Computers & Graphics 57 (2016) 12–30

Contents lists available at ScienceDirect

Computers & Graphics

journal homepage: www.elsevier.com/locate/cag

Survey Paper

Shape context based mesh saliency detection and its applications: A survey ☆

Xianyong Liu [a], Ligang Liu [b], Weijie Song [c], Yanping Liu [d], Lizhuang Ma [a,*]

[a] School of Computer Science and Technology, Shanghai Jiao Tong University, Shanghai, China
[b] School of Mathematical Sciences, University of Science and Technology of China, Hefei, China
[c] School of Science, Northwestern Polytechnical University, Xi'an, China
[d] School of Marine Science and Technology, Zhejiang Ocean University, Zhoushan, China

ARTICLE INFO

Article history:
Received 30 October 2015
Received in revised form
1 March 2016
Accepted 9 March 2016
Available online 18 March 2016

Keywords:
The human visual system
Geometric analysis
Mesh saliency
Shape descriptor
Shape spectrum

ABSTRACT

Mesh saliency was introduced and joined the community of computer graphics ten years ago, which can benefit various applications, for instance, mesh reduction, mesh segmentation, self-similarity matching, scan integration, volume rendering, 3D printing, etc. Before, saliency detection had been successfully applied to image processing and pattern recognition to study how the world is perceptually intelligible for robots. In contrast to color of images and coherence of videos, geometric signals are defined with two-dimensional manifolds whose discrete representation is irregular, leading differences to the nature and difficulties to the solution of mesh saliency. To tackle the challenge, the last decade has witnessed significant advances in mesh saliency detection. However, a survey of recent advances in mesh saliency detection as well as its applications does not yet exist to date. This paper provides a first and comprehensive reference source of shape context based mesh saliency for researchers from a wide range of domains, including but not limited to computer graphics and vision. It reviews main contributions, advantages, drawbacks, and applications of known mesh saliency detection methods and discusses current trends and outlook for future study.

1. Introduction

> "*Seeing comes before words* [1]." "*Most of what our eyes take in is filtered, as we cannot process all that is within the field of our vision* [2]." John Berger

The human visual system (HVS) is a significant part of how we process sensorial information and a key factor in how we learn. In order to reveal the brain's perception processing mechanisms, over the past decades, a large number of neurophysiological and cognitive neuroscience researches have provided in-depth and detailed experimental data and theoretical models. For example, *Science*, *Nature*, and *Neuron* published several remarkable findings [3–6] relevant to the behaviors within the process of the HVS. The visual media, including digital images, videos, and three-dimensional (3D) models, contains superfluous visual information, of which the portion that is visually interesting is filtered and referred to *saliency* [2]. *Saliency detection*, thereby, imitates the ways of seeing and becomes an interdisciplinary scientific study of theoretical computer science and the human perception.

This paper, in particular, surveys recent developments in *shape context based* mesh saliency. By "mesh saliency" we mean enabling a machine system to automatically reason about which points or regions of a 3D *polygonal mesh* are perceptually important [7], so the artificial intelligence can observe like the humans. The term of "perceptually" distinguishes the topic surveyed from others detecting *interest points* [8] or *keypoints* [9], though they do have overlaps (see Section 7.1).

Moreover, the task of mesh saliency is relevant to a relatively mature research area called *visual saliency* in image analysis (see Section 7.2), as both are related to the nature of the HVS but from different perspectives. Compared to 2D images, 3D meshes have several deviations, leading challenges to mesh saliency detection. For instance, (i) 2D images and videos have regular discrete representations, which are irregular for 3D meshes [10]; (ii) 3D meshes provide more geometrical (depth) information which are hidden from static 2D images, also 3D shapes are believed to encode surface metric dimensions unambiguously [11]; (iii) features extracted from 3D models are not affected by scale, rotation, and illumination [12]; (iv) salient objects (or regions) are

☆This article was recommended for publication by Wenping Wang.

* Corresponding author.
E-mail address: ma-lz@cs.sjtu.edu.cn (L. Ma).

http://dx.doi.org/10.1016/j.cag.2016.03.001
0097-8493/

empirically centered of 2D images, while such an assumption does not hold for 3D meshes. Although the discrepancies between 2D images and 3D meshes are noticeable, when we look for a solution to the issue of saliency detection the principals get intimately tied up – saliencies originate from visual *distinctness*, *rarity*, or *surprise*, and often are attributed to variations in data attributes at multi-scales and multi-levels [13,14].

The motivation of this work is three-fold. First, the present advances in modeling, digitizing, and visualizing techniques have led to an increasing amount of 3D models, both on the Internet and in domain-specific databases. As a result, there is a growing need for the task of 3D geometric analysis and processing, where the use of mesh saliency has been found within a wide range of applications. Second, since Lee et al. [15] firstly introduced mesh saliency into the graphics community ten years ago, the last decade has witnessed significant progress toward the development of new methods and applications. Meanwhile, we strongly believe that the long and challenging journey to mesh saliency detection has just started, as the highly relevant studies such as visual saliency and scene saliency are drawing more and more attention. Third, so far there does not yet exist a survey of recent mesh saliency detection as well as its applications. To avoid that the wheel would be reinvented, the paper provides a first and comprehensive source of shape context based mesh saliency.

So, the paper first reviews 35 representative methods published from 2005 to 2015. Then, with 7 examples, it demonstrates how mesh saliency significantly improves mesh simplification, mesh segmentation, mesh resizing, normal enhancement, volume rendering, 3D printing, and scan integration. Next, 12 feasible ways related to qualitative evaluation and quantitative evaluation of mesh saliency are discussed. Lastly and more importantly, with 4 key insights, the survey sheds light on current trends and outlook for future study.

In particular, we noticed that the majority methods of mesh saliency detection are local contrast based and for capturing "*saliency seeds*", i.e., the most representative salient elements. Consequently, often the saliencies detected are too dispersed. In other words, their scale is small, and their number is large. Even, the structure of 3D models is somewhat "damaged". The reason is that the recent study on mesh saliency does not fully appreciate the *two-stage characteristic* of the HVS, which has been well discussed in other areas such as neurophysiology and cognitive neuroscience. For example, literatures [3–6] showed the role of the bottom-up and top-down visual attention selection in the process of the HVS. Remarkably, Koch and Poggio [16] pointed out that the HVS is essentially served by a two-stage prediction network of the brain, measuring the local contrast first and then the global contrast. The finding, certainly, inspired Cheng et al. [13] to investigate how the network guides visual saliency detection in image processing. Such a useful comprehension to the HVS, in turn, becomes the fundamental criterion for classifying mesh saliency methods in the paper. To make an intuitive test of the comprehension, at the beginning of the work a user study was conducted, and the statistics, which further prove that the criterion we set is reasonable, are shown in Fig. 1. Also, the work informs that developing global contrast based means for perceiving "*saliency structures*", i.e., the most representative salient components, is the next step in mesh saliency detection.

Finally, it is worthwhile mentioning that in the recent years range images of 3D point cloud are gaining more attention researchers paid to computer graphics and vision, especially along with the emergence of low cost 3D acquisition/virtual reality (VR) devices. Thus, it is seen that in computer graphics the interest of saliency detection grows from single objects to larger scenes, where mesh saliency can facilitate and cooperate with the studies of *scene saliency* and *object recognition* (see the latest survey in [11]), which at present become more promising research directions (see more discussions in Sections 7.3 and 7.4).

