A SALIENCY DETECTION BASED METHOD FOR 3D SURFACE SIMPLIFICATION

Yitian Zhao, Yonghuai Liu, Ran Song and Min Zhang

Deparment of Computer Science, Aberystwyth University,UK

ABSTRACT

To accelerate the processing for the integration, registration, representation and recognition of point clouds, it is of growing necessity to simplify the surface of 3-D models. Simplification is an approach to vary the levels of visual details as appropriate, thereby improving on the overall performance of applications. This paper proposes a saliency detection based points sampling method for mesh simplification. By generating and enhancing the saliency map, the regions which are visually important can be located. For the mesh simplification, the local details are captured by the saliency, while for the overall shape, the approach voxelizes the model and samples points in terms of the entropy of the shape index of vertices in voxels. We present a number of results to show that the method significantly simplifies the surface without distortion and loss of local details.

Index Terms— simplification, saliency, surface, mesh, vertex.

1. INTRODUCTION

Laser range scanners have become more and more popular for 3D measurement. The output of laser range scanners is a set of structured data points with or without reflectance strength information, depicting the reflectance characteristics of the 3D objects of interest. The structured data points can easily be triangulated and rendered as range images. However, most of the range scanners usually generate huge amounts of data points. A complete model could contain millions of points and often leads to expensive processing time. This creates a number of challenges, for storage, editing and transmission. Thus efficient and accurate simplification of point cloud is necessary and the simplified point cloud has become a powerful alternative to the original data.

In the past years the evaluations of point feature detectors in the matching, recognition or texture classification in the 2D domain have been proposed. Moreover, the techniques of interest points detection is expanded to the human visual system area: Itti et al. [1] created a computational model of saliency-based spatial attention derived from a biologically plausible architecture. The saliency maps for features of luminance, color and orientation at different scales were computed. Hu et al. [2] proposed saliency maps by thresholding the color, intensity and orientation maps. In this method, the histogram entropy thresholding analysis was employed.

Fig. 1. Surface saliency:Image (a) and (c) shows the $buddha$ and $duck$ and image (b) and (d) shows the detected surface saliency

More recently, the selection of interest points is exploited to the notion of saliency in the 3D domain. Lee et al. [3] defined a mesh saliency for the mesh simplification. This approach is using a center-surround operator on the local curvature as the discriminative feature. Pauly et al. [4] presented a multi-scale technique for extracting line-type features on a point-sampled geometry. A measure of surface variation and persistence of feature-points over different scales was used. Shilane et al. [5] defined distinction as the retrieval performance of a local shape descriptor. After a training phase, the retrieval performance of each descriptor has been evaluated and only the most distinctive are retained. Frome et al. [6] introduced regional point descriptors, which contain 3D shape and harmonic shape context. This approach extracts the points by random sub-sampling the whole set of points. Owing to its efficiency of visual persuasion in traditional art and technical illustrations, visual saliency has now been widely used in many computer graphics applications, including saliency guided shape enhancement [7]. Miao et al. [8] proposed a normal perturbation technique to enhance the visually salient features of 3Dshapes explicitly. This method demonstrate that saliency guided shading scheme can improve the depiction of the underlying shape and the perception of its salient features.

This paper proposes a novel sampling method for 3D surface simplification based on the surface saliency. Fig. 1 shows the surface saliency on $buddha$ and $duck$ with "warm colors" (yellow and red). We propose a hybrid approach which combines the saliency detection and saliency enhancement. The proposed method first calculates the surface saliency then enhances the saliency map using the Retinex theory [9] which is based on the human visual system. As the human visual attention is found to be useful for various vision processing tasks, now it is becoming more and more important in computer vision research. The saliency region can be used to find the objects that are important in 3D surface. The benefits of saliency detection in the 3D domain include mesh simplification, registration, segmentation, compression, etc. The local geometry of the surface is changed by most of the existing simplification methods but not by our method. The surface details will be detected as salient regions and represented with more points, while the overall geometry will be represented with uniformly sampled points. Experiments are presented in this paper to demonstrate the effectiveness of the proposed method.

