

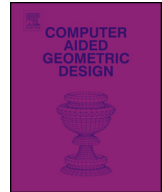


ELSEVIER

Contents lists available at ScienceDirect

Computer Aided Geometric Design

www.elsevier.com/locate/cagd



Multi-scale mesh saliency based on low-rank and sparse analysis in shape feature space

Shengfa Wang^a, Nannan Li^b, Shuai Li^{c,*}, Zhongxuan Luo^a, Zhixun Su^b,
Hong Qin^d

^a School of Software Technology, Dalian University of Technology, China

^b School of Mathematical Sciences, Dalian University of Technology, China

^c State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, China

^d Department of Computer Science, Stony Brook University, United States

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Saliency
Low-rank and sparse analysis
Shape feature
Structure

ABSTRACT

This paper advocates a novel multi-scale mesh saliency method using the powerful low-rank and sparse analysis in shape feature space. The technical core of our approach is a new shape descriptor that embraces both local geometry information and global structure information in an integrated way. Our shape descriptor is organized in a layered and nested structure, enabling both multi-scale and multi-level functionalities. Upon devising our novel shape descriptor, the remaining challenge is to accurately capture sub-region (or sub-part) saliency from 3D geometric models. Towards this goal, we exploit our novel shape descriptor to define local-to-global shape context in a vertex-wise fashion and concatenate all the shape contexts to form a feature space, which encodes both local geometry feature and global structure feature. It then paves the way for us to employ the powerful low-rank and sparse analysis in the feature space, because the low-rank components emphasize much more on stronger patch/part similarities, and the sparse components correspond to their differences. By focusing on the sparse components, we develop a versatile, structure-sensitive saliency detection framework, which can distinguish local geometry saliency and global structure saliency in various 3D geometric models. Our extensive experiments have exhibited many attractive properties of our novel shape descriptor, including: being suitable for perception-driven analysis, being structure-sensitive, multi-scale, discriminative, and effectively capturing the intrinsic characteristic of the underlying geometry.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Visual saliency is an important and fundamental research topic in psychology and neuroscience to investigate the mechanism of human visual systems, and is also an attractive topic in computer vision and computer graphics. Saliency together with its effective and accurate detection depends on the fact that meaningful shape context and its corresponding physical location in an object always attract people's attention. Saliency can also be treated as certain type of broadly-defined features (Gal and Cohen-Or, 2006). Although much excellent work has been carried out in signal processing and computer

* Corresponding author.

E-mail address: lishuai@buaa.edu.cn (S. Li).

<http://dx.doi.org/10.1016/j.cagd.2015.03.003>

0167-8396/© 2015 Elsevier B.V. All rights reserved.

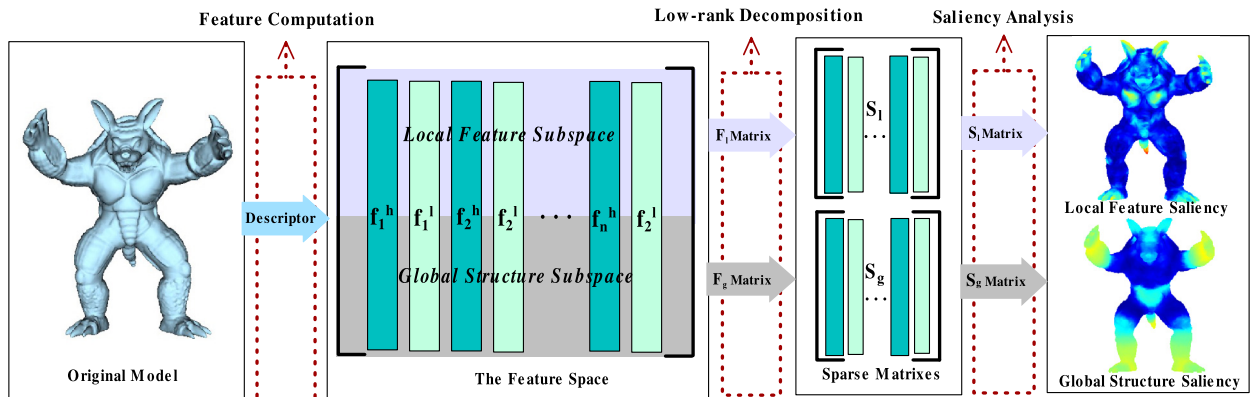


Fig. 1. The functional pipeline of our saliency detection. We first use the proposed structure-aware descriptor to construct a feature space, which consists of the local feature subspace and the global structure space. A local feature matrix and a global structure matrix are extracted from the corresponding subspace. Then, we explore the low-rank decomposition and sparse representation to decompose the matrices into low-rank parts and sparse parts. Finally, the local feature saliency and the global structure saliency are obtained by analyzing the corresponding sparse matrices.

vision, there is still a limited research attention paid to the utility of perception-inspired metrics for saliency definition and its effective detection in geometry processing of 3D surfaces.

Moreover, most of the existing approaches on 3D mesh saliency detect the saliency according to the local contrast (Gal and Cohen-Or, 2006; Kim and Varshney, 2006; Kim et al., 2010). Recently, the multi-scale computation (Lee et al., 2005; Shilane and Funkhouser, 2007; Leifman et al., 2010) and the global rarity (Wu et al., 2013) are proposed to detect the saliency in a larger region. Nonetheless, both of them consider only the local geometric features of the salient (i.e., distinctive) differences in different scale regions. On the other hand, the global structures, which depict global shapes, have not yet to be fully considered when detecting mesh saliency.