2. Overview

2.1. Notations

To make this survey self-contained, here we introduce some basic notations. $\mathcal{S}$ represents a 3D smooth surface that is a compact 2-dimensional Riemannian manifold. In differential geometry, "sharpness" of a point on curves or surfaces is measured by *curvature*. In the case of surfaces, this measure is the so-called *normal curvature*, which smoothly varies within $[k_1, k_2]$ where the values are the *principal curvatures*. The *mean curvature* and *Gauss curvature*, thus, are calculated as $(k_1+k_2)/2$ and $k_1 \cdot k_2$, respectively. In the paper, a surface $\mathcal{S}$ is discretized and approximated by a triangulated mesh $\mathcal{M}\{\mathrm{V}, \mathrm{E}, \mathrm{T}\}$, wherein V, E, and T are the set of vertices, edges, and triangles, respectively. In the sequel, $\mathbf{n}(\cdot)$ represents the normal of a vertex $\mathbf{v} \in \mathrm{V}$, an edge $\mathbf{e} \in \mathrm{E}$, or a triangle $\mathbf{t} \in \mathrm{T}$.

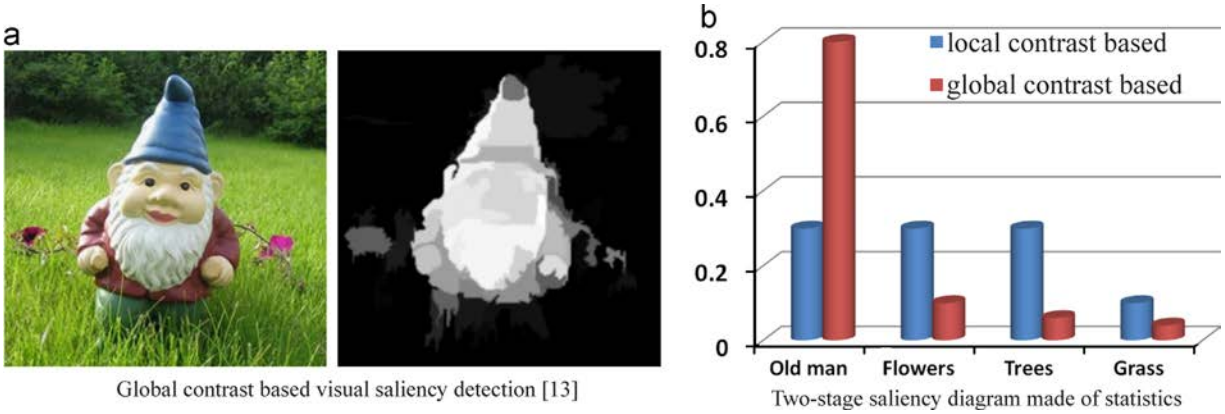

Fig. 1. An illustration of the two-stage characteristic of the human visual system (HVS) [16,3–6]: a fast, pre-attentive, bottom-up, local contrast based saliency extraction; followed by a slower, task dependent, top-down, global contrast based saliency extraction. (a) When we looked at the leftmost image, in the first stage the eye gaze moved quickly throughout the old man, flowers, trees, and grass, searching for regions that are locally salient; then in the second-stage the gaze clung to the old man, which among what had been observed was more globally interesting. (b) A quantitative verification of the result in (a), with a user study participated by 150 students aged 17–29. First, the original image was shown to each participant for about 7 s. Then, everyone was queried "which object of the image did you mainly pay attention to at the beginning and at the end of your looking?". The right diagram shows statistics of the two choices, where the cylinders indicate the percentile scores of participants who chose the old man, flowers, trees, and grass, respectively.

Given a mesh $\mathcal{M}$, a global (resp. local) *shape descriptor* Γ, encoding the intrinsic shape contexts of $\mathcal{M}$, is a transformation function mapping the entire (resp. partial) mesh into a k-dimensional feature space $\mathbb{R}^k$. Specifically, a mesh saliency map Γ_s, discussed in the paper, represents a real-valued global shape descriptor that is parameterized by points or regions.

The 2D Gaussian smoothing kernel is defined as $G(\mathbf{x},\sigma)=\frac{1}{2\pi(\sigma)^2}\exp[-\langle\mathbf{x},\mathbf{x}\rangle/2\sigma^2]$, where $\langle\cdot,\cdot\rangle$ stands for the inner product, σ determines standard deviation of the Gaussian distribution, and $\mathbf{x}$ is the distribution variable in the domain. We refer to σ as the *inner scale* or shortly *scale*. Since the Laplacian commutes with convolution, given a geometric signal described by a shape descriptor Γ, it follows that $\nabla^2(G(\mathbf{x},\sigma)*\Gamma)=(\nabla^2G(\mathbf{x},\sigma))*\Gamma$, where ∇^2 (or Δ) is the Laplace–Beltrami operator (LBO), $*$ denotes the convolution operator, and $\nabla^2G(\mathbf{x},\sigma)$ is the well-known Laplacian of Gaussian that can be approximated by the differences of two Gaussian functions (DoG), i.e., $-(\alpha-1)\sigma^2\nabla^2G(\mathbf{x},\sigma)\approx G(\mathbf{x},\sigma)-G(\mathbf{x},k\sigma)$, with k being an infinitesimally small constant.

At any vertex $\mathbf{v}$, its neighborhood $\mathrm{N}(\mathbf{v},\sigma_c)=\{\mathbf{u}|\,\|\mathbf{u}-\mathbf{v}\|<\sigma_c\}$ is a set of vertices within the distance σ_c measured under either Euclidean or geodesic distance. In the domain where a surface descriptor Γ is known, the Gaussian convolution $G_w(\mathbf{v},\mathbf{x},\sigma_c)$ is defined as

$$\frac{\sum_{\mathbf{u}\in\mathrm{N}(\mathbf{v},\sigma_c)}\Gamma(\mathbf{u})G(\mathbf{x}(\mathbf{u},\mathbf{v}),\sigma_c)}{\sum_{\mathbf{u}\in\mathrm{N}(\mathbf{v},\sigma)}G(\mathbf{x}(\mathbf{u},\mathbf{v}),\sigma_c)},$$

where σ_c is the scale. By analogy, the bilateral Gaussian convolution $G_b(\mathbf{v},\mathbf{x},\sigma_c,\mathbf{y},\sigma_s)$ is formulated as

$$\frac{\sum_{\mathbf{u}\in\mathrm{N}(\mathbf{v},\sigma)}\Gamma(\mathbf{u})G(\mathbf{x}(\mathbf{u},\mathbf{v}),\sigma_c)G(\mathbf{y}(\mathbf{u},\mathbf{v}),\sigma_s)}{\sum_{\mathbf{u}\in\mathrm{N}(\mathbf{v},\sigma)}G(\mathbf{x}(\mathbf{u},\mathbf{v}),\sigma_c)G(\mathbf{y}(\mathbf{u},\mathbf{v}),\sigma_s)},$$

where $\mathbf{y}$ and σ_s are similar to $\mathbf{x}$ and σ_c, respectively. Therefore, a *multi-scale* DoG feature space generated by Gaussian convolution is specified as $\{G_w(\mathbf{v},\mathbf{x},\sigma_c^i)-G_w(\mathbf{v},\mathbf{x},k\sigma_c^i)\}|_{i=1}^{s}$, with s being the number of scales.

Besides the *multi-scale* technique above, likewise Gaussian pyramid in imaging that breaks down an image into successively smaller groups of pixels, the *multi-level* method offers another way for analyzing 3D shapes, such that a 3D object is viewed over several different *level of details* (LOD). For instance [17], by successively removing collapse edges, a mesh $\mathcal{M}$ is remeshed at different levels, obtaining a series of meshes $\{\mathcal{M}^i\}|_{i=1}^{l}$, with l being the number of levels.

