

Une collection de modèles 3D avec vérité-terrain pour l'évaluation objective des algorithmes de segmentation

Halim Benhabiles¹, Jean-Philippe Vandeborre^{1,2}, Guillaume Lavoué³ et Mohamed Daoudi^{1,2}

¹ LIFL (UMR USTL/CNRS 8022), Université de Lille, France

² Institut TELECOM ; TELECOM Lille 1, France

³ LIRIS UMR 5205 CNRS, INSA-Lyon, F-69621, France

prenom.nom@lifl.fr

glavoue@liris.cnrs.fr

Résumé

Dans cet article nous présentons une nouvelle méthode d'évaluation des algorithmes de segmentation de maillages 3D basée sur une vérité-terrain. Cette dernière est représentée par une collection de modèles 3D de différentes classes (animaux, humanoïdes, etc.) accompagnés de leurs segmentations manuelles effectuées par des opérateurs humains. Nous définissons également une mesure qui permet de quantifier la consistance entre la vérité-terrain et les segmentations obtenues par des algorithmes automatiques afin d'évaluer leur qualité d'une manière objective. Cette mesure consiste à calculer l'écart géométrique entre deux segmentations. Notre approche est illustrée par l'évaluation de deux algorithmes récents de segmentation automatique.

Mots clefs

Maillage 3D, Segmentation, Évaluation.

1 Introduction

La segmentation de maillage consiste à subdiviser une surface polygonale en différentes parties présentant des caractéristiques identiques, soit d'un point de vue strictement géométrique, soit d'un point de vue plus sémantique. Cette opération est devenue un outil de prétraitement nécessaire à de nombreuses applications telles que la reconnaissance de forme [1], la compression [1], la déformation [2], le plaquage de texture [3], etc.

Selon Attene et al. [4], les techniques de segmentation de maillage peuvent être classifiées en deux catégories : les méthodes géométriques et les méthodes sémantiques. Dans le cas des méthodes géométriques, les algorithmes se basent sur l'information géométrique de bas niveau, telle que la courbure [5] pour définir des segments (c'est-à-dire des régions) qui sont géométriquement homogènes. Par contre, les algorithmes dits sémantiques visent à distinguer les régions d'un point de vue perceptuel, en suivant des notions de haut niveau définies dans la théorie de la perception humaine [6].

Ce type d'approche est particulièrement utile pour les applications de déformation, et d'animation d'objets où la décomposition doit être significative.

Il existe dans la littérature un grand nombre d'algorithmes de segmentation permettant le traitement des modèles 3D. Cependant, jusqu'à maintenant, aucun outil automatique n'a été proposé pour leur évaluation sachant que cette tâche est nécessaire non seulement aux chercheurs pour comparer un nouvel algorithme à ceux existants déjà, mais aussi aux utilisateurs afin qu'ils puissent choisir un algorithme et régler ses paramètres en fonction du problème à résoudre.

Attene et al. [4] ont fourni un ensemble de critères à travers lesquels ces méthodes peuvent être évaluées ; on peut citer notamment : la sensibilité à la pose, les paramètres de contrôle, et l'extraction de segment correct.

- La sensibilité à la pose : pour certaines applications telles que l'extraction de squelette, la métamorphose et la reconnaissance de forme, il est important que les modèles des objets similaires sous différentes poses soient segmentés de la même manière.
- Les paramètres de contrôle : le type et le nombre de paramètres de contrôle offrent une idée sur l'interaction nécessaire pour produire une bonne segmentation. Plus le nombre de paramètres à régler est important plus l'outil de segmentation est complexe.
- L'extraction de segment correct : cela consiste à déterminer d'une manière visuelle si le segment d'un modèle donné est correct. En réalité cela dépend de l'application, des connaissances de l'environnement et du point de vue de l'observateur.

Les critères précédents sont très importants mais ne permettent pas de valider la pertinence des régions extraites et de mesurer leur rapprochement de la perception visuelle humaine par exemple. Ces critères ne sont pas donc suffisants pour quantifier la qualité de la segmentation. C'est pourquoi nous proposons et implémentons une nouvelle

méthode automatique d'évaluation de la segmentation de maillages tridimensionnels basée sur une vérité-terrain. Cette vérité-terrain est composée d'une ou plusieurs segmentations manuelles pour différents modèles 3D. Nous proposons également une mesure de distance entre segmentations qui évalue les méthodes de segmentation existantes et à venir en les confrontant à la vérité-terrain. Ce travail s'inscrit dans le projet ANR MADRAS¹ (3D models and dynamic models representation and segmentation).

