LETTER

Skeleton Modulated Topological Perception Map for Rapid Viewpoint Selection*

Zhenfeng SHI[†,††a)], **Liyang YU**[††b)], **Ahmed A. ABD EL-LATIF**[††c)], *Nonmembers*, *and* **Xiamu NIU**[††d)], *Member*

SUMMARY Incorporating insights from human visual perception into 3D object processing has become an important research field in computer graphics during the past decades. Many computational models for different applications have been proposed, such as mesh saliency, mesh roughness and mesh skeleton. In this letter, we present a novel Skeleton Modulated Topological Visual Perception Map (SMTPM) integrated with visual attention and visual masking mechanism. A new skeletonisation map is presented and used to modulate the weight of saliency and roughness. Inspired by salient viewpoint selection, a new Loop subdivision stencil decision based rapid viewpoint selection algorithm using our new visual perception is also proposed. Experimental results show that the SMTPM scheme can capture more richer visual perception information and our rapid viewpoint selection achieves high efficiency.

key words: *visual attention, visual masking, topological perception, viewpoint selection*

1. Introduction

With the development of hardware and computer graphics, visual perception based 3D object processing technology has attracted many research works, such as mesh saliency, mesh roughness and mesh skeleton.

The visual attention mechanism is one of the most often exploited visual perception characteristics in 3D mesh perception or scene analysis applications. During the last two decades, a particular strategy consisting of two attention mechanisms are developed. One is bottom-up and the other is top-down. Feature Integration Theory (FIT) [1] suggests that visual information is analyzed in parallel from different maps. Itti et al. have maintained that visual attention is saliency-dependent [2]. According to the saliency model proposed in [2], Lee et al. have developed a computational model of mesh saliency [3]. It can capture the visually interested regions on a mesh. They have used saliency map successfully in mesh simplification and viewpoint selection.

Masking is a robust perceptual phenomenon and has been studied by physiologists and psychologists for many years [4]. The Visible Difference Predictor (VDP), as a complex numerical models relying on some psychophysical and physiological evidences like the masking effect, has developed by Daly [5]. In the field of computer graphics, to measure the quality of a watermarked mesh, Corsini et al. [6] propose a perceptual metric based on global roughness variation. Lavoué has developed a computational model linked with the concept of visual masking and named it as mesh roughness [7]. This concept of roughness is quite relevant to 3D perception.

The theories of topological perception and global precedence developed by Navon [8] indicate that global perception is prior to local perception. Recently, some experiments conducted by Chen et al. [9] have been approved by many researchers in the field of cognition. Skeleton of 3D object can describe the topological feature and visual shape. Topological perception might benefit from the technology of skeleton extraction. Oscar et al. [10] present a simple and robust geometric contraction method which works directly on the mesh domain. Inspired by the work [10], Cao et al. [11] use Laplacian-based contraction to successfully extract a curve skeleton. Other comprehensive survey of curve-skeleton extraction is presented in [12].

Our goal in this paper is to bring multiple mechanism of HVS to a unified perceptual model and apply it to 3D object general-purpose processing. Mesh saliency and roughness belong to local perception because they only model the stimuli from local window or region. According to the theory of topological perception, our goal is to integrate the skeleton with mesh saliency and roughness. Here, skeleton is used to modulate the final visual perception property composed of saliency and roughness features. Viewpoint selection application will be used to validate the effectiveness of the proposed unified perception model in this paper.

2. Skeleton Modulated Topological Perception Map

2.1 Algorithm Overview

In this paper, we present a novel robust visual perception map named as Skeleton Modulated Topological Perceptual Model (SMTPM) for 3D object processing. In order to understand the scheme more clearly, we overview it systematically. Figure 1 illustrates the workflow of our proposed SMTPM.

Firstly, curvature estimation. Based on Normal Cy-

Manuscript received December 5, 2011.

[†]The author is with the Department of Mathematics, Harbin Institute of Technology, No.92, West Da-Zhi Street, Harbin, China.

[††]The authors are with the School of Computer Science and Technology, Harbin Institute of Technology, No.92, West Da-Zhi Street, Harbin, China.

