3D Head Models Retrieval Based on Hierarchical Facial Region Similarity

Horace H S Ip
Image Computing Group, Department of Computer Science
Centre for Innovative Applications of Internet and Multimedia Technologies (AIMtech Centre)
City University of Hong Kong
Tat Chee Avenue, Kowloon, Hong Kong

e-mail: cship@cityu.edu.hk

William Y F Wong
Image Computing Group, Department of Computer Science
City University of Hong Kong
Tat Chee Avenue, Kowloon, Hong Kong

e-mail: fyfwong@cs.cityu.edu.hk

Abstract

This paper presents a technique for 3D head models retrieval. The approach combines a 3D shape representation scheme and hierarchical indexing of 3D models based on facial region similarity. The proposed shape similarity measure is based on comparing 3D model shape signatures computed from the Extended Gaussian Images of polygon normal. The technique is made highly efficient and scalable by partitioning the 3D head model into distinctive facial regions and building a hierarchical index for the head model database. In our database, there are over 1,000 models and all the head models are represented up to about 3,000 polygons. Furthermore, we have developed a novel user interface for specifying the visual queries and to interact with the retrieval system. We have desmonstrated that our approach performs similarly with Eigenheads but computationally more efficient by several orders of magnitudes. This makes our approach a practical solution for large model databases.

1. Introduction

With the general availability of 3D digitizers and scanners, one can readily build up large collections of 3D graphical models for different applications e.g. in CAD/CAM, games design, computer animations, and manufacturing. The retrieval of specific 3D models from a database of 3D objects becomes an issue to be tackled for both the manufacturing and entertainment industry. In computer vision, previous work on 3D model retrieval has been concentrated on 3D model invocation and pose estimation from 2D images within the context of scene interpretation [10-15], but the models involved were mainly rigid objects or CAD generated models. However, the need to retrieve non-rigid models and models of free-form surfaces is increasing particularly within the entertainment and graphics design industry.

Previous works [1,2,5,7] on model retrieval and instantiation are mainly related to 3D objects recognition and representation for rigid geometric objects, e.g. machine parts, airplanes and vehicles. In this work, we focus on developing retrieval techniques for a class of deformable 3D object, e.g. the human heads. However, it is expected that the techniques may also be generalized to 3D head models of other species. The technique computes model shape signatures based on an analysis of the distributions of the polygon normal of 3D objects as represented by the Extended Gaussian images (EGI) [9]. It is well known that the normal of a polygon defines its orientation in 3D space and an analysis of how these normals are distributed in 3D space can give an indication of the global shape of a 3D object. Previous work has discussed the applications of polygon normals and EGI for various applications, for example, in image security of watermarking [8].

A survey of the literature reviews that almost no work has been done on indexing 3D head models. Work related to indexing and retrieval of 3D objects can be found in [2, 4, 5, 6]. Relational indexing [2] and Topological surface indexing [6] are similar in concept. The idea of both is to model a 3D object from basic primitives. They define their own structure of primitive model and a 3D model is represented using graph-theoretic techniques. Both techniques apply only on simple CAD-based objects. Syeda-Mahmood [4] applied 3D object invariants to poses from 2D images. In this case, only 2D technical drawings are indexed, but the indexing is used on 3D application. Content-based 3D neuroradiological image retrieval system [5] is a good reference for 3D models indexing for medical uses. In [2, 4, 5, 6], the type of indexing is static, that means it is hand-crafted, off-line and cannot be altered easily or updated upon user interactions.

Obviously, 3D objects cannot be easily and precisely described with only textual query, therefore, visual query are essential for the retrieval and access of 3D objects efficiently. In [3], a 3D indexing language is developed to demonstrate how to establish an indexing language to describe 3D objects. However, it is only useful for simple 3D objects with limited variance of object appearance.

The organization of this paper is as follows. Sections 2 – 3 describes 3D shape representation based on polygon normal and presents the construction of a shape signature that describes 3D model shape for similarity matching. Section 4 presents an indexing scheme based on the shape signature of facial regions. An implementation of our approach is shown in section 5, which demonstrates how to makes use of indexing with visual query for retrieval. Section 6 compares the performance of this algorithm with other approaches, e.g. Eigenfaces [16]. The final section presents the conclusion.

2. Representation of 3D head models

The head models used in our database were created by using 3SPACE FastSCAN hand held laser 3D scanner. Scanner records the person's data from definite positions with fixed distance in order to normalize all the head models with a common reference system axis as shown in Fig. 1. The intersection of the system axis is mapped to the center of Gaussian sphere, so that all the EGI transformed from head models are consistent. During scanning, people are required to wear a swim cap to cover their hair, it is because hairs are black in color which cannot be scanned. And they all present a natural facial expression.