Cet article est organisé comme suit. La section 2 présente un état de l'art de l'évaluation de la segmentation. Dans la section 3, nous décrivons en détail la construction de notre collection de modèles tridimensionnels segmentés. Dans la section 4, nous définissons une mesure permettant le calcul de l'écart entre la segmentation de référence (vérité-terrain) et celle obtenue par un algorithme. La section 5, met en valeur la vérité-terrain en l'utilisant pour évaluer la performance des algorithmes de segmentation de Tierny et al. [7], et de Lavoué et al. [8].

2 État de l'art de l'évaluation de la segmentation

Contrairement à l'évaluation de la segmentation de maillages polygonaux 3D, celle des images 2D présente des travaux très avancés.

2.1 Évaluation de la segmentation d'images 2D

Zhang [9] offre une étude sur les différentes méthodes d'évaluation de la segmentation d'images 2D. Selon lui, les différentes méthodes peuvent être classées en trois groupes :

- Méthodes analytiques ;
- Méthodes empiriques avec qualité (*empirical goodness*) ;
- Méthodes empiriques avec écart (*empirical discrepancy*).

Méthodes analytiques

Les méthodes analytiques traitent directement les algorithmes de segmentation en considérant leur utilité, leur complexité, les critères sur lesquels ils sont basés, etc. L'utilisation des méthodes analytiques pour évaluer les algorithmes de segmentation évite l'implémentation concrète de ces derniers. Cependant les propriétés de ces algorithmes ne peuvent être obtenues par une simple étude analytique.

Méthodes empiriques avec qualité (sans vérité terrain)

Ces méthodes évaluent la performance des algorithmes en jugeant la qualité des images segmentées. Pour effectuer cette tâche, un ensemble de mesures de qualité doit être défini. Ces mesures sont établies à l'aide de l'intuition humaine en fonction des conditions qui doivent être satisfaites par une segmentation idéale. Ce type de méthodes fait la distinction entre les différents algorithmes de segmentation, en déterminant simplement la qualité de l'image segmentée sans avoir acquis a priori une connaissance sur la segmentation correcte.

Méthodes empiriques avec écart (avec vérité terrain)

Dans ce type de méthode, un ensemble d'images références, qui représente la segmentation idéale, est tout d'abord créé. Cet ensemble d'images, pouvant être segmentées manuellement par des experts du domaine, constitue ce qu'on appelle une vérité-terrain. L'objectif est de mesurer l'écart entre la segmentation référence et celle obtenue par un algorithme à évaluer. Un grand écart implique une grande erreur dans l'image segmentée par rapport à l'image référence, et ceci indique que l'algorithme de segmentation utilisé est peu performant.

Martin et al. [10] ont proposé une méthode d'évaluation de ce type. Ils ont créé une base de données publique contenant des images de références accompagnées de leurs segmentations obtenues par des opérateurs humains. Ils ont aussi défini une mesure basée sur le calcul d'erreur de raffinement d'un pixel dans un segment par rapport à un autre. Cette mesure quantifie la consistance entre les segmentations et évalue ces dernières d'une manière objective.

2.2 Évaluation de la segmentation de maillages 3D

Très récemment, Delest et al. [11] ont montré l'intérêt d'une vérité-terrain pour la segmentation de maillages 3D. Ils proposent également quelques réflexions pour sa mise en œuvre. Les critères proposés par Attene et al. [4] concernant l'évaluation des algorithmes de segmentation de maillages 3D (voir Introduction), peuvent être classés dans le groupe des méthodes empiriques avec qualité.

3 Collection de modèles tridimensionnels avec leur vérité-terrain

Notre collection contient vingt modèles polygonaux tridimensionnels, regroupés en cinq classes (*animal, human, furniture, misc, et hand*), contenant chacune d'entre elles quatre modèles. La figure 1 présente les modèles de chaque classe.

Cette variété de modèles s'avère utile pour étudier le comportement des différents algorithmes de segmentation et pour ne pas privilégier certains algorithmes par rapport à d'autres, ce en évitant de se limiter à une seule catégorie

1. <http://www-rech.telecom-lille1.eu/madras/>

de modèles.