2. SALIENCY DETECTION

2.1. Saliency detection

In this section we propose an novel algorithm for the 3D surface saliency detection. Let I be a point set with dimension $3 * W$. Denotes $I = \{P_i\}$ and P_1, P_2 are two points position in I, where $P_1 = (x_1, y_1, z_1)$ and $P_2 = (x_2, y_2, z_2)$. The saliency map can be computed as follows:

- Apply a Gaussian filter on the input points set.
- Estimate the shape index ϑ, Gaussian curvature K and mean curvature H, where $\vartheta = \frac{2}{\pi} \arctan \frac{k_2+k_1}{k_2-k_1}$, $K = k_1 * k_2$ and $H = \frac{k_1+k_2}{2}$. k_1 and k_2 are the principal curvatures which are estimated during the method in [10].

- Set all elements of the Saliency Map SM to 0.
- Start updating saliency values:
  n_1 is a random number where $1 \le n_1 \le W$, n_2 is a random number where $1 \le n_2 \le W$, the saliency can be estimated as follows:

$$SM_{n_2} = SM_{n_2} + S(\vartheta_{n_1}, \vartheta_{n_2}) \tag{1}$$

  where $S(\vartheta_{n_1}, \vartheta_{n_2})$ is the attention value of R_{n_1} due to R_{n_2} is given by the function

$$S(\vartheta_{n_1}, \vartheta_{n_2}) = \frac{|\vartheta_{n_1} - \vartheta_{n_2}|}{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}} \tag{2}$$

- Applying the following steps to each of ϑ, K and H and generate saliency maps SM_ϑ, SM_K and SM_H:
  $SM_\vartheta = SM(\vartheta, n_1, n_2)$
  $SM_K = SM(K, n_1, n_2)$
  $SM_H = SM(H, n_1, n_2)$

- Apply Median filter on each saliency map.
- Generate final combined saliency map:

$$SM = \sqrt{{SM_\vartheta}^2 + {SM_K}^2 + {SM_H}^2} \tag{3}$$

2.2. Retinex based saliency enhancement

Most 3-D images are captured by cameras and scanners, thus they usually contain noise from various sources or the images are visually undesirable. Retinex theory deals with the removal of unfavourable illumination effects from images in order to improve their quality. The theoretic foundation of the retinex is that an image $I(x, y)$ is regarded as a product $I(x, y) = L(x, y) * R(x, y)$, where $L(x, y)$ is the illumination image and $R(x, y)$ is the reflectance image. Generally, $L(x, y)$ is determined by the illumination source and $R(x, y)$ is determined by the characteristics of the imaged object. This paper presents a Retinex theory based bilateral filter[11] to enhance the saliency map. L is estimated by a shift invariant Gaussian filter. The bilateral filtering can be described as follows:

$$h(x) = k^{-1}(x) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_n) g(x_n, x) s(f(x_n), f(x)) dx_n \tag{4}$$

where the normalization

$$k(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x_n, x) s(f(x_n), f(x)) dx_n \tag{5}$$

where $g(x_n, x)$ measures the geometric closeness between the center of neighbourhood x and a nearby pixel x_n and it can be expressed as $g(x_n, x) = e^{-\frac{1}{2}(\frac{d(x_n,x)}{\sigma_d})^2}$ where $d(x_n, x)$is the Euclidean distance between x_n and x, σ_d denotes the geometric spread chosen based on the desired amount of low-passing filtering. $s(f(x_n), f(x))$ measures the photometric similarity between the x and x_n. The similarity function s can be defined as:

$$s(f(x_n), f(x)) = e^{-\frac{1}{2}(\frac{d(f(x_n), f(x))}{\sigma_r})^2} \tag{6}$$

where $d(f(x_n), f(x))$ is the distance between the two shape index values $f(x_n)$ and $f(x)$. σ_d is the photometric spread in the image range that is set to achieve the desired amount of combination of shape index values. By estimating the illumination L, the informative saliency map $R(x, y) = log(I(x, y) + 1) - log(L(x, y) + 1)$ where $I(x, y)$ is the input saliency map SM. As the saliency map which has been estimated in the previous section is view point independent, we transformed it into view point dependent: $SM(i, j) \Leftrightarrow SM(x(i, j), y(i, j), z(i, j))$. $L(x, y)$ is the illumination image.