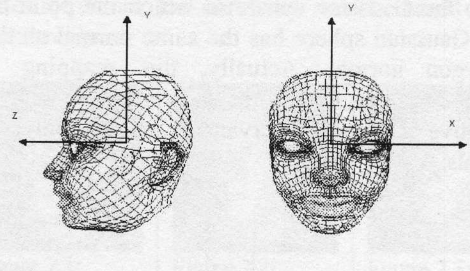

Figure 1: *Normalization of head models on x, y and z axis*

In our head models database, all the models are composed of 3,118 polygons with a total of 2,954 vertices. Each polygon of the model is formed by either 3 or 4 vertices. Fig. 2.1 – Fig. 2.5 show five samples on our database. These five sample models will be used as examples in the following sections to illustrate our approach.

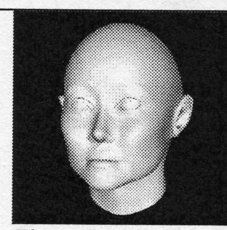

Figure 2.1a: *A female head model*

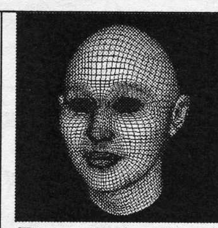

Figure 2.1b: Wire frame of Figure 2.1a

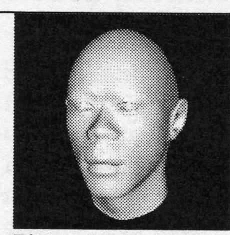

Figure 2.2: *A thin and long head model*

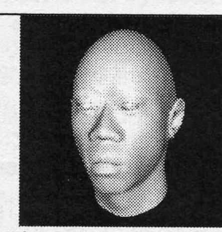

Figure 2.3: *Another thin and long head model*

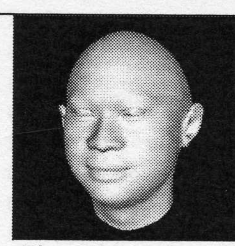

Figure 2.4: *A fat and short head model*

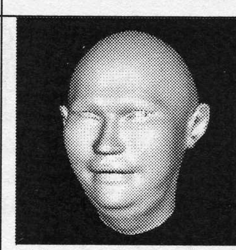

Figure 2.5: *Another fat and short head model*

3. From Extended Gaussian Image to model signature

To form the EGI of a head model, the normal vector of each polygon of a head can be mapped onto the Gaussian sphere in which the point on the Gaussian sphere has the same normal as the polygon normal. Actually, this mapping is unique if the head model is convex with a positive Gaussian curvature at all polygon points.

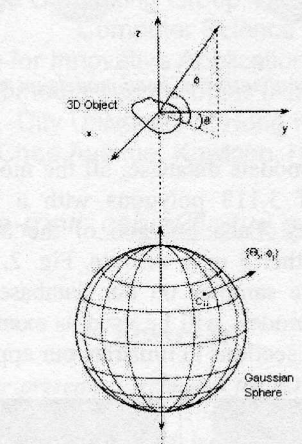

Figure 3: *Conversion from polygon normal to EGI*

Next, the Gaussian sphere is divided into cells. Each cell corresponds to a range of orientations. In practice, we choose the (θ, ϕ) representation as a sampling scheme for the discretization of the Gaussian sphere as longitude (θ) and latitude (ϕ). Refer to Fig. 3, all points in the cell c_{ij} of the EGI are assumed to have the same range of spherical angles (θ_i, ϕ_j). It means any normal, if its longitude (θ) and latitude (ϕ) are within (θ_i, ϕ_j), it belongs to cell c_{ij}.

After constructing the EGI for the head model, we then map the cells of sphere onto a rectangular array to form a 1D shape signature of the 3D models.

