2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing

A Perceptual Metric Based on Salient Information Entropy for 3D Mesh Distortion

Zhenfeng Shi
School of Computer Science and Technology
Harbin Institute of Technology
150080, Harbin, China
zhenfeng.shi@ict.hit.edu.cn

Qiong.Li
School of Computer Science and Technology
Harbin Institute of Technology
150080, Harbin, China
qiong.li@ict.hit.edu.cn

Xiamu Niu
School of Computer Science and Technology
Harbin Institute of Technology
150080, Harbin, China
xiamu.niu@ict.hit.edu.cn

Abstract **Visual degradation and distortion may be introduced during the 3D mesh simplification. The critical issue in all mesh simplification methods is how to evaluate it. Some geometric metrics without the consideration of human visual perception and perceptual metrics based on subjective test have been proposed. It is urgent to develop an objective evaluation metric with the consideration of human visual perception. We used mesh saliency integrating with Just-Noticeable Difference(JND) methodology to define the salient information entropy based on the theory of information and presented a perceptual metric based on it for measuring the visual degradation and distortion introduced by 3D mesh simplification. Experimental results show that our perceptual metric can evaluate the visual degradation effectively and the essential of 3D mesh simplification is decreasing the visual redundancy and average visual uncertainty.**

Keywords saliency; HVS; perception; distortion; entropy

I. INTRODUCTION

3D triangular meshes are usually used as one of the most popular representations in computer graphics applications. The representation precision depends on the total numbers of triangle vertices and faces. Much more triangles were generated through the use of 3D-scanners, reconstruction from images and other techniques. However, in previous perceptual experiments[1], it was observed that when the number of triangle vertices is larger than a certain minimum number, the extra more detail of geometry is not perceptible to the human visual system(HVS) because of the limit of its resolution[2][3]. Given a certain visual threshold that the HVS cannot be noticed, two 3D meshes which may be different in geometry properties can be visual similarly. In order to produce a noticeable variation in human visual experience the minimum change in perceptual value is stated by Weber's Law[4]. The Just-Noticeable-Difference (JND) is computed as the variation relative to the original value, and it is the certain visual threshold that maybe noticeable minimum variation by which the stimulus intensity must be changed in order to be noticeable to human perception[5]. Any stimulus greater than the threshold will be noticed by human visual system. In 3D object application, an alternative usage to represent the stimulus is proposed by I.Cheng[5]. Experiments show that the alternative approach based on JND presents an accurate estimation of visual quality for evaluating the performance of different simplification algorithms, but it is experimental.

Many mesh simplification, LOD and other multi-resolution algorithms have been proposed in the last twenty years[6][7][8][9]. They usually simplify 3D triangular mesh either by face merging or by incremental decimation or by remeshing and aimed to transform the given triangular mesh into a new mesh with fewer vertices, faces and edges, but with the same shape and visual effect. These algorithms or technologies differ especially in the distortion metrics to guide the simplification process. However, these simplification operators maybe introduce some visual degradations. The common key problem in all mesh simplification algorithms is how to evaluate it. Geometric measurement without taking human visual perception into consideration is the most popular metric in these existing algorithms. Since the 3D mesh is seen by human beings, some metrics based on human visual perception were proposed and demonstrated to be more practical and accurate. However, there is no systematic and theoretical method for evaluating the visual or perceptual distortion.

In 3D mesh perception or scene analysis applications one of the most often exploited visual perception attributes is visual attention. Saliency is a very important feature in low-level human visual system cues. Koch and Ullman suggested a model that salient image locations will be distinct from their surroundings[10]. Itti et.al maintain that visual attention is saliency-dependent[11]. Recently, Lee et.al developed a model of mesh saliency that be considered as a perception-inspired measure of regional importance from surface curvatures using center-surround filters with Gaussian-weighted curvatures[12]. M. Feixas et.al proposed a new definition of mesh saliency based on polygonal mutual information (PMI) and built a unified information-theoretic framework for viewpoint selection and mesh saliency[13].

