Vis Comput (2016) 32:1121–1132
DOI 10.1007/s00371-015-1184-x

ORIGINAL ARTICLE

Mesh saliency detection via double absorbing Markov chain in feature space

Xiuping Liu[1] · **Pingping Tao**[1] · **Junjie Cao**[1] · **He Chen**[1] · **Changqing Zou**[2,3]

Published online: 25 November 2015

Abstract We propose a mesh saliency detection approach using absorbing Markov chain. Unlike most of the existing methods based on some center-surround operator, our method employs feature variance to obtain insignificant regions and considers both background and foreground cues. Firstly, we partition an input mesh into a set of segments using Ncuts algorithm and then each segment is over segmented into patches based on Zernike coefficients. Afterwards, some background patches are selected by computing feature variance within the segments. Secondly, the absorbed time of each node is calculated via absorbing Markov chain with the background patches as absorbing nodes, which gives a preliminary saliency measure. Thirdly, a refined saliency result is generated in a similar way but with foreground nodes extracted from the preliminary saliency map as absorbing nodes, which inhibits the background and efficiently enhances salient foreground regions. Finally, a Laplacian-based smoothing procedure is utilized to spread the patch saliency to each vertex. Experimental results demonstrate that our scheme performs competitively against the state-of-the-art approaches.

Keywords Mesh saliency · Absorbing Markov chain · Feature space · Foreground cues

✉ Junjie Cao
jjcao1231@gmail.com

[1] School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China

[2] College of Computer Science and Technology, Hengyang Normal University, Hengyang 421000, China

[3] School of Computing Science, Simon Fraser University, Burnaby V5A 1S6, Canada

1 Introduction

With the rapid development of digital scanners and modeling softwares, 3D mesh models are abundant in Internet. It is worth considering that whether all the information in a 3D shape is important for further mesh processing. It is known that human beings tend to identify important regions when an object comes into their eyes, and the recognition of important regions is what 3D mesh saliency concentrates on. Besides, 3D mesh saliency has received considerable attentions from graphics researchers due to its great value in graphics and geometric computing, such as rendering [1], mesh smoothing [2,3], mesh simplification [4,5], viewpoint selection [6,7], shape matching [8,9] and so forth. Visual saliency has also been applied in 3D printing [10], which greatly reduces the printing time.

Recently, a set of new methods have been proposed about mesh saliency computation. Most of them are proposed to capture interesting regions according to the some local contrast operators [4,11], and some multi-scale operators [2,8,12]. Lee et al. [4] advocate 3D mesh saliency by means of center-surrounding contrast on Gaussian-weighted mean curvatures. However, local changes of the curvature may greatly influence the result. Wu et al. [2] propose a novel approach to estimate mesh saliency, which combines of both the local contrast and global rarity considerations. They define the sum of feature distances between every two vertices as global rarity, which may result in blurring of saliency of some distinctive regions through the direct summation of descriptors' distances. Shilane and Funkhouser [8] obtain distinct regions using a shape-based search approach in a training database. The shape saliency is changed with the training database inevitably. Leifman et al. [12] consider regions close to focus points to be more informative. But the selection of the focus points is sensitive to the choice of

different parameters. Tao et al. [11] estimate the saliency of regions according to the relevance to background seeds. They may generate incorrect saliency map when some unsalient patches are identified as salient seeds using a local contrast mechanism.

In this paper, a more faithful and robust mesh saliency detection method is proposed based upon absorbing Markov chain. It is known that human attention is firstly attracted by the most representative salient elements and then the visual attention will be transferred to other regions [11,13,14]. In our model, four stages are adopted to imitate the process. Taking a mesh as input, we first over-segment it into patches based upon Ncuts and Wu et al. [2]. Then, feature variance is introduced to capture the background seeds which are defined as the patches belonging to the segment with minimum feature variance. We utilize the expected time to be absorbed on a graph model which is constructed in a descriptor space to measure the saliency value of each node. We further present a procedure to enhance salient foreground regions and simultaneously inhibit the background regions via foreground cues. Here, the foreground seeds represent the nodes with higher saliency value. Finally, a Laplacian-based smoothing strategy is adopted to spread the patch saliency to vertex saliency. We summarize the contributions of this paper as follows:

- Absorbing Markov chain is applied to describe the transformation of visual attention in mesh saliency detection.
- Feature variance is introduced to extract unsaliency seeds, which is preferred over the local contrast mechanism.
- Our model takes the diversity of foreground nodes into consideration, which efficiently highlights the salient region and suppresses the background region.

2 Related work

Visual saliency is a hot research topic in computer graphics, which is used to identify perceptually important information correlate to human vision systems. Recently, many effective saliency detection models have been developed, which makes the research improve greatly.

Early works on saliency of a 3D shape concentrate on calculating saliency in its 2D projection. Guy and Medioni [15] compute a saliency map based upon edges in a 2D image, and extend it to 3D meshes. Yee et al. [1] employ the method of [13] to calculate a saliency measure based on projecting a 3D dynamic scene to coarsely rendered 2D model. Mantiuk et al. [16] compress a 3D scene animation with a 2D saliency algorithm. Frintrop et al. [17] present a saliency result for fast detecting objects in 3D data.

Besides, some graphics researchers determine saliency based on 3D structure directly. Shilane and Funkhouser [8] compute the distinctive regions of an object by performing a shape-based search. Feixas et al. [6] developed an approach for computing mesh saliency based on viewpoint using mutual information between polygons. Castellani et al. [18] take a scheme for locating and matching interesting points in terms of multi-view of a mesh. Leifman et al. [12] detect the interest regions based on the distance to the foci of attention and apply it to viewpoint selection. Song et al. [19] present a method to detect points of interest on 3D shapes from the point view of spectral.

Moreover, many saliency detection algorithms of 3D shapes are evolved from image saliency. Inspired by image saliency work [13], Lee et al. [4] define mesh saliency of an object in a multi-scale manner. Cheng et al. [20] perform a global contrast-based detection model via spatially weighted distances of region features. Similar to global contrast method in [20], Wu et al. [2] measure mesh saliency based on local contrast and global rarity. However, it is difficult to obtain faithful results by tuning the parameters. Enlightened by the spectral residual analysis taken in Hou and Zhang [21], Song et al. [22] introduce the log-Laplacian spectrum to analyse the attributes of a mesh. They transform the spectral residual from the frequency domain to the spatial domain at multi-scale to localise the salient areas. Tao et al. [11] extend the idea of Yang et al. [23] to mesh saliency detection. They cast it into a graph-based manifold ranking problem and each ranked label characterizes the difference with background.

Besides the methods documented above, random walk has been employed to detect saliency in images [24–27]. Jiang et al. [24] present a method to identify the significant regions according to the random walk time of each element to the boundary of image. Illuminated by [24], we attempt to perform saliency detection process as a random walk issue on 3D geometry. However, considering the technical essence of mesh saliency detection is different from image detection, some hurdles should be tackled during the procedure, such as how to find regions in a 3D surface corresponding to the image boundary, how to connect the elements to capture the underlying geometry and how to improve the saliency result.

