

Une approche Bayésienne pour l’indexation de modèles 3D basée sur des vues caractéristiques

T. Filali Ansary J.P. Vandeborre M. Daoudi

MIIRE INT / LIFL

ENIC Télécom Lille 1

Cité Scientifique, rue G. Marconi, 59658 Villeneuve d'Ascq cedex – France

{filali, vandeborre, daoudi}@enic.fr

Résumé

La gestion de grandes bases de données de modèles tridimensionnels (utilisés dans les applications de CAD/CAO, de visualisation, de jeux etc.) est un domaine de grande importance. En effet, pouvoir caractériser les modèles et les retrouver facilement sont des problèmes essentiels pour les concepteurs comme pour les utilisateurs. Dans ce cadre, deux grandes approches existent : la recherche par l'exemple d'un modèle tridimensionnel, et la recherche par une vue 2D. C'est ce dernier point que nous abordons dans cette communication. Nous nous intéressons donc à la détermination du nombre de vues caractéristiques d'un modèle tridimensionnel, et à l'indexation de ces modèles grâce à une approche probabiliste bayésienne sur les vues caractéristiques. Nous illustrons nos résultats à l'aide d'une base de modèles tridimensionnels fournie par Renault dans le cadre du projet SEMANTIC-3D.

Mots clefs

Modèles 3D, indexation 2D/3D, vues caractéristiques, probabilités..

1 Introduction

Les avancées récentes en imagerie tridimensionnelle (les outils d'acquisition, les modeleurs, les cartes graphiques, etc.) rendent possibles la création et le stockage de modèles tridimensionnels à l'échelle du Web et ceci pour de nombreuses applications telles que l'héritage culturel, le médical, l'industrie mécanique, les jeux, etc. Avec l'expansion d'Internet qui se poursuit, le développement et la multiplication des réseaux d'accès à haut débit, les objets tridimensionnels feront de plus en plus partie intégrante des bases de données distribuées liées aux services multimédia.

Dans cet article, nous nous intéressons à deux problèmes:

la sélection des vues caractéristiques d'un modèle tridimensionnel, et l'indexation d'une collection de modèles tridimensionnels à partir de ces vues.

Abbasi et Mokhtarian [1] proposent une méthode de sélection du nombre de vues optimales d'un objet tridimensionnel qui consiste à éliminer les vues qui sont similaires au sens d'une distance entre CSS (Curvature Scale Space) des contours de ces vues, et sélectionner le nombre minimal de vues en utilisant un algorithme d'optimisation. Le nombre de vues varie entre cinq et vingt-cinq selon le degré de complexité de l'objet.

Mahmoudi et Daoudi [2] proposent également d'utiliser les CSS des contours des vues extraites des modèles. Les CSS sont organisés sous forme d'une structure d'arbre dite *M-Tree*. Cette étape est importante pour l'accélération de la recherche et la gestion de la mémoire vu le nombre important de vues associées au modèle tridimensionnel. Les résultats obtenus sur une base de cent modèles montrent l'intérêt de cette approche.

Chen et Stockman [3] et Yi et al. [4] proposent une méthode basée sur l'approche bayésienne. Formellement, cela revient à calculer la probabilité *a posteriori* de reconnaître le modèle M_i quand on observe les descripteurs m_k , c'est-à-dire $P(M_i / m_k)$.

2 Sélection des vues caractéristiques

Dans ce paragraphe, nous présentons notre algorithme de sélection des vues caractéristiques d'un modèle tridimensionnel. Nous commençons par le calcul de l'ensemble des vues parmi lesquelles seront sélectionnées les vues caractéristiques grâce à un descripteur de courbures, puis nous vérifions deux critères essentiels à notre démarche. Le premier est que toute vue du modèle doit avoir une vue caractéristique qui la représente. Le seconde concerne le fait que chaque vue caractéristique représente au moins une vue qui n'est pas représentée par une autre vue caractéristique.

2.1 Calcul de l'ensemble des vues

Soit $Db = \{M_1, M_2, \dots, M_N\}$ une collection composée de N modèles tridimensionnels. On désire représenter chaque modèle de cette collection par un ensemble de vues. Ces vues doivent être équitablement réparties sur la surface de la sphère englobant l'objet. Dans l'implémentation courante le nombre de vue est de 80. Pour cela on a subdivisé un icosaèdre une fois en utilisant le « *Icosahedron Loop subdivisions scheme* ». Nous avons placé une caméra sur chacune des 80 faces en les orientons vers le centre de l'objet.

Nous appelons $V_M = \{V_M^1, V_M^2, \dots, V_M^v\}$ l'ensemble des vues du modèle tridimensionnel M , où v est le nombre total de vues.