In this paper, we defined the salient information entropy which can be used to describe the visual redundancy and the average of visual uncertainty. A perceptual metric SIDM(Salient Information Distortion Metric) based on the salient information entropy and used to evaluate the visual degradation and distortion introduced by mesh simplification. Experiment results show that SIDM can evaluate the visual degradation and distortion. We also

978-0-7695-4222-5/10 $26.00 © 2010 IEEE
DOI 10.1109/IIHMSP.2010.143

find that the essential of mesh simplification is decreasing the visual redundancy and average visual uncertainty.

II. THE SALIENT INFORMATION ENTROPY

Maximizing the visual similarity or perceptual similarity and minimizing the visual degradation are the goal of all mesh simplification methods. However, fewer consider directly with the characteristics of HVS, such as visual attention, visual masking, and visual adaptation. A challenging problem is how to describe the visual redundancy and define the “maximum” data for a given 3D mesh without visual degradation. To address this problem, some efficient tools or algorithms were provided to measure the loss quality or the visual differences between the original 3D object and the simplified objects because these objects are often to be seen by human beings. In this section, we will give a perceptual distortion metric based on salient information entropy for visual degradation.

A. *Mesh Saliency Computation*

We used the mesh saliency algorithm based on Gaussian-weighted center-surround mechanism described in[12].

B. *Mesh Saliency with JND*

According to Webber’s Law, we define $\Psi_{\max}(s)$ as the maximum perception distance of the original mesh and $\Delta\Psi(s)$ as the perception change when removing one vertex from the original 3D mesh. Suppose the original saliency value s of the 3D mesh is in $[a,b]$, Obviously, $\Psi_{\max}(s)=b$. We introduce the JND threshold $K=0.1$ adopted by[14] to Webber’s equation. Then we get the following expression

$$0.1=K=\frac{\Delta\Psi(s)}{\Psi_{\max}(s)}=\frac{\Delta\Psi(s)}{b} \tag{1}$$

According to (1), we get the saliency value corresponding to the JND. Then, the JND saliency value s_{JND} can be deduced as following:

$$s_{JND}=b*0.1 \tag{2}.$$

Due to the limit of the HVS some saliency value less than a certain value s_{JND} is not perceptible. In other words, any change over these mesh vertices with saliency value greater than or equal to s_{JND} will be perceptible to the HVS under the human visual perception threshold.

C. *Salient Information Entropy*

The salient information entropy is a kind of information entropy with the consideration of human visual perception. Given a saliency map of the 3D mesh M with C_v vertices, for convenience, it is desirable to normalize the saliency values to the interval $[0,1]$. First, the mean saliency is estimated with (3)

$$\mu_s=\frac{1}{C_v}\sum_{i=1}^{C_v}s_i \tag{3}$$

Second, we compute the square root of variance which is given by

$$\sigma_s=\left(\frac{1}{C_v}\sum_{i=1}^{C_v}\left(s_i-\mu_s\right)^2\right)^{\frac{1}{2}} \tag{4}$$

Finally, we map the s_i from interval $[a,b]$ to s_i' in the interval $[0,1]$ using the following non-linear function

$$s'=\frac{1-e^{-\frac{(s-a)}{2*\sigma_s}}}{1-e^{-\frac{(b-a)}{2*\sigma_s}}},\ s\in[a,b] \tag{5}$$

Substituting the s_{JND} from (2) into (5), we can get the normalized JND saliency value s'_{JND}:

$$s'_{JND}=\frac{1-e^{-\frac{(s-a)}{2*\sigma_s}}}{1-e^{-\frac{(b-a)}{2*\sigma_s}}}=\frac{1-e^{-\frac{(b*0.1-a)}{2*\sigma_s}}}{1-e^{-\frac{(b-a)}{2*\sigma_s}}} \tag{6}$$