In the continuous case, given that f is a real-valued function ($f\in C^2$, defined on $\mathcal{S}$) the LBO is defined as $\Delta f:=div(grad\,f)$, with $grad\,f$ being the gradient of f and div the divergence on the surface. From a viewpoint of frequency, the spectra of the LBO contain intrinsic shape information. Thus, they are called "Shape-DNA" [18], being defined to be the family of eigenvalues of the Helmholtz equation, $\Delta f=\lambda f$, and consisting of a diverging sequence $0\le\lambda_1\le\lambda_2\cdots\uparrow+\infty$. In the discrete case, the LBO is denoted by $\mathbf{L}$ and often expressed using positive semi-definite (PSD) matrices [19], that is, the eigenvalues of $\mathbf{L}$ are non-negative.

2.2. Structure of survey

Before attention from computer graphics was paid to mesh saliency detection, the nature of the HVS had been widely studied and explored in different areas, such as psychophysics [3,4], neuroscience [5,6,20], biology [21,22], signal processing [23], computer vision [13], applied perception [24,25], and cognition science [26,27]. These prior studies, together, formed theoretical foundations of the research of mesh saliency, as illustrated by the first layer in Fig. 2. Although the first layer is skipped by the paper, it lays the background of this study to understand what the essential mechanism of the HVS is, and how the natural abilities associated with the HVS interpret 3D shapes or projected 2D images.

According to the conventional wisdom of the HVS explained in Section 1, this survey divides the type of mesh saliency methods into two categories, namely local contrast based (*LC*, see Section 3) and global contrast based (*GC*, see Section 4). The former captures

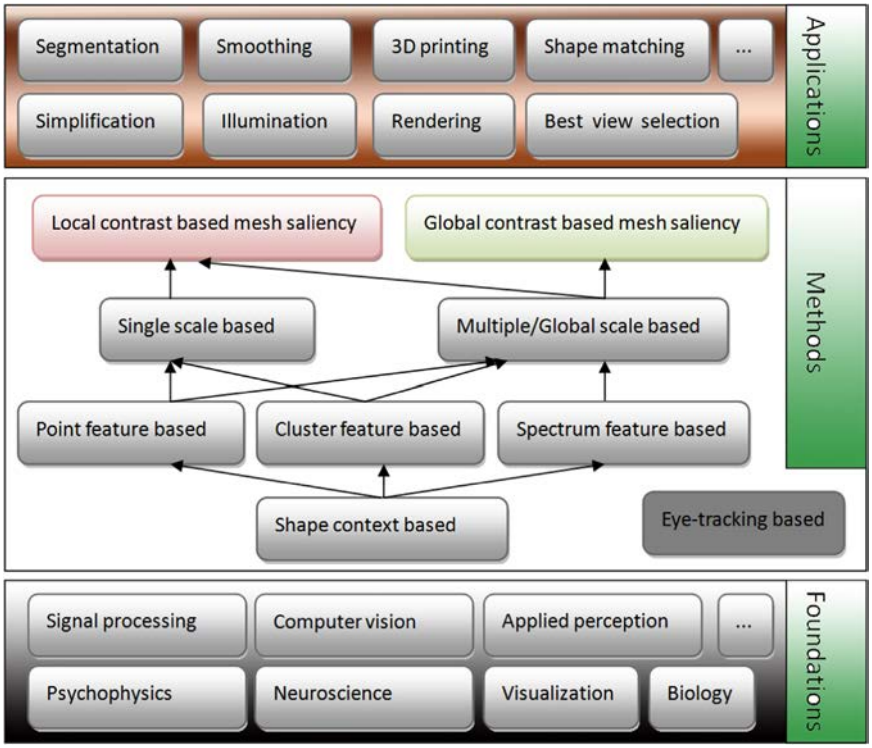

Fig. 2. The structure of this survey is constituted of three layers. From bottom to top: the first layer lists research foundations related to the topic of mesh saliency; the second layer is the taxonomy of mesh saliency detection methods; the third layer exhibits applications to which mesh saliency has been successfully applied.

"*saliency seeds*" [28], i.e., the most representative salient elements of surfaces, by contrast the latter detects "*saliency structures*" [29], i.e., the most representative salient components of meshes. So far, the majority of known approaches belong to the first category, as they are based on the exploitation of the first-stage of the HVS. However, recent advances in saliency detection with 2D images [13,30] and 3D models [31,32] manifest that the second category is meaningful and an emerging research [29], given that, compared to small-scale salient elements, large-scale salient structures are more advantageous for measuring and predicting visual fidelity when looking at objects.

The second layer in Fig. 2 demonstrates the taxonomy of existing mesh saliency methods and reflects our mental map of the known techniques in mesh saliency. Let symbol **Sly** be a saliency detection method. In the paper, three rules are designed to return the classification to which **Sly** belongs.

(1) **type**(**Sly**)={local contrast based (*LC*), global contrast based (*GC*)}.
(2) **scale**(**Sly**)={single scale based (*SS*), multiple scale based (*MS*), global scale based (*GS*)}.
(3) **signal**(**Sly**)={point feature based (*PF*), cluster feature based (*CF*), spectrum feature based (*SF*)}.

In fact, as indicated by the arrows in the taxonomy, these rules are not isolated but linked in a hierarchical relationship. Accordingly, the paper organizes mesh saliency detection methods in a *top-down manner*, i.e., they are ordered by **type** first, then by **scale**.

Note that a method **Sly** may take advantage of a combination of measures in different scales. In the case, for the sake of simplification, the fourth rule of "*lifting principal*" is adopted to assist in classifying the method. That is, **scale** (**Sly**) $= \max(SS, MS, GS)$, with respect to $SS < MS < GS$.

Also, it is worth mentioning that the survey only concentrates on methods whose source signal of saliency information is *shape context based*. As for the term "shape", to be on the same page we borrow the formal definition given by Dryden and Mardia in [33]: "all the geometric information that remains when location, scale, and rotational effects (i.e., Euclidean transformations) are filtered out from an object". In fact, besides using intrinsic shape contexts, the eye fixation data is another important type of signals for mesh saliency detection, which can be collected by using physical eye-tracking devices or laborious annotations. Such a method is often regarded as an auxiliary tool for many researches such as applied perception [34,25] and computer animation [35]. Considering a case that a subjective evaluation of the method not only is expensive but also seems unconscionable for automatic systems, methods of the *eye-tracking based* class are out of the scope of the survey.

The third layer in Fig. 2 will be viewed in Section 5, showing typical applications where mesh saliency has been successfully applied. Section 6 presents qualitative and quantitative evaluation of mesh saliency methods. Section 7 discusses current trends and outlook of mesh saliency detection. Conclusions are made in Section 8.

3. Local contrast based (*LC*) methods

The idea of local scale based methods was introduced by Koch and Ullman in [36], who had suggested that salient regions should be those that are distinctive from their immediate surroundings in different scales. Thus, the distinctness of a point or region can be measured by taking deviations where a wider *surround* filter is subtracted from a narrower *center* filter. Further, an *LC* method falls into three subgroups, namely *GS*, *MS*, and *SS*, by checking the following criteria:

(i) if the band width of a filter reaches the diameter of input domain, the method is global scale based (*GS*);
(ii) otherwise, if the deviations take place more than once, the method is multiple scale based (*MS*);
(iii) otherwise, the method is single scale based (*SS*).

Also, it aims to detect the most representative salient elements of a 3D object, which often are small-scale and dispersed.

3.1. Single scale based (SS) methods

Surface roughness variation was regarded as a metric for detecting salient region in [37,38], by making statistical considerations about the dihedral angles associated to faces and vertices of a mesh. Even though this measure is perception-inspired, it is, specifically, more suitable for assessing the degree of distortions produced by watermarking 3D meshes.