2.2 Discrimination des vues caractéristiques

Parmi l'ensemble de vues ainsi obtenues, nous devons sélectionner celles qui caractérisent effectivement le modèle tridimensionnel en question en fonction d'un descripteur de ces vues. Le descripteur que nous avons choisi, même si notre méthode est **indépendante** de celui-ci, est l'histogramme de courbures.

Histogramme de courbures d'une vue

Nous utilisons l'histogramme de forme tridimensionnelle introduit par Koenderink et van Doorn [5] pour décrire une vue. Pour cela, la vue est tout d'abord transformée en niveaux de gris afin d'obtenir une surface tridimensionnelle dans laquelle chaque pixel de la vue est un point de la surface. La coordonnée Z de chaque point est égale au niveau de gris du pixel considéré.

L'index de courbure s'exprime par :

$$I_p = \frac{2}{\pi} \arctan \frac{k_1 + k_2}{k_1 - k_2} \quad \text{avec } k_1 \geq k_2$$

Cet index possède des propriétés intéressantes : il prend ses valeurs dans l'intervalle $[-1 ; +1]$, il est invariant aux transformations euclidiennes et aux homothéties, mais il n'est pas défini pour les surfaces planes. Cet index de forme tridimensionnel a notamment été utilisé pour l'indexation d'images fixes [6], d'images de profondeurs [7], et de modèles tridimensionnels [8].

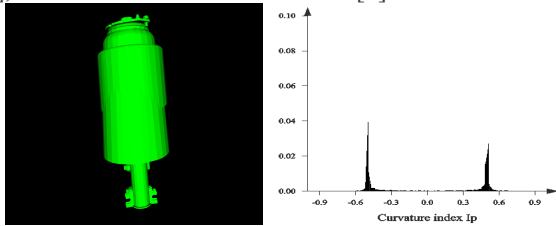


Figure 1 – exemple d'histogramme de courbures

Comparaison de vues

Puisqu'une vue est décrite par son histogramme de courbures, pour comparer deux d'entre elles, il suffit de comparer leurs histogrammes respectifs. Il existe pour cela de nombreuses méthodes : distance L_N de Minkowski, distance de Kolmogorov-Smirnov etc. Nous avons choisi d'utiliser la distance L_1 pour sa simplicité de mise en œuvre et les bons résultats qu'elle donne dans notre approche :

$$D_{L_1}(f_1, f_2) = \left(\int_{-\infty}^{+\infty} |f_1 - f_2| \right)$$

Sélection des vues caractéristiques

Il s'agit maintenant de réduire l'ensemble des vues d'un modèle M à un ensemble ne représentant que les vues les plus importantes, c'est à dire l'ensemble des vues caractéristiques V_{C_M} .

On dit qu'une vue V_M^k est une vue caractéristique du modèle M pour une distance ε , si la distance entre cette vue et toutes les autres vues caractéristiques de M est supérieure à ε . Ce qui peut s'écrire également :

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

Avec $D_{V_M^j, V_{C_M}^k}$ la distance entre les vues V_M^j et $V_{C_M}^k$, c'est-à-dire la distance entre les histogrammes de courbures de ces vues.

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 , nous avons adapté l'algorithme précédent en prenant en compte un intervalle de ces distances de 0 à 1 avec un certains pas. L'ensemble final des vues caractéristiques est alors l'union de tous les ensembles de vues caractéristiques pour chaque ε dans l'intervalle $]0..1[$.

2.3 Propriétés de l'algorithme de sélection des vues

Afin d'obtenir les vues caractéristiques réellement discriminantes pour un modèle tridimensionnel M , il nous faut vérifier deux critères :

- toutes les vues du modèle doivent être représentées au moins une fois dans l'ensemble des vues caractéristiques :

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

avec \mathfrak{R} une **application** associant à chaque élément de

V_M un élément de V_{C_M} ;

- les vues caractéristiques doivent être non-redondantes

Soit Vr_M^j l'ensemble des vues représentés par la vue caractéristique Vc_M^j . Vc_M^j est redondante s'il existe un ensemble de vie caractéristique tel que l'union de ses vues représentés inclus Vr_M^j .

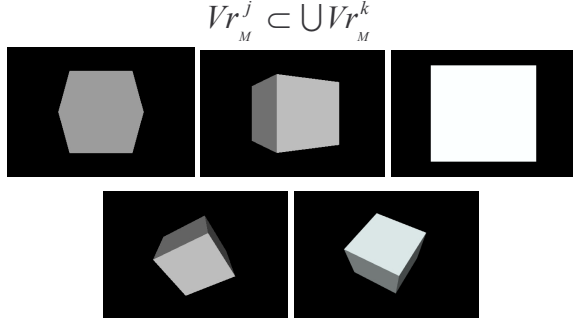


Figure 2 – les vues caractéristiques d'un cube

3 Approche probabiliste d'indexation 3D

Soit $Db = \{M_1, M_2, \dots, M_N\}$ une collection de N modèles tridimensionnels. Chaque modèle est représenté par un ensemble de vues caractéristiques $V = \{Vc_1, Vc_2, \dots, Vc_{\hat{v}}\}$, avec \hat{v} le nombre de vues caractéristiques. Et soit une image requête Q . Nous souhaitons retrouver le modèle $M_i \in Db$ dont l'une des vues caractéristiques est la plus proche de l'image requête. Ce modèle est celui qui a la plus grande probabilité $P(M_i, Vc_{M_i}^j / Q)$. C'est la probabilité que l'image Q