In (6), $s'_{JND}\in[0,1]$. When s_i' is less than the s'_{JND}, the extra detail is not perceptible to the HVS. Any 3D mesh vertex with saliency value s_i' under the JND threshold value s'_{JND} will not contribute to the perceptual or visual information. So the typically transformed saliency value interval $[0,1]$ is too large for our perceptual purposes. More appropriate is the saliency values in the sub-interval $[s'_{JND},1]$. Hence, we let $s_i'=0$ if $s_i'<s'_{JND}$. For a given 3D mesh the saliency map can be thought as a kind of perceptual or visual information distribution, and the transformed saliency values s_i' are interpreted as random variable series. Then we use N which may be equal to 2% of the numbers of vertices or more than 2% to subdivide the saliency s' sub-interval $[0,1]$ to N equational sub-ranges called as $\Re_i'$ denoted by following interval

$$\Re_i'=\left[\frac{1}{N}(i-1),\frac{1}{N}i\right), \tag{7}$$

where $1\le i\le N-1$ and $\Re_N'=\left[\frac{N-1}{N},1\right]$. The perceptual or visual information distribution probabilities

p_i are computed from the total sum of saliency values that s' values will fall into sub-range $\Re_i'$ by (8)

$$p_i = \frac{\sum_{s_k' \in \Re_i'} s_k'}{\sum_{s_j' \in [s'_{JND},1]} s_j'}, \quad 1 \le i \le N, \tag{8}$$

We call the vector $(p_1, p_2, \cdots, p_N)$ or for short $\langle p_i \rangle$ as the salient information probability distribution.

The salient information entropy H_S of a 3D mesh M is a function of JND threshold K and mesh M and counts of sub-intervals N, and it can be finally computed via (9)

$$H_S = H_S(M, K, N) = \sum_{i=1}^{N} p_i \log \frac{1}{p_i} \tag{9}$$

D. Salient Information Distortion Metric(SIDM)

We use the salient information distribution from equation (8) and the saliency entropy (9) to define the salient information distortion metric(SIDM). Let the source salient information probability distribution is $\langle p_i \rangle$ and the target is $\langle q_j \rangle$, then the salient information distortion over the finite N sub-intervals data value is defined as

$$SIDM(p,q) = \left\| \frac{H_s(p) - H_s(q)}{H_s(p)} \right\| = \left\| 1 - \frac{H_s(q)}{H_s(p)} \right\| \tag{10}$$

where $H_s(p)$ *is the salient information entropy based on the distribution* $\langle p_i \rangle$ and can be computed by (9). *Obviously,* $0 \le SIDM \le 1$ *with* $SIDM = 0$ *if and only if* $p \equiv q$ *and with* $SIDM = 1$ *if and only if* $H_s(q) = 0$. So, $SIDM$ provide an excellent distortion metric for *evaluating the visual degradation or distortion between two meshes with different resolutions but from the same original mesh.*

III. Experimental Results and Analysis

In order to obtain some 3D triangular meshes with different resolutions, we use the open source software MeshLab V1.1.1 to generate some triangular meshes through the quadric edge collapse decimation filter. For an initial 3D triangular mesh, we decimate the new triangular with step down by 5%. Then, including the initial source mesh we obtain twenty 3D meshes with different resolution.

A. Salient Information Entropy Vs. Distortion

We choose the 3D triangular mesh (dinosaur) and 3D face triangular mesh (mannequin) to verify the relation between salient information entropy and distortion introduced by mesh simplification operators described in figure 2. In the dinosaur model we choose the sub-interval numbers $N = 1124$ which is 2% of the number of vertices for the source mesh. JND threshold K is introduced in the processing of normalization, and normalized saliency threshold s'_{JND} can be obtained by using the (6). The probability of some vertices with greater saliency value is very small, the stimulus variation generated to the human retina is not enough to capture the perception variation from statistical view. In other words, fewer vertices with higher saliency value will generate more stimulus than its original status and contribute more perceptual information after the normalization operator. Figure 1 gives the histogram distribution for original saliency map and non-linear normalized saliency map. From the (a) and (b) or (c) and (d) in figure 1 we observed that non-linear normalized operator makes the histogram distribution well-proportioned.

