Indexing and Retrieval of 3D Models by Unsupervised Clustering with Hierarchical SOM

Hau-San Wong*, Kent K. T. Cheung*, Yang Sha* and Horace H. S. Ip*,+
* Image Computing Group, Department of Computer Science, City University of Hong Kong
+ Centre for Innovative Applications of Internet and Multimedia Technologies (AIMtech) Center, City University of Hong Kong
email: cshswong@cityu.edu.hk, cskent@cityu .edu.hk, 50095477@student.cityu.edu.hk, cship@cityu.edu.hk

Abstract

A hierarchical indexing structure for 3D model retrieval based on the Hierarchical Self Organizing Map (HSOM) is proposed. The proposed approach organizes the database into a hierarchy so that head models are partitioned by coarse features initially and finer scale features are used in lower levels. The aim is to traverse a small subset of the database during retrieval. This is made possible by exploiting the multi-resolution capability of spherical wavelet features to successively approximate the salient characteristics of the head models, which are encoded in the form of weight vectors associated with the nodes at different levels (from coarse to fine) of the HSOM. To avoid premature commitment to a possibly erroneous model class, search is propagated from a subset of nodes at each level, which is selected based on a fuzzy membership measure between the query feature vector and weight vector, instead of taking the winner-take-all approach. Experiments show that, in addition to efficiency improvement, model retrieval based on the HSOM approach is able to achieve a much higher accuracy compared with the case where no indexing is performed.

1. Introduction

The general availability of 3D digitizers and scanners facilitates the construction of large collections of 3D graphical models for different applications e.g. in CAD/CAM, game design, computer animations, and manufacturing, This in turn leads to the proliferation of research in 3D model retrieval.

In previous works of 3D model retrieval, the focus is mainly on the adoption of global shape features for model description. [1-4]. To take into account of local shape variations as well, topological relationships between the model components were considered [5]. However, this method is computationally intensive due to the need to perform graph matching. Efficient indexing of 3D models has been considered to a limited extent [6], but this is mostly applied to primitive shape features only.

This paper proposes a hierarchical indexing structure for 3D model retrieval that takes into account both global and local features Given this hierarchical arrangement, only a small subset of the model database needs to be searched in retrieval applications. Specifically, multi-resolution 3D model features are derived from a set of spherical wavelet basis functions, and features associated with different resolutions are mapped to the corresponding levels of a Hierarchical Self Organizing Map (HSOM) [7]. The first level weights represent the coarse resolution wavelet features, while node weights at successive levels encode increasingly finer details.

To avoid premature commitment to a possibly erroneous model class, search is propagated from a *subset* of candidate nodes at each level of the HSOM, instead of a single node as in the case of winner-take-all competition. The subset of candidate nodes is selected based on a fuzzy membership measure between the query feature vector and the weight vectors at each level.

2. Feature Extraction

Adopting a particular feature vector representation for each 3-D model is equivalent to the specification of a set of basis functions for describing the models in the database. In addition, the set of basis functions to be adopted should satisfy the successive approximation

0-7695-2128-2/04 $20.00 (C) 2004 IEEE

property. Specifically, we assume that a given 3-D model is represented as the function $f(\mathbf{z})$, where **z** is the position vector associated with a suitable parameterization of the surface.

Suppose further that there exist a set of basis functions $g_1(\mathbf{z}),\ldots\ldots,g_S(\mathbf{z})$, such that each model in the database can be adequately represented by this function set as follows

$$f(\mathbf{z})=\sum_{s=1}^{S} v_s g_s(\mathbf{z}) \tag{1}$$

and that the following successive approximation property is satisfied:

$$\left\| f(\mathbf{z})-\sum_{s=1}^{p+1} v_s g_s(\mathbf{z}) \right\| \le \left\| f(\mathbf{z})-\sum_{s=1}^{p} v_s g_s(\mathbf{z}) \right\| \tag{2}$$

Thus, the model characterizing function $f(\mathbf{z})$ is approximated by the basis functions and their associated coefficients v_s. Therefore, given the basis functions $g_1(\mathbf{z}),\ldots\ldots,g_S(\mathbf{z})$, each of the 3-D models in the database can be represented as the feature vector $\mathbf{v}=[v_1,\ldots\ldots,v_S]^T$, which serves as the index for the database. Based on the above successive approximation formulation, a hierarchical indexing structure can be established.