Our motivation in this paper is to empower 3D mesh saliency with novel techniques that also enable more meaningful applications in graphics and geometric computing. We observe that, only detecting the saliency of local features, whether in local scale or global scale, is not discriminative enough at present to express the perceptual importance of global structures, such as the limbs in certain animal models. Moreover, generic shape features in a global setting are also relevant to shape frequency. From the point view of shape frequency, a 3D shape can be decomposed into a shape spectrum that spans across high frequency information to low frequency information, which collectively forms the “shape-DNA” (Reuter et al., 2006; Bronstein et al., 2010). The high frequency information is more sensitive to the fine features with geometric details, and the low frequency information contributes more to the base shape, which is much more stable. The novel method to be presented in this paper has been built upon our strong belief that effectively and accurately recognizing the global importance of different frequency information will help us quantify the saliency of 3D models. Equally important, at the computational level, our work has also been strongly inspired by the successful application of low-rank and sparse analysis for image processing (Yan et al., 2010; Zhang and Li, 2012), we have found that the feature space spanned by the proposed shape descriptor can be decomposed into the low-rank part and the sparse part respectively, which means that we could also make use of the low-rank and sparse analysis for mesh saliency and its effective detection.

To handle the aforementioned problems, we propose a versatile saliency detection framework based on the low-rank and sparse analysis, as shown in Fig. 1. Towards this goal, we first propose a new structure-aware descriptor that integrates both multi-scale and multi-level information to capture the intrinsic characteristic of the underlying geometry. Then a low-rank and sparse modeling is explored in a feature space spanned by the shape descriptors to exhibit the sparse characteristic, which are further reformulated into different saliency, including local feature saliency and global structure saliency. The former captures the salient regions of local scope that is related to local features, and the later captures the salient structures from the global point of view, such as the limbs of animal models.

2. Background and related work

The concept of saliency originates from the area of computer vision, and in recent years we have seen many research progresses in saliency and its effective detection. The technical essence of saliency detection is to identify automatically the important sensory information that is pertinent to a human vision system. Also, the newly-proposed and promising method, namely low-rank and sparsity decomposition, is popular and powerful for saliency analysis.

Earlier research works on saliency analysis concentrate more on the saliency measurement in a local scale. From the point of view of fundamental geometric perception, the intuition of salient regions should be those that are distinctive from their immediate surroundings, as suggested by Koch and Ullman (1987). Such type of definition is so straightforward that it tends to ignore the object’s details mechanically. To address this problem, Einhäuser and König (2003) performed the detection process via the combinatorial use of color and texture information. Similarly, Ma and Zhang (2003) defined a perceived field to set up the unit, from which the contrast value based on color and intensity was calculated. Other feature

attributes, such as luminance (Le Meur et al., 2006), edge orientation (Walther and Koch, 2006), contour information (Itti et al., 1998) and their combinations (Liu et al., 2011; Goferman et al., 2012; Perazzi et al., 2012) are commonly used in previously proposed saliency frameworks. Inspired by the work of saliency detection in images, Lee et al. (2005) were among the first to introduce the perception of saliency into mesh models, which highlights the interesting regions by introducing a center-surrounding operator on Gaussian-weighted mean curvatures. Following the first work of saliency on meshes (Lee et al., 2005) are two works on saliency-guided mesh editing. Kim and Varshney (2006) enhanced the selected regions of a volume by a visual-saliency-based operator. Gal and Cohen-Or (2006) defined the saliency grade of a cluster by considering the area of the patch, the curvature, and the curvature variance in the cluster. In addition to research works on the saliency-guided applications (Feixas et al., 2009; Kim and Varshney, 2008), Kim et al. (2010) presented a user study that performs comparison between the mesh saliency approaches and recorded human eye movements. However, these methods overlook the fact that human eyes oftentimes catch an object from the global perspective, which should be incorporated into the conception and measurement of saliency.

Later on, researchers started to define and perceive saliency in a larger scope, whereby the central idea is to set up one kind of measurement that makes the eye-catching regions stand out from a global scope. Several methods (Chen et al., 2007; Torralba et al., 2006) first obtained saliency by globally comparing the local statistics. However, the repeated patterns of such sub-parts in a global setting greatly deteriorated their rarity and uniqueness measurement. Recent methods (Cheng et al., 2011; Duan et al., 2011) measured global contrast-based saliency based on spatially weighted feature dissimilarities. And Perazzi et al. (2012) conducted saliency estimation using two Gaussian filters. In addition, priors or heuristics regarding the likely positions of foreground and background have been shown to be effective in recent methods (Borji et al., 2012; Wei et al., 2012). Enlightened by the global consideration taken in images, Wu et al. (2013) first gave the definition of global rarity on meshes by directly summing up the distances between every two indexed vertices. However, the direct summation of descriptors' distances may give rise to the averaging and blurring effect of the distinctive properties of certain regions. Spectral analysis is also an effective tool used in image saliency detection (Hou and Zhang, 2007). Song et al. (2014) obtained good saliency results by analyzing the spectral attributes of the log-Laplacian spectrum of a mesh. However, there is no approach could capture the global structure saliency, which will help to understand the global shape well. Such deficiency calls for new and more powerful analytical tools for saliency detection that can depict global shapes.