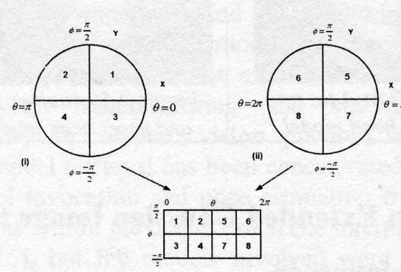

Figure 4: *Mapping from Gaussian sphere to rectangular array*

In Fig. 4, we uniquely map points on a Gaussian sphere to a rectangular array. Without loss of generality, we simplify the discussion by dividing the Gaussian sphere into 8 cells. Fig. 4 (i) and Fig. 4 (ii) show the front and back hemisphere respectively. The numbers on the hemispheres indicate the corresponding positions on the array shown in Fig. 4 (iii). In our experiments, the column and the row of the array are set to 20 and 10 respectively, and the total number of cells is equal to 200. Using this array bin size, experiments show that two head models would have the same EGI if the variations of corresponding vertexes coordinates are within 2.7%. That means if a person's head is scanned twice, the two head models created may have difference in shape due to digitization noise. However, if the vertexes coordinates between them are within the above percentage, they will yield the same EGI and hence the same shape signature. At the same time, this number also provides the maximum samplings on the Gaussian sphere in order to obtain the highest precision on head models.

Each cell on the rectangular array counts the total number of normals belonging to it. After counting, a 1D signature can be constructed. Fig. 5.1 – Fig. 5.2 show the corresponding 1D signature for the head models Fig. 2.2 – Fig. 2.4 respectively, where x-axis is the EGI for the head model, y-axis is the total number of normal. In practice, the difference of any two signatures can be efficiently quantified by histogram differencing as defined in equation (1):

$$\textit{Similarity Measure } (SM) = \frac{\sum |E_1(x,y) - E_2(x,y)|}{\sum (E_1(x,y) + E_2(x,y))} \quad \textbf{(1)}$$

where $E_1(x,y)$ and $E_2(x,y)$ mean the EGI of two models with coordinates (x,y) respectively, in which $0 \leq x <$ EGI Columns and $0 \leq y <$EGI Rows.

The minimum and maximum values of the Similarity Measure are 0 and 1 respectively, in which a lower value means the two signatures are more similar. Table 1 compares the similarity between the above 5 signatures. Fig. 6 shows the relation of EGI samplings and corresponding Similarity Measure for head model Fig. 2.1.

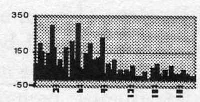

Figure 5.1: *Shape Signature for Figure 2.2*

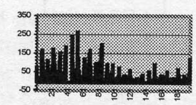

Figure 5.2: *Shape Signature for Figure 2.3*

Table 1. Similarity Measure between the head models Fig. 2.1 – Fig. 2.5

Similarity Measure	Fig. 2.1	Fig. 2.2	Fig. 2.3	Fig. 2.4	Fig. 2.5
Fig. 2.1	\	0.228	0.218	0.251	0.279
Fig. 2.2	0.228	\	0.163	0.284	0.307
Fig. 2.3	0.218	0.163	\	0.289	0.312
Fig. 2.4	0.251	0.284	0.289	\	0.149
Fig. 2.5	0.279	0.307	0.312	0.149	\

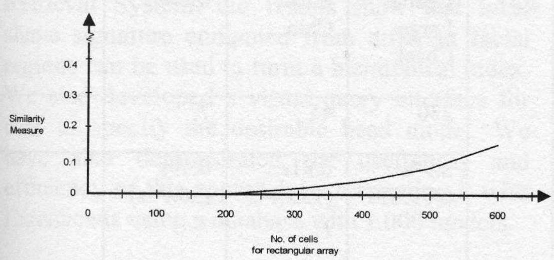

Figure 6:: *Relation between EGI sampling and Similarity Measure for head model Fig. 2.1.*

4. 3D model retrieval based on facial regions similarity

An human head can in general be divided into several distinctive facial regions based on gross facial features such as chin, nose, mouth and eyes, we designed a hierarchical index structure based on the polygon normal distribution due to these facial features. Each search of the index structure will further narrow down the remaining model samples for efficient retrieval. In our case, the top levels indexes require facial features that composed of 1,000 to 1,500 polygonal surfaces (and hence polygon normals) for coarse filtering and low level indexes make use of facial features that composed of about 50 to 200 polygons for fine selection. In fact, coarse level indexes are built by a union of several individual fine facial features, these coarse indexes will subsequently be split into smaller subsets called fine facial features indexes. Fig. 7.1 – Fig. 7.3 show facial features of forehead, nose and nose bone respectively. Fig. 8 shows the hierarchical index structure for our head models based on facial regions.