In this paper, the spherical wavelet coefficients [8] are chosen as the basis function sets. Specifically, different wavelet scales are chosen to characterize head model features under different resolutions. The spherical wavelet coefficients had been shown to be useful for multi-resolution object recognition [9].

3. Hierarchical Index Structure

3.1 Hierarchical SOM

A self-organizing map represents a set of training data using a small set of *I* clusters, with the *i*-th cluster characterized by a node n_i with associated weight vector $\mathbf{w}_i$. In addition, a neighborhood relationship (in the form of a 2-D grid) is defined between these nodes such that similar clusters are located close to each other. The hierarchical SOM is an enhancement of the original architecture such that the set of data associated with a particular weight vector is further partitioned using a secondary SOM into a finer set of clusters. This recursive splitting process is repeated for the secondary set of weight vectors to facilitate a hierarchical data representation.

Given the successive approximation property of the feature vectors for the 3-D model, where a model can be approximately represented by a subset of basis coefficients and the representation can be further refined by the inclusion of additional coefficients, a corresponding HSOM architecture is defined as follows: A set of sub-vectors $\mathbf{v}^k=\left[v_1^k,\ldots\ldots,v_{S_k}^k\right]\in\mathbf{R}^{S_k}$ with dimensionalities $1\le S_k\le S_K$, $k=1,\ldots\ldots,K$ are extracted from the full feature vector $\mathbf{v}\in\mathbf{R}^S$ for each 3-D model. In other words, there are K levels of hierarchies in the indexing structure. The sub-vectors corresponding to the coarse resolution wavelet coefficients are then used to form the clusters with associated weight vectors $\mathbf{w}_i^k\in\mathbf{R}^{S_k}$ in the first level. For each of the clusters in the first level, the second set of sub-vectors are adopted to form secondary clusters in the next hierarchical level. This process is repeated until the last set of sub-vectors corresponding to the finest resolution level are adopted at the leaf clusters with associated weight vectors $\mathbf{w}_i^K\in\mathbf{R}^{S_K}$. The hierarchical indexing structure is shown in Fig. 1.

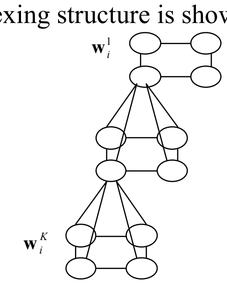

Figure 1. HSOM architecture for 3-D model retrieval

3.2 Fuzzy HSOM

Given a query model $\mathbf{v}_q$, we evaluate its fuzzy membership value U_{qi}^k with respect to a node n_i at a particular level k based on the following fuzzy membership function

$$U_{qi}^k = 1 \Bigg/ \frac{\left\|\mathbf{v}_q^k-\mathbf{w}_i^k\right\|^{\frac{2}{m-1}}}{\sum_{j=1}^{I}\left\|\mathbf{v}_q^k-\mathbf{w}_j^k\right\|^{\frac{2}{m-1}}} \tag{3}$$

where $\|\bullet\|$ is the Euclidean distance, and *m* is an exponent which determines the degree of fuzziness.

Given the set of fuzzy membership values, we construct a subset of candidate nodes B_q^1 which consist of those with associated membership values U_{qi}^1 greater than a particular threshold U_T at the first level:

$$B_q^1=\{n_i^1 : U_{qi}^1\ge U_T\} \tag{4}$$

where $\overline{\mathbf{w}_i^k}$ is the average weight vectors of the neighbouring nodes and α is a weighting factor which determines the influence of the neighbourhood.

Figure 2 shows the top-right corner and the bottom-left corners of the first level sub-map of the Fuzzy HSOM, alongside with two second level sub-maps of two selected nodes. We can observe the gradual variations of the encoded models across the nodes due to the neighbourhood preservation capability of SOM.