Besides the straightforward comparison of local statistics as documented above, the low-rank and sparse analysis algorithm in machine learning has been employed to detect saliency in images (Yan et al., 2010; Lin et al., 2009; Zhang and Li, 2012). Methods (Zhang and Li, 2012; Sun et al., 2010) based on low-rank decomposition typically employ learning based feature transformations to represent their similar attributes. There exists strong correlation among the proposed representations of the redundant parts, namely non-salient background, and what are left out are the novel parts that cannot be represented well, which are the so-called saliency. The basic principle of low-rank and sparse analysis (Yan et al., 2010; Zhang and Li, 2012) is that the common patterns should be suppressed, while at the same time novel features should be kept and highlighted. These methods, to certain extent, can well depict the region of interest from the global point of view, but they usually fail to identify the relatively small-scale salient elements and the repetitive texture-pattern saliency. One key reason is the lack of an intrinsic and informative attribute descriptor to serve as the structural feature carrier for saliency measurement.

3. Structure-aware descriptor

We first introduce a powerful structure-aware descriptor to characterize the shapes. Then, we use the proposed descriptor to span a feature space, in which saliency detection could be conducted. Our descriptor is powerful in two aspects: both the scale information (from local to global region) and the level information (from high to low frequency) are encoded. For the convenience of technical discussion in the following sections, we assume that the input model \mathbf{M} is a triangular mesh with vertices $\mathbf{V} = (v_1, v_2, \dots, v_n)^T$.

3.1. Multi-level shape expression

From the spectral graph theory (Reuter et al., 2006; Bronstein et al., 2010), we know that the Laplace–Beltrami spectra are considered as ‘Shape-DNA’. The collection of different frequency (multi-level) information would help us understand the shape in a better and more intrinsic way. In order to capture the multi-level information, the model should first be decomposed into different frequency information. Then, we reconstruct models of different levels using different frequency information

$$\mathbf{V}^m = \sum_{k=1}^m \tilde{x}_k \phi_k, \quad (1)$$

where \mathbf{V}^m are the vertices of the m -th level model \mathbf{M}^m , $\{\tilde{x}_k\}$ are the coefficients corresponding to the first m low frequency eigenbases $\{\phi_k\}$, and the Laplacian matrix is constructed using the cotangent weight scheme (Meyer et al., 2002; Rong et al., 2008). The more levels of the model are utilized, the more information we will obtain. A proper tradeoff should be made between the level of information and efficiency. In this paper, we choose to use two levels, the model with high

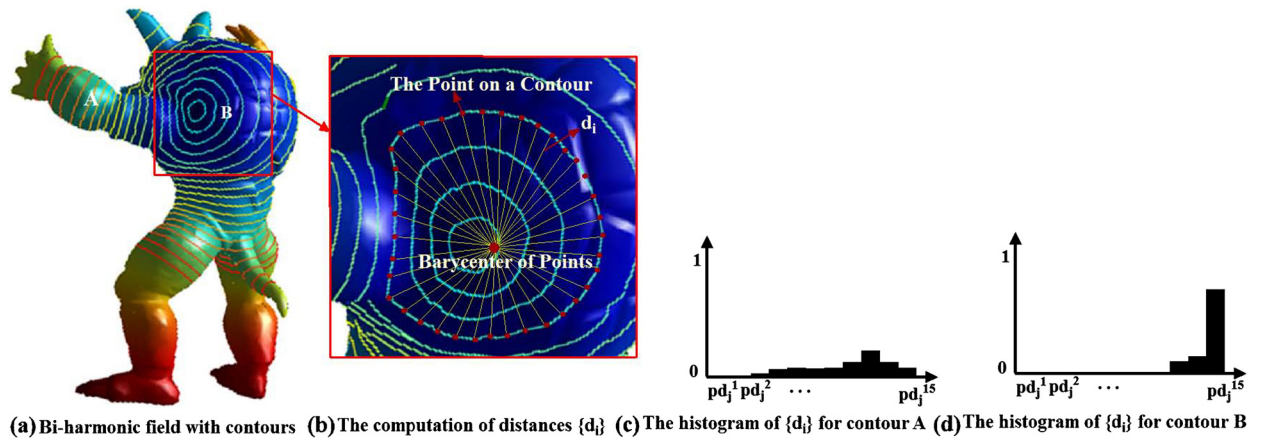


Fig. 2. The construction of the shape context. (a) A bi-harmonic distance field. (b) The Euclidean distances between the points on the contour and their barycenter are computed. ((c) and (d)) The probability distribution histograms for two different contours with the same perimeter.

frequency information and the one with low frequency information. It may be noted that, multiple levels could be easily accommodated, the strong motivation for us to only use two levels in this paper is solely due to the efficiency reason, and it works well for all the experiments. For the high level, we use the original model \mathbf{M} . For the low level, a low frequency model \mathbf{M}^l is constructed in Eq. (1) (we use the m -th level model \mathbf{M}^m , where m is set to be 200 empirically for all examples). The two levels of models combine to form a multi-level expression, which makes the descriptor contain more information, and allows sufficient redundancy in order to empower the sparse analysis tool.

3.2. Multi-scale shape context

For each level of a model, we construct a multi-scale shape context to describe the shape characteristics by considering the contours on a distance field. Here, we choose to utilize bi-harmonic distance (Lipman et al., 2010) to obtain the distance field for two reasons: (1) the bi-harmonic distance is widely used in geometry processing because of its good properties, such as robustness, parameter-free and versatility; (2) we could reuse the eigenvalue and eigenbasis decomposition in the multi-level expression from the prior step. The bi-harmonic distance between vertex v_i and v_j can be expressed as

$$D(i, j)^2 = \sum_{k=1}^m \frac{(\phi_k(i) - \phi_k(j))^2}{\lambda_k^2}, \quad (2)$$

where $\{\lambda_k\}$ and $\{\phi_k(\cdot)\}$ are the first m eigenvalues and eigenbases respectively.

