HAL
open science

3D model retrieval based on depth line descriptor

Mohamed Chaouch, Anne Verroust-Blondet

To cite this version:

Mohamed Chaouch, Anne Verroust-Blondet. 3D model retrieval based on depth line descriptor. ICME 2007 - IEEE International Conference on Multimedia and Expo, Jul 2007, Beijing, China. pp.599 - 602, 10.1109/ICME.2007.4284721 . hal-00802588

HAL Id: hal-00802588

https://inria.hal.science/hal-00802588v1

Submitted on 20 Mar 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

3D MODEL RETRIEVAL BASED ON DEPTH LINE DESCRIPTOR

Mohamed Chaouch and Anne Verroust-Blondet

INRIA Rocquencourt Domaine de Voluceau, B.P. 105 78153 Le Chesnay Cedex, FRANCE
Email: {mohamed.chaouch, anne.verroust}@inria.fr.

ABSTRACT

In this paper, we propose a novel 2D/3D approach for 3D model matching and retrieving. Each model is represented by a set of depth lines which will be afterward transformed into sequences. The depth sequence information provides a more accurate description of 3D shape boundaries than using other 2D shape descriptors. Retrieval is performed when dynamic programming distance (DPD) is used to compare the depth line descriptors. The DPD leads to an accurate matching of sequences even in the presence of local shifting on the shape. Experimentally, we show absolute improvement in retrieval performance on the Princeton 3D Shape Benchmark database.

1. INTRODUCTION

The 3D shape retrieval has recently been a very active area of research. We refer the reader to a recent good survey [1] and comprehensive comparative studies of 3D retrieval algorithms, reported in [2, 3, 4, 5]. The 2D/3D shape retrieval methods are based on the computation of intermediate representations corresponding to the rendered images from multiple viewpoints, considering that two models are similar when they look similar from all viewing angles. The 3D model is then indirectly represented by various 2D-shape descriptors associated with 2D views so that the 3D-shape matching is transformed into similarity measuring between 2D images. Silhouettes and depth-buffer images are the most widespread projections. A silhouette indicates the region of a 2D-image that contains the projection of the visible points of the 3D object whereas a depth-buffer image contains the information of the distance between the object and the viewing plane in its pixels.

In [6], the rendered silhouettes are encoded by their Zernike moments and Fourier descriptors. The dissimilarity of two 3D models is defined as the minimum of the sum of the distances between features extracted over all rotations and all pairs of vertices on the corresponding dodecahedrons. Filali Ansary et al. [7] introduce a novel probabilistic Bayesian using a characteristic views selection. In [8], Mahmoudi and Daoudi use curvature scale space (CSS) descriptor which provides compact and highly discriminating representations.

Most of the 2D/3D approaches based on depth images [9, 10, 11] use the two-dimensional discrete Fourier transform (2D-DFT) as a 2D-shape signature. They differ mainly in the similarity estimation technique used and in the number of views retained. Recently, Chaouch et al. [10] have introduced a technique to enhance 2D/3D shape descriptors. To take into account the dispersion of information in the views, they associate to each view a relevance index. This approach can be applied to any type of 2D/3D descriptors, in particular to the Silhouette-based and the Depth buffer-based descriptors using the Fourier-transform, presented by Vranic [11].

Our approach is based on depth images but differs with the previous ones. We introduce a new descriptor based on depth lines and a new similarity measure: first, we create sequencing of lines of the depth images in order to represent a 3D shape by a collection of sequences without loss of information; secondly, we use a Dynamic Programming distance contrarily to the previous techniques that adopt generally the Euclidean distance. In Section 2, we first describe efficient extraction of the depth sequences from a given 3D model. In Section 3, we present the proposed 3D shape matching with dynamic programming. Experimental results are given in Section 4 and we conclude in Section 5.

2. A DEPTH LINE DESCRIPTOR

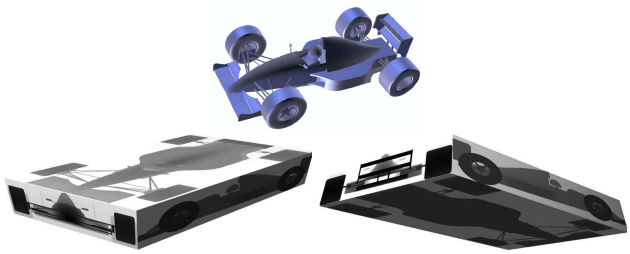

Fig. 1. Two views of the bounding box of the race car model with their associated depth-buffer images.