3 Absorbing Markov chain

3.1 Mesh saliency and absorbing Markov chain

In this paper, we formulate mesh saliency detection via absorbing Markov chain on a graph model. After oversegmenting a mesh into some big segments and further into a set of smaller patches, the mesh is represented by a graph whose nodes are these patches and every two nodes close in feature space of the patches are connected by an edge. We consider the absorbing Markov chain which includes two kinds of nodes (i.e., absorbing nodes and transient nodes) to

measure mesh saliency. Patches in the segments with the minimum feature variance are selected as unsalient seed patches. Once these seed patches are given, some virtual nodes, which share the same features with the seed patches, are introduced as absorbing nodes. The nodes represent the seed patches and other patches are all transient nodes. We compute the expected time to absorption (i.e., the absorbed time) for each transient node. The absorbed time from each transient node to the absorbing nodes measures the node's global similarity with all absorbing nodes. In other words, the absorbed time measures the global similarity with all unsalient patches of the patch. For the patches having similar appearance with the unsalient seed patches, the random walk starting in their corresponding nodes can easily reach the absorbing nodes and thus has shorter absorbed time, while patches having great contrast to the seed patches will have longer absorbed time. So salient patches can be consistently separated from the unsalient patches via the absorbed time. Based on the above observation, we utilize absorbing Markov chain to detect mesh saliency.

3.2 Principles of absorbing Markov chain

In this section, some of the basic knowledge on Markov chains is described as follows: Given a set of states $S = \{s_1, \ldots, s_l, s_{l+1}, \ldots, s_m\}$, the chain can successively walk from one state to another. Translation probability denoted by p_{ij} is the probability of moving from state s_i to state s_j, which is irrelevant to the state before s_i. The chain can be completely described by the $m*m$ transition matrix $\mathbf{P}$. An ergodic Markov chain is one that can start from any state to any state, which is not necessarily in one step. A state s_i of a Markov chain is named absorbing state when $p_{ii} = 1$. A Markov chain is absorbing on condition that it has at least one absorbing state, and the other state is called transient state. For each one of the transient states, it is likely to be absorbed by one absorbing state, which is not necessarily in a single step.

Given an absorbing chain consisted of m absorbing states and n transient states, renumber the states such that all the transient states are ranked to front. Then, the canonical form of transition matrix $\mathbf{P}$ is expressed as follows:

$$\mathbf{P} = \begin{pmatrix} \mathbf{Q} & \mathbf{R} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}, \tag{1}$$

where $\mathbf{Q} \in [0, 1]_{n*n}$ is the transition matrix, the element of which is the probability of two transient states, $\mathbf{R} \in [0, 1]_{n*m}$ denotes the probabilities from any transient state to any absorbing state, $\mathbf{0}$ is the $m * n$ zero matrix and $\mathbf{I}$ is the $m * m$ identity matrix.

Given the canonical form of transition matrix $\mathbf{P}$ to an absorbing chain, we can extract $\mathbf{Q}$ and obtain the fundamental matrix:

$$\mathbf{N} = (\mathbf{I} - \mathbf{Q})^{-1}. \tag{2}$$

The element n_{ij} of $\mathbf{N}$ is the expected number of times that the chain spends from the transient state s_i to the transient state s_j, and the sum $\sum_j n_{ij}$ gives the expected number of times until the state s_i is absorbed. Let $\mathbf{c}$ denote n dimensional column vector in which all of the entries are 1. Then, the time of absorption of each transient state is given by

$$\mathbf{y} = \mathbf{N} * \mathbf{c}. \tag{3}$$

4 Approach

4.1 Algorithm overview

We design four main steps to generate mesh saliency regions of 3D geometric models, as shown in Fig. 1. First, we segment the input mesh into segments based upon Ncuts and each segment is partitioned into primitive patches based on Zernike coefficients (Sect. 4.2). Then, the feature variance of per segment is calculated using all patches within it (Sect. 4.3). We exploit all the patches whose segment has the minimum data variance as background cues in our model. With a graph constructed in the feature space (Sect. 4.4), the saliency map is estimated via the expected absorbed time in an absorbing Markov chain (Sect. 4.5). We describe the manner of refining the saliency map via the foreground cues in detail (Sect. 4.6). Finally, a smooth vertex-based saliency map is achieved by

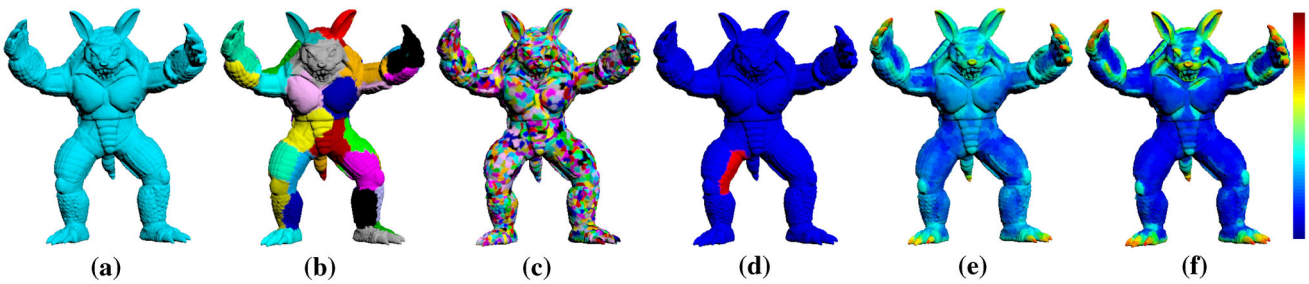

Fig. 1 The functional pipeline of the proposed saliency detection model. **a** Original mesh. **b** Segmentation by Ncuts. **c** Over-segmentation. **d** Some background nodes (*red*). **e** Saliency map by the background cues. **f** Saliency map by the foreground cues

Algorithm 1: Mesh saliency detection based on absorbing Markov chain

Input: A mesh and required parameters.
Steps 1: Segment the mesh into segments based on Ncuts segmentation and then oversegment each segment into patches.
Steps 2: Choose some patches based on the feature variance of segments as seed patches, and duplicate these nodes as absorbing nodes.
Steps 3: Construct a graph $\overline{G}$ with all patches as transient nodes plus (virtual) absorbing nodes.
Steps 4: Compute the transition matrix $\overline{\mathbf{Q}}$ and then compute the fundamental matrix $\overline{\mathbf{N}}$, thereby obtain the saliency map $\mathbf{S}_b$.
Steps 5: Bi-segment $\mathbf{S}_b$ to form salient foreground nodes and compute the saliency map $\mathbf{S}_f$.
Steps 6: Spread the patch saliency $\mathbf{S}_f$ to each vertex by Laplacian-based smoothing process.
Output: A saliency map.

a Laplace procedure among the patches. The pseudo-code of our method is shown in Algorithm 1.