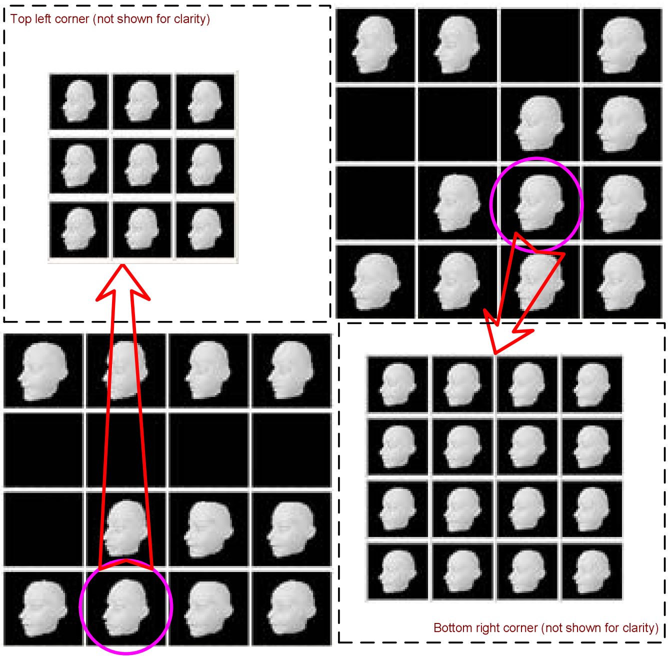

Figure 2. The FuzzyHSOM: Two corners of the first level sub-map and two second levels sub-maps of the indicated nodes

In each succeeding level k, the corresponding map can be considered as being partitioned into a set of submaps, each of which corresponds to a parent node in the previous level. The nodes in the submap thus serve to encode the possible variations of the models in its associated parent model class. As a result, instead of performing fuzzy matching for all nodes, we need only to construct the candidate node subset B_q^k by considering submaps emanating from the candidate node subset B_q^{k-1} in the previous layer:

$$B_q^k = \{n_i^k : U_{qi}^k \geq U_T \text{ and } n_{P(i)}^{k-1} \in B_q^{k-1}\} \quad (5)$$

where $P(i)$ maps the node i in level k to its parent in level $k-1$. The final search set thus consists of all those models which are assigned to the candidate node set B_q^K at the last level K. In this way, the problem of premature commitment to a possibly incorrect model class at the initial levels could be alleviated.

For HSOM training, a set of training examples, $\mathbf{v}_r$ is used to update the weight vectors $\mathbf{w}_i^k$ in each epoch based on the previous set of membership values as follows:

$$\mathbf{w}_i^k(t+1) = \frac{\sum_{r=1}^{R} U_{ri}^m(t)\mathbf{v}_r^k + \alpha\overline{\mathbf{w}_i^k(t)}}{\sum_{r=1}^{R} U_{ri}^m(t) + \alpha} \quad (6)$$

4. Experiments

We have applied our proposed approach to a 3D head model database. The set of models is categorized into six classes, each of which consists of 50 models. Another set of models of equal quantity are used as sample queries. In general, a hierarchical indexing structure runs the risk of omitting relevant records by searching only a subset of the database. To see if our fuzzy categorization approach could overcome this problem, we compare the retrieval results using the proposed hierarchical structure with those based on a full search of the database.

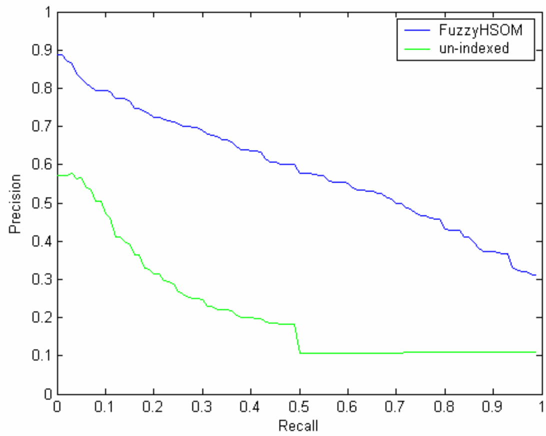

Figure 3. Recall-Precision Graph of HSOM and full search

Figure 3 shows the two recall-precision curves for retrieval based on FuzzyHSOM and full search respectively. We can observe that retrieval results based on FuzzyHSOM is significantly better than the case where no hierarchical indexing is performed.

Out of 300 models in the database, 133.5 models are matched with the query during retrieval. Of the 300 sample queries we have tested, the maximum number of models matched is 168. Comparing with a retrieval engine without any indexing, we save over 55% of matching time on average but still save 44% in the worst case.

Figure 4 shows a sample query and the top 5 retrieved models using FuzzyHSOM. As a comparison, Figure 5 shows the top 5 ranked models based on full search.