Our 3D shape retrieval system compares 3D models based on their visual similarity using depth lines extracted from depth images. The process first normalizes and scales 3D model into a bounding box. Then, it computes the set of $N \times N$ depth-buffer images associated to the six faces of the bounding box (see Fig. 1). The system then generates $2 \times N$ depth lines per image, considering each depth image as a collection of N horizontal and N vertical depth lines. Finally, each

depth line is represented by a sequence of states using our proposed sequencing method (see Fig. 2). The shape descriptor consists in the set of $6 \times 2 \times N$ sequences. The sequence extraction process is described step by step in the following.

2.1. Pose Normalization

A 3D object is generally given in arbitrary orientation, scale and position in the 3D space. As most of the 2D/3D shape descriptors do not satisfy the geometrical invariance, pose normalization is then necessary before the 3D object feature extraction. Some of the previous methods employed the Principal Component Analysis (PCA) as normalization step. PCA generates an alignment of a 3D-mesh model into a canonical coordinate system frame by translating, rotating, reflecting, and scaling the set of vertices. We have retained here the "Continuous" PCA [12] (CPCA) because it appears to be more complete and the most stable of all the approaches we have studied.

2.2. Depth Lines Extraction

The bounding box is used in rendering algorithms in order to generate depth-buffer images from a model. It is the minimal box, with the faces parallel to the coordinate hyper-planes that completely encompasses a 3D model. Each 3D model is projected perpendicularly on the planes of own bounding box, in order to generate six $N \times N$ depth-buffer images $\mathbf{v_i}$, $i \in \{1, ..., 6\}$. The rows and the columns of each image are decomposed into $6 \times 2 \times N$ depth lines of size N.

2.3. Sequencing of Depth Line

The challenge of the sequencing method lies in the definition of states, which need to be efficient and robust representations of the shape information. It is possible to assimilate depth lines to 2D curves, thus the space points are taken as sequence elements. However, this sequencing technique has some disadvantages. When the depth line has several parts (e.g., contains disconnected background regions), there is a bad resolution in the matching process due to the fact that the dynamic programming algorithm matches empty samples. It is then interesting to discover, analyze and preserve the content of a depth line.

The different regions present on a depth line are classified into two categories:

Background regions: they represent the background of the projected plane. Sometimes, a 3D model might be rendered into several separated parts and also certain parts might contain holes. When the situation occurs, two types of background points can be rendered: interior points and exterior points. We want to distinguish these two background states in our sequencing.

Projected object regions: they contain the projections of the visible points of 3D model. For these regions, we adopt a "derivative" description that does not consider the depths of the visible points or "the attributes of pixels", but rather considers the first derivative of the sequences. While there are sophisticated methods for estimating derivatives, we use the following estimate for simplicity $D = dl(p) - dl(p+1)$. This is the slope of the line segment between the depth at p and the depth at its right neighbor. To represent this estimate at each point, we introduce three states, Decreased, Increased and Constant, according to the sign of its derivative (negative or positive or null). We obtain information about the shape by considering the space derivative properties of the depth line.

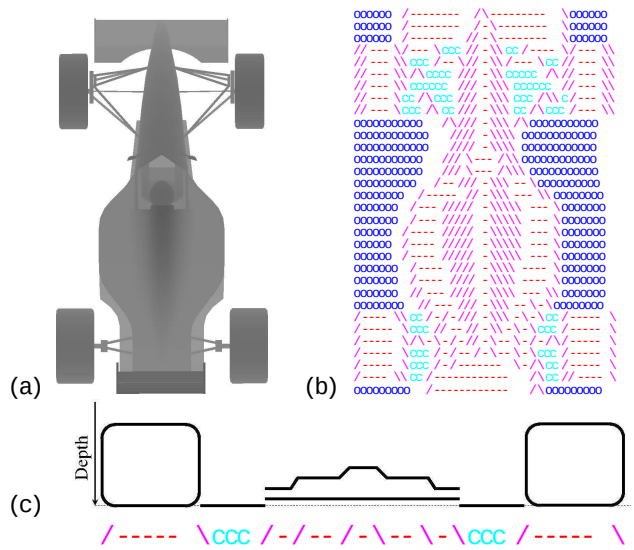

Fig. 2. (a). Depth image, (b). The horizontal sequences computed from the depth lines, (c). Sequencing of a depth line.