For each vertex, we construct a corresponding bi-harmonic distance field using the distance in Eq. (2). Then, a set of equal-distance contours can be obtained (as shown in Fig. 2(a)), which can characterize the shape structure well. However, any single contour could not completely define a shape structure. To address the above problem, we further explore the statistics to complete our descriptor. For each contour, we compute the Euclidean distances between the points on the contour and their barycenter (see Fig. 2(b)), and further evaluate the probability distribution histogram for these normalized Euclidean distances (see Fig. 2(c) and (d)), and we can notice that the contours with the same perimeter also have distinguishing characters. Finally, we can define the multi-scale shape context of vertex v_i as

$$\mathbf{f}_i = [m_i, p_1, pd_1, p_2, pd_2, \dots, p_k, pd_k]^T, \quad (3)$$

where m_i is the local metric (we use mean curvature here), p_j is the normalized perimeter of the j -th contour (using the radius of bounding box), pd_j is the corresponding probability distribution histogram of the distances $\{d_i\}$, and k is the number of contours. Empirically, we set the histogram $pd_j = \{pd_j^b\}_{b=1..15}$ and $k = 10$ in the interval $\max\{D(\cdot, \cdot)\}/40$ for all our examples. This shape context contains both local geometric information and global shape information, that ensure to well describe a shape in a hierarchical fashion.

3.3. Shape descriptor and feature space

Now, it sets the stage for us to construct our shape descriptor using both the multi-level expression and the multi-scale shape context. For each vertex v_i , the descriptor can be expressed as $\mathbf{F}_i = [\mathbf{f}_i^h, \mathbf{f}_i^l]$, where \mathbf{f}_i^h is the shape context calculated on the high level model \mathbf{M} (both high frequency and low frequency information) and \mathbf{f}_i^l is the shape context calculated on the low level model \mathbf{M}^l (low frequency information only). The proposed shape descriptor integrates both multi-scale (local and global) and multi-level (low and high frequency) information. It is a new class of shape descriptors

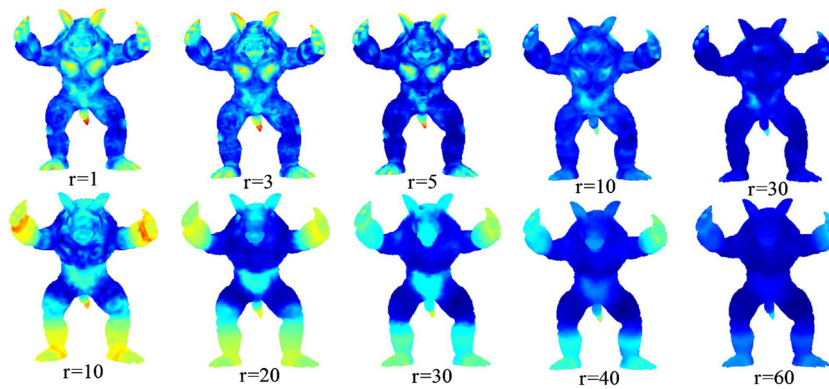


Fig. 3. The saliency results with different rank constraints. The top row is the local feature saliency. The bottom row is the global structure saliency.

with many attractive properties, including rotation-invariant, scale-invariant, sampling-insensitive, structure-aware, multi-scale, and discriminative. Therefore, it is suitable for shape analysis, and can effectively capture the intrinsic characteristic of the underlying geometry.

Assemble the descriptor of each vertex to span a feature space, which can be reorganized in the form of matrix $\mathbf{F} = [\mathbf{F}_1; \mathbf{F}_2; \dots; \mathbf{F}_n]$. The feature space contains both local and global information. As we know from the above construction, the front components of shape contexts mainly represent the local characters, and the rear components of shape contexts put more emphasis on the global structures. Therefore, we further divide the feature space \mathbf{F} into two subspaces: \mathbf{F}_l and \mathbf{F}_g , where \mathbf{F}_l is the local feature subspace and \mathbf{F}_g is the global structure subspace. Here, the local feature subspace consists of components of the first five dimensions in the feature space, and the global structure consists of the rest components in the feature space. The abundant geometric information affords us to detect various saliency in the space. The former is mainly utilized to detect the local feature saliency, and the latter is used primarily for global structure saliency detection.

4. Saliency detection based on low-rank and sparse modeling

Our structure-aware descriptor contains sufficient amount of information to describe the shapes, and there is also a large amount of repeated and redundant information, which can be considered as the non-salient background. Based on the descriptor, we primarily concentrate on the saliency detection by exploiting the powerful capability of low-rank approximation and sparse representation in the feature space. The main process is shown in Fig. 1. Once the feature space is constructed using the proposed structure-aware descriptor, a low-rank and sparse modeling is explored to capture both the local feature saliency and the global structure saliency.