It can be observed that the retrieved models using FuzzyHSOM are visually similar to the query. On the other hand, for the retrieved list based on full search, it is seen that some models which are not visually similar to the query also appears in the top ranked list, most noticeably the models in rank 2 and 4. This can be explained by the subtle variations of the characteristics of the models across classes, which causes the appearance of many non-relevant models in the top ranked list if no hierarchical indexing is performed to suitably constrain the search domain.

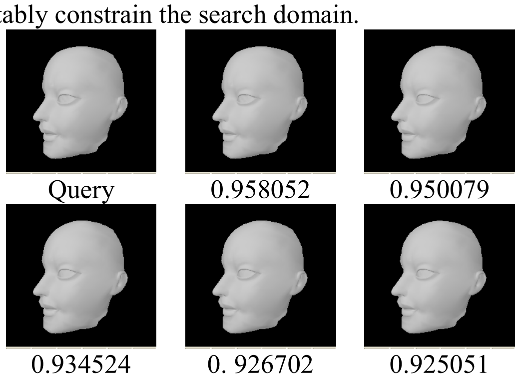

Figure 4. Sample query results based on HSOM

5. Conclusion

We propose a hierarchical indexing structure for 3D model retrieval based on the HSOM unsupervised learning process to facilitate efficient searching of the database. Coarse resolution spherical wavelet features are first adopted to delineate the global features of the models at the top HSOM level, while the fine resolution features are used to characterize the possible model variations within a parent class at the lower levels. In addition to efficiency improvement, it is observed in experimental studies that, when the variations of model characteristics are subtle across classes, the proposed approach is capable to improve the retrieval accuracy as well by constraining the search domain in the retrieval process

Acknowledgements

The work described in this paper was partially supported by a grant from the Research Grants Council of Hong Kong Special Administrative Region, China [Project No. CityU 1197/03E] and a grant from City University of Hong Kong [Project No. 7001417].

Figure 5. Sample query results based on full search

References

[1] E. Paquet, M. Rioux, A. Murching, T. Naveen and A. Tabatabai, "Description of Shape Information for 2-D and 3-D Objects", Signal Processing: Image Communication, vol. 16, 103-122 (2000).

[2] D. V. Vranic, D. Saupe, and J. Richter, "Tools for 3D-object Retrieval: Karhunen-Loeve Transform and Spherical Harmonics", Proc. IEEE 2001 Workshop Multimedia Signal Processing, Cannes, France, pp. 293-298, Oct. 2001.

[3] H. H. S. Ip and W. Y. F. Wong, "3D Head Models Retrieval Based on Hierarchical Facial Region Similarity", 15th Vision Interface, Calgary, Canada, pp. 314-319, May, 2002.

[4] Kent K. T. Cheung, Hau-Sang Wong, Horace H. S. Ip, "3D Graphical Model Retrieval by Evolutoinary Computation-based Adaptive Histogram Binning", Proc. Of 6th Intil. Conf. on Visual Information Systems (VISual'2003) and 9th Intl. Conf. on Distributed Multimedia Systems (DMS'2003), pp. 420-425, Miami, USA, 24-26 Sep., 2003.

[5] M. Hilaga, Y. Shinagawa, T. Kohmura and T. Kunii, "Topology Matching for Fully Automatic Similarity Estimation of 3D Shapes", Proc. ACM SIGGRAPH '01, Log Angeles, USA, pp. 203-212, Aug., 2001.

[6] D. Keim, "Efficient Geometry-based Similarity Search for 3D Spatial Databases", Proc. 1999 ACM SIGMOD Intl. Conf. on Management of Data and Symposium on Principles of Database Systems, Philadelphia, USA, pp. 419-430, June, 1999.

[7] J. Lampinen and E. Oja, "Clustering properties of hierarchical self-organizing maps," Journal of Mathematical Imaging and Vision, vol. 2, no. 2-3, pp. 261-272, 1992.

[8] J. P. Antoine, L. Demanet, L. Jacques, P. Vandergheynst, "Wavelets on the sphere : Implementation and Approximation", Appl. Comput. Harmon. Anal. 13, 177-200 (2002).

[9] L. Pastor, A. Rodrguez, J. M. Espadero┌ , L. Rincon, "3D Wavelet-based Multiresolution Object Representation", Pattern Recognition, 34(12), 2497-2513 (2001).