Ces modèles proviennent des bases de donnée GAMMA² de l'INRIA et Princeton [12] publiques contenant des milliers de modèles tridimensionnels offerts gratuitement dans le but d'encourager la recherche et le développement des différents algorithmes de classification, de clustering, de reconnaissance de forme, etc.

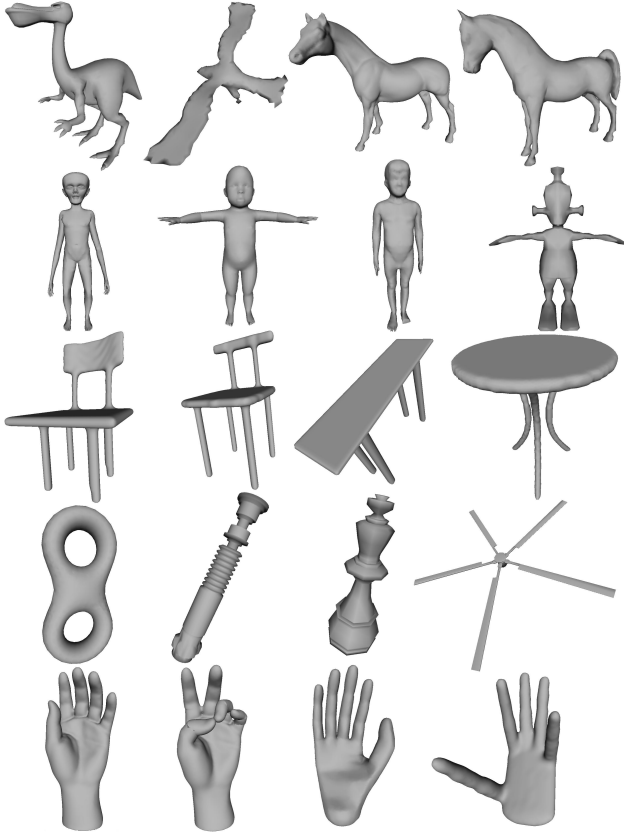


Figure 1 – Modèles de la collection.

3.1 Outil de segmentation manuelle

Afin de faciliter l'obtention de segmentations manuelles auprès d'un grand nombre de personnes, nous utilisons l'application MeshLab³ qui permet de segmenter un modèle 3D à l'aide de couleurs. Chaque segment est distingué des autres par une couleur différente.

L'application MeshLab a pour objectif de faciliter le traitement des modèles tridimensionnels en fournissant un ensemble d'outils dont l'interface de coloriage d'objet 3D, des fonctions de base telles que le zoom, la rotation, le déplacement, le type d'affichage et le rendu permettant à l'utilisateur de définir la zone de coloriage du maillage. Il permet aussi l'import et l'export de plusieurs formats de fichiers.

2. <http://www-c.inria.fr/gamma/disclaimer.php>

3. <http://meshlab.sourceforge.net/>

3.2 Protocole de segmentation des modèles

Pour segmenter les maillages de la collection, nous avons fait appel à des volontaires. Ils ont librement segmenté les modèles et aucune condition n'a été imposée sur la manière avec laquelle ils ont effectué cette tâche. Les modèles fournis à un volontaire pour la segmentation étaient choisis aléatoirement tout en prenant en compte ceux qui étaient déjà segmentés. La figure 2 présente quelques modèles segmentés manuellement de la vérité-terrain.

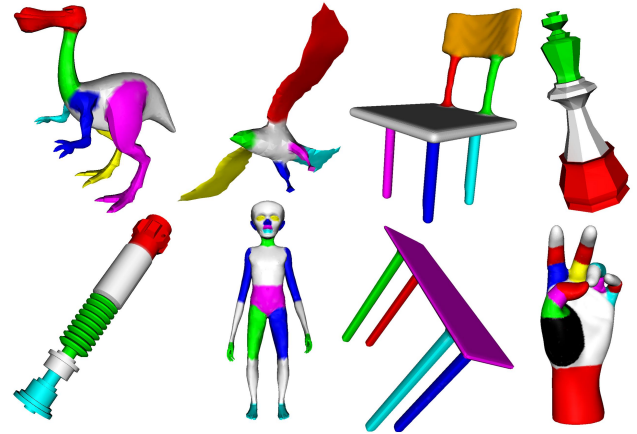


Figure 2 – Vérité-terrain de quelques modèles de la collection.