corresponde à la vue caractéristique j du modèle M_i . Soit H l'ensemble de toutes les hypothèses possibles pour une image requête $H = \{h_1 \vee h_2 \vee \dots \vee h_{\hat{v}}\}$ pour un modèle donné. Une hypothèse h_k correspond à ce que la vue k du modèle corresponde à l'image requête Q . Le signe \vee représente le *ou logique*. Notons également que si une hypothèse h_k est vraie, toutes les autres hypothèses sont fausses.

$P(M_i, Vc_{M_i}^j / Q)$ peut être exprimée par $P(M_i, Vc_{M_i}^j / H)$.

Le modèle le plus proche est celui qui contient la vue ayant la plus grande probabilité. En utilisant le théorème de Bayes, nous avons :

$$P(M_i, Vc_{M_i}^j / H) = \frac{P(H, Vc_{M_i}^j / M_i)P(M_i)}{P(H)}$$

Nous avons également :

$$P(H, Vc_{M_i}^j / M_i) = \sum_{k=1}^{\hat{v}} P(h_k, Vc_{M_i}^j / M_i)$$

La somme $\sum_{k=1}^{\hat{v}} P(h_k, Vc_{M_i}^j / M_i)$ peut être réduite à la seule hypothèse vraie $P(h_j, Vc_{M_i}^j / M_i)$.

En intégrant cette remarque, nous trouvons :

$$P(M_i, Vc_{M_i}^j / H) = \frac{P(h_j / Vc_{M_i}^j, M_i)P(Vc_{M_i}^j / M_i)P(M_i)}{\sum_{i=1}^N \sum_{j=1}^{\hat{v}} P(h_j / Vc_{M_i}^j, M_i)P(Vc_{M_i}^j / M_i)P(M_i)}$$

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

$$P(M_i) = \alpha e^{\left(-\alpha \frac{N(Vc_{M_i})}{N(Vc)}\right)}$$

Où $N(Vc_{M_i})$ est le nombre de vues caractéristiques du modèle M_i , et $N(Vc)$ est le nombre total de vues caractéristiques pour l'ensemble des modèles de la collection. α est un paramètre pour réduire l'effet de la probabilité. Intuitivement, une vue caractéristique peut sembler plus probable si elle représente un plus grand nombre de vues.

La probabilité d'une vue caractéristique sachant le modèle s'exprime par :

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

Où $N(Vr_{M_i}^j)$ est le nombre de vues représentées par la vue caractéristique j du modèle M_i , et $N(Vr)$ est le nombre total de vues représentées pour l'ensemble des modèles de la collection. Le coefficient β est introduit pour réduire l'effet de la probabilité des vues sur le résultat de la distance entre les histogrammes. Nous avons utilisé la valeur $\alpha = \beta = 1/100$ qui a donné les meilleurs résultats lors de nos expériences.

La valeur $P(h_j / Vc_{M_i}^j, M_i)$ est la probabilité que, sachant qu'on observe la vue caractéristique j du modèle M_i , cette même vue corresponde à la vue requête.

$$P(h_j / Vc_{M_i}^j, M_i) = e^{(-D_{h_q, h_{Vc_{M_i}^j}}^j)}$$

Avec $D_{h_q, h_{Vc_{M_i}^j}}$ la distance de Minkowski entre les

histogrammes de courbures de Q et celui de la vue caractéristique $Vc_{M_i}^j$ du modèle tridimensionnel M_i .

4 Expériences et résultats

Dans ce paragraphe, nous exposons les résultats obtenus à l'aide d'une collection de modèles tridimensionnels fournie par Renault, partenaire du projet SEMANTIC-3D. Nous commençons donc pas une présentation de la collection de modèles avec lesquels nous avons effectué nos tests, puis nous exposons nos résultats sur cette collection en terme de sélection de vues caractéristiques, ainsi qu'en terme de performances de notre approche d'indexation probabiliste.

4.1 Présentation de la base

La base de modèles tridimensionnels du projet SEMANTIC-3D fournie par Renault. Pour effectuer nos tests, nous avons prélevé une partie de ces modèles pour former une collection de 132 modèles. Ces modèles ont été classés en 14 classes différentes selon le jugement de deux personnes adultes.