4.2 Over-segmentation and feature descriptor

Similar to the image segmentation approach in [28,29], we perform over-segmentation for each mesh considering between segmentation quality and computational cost. We employ the normalized cuts (Ncuts) [30] to generate a set of segments for each mesh. The number of segments is 80 roughly in our numerical experiments, influenced by the complexity of the shape. Then, the method of Wu et al. [2] is employed to partition per segment into primitive patches.

Inspired by Wu et al. [2] and Tao et al. [11], we adopt the descriptor of Zernike coefficients to compute mesh saliency. Maximo et al. [31] introduce a robust surface descriptor which is insensitive to triangle quality. The Zernike-basis expansion of the heightmap of the surface surrounding a vertex is used as its descriptor, which turns out to be rotated invariantly. It describes the local shape around the vertex. To compute the heightmap, a square sub-region of the tangent plane with side length r is considered. The descriptor can depict a wider range of surface shape with a bigger radius r. As shown in Tao et al. [11], we use $r = 3.0l$ in our implementation, where $l = 0.5\,\%$ of the longest diagonal of the mesh's bounding box. For each patch, its descriptor is defined as the mean of descriptors of the vertices in this patch. We compare Zernike-based patch descriptor and patch descriptor based on curvature histogram in Fig. 2. Taking the top row as an example, the hand, knee and facial regions of Armadillo are considered significant using Zernike-based patch descriptor, which is coherent with human visual system. However, the saliency map using the patch descriptor based on curvature histogram is not. Moreover, the saliency maps using the Zernike-based patch descriptor are more close to the pseudo-ground truth than that using the patch descriptor based on curvature histogram.

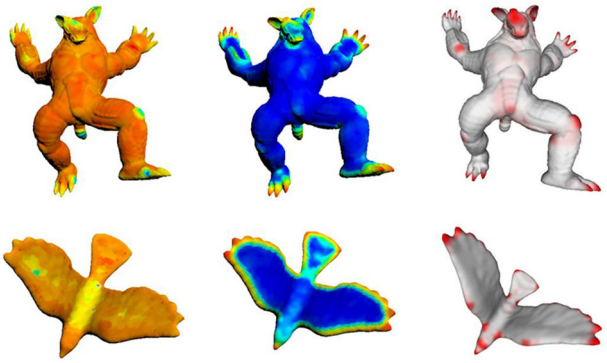

Fig. 2 Mesh saliency via absorbing Markov chain with patch descriptor based on curvature histogram (the *left column*), Zernike-based patch descriptor (the *middle column*) and the pseudo-ground truth (the *right column*)

4.3 Background nodes estimation

To generate stable background nodes for the absorbing Markov chain, we estimate the feature variance of each segment p using all the features of the patches within it:

$$\mathbf{S}_v(p) = \frac{1}{k}\sum_{i=1}^{k} d^2(z_{p_i}, z), \tag{4}$$

where k is the number of patches within the segment p, z_{p_i} is the Zernike descriptor of patch p_i, z denotes the mean Zernike descriptor of all patches in it and $d(.,.)$ is the Euclidean distance between the feature descriptors. We choose all the patches belonging to the segment with minimum feature variance as the seed patches. Thus, it avoids the seed patches scatter in a mesh, which enhances robustness of background cues. Seed nodes generated by the method of Tao et al. [11] are mainly located on the legs (39 red nodes in the leg and 15 red nodes in the body for the model), therefore, generating many (virtual) absorbing nodes corresponding to these red nodes in the leg. Thus, the absorbed time from the nodes on the leg is less than that from nodes on the body (see Fig. 3). And the body is detected as significant region, which is not coherent with human perception. However, the saliency map using our seed nodes achieves more faithful result.

4.4 Graph representation

As stated in [11], mesh saliency is different from image saliency. Image saliency approach is designed to separate the foreground object from background regions, while mesh saliency is to obtain perceptually important points and regions of shapes. Thus, the graph $G(V, E)$ is constructed

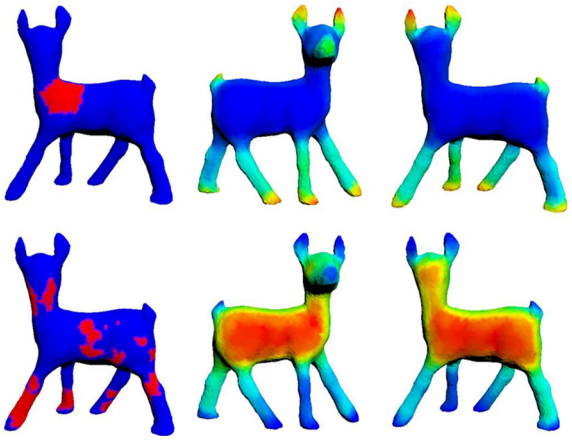

Fig. 3 Saliency generated by Markov random walk using our seed nodes (the *upper row*) is more robust than that generated using seed nodes in [11] (the *bottom row*). The *left column* shows the input model with the seed patches (*red*). The other *two columns* show the final saliency result using the corresponding seed nodes in *left column* as background cues

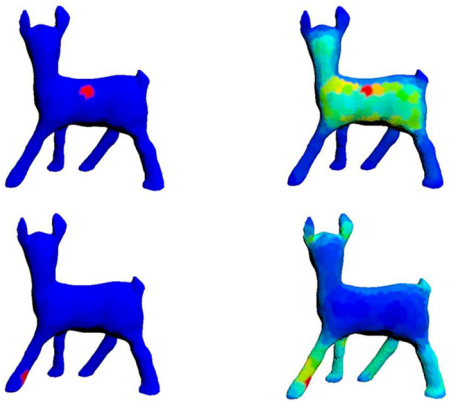

Fig. 4 Visualization of two typical *rows* of the transition matrix $\overline{\mathbf{Q}}$. Current nodes, *colored in red*, and the transition probability from them to their neighbors are illustrated in the *left* and *right columns* respectively

with patches V as nodes and edges E depicting the connectivity between patches in the descriptor space, which reflect the shape structure information. Each node is connected to the nodes within the radius s in the feature space, where s is two-thirds of the maximal feature distance. The weight of the edges encodes the affine relation. The higher value represents stronger connect relation between two nodes. In this work, the weight w_{ij} is defined as:

$$w_{ij} = e^{-||z_i - z_j||/\sigma^2} \quad i, j \in V, \tag{5}$$

where z_i and z_j denote the patch descriptor corresponding to two adjacent nodes i and j, and σ is a constant which controls the strength of the weight. Thus, the affinity matrix $\mathbf{W} = (w_{ij})_{n*n}$ corresponding to the graph $G(V, E)$ is obtained naturally.

