vue caractéristique – pour éviter le problème de sur-représentation.

Cette idée vient simplement du fait que toutes les vues d'un modèle 3D n'ont pas la même importance dans la caractérisation de ce modèle. En effet, certaines vues contiennent plus d'informations pertinentes que d'autres.

3.1 Génération de l'ensemble des vues initiales

Pour créer l'ensemble des vues initiales d'un modèle 3D de la collection, des vues – projections perspectives – sont calculées à partir de différents points de vue. Ces points de vue sont équi-répartis sur la sphère englobante de l'objet 3D grâce à une facétisation régulière en triangles de celle-ci. La direction de visée pointe sur le centre de masse du modèle 3D. Dans notre implémentation, nous utilisons 80 vues initiales.

Nous appelons $V_M = \{V_M^1, V_M^2, \dots, V_M^v\}$ l'ensemble des v vues initiales d'un modèle 3D M .

3.2 Réduction de l'ensemble des vues initiales

Parmi l'ensemble de vues initiales ainsi obtenu, nous devons sélectionner celles qui caractérisent effectivement le

modèle tridimensionnel en question en fonction d'un descripteur de ces vues, c'est-à-dire l'ensemble des vues caractéristiques V_{cM} . Le descripteur que nous avons choisi dans notre implémentation, même si notre méthode est *indépendante* de celui-ci, est les moments de Zernike 2D (voir paragraphe 5).

Une vue V_M^k du modèle M est considérée caractéristique pour une distance ε , si la distance, au sens du descripteur 2D choisi, entre cette vue et toutes les autres vues caractéristiques de M est supérieure à ε . Ce qui peut se formaliser de la manière suivante :

$$V_M^j \in V_{cM} \Leftrightarrow \forall V_{cM}^k \in V_{cM}, D_{V_M^j, V_{cM}^k} > \varepsilon$$

Avec $D_{V_M^j, V_{cM}^k}$ la distance entre les descripteurs des vues V_M^j et V_{cM}^k .

Cependant, le choix de la distance seuil ε est important et dépend de la complexité du modèle tridimensionnel, information qui n'est *a priori* pas connue.

Pour résoudre le problème de la détermination de la distance seuil ε , l'algorithme précédent a été adapté en prenant en compte un intervalle de ces distances de 0 à 1 avec un pas de 0,001. Cela signifie qu'une vue est considérée importante si elle revient fréquemment dans l'ensemble des vues caractéristiques pour différents ε .

Finalement, l'ensemble des vues caractéristiques d'un modèle M est l'ensemble des vues qui sont fréquemment considérés caractéristiques pour les différentes valeurs de ε .

3.3 Propriétés finales de l'algorithme de sélection des vues

L'ensemble des vues ainsi obtenu doit enfin respecter deux critères pour éviter la sous et la sur-représentation. Ces deux critères vont donc permettre de minimiser le nombre de vues caractéristiques réel.

Les deux critères sont les suivants :

- Chaque vue initiale du modèle M doit être représenté par au moins une vue caractéristique. Cela signifie que, si nous considérons que notre algorithme est une fonction \mathfrak{R} entre l'ensemble de vues initiales et l'ensemble des vues caractéristiques, cette fonction doit être une *application* associant chaque élément de V_M avec au moins un élément de V_{cM} :

$$\forall V_M^j \in V_M, \exists V_{cM}^k \text{ tel que } \mathfrak{R}(V_M^j) = V_{cM}^k$$

- Maintenant que nous sommes sûrs que chaque vue initiale est représentée par une vue caractéristique, nous souhaitons réduire cet ensemble à son minimum en éliminant les redondances qui auraient pu être créées.

Soit Vr_M^j l'ensemble des vues représentées par la vue caractéristique V_{cM}^j . Une vue caractéristique V_{cM}^j est considérée redondante s'il existe un ensemble de vues caractéristiques pour lesquelles l'union des vues représentées inclue Vr_M^j :

$$Vr_M^j \subset \bigcup Vr_M^k$$

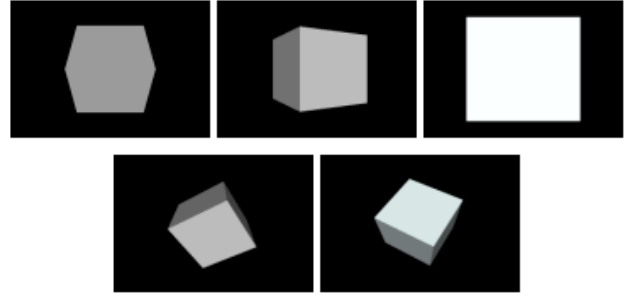


Figure 1 – Les vues caractéristiques d'un cube.

