

A COGNITIVE PSYCHOLOGY-BASED APPROACH FOR 3-D SHAPE RETRIEVAL

Hsueh-Yi Sean Lin[†], Hong-Yuan Mark Liao^{†‡}, and Ja-Chen Lin[†]

[†]Department of Computer Science
National Chiao Tung University
Hsinchu, Taiwan 300

[‡]Institute of Information Science
Academia Sinica
Taipei, Taiwan 115

ABSTRACT

In this paper, we incorporate a set of principles originated from cognitive psychology into the design of 3-D shape analysis and retrieval algorithms. Based on the “visual salience-guided mesh decomposition” we previously proposed, a 3-D mesh-based shape is first broken up into parts such that human visual perception on parts can be appropriately mimicked. Next, the decomposed parts are individually analyzed and quantified according to the properties of visual salience. To establish the indices of 3-D meshes for shape matching, spherical parameterization is adopted to map the decomposed parts onto the surface of a unit sphere. In this way, the dissimilarity degree between the query acquired from users and a model in database can be calculated. The experimental results have shown that the matching performance of the proposed scheme is indeed efficient and powerful.

1. INTRODUCTION

With the fast growth of the Internet related technologies, nowadays users can easily download and distribute multimedia contents through the network channels. Due to the proliferation of multimedia contents, it is an urgent need to develop an efficient indexing mechanism to assist users in retrieval of the requested contents. In the past decade, a large number of search engines have been proposed for the retrieval of multimedia contents. Among the different types of multimedia retrieval techniques, 3-D shape retrieval has recently become a popular research topic due to the fast development of graphics hardware and software. Most of the existing 3-D shape retrieval techniques extract features based on statistical analysis. Examples of some statistics-based approaches include shape distribution [5] and higher-order moment analysis [9]. The advantage of applying a statistics-based approach is its efficiency in executing feature extraction. However, a statistics-based approach can only characterize the global information but it cannot describe the fine structure of a 3-D shape. Therefore, most

This work was supported by NSC (No. 93-2422-H-001-0003) and DOIT (No. 93-EC-17-A-02-S1-032) of Taiwan.

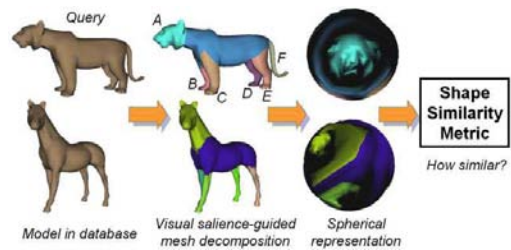


Fig. 1. The flowchart of the proposed cognitive psychology-based scheme for 3-D shape retrieval.

of the existing methods adopt a statistics-based approach to conduct coarse search. On the other hand, a structure-based approach [3] characterizes the structural information of a 3-D shape. Therefore, it is able to be used as the basis for fine search. However, a structure-based approach is very time-consuming since most of the computing power would be consumed in calculating relational information. Furthermore, the existing 3-D shape analysis and retrieval algorithms are lack of explicit link to psychological principles while cognitive psychologists have found a set of principles (or properties) that are fairly important in perception of a form or a shape.

In this paper, we incorporate a set of principles which are originated from cognitive psychology into the design of 3-D shape retrieval system. Based on the “visual salience-guided mesh decomposition” mechanism we previously developed [8], a 3-D mesh is first broken up into parts based on visual salience such that the human visual model can be appropriately mimicked. Next, the decomposed parts are individually analyzed and quantified according to the psychological theory of visual salience [2]. To properly represent 3-D meshes for shape matching, the spherical parameterization scheme proposed in [6] is adopted to map the decomposed parts onto the surface of a unit sphere. With this representation, the dissimilarity between the query and the models pre-stored in the database can be easily compared within a canonical coordinate system. Fig. 1 shows the flowchart of the proposed cognitive psychology-based scheme for 3-D shape retrieval.

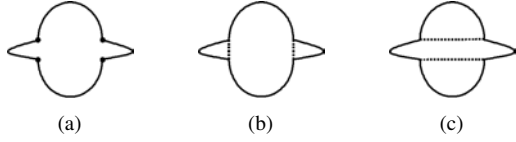


Fig. 2. Illustration of (a) part boundaries and (b)-(c) part cuts on a 2-D silhouette (Re-sketched from [2]).

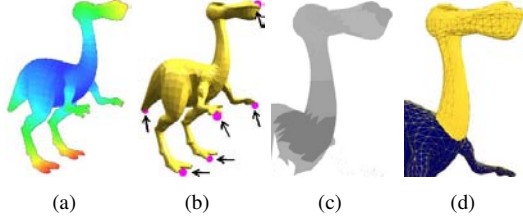


Fig. 3. The visual salience-guided mesh decomposition process [8]: (a) protrusion characterization; (b) choosing salient representatives; (c)-(d) boundary strength characterization and generating the boundary for certain part.