Notre collection est en constante évolution en termes de nombre de segmentations par modèle. Il est important d'avoir plusieurs segmentations par maillage, car deux sujets n'ont pas nécessairement le même point de vue sur la segmentation d'un modèle donné. Ceci est expliqué par l'absence de règles fixées aux opérateurs, ce qui peut entraîner une différence de granularité au niveau de la façon dont ils vont découper un modèle en sous-objets. La figure 3 montre l'exemple d'un modèle 3D découpé par deux personnes différentes. On peut remarquer qu'il y a deux segmentations différentes, et que la première (cheval de gauche) est effectuée d'une manière plus fine que la deuxième (cheval de droite).

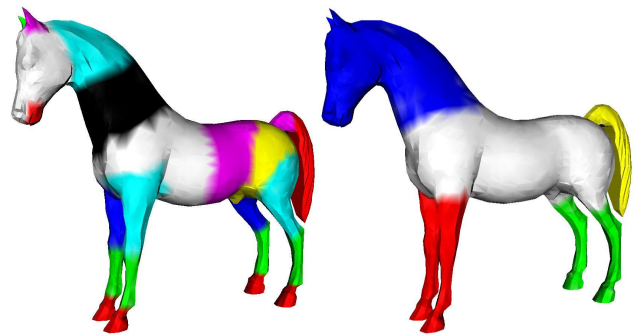


Figure 3 – Segmentations manuelles d'un cheval par deux personnes différentes.

Les résultats présentés dans cet article ont été générés en utilisant notre première version de vérité-terrain contenant 20 segmentations, une pour chaque modèle, effectuées par 11 personnes.

4 Mesure de distance entre deux segmentations d'un maillage 3D

Nous proposons ici une mesure de distance entre deux segmentations d'un maillage 3D, basée sur l'écart géométrique entre les régions générées. Cette mesure nous permettra d'évaluer, d'une manière objective, la qualité de la segmentation d'un maillage obtenue par des algorithmes automatiques en mesurant sa distance par rapport à la vérité-terrain.

Aspert et al. [13] ont proposé une méthode qui estime l'erreur entre deux maillages 3D, basée sur une approximation de la distance de Hausdorff.

Définition la distance de Hausdorff entre deux ensembles de points est la distance maximale d'un point d'un sous-ensemble au point le plus proche de l'autre sous-ensemble.

Si on considère que A et B sont deux ensembles de points, alors la distance de Hausdorff est définie comme suit :

$$dh(A, B) = \max(d(A, B), d(B, A)) \quad (1)$$

car cette distance n'est pas symétrique.

$$d(A, B) = \max_{a \in A} d(a, B) \quad (2)$$

où $d(a, B)$ est la distance entre un point a et l'ensemble de points $B = \{b_1, \dots, b_n\}$ est définie par :

$$d(a, B) = \min_{b \in B} d(a, b) \quad (3)$$

La distance entre deux points a et b étant la distance Euclidienne : $d(a, b) = \|a - b\|$.

De manière similaire, Dubuisson et Jain [14] ont proposé pour l'imagerie 2D, une version modifiée de la distance de Hausdorff pour estimer l'erreur entre deux images, et montré qu'elle donne de meilleurs résultats que la distance de Hausdorff. Pour obtenir la version modifiée de Hausdorff, il suffit de remplacer l'équation 2 par l'équation suivante :

$$d(A, B) = \frac{1}{N_a} \sum_{a \in A} d(a, B) \quad (4)$$

Nous avons donc utilisé les deux distances et nous avons comparé les résultats obtenus.

Application des deux distances entre deux segmentations d'un maillage 3D

Soit M un modèle tridimensionnel, $S = \{s_1, \dots, s_k\}$ sa segmentation manuelle (segmentation de référence), et

$S' = \{s'_1, \dots, s'_l\}$ sa segmentation automatique.