4.2 Performances de l'indexation

Pour mesurer objectivement les performances de notre approche nous avons utilisé différentes mesures de performances: les méthodes First Tier (FT), Second Tier (ST), la méthode Nearest Neighbor (NN), ainsi que les courbes de rappel-précision.

Nous avons testé les performances de notre méthode avec et sans l'utilisation de l'approche probabiliste. La Table 1 montre les performances en terme de FT, ST et NN de notre méthode. La figure 3 montre les courbes de rappel-précision. Nous remarquons l'apport de l'approche probabiliste qui améliore la qualité des résultats. Pour produire ces résultats, nous avons sélectionné un modèle aléatoirement de chaque classe. Nous avons pris 5 vues aléatoires de chaque modèle. Ces vues ont été utilisées comme requêtes pour notre système. Les résultats sont la moyenne de 70 requêtes.

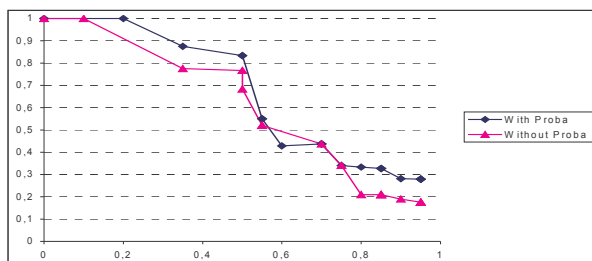


Figure 3 – les courbes de rappel-précision

	FT	ST	NN
Sans proba.	28,8	40,65	51,46
Avec proba.	27,84	37,26	47,18

Tableau 1– performance de l'indexation

Les résultats obtenus montrent bien que, dans la grande majorité des cas, notre méthode d'indexation probabiliste donne de meilleurs résultats que la simple comparaison des histogrammes de courbures des vues deux à deux.

5 Conclusion

Dans ce article, nous nous sommes intéressés à l'indexation de modèles tridimensionnels, et plus particulièrement à l'indexation à partir de vues de ces modèles. Nous avons notamment développé deux propositions pour résoudre les problèmes qui se posent à l'indexation de ces modèles par leurs vues caractéristiques. Nous avons proposé une méthode de sélection "optimale" des vues caractéristiques qui consiste à discriminer les

vues réellement déterminantes parmi un ensemble de vues obtenues de manière systématique. Cette sélection se fait grâce à un descripteur de forme, habituellement utilisé pour les données purement tridimensionnelles, que nous avons adapté pour les images.

Nous avons également proposé un algorithme d'indexation qui utilise les vues caractéristiques précédemment sélectionnées, et s'appuie sur une approche probabiliste bayésienne. Elle prend, en outre, en considération le fait de la non-équiprobabilité d'apparition des modèles et des vues dans la collection de modèles tridimensionnels.

Notre algorithme de sélection de vues caractéristiques nous permet de caractériser un modèle tridimensionnel avec un petit nombre de vues. Notre approche d'indexation probabiliste donne également de bons résultats.

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] S. Abbasi, and F. Mokhtarian, "Affine-Similar Shape Retrieval: Application to Multi-View 3-D Object Recognition", IEEE Transactions on Image Processing, Volume 10, Number 1, pp. 131-139, 2001.
- [2] S. Mahmoudi et M. Daoudi, "Une nouvelle méthode d'indexation 3D", 13ème Congrès Francophone de Reconnaissance des Formes et Intelligence Artificielle (RFIA2002), volume 1, pp. 19-27, Angers, France, 8-9 janvier 2002.
- [3] J-L. Chen and G. Stockman, "3D Free-Form Object Recognition Using Indexing by Contour Feature", Computer Vision and Image Understanding, Volume 71, Number 3, pp.334-355, 1998.
- [4] J.H. Yi and D.M. Chelberg, "Model-Based 3D Object Recognition Using Bayesian Indexing", in Computer Vision and Image Understanding Volume 69, Number 1, January, pp.87-105,1998.
- [5] J.J. Koenderink and A.J. van Doorn, "Surface shape and curvature scales", Image and Vision Computing, Volume 10, Number 8, pp. 557-565, 1992.
- [6] C. Nastar, "The image Shape Spectrum for Image Retrieval", rapport technique INRIA numéro 3206, 1997.
- [7] C. Dorai and A.K. Jain, "Shape Spectrum Based View Grouping and Matching of 3D Free-Form Objects", IEEE Transactions on PAMI, Volume 19, Number 10, pp.1139-1146, 1997.
- [8] J.P. Vandeborre, V. Couillet and M. Daoudi, "A Practical Approach for 3D Model Indexing by combining Local and Global Invariants", 3D Data Processing Visualization and Transmission (3DPVT), pp. 644-647, Padova, Italy, june 2002.