2. A COGNITIVE PSYCHOLOGY-BASED APPROACH FOR 3-D SHAPE RETRIEVAL

According to Hoffman and Singh’s theory [2], there are at least three factors that determine the salience of a part: the protrusion, the boundary strength, and the relative size. Based on different visual salience perceived by the human visual system, a 2-D silhouette or a 3-D shape may have different interpretations and decomposition results. As shown in Fig. 2(b), for example, the 2-D silhouette shown in Fig. 2(a) might be interpreted as an alien’s head with a pair of protrusive ears when the salience of a part is defined primarily by its protrusion. As shown in Fig. 2(c), on the other hand, the same 2-D silhouette might be interpreted as a UFO when the part salience is defined primarily by its relative size. As a result, it is apparent that the process of defining the salience of a part does influence the judgment of the human visual processes. We, therefore, propose a new 3-D shape retrieval system based on the psychology-based rules proposed by Hoffman and Singh [2].

2.1. Visual Salience-Guided Mesh Decomposition

In [8], we proposed a “visual salience-guided mesh decomposition” scheme, which uses two visually salient features (i.e., protrusion and boundary strength) to guide the 3-D mesh decomposition process. Since the features adopted are closely related to Hoffman and Singh’s theory of part salience, the decomposition process can appropriately mimic the function of a human visual system. We apply this algorithm to break up a 3-D mesh into parts. Fig. 3 shows the visual salience-guided mesh decomposition process [8].

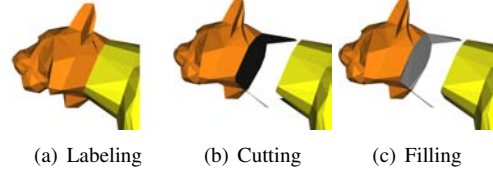


Fig. 4. The protrusion of a 3-D part can be obtained by first cutting the decomposed part, applying the filling holes algorithm [4] to form its base, and then calculating the ratio of the area of the part surface to the area of the base.

Table 1. Part salience of the query ‘tiger’ shown in Fig. 1.

Part	Protrusion	Boundary strength	Relative size
A	4.658243	-0.992333	0.171755
B	5.012940	-0.583065	0.063972
C	5.407292	-1.266561	0.067655
D	5.596881	-2.180321	0.106905
E	5.315072	-1.791611	0.110288
F	30.328653	-1.456575	0.062332

2.2. Part Salience Characterization

2.2.1. Protrusion of A Part

The protrusion of a 3-D part can be quantified as the ratio of the area of the part surface to the area of its base surface [2]. For a 3-D mesh, the base of a part is referred to the minimal surface formed by the boundary polygon of the part. We, therefore, apply the filling holes algorithm described in [4] to form such a base surface (as illustrated in Fig. 4). In this way, the protrusion of a 3-D part represented in the form of a mesh can be obtained.

2.2.2. Strength of A Part Boundary

This factor can be characterized by the turning normals at crease boundaries [2]. For 3-D mesh representation, the dihedral angle formed by two adjacent faces has been commonly used in calculating local curvature. We, therefore, quantify the boundary strength as the minimum dihedral angle along the boundary polygon of a 3-D part.

2.2.3. Relative Size of A Part

The relative size of a 3-D part can be quantified as the ratio of the volume of a part to the volume of the whole object [2]. However, calculating the volume of a 3-D shape is a tedious work, especially for an open surface. To simplify this task, we propose to characterize the relative size of a part as an area ratio. Table 1 lists the part salience extracted from the tiger model using the methods described above.

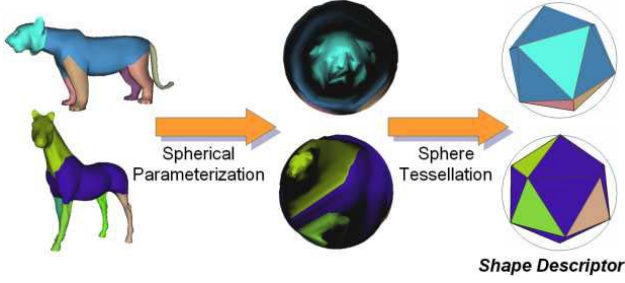


Fig. 5. Illustration of the relation-based shape descriptor construction. The original 3D shape is first parameterized onto the spherical domain. Next, the spherical domain is tessellated as an icosahedron. Based on the resulting icosahedron, the shape descriptor stores the geometrical relations among the parts.