4 Approche probabiliste de l'indexation

Après les étapes décrites précédemment, chaque modèle de la collection D_b est représenté par un ensemble de vues caractéristiques $V_c = \{V_{c1}, V_{c2}, \dots, V_{cN}\}$, où N est le nombre de vues de l'ensemble. De plus, à chaque vue caractéristique correspond un nombre de vues qu'elle représente nommé V_r .

Considérant un modèle 3D requête Q , nous souhaitons trouver le modèle $M_i \in D_b$ qui est le plus proche du modèle requête Q . Ce modèle est celui qui a la plus grande probabilité $P(M_i/Q)$.

Sachant que chaque modèle est représenté par ses vues caractéristiques, la probabilité $P(M_i/Q)$ peut être écrite de la façon suivante :

$$P(M_i/Q) = \sum_{k=1}^K P(M_i|V_{cQ}^k)P(V_{cQ}^k/Q)$$

Avec K le nombre de vues caractéristiques du modèle Q . Soit H l'ensemble de toutes les hypothèses de correspondance possibles entre la vue requête V_{cQ}^k et le modèle M_i :

$$H = \{h_1^k \vee h_2^k \vee \dots \vee h_N^k\}$$

Une hypothèse h_p^k signifie que la vue p du modèle est la vue requête V_{cQ}^k . Le symbole \vee représente l'opérateur *ou logique*. Notons également que si une hypothèse h_p^k est vraie, alors toutes les autres hypothèses de l'ensemble sont fausses.

Par conséquent, $P(M_i|V_{cQ}^k)$ peut également s'écrire $P(M_i|H^k)$. Nous avons alors le développement suivant :

$$P(M_i|H^k) = \sum_{j=1}^N \sum_{p=1}^{N'} P(M_i, V_{cM_i}^j | h_p^k)$$

Sachant qu'il est impossible d'observer dans la même scène deux vues caractéristiques du même modèle 3D, nous avons :

$$P(M_i, V_{cM_i}^j | h_p^k) = 0 \text{ pour chaque } j \neq p$$

Alors :

$$P(M_i|H^k) = \sum_{j=1}^N P(M_i, V_{cM_i}^j | h_j^k)$$

Cette somme peut être réduite à la seule hypothèse vraie $P(M_i, Vc_{M_i}^j | H_j^k)$. Dans les faits, une vue caractéristique d'un modèle requête Q peut seulement correspondre à une et une seule vue caractéristique du modèle M_i . Nous choisissons la vue caractéristique avec la probabilité la plus grande :

$$P(M_i|Q) = \sum_{k=1}^K \max_j (P(M_i, Vc_{M_i}^j | h_j^k)) P(Vc_Q^k | Q)$$

Finalement :

$$P(M_i, Vc_{M_i}^j | h_j^k) = \sum_{k=1}^K P(Vc_Q^k | Q) \max_j \left(\frac{P(h_j^k | Vc_{M_i}^j, M_i) P(Vc_{M_i}^j | M_i) P(M_i)}{\sum_{i=1}^N \sum_{k=1}^K P(h_j^k | Vc_{M_i}^j, M_i) P(Vc_{M_i}^j | M_i) P(M_i)} \right)$$

avec $P(M_i)$ la probabilité d'observer le modèle M_i .

$$P(M_i) = \frac{e^{(-\alpha \cdot N(Vc_{M_i}^j) / N(Vc))}}{\sum_{j=1}^N e^{(-\alpha \cdot N(Vc_{M_i}^j) / N(Vc))}}$$

Où $N(Vc_{M_i}^j)$ est le nombre de vues caractéristiques du modèle M_i et $N(Vc)$ le nombre total de vues caractéristiques de l'ensemble des modèles de la collection D_b . Le coefficient α est un paramètre de pondération de l'effet de la probabilité $P(M_i)$. L'algorithme est conçu de telle sorte que, plus le nombre de vues caractéristiques pour un modèle est grand, plus ce dernier est complexe. En effet, les objets simples (comme le cube) peuvent être à la base de modèles 3D plus complexes.