*This work is supported by the National Natural Science Foundation of China (Project Number: 60832010, 61100187), the Fundamental Research Funds for the Central Universities (Grant No. HIT. NSRIF. 2010046).

a) E-mail: szf@hit.edu.cn
b) E-mail: liyang.yu@ict.hit.edu.cn
c) E-mail: ahmed_rahiem@yahoo.com
d) E-mail: xm.niu@hit.edu.cn

DOI: 10.1587/transinf.E95.D.2585

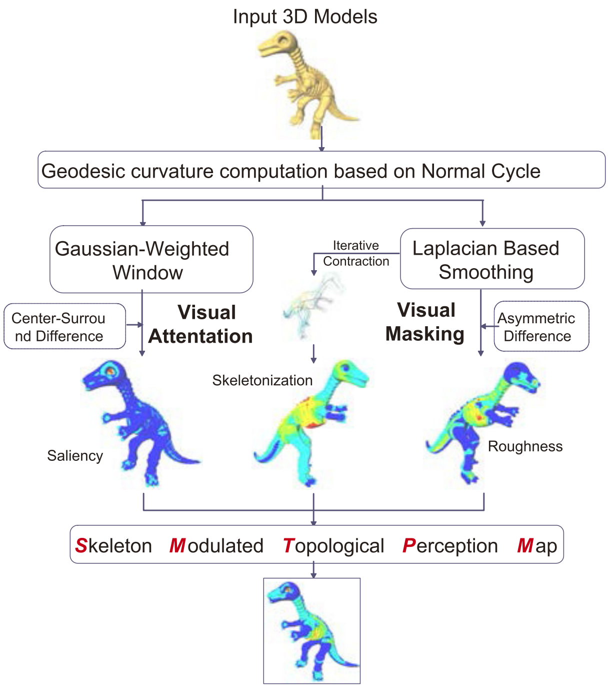

Fig. 1 Overview of our proposed scheme.

cle [13], we implement the estimation of curvature tensor as [7] for each vertex of the mesh. Then, we extract the principal curvature values kmin and kmax.

Secondly, saliency and roughness computation. Based on the framework of computational model of mesh saliency [3], we use the mean curvature generated from curvature estimation as a substitute for Taubin mean curvature [3], then the mesh saliency is computed. Mesh roughness estimation is the same as [7].

Thirdly, skeletonisation map computation. We implement the skeleton extraction based on iterative contraction presented by Oscar et al. [10] and Cao et al. [11]. Based on the skeleton, we compute the shortest distance from each vertex of mesh to the skeleton.

Finally, feature map fusion. We use a new fusion style based on three maps (mesh saliency, mesh roughness and mesh skeletonisation) to define the SMTPM.

2.2 Skeletonisation Map Computation

Let the final curve-skeleton of M_0 with N vertices $V = \{v_i | v_i \in R^3, 1 \leq i \leq N\}$ be $S = Skel(M_0)$, then contraction based curve-skeleton extraction aims to design transformation C for extracting the skeleton quickly. Laplacian-based contraction provide a rapid and effective transformation for extracting a final curve-skeleton. In this paper, we use the Laplacian-based contraction described in [11].

For any vertex $v_i \in V$, we use Eq. (1) to denote the minimal distance from the jth new contracted position $v_i^{(j)}$ to the final curve-skeleton.

$$d_i^{(j)} = \min\left\{\left|v_i^{(j)} - v_{skel}\right| v_{skel} \in S = Skel(M_0)\right\}, \tag{1}$$

where $j = 0, 1, 2, \cdots$. During the extraction of curve-skeleton, the distance series $\left\{d_i^{(j)}\right\}_j$ of vertex v_i can be obtained. We map $\left\{d_i^{(j)}\right\}_j$ to the corresponding vertices of the original mesh. In this paper, we use the original mesh to calculate the minimal distance and define the skeletonisation map.

2.3 Definition of the SMTPM

Mesh saliency and mesh roughness mainly focus on features of local regions and can not well illustrate the global topological perception of 3D object. Mesh skeleton describes the main information of the overall topological shape and visual shape. Therefore, we incorporate them together into a new perception map for 3D object processing.

Let S, R and B be the saliency map, roughness map and skeletonisation map of a 3D mesh M respectively. Supposed that mesh M has N vertices $V = \{v_i | v_i \in R^3, 1 \leq i \leq N\}$. We first normalize each map to unit interval [0, 1] and use $S(v_i)$, $R(v_i)$ and $B(v_i)$ to denote each perceptual feature of vertex v_i respectively. Then, the final skeleton modulated topological perception of each vertex v_i is defined as the Eq. (2):

$$\begin{aligned} &SMTP(v_i) \\ &\quad = ((1 - B(v_i)) * S(v_i)^{\alpha} + B(v_i) * R(v_i)^{\alpha})^{1/\alpha}, \end{aligned} \tag{2}$$

where α is greater than zero.