The method in [39] defined a local surface patch, which was associated with and represented by a shape descriptor. For each patch, a representative point at the center of its mass is chosen and associated with the highest curvature across the patch. Thus, a salient geometric feature is constructed by clustering together a set of descriptors that are interesting enough in the sense that they have a high curvature relative to their surroundings and a high variance of curvature values. When estimating the saliency of a patch, the area and the number of local minimum(s) or maximum(s) curvatures are considered as well. However, this means does not function well in online mode as the overhead is high in both time and space due to the involved implicit fitting scheme.

A learning based approach was introduced by Atmosukarto and Shapiro [40], where the absolute Gaussian curvature is used as low-level surface property that is convolved with a Gaussian filter to reduce noises. In the training step, for each vertex a histogram of neighboring low-level features is constructed. Then, for those manually marked points their histograms are saved and used for training. After that, the classifier is able to label each of the points of any 3D object as either salient or non-salient, providing a score of confidence for its decision. However, usually the supervised learning method is time consuming, thus its popularity is impaired.

Miao and Feng [41] proposed the drawing of perceptual saliency extremum lines that are the loci of extremum points of a function, describing the perceptual saliency along the principal curvature directions on the underlying 3D shape. For a vertex $\mathbf{v}$, the perceptual saliency measure is expressed as the local projection heights between $\mathbf{v}$ and its neighbors. Thus, a center-surrounding bilateral filter operator on Gaussian-weighted average of local projection heights between $\mathbf{v}$ and its neighbors is defined as $G_b(\mathbf{v}, \|\mathbf{u}-\mathbf{v}\|_2, \sigma_c, |\langle \mathbf{u}-\mathbf{v}, \mathbf{n}(\mathbf{v})\rangle|, \sigma_s)$. This method bears some interesting similarities with [42] in that the local projections or motions of vertices are more adequate than the conventional surface curvatures for assessing visual saliency regions.

Inspired by Harris and Stephens who had proposed an interest points detector for images [44], Pratikakis et al. [43,45] extended the Harris operator into 3D cases. Instead of exploiting the Hessian operator immediately, a continuous Gaussian function is applied to the integration of the derivatives. For a vertex $\mathbf{v}$, its Harris response is computed as $\Gamma_s(\mathbf{v}) = \det(\mathbf{A}) - \beta \cdot (\mathrm{tr}(\mathbf{A}))^2$, where $\mathbf{A}$ is the integrated matrix produced by the 3D Harris operator, and β is a range parameter. The 3D Harris operator is simply single scale based, whilst it can achieve results that are comparable to more complicated methods such as [7] (see Fig. 3). However, likewise the work in [39], the 3D Harris operator relies on the procedure of

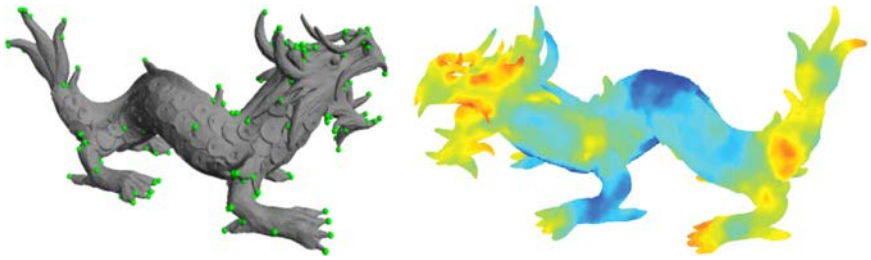

Fig. 3. In contrast to the global scale based method in [7] that captures regional importance (right, the warm is high and cold is low), the single scale based 3D Harris in [43] detects saliencies in the form of points (left, the marked spheres). Yet, in their results the locations with more geometric information and larger curvature are rather similar.

local implicit fitting that usually demands tremendous computation time and memory.

3.2. Multiple scale based (MS) methods

Lee et al. [15] presented a *center-surround operator* for detecting mesh saliency in a scale-dependent manner. There are five scales used, and the surface mean curvature is employed as source signal of filtering. The saliency of a vertex $\mathbf{v}$ at a single scale level is the absolute difference of two Gaussian filters. Hence, the final mesh saliency of $\mathbf{v}$ is evaluated by adding the saliency maps at all scales. The main disadvantage is that the Gaussian-weighted difference of fine and coarse scales can result in the same saliency values for two opposite and symmetric vertices because of the absolute difference in the computation.

Assuming that saliency map over a given mesh is a real-valued Morse function, Liu et al. [46] defined *salient critical points* as the points where the Morse function's gradient vanishes, namely, the minima, maxima, and saddles. The existing method in [15] is employed for detecting mesh saliency, while the saliency of a vertex $\mathbf{v}$ is changed to be the Gaussian-weighted average of the mean curvature difference between $\mathbf{v}$ and its neighboring vertices. The advantage of this method is that the bilateral filtering is applied to anisotropically smoothing the Morse function, so the number of critical points can be much lower than that yielded by using mean curvature function directly. The major drawback in this method's current implementation is that the number of salient critical points is dependent on the iteration number that cannot be accurately controlled.

Given a mesh $\mathcal{M}$, the method presented by Castellani et al. in [42] remeshes $\mathcal{M}$ at l different levels of decimation. Hence, salient point detection is then composed of two phases, namely intra-octave and inter-octave phases. In the first phase, instead of regarding surface curvatures as perceptual signals of saliency, it modifies the method in [15] by evaluating how significant the motion of a vertex along its surface normal is after filtering. At the ith-scale σ_i, for a point $\mathbf{v}$ of an octave mesh $\mathcal{M}^j$ the intra-octave saliency map is defined as

$$\Gamma_j^i(\mathbf{v}) = |\langle \mathbf{n}(\mathbf{v}), G_w(\mathbf{v},\mathbf{v},2\sigma_i) - G_w(\mathbf{v},\mathbf{v},\sigma_i)\rangle|.$$

In the second phase, there are five levels of decimation used for validating the points of each octave $\mathcal{M}^j$. Only those appearing at least in three octaves are retained. Such a "joint multi-scale" paradigm is robust to noises and insensitive to mesh resolution change, making it suitable for dealing with scanned point sets.

Euclidean distance metric is widely used with the DoG operator for approximating a time-varying diffusion process, which seems indiscreet in 3D cases. Thus, Zou et al. [9] introduced salient keypoint detection by searching local extrema in *Geodesic Scale Space* (GSS). If the geometric attribute value associated with a point $\mathbf{v}$ at scale σ_i is larger or smaller than any other vertices in $N(\mathbf{v},\sigma_i)\bigcup N(\mathbf{v},\sigma_{i-1})\bigcup N(\mathbf{v},\sigma_{i+1})$, $\mathbf{v}$ is selected as a local extremum. GSS is helpful for mesh saliency detection in the sense that ambiguous neighborhood produced by Euclidean measure usually leads to ambiguous saliency. However, GSS suffers from poor performance as it often requires a quadratic fitting performed around the neighborhood of each vertex.

The method called *MeshDoG* was introduced by Zaharescu et al. [48], where the scale space of a geometric property Γ is built progressively: $\Gamma_0 = \Gamma$, $\Gamma_1 = \Gamma_0 * G_w(\cdot,\cdot,\sigma)$, $\Gamma_2 = \Gamma_1 * G_w(\cdot,\cdot,\sigma)$, etc. Convolved functions are subtracted: $DoG_1 = \Gamma_1 - \Gamma_0$, $DoG_2 = \Gamma_2 - \Gamma_1$, etc., in order to obtain the difference of Gaussian operator. Since the same Gaussian kernel is used for convolutions, there is no need for normalization. Moreover, the *corner* characteristic is applied to refining the points of interest, which is defined as $\lambda_{\max}/\lambda_{\min}$ where the values are the eigenvalues of the Hessian matrix of the finally convolved function at each vertex. Later on, the computation of this method is improved in [49], by relaxing the constraint that the mesh vertices must correspond to a regular sampling of the underlying continuous surface. In addition, the gradient computation is cast into a least square minimization problem that can be efficiently estimated using a linear solver.