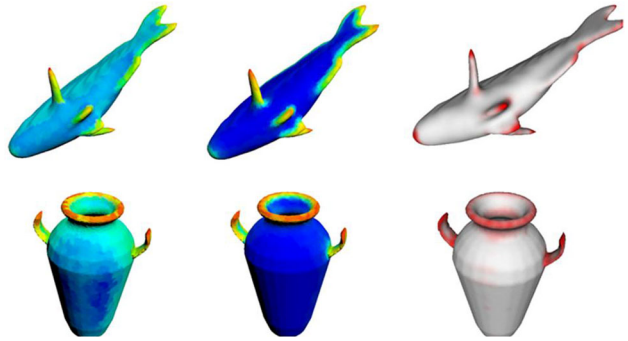

Fig. 5 Mesh saliency refined with the foreground nodes. From *left* to *right*: saliency map via absorbing Markov chain with background cues, saliency map with the foreground cues, the pseudo-ground truth [32]

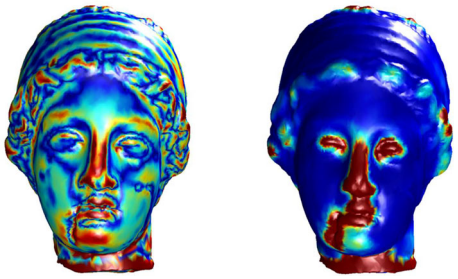

Fig. 6 Comparisons of curvature map (*left*) and saliency map (*right*)

In the graph $G(V, E)$, we construct the (virtual) absorbing nodes by duplicating the seed patches, and all nodes in V are considered as the transient nodes. Then, a new graph $\overline{G}(\overline{V}, \overline{E})$ is built. Here, $\overline{V}$ denotes both the (virtual) absorbing nodes and all nodes in V. $\overline{E}$ represents all edges in E and relationship between the (virtual) absorbing nodes and the transient nodes. The affinity matrix $\mathbf{A} = (a_{ij})_{n*m}$ depicts the weights between the transient nodes and the (virtual) absorbing nodes, the element of which is defined as:

$$a_{ij} = \begin{cases} w_{ij_1} & j_1 \in N(i) \\ 1 & j_1 = i \\ 0 & \text{otherwise}, \end{cases} \tag{6}$$

Table 1 Run times for computing mesh saliency

Model	Verts (K)	Descriptor computation (s)	Over segmentation (s)	Nodes selection (s)	Saliency detection (s)	Saliency refinement (s)
Hand	7	12.488	5.23	0.38	0.034	0.040
Bird	2	10.326	1.59	0.08	0.004	0.003
Vase	15	25.19	20.63	1.203	0.151	0.160
Gargoyle	25	50.834	50.55	3.06	0.62	0.66
Human	15	34.115	15.71	1.38	0.17	0.17

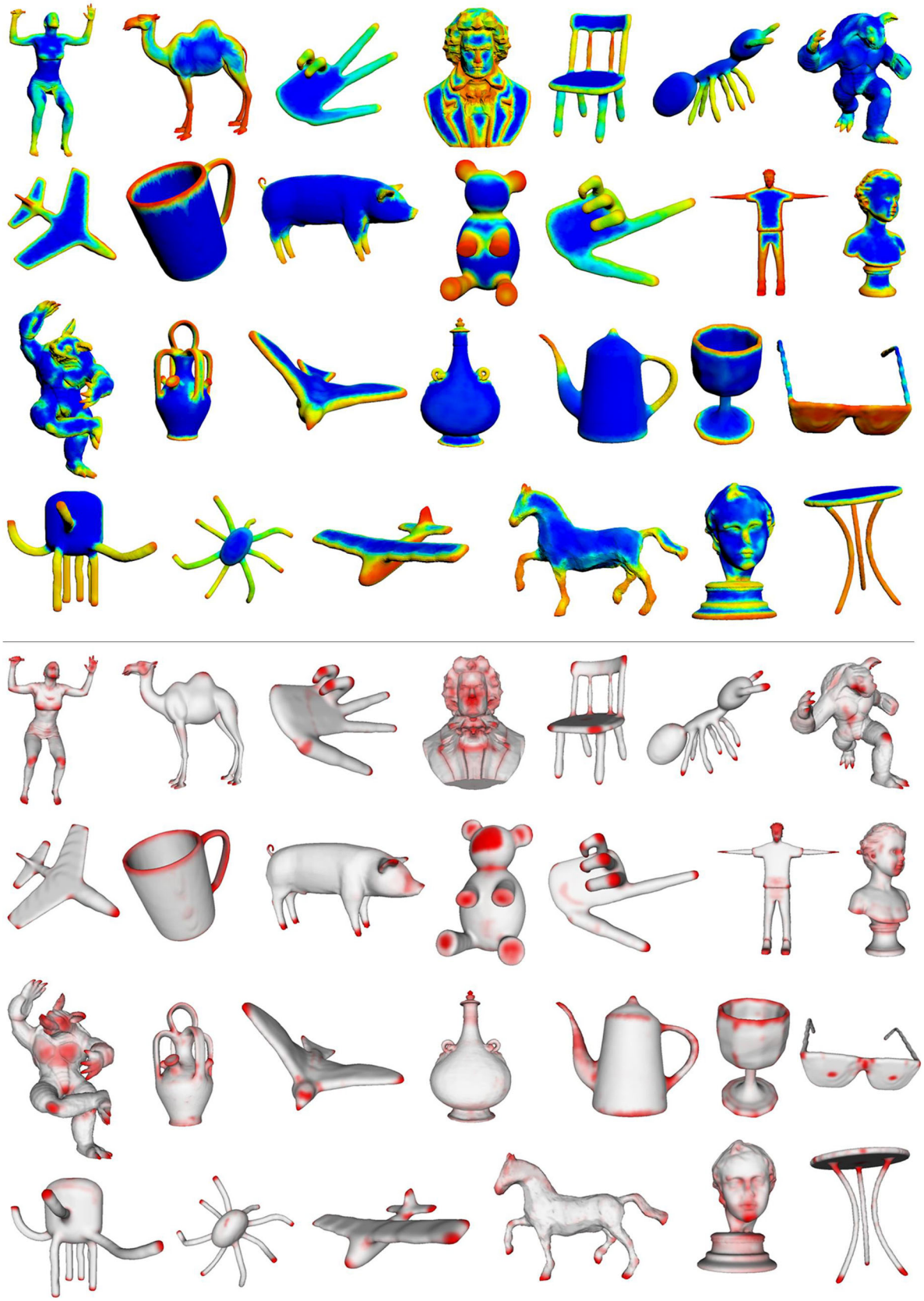

Fig. 7 Experiments of mesh saliency. Our salient results are listed on the *top part* and the corresponding pseudo-ground truth Chen et al. [32] is shown on the *bottom part*

where $N(i)$ denotes the neighborhood of transient node i in the feature space, and j_1 denotes the seed patch corresponding to the (virtual) absorbing node j. Then, the canonical form of transition matrix $\overline{\mathbf{P}}$ of the graph model $\overline{G}(\overline{V}, \overline{E})$ is

$$\overline{\mathbf{P}} = \begin{pmatrix} \overline{\mathbf{Q}} & \overline{\mathbf{R}} \\ \mathbf{0} & \bar{\mathbf{I}} \end{pmatrix}, \tag{7}$$

where $\overline{\mathbf{P}}$ is a row-normalized matrix, $\overline{\mathbf{Q}} = \mathbf{D}^{-1}\mathbf{W}$, $\overline{\mathbf{R}} = \mathbf{D}^{-1}\mathbf{A}$, $\mathbf{0}$ is the $m * n$ zero matrix, $\bar{\mathbf{I}}$ is the $m * m$ identity matrix, and $\mathbf{D} = \text{diag}\{d_{ii}\}_{n*n}$, the element of which is $d_{ii} = \sum_{j=1}^{n} w_{ij} + \sum_{j=1}^{m} a_{ij}$. A visualization of $\overline{\mathbf{Q}}$ is illustrated in Fig. 4. Therefore, the fundamental matrix $\overline{\mathbf{N}}$ is calculated by Eq. 2.