D'autre part, nous avons l'égalité suivante :

$$P(Vc_{M_i}^j | M_i) = \frac{1 - \beta e^{(-\beta \cdot N(Vr_{M_i}^j) / N(Vr_{M_i}))}}{\sum_{j=1}^N (1 - \beta e^{(-\beta \cdot N(Vr_{M_i}^j) / N(Vr_{M_i}))})}$$

Où $N(Vr_{M_i}^j)$ est le nombre de vues représentées par la vue caractéristique j du modèle M_i , et $N(Vr_{M_i})$ est le nombre total de vues représentées par le modèle M_i . Le coefficient β est introduit ici pour limiter l'effet de la probabilité d'apparition de la vue. Dans notre implémentation, nous utilisons la valeur 1/100 pour α comme pour β . Ces valeurs ont donné les meilleurs résultats durant nos phases de test. Nous pouvons noter également que plus le nombre de vues représentées $N(Vr_{M_i}^j)$ est grand, plus la vue caractéristique $Vc_{M_i}^j$ est importante, et, par conséquent, mieux elle représente le modèle 3D M_i .

La valeur $P(h_j^k | Vc_{M_i}^j, M_i)$ est la probabilité que, sachant qu'on observe la vue caractéristique j du modèle M_i , cette vue est la vue requête k du modèle Q :

$$P(h_j^k | Vc_{M_i}^j, M_i) = \frac{e^{-D(h_q, h_{V_{M_i}^j})}}{\sum_{j=1}^N e^{-D(h_q, h_{V_{M_i}^j})}}$$

Avec $D(h_q, h_{V_{M_i}^j})$ la distance entre les descripteurs de la vue requête k du modèle Q et de la vue caractéristique $Vc_{M_i}^j$ du modèle 3D M_i .

Cette probabilité P est utilisée dans notre processus d'indexation.

Comme cela a déjà été souligné, les procédés de sélection des vues caractéristiques optimales et d'indexation sont indépendants du descripteur de vues utilisé. Le paragraphe suivant détaille celui que nous avons utilisé pour nos tests pour ses excellents résultats : les moments de Zernike 2D.

5 Moments de Zernike 2D

Les moments de Zernike sont des moments orthogonaux complexe dont la magnitude possède une invariance par rapport à la rotation. Les moments de Zernike sont définis à l'intérieur du cercle unitaire, et le polynôme radial orthogonal est défini comme suit :

$$\Re_{mn}(P) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \frac{(n-s)}{s! \binom{n+|m|}{2} - s! \binom{n-|m|}{2} - s!} P^{n-2s}$$

avec n et m des entiers strictement positifs respectant les contraintes suivantes : $n - |m|$ est pair et $|m| \leq n$. Le couple (n, m) de la fonction de Zernike basique $V_{nm}(\rho, \phi)$ définie sur le disque unitaire est :

$$V_{nm}(\rho, \phi) = \Re_{mn}(P) \exp(jm\phi), \rho \leq 1$$

Le moment de Zernike d'une image est défini par :

$$Z_{mn} = \frac{n+1}{\pi} \int \int_{unitdisk} V_{nm}^*(\rho, \phi) f(\rho, \phi)$$

avec V_{nm}^* un conjugué complexe de V_{nm} .

Les moments de Zernike ont les propriétés suivantes : la magnitude des moments de Zernike est invariante aux rotations ; ils sont robustes au bruit et aux légères variations des formes ; il n'y a pas de redondance d'information car leurs bases sont orthogonales.

Une image est mieux décrite par un petit ensemble de ses moments de Zernike que par tout autre type de moment comme les moments géométriques, les moments de Legendre, les moments rotationnels ou les moments complexes au sens des moindres-carrés. Les moments de Zernike d'ordre faible représentent la forme globale, alors que les ordres plus élevés représentent les détails de la forme caractérisée.

Ces moments sont uniquement invariants par rapport à la rotation. Pour obtenir une invariance par rapport à la mise à l'échelle et la translation, l'image subit tout d'abord une étape de normalisation. Les moments de Zernike, invariants à la rotation, sont ensuite extraits de cette image normalisée. L'invariance par rapport à la mise à l'échelle est obtenue en agrandissant puis en rétrécissant la forme de telle sorte que son moment d'ordre zéro m_{00} soit égal à une valeur β prédéfinie. L'invariance par rapport à la translation est enfin obtenue en déplacement l'origine vers le centre juste avant de calculer les moments. Pour résumer, une fonction d'image $f(x, y)$ est normalisée par rapport à la

mise à l'échelle et la translation en la transformant en une fonction $g(x, y)$ où

$$g(x, y) = f(x/a + \bar{x}, y/a + \bar{y})$$

avec (\bar{x}, \bar{y}) le centre de $f(x, y)$ et $a = \sqrt{\beta/m_{00}}$.