Each depth line is encoded in a set of N states called sequence of observations. Each such observation is represented by one of the five characters, o, c, /, -, \, representing the proposed states that compose the sequence: exterior-background, interior-background, increased-depth, constant-depth and decreased-depth (see Fig. 2.c).

3. 3D MODEL RETRIEVAL USING DYNAMIC PROGRAMMING

For some sequence similarity matching, a very simple measure, such as Hamming distance, will suffice. It is the number of corresponding state positions that differ. However, this metric turned out not to be optimal in our sequence similarity computing. It is essentially due to the fact that two similar sequences often have approximately the same overall component shapes, but these shapes do not line up in X-axis. Thus, it is necessary to find the best alignment between two sequences before computing their similarity.

As one can see on Fig. 3, the Hamming distance does not tolerate phase shifting of subsequences of the shape, thus, similar shapes that are not perfectly aligned along the X-axis lead to counter intuitive high distance values. To overcome

the limits of the Hamming metric, we introduce the Dynamic Programming distance to compare the shape descriptors, thus local deformations of objects' boundaries can be matched.

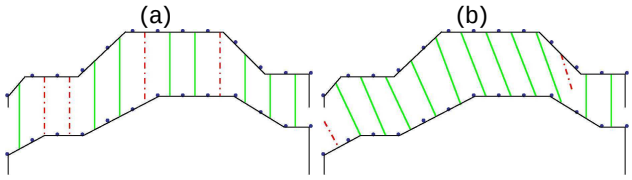

Fig. 3. Comparing the distances of two depth sequences by Hamming distance and dynamic programming distance while the sequences have an overall similar shape, but, they are not perfectly aligned along the X-axis. (a) 4 mismatches -·-·-, Hamming distance produces a pessimistic dissimilarity measure. (b) 2 mismatches -·-·-, a nonlinear alignment allows a more intuitive distance measure to be calculated.

3.1. Dynamic Programming

In diverse domains, the dynamic programming has been developed to align two signals, two sequences or two series with different dynamics. While there are many algorithms in the literature, Dynamic Time Warping (DTW), Needleman-Wunsch algorithm and Smith-Waterman algorithm account for the majority of the literature. DTW has been well-adopted for curve alignment [13] and 2D shape matching [14]. The Needleman-Wunsch distance has already proven to yield superior performance for the retrieval of sequences, and we show that this is also the case for 3D shape matching.

The Needleman-Wunsch algorithm [15], which calculates a "global" similarity distance between two sequences, has been used here. It computes the global alignments between two sequences, which produce the optimal score. Global scores require the alignment to begin at the beginning of each sequence and extend to the end of each sequence.

To align two sequences, $A = a_0 a_1 \ldots a_{N-1}$ and $B = b_0 b_1 \ldots b_{N-1}$ using dynamic programming, we first construct a N-by-N match matrix $\mathbf{M}$ using the match operators (adjacent transposition C and substitution S) where the (i^{th},j^{th}) element of matrix contains the similarity between sequences elements a_i and b_j. Then, we construct recursively the $(N+1)$-by-$(N+1)$ matrix $\mathbf{H}$ (Fig. 4) using the match matrix $\mathbf{M}$ and the gap operators (insertion I or deletion D). When $N = 32$, the computational cost of the Needleman-Wunsch algorithm is not prohibitive.

3.2. Measuring Similarity between two 3D Models

To compare two 3D models $O1$ and $O2$, we generate, $12 \times N$ observation sequences $\mathbf{S^1_{i,j,k}}$ and $\mathbf{S^2_{i,j,k}}$ from the six rendered images. Then, we compute the dynamic programming distance (DPD) between each sequence of model $O1$ and the corresponding sequence of model $O2$. Finally, the dissimilarity between a pair of 3D models, $O1$ and $O2$, are computed as the sum of all DPDs. Therefore, it is defined as follows:

$$\Delta(O1, O2) = \sum_{i=1}^{6} \sum_{j=r,c} \sum_{k=0}^{N-1} DPD(\mathbf{S^1_{i,j,k}}, \mathbf{S^2_{i,j,k}}). \quad (1)$$

where i denotes the index of the rendered images, k the order of lines in the depth image, j indicates if the depth line is a row or a column and DPD is computed as in [15].