– **Distance de Hausdorff entre S et S'** : il s'agit de la distance entre les deux segments les plus éloignés.

$$dh(S, S') = \max(d(S, S'), d(S', S)) \quad (5)$$

$$d(S, S') = \max_{s \in S} d(s, S') \quad (6)$$

où $d(s, S')$ est la distance entre un segment s et l'ensemble de segments S' définie par :

$$d(s, S') = \min_{s' \in S'} d(s, s') \quad (7)$$

La distance entre deux segments (s, s') , n'est que la distance entre deux ensembles de points (voir équation 2).

– **Distance de Hausdorff modifiée entre S et S'** : il s'agit de la moyenne des distances entre les segments appariés deux à deux, de telle sorte que deux segments s et s' ne peuvent être regroupés que si leur distance est inférieure à celle de l'un des deux avec le reste des autres segments.

Cette distance est obtenue en remplaçant l'équation 6 de la distance de Hausdorff entre deux segmentations par :

$$d(S, S') = \frac{1}{N_S} \sum_{s \in S} d(s, S') \quad (8)$$

Dans le cas de Hausdorff modifiée, la distance entre deux segments $d(s, s')$ est calculée en appliquant l'équation 4.

Remarquons que les deux distances n'imposent aucune condition sur le nombre de segments entre S et S' d'un modèle M . Dans le cas où les deux segmentations d'un modèle sont identiques ($S=S'$), les deux distances, que ce soit Hausdorff ou Hausdorff modifiée, sont nulles, puisqu'elles sont toutes les deux basées sur la distance Euclidienne. Cette dernière étant nulle entre deux points identiques.

5 Résultats et tests d'évaluation

Dans cette section, nous utilisons la collection vérité-terrain et les deux distances définies précédemment pour évaluer la performance des deux algorithmes de segmentation, à savoir, celui de Tierny et al. [7] et celui de Lavoué et al. [8]. L'algorithme de Tierny et al. [7] est basé sur l'extraction d'un squelette qui est utilisé pour délimiter le noyau de l'objet et identifier les surfaces de jonction. Ceci donne comme résultat une segmentation grossière qui est raffinée à l'aide d'un schéma hiérarchique basé sur la topologie du modèle pour donner une segmentation plus fine (voir figure 4). Nous avons donc comparé notre vérité-terrain avec les deux niveaux de segmentation.

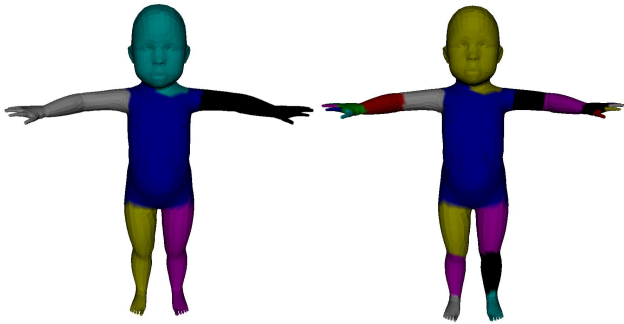


Figure 4 – *Modèle bébé avec deux niveaux de segmentation ; Tierny et al. [7].*

Lavoué et al. [8] utilisent un algorithme de croissance de région basé sur l’analyse des tenseurs de courbure. L’algorithme regroupe les sommets du maillage en régions de courbures similaires.

5.1 Présentation des tests

Nous présentons sous forme d’histogrammes (figure 5), l’ensemble des tests d’exécution de notre algorithme pour le calcul de l’écart entre les segmentations vérité-terrain et celles obtenues par deux méthodes automatiques.

5.2 Analyse des résultats

Comme nous pouvons le constater, la distance, que ce soit Hausdorff ou Hausdorff modifiée, obtenue avec la segmentation de Tierny et al. [7] de niveau 1 (grossière) est inférieure à celle de niveau 2 (fine) et à celle obtenue par l’algorithme de Lavoué et al [8] dans la plupart des modèles, ce qui signifie à première vue que la segmentation de niveau 1 est plus proche de la vérité-terrain.

Sachant qu’on a une seule segmentation de référence par modèle, on ne peut pas conclure que la segmentation de niveau 1 est la meilleure. Néanmoins, on peut dire que les personnes qui ont participé à la réalisation de cette première version de vérité-terrain ont segmenté les modèles d’une manière plutôt grossière.