3. SMTPM Based Rapid Viewpoint Selection

3.1 Definition of the Best and Worst Viewpoint

The term good view in computer graphics is difficult to define precisely. Up to now, there is no consensus about what a good viewpoint is. However, it seems that the best viewpoint is the one that obtains the maximum information of a scene. Therefore, a good viewpoint must help user to capture as much information of the object or scene represented as possible.

In this paper, we develop a method for automatically selecting a good viewpoint so as to maximize the sum of features of visible regions of the object according to SMTPM. For a given viewpoint vp, supposed that $F(vp)$ is the set of vertices visible from vp, and SMTPM is the feature map of 3D mesh. Then, we compute the sum of visible SMTPM from vp as: $SMTPM_{sum}(vp) = \sum_{v \in F(vp)} SMTPM(v)$.

Then, the best viewpoint vp_{best} and the worst viewpoint vp_{worst} satisfy Eq. (3):

$$\begin{aligned} &SMTPM_{sum}(vp_{best}) = \underset{vp}{argmax}\{SMTPM_{sum}(vp)\} \\ &SMTPM_{sum}(vp_{worst}) = \underset{vp}{\arg\min}\{SMTPM_{sum}(vp)\} \end{aligned} . \tag{3}$$

3.2 Iterative Subdivision Based Rapid Viewpoint Selection

In this paper, we design an iterative based stepwise refinement method to help us select the best and worst viewpoints

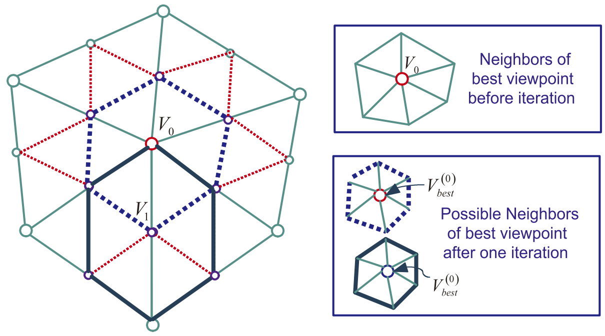

Fig. 2 Loop subdivision stencil based iterative viewpoint selection.

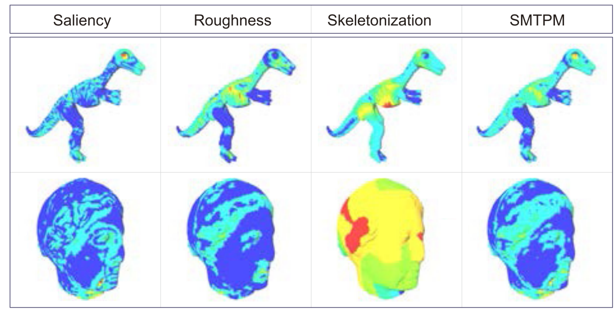

Fig. 3 Different Maps for 3D meshes: dinosaur (5025 vertices).

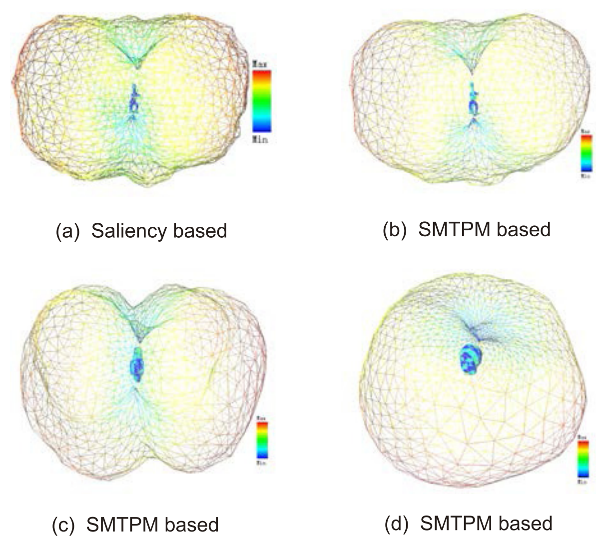

Fig. 4 The wireframe mesh around each model shows the magnitude of the visible saliency or SMTPM sum when we see it from each direction.

rapidly. The following illustrates the algorithm of the best viewpoint selection. It is similar to the worst case.

Firstly, we sample some vertices over viewpoint space.