Pour obtenir l'invariance par rapport à la translation, il est nécessaire de connaître le centre de l'objet dans l'image. Dans notre système, il est facile de le calculer car nous utilisons des images binaires (la forme est composée de pixels noirs, et le fond de pixels blancs). Notons que, dans le cas d'une image binaire, m_{00} est le nombre total de pixel appartenant à la forme dans l'image.

Plus de détails sur les moments de Zenike sont disponibles notamment dans [15] et [16].

6 Expériences et résultats

Pour mesurer les performances de nos algorithmes, nous avons classifié notre collection de 760 modèles tridimensionnels – pièces mécaniques fournies par Renault – en 76 classes en fonction de ressemblances visuelles. La figure 2 montre des exemples de modèles de différentes classes.

Nous utilisons différents critères de mesure de performances pour évaluer notre méthode le plus objectivement possible : le critère *First Tier* (FT), le critère *Second Tier* (ST), et le critère *Nearest Neighbor* (NN), ainsi que des courbes rappel-précision moyennes sur l'ensemble de la collection.



Figure 2 – Exemples de modèles 3D de différentes classes.

Dans notre implémentation actuelle, les moments de Zernike sont extraits des vues à partir du second ordre. Nous extrayons jusqu'aux moments de Zernike du douzième ordre correspondants à 49 caractéristiques. Pour comparer deux moments de Zernike, la distance euclidienne est utilisée.

Nous testons également notre méthode avec et sans l'utilisation des probabilités décrites au paragraphe 4. De plus, nous comparons notre système avec trois descripteurs purement 3D classiques : les histogrammes de courbures de la surface maillée 3D [17], les distributions de distances entre points du maillage 3D [18], et la combinaison de ces deux descripteurs [1].

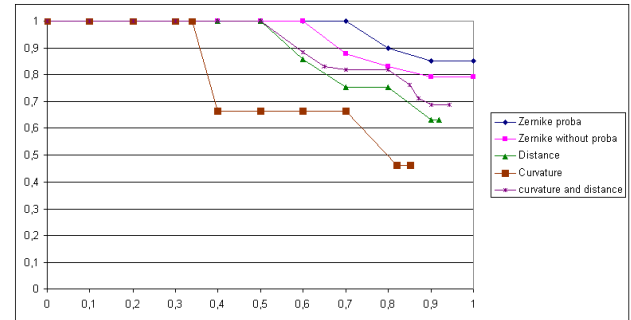


Figure 3 – Courbes rappel-précision pour les différents descripteurs.

La figure 3 montre les courbes rappel-précision moyennes – calculées sur les trente premiers résultats retournés – pour les deux méthodes que nous avons développées et pour les trois descripteurs classiques précédemment cités. Grâce à ces courbes, nous pouvons remarquer que l'histogramme de courbures donne les moins bons résultats. Ceci est principalement dû au fait que les pièces automobiles – pièces de CAO par excellence – sont en majorité constituées de plans et n'offrent pas conséquent que peu d'information de courbures. La distribution de distance donne des résultats de meilleure qualité, et la combinaison de ces deux descripteurs améliore sensiblement l'efficacité du système. Enfin, nous pouvons noter que notre méthode donne les meilleurs résultats, qu'elle soit utilisée avec ou sans l'aspect probabiliste exposé au paragraphe 4. Pour terminer cette comparaison, le tableau 1 montre les performances au sens ST, ST et NN des différentes méthodes.

| méthodes | performances | | |
|------------------------|--------------|-------|-------|
| | FT | ST | NN |
| 2D Zernike avec proba | 74.67 | 93.18 | 79.11 |
| 2D Zernike sans proba | 68.11 | 88.35 | 77.23 |
| courbures et distances | 59.62 | 82.13 | 69.11 |
| distances | 55.79 | 79.48 | 63.29 |
| courbures | 37.70 | 65.19 | 51.00 |

Tableau 1 – Performances au sens FT/ST/NN.

7 Conclusion

Cet article introduit une méthode efficace d'indexation de modèles tridimensionnels indépendamment de leurs représentations (maillage de polygones, surfaces implicites, représentations CSG, etc.) puisqu'elle s'appuie sur des comparaisons de vues caractéristiques bi-dimensionnelles.

Nous avons tout d'abord présenté un algorithme, indépendant du descripteur 2D utilisé, pour extraire les vues caractéristiques d'un modèle 3D. Cette méthode est indépendante du descripteur de vues 2D utilisé – même si elle donne ses meilleurs résultats avec les moments de Zernike – et permet de caractériser un modèle 3D avec un petit nombre de vues, appelées *vues caractéristiques*. Nous avons ensuite présenté une méthode probabiliste bayésienne pour l'indexation de ces modèles 3D. Les résultats présentés attestent des bonnes performances de notre méthode notamment face à quelques descripteurs classiques purement 3D.