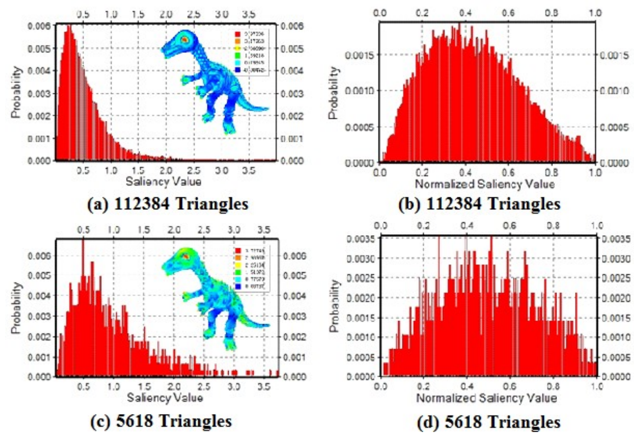
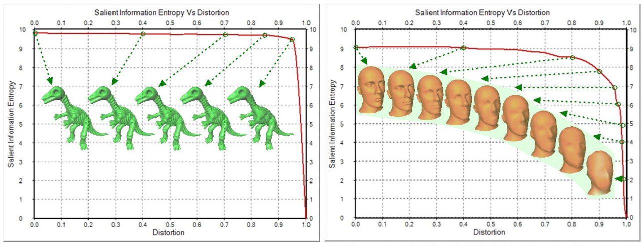


Figure 1. Salient information histogram with parameter $N = 1124$.

We can calculate the salient information entropy after the normalization process. Figure 2 shows that the salient information entropy will hold the line or introduce very tiny change in a certain normalized sub-interval during the mesh simplification. In this paper, we use the salient information entropy to express the visual or perceptual average uncertainty. According to figure 2 we know the case that during a certain data processing the perceptual average uncertainty will keep invariable. In other words, a certain 3D mesh simplification processing in the prior stage will not introduce perceptual variation, or the difference is not noticeable to HVS. Some perceptual redundancy can be removed in this stage. In 3D face triangular mesh we choose the sub-interval numbers $N = 674$ which is 10% of the number of vertices of the source face model. The same result can be obtained and figure 2 (b) illustrated it.

From the figure 2 we observed that an appropriate number N or the percent of number of vertices of the source mesh is very important in order to obtain the perceptual redundancy range. The percentage of the source mesh can be choosing smaller with the increasing number of the mesh vertices. Bigger percentage is needed to select for the mesh with fewer vertices.

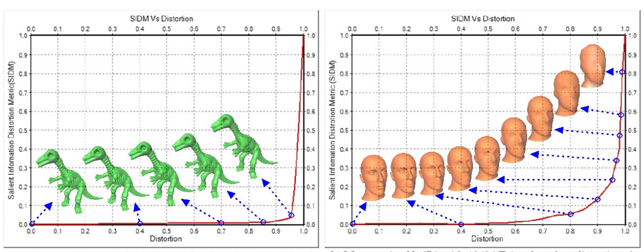


(a) Dinosaur: N=1124, with 112384 Triangles when distortion=0. (b) Mannequin: N=674, with 13424 Triangles when distortion=0.

Figure 2. Relation between salient information entropy and distortion introduced by mesh simplification.

B. *SIDM Vs. Distortion*

How to evaluate the similarity between two 3D meshes is very important in many applications. In this paper, our experiments based on the relative salient information similarity show that it is a very efficient tool to evaluate the simplification distortion. In figure 3, when we remove the 80% of vertices number of the dinosaur model SIDM gives an very good result that the similarity is far less than 0.05. The SIDM value is less than 0.05 even if we removed about over 95% of vertices. Obviously, these dinosaur models are very similar with each other from human visual perception. So, SIDM is very effective for mesh distortion metric.



(a) Dinosaur: N=1124, with 112384 Triangles when distortion=0. (b) Mannequin: N=674, with 13424 Triangles when distortion=0.