SIFT [50] is a well known operator for extracting distinctive features of 2D images. Godil and Wagan [51] extended 2D SIFT into 3D cases for identifying the local salient points of surfaces. Nevertheless, we are not convinced of how useful the method is, partially because it requires a pre-processing where a surface mesh should be voxelized, given that so far voxelizing a given two-dimensional manifold is still a challenging problem. By contrast, Ohbuchi et al. [52] applied the SIFT operator to projected 2D images of a 3D model viewed from uniformly sampled locations on a view sphere. Therefore, the salient map of the model is constructed by integrating the detected visual local features into a histogram using the Bag-Of-Features approach. This method, essentially, is an imaging-based visual saliency technique. Thus, some saliencies in spatial space might be overlooked in imaging space.

The spectral analysis tool called *admissible diffusion wavelets* (ADW) was proposed [53], where admissible wavelets are formulated as differences of adjacent scaling functions with zero means, such that the expensive QR decomposition is avoided. By the ADW, given a mesh $\mathcal{M}$, a function $f \in C^2(\mathcal{S})$, and a vertex $\mathbf{u} \in V(\mathcal{M})$, the scaling coefficient $S_f(j,\mathbf{u})$ is a smoothed representation of the function f at a scale j, and the wavelet coefficient $W_f(j,\mathbf{u})$ records the residual detailed information of f with respect to the scale j. Since the wavelet operator is closely related to the discrete Laplace operator $\mathbf{L}$, the ADW have similar effects with the difference of two Gaussian convoluted scalar fields. In short, for a given function defined at the vertices of a mesh, its wavelet coefficients encode the saliency information of multi-scale details. It is advantageous that a choice of the function is not fixed. Thus, a saliency map in terms of the ADW can be computed from other functions on the shape, such as texture and vertex coordinates.

The *Curvature Scale-Space 3D* (CS3) was presented in [47] (see Fig. 4). Let $f^j : V \to \mathbb{R}$ denote the smoothed discrete surface signal (curvatures) at level j, and define $\mathbf{F}^j = (f^j(\mathbf{v}_1),\ldots,f^j(\mathbf{v}_n))$, where

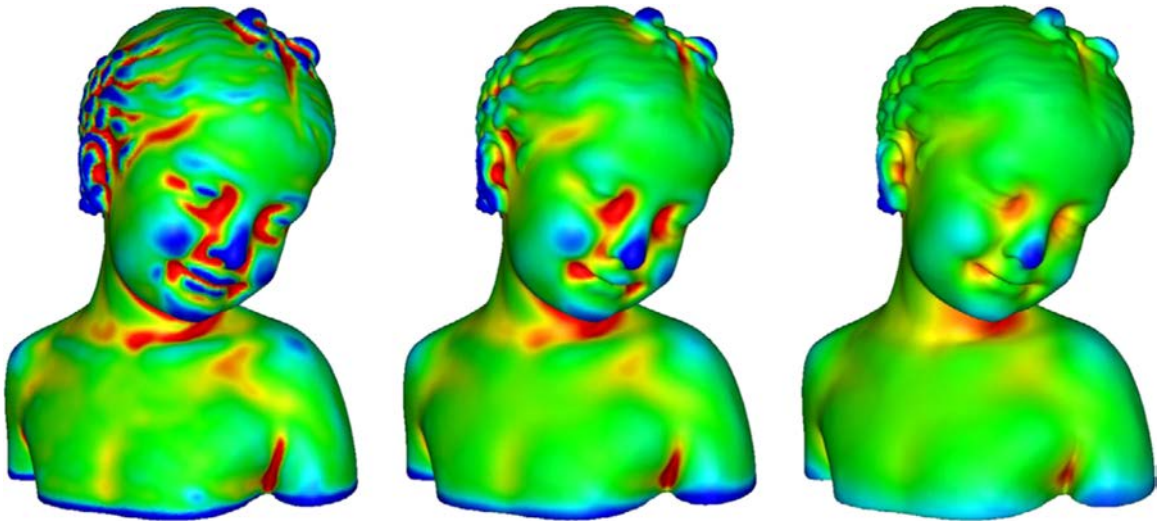

Fig. 4. The *Curvature Scale-Space 3D* (CS3) [47] representation holds a stack of Gaussian smoothed surface curvatures (from left to right, the scale is 3.0, 7.5, and 13.8, respectively) that can be used directly in multi-scale feature extraction and descriptor computations. It has been verified in [15] that salient features extracted in the multi-scale manner are meaningful to the human vision and cognition.

$n = |\mathrm{V}(\mathcal{M})|$. The implicit surface smoothing scheme is employed to obtain the smoothed surface signal, $\mathbf{F}^{j+1}$, at level $j+1$, by solving the following sparse system of linear equations

$$(\mathbf{I} - t_j\mathbf{L})\mathbf{F}^{j+1} = \mathbf{F}^{j},$$

where t_j is the time step , and $\mathbf{I}$ and $\mathbf{L}$ are the matrices of identity and discrete LBO, respectively. Thus, the multi-level and multi-scale representation of the surface signal of $\mathcal{M}$ is then given by the sequence $(\mathbf{F}^0, \ldots, \mathbf{F}^{l-1})$. In addition, the scale parameter t_j for each level j can be defined in terms of the variance of the transfer function at that level. Thus, the obtained sequence of scales, $(t_0, \ldots, t_{l-1})$, together with the stack of smoothed signals, $(\mathbf{F}^0, \ldots, \mathbf{F}^{l-1})$, define the CS3 representation. Finally, the differences between the smoothed signals at consecutive levels can be used to approximate the Laplacian of the input signal, whose local extrema can be regarded as saliencies of models. The major advantage of CS3 is the ability of performing automatic scale selection, which is indispensable to most partial shape matching tasks.

Jia et al. [54] proposed a region descriptor based on its vulnerability to a resizing direction and used this descriptor to compute the region's saliency based on its contrast to neighboring regions. The algorithm runs in three steps. First, a method based on face clustering is used to merge neighboring triangles into representative clusters, thus the contraction of the dual edge builds a hierarchical segmentation tree. Second, for a region r, its saliency value is based on its contrast to the neighboring regions as

$$\Gamma(r) = \frac{1}{n}\sum_{r_i \in N(r,\sigma)} g_i^{-1} w(r_i) d(r, r_i),$$

where $N(r,\sigma)$ denotes the set of regions within the Euclidean distance σ to the region r, n is the number of regions in $N(r,\sigma)$, $w(\cdot)$ represents the area of a region, $g_i = \|\mathbf{v}_r - \mathbf{v}_{r_i}\|_2$ with $\mathbf{v}_r$ and $\mathbf{v}_{r_i}$ respectively being the center point of regions r and r_i, and $d(\cdot,\cdot)$ is the absolute difference between vulnerabilities of two regions. Third, the final mesh saliency map is constructed by adding three saliency maps after applying the non-linear normalization of suppression. The method successfully overcomes the issue that existing saliency maps cannot be used for mesh resizing because they are computed without considering the resizing direction.