4.5 Saliency detection via background cues

The absorbed time $\mathbf{y}$ is obtained using Eq. 3 and normalized it to the range [0, 1]. The saliency map $\mathbf{S}_b$ can be written as:

$$\mathbf{S}_b(i) = \mathbf{y}^*(i) \quad i = 1, 2, \ldots, n, \tag{8}$$

where i indexes a transient node on the absorbing Markov chain, and $\mathbf{y}^*$ denotes the normalized vector. As illustrated in the first column of Fig. 5, most salient region can be highlighted through the random walk framework. However, some background region may not be adequately inhibited. To alleviate this problem, the saliency result is further improved via the foreground cues.

4.6 Saliency refinement via foreground cues

Because the expected absorbed number of times depend on the appearance weights on the path and their spatial distance on the graph model, some background region close to foreground nodes may be salient. Taking diversity of the foreground nodes into consideration, we present a process to suppress the background region based upon the foreground cues.

The saliency map obtained by Eq. 8 is binary segmented using an adaptive threshold, from which the nodes of foreground regions could be selected as the seed nodes. We expect that the selected nodes could cover perceptual important points and regions as much as possible. In our experiment, the threshold β is set to be the mean value of the entire saliency map.

Given the seed nodes, we duplicate these nodes as the virtual absorbing states, while all nodes in the mesh as the transient states in the Markov chain. Similar to the former stage, the absorbed time $\mathbf{y}$ is computed using Eq. 3 when the fundamental matrix is formed. We normalize the complement of $\mathbf{y}$ as the saliency map, and then denote it as $\mathbf{S}_f$. Finally, a Laplacian-based smoothing procedure is employed to spread the patch saliency to each vertex.

Figure 5 shows two examples where some insignificant regions are better suppressed using foreground cues, and then the salient regions are highlighted. The results are highly consistent with pseudo-ground truth. Note that although high curvature regions tend to be more salient, high curvature does not mean high saliency. As shown in Fig. 6, the saliency of many high curvature regions of the hair is actually low.

5 Experimental results

In this section, we evaluate the proposed method on a variety of object shapes from the Stanford 3D Scanning Repository, the Princeton Segmentation Benchmark [33] and the Watertight Track of the 2007 SHREC Shape-based Retrieval Contest. We present its comparisons with six state-of-the-art mesh saliency detection methods: Lee et al. [4], Wu et al. [2], Shilane and Funkhouser [8], Leifman et al. [12], Tao et al. [11], Song et al. [22] and pseudo-ground truth. The data from [32] are used as pseudo-ground truth. It is collected from an online user study, which utilizes a regression model trained by a leave-one-out procedure for meshes of the same categories. Our algorithm is implemented on a PC with Intel(R) Core(Tm) i7-4790K CPU @ 4.00 GHz and 32 GB RAM. Table 1 shows the performance of our method. The run time of our algorithm mainly consumes at the preprocessing stage, including descriptor computation

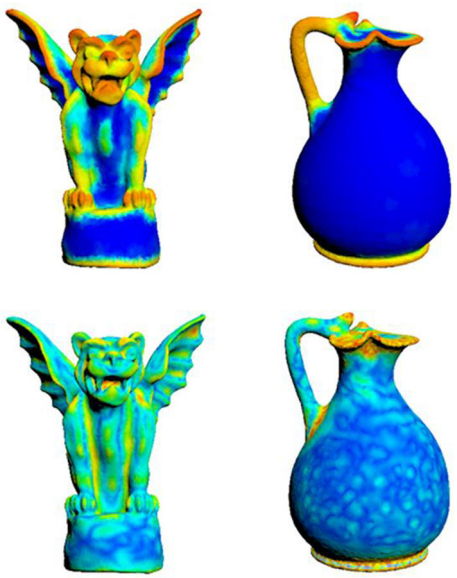

Fig. 8 Saliency results of our method (the *top row*) and the method in [4] (the *bottom row*). Our results are less influenced by local changes of the curvature. Furthermore, our method generates large meaningful salient regions

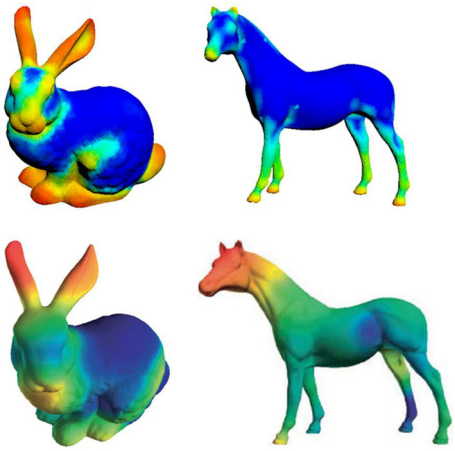

Fig. 9 Saliency results of our method (the *top row*) and the method in [8] (the *bottom row*). Note the differences in the face areas and the legs

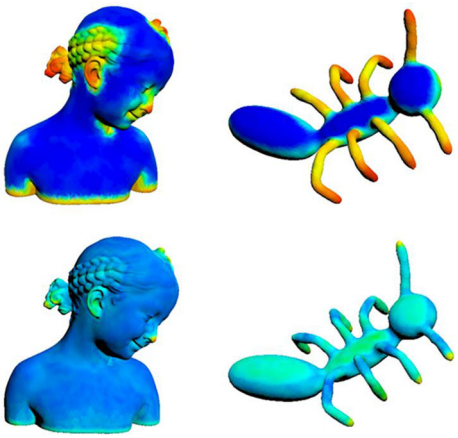

Fig. 10 Saliency results of our method (the *top row*) and the method in [2] (the *bottom row*). Note the differences in the antennae and legs of the Ant model, and the hair of the Girl model

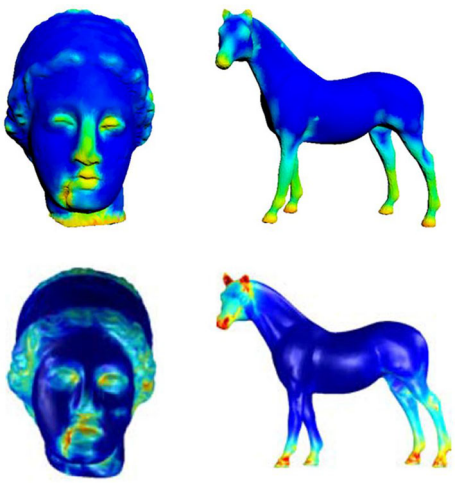

Fig. 11 Saliency results of our method (the *top row*) and the method in [12] (the *bottom row*). Taking the face areas of Venus as an example, some salient region could miss using [12] (such as the nose)