4.1. Low-rank and sparse modeling

Generally speaking, when using the low-rank property to help detect salient objects, the most critical principle to achieve desirable performance is to make sure that the salient regions are presented with strong anisotropic strength, in other words, feature representations of salient object should be regarded as sparse components by low-rank decomposition. From the viewpoint of the matrix decomposition, the feature matrix \mathbf{F} can be divided into a low-rank component and a sparse component $\mathbf{F} = \mathbf{L} + \mathbf{S}$, with \mathbf{L} and \mathbf{S} corresponding to the correlated elements and independent elements respectively. The traditional low-rank decomposition is NP-hard, it is proved that this problem can be approximated by the nuclear norm $\|\mathbf{L}\|_*$ and L1-norm $\|\mathbf{S}\|_1$ using the following formulation:

$$(\mathbf{L}^*, \mathbf{S}^*) = \arg \min_{\mathbf{L}, \mathbf{S}} (\|\mathbf{L}\|_* + \lambda \|\mathbf{S}\|_1). \quad (4)$$

And this newly-formulated problem can be solved using several methods, such as randomized approximate matrix decomposition (Zhou and Tao, 2011). Because of the rareness attribute of salient object, values remaining in sparse matrix \mathbf{S} reflect the saliency degree of its corresponding parts. More residual in matrix \mathbf{S} means higher saliency value and vice versa.

Traditional methods tend to solve the aforementioned low-rank decomposition problem by minimizing the sum of the kernel norm and the L_1 norm without the explicit control on the rank level (Yan et al., 2010; Shen and Wu, 2012). In sharp contrast, we define our low-rank and sparse model as:

$$\mathbf{F} = \mathbf{L} + \mathbf{S} + \mathbf{G}, \text{ s.t. } \text{rank}(\mathbf{L}) \leq r, \text{card}(\mathbf{S}) \leq c, \quad (5)$$

Table 1

Running time for computing mesh saliency (s), including the computing time of descriptors (including shape expression, bi-harmonic distance and shape context), local feature saliency (Loc) and global structure saliency (Glo).

Models (# V)	Descriptor			Saliency	
	Exp	Dist	Cont	Loc	Glo
Dinosaur (14k)	2.02	3.19	4.10	2.85	4.33
Skull (20k)	3.21	4.72	5.65	3.57	5.54
Venus (22k)	3.73	5.16	5.98	3.9	5.77
Lion (32k)	5.01	10.31	11.58	5.28	8.11
Armadillo (34k)	5.65	11.68	12.65	5.39	7.19

where \mathbf{G} is the error matrix. This low-rank decomposition problem can be solved approximately using the fix-one-and-solve-another strategy iteratively as follows:

$$\begin{cases} \mathbf{L}_t = \arg \min_{\text{rank}(\mathbf{L}) \leq r} \|\mathbf{F} - \mathbf{L} - \mathbf{S}_{t-1}\|_F^2 \\ \mathbf{S}_t = \arg \min_{\text{card}(\mathbf{S}) \leq c} \|\mathbf{F} - \mathbf{L}_t - \mathbf{S}\|_F^2 \end{cases}, \quad (6)$$

where $\|\cdot\|_F$ is the Frobenius norm, the rank constraint r and the cardinality constraint c are used to explicitly control the low-rank degree and the sparse degree respectively. As for numerical computation, we employ the GoDec method (Zhou and Tao, 2011) to efficiently accelerate this optimization process (and it may be noted that, its convergence can be guaranteed).

4.2. Multi-scale saliency detection

As mentioned before, apart from the highly discriminative power for salient objects and non-salient parts, another advantage of our descriptor is its simultaneous integration of both local and global information. The feature space \mathbf{F} is then organized into the local feature subspace \mathbf{F}_l and the global structure subspace \mathbf{F}_g .

For the local feature subspace \mathbf{F}_l , we apply the low-rank and sparse model in Eq. (5) to evaluate the sparse matrix \mathbf{S}_l . Then, the local feature saliency of each vertex is evaluated using $L1$ -norm of columns in matrix \mathbf{S}_l

$$\mathcal{S}_l(v_i) = \sum_{j=1}^{d_l} (|\mathbf{S}_l(j, 2i)| + |\mathbf{S}_l(j, 2i-1)|), \quad (7)$$

where $\mathcal{S}_l(v_i)$ indicates the local feature saliency of vertex v_i , d_l is the dimension of the local feature subspace, and $\mathbf{S}_l(\cdot, \cdot)$ is the element of sparse matrix \mathbf{S}_l .

Similarly, the global structure saliency can be evaluated using the sparse matrix \mathbf{S}_g obtained in the global structure subspace. The global structure saliency of vertex v_i is

$$\mathcal{S}_g(v_i) = \sum_{j=1}^{d_g} (|\mathbf{S}_g(j, 2i)| + |\mathbf{S}_g(j, 2i-1)|), \quad (8)$$

where d_g is the dimension of the global structure subspace and $\mathbf{S}_g(\cdot, \cdot)$ is the element of sparse matrix \mathbf{S}_g .

5. Experimental results

We now demonstrate the performance in this part. All the experiments throughout the paper are conducted on a 3.5 GHz Intel(R) Core(TM) i7 computer with 16G memory. Table 1 shows our computing time, including the computing time of descriptors, local feature saliency and global structure saliency.

5.1. Parameter selection

There are several parameters in our saliency detection framework, however, most of them are fixed or set empirically. The low level model M^l is reconstructed using the first 200 low frequency eigenbases, and the multi-scale shape context is constructed with $k = 10$. Especially, the rank constraint r is used to intrinsically emphasize the stronger similarities, meanwhile, the cardinality constraint c provides a flexible way to control the uniqueness degree of the saliency to be detected. In our experiments, we only release the rank constraint r , and set the cardinality constraint k with a hard threshold $5n$ to alleviate the complexity. Users can adjust r to obtain different saliency results. Fig. 3 shows the results of local feature saliency and global structure saliency with different rank constraints, respectively.