Secondly, we compute the best viewpoint from the N initial viewpoints. The final best viewpoint might be located in the 1-ring neighbor region N_{v_0} if the v_0 vertex is the best viewpoint for all N initial viewpoints.

Thirdly, we subdivide each of the neighbor triangles in N_{v_0} into four sub-triangles. Each newly generated vertex can be achieved by using Loop subdivision stencil [14].

Fourthly, we use $\tilde{N}_{v_0}$ to denote the new 1-ring neighbors which are composed of those new vertices. The SMTPM value of each new vertex in $\tilde{N}_{v_0}$ is computed. If the SMTPM values of all vertices are less than v_0, we can take v_0 as the new maximum iterative viewpoint. Otherwise, we take v_1 with the maximum SMTPM as the new viewpoint v_0 described in the second step. Figure 2 shows the subdivision rules and the iterative viewpoint selection principles used in the third and the fourth steps.

Finally, let the current best viewpoint be the vertex $vp^{(k)}$, and the next possible best viewpoint is $vp^{(k+1)}$. The final best viewpoint can be obtained and the recursive iterative subdivision operation can stop if $\left|SMTPM_{sum}\left(vp^{(k)}\right) - SMTPM_{sum}\left(vp^{(k+1)}\right)\right| < \varepsilon$, where ε is an interactive control precision.

4. Experimental Results and Analysis

4.1 Skeleton Modulated Topological Perception Map

In this paper, we use skeletonisation map as a topological perception expression to modulate the whole visual perception information. Based on the skeleton extraction work by [11], Fig. 3 shows the SMTPM and skeleton for 3D objects: dinosaur and venus. The SMTPM provides much important visual perception information for 3D mesh. For all experiments in this paper, we configure the same ε as suggested in [3] and the same local window size r as mentioned in [7], the parameter $\alpha = 3.0$ is selected for Eq. (2). We can conclude that the salient and rough region is retained while these three feature maps are integrated together. Therefore, the visual attention and visual masking mechanism are well modeled with our proposed scheme. Figure 3 demonstrates that the SMTPM is effective for describing human visual perception information.

4.2 The Best Viewpoint Selection

To validate SMTPM for viewpoint selection, we implement mesh saliency and salient viewpoint selection presented in [3]. We only show the best viewpoint selection in this paper. Figure 4 shows the magnitude of the visible saliency and the SMTPM sum when we see it from each direction. The warmer colors(reds and yellows) in Fig. 4 show that we can capture more visual perceptual information from these viewpoints. On the contrary, the cooler colors(greens and blues) show that we can obtain less information when locating viewpoint at these vertices.

In order to further compare the saliency based with the SMTPM based viewpoint selection clearly, Fig. 5 illustrates the best viewpoint for 3D objects: dinosaur (5026 vertices), rock-arm (10000 vertices) and venus (33591 vertices). The first row shows the results of best viewpoint selected from mesh saliency [3] and the second row is the results based on SMPTM. Obviously, SMTPM based best viewpoint is more effective than the saliency based. In Fig. 5, the stomach of

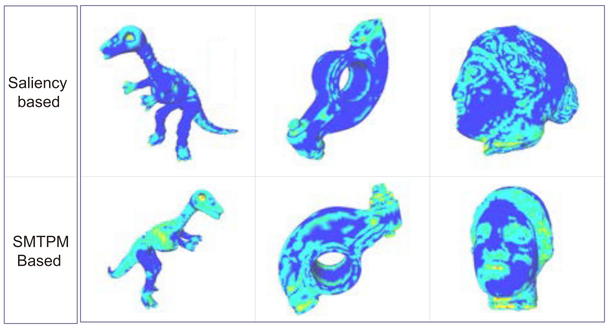

Fig. 5 Selected best viewpoints of 3D object.

Table 1 Processing time of saliency based and SMTPM based best viewpoint selection. (Unit: second)

3D Model	Rocker	Venus	Horse	Dinosaur	Armadillo
#verts	10 K	33 K	48 K	56 K	172 K
Saliency	13.2	46.5	65.1	81.3	241.7
SMTPM	5.3	17.6	28.2	36.7	102.5

dinosaur is occluded by the front left leg in saliency based best viewpoint, but SMTPM based best viewpoint is not the same. The right eye of dinosaur has almost the same salient feature as the left one and the right stomach has more rough regions than the left according to [7]. Therefore, SMTPM based best viewpoint provide more visual perception information than saliency based. For venus object, there are two important scars on the lower jaw, Saliency based viewpoint cannot provide the two important visual information, but SMTPM based viewpoint can well capture them and provide almost the whole front face for observers. Therefore, we can conclude that SMTPM based viewpoint selection has the capability of providing more visual perception information as far as possible.