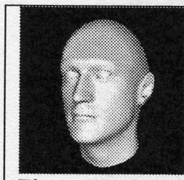

Figure 7.1: *Coarse upper face index*

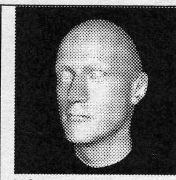

Figure 7.2: *Nose feature index*

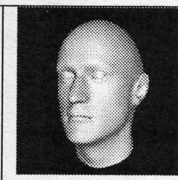

Figure 7.3: *Nose bone index*

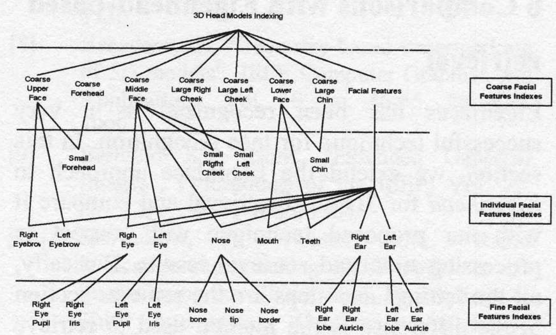

Figure 8: *Hierarchical index structure for head models based on facial regions*

5. A 3D head model retrieval system

This section describes a system which has been built to evaluate this approach. It aims to search for similar head models from a query sample. A user can define the query sample through an interactive interface to manipulate a 3D head model and to edit the shape or feature size of a generic head model to the desirable head shape query. The system database consists of over 1,000 head models. After a user submitted his/her query sample, a number of models are selected based on the similarity measure presented in section 2. User can further explore the remaining models and choose one of them for further modifications iteratively until the target group of models are found. Fig. 9 shows the interface of our Head Models Retrieval System.

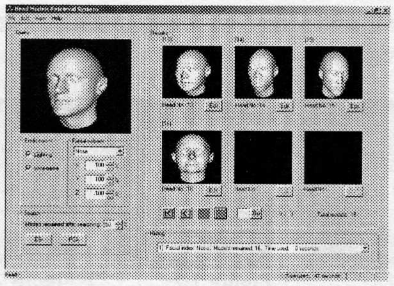

Figure 9: *The user interface design of a Head Models Retrieval System*

6 Comparisons with Eigenhead-based retrieval

Eigenfaces has been recognized as a very successful technique for face recognition. In this section, we extend the Eigenface approach to *Eigenhead* for 3D head retrieval and compare it with our proposed technique with respect to processing time and retrieval results. Basically, all the settings and steps are the same as section 5, the difference is the method used to retrieve similar head models. First, we collect the components of normals for all the head models in database and perform principal component analysis. It is defined as the head space. After the user has constructed the query on the Head Models Retrieval System, we analysis the components of normals of this query and project them onto the head space. By analysing the corresponding distribution in thehead space, we compute the similarity with the head models. Table 2 shows the successive refinement of retrieved models for searching similar head models of Fig. 2.1 – Fig. 2.5 within a 1,000 models database using different users defined queries using our facial regions similarity and the Eignheads approach respectively.

Table 2: *Experiment results showing successive refinement of retrieved models for searching similar head models of Fig. 2.1 – Fig. 2.5 within a 1,000 models database by Facial regions similarity and Eigenheads*

Fig.	**Searching results:**			
	First search	**Second search**	**Third search**	**Fourth search**
	Index type:			
	Number of retrieved models which have similarity measure smaller than a given threshold value:			
	Processing time (sec):			
2.1)	Coarse Middle Face	Nose	Nose Border	
	530	250	25	
	4.25s (466.21s)	1.98s (178.09s)	0.88s (34.18s)	
2.2)	Coarse Large Cheek	Coarse Large Chin	Nose	Nose Bone (Nose Tip)
	480	240	105	17 (15)
	4.11s (405.22s)	2.12s (140.42s)	1.33s (38.17s)	0.36s (6.93s)
2.3)	Coarse Large Cheek	Small Chin	Mouth	
	230	45	7	
	3.97s (399.02s)	0.91s (133.44s)	0.19s (56.79s)	
2.4)	Coarse Middle Face	Nose	Nose Border	Eyes
	550	310	125	35
	4.03s (420.11s)	3.88s (140.89s)	2.89s (112.80s)	1.22s (14.23s)
2.5)	Coarse Middle Face	Nose	Nose Border	Small Chin
	660	420 (300)	130 (150)	30
	5.17s (530.83s)	3.76s (184.11s)	1.21s (61.02s)	0.33s (16.52s)

*Remarks: ( ) brackets mean value obtained by Eigenheads, otherwise, both approaches obtain the same value.