2.3. A Spherical Representation for 3-D Shape

In the previous section, we have described the method of how the set of human perception-related features was quantified. However, the set of features is still not enough to achieve a “thorough” description of a 3D shape. As a result the retrieval task based on the above mentioned feature set will not receive a satisfactory result. Under these circumstances, one has to “add” some relational feature such as relative position of parts, into the feature set so that the description of a 3D shape can be more precise. In what follows, we shall use the so-called spherical parameterization to represent the geometrical relations among the parts of a 3D shape. With this parameterization, a part of a 3D shape will correspond to a point on the sphere after performing spherical parameterization. In our investigation, we found that the spherical parameterization algorithm proposed by Praun and Hoppe [6] is very suitable for this task because their algorithm can generate a valid sphere embedding with less undersampling effect. After the spherical parameterization process is done, we quantize the spherical domain by applying spherical coordinate tessellation. Based on the tessellated sphere, four shape descriptors are constructed such that part relation and part salience (i.e., protrusion, relative size, and boundary strength) can be recorded, respectively. In our current implementation, the sphere tessellation process is achieved by generating an icosahedron. Fig. 5 shows the results after establishing the relation-based shape descriptors. In [1], Hebert *et al.* used a similar spherical representation scheme to map the curvature distribution of a 3D surface onto a unit sphere, on which a 3D shape dissimilarity metric is defined. Here, in contrast to their intrinsic spherical representation, four shape descriptors are constructed based on part relation, protrusion, boundary strength, and relative size, respectively.

2.4. 3-D Shape Dissimilarity Metric Defined on Sphere

Since the spherical parameterization process creates a one-to-one mapping between the points on the surface of a 3D shape and those on a representation sphere, the shape descriptor mentioned in Section 2.3 can be regarded as a function defined on the representation sphere. It is obvious that some existing mathematical analysis tools (e.g., [7]) can be utilized to perform shape analysis and calculate shape dissimilarity. However, the above mentioned tools do not have any direct correlation with the function of a human visual system. In this paper, we propose to compare the shapes of two distinct targets directly in the proposed spherical domain. The reason is twofold: First, the proposed shape analysis algorithms (including mesh decomposition and part salience characterization) have been explicitly considered a natural way to mimic human visual system. Second, mathematicians usually favor elegant formulation (e.g., frequency decomposition) and beautiful mathematical properties (e.g., invariance) while cognitive psychologists intend to account for whether a proposed mechanism is close to the way a human visual system operates. In this work, we try to build a 3D shape retrieval system that makes both mathematicians and cognitive psychologists feel comfortable. In what follows, we shall describe the proposed dissimilarity metric in more detail.

Let $SD^{prop}(A)$ and $SD^{prop}(B)$ denote the shape descriptors of two mesh-based shapes A and B , respectively. The superscript *prop* indicates certain property of interest (either part relation or part salience) used to perform the retrieval task. Now, the distance between A and B subject to certain spherical rotation R is defined as

$$d(A, B, R, prop) = \|SD_I^{prop}(A) - SD_R^{prop}(B)\|, \quad (1)$$

where $\|\cdot\|$ denotes the L_2 -norm difference over the tessellated sphere and I is the identity matrix. Under these circumstances, the dissimilarity between A and B becomes

$$Dissimilarity(A, B, prop) = \min_R d(A, B, R, prop). \quad (2)$$

Eq. (2) tries to minimize $d(A, B, R, prop)$ over all possible spherical rotations R . From Eqs. (1)-(2), it is apparent that the proposed shape matching algorithm is intended for continuously rotating one of the shapes until the difference of the proposed properties perceived by the human visual system is minimized. Note that the number of rotations for which $d(A, B, R, prop)$ should be evaluated is confined by the facet number of a tessellated sphere. Since both A and B are projected onto the tessellated sphere, the comparisons should be directed to finding the best correlation between the instances of $SD(A)$ and those of $SD(B)$. Moreover, we also construct a set of correspondence tables based on the above mentioned shape descriptor for shape matching.