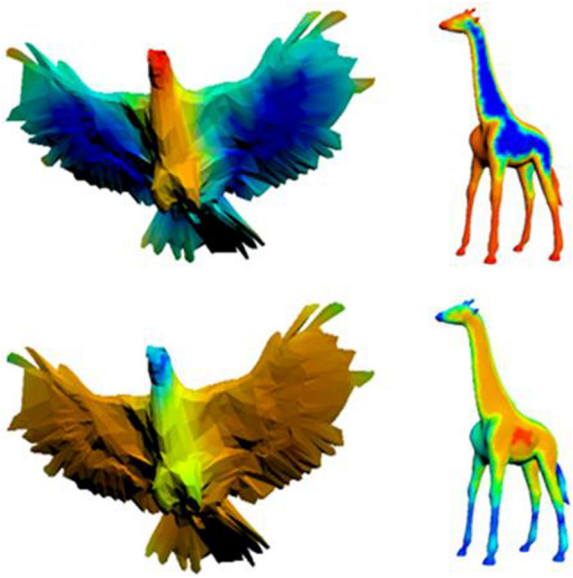

Fig. 12 Saliency results of our method (the *top row*) and the method in [11] (the *bottom row*). Note differences in the head of the bird, and the feet and the tail of the deer model

Fig. 13 Saliency results of our method (the *top row*) and the method in [22] (the *bottom row*). Note differences in lens of the Glasses model and the face areas of the Armadillo

and over-segmentation. For a model with 15 K vertices, the overall time requires about 53.96 s. The two saliency maps computing take about 0.16 and 0.165 s, respectively.

5.1 Algorithm performance

Figure 7 shows comparisons of our saliency results with the corresponding pseudo-ground truth. As can be seen, our

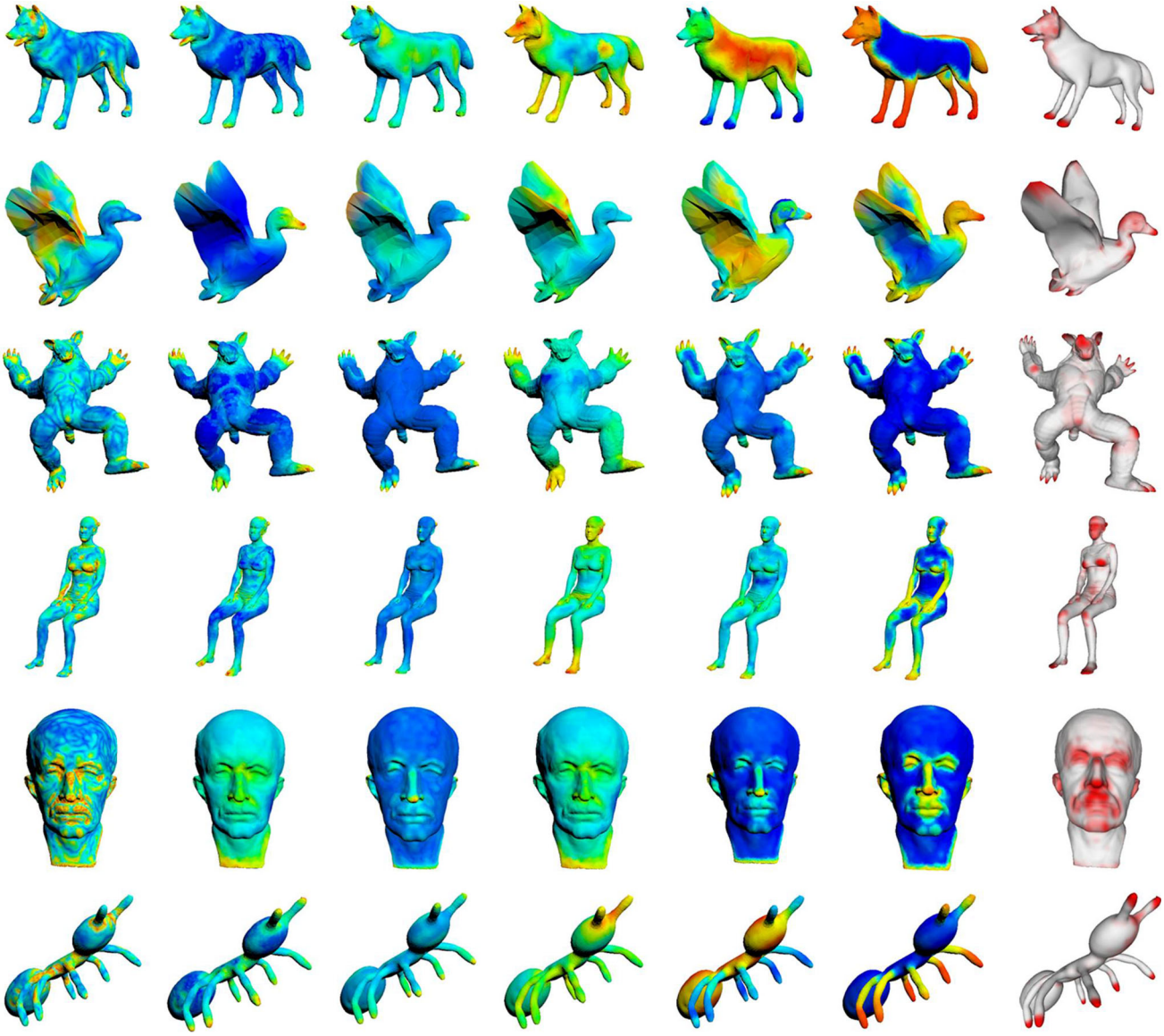

Fig. 14 Mesh saliency results of Lee et al. [4] (the *left column*), Leifman et al. [12] (the *second column*), Wu et al. [2] (the *third column*), Song et al. [22] (the *fourth column*), Tao et al. [11] (the *fifth column*), our method (the *sixth column*), and the pseudo-ground truth (the *right column*)

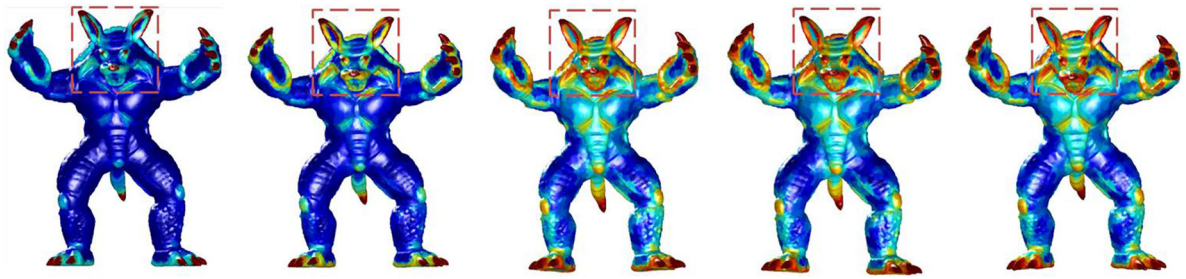

Fig. 15 Comparisons of preliminary saliency results (without saliency refinement) with different σ. The results shown from *left* to *right* are obtained with σ of $10^{-2} * 0.3$, $10^{-2} * 0.5$, $10^{-2} * 2/3$, $10^{-2} * 0.75$ and $10^{-2} * 1.0$, respectively

results are to a large extent consistent with the pseudo-ground truth.