Fig. 2(b) presents the surface saliency. It can be seen that the regions with more geometry information have been indicated with ”warm colors”. The nose, eyes, mouth and ears are detected on $face$, the bubbles of the $lobster$ are also highlighted.

3. SALIENT POINTS SELECTION AND SURFACE SIMPLIFICATION

In this paper, we use the detected saliency to simplify surface. Our method voxelizes the surface and selects one or more

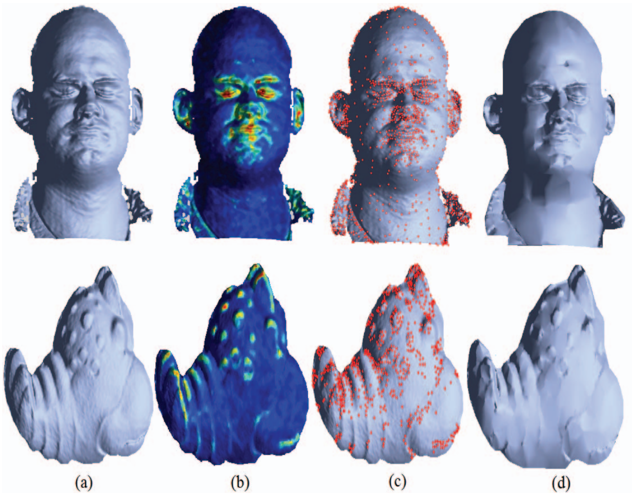

Fig. 2. (a): Input model $face$ and $lobster$.(b): Surface saliency. (c): Selected salient points, 5% points selected. (d): Simplified surface

points from each voxel based on the entropy of the saliency map of vertices and restricted to the original model's surface. The larger the variation, the larger the entropy, the more detail the voxel contains. Thus the entropy can be used to guide the points sampling. If a voxel contains large saliency, more points will be selected. This algorithm consists of four steps:

- The first step voxelizes the model by the bounding boxes and the model can be divided into several sub boxes. The resolution of the voxelization is specified by the user. In this paper, the resolution is 8*8*8.
- Based on the histogram of saliency map, the entropy can be calculated $H(X) = \sum_{i=1}^{n} - P(r_i) * log_2(P(r_i))$ where $P(r_i)$ is the probability of vertex r_i in a local voxel, n denotes the total number of the vertices.
- The third step chooses one or more points from each sub box based on the estimated entropies. The number of selected points is depended on the simplification rate.
- The minimum distances between the samples are used to address the possibility that the generated samples are close to the boundary between two or more adjacent voxels which might be too close to each other.
- In order to achieve a better simplification result, the sampled points which are from the saliency region and non-saliency region are combined. By denoting the points from saliency region as L_s and the points from non-saliency region as G_s, the simplification approach is defined: $SIM = T \cdot L_s + (1 - T)G_s$, where T denotes the weight of the points which are sampled from saliency region and $T \in [0, 1]$. According the evolution of mesh errors with different parameter T in [13], we set $T = 0.65$.

Fig. 2(c) presents the sampled points, it can be seen that more points are sampled in the saliency regions. Fig. 2(d) shows the simplification results with 90% simplification rate, which means 90% points are removed from the original data. It can be observed that the simplified surface still contains local details and overall shape.