We provide an automatical rank-evaluation to estimate the appropriate r for a given model (this is the default setting, if there is no special explanation). Firstly, we cluster the column vectors of pending subspace \mathbf{F}_l (or \mathbf{F}_g) into several groups

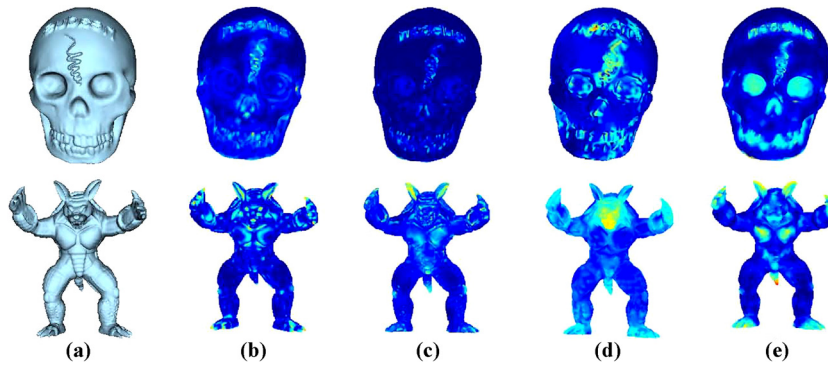


Fig. 4. A comparison of the local saliency results produced by related methods. (a) Original models. (b) The method of Lee et al. (2005). (c) The method of Wu et al. (2013). (d) The method of Song et al. (2014). (e) Our method ($r = 2, 5$).

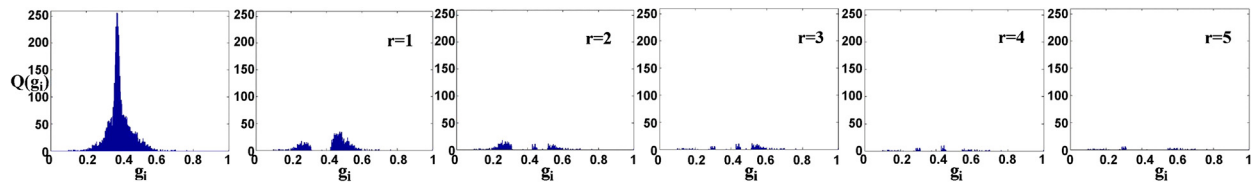


Fig. 5. The illustration of rank evaluation. The number of iterations is the evaluated rank constraint.

$\{g_i\}$ using k-means (Ferreira et al., 2013). Then, we represent the number of vectors in each group using function $Q(g_i)$. Thirdly, we select the group g_k with $\max\{Q(g_i)\}$, and formulate our evaluation equation as

$$Q'(g_i) = Q(g_i) - Q(g_k) * \exp\left(\frac{-\|g_k - g_i\|^2}{2\delta^2}\right), \quad (9)$$

where $\|g_k - g_i\|$ is the distance between group g_k and g_i (the Euclidean distance between two vectors that represent the barycenters of the groups), and δ is the Gaussian weight (we set $\delta = 0.1$ in our experiments). Fourthly, we shall repeat the third step until $\max\{Q(g_i)\}$ approximates zero, and the number of iterations is the rank constraint that we have evaluated. For the purpose of convenience, we illustrate this process in one dimension using histogram (see Fig. 5). Although this is not an accurate calculation of the rank constraint, it provides a feasible reference, and we can find the appropriate rank constraint approaches to the estimated rank.

5.2. Comparisons with other approaches

To our best knowledge, we are the first to propose a saliency framework that can detect both local feature saliency and global structure saliency. There are several saliency detection methods that can detect local feature saliency from different points of view, but no global structure saliency is proposed. We compare our local feature saliency with three excellent saliency detection schemes proposed in Lee et al. (2005), Wu et al. (2013), Song et al. (2014). The method in Lee et al. (2005) can obtain reasonable saliency results with low-level human visual attention in a multi-scale way. But since they focus mainly on the local feature attribute such as mean curvature, this method can only detect local saliency with different scales. The method in Wu et al. (2013) detects the saliency using both the local contrast and global rarity. The method in Song et al. (2014) achieves better results in a multi-scale way, however, high level cues may be lost. Moreover, all of them are essentially the saliency detection based on local features, but at different aspects. Fig. 4 shows the comparisons between our method and above two methods. Since our method integrates both multi-scale information and multi-level information, the result of our method is highly competitive, while having the capability of capturing more local features, such as the eyes of Skull model. Moreover, our method can also detect the global structure saliency (see Fig. 3), which can help understand the shape better from the aspect of global structures. More results are demonstrated in Fig. 6.

5.3. Applications

The detected saliency can be further used to help several applications with improved results, such as mesh smoothing, adaptive remeshing, mesh simplification, structure-driven mesh segmentation and other feature-sensitive (or structure-aware) tasks. An extensive survey is beyond the technical scope of this paper, and in the interest of space we can only demonstrate the mesh smoothing and mesh segmentation to reveal the effectiveness of our local feature saliency and global structure saliency respectively.



Fig. 6. Mesh saliency results on various models. The top row shows local feature saliency (from left to right $r = 3, 3, 4, 4, 3$, respectively), and the bottom row shows global structure saliency (from left to right $r = 24, 23, 30, 25, 24$, respectively).