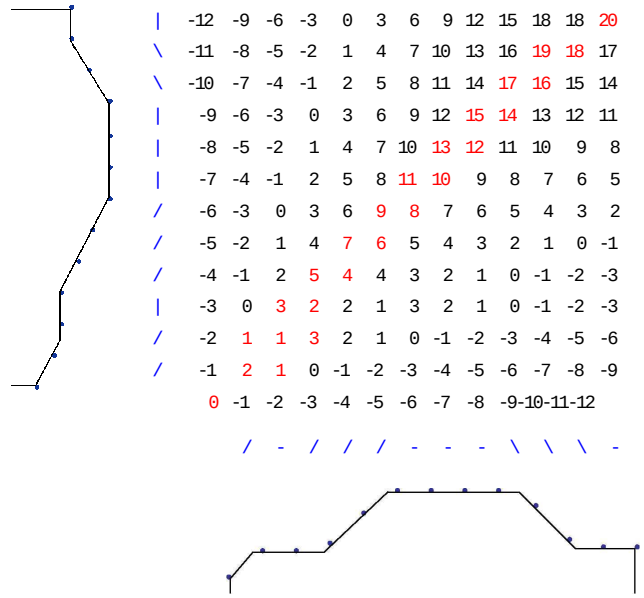

Fig. 4. Matching of the two observation sequences corresponding to the depth lines of Fig. 3 using Needleman-Wunsch algorithm [15].

4. EXPERIMENTAL RESULTS

We made our tests on the Test Princeton 3D Shape Benchmark database [2] (907 models categorized within 92 distinct classes). The experimental results presented here were obtained with the following parameters: N = 32 for the different rendered image size; (C, S, I, D) = (-2, 1, 1, 1) for Needleman-Wunsch Algorithm. To compare objectively the retrieval effectiveness of the proposed approaches, we compute Precision-Recall diagrams commonly used in information search (the query is not counted in the answer as in [11]) and four quantitative measures for evaluating query results (see [2] for a description of this measures): **(1)** The Nearest Neighbor (NN), **(2)** The First Tier (FT), **(3)** the Second Tier (ST), **(4)** The Discounted Cumulative Gain (DCG).

In Fig. 5, we compare our approach against the original depth buffer-based approach (DBA [11]) and in Fig. 6 we show some examples. Our method DLA clearly outperforms the DBA. When using the Hamming distance, our method based on the sequencing of depth lines provides more accurate retrieval results than the DBA, and by introducing the dynamic programming distance, the performance gain is increased. Moreover, if we compare these measures with the results listed in table 4 of [2], our approach produces 1%, 1.5%,

1.5% and 1% better than the top ranked Light Field Descriptor (LFD) [6], in term of NN, FT, ST and DCG measures.

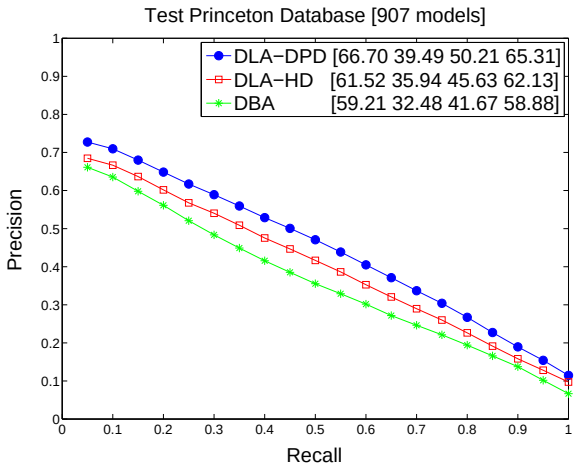

Fig. 5. Average Precision-recall curves using the proposed depth line-based approach DLA (with dynamic programming distance DPD and Hamming distance HD) and the depth buffer-based approach DBA (with the same bounding box and 6 views). The mean NN, FT, ST and DCG values are given in the legends.

5. CONCLUSION

In this paper, we have presented a new 3D descriptor based on depth lines and an associated similarity measure using dynamic programming. We have evaluated its performances on the Princeton Shape Benchmark database. The depth-line approach has proved to be efficient in space and accuracy for 3D retrieval, even if it used with the Hamming distance. Work in progress includes the test of other dynamic programming algorithms and other values of N to obtain an optimal ratio time/accuracy.