Instead of using conventional surface curvatures as [15], Miao et al. [55] proposed to compute the relief hight of each vertex over a given mesh, by assuming that $\mathcal{S}$, approximated by the mesh, is composed of a smooth base surface $\mathcal{B} \subset \mathbb{R}^3$ and a height function $f : \mathcal{S} \to \mathbb{R}$. Therefore, the relief height values can be estimated by minimizing the energy function as follows

$$\min \sum_{\mathbf{e} \in \mathrm{E}(\mathcal{M})} \left(f(\mathbf{v}_{\mathbf{e}_0}) - f(\mathbf{v}_{\mathbf{e}_1}) - \langle \mathbf{v}_{\mathbf{e}_0} - \mathbf{v}_{\mathbf{e}_1}, \mathbf{n}_{\mathcal{B}}(\mathbf{e})\rangle\right)^2,$$

where $\mathbf{v}_{\mathbf{e}_0}$ and $\mathbf{v}_{\mathbf{e}_1}$ are terminals of $\mathbf{e}$, and $\mathbf{n}_{\mathcal{B}}(\cdot)$ is the base's normal. Unfortunately, this work does not achieve satisfactory, even somewhat noisy, experimental results. Because the relief analysis is particularly suitable for surfaces abundant in detail, for instance, the challenging archaeological reliefs.

More recently, a method [56] was introduced to treat a polygonal mesh as a 3D point set. Likewise the one in [42], the method detects fine-scale salient features as well. Nevertheless, it differs that the distinctness measure is done by computing a per-region contrast from cluster uniqueness and spatial distribution, thus cutting computational costs. In the adaptive fuzzy clustering step, it assigns each point to a cluster based on the point's closeness to others in terms of both spatial distance in Euclidean space and geometric distance in a descriptor space, then computes probabilities that a point belongs to all clusters. Note that an adjacency graph of clusters, where spatially closed regions are connected by an edge, is constructed to speed up the clustering. For each of the clusters obtained, the *uniqueness* and *distribution* are computed. Finally, each point is assigned a saliency value which is a linear combination of saliency from all clusters, weighted by the probability that the point belongs to a cluster. Without topological information of shapes, the method achieves results that are comparable to state-of-the-art. However, the expensive clustering strategy limits its uses to point clouds representing a real larger indoor or outdoor scene.

3.3. Global scale based (GS) methods

The idea behind the learning based methods in [57,58] is that salient regions of an object are those that are consistent within a class while distinguishing from other classes. In the training stage, n local descriptors are randomly chosen across each shape. Thus, given a local descriptor Γ of a mesh, its retrieval performance is computed and normalized between 0 and 1, evaluating the quality of its retrieval list via the Discounted Cumulative Gain. This method is specifically designed for content-based shape matching and retrieval. It depends on the database under consideration and how the training pool is partitioned into object classes. In addition,

the scale (size of the region) covered by each spherical shape descriptor has an impact on the results.

Tao et al. [14] presented a method for detecting mesh saliency, in which a semi-supervised mode is adopted. First, via Zernike coefficients a descriptor is calculated for every patch determined by oversegmenting a given mesh. Second, through computing the local distinctness by a center-surround operator, patches with small or high local distinctness are named as background or foreground patches, respectively. Third, the saliency of patches is estimated based on their relevances to some of the most unsalient background patches via manifold ranking. The method outperforms many others because of its high robustness against noises, while suffering from the involved ray-tracing conflict test for computing Zernike coefficients (See Table 1 in [14]). Another limitation mentioned by the authors is that it does not incorporate any high-level priors.

Placing a mesh at the center of a closed scene, Feixas et al. [24] proposed an information-theoretic framework, in which a new definition of mesh saliency was given based on Polygonal Mutual Information (PMI), representing the degree of correlation between a polygon and the set of viewpoints. Hence, the polygonal saliency of a polygon $\mathbf{t}_i$ is specified as

$$\Gamma_s(\mathbf{t}_i) = \frac{1}{n}\sum_{j=1}^{n} d(\mathbf{t}_i, \mathbf{t}_j) \geq 0,$$

where $\mathbf{t}_j$ is a neighbor polygon of $\mathbf{t}_i$, n is the number of neighbor polygons of $\mathbf{t}_i$, and $d(\cdot,\cdot)$ is the dissimilarity measure in terms of Jensen–Shannon divergence. In short, a polygon at the center of a smooth region will have probably low saliency since the polygons of this region will present small visibility differences with respect to the set of viewpoints. This method is specifically suitable for viewpoint selection as the above definition is view-based, however the Raycasting in computing PMI degrades the performance.

An *intrinsic geometric scale space* (IGSS) was constructed by Zou et al. [59] through the surface Ricci flow that conformally deforms the Riemannian metric on a surface according to the induced Gaussian curvature, such that the curvature evolves in a heat diffusion fashion [60,61] that provides a natural multi-scale characterization of the neighborhoods of a given point. In practice, the IGSS is represented as a time-varying Gaussian function $g(\mathrm{V}, t)$ defined on the vertex set V of a triangular mesh $\mathcal{M}$, with $g(\mathbf{v}_i, 0)$ being the initial discrete curvature at $\mathbf{v}_i \in \mathrm{V}$ and $t \in [0, \infty)$ the time or scale parameter. Once a feature point $\mathbf{v}$ is detected as the local extremum in the IGSS, it is natural to define the magnitude of the scale-normalized Laplacian $(t\Delta g_{v,t})$ at $\mathbf{v}$ as the measure of feature strength, i.e., the saliency of $\mathbf{v}$ within its scale-dependent neighborhood. Because the IGSS purely relies on intrinsic properties of shapes, this method is generally more robust to noises and insensitive to resolution changes than the ones based on the mean curvature flow.

Sun et al. [62] presented a point signature called *Heat Kernel Signature* (HKS) by considering heat kernel's restriction to its temporal domain. Given a mesh $\mathcal{M}$ the heat kernel, containing a large amount of redundant information of the mesh, is defined as a function $g_t(\mathbf{v}, \mathbf{u}) : \mathbb{R}^+ \times \mathcal{M} \times \mathcal{M} \to \mathbb{R}$. Thus, the single point signature $\{g_t(\mathbf{v}, \cdot)_{t>0}\}$ is used for defining HKS over the temporal domain as

$$HKS(\mathbf{v}) : \mathbb{R}^+ \to \mathbb{R}, \quad HKS(\mathbf{v}, t) = g_t(\mathbf{v}, \mathbf{v}).$$

The computation of HKS involves the application of the LBO, since the response of the LBO is a positive semi-definite matrix whose size is the number of vertices in the model. This matrix is then decomposed into its eigenvalues and eigenvectors, denoted as λ_i and $\mathbf{w}_i$, respectively. Therefore, we have $g_t(\mathbf{v}, \mathbf{v}) = \sum_{i=0}^{n} (e^{-\lambda_i t}\langle \mathbf{w}_i(\mathbf{v}), \mathbf{w}_i(\mathbf{v})\rangle)$, where n is the number of eigenvectors. Via configuring t and n with specific values, those salient points of a 3D model, especially tips of extremities and protrusions, are assumed to correspond to the local maxima of HKS. In contrast to CS3 [47], HKS is mainly degraded by its less control to performing automatic scale selection.

Hu et al. [64] and Ruggeri et al. [65] showed that the task of extracting salient geometric feature points can be effectively resolved in the frequency space of the LBO, where low frequencies of the Laplace–Beltrami spectrum are related to information about the global structure and features of the input surface, while high frequencies reflect fine changes in its shape. In short, the authors turned to find the critical points of the eigenfunctions that are defined over an input mesh $\mathcal{M}$ as the solutions of the following eigenvalue problem

$$f : \mathcal{M} \to \mathbb{R}, \quad s.t., \Delta f = \lambda f, \quad \lambda \in \mathbb{R}.$$

This method is theoretically plausible [66], and it generates salient points that are intrinsically defined by the input shape and invariant to isometric deformations. Also, the technical core of this method motivated the rewarding work carried out by Song et al. [67,7].