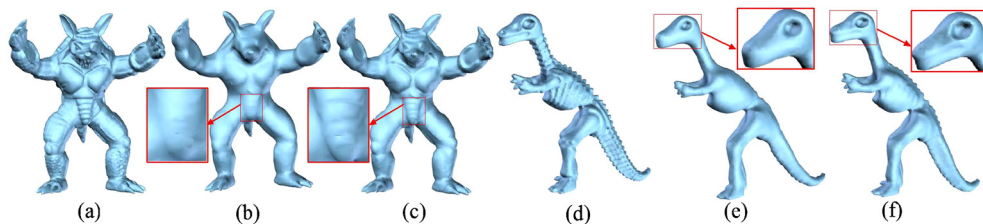


Fig. 7. Mesh smoothing results. (a) and (d): Original models. (b) and (e): Laplacian smoothing. (c) and (f): Our smoothing results based on saliency detection (the Laplacian smoothing with our local feature saliency).

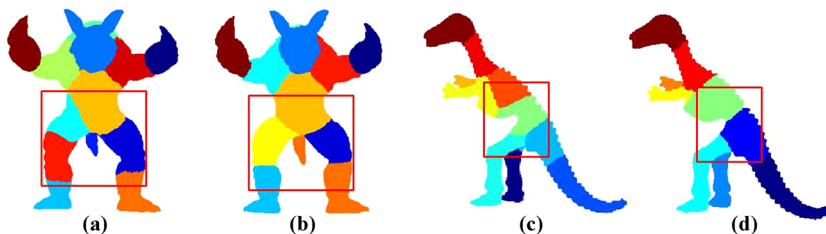


Fig. 8. Mesh segmentation results. (a) and (c): k-Means segmentation. (b) and (d): Our segmentation results based on saliency detection (the global structure saliency guided k-means segmentation).

Mesh smoothing appears to be a mature topic in graphics, and there are plenty of research works being done already. However, most of them are based on local geometric metrics, such as curvature and normal (Desbrun et al., 1999), and little is focused on perception-based metrics, such as saliency. Hence, our work provides a potential application in mesh processing from the perspective of visual perception. We use the Laplacian mesh smoothing (Nealen et al., 2006) with constraint of the local feature saliency (i.e., the saliency values are treated as soft constraints) to show the effectiveness of our local feature saliency. Fig. 7 shows that our saliency-guided mesh smoothing removes the non-salient details, while preserving the visually significant features.

There are also a large body of literatures on mesh segmentation, such as k-means (Shlafman et al., 2002), region growing (Ji et al., 2006), etc. Take k-means for example, we use our global structure saliency to guide the k-means segmentation (Shlafman et al., 2002) (i.e., the saliency values work together with geometry information). Our saliency-guided segmentation is sensitive to the visually significant structures. From Fig. 8, we could clearly see that the boundaries of different segmented parts are closely related to the structures, and our method can improve the segmentation results effectively. The global structure saliency can also be used to assist other segmentations, such as structure-driven co-segmentation, which deserves our further research in the near future.

6. Conclusion

In this paper, we have presented a versatile method to detect multi-scale saliency of 3D models, including the local feature saliency and the global structure saliency. The critical and novel technical elements include the structure-aware shape

descriptor embracing both multi-scale and multi-level information, the feature space that consists of a local feature subspace and a global structure subspace, and the low-rank approximation and sparse representation based saliency detection. Comprehensive experiments and extensive comparisons with other state-of-the-art methods have demonstrated some key advantages of our method in terms of flexibility, reliability, robustness, and versatility.

Acknowledgement

This work is partially supported by NSFC grants (61300083, 61272371, 61432003, 61173103, 61370143), the open funding project of State Key Laboratory of Virtual Reality Technology and Systems (61190120, 61190125), and the US National Science Foundation grants IIS-(0949467, 1047715, 1049448).