Figure 8 compares our results with those of [4]. Our results are less influenced by frequent local changes of curvature. For instance, the hand and facial regions of the gargoyle and the handle of the vase are considered significant by our algorithm, whereas the results captured by [4] are disjointed.

Figure 9 compares our results with those from [8]. Our method is able to detect some small salient regions, such as the ears, eyes and mouth of the horse model. However, [8] can only mark the entire head of the horse as distinct, since the smallest regions they employ to estimate saliency have radius 0.25 times the radius of the entire object.

Figure 10 compares our results with those of [2], in which local contrast and global rarity are considered. However, they may lead to the missing of some significant regions because the concept of global rarity is expressed as the sum of feature distances between every two vertices. Our method marks the facial features and the hair of the girl, and the antennae and legs of the ant, while they are not captured by [2].

Figure 11 compares our results with those of [12]. Both 20 % of the most distinct vertices and the extreme vertices are considered as focus points in [12], while some saliency regions that are not close to focus points are missed. Note that our algorithm captures the nose of the Venus and the eyes of the horse, but [12] fails to achieve this.

Figure 12 compares our results with those of [11]. The head, legs and tail of the deer and the head of the bird are detected as salient by our method, while the results of the competing method are disorganized.

Figure 13 compares our results with those of [22]. Although in [22] the entire head of the Armadillo is marked, we detect more detailed regions, such as the ears, the mouth, and the eyes. Our algorithm captures the knees, which are not captured by [22]. In addition, their approach fails to mark the lens of the Glass, while it is captured by our method.

Figure 14 illustrates more comparison results. Our method detects the salient regions of the models more close to the pseudo-ground truth. However, the results of others, to a certain extent, are confused.

5.2 Influence of parameters on mesh saliency detection

Parameter σ As the value of σ influences the graph construction, we also show saliency results (without saliency refinement) using different σ. As shown in Fig. 15, the saliency results in the last three columns perform better than the first two columns. Besides, diffusion time of $\sigma = 10^{-2} * 2/3$ is shorter than the last two columns. Thus, we use $\sigma = 10^{-2} * 2/3$ empirically.

Sampling and number of segments The locations of the seed patches are influenced by the number of segments and different sampling densities. However, the saliency maps of our method are robust to these factors (see Figs. 16, 17). In all the experiments, we segment the input model into 80 segments which will be further divided into 3000 patches roughly.

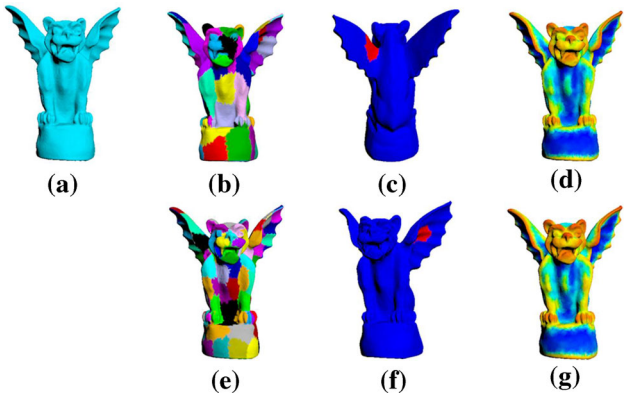

Fig. 16 Our method is insensitive to the number of segments. **a** Input mesh. **b**, **e** Are segmentations with 84 and 164 segments, respectively. The background seeds (*red*) and saliency maps generated using **b** and **e** are shown in the rest *two columns*

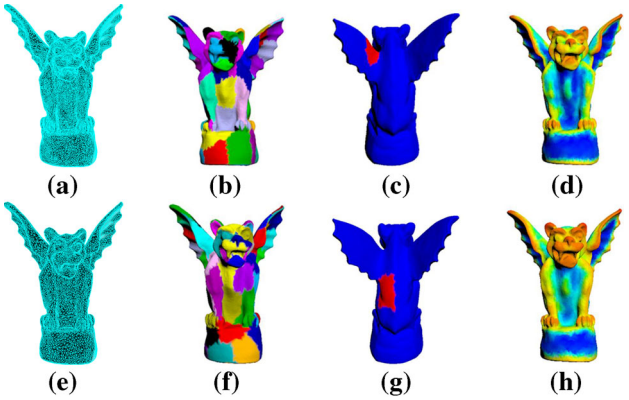

Fig. 17 Our method is sampling insensitive. The models in **a** and **e** have 25 and 12 K vertices, respectively. **b**–**d** Illustrate the segmentation, background seeds (*red*) and the saliency map of **a**. And **f**–**h** show those of **e**

6 Conclusion

In this paper, we propose a novel mesh saliency detection method based on random walk framework. Unlike most of the prior methods focusing on local contrast metrics, we utilize data variance to capture the robust background nodes in our model. Using the time property of absorbing Markov chains, a two-stage process with background and foreground cues is adopted to generate the saliency maps based on the graph constructed in feature space. The first stage aims to localize salient regions roughly. Moreover, to separate the foreground salient regions from the background regions better, the second stage is applied. Furthermore, the proposed method runs on a variety of shapes, which validates convincing results with comparisons against state-of-the-art approaches.

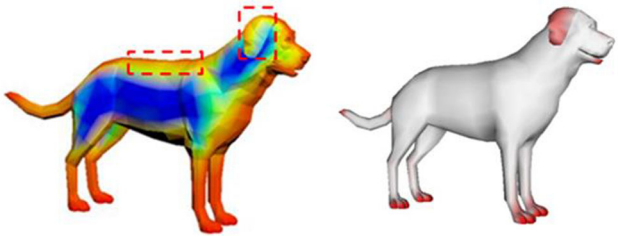

Fig. 18 Failure case. Saliency map of our method (*left*) and the pseudo-ground truth [32] (*right*)

However, our scheme does not take shape extremities suggested by Leifman et al. [12] into consideration. As shown in Fig. 18, our method cannot capture the whole ears and it marks the spine as saliency region, which is not in accord with the pseudo-ground truth provided by Chen et al. [32]. In the future, we plan to fuse shape extremities into our method.

Acknowledgments The authors sincerely thank reviewers for their valuable comments. Xiuping Liu is supported by the NSFC (Nos. 61173102, 61370143). Junjie Cao is supported by the NSFC (Nos. 61363048, 61262050). Changqing Zou is supported by the NSFC (No. 61502153).