Figure 3. Salient information Distortion Metric

IV. CONCLUSIONS AND FEATURE WORK

In this paper, we have developed a model for 3D mesh salient information entropy which takes the consideration with mesh saliency, the JND methodology and the theory of information. At the same time we discussed its computation and the relation between the entropy and the distortion introduced by mesh simplification. We also developed a very important and efficient but simple distortion metric for 3D meshes with different resolutions derived from one source object. The application of our salient information entropy to express the essential of mesh simplification has given us very effective results. Mesh simplification is the process of decreasing the visual or perceptual redundancy of 3D mesh. In a certain range the salient information entropy will hold the line or keep almost invariable. The stimulus variation in this range generated to the human retina is not enough to be noticeable for HVS. Figure 2 can illustrate it effectively.

Mesh visual or perceptual information promised many further researches. We are currently defining salient information entropy using the mesh saliency and JND. However, human attenuation is only one characteristic of HVS, some other important characteristic such as visual masking and visual adaption should be considered for mesh visual or perceptual information. It should be possible to improve this by developing better visual or perceptual information entropy and better metrics with consideration of multi-features of HVS.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China (Project Number: 60703011, 60832010, 60671064), the Chinese national 863 Program (Project Number: 2007AA01Z458), The Research Fund for the Doctoral Program of Higher Education (RFDP: 20070213047) and the Fundamental Research Funds for the Central Universities (Grant No.HIT.NSRIF.2009050).

REFERENCES

[1] Y. Pan, I. Cheng and A. Basu, "Quantitative Metric for Estimating Perceptual Quality of 3D Objects," IEEE Trans. on Multimedia vol.7, no.2, 2005, pp. 269–279.

[2] I. Cheng and A. Basu. "Reliability and judging fatigue reduction in 3D perceptual quality estimation", Proceedings of the 2nd International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'04), 2004.

[3] S. Krinidis and I. Pitas, "Statistical Analysis of Human Facial Expressions", Journal of Information Hiding and Multimedia Signal Processing, Vol.1, No.3,, July 2010, pp.241-260.

[4] R. Gonzalez and R. Woods, "Digital Image Processing," second edition, 2002, Prentice Hall.

[5] I. Cheng, R. Shen, X.D. Yang and P. Boulanger, "Perceptual Analysis of Level-of-Detail: The JND Approach". Proceedings of the Eighth IEEE International Symposium on Multimedia (ISM'06), 2006.

[6] M. Garland and PS. Heckbert, "Surface simplification using quadric error metrics". ACM SIGGRAPH, 1997, pp. 209–216.

[7] J. Cohen, A. Varshney, D. Manocha, G. Turk and H. Weber, "Simplification Envelopes", In Proceedings of SIGGRAPH, 1996, pp. 119–128.

[8] H. Hoppe, "Progressive Meshes," In Proceedings of SIGGRAPH, 1996, pp. 99–108.

[9] M. Garland, A. Willmott and P. Heckbert, "Hierarchical face clustering on polygonal surfaces". ACM Symposium on Interactive 3D Graphics, 2001, pp. 49–58.

[10] C.Koch, and S. Ullman, "Shifts in selective visual attention: towards the underlying neural circuitry". Human Neurobiology 4, 1985, pp. 219–227.

[11] L. Itti, C. Koch, and E.Niebur, "A model of saliency based visual attention for rapid scene analysis". IEEE Transactions on Pattern Analysis and Machine Intelligence vol.20, no.11, 1998, pp. 1254–1259.

[12] C. H. Lee, A. Varshney, and D. W. Jacobs, "Mesh saliency", ACM Trans. Graph. Vol.24, no.3, 2005, pp. 659–666.

[13] M. Feixas, M. Sbert, and F. Gonz´ alez, "A unified information-theoretic framework for viewpoint selection and mesh saliency". ACMTrans. Appl. Percpt. 6, 1, Article 1 (February 2009).

[14] I. Cheng and P. Boulanger, "A 3D Perceptual Metric Using Just-Noticeable-Difference", Eurographics 2005, Aug/Sep Dublin Ireland.