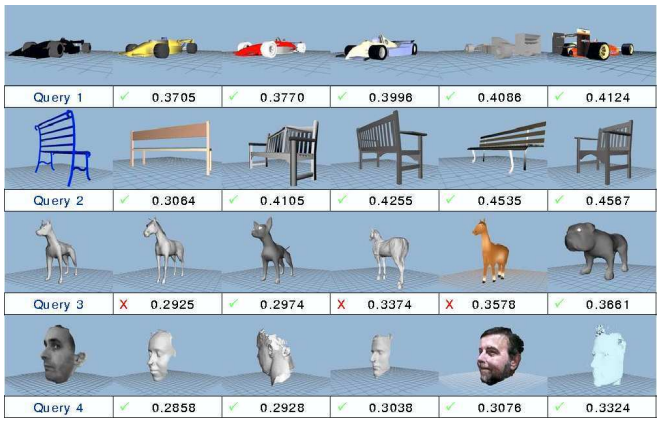

Fig. 6. Examples of similarity search. For each query, we show the top 5 objects matched with DLA approach. The similarities between the query models and the retrieved models are given below corresponding images. ✓ and x indicate that the retrieved models belong or don't belong to the query's class, respectively.

6. REFERENCES

[1] J.W.H. Tangelder and R.C. Veltkamp, "A survey of content based 3D shape retrieval methods," in *SMI'04*, Genova, Italy, June 2004, pp. 145–156.

[2] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser, "The Princeton shape benchmark," in *SMI'04*, Genova, Italy, June 2004, pp. 167–178.

[3] T. Zaharia and F. Prêteux, "3D versus 2D/3D shape descriptors: A comparative study," in *SPIE Conf. on Image Processing: Algorithms and Systems*, San Jose, CA, USA, Jan. 2004, vol. 5298.

[4] B. Bustos, D. A. Keim, T. Schreck, and D. Vranic, "An experimental comparison of feature-based 3D retrieval methods," in *3DPVT'04*, Thessaloniki, Greece, Sept. 2004.

[5] A. Del Bimbo and P. Pala, "Content-based retrieval of 3D models," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 2, no. 1, 2006.

[6] D.Y. Chen, X.P. Tian, Y.T. Shen, and M. Ouhyoung, "On visual similarity based 3D model retrieval," *Computer graphics forum*, vol. 22, no. 3, pp. 223–232, Sept. 2003.

[7] T. Filali Ansary, M. Daoudi, and J-P. Vandeborre, "3D model retrieval based on adaptive views clustering," in *ICAPR'05*, Bath, UK, Aug. 2005.

[8] S. Mahmoudi and M. Daoudi, "3D models retrieval by using characteristic views," in *ICPR'02*, Québec, Canada, Aug. 2002, pp. 11–15.

[9] R. Ohbuchi, M. Nakazawa, and T. Takei, "Retrieving 3D shapes based on their appearance," in *MIR'03*, Berkeley, CA, USA, Nov. 2003.

[10] M. Chaouch and A. Verroust-Blondet, "Enhanced 2D/3D approaches based on relevance index for 3D-shape retrieval," in *SMI'06*, Matsushima, Japan, June 2006.

[11] D.V. Vranic, *3D Model Retrieval*, Ph.D. thesis, U. of Leipzig, 2004.

[12] D. Vranic, D. Saupe, and J. Richter, "Tools for 3D-object retrieval: Karhunen-Loeve transform and spherical harmonics," in *2001 Workshop Multimedia Signal Processing*, Cannes, France, Oct. 2001.

[13] Kongming Wang and Theo Gasser, "Alignement of curves by dynamic time warping," *The Annals of Statistics*, vol. 25, no. 3, pp. 1251–1276, 1997.

[14] Eamonn J. Keogh, Li Wei, Xiaopeng Xi, Sang-Hee Lee, and Michail Vlachos, "Lb_keogh supports exact indexing of shapes under rotation invariance with arbitrary representations and distance measures.," in *VLDB*, 2006, pp. 882–893.

[15] S. Needleman and C. Wunsch, "A general method applicable to the search for similarities in the amino acid sequence of two proteins," *Journal of Molecular Biology*, vol. 48, no. 3, pp. 443–453, 1970.