Des tests sur des collections de très grande taille sont actuellement en cours et donnent des résultats prometteurs, notamment sur la base de modèles 3D de référence dans le domaine : la base du *Princeton Shape Retrieval and Analysis Group* [19].

Enfin, une interface web permet de tester notre moteur de recherche : <http://www-rech.enic.fr/3dretrieval>.

Remerciements

Ce travail est soutenu par le Ministère Français de la Recherche et le RNRT (Réseau National de Recherche en Télécommunications) dans le cadre du projet national SEMANTIC-3D (<http://www.semantic-3d.net>).

Références

- [1] J.P. Vandeborrel, V. Couillet, et M. Daoudi. A practical approach for 3D model indexing by combining local and global invariants. Dans *IEEE 3D Data Processing, Visualization and Transmission*, pages 644–647, June 2002.
- [2] H. Sundar, D. Silver, N. Gagvani, et S. Dickinson. Skeleton based shape matching and retrieval. Dans *IEEE proceedings of the Shape Modeling International*, pages 130–139, 2003.
- [3] M. Kazhdan, T. Funkhouser, et S. Rusinkiewicz. Rotation invariant spherical harmonic representation of 3d shape descriptors. Dans *Symposium on Geometric Processing*, 2003.
- [4] T. Tung et F. Schmitt. Indexation d'objets 3d par graphe de reeb multirésolution augmenté. Dans *Compression et Représentation des Signaux Audiovisuels (CORESA)*, mai 2004.
- [5] R. Ohbuchi, M. Nakazawa, et T. Takei. Retrieving 3D shapes based on their appearance. Dans *Proc. ACM SIGMM Workshop on Multimedia Information Retrieval*, pages 39–46, November 2003.
- [6] S. Abbasi et F. Mokhtarian. Affine-similar shape retrieval : Application to multi-view 3D object recognition. *IEEE Transactions on Image Processing*, 10 :131–139, 2001.
- [7] S. Mahmoudi et M. Daoudi. Une nouvelle méthode d'indexation 3D. Dans *13ème Congrès Francophone de Reconnaissance des Formes et Intelligence Artificielle*, volume 1, pages 19–27, janvier 2002.
- [8] C. Dorai et A.K. Jain. Shape spectrum based view grouping and matching of 3D free-form objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10 :1139–1146, 1997.
- [9] J.L. Chen et G. Stockman. 3D free-form object recognition using indexing by contour feature. *Computer Vision and Image Understanding*, 71(3) :334–355, 1998.
- [10] J.H. Yi et D.M. Chelberg. Model-based 3D object recognition using bayesian indexing. *Computer Vision and Image Understanding*, 69(1) :87–105, 1998.
- [11] D.Y. Chen et M. Ouhyoung. A 3D model alignment and retrieval system. Dans *Proceedings of International Computer Symposium, Workshop on Multimedia Technologies*, volume 2, pages 1436–1443, December 2002.
- [12] D.Y. Chen, X.P. Tian, Y.T. Shen, et M. Ouhyoung. On visual similarity based 3D model retrieval. Dans *Eurographics*, volume 22, 2003.
- [13] J.W.H. Tangelder et R.C. Veltkamp. A survey of content based 3d shape retrieval methods. Dans *Shape Modeling International*, pages 145–156, 2004.
- [14] T. Zaharia et F. Prêteux. Descripteurs de forme : Etude comparée des approches 3D et 2D/3D. Dans *Actes 14ème Congrès Francophone AFRIF-AFIA Reconnaissance des Formes et Intelligence Artificielle (RFIA'2004)*, January 2004.
- [15] W.Y. Kim et Y.S. Kim. A region-based shape descriptor using Zernike moments. *Signal Processing : Image Communication*, 16 :95–100, 2000.
- [16] A. Khotanzad et Y.H. Hong. Invariant image recognition by Zernike moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(5), May 90.
- [17] J. Koenderink et A.J. vanDoorn. Surface shape and curvatures scales. *Image and Vision Computing*, 10 :557–565, October 1992.
- [18] R. Osada, T. Funkhouser, B. Chazells, et D. Dobkin. Matching 3D models with shape distributions. Dans *Shape Modeling International*, pages 154–166, May 2001.
- [19] P. Shilane, P. Min, M. Kazhdan, et T. Funkhouser. The princeton shape benchmark. Dans *Shape Modeling International*, June 2004.