4.3 Runtime Analysis for Rapid Viewpoint Selection

We conduct experiments for selecting the best viewpoint with Loop subdivision stencil based iterative decision in order to demonstrate the efficiency of our algorithm. The time to compute the best viewpoint depends on the initial viewpoints set. One triangulated 3D cube bounding box with eight vertices is used as the initial viewpoints to be selected. The threshold of stopping iteration is 0.1. We can conclude that the proposed rapid viewpoint selection scheme is very efficient. Table 1 details the processing times for different 3D meshes on 2.2 GHz Pentium PC with double kernel and 4 GB RAM.

5. Conclusions and Future Research

This paper presents a new visual perception map named Skeleton Modulated Topological Perception Map (SMTPM). Based on mesh saliency and mesh roughness, the SMTPM integrate the two important visual perception information with the weight of skeletonisation map. Our proposed scheme can capture richer visual perception information than saliency map. A new rapid viewpoint selection based iterative decision using Loop subdivision stencil is also presented. Benefit from the SMTPM and the rapid iterative subdivision stencil, the best and worst viewpoint are accurately selected. Experimental results demonstrate the efficiency and effectiveness of our scheme. We can also apply the SMTPM to other 3D object processing, such as mesh simplification, mesh segmentation, mesh content authentication.

Acknowledgements

We would like to thank Guillaume Lavoué for the program of roughness map calculation. We also would like to thank all anonymous reviewers for their helpful comments and suggestions, which will significantly improve our paper.

References

[1] A.M. Treisman and G. Gelade, "A feature-integration theory of attention," Cognitive Psychology, vol.12, no.1, pp.97–136, 1980.
[2] L. Itti, C. Koch, and E. Niebur, "A model of saliency based visual attention for rapid scene analysis," IEEE Trans. Pattern Anal. Mach. Intell., vol.20, no.11, pp.1254–1259, Nov. 1998.
[3] C.H. Lee, A. Varshney, and D.W. Jacobs, "Mesh saliency," ACM Trans. Graphics, vol.24, no.3, pp.659–666, 2005.
[4] C. Harmon and B. Julesz, "Masking in visual recognition: effects of two-dimensional filtered noise," Science, vol.180, pp.1194–1197, 1973.
[5] S. Daly, The visible differences predictor: An algorithm for the assessment of image fidelity, pp.179–206, MIT Press, Cambridge, MA, 1993.
[6] M. Corsini, E.D. Gelasca, T. Ebrahimi, and M. Barni, "Watermarked 3d mesh quality assessment," IEEE Trans. Multimedia, vol.9, no.2, pp.247–256, Feb. 2007.
[7] G. Lavoué, "A local roughness measure for 3D meshes and its application to visual masking," ACM Trans. Applied Perception, vol.5, no.4, Article 21, Jan. 2009.
[8] D. Navon, "Forest before trees: The procedence of global features in visual perception," Cognitive Psychology, vol.9, no.2, pp.353–383, 1977.
[9] L. Chen, S. Zhang, and M.V. Srinivasan, "Global perception in small brains: Topological pattern recognition in honey bees," Proc. American National Acacdemy of Science, vol.100, no.11, pp.6884–6889, 2003.
[10] O.K.-C. Au, C.-L. Tai, H.-K. Chu, D. Cohen-Or, and T.-Y. Lee, "Skeleton extraction by mesh contraction," ACM Trans. Graph, vol.27, no.3, Article 44, 2008.
[11] J. Cao, A. Tagliasacchi, M. Olson, H. Zhang, and Z. Su, "Point cloud skeletons via laplacian-based contraction," Proc. IEEE Shape Modeling International, pp.187–197, 2010.
[12] N.D. Cornea, D. Silver, and P. Min, "Curve-skeleton properties, applications, and algorithms," IEEE Trans. Vis. Comput. Graphics, vol.13, no.3, pp.530–548, 2007.
[13] D. Cohen-Steiner and J. Morvan, "Restricted delaunay triangulations and normal cycle," 19th Annu. ACM Sympos. Comput. Geom., 2003.
[14] C.T. Loop, Smooth subdivision surface based on triangle, Master's thesis, University of Utah, 1987.