Referring to Table 2, we can observe that the results are similar with respect to the feature indexes used as well as models remained after each search. The main difference is the processing time required. On the other hand, experiments also show that there is no difference in the retrieval results by both methods for the top five most similar head models using Fig. 2.1 and Fig. 2.5 respectively. Actually, both approaches can search for target models for users, but Eigenheads required significantly more computational time, especially when the number of models in the database is large as in this case.

7 Conclusion

In this paper, we propose a representation scheme that supports indexing and retrieval of 3D head models. First, we present the construction of a compact shape signature which allows similarity between 3D models to be computed efficiently. Then, we categorize these models based on their facial regions in groups or sub-divide them into small parts as coarse and fine indexes respectively. Through the application of a prototype Head Models Retrieval System, our results show that local shape signature computed from different facial regions can be used to form a hierarchical index. We also developed a visual query interface for user to specify the desirable head model. We have also demonstrated the usefulness and efficiency of this approach by comparing it with Eigenheads using a database with 1,000 models.

Acknowledgements

The authors would like to thank the students who have acted as the models for constructing 3D head models in our database. This work is part-sponsored by HKSAR RGC Research Grants No. CityU1115/99E and CityU 1150/01E.

Reference

[1] Eric Paquet and Marc Rioux, "Content-Based Access of VRML Libraries", Proceedings of IAPR Workshop on Multimedia Information Analysis and Retrieval (MINAR'98), Lectures Notes in Computer Science 1464, pp. 20-32, 1998.

[2] Coista, M.S., Shapriro, L.G., "Scene analysis using appearance-based models and relational indexing", Proceedings of International Symposium on Computer Vision, pp. 103-108, 1995.

[3] Horikoshi, T., Kasahara, H., "3-D shape indexing language", Proceedings of Ninth Annual International Phoneix Conference on Computers and Communications, pp. 493-499, 1990.

[4] Syeda-Mahmood, Tanveer, "Indexing of technical line drawing databases", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 21, Issue 8, pp. 737-751, 1999.

[5] Yanxi Liu, Rothfus, W.E., Kanade, T., "Content-based 3D neuroradiologic image retrieval: preliminary results", Proceedings of IEEE International Workshop on Content-Based Access of Image and Video Database, pp. 91-100, 1998.

[6] Villalobos, L., Merat, F.L., "3D modeling and indexing for CAD-based object recognition", Proceedings of IEEE International Conference on Rogbotics and Automation, Vol. 3, pp. 1965-1972, 1994.

[7] Dorai, C., Jain, A.K., "COSMOS-A representation scheme for 3D free-form objects", IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1115-1130.

[8] Benedens, O., "Geometry-based watermarking of 3D models", IEEE Computer Graphics and Applications, Vol. 19, pp. 46-55, Jan.-Feb. 1999.

[9] Berthold K.P. Horn, "Extended Gaussian Images", Proceedings of the IEEE, Vol. 72, No. 12, pp. 1671-1686, 1984.

[10] McAndrew, P., Wallace, A.M., "Rapid invocation and matching of 3D models to 2D images using curvilinear data", Proceedings of Third International Conference on Image Processing and its Applications 1989, pp. 83-87, 1989.

[11] Lingxiang Zhou, Weikang Gu, "A passive approach toward interactive 3D scene modeling", Proceedings of Fourth International Conference on Signal Processing Proceedings, ICSP '98, Vol. 2, pp. 865-868, 1998.

[12] Hermiston, E.J., Booth, D.M., Foulkes, S.B., Reno, A.L., "Pose estimation and recognition of ground vehicles in aerial reconnaissance imagery", Proceedings of 14th International Conference on Pattern recognition (ICPR'98), Vol. 1, pp. 578-582, 1998.

[13] Wunsch, P., Hirzinger, G., "Registration of CAD-models to images by iterative inverse perspective matching", Proceedings of the 13th International Conference on Pattern Recognition (ICPR'96), Vol. 1, pp. 78-83, 1996.

[14] Kayanuma, M., Hagiwara, M., "A new method to detect object and estimate the position and the orientation from an image using a 3-D model have feature points", Proceedings of IEEE International Conference on Systems, Man, and Cybernetics, 1999, Vol. 4, pp. 931-936, 1999.

[15] Wong, A.K.C., Rong, L., Liang, X., "Robotic vision: 3D object recognition and pose determination", Proceedings of International Conference on Intelligent Robots and Systems, Vol. 2, pp. 1202-1209, 1998.

[16] M.Turk, A.Pentland. "Eigenfaces for Recognition", Journal of cognitive neuroscience, 3(1), pp. 71-86, 1991.