The method called *Heat Walk* was introduced by Benjamin et al. [68] based on heat flow process. Owing to the fact that the heat kernel controls the geometry-dependent propagation of heat across the shape, the salient regions of 3D surfaces, called *accumulators*, can be considered as places where the heat is dissipated very slow compared to the rate at which they receive it. By contrast, the unsalient regions, called *dissipators*, tend to dissipate heat faster than they receive it. This approach is helpful in capturing regional saliencies, and the experimental results showed that it has many nice properties, for example, the resistance to random noise, topological noise, and short circuits. However, evaluating the importance between the extracted accumulator regions was not given.

Diffusion process was also used by Leifman et al. [69], who introduced an algorithm according to observations that people are drawn to *differences*, *extremities*, and *grouping close items*. In terms of the first consideration, a 2D histogram of spin images [70] is built for each vertex first, and then a Gaussian pyramid is used to discretize the continuous diffusion process to measure the dissimilarity between histograms. Finally, a single scale based computation is employed to evaluate the distinctness of vertices, followed by a multiple scale computation to decrease the importance of 3D textures. In the second consideration, intuitively, an extreme vertex is a vertex that resides on the "tips" of the object. In the last step, 20% of vertices with the highest distinctness values are chosen as focus points, of which the maxima of the distinctness and the extremity are taken into account the patch association. The assumption that objects are limb-like is a limitation of this approach. Hereafter, when it drives best-view selection, the evaluator prefers a "side" view rather than a "natural" view.

A hybrid approach which combines the saliency detection and enhancement was presented by Zhao et al. in [71,72]. The algorithm contains three steps. Firstly, a model is voxelized by the bounding box which can be divided into several sub boxes. The scale of the voxelization is specified by the user, thus a surface patch is the one that is located within a voxel. Secondly, a dissimilarity measure between a pair of patches from two randomly chosen sub-voxels is defined by estimating the shape index image of the model as $\frac{2}{\pi}\arctan\frac{k_2+k_1}{k_2-k_1}$, where k_1 and k_2 are the principal curvatures. Lastly, an overall dissimilarity map is iteratively updated, according to that if a patch is significantly different from patch $\mathcal{P}$, then all other patches similar to the patch will be highly likely different from $\mathcal{P}$ as well. The advantage of this method is that by enhancing mesh saliency it becomes easier to locate the meaningful region completely, while it suffers from the voxelization and intensive computation if all patches are compared.

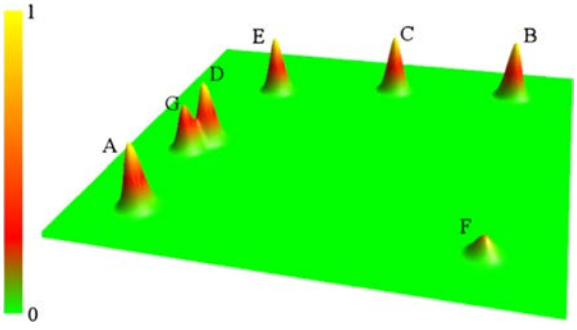

Fig. 5. An example of *global rarity* motivating the notion of topological saliency in [63]: though peak F remains a lone peak for a significantly larger part of the domain than the other peaks (A–E, G), it is not considered to be perceptually important by existing means, and such scenarios are common when studying medical data obtained using diffuse optimal tomography. To tackle the challenge, topological saliency considers the presence or absence of other topological features within the neighborhood of critical points while measuring their importance.

Global rarity is persuasive for visual attentions, as psychophysical studies [73] have shown that the eye gaze is more interested in less frequent features while suppressing repeated features of 3D surfaces. Unfortunately, it was overlooked by previous perception-independent methods. An intuitive example is shown in Fig. 5, which motivated the work in [63]. For a critical point $\mathbf{p}_i$ on $\mathcal{M}$, the metric of topological saliency is defined as follows

$$\Gamma_\sigma(\mathbf{p}_i) = \frac{w_i^i P(\mathbf{p}_i)}{\sum_{\mathbf{p}_j \in \mathrm{N}(\mathbf{p}_i,\sigma)} w_j^i P(\mathbf{p}_j)}, \tag{1}$$

where σ governs neighborhood size, $P(\cdot)$ denotes the topological persistence, w_j^i is a weighting function encoding geometric information for feature point $\mathbf{p}_j$ with respect to $\mathbf{p}_i$. Once topological saliency plot is done as σ increases from 0 to the diameter of input domain $\mathcal{M}$, the order of saliency of features can be decided. This work firstly combined geometric and topological information for mesh saliency detection.

The rarity determined by topological saliency above is sensitive to topology of shapes, instead Wu et al. [74] suggested to measure it using geometric information of Zernike moments. Similar to the model of Cheng et al. [13], the rarity of a vertex can be evaluated by its contrast to all other vertices. As for a vertex, the more similar vertices and the closer to other vertices in the feature space, the weaker its global saliency becomes. Therefore, for a given vertex $\mathbf{v}$, the global rarity is defined as $\sum_{\mathbf{u} \in \mathrm{V}(\mathcal{M})} d(\mathbf{u},\mathbf{v})$, where $d(\cdot,\cdot)$ is the dissimilarity of Zernike coefficients between two vertices. Computing Zernike moments by Raycasting results in the major shortcoming of this method.

Temporal variations of the geometry were not well integrated into mesh saliency detection until the*motion-saliency model* introduced in [75], which, along with surface curvature, visibility, and rarity, are important ingredients of the HVS. The model contains two sub-processes. In the first process, the differential velocities of vertices and their connectivity information are used to group vertices of similar motions into the same cluster. Thus, the generated clusters should contain vertices that have very close motions throughout the animation. In the second process, for each cluster a representative vertex, called a *cluster head*, is chosen. Hence, the motion state of a cluster head is extracted via its velocity relative to the whole object. After that, the motions of the cluster heads are analyzed throughout the animation to calculate the saliency of their clusters. Though this work firstly studied temporal variations in mesh saliency, we were aware that its use is limited because saliencies captured are almost the parts where motion changes happen in scenes.

Rewardingly, in [67] it was demonstrated that the irregularity of the spectrum of a 3D mesh is highly related to the saliency of the mesh. The frequencies of a triangular mesh $\mathcal{M}$ are specified as the eigenvalues of the *Laplacian matrix*, with which the Laplacian spectrum is $\mathbf{H} = \{\lambda_i, 1 \le i \le |\mathrm{V}(\mathcal{M})|, 0 \le \lambda_i\}$, where λ_i are the eigenvalues, and i is called *frequency index*. First, the spectral irregularity is computed. After that, inspired by HKS [62] a curvature-weighted shape descriptor called *irregularity kernel signature* (IKS) is proposed to deliver and reorganize the saliency information into the spatial domain. The descriptor represents a saliency map of the mesh, with which the *globally salient points* and the *stable locally salient points* are detected.

Instead of using the spectrum of LBO directly, the log-Laplacian spectrum is used in [7]. That is, $\mathbf{A} = \log(|\mathbf{H}|)$, where log denotes the natural logarithm acting as a spectral re-distributor. In order to compute spectral deviations from the norm, an approach with no-learning is adopted. Therefore, spectral deviation can now be computed as the spectral irregularity $\mathbf{R} = |\mathbf{A} - \mathbf{J}_n * \mathbf{A}|$, where $\mathbf{J}_n = \frac{1}{n}[1\ 1\ \cdots\ 1]$ ($n=9$, for example). To bring the spectral representation back to the spatial domain, a composition $\mathbf{S} = \mathbf{W}\mathbf{D}\mathbf{W}^T \cdot \mathbf{B}$ is performed, where $\mathbf{D} = Diag\{\exp(\mathbf{R})\}$, $\mathbf{W}$ is the eigenvector matrix, and $\mathbf{B}$ is the distance-weighted adjacency matrix. Therefore, the saliency map $\Gamma_s(\mathbf{v}_i)$ for vertex $\mathbf{v}_i$ is derived by summing $\mathbf{S}$ along the ith row. This method is theoretically sound in terms of Fourier transformation and achieves good results, unfortunately it suffers from inability for perceiving high-level saliency (see Fig. 6).