4. EXPERIMENTAL RESULTS

In order to evaluate the effect of the simplification criteria, our algorithm (RSim) has been compared with the geometric QSim [12] algorithm which uses the best half-edge collapse and the Retinex theory based mesh simplification (SSim) [13]. The QSim and SSim algorithms were chosen because of the high-quality of its approximations and its code is freely available. Although the QSim algorithm was proposed some time ago, it still achieves competitive results. For the comparative study, a publicly available range image database hosted by the Signal Analysis and Machine Perception Laboratory at Ohio State University was used [14]. To this end, several experiments with meshes of differing complexities were performed. The visual and geometric errors are the differences between the original and simplified mesh, which were measured using root mean squared error (RMSE) and Metro error [15].

Table 1. RMSE ($*10^{-3}$) and Metro Errors measured using QSim, SSim and RSim

	RMSE			Metro		
Model	Qsim	SSim	RSim	QSim	SSim	RSim
face	16.27	14.22	11.63	0.39	0.21	0.16
lobster	25.42	18.21	15.34	0.57	0.42	0.32
duck	15.33	14.21	11.21	0.30	0.23	0.18
buddha	17.85	12.12	10.12	0.58	0.42	0.29
bird	16.22	11.02	9.42	0.41	0.33	0.23

Fig. 3 presents the qualitative results of a 90% model simplification rate using QSim, SSim and RSim, respectively. Table 1 shows the RMSE and Metro errors of QSim, SSim and RSim for all models. From Fig. 3 it can be observed that the simplification results of three methods contain the major topology characteristics of the initial models. Apparently, SSim and RSim for all the models in retaining local details are much better than the QSim, but careful observation at Table 1. shows that the errors of RSim are smaller than the QSim and SSim for both RMSE and Metro mesh error measurements. For example, for the $bird$, the face and neck are over smoothed by QSim, but the characteristics of face and neck are still preserved well by SSim and RSim. The RMSE and Metro errors of RSim have been reduced by 41.92% and 43.9% when compares with QSim. While the corresponding reductive errors compares with SSim are 14.52% and 30.3%. For model $buddha$, the hair and eyebrow are preserved better by SSim and RSim, but the hair is removed completely by QSim, confirmed the fact that, the errors for RSim are much lower than those of QSim. Fig. 3(e)-(h) show snapshots of the small geometry features by preserved simplified surface, RSim preserves the shapes of the wing, neck and ears of the

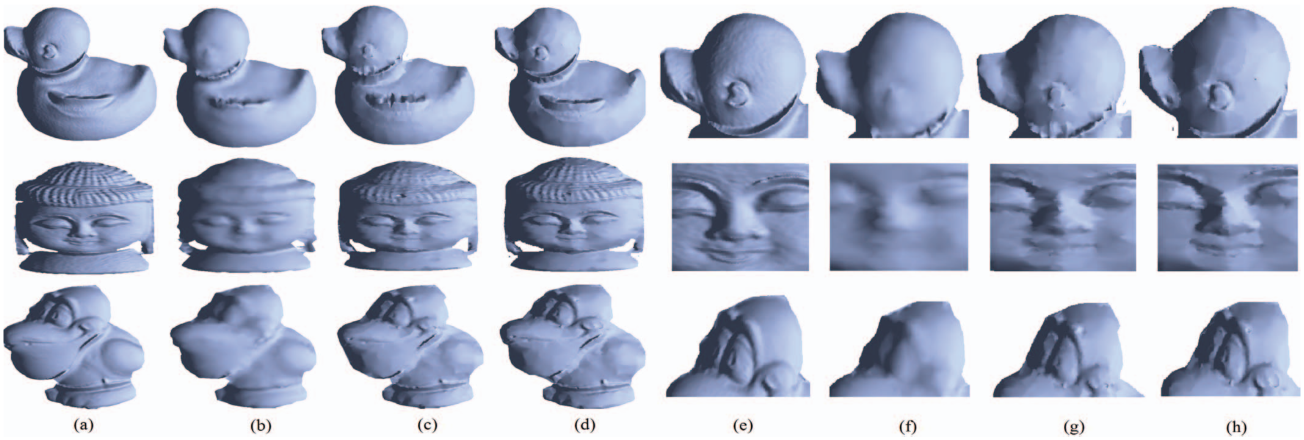

Fig. 3. Simplification by using QSim, SSim and RSim with simplification rate 90% on duck, buddha and bird. (a): Original; (b): QSim; (c): SSim; (d): RSim; (e)-(h): The snapshot of the local features' reservation by using QSim, SSim and RSim respectively