On remarque également une contradiction entre les résultats obtenus par la distance de Hausdorff, et sa version modifiée. Dans le modèle *Alien* par exemple, la distance de Hausdorff ordonne les trois algorithmes comme suit :

Lavoué et al. [8], suivi de Tierny et al. [7] (segmentation grossière), suivi de Tierny et al. [7] (segmentation fine). Alors que Hausdorff modifiée, les ordonne dans l’ordre inverse.

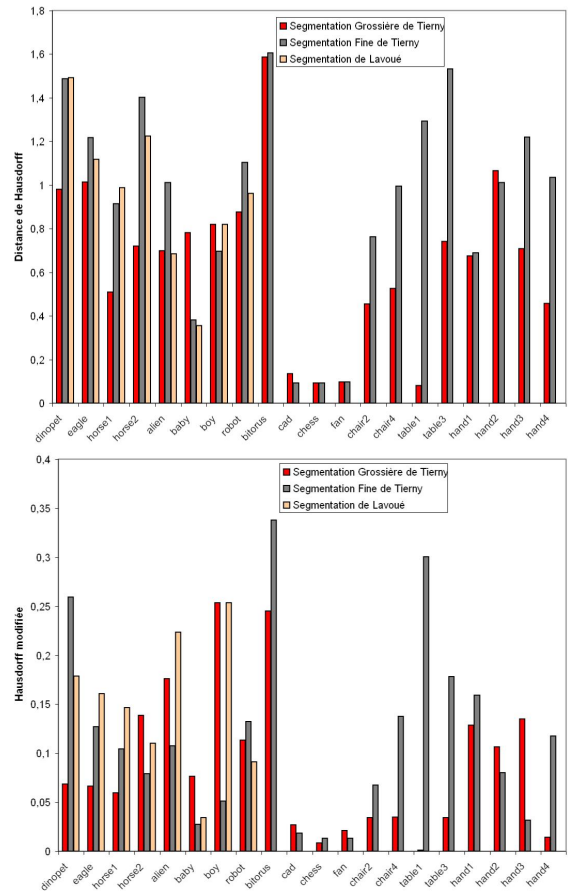


Figure 5 – *Résultats des distances entre la segmentation manuelle et automatique.*

Afin d’expliquer ces contradictions nous présentons le modèle *Alien* et nous le comparons avec sa vérité-terrain d’une manière visuelle (figure 6).

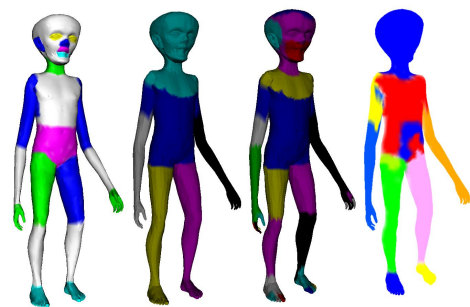


Figure 6 – *Vérité-terrain et segmentations automatiques obtenues par l’algorithme de Tierny et al. [7](niveaux 1 et 2) et par l’algorithme de Lavoué et al. [8], présentées de gauche à droite.*

D’un point de vue sémantique les trois segmentations sont correctes puisqu’elles segmentent le modèle en parties significatives. Néanmoins, il est clair que si on prend comme critères de comparaison le nombre et les détails

des différentes parties significatives du modèle ainsi que la continuité des frontières de chaque segment, la segmentation de niveau 2 sera plus proche de la vérité-terrain que les deux autres, ce qui signifie que la distance de Hausdorff modifiée est celle qui donne les résultats les plus justes.

Ceci est expliqué par le fait que, théoriquement, la distance de Hausdorff entre deux segmentations d'un modèle est égale à la distance entre les deux segments les plus éloignés. Elle ne prend pas en considération le reste des segments bien que ces derniers puissent être proches les uns des autres. A l'inverse, la distance de Hausdorff modifiée calcule une moyenne des différents segments et ne néglige donc aucun de ces derniers.

Cela confirme les tests de Dubuisson et Jain [14] (voir section 4) obtenus sur le calcul d'erreur entre deux images 2D, et nous pouvons déduire à notre tour que la distance de Hausdorff modifiée appliquée aux maillages est également mieux adaptée au calcul de l'écart entre les segments de références et les segments automatiques que la distance de Hausdorff classique. Cependant, cette distance ne prend pas en compte le raffinement. Idéalement, on souhaiterait que la distance entre deux segmentations, dont l'une est un raffinement de l'autre, soit nulle. Comme le montre la figure 7, la segmentation de Tierny et al. [7] de niveau 2 n'est qu'un raffinement de la segmentation de niveau 1, alors que leurs distances par rapport à la vérité-terrain ne sont pas les mêmes.