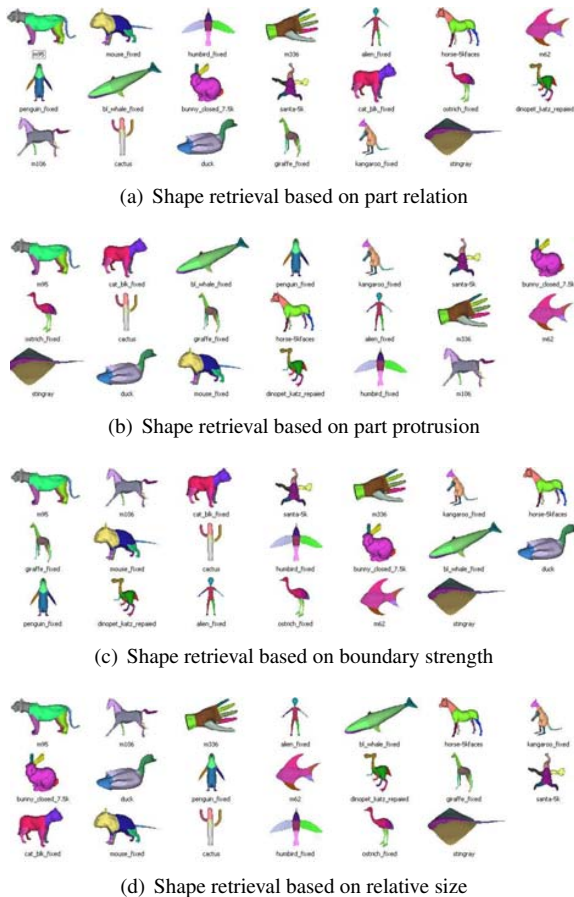


Fig. 6. Results of (a) retrieval by components and (b)-(d) retrieval by visually salient components.

3. EXPERIMENTAL RESULTS

A series of experiments were conducted to test the effectiveness of the proposed method. The database used in our experiments consists of 20 triangulated meshes. In addition, in order to properly apply spherical parameterization, all these models were type of genus-0 and closed meshes. As shown in Fig. 6(a), based on the relation features, those models that possess similar part relation/organization to that of the query ‘tiger’ were retrieved from the databa shown in Fig. 6(b), based on the protrusion features, the models possessing similar part protrusion degree to that of the query were precisely retrieved from the database. Fig. 6(c) shows that with boundary strength-based features, our system could precisely retrieve the model whose parts and the remainder (body) form similar angles (i.e., boundary strength). Fig. 6(d) shows that based on relative size, the models whose parts possess similar area ratio to that of the query could be precisely retrieved from the database.

4. SUMMARY AND FUTURE WORK

The proposed cognitive psychology-based scheme for 3D shape retrieval has two remarkable features: (1) a 3D shape retrieval system intended for mimicking human visual perception and recognition is achieved by incorporating Hoffman and Singh’s theory of part salience in the design of shape analysis algorithms; (2) a coarse-to-fine shape searching scheme is achieved by consideration of part relation and visual salience, respectively. In future work, we shall conduct a more thorough investigation of sphere tessellation approaches for establishing the proposed shape descriptor. Moreover, use of combined features to perform the retrieval task will be the main subject for our future work.

5. REFERENCES

- [1] M. Hebert, K. Ikeuchi, and H. Delingette, “A Spherical Representation for Recognition of Free-Form Surfaces,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 17, No. 7, pp. 681-690, 1995.
- [2] D. D. Hoffman and M. Singh, “Salience of Visual Parts,” *Cognition*, Vol. 63, pp. 29-78, 1997.
- [3] M. Hilaga, Y. Shinagawa, T. Kohmura, and T. L. Kunii, “Topology Matching for Fully Automatic Similarity Estimation of 3-D Shapes,” *Proc. ACM SIGGRAPH*, pp. 203-212, Los Angeles, CA, 2001.
- [4] P. Liepa, “Filling Holes in Meshes,” *Eurographics Symposium on Geometry Processing*, pp. 200-205, 2003.
- [5] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin, “Shape Distributions,” *ACM Trans. on Graphics*, Vol. 21, No. 4, pp. 807-832, Oct 2002.
- [6] E. Praun and H. Hoppe, “Spherical Parameterization and Remeshing,” *ACM Trans. on Graphics*, Vol. 22, No. 3, pp. 340-349, 2003 (Special issue: *Proc. ACM SIGGRAPH 2003*).
- [7] Fast Spherical Harmonic Transforms: SpharmonicKit 2.5 (<http://www.cs.dartmouth.edu/~geelong/sphere/>), 1998.
- [8] H.-Y. S. Lin, H.-Y. M. Liao, and J.-C. Lin, “Visual Salience-Guided Mesh Decomposition,” *Proc. IEEE Int. Workshop on Multimedia Signal Processing*, Siena, Italy, Sept 2004.
- [9] C. Zhang and T. Chen, “Efficient Feature Extraction for 2-D/3-D Objects in Mesh Representation,” *Proc. IEEE Int. Conference on Image Processing*, Vol. 3, pp. 935-938, Thessaloniki, Greece, 2001.