References

1. Yee, H., Pattanaik, S., Greenberg, D.: Spatiotemporal sensitivity and visual attention for efficient rendering of dynamic environments. ACM Trans. Graph. **20**(1), 39–65 (2001)
2. Wu, J., Shen, X., Zhu, W., Liu, L.: Mesh saliency with global rarity. Graph. Models **75**(5), 255–264 (2013)
3. Li, Z., Ma, L., Jin, X., Zheng, Z.: A new feature-preserving mesh-smoothing algorithm. Vis. Comput. **25**(2), 139–148 (2009)
4. Lee, C., Varshney, A., Jacobs, D.: Mesh saliency. ACM Trans. Graph. **24**(3), 659–666 (2005)
5. Yang, B., Li, Frederick, W., Wang, X., Xu, M., Liang, X., Jiang, Z., Jiang, Y.: Visual saliency guided textured model simplification. Vis. Comput. 1–18 (2015)
6. Feixas, M., Sbert, M., González, F.: A unified information-theoretic framework for viewpoint selection and mesh saliency. ACM Trans. Appl. Percept. **6**(1), 1–23 (2009)
7. Zhao, S., Ooi, W.: Modeling 3D synthetic view dissimilarity. Vis. Comput. (2015)
8. Shilane, P., Funkhouser, T.: Distinctive regions of 3D surfaces. ACM Trans. Graph. **26**(2), 7 (2007)
9. Hu, J., Hua, J.: Salient spectral geometric features for shape matching and retrieval. Vis. Comput. **25**(5–7), 667–675 (2009)
10. Wang, W., Chao, H., Tong, J., Yang, Z., Tong, X., Li, H., Liu, X., Liu, L.: Saliency-preserving slicing optimization for effective 3D printing. Comput. Graph. Forum
11. Tao, P., Cao, J., Li, S., Liu, X., Liu, L.: Mesh saliency via ranking unsalient patches in a descriptor space. In: IEEE International Conference on Shape Modeling and Applications, pp. 264–274 (2014)
12. Leifman, G., Shtrom, E., Tal, A.: Surface regions of interest for viewpoint selection. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 414–421 (2012)
13. Itti, L., Koch, C., Niebur, E.: A model of saliency-based visual attention for rapid scene analysis. IEEE Trans. Pattern Anal. Mach. Intell. **20**(11), 1254–1259 (1998)
14. Liu, R., Cao, J., Lin, Z., Shan, S.: Adaptive partial differential equation learning for visual saliency detection. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 3866–3873 (2014)
15. Guy, G., Medioni, G.: Inference of surfaces, 3D curves, and junctions from sparse, noisy, 3D data. IEEE Trans. Pattern Anal. Mach. Intell. **19**(11), 1265–1277 (1997)
16. Mantiuk, R., Myszkowski, K., Pattanaik, S.: Attention guided MPEG compression for computer animations. In: Proceedings of the 19th Spring Conference on Computer Graphics, pp. 239–244 (2003)
17. Frintrop, S., Nüchter, A., Surmann, H.: Visual attention for object recognition in spatial 3D data. In: Attention and Performance in Computational Vision, pp. 168–182 (2005)
18. Castellani, U., Cristani, M., Fantoni, S., Murino, V.: Sparse points matching by combining 3D mesh saliency with statistical descriptors. Comput. Graph. Forum **27**(2), 643–652 (2008)
19. Song, R., Liu, Y., Martin, R., Rosin, P.: 3D point of interest detection via spectral irregularity diffusion. Vis. Comput. **29**(8), 695–705 (2013)
20. Cheng, M., Zhang, G., Mitra, N., Huang, X., Hu, S.: Global contrast based salient region detection. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 409–416 (2011)
21. Hou, X., Zhang, L.: Saliency detection: a spectral residual approach. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8 (2007)
22. Song, R., Liu, Y., Martin, R., Rosin, P.: Mesh saliency via spectral processing. ACM Trans. Graph. **33**(1), 1–17 (2014)
23. Yang, C., Zhang, L., Lu, H., Ruan, X., Yang, M.-H.: Saliency detection via graph-based manifold ranking. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 3166–3173 (2013)
24. Jiang, B., Zhang, L., Lu, H., Yang, M.: Saliency detection via absorbing markov chain. In: IEEE International Conference on Computer Vision, pp. 1665–1672 (2013)
25. Harel, J., Koch, C., Perona, P.: Graph-based visual saliency. In: Advances in Neural Information Processing Systems, pp. 545–552 (2006)
26. Gopalakrishnan, V., Hu, Y., Rajan, D.: Random walks on graphs for salient object detection in images. IEEE Trans. Image Process. **19**(12), 3232–3242 (2010)
27. Zhang, W., Xiong, Q., Shi, W., Chen., S.: Region saliency detection via multi-feature on absorbing Markov chain. Vis. Comput. (2015)
28. Shi, J., Malik, J.: Normalized cuts and image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **22**(8), 888–905 (2000)
29. Ren, X., Malik, J.: Learning a classification model for segmentation. In: IEEE International Conference on Computer Vision, vol. 2, pp. 10–17 (2003)
30. Aleksey, G., Thomas, F.: Randomized cuts for 3D mesh analysis. ACM Trans. Graph. **27**(5), 145–158 (2008)
31. Maximo, A., Patro, R., Varshney, A., Farias, R.: A robust and rotationally invariant local surface descriptor with applications to non-local mesh processing. Graph. Models **73**(5), 231–242 (2011)
32. Chen, X., Saparov, A., Pang, B., Funkhouser, T.: Schelling points on 3D surface meshes. ACM Trans. Graph. **31**(4), 29:1–29:12 (2012)
33. Chen, X., Golovinskiy, A., Funkhouser, T.: A benchmark for 3D mesh segmentation. ACM Trans. Graph. **28**(3), 73:1–73:12 (2009)

Xiuping Liu is a Professor in School of Mathematical Sciences at Dalian University of Technology, China. She received Ph.D. degree in computational mathematics from Dalian University of Technology. Her research interests include shape modeling.

Pingping Tao is a Ph.D. student in computational mathematics, Dalian University of Technology. She received M.S. degree from University of Science and Technology of China. Her research interest is computer graphics.

Junjie Cao is a lecturer in School of Mathematical Sciences at Dalian University of Technology. He received Ph.D. degree in computational mathematics from Dalian University of Technology. His research interests include shape modeling.

He Chen is a graduate student in computational mathematics, Dalian University of Technology. He received bachelor's degree of electrical engineering from Hunan University, and turns to study computational-mathematics in his postgraduate education. His research interests include shape modeling and machine learning.

Changqing Zou received the BE degree from the Harbin Institute of Technology, China, in 2005, and the ME degree from the Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, China, in 2008. He was with the multimedia lab of Shenzhen Institutes of Advanced Technology, CAS as a Ph.D. student from Aug. 2010 to Jan. 2015. He is now a postdoctoral research fellow in SFU, Canada, also an assistant professor in Hengyang Normal University, China. His interests include computer graphics and computer vision, and multimedia.