References

- Borji, A., Sihite, D.N., Itti, L., 2012. Salient object detection: a benchmark. In: *European Conference on Computer Vision*, pp. 414–429.
- Bronstein, A.M., Bronstein, M.M., Guibas, L.J., Ovsjanikov, M., 2010. Shape google: geometric words and expressions for invariant shape retrieval. *ACM Trans. Graph.* 30 (1), 1–20.
- Chen, X., Huo, H., Fang, T., Li, D., 2007. New approach to texture saliency based on intrinsic relationship among texture features. In: *Remote Sensing and GIS Data Processing and Applications, and Innovative Multispectral Technology and Applications*.
- Cheng, M., Zhang, G., Mitra, N., Huang, X., Hu, S., 2011. Global contrast based salient region detection. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 409–416.
- Desbrun, M., Meyer, M., Schröder, P., Barr, A.H., 1999. Implicit fairing of irregular meshes using diffusion and curvature flow. In: *SIGGRAPH*, pp. 317–324.
- Duan, L., Wu, C., Miao, J., Qing, L., Fu, Y., 2011. Visual saliency detection by spatially weighted dissimilarity. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 473–480.
- Einhüser, W., König, P., 2003. Does luminance-contrast contribute to a saliency map for overt visual attention. *Eur. J. Neurosci.* 17 (5), 1089–1097.
- Feixas, M., Sbert, M., González, F., 2009. A unified information-theoretic framework for viewpoint selection and mesh saliency. *ACM Trans. Appl. Percept.* 6 (1), 1–23.
- Ferreira, N., Klosowski, J.T., Scheidegger, C.E., Silva, C.T., 2013. Vector field k-means: clustering trajectories by fitting multiple vector fields. *Comput. Graph. Forum* 32 (3), 201–210.
- Gal, R., Cohen-Or, D., 2006. Salient geometric features for partial shape matching and similarity. *ACM Trans. Graph.* 25 (1), 130–150.
- Goferman, S., Zelnik-Manor, L., Tal, A., 2012. Context-aware saliency detection. *IEEE Trans. Pattern Anal. Mach. Intell.* 34 (10), 1915–1926.
- Hou, X., Zhang, L., 2007. Saliency detection: a spectral residual approach, pp. 1–8.
- Itti, L., Koch, C., Niebur, E., 1998. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* 20 (11), 1254–1259.
- Ji, Z., Liu, L., Chen, Z., Wang, G., 2006. Easy mesh cutting. *Comput. Graph. Forum* 25 (3), 283–291.
- Kim, Y., Varshney, A., 2006. Saliency-guided enhancement for volume visualization. *IEEE Trans. Vis. Comput. Graph.* 12 (5), 925–932.
- Kim, Y., Varshney, A., 2008. Persuading visual attention through geometry. *IEEE Trans. Vis. Comput. Graph.* 14 (4), 772–782.
- Kim, Y., Varshney, A., Jacobs, D.W., Guimbretière, F., 2010. Mesh saliency and human eye fixations. *ACM Trans. Appl. Percept.* 7 (2), 1–13.
- Koch, C., Ullman, S., 1987. Shifts in selective visual attention: towards the underlying neural circuitry. In: *Matters of Intelligence*, pp. 115–141.
- Lee, C.H., Varshney, A., Jacobs, D.W., 2005. Mesh saliency. *ACM Trans. Graph.* 24 (3), 659–666.
- Leifman, G., Shtrom, E., Tal, A., 2010. Surface regions of interest for viewpoint selection. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 414–421.
- Le Meur, O., Le Callet, P., Barba, D., Thoreau, D., 2006. A coherent computational approach to model bottom-up visual attention. *IEEE Trans. Pattern Anal. Mach. Intell.* 28 (5), 802–817.
- Lin, Z., Chen, M., Wu, L., Ma, Y., 2009. The augmented Lagrange multiplier method for exact recovery of corrupted low rank matrices. *UIUC Technical Report*.
- Lipman, Y., Rustamov, R., Funkhouser, T., 2010. Biharmonic distance. *ACM Trans. Graph.* 29 (3), 1–11.
- Liu, T., Yuan, Z., Sun, J., Wang, J., Zheng, N., Tang, X., Shum, H., 2011. Learning to detect a salient object. *IEEE Trans. Pattern Anal. Mach. Intell.* 33 (2), 353–367.
- Ma, Y., Zhang, H., 2003. Contrast-based image attention analysis by using fuzzy growing. In: *ACM International Conference on Multimedia*, pp. 374–381.
- Meyer, M., Desbrun, M., Schröder, P., Barr, A.H., 2002. Discrete differential-geometry operators for triangulated 2-manifolds. In: *VisMath*, pp. 1–26.
- Nealen, A., Igarashi, T., Sorkine, O., Alexa, M., 2006. Laplacian mesh optimization. In: *International Conference on Computer Graphics and Interactive Techniques in Australasia and Southeast Asia*, pp. 381–389.
- Perazzi, F., Krahenbuhl, P., Pritch, Y., Hornung, A., 2012. Saliency filters: contrast based filtering for salient region detection. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 733–740.
- Reuter, M., Wolter, F.E., Peinecke, N., 2006. Laplace–Beltrami spectra as ‘shape-DNA’ of surfaces and solids. *Comput. Aided Des.* 38 (4), 342–366.
- Rong, G., Cao, Y., Guo, X., 2008. Spectral mesh deformation. *Vis. Comput.* 24 (7–9), 787–796.
- Shen, X., Wu, Y., 2012. A unified approach to salient object detection via low rank matrix recovery. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 853–860.
- Shilane, P., Funkhouser, T., 2007. Distinctive regions of 3d surfaces. *ACM Trans. Graph.* 26 (2). Article 7, 1–16 p.
- Shlafman, S., Tal, A., Katz, S., 2002. Metamorphosis of polyhedral surfaces using decomposition. *Comput. Graph. Forum* 21 (3), 219–228.
- Song, R., Liu, Y., Martin, R., Ronsion, P., 2014. Mesh saliency via spectral processing. *ACM Trans. Graph.* 33 (1), 1–17.
- Sun, X., Yao, H., Ji, R., Xu, P., Liu, X., Liu, S., 2010. Saliency detection based on short-term sparse representation. In: *IEEE International Conference on Image Processing*, pp. 1101–1104.
- Torralba, A., Castelhano, M.S., Oliva, A., Henderson, J.M., 2006. Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. *Psychol. Rev.* 113 (4), 766–786.
- Walther, D., Koch, C., 2006. Modeling attention to salient proto-objects. *Neural Netw.* 19 (9), 1395–1407.
- Wei, Y., Wen, F., Zhu, W., Sun, J., 2012. Geodesic saliency using background priors. In: *European Conference on Computer Vision*, pp. 29–42.
- Wu, J., Shen, X., Zhu, W., Liu, L., 2013. Mesh saliency with global rarity. *Graph. Models* 75 (5), 255–264.
- Yan, J., Zhu, M., Liu, H., Liu, Y., 2010. Visual saliency detection via sparsity pursuit. *IEEE Signal Process. Lett.* 17 (8), 659–666.
- Zhang, Q., Li, B., 2012. Mining discriminative components with low-rank and sparsity constraints for face recognition. In: *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1469–1477.
- Zhou, T., Tao, D., 2011. GoDec: randomized low-rank and sparse matrix decomposition in noisy case. In: *IEEE International Conference on Machine Learning*, pp. 33–40.