6 Conclusion

Dans cet article, nous avons proposé un outil automatique basé sur la notion de vérité-terrain permettant l'évaluation des différents algorithmes de segmentation. Nous avons créé une collection de modèles segmentés manuellement ; les différents modèles et leur vérité-terrain sont disponibles en ligne⁴. Nous avons également utilisé deux distances permettant la mesure de l'écart entre segmentations. L'objectif principal de notre méthode est d'évaluer les algorithmes de segmentation d'une manière objective, ce en quantifiant la consistance entre nos segmentations de référence et les segmentations obtenues par ces algorithmes. La pertinence de cette approche a été démontrée à l'issue de l'évaluation de deux algorithmes récents.

En vue de travaux futurs, nous souhaitons accroître notre collection en terme de nombre de modèles et de vérité-terrain, et proposer d'autres distances qui prennent en compte le raffinement entre les segmentations d'un maillage 3D.

Remerciements

Ce travail a bénéficié d'une aide de l'ANR (Agence Nationale de la Recherche) à travers le projet MADRAS (ANR-07-MDCO-015).

4. <http://www-rech.telecom-lille1.eu/madras/>

Références

- [1] E. Zeckerberger, A. Tal, et S. Shlafman. Polyhedral surface decomposition with applications. *Computers and Graphics*, 26(5) :733–743, 2002.
- [2] S. Katz et A. Tal. Hierarchical mesh decomposition using fuzzy clustering and cuts. *ACM Transactions on Graphics (SIGGRAPH)*, 22(3) :954–961, 2003.
- [3] A. Sheffer, E. Praun, et K. Rose. Mesh parameterization methods and their applications. *Foundations and Trends in Computer Graphics and Vision (FTCGV)*, 2(2) :64, 2007.
- [4] M. Attene, S. Katz, M. Mortara, G. Patané, M. Spagnuolo, et A. Tal. Mesh segmentation, a comparative study. *IEEE International Conference on Shape Modeling and Applications*, pages 7–7, 2006.
- [5] G. Lavoué, F. Dupont, et A. Baskurt. A new cad mesh segmentation method, based on curvature tensor analysis. *Computer Aided Design*, 37(10) :975–987, 2005.
- [6] I. Biederman. Recognition-by-components : A theory of human image understanding. *Psychological Review*, 94 :115–147, 1987.
- [7] J. Tierny, J. P. Vandeborre, et M. Daoudi. Topology driven 3D mesh hierarchical segmentation. Dans *Shape Modeling International, Lyon, France, 2007*.
- [8] G. Lavoué, F. Dupont, et A. Baskurt. A new cad mesh segmentation method, based on curvature tensor analysis. *Computer Aided Design*, 37(10) :975–987, 2005.
- [9] Y.J. Zhang. A survey on evaluation methods for image segmentation. *Pattern Recognition*, 29(8) :1335–1346, 1996.
- [10] D. Martin, C. Fowlkes, D. tal, et J. Malik. A database of human segmented natural images and its application to evaluating algorithms and measuring ecological statistics. *International Conference On Computer Vision*, 2 :416–423, 2001.
- [11] S. Delest, R. Boné, et H. Cardot. Réflexions sur la mise en place d'une vérité terrain pour la segmentation de maillages polygonaux 3d en parties significatives. Rapport technique, GTMG : Groupe de Travail en Modélisation Géométrique, 2008.
- [12] P. Shilane, P. Min, P. Kazhdan, et M. Funkhouser. The princeton shape benchmark. *Shape modelling international, IEEE Computer*, pages 167–178, 2004.
- [13] N. Aspert, D. Santa-Cruz, et T. Ebrahimi. Measuring errors between surfaces using the hausdorff distance. *Multimedia and Expo. ICME '02. Proceedings*, 1 :705–708, 2002.
- [14] M.P. Dubuisson et A.K. Jain. A modified hausdorff distance for object matching. *Proceedings of the 12th IAPR International Conference on Pattern Recognition*, 1 :566–568, 1994.