duck much better than SSim and QSim that have blur and distortion. The *bird* surface are preserved well with RSim, but the shape of eyes are over smoothed by QSim. For the surface of the *buddha*, the hair is removed completely by QSim, but the outline of the curls is still preserved well by RSim. The experimental results demonstrated that our method can get better results since they contain more information about the model. All results show that RSim can preserve small geometrical features faithfully.

5. CONCLUSIONS

This paper presents a novel approach for mesh simplification by combining surface saliency detection and saliency enhancement. The method aims at producing the saliency map to simplify the surface, also preserve the local details which are visually important. Compared with the other simplification algorithms such as edge collapse, a number of experimental results have illustrated that the proposed method can get better results since they contain more information about the model, the method not only can represent the overall shape fidelity and but also has the capability to retain the topology and small features well. Our current method of computing saliency takes a not very short time. It should be possible to significantly speed up by some optimization methods.

6. REFERENCES

[1] L. Itti, C. Koch, and E. Niebur,"A model of saliency-based visual attention for rapid scene analysis", *PAMI*, 1998 pp. 1254-1259.

[2] Y. Hu, X. Xie, W. Ma, L. Chia, and D. Rajan," Salient region detection using weighted feature maps based on the human visual attention model", *Springer Lecture Notes in Computer Science*, 2004, pp. 993-1000.

[3] C. Lee, A. Varshney and D. Jacobs," Mesh saliency", *In ACM SIGGRAPH*, 2005.

[4] M. Pauly, R. Keiser and M. Gross," Multi-scale feature extraction on point-sampled surfaces", *Computer Graphics Forum* 22, 2003.

[5] P. Shilane and T. Funkhouser, "Selecting distinctive 3D shape descriptors for similarity retrieval", *In SMI*, 2006.

[6] A. Frome, D. Huber, R. Kolluri, T. Bulow and J. Mailk,"Recognizing objects in range data using regional point descriptors", *In ECCV*, 2004.

[7] Y. Kim and A. Varshney," Persuading visual attention through geometry",*IEEE Transactions on Visualization and Computer Graphics*, 2008, pp.772-82.

[8] Y. Miao, J. Feng and R. Pajarola,"Visual saliency guided normal enhancement technique for 3D shape depiction", *Computers & Graphics*, 2011, pp.706-712

[9] E. Land, "An alternative technique for the computation of the designator in the retinex theory of color vision", *Proc. Natl. Acad. Sci*, 83, 1986, pp. 3078–3080.

[10] G. Taubin, "Estimating the tensor of curvature of a surface from a polyhedral approximation", *ICCV*, 4(2), October 1986, pp. 107–114.

[11] M. Elad, "Retinex by two bilateral filters", *In Proceedings of the Scale-Space conference,*, 3459, 2005, pp. 217–229.

[12] M. Garland and P. Heckbert,"Surface simplification using quadric error metrics", *SIGGRAPH*, 1997, pp. 209-221.

[13] Y. Zhao and Y. Liu and R. Song and M. Zhang "A Retinex theory based points sampling method for mesh simplification", *7th International Symposium on Image and Signal Processing and Analysis* , 2011, pp.230-235.

[14] http://www2.ohio-state.osu.edu/data/3DDB/RID/index.htm

[15] P. Cignoni, C. Rocchini, and R. Scopigno,"Metro: measuring error on simplified surfaces", *Computer Graphics forum*,1998, pp. 49-56.