4. Global contrast based (*GC*) methods

In contrast to local contrast based methods, global contrast based methods aim at detecting the most representative salient components of a 3D mesh, which often appear large-scale and robust.

CRF-based mesh saliency detection was presented in [77], which consists of two steps. In the first step, a multi-scale representation of a given mesh is constructed by proposing a rank-based Gaussian filtering. The filtering not only utilizes geodesic to reduce neighborhood ambiguities produced by Euclidean measure but also replaces Gaussian-weighted position with surface curvature for higher robustness against scanning noise. Therefore, at each scale, a 3D vector representing the displacement of point $\mathbf{p}$ from its original position on a mesh after the filtering is obtained, which is projected onto $\mathbf{n}(\mathbf{p})$ to reduce a scalar quantity. In the second step, unlike [15] simply sums over the multi-scale information, the Conditional Random Field (CRF) framework is incorporated, i.e., a label assignment $\mathbf{s} = \{s_\mathbf{p}, \forall \mathbf{p} \in \mathcal{M}\}$ is defined, where the label set comprises the scale indices. Solving this CRF in accordance with the *maximum a posteriori probability* criterion via belief propagation (BP) assigns a label to each point on the mesh. Therefore, points assigned the same principal scale comprise the non-salient regions, while the left points comprise the salient regions. Since all points are identified as either salient or non-salient, therefore compared to most existing mesh saliency detection methods the CRF-based approach is more effective in capturing stable saliencies, while a limitation is that the accuracy of final results is somewhat sacrificed.

Sipiran and Bustos [32] introduced the term *key-component* as a region on a 3D mesh where there are a lot of discriminative local features. This method consists of three steps. First, the effective and efficient 3D Harris operator [43,45] is employed for detecting keypoints. Second, let $S = \{\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_n\}$ be the set of keypoints previously detected, partitions $S_i \subset S$ are found by fulfilling four constraints in order to group them according to their closeness. The partitioning is done by solving a Multidimensional Scaling problem that a set $P \subset \mathbb{R}^2$, in which Euclidean distances between

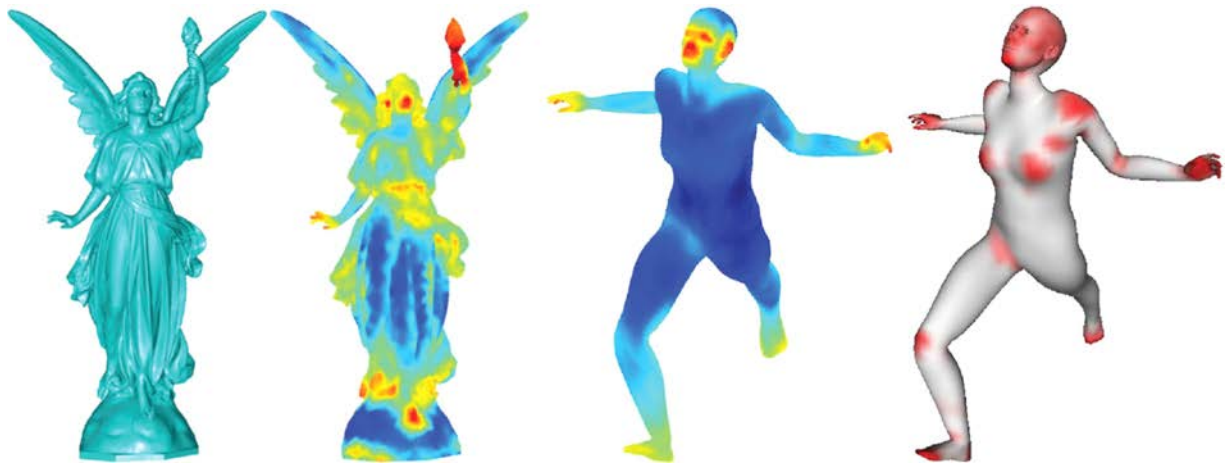

Fig. 6. An illustration of mesh saliency detection method based on spectral processing in [7]. The method is capable of dealing with meshes containing a large number of vertices, e.g., Lucy model is of 14*M* vertices. However, it fails to detect the breasts as salient in Woman model (ground truth of the rightmost model is labeled and colored in red). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

Fig. 7. A comparison between local salient elements (top row) and global salient structures (bottom row) [29].

elements in P approximately preserve the geodesic distances between elements in S, is computed. Finally, each cluster will generate a component comprising the region of the mesh where the keypoints are located. The advantage of the method is that the number of key-components should be much less than the number of conventional salient seeds, so using the results in subsequent tasks would be more efficient. While a large portion of the final results are repeatable components, which does not comply with the principal of rarity of the HVS.

Very recently, the *low-rank and sparse analysis* was utilized by Wang et al. [29] for detecting mesh saliency, based on a strong belief [7] that effectively and accurately recognizing the global importance of different frequency information will help quantify the saliency of 3D models. To do so, a powerful structure-aware descriptor was introduced, with which the original model $\mathcal{M}$ and a constructed low frequency model $\mathcal{M}^l$ are used to represent the high level and low level information, respectively, forming a multi-level expression of a 3D model. At each level, a multi-scale shape context, describing the shape characteristics, is constructed by considering the contours on a vertex-wise bi-harmonic distance field. Given a vertex $\mathbf{v}_i$, its multi-scale shape context is defined as $f_i = [m, p_1, pd_1, \ldots, p_k, pd_k]$, where m is the local metric (e.g., curvature), p_i is the normalized perimeter of the ith contour, and pd_i is the corresponding probability distribution histogram of the distances $\{d_i\}$. Thus, the vertex-wise shape descriptor is expressed as $F_i = [f_i^h, f_i^l]$ in the shape contexts of $\mathcal{M}$ and $\mathcal{M}^l$. The descriptor of each vertex is assembled to span a feature space, which can be reorganized in the form of matrix $\mathbf{F} = [F_i, \ldots, F_n]$. Therefore, we further divide the feature space $\mathbf{F}$ into two subspaces of $\mathbf{F}_l$ and $\mathbf{F}_g$, where $\mathbf{F}_l$ is the local feature subspace and $\mathbf{F}_g$ is the global structure subspace. It is proven that the first subspace is low-rank and the second is sparse, which, in turn, characterize local feature saliencies and global structure saliencies, respectively (see Fig. 7).

5. Mesh saliency detection applications

Table 2 summarizes state-of-the-art work of mesh saliency detection that have been elaborated in the previous two sections, also applications where the work itself has tested are enumerated. Though the range of the applications is wide, there does not exist one method that is superior to other measures in all aspects, as some target a specific application. Besides the applications listed by the table, others include mesh denoising [78], mesh remeshing [79], icon generation [58], abstract caricature [80], facial expression analysis [81], data visualization [82,83], and 3D printing [84]. In this section, a few typical works are briefed, meanwhile we refer the readers to the references in the table.

5.1. Mesh simplification

In general, small high-curvature details in the middle of largely flat regions are likely to be decimated by most mesh simplification methods, while it is desirable that these small surfaces can be perceived as they are more visually prominent. For instance, in a bust model mesh, the back will be simplified more than the front, since there are a lot of salient regions in the face such as the eyes, the nose and the mouth.

In order to meet the requirement, Castelló et al. [76] included viewpoint saliency in the viewpoint-driven simplification scheme by using a criterion to weight the importance of a viewpoint. By quantifying the variation in the shape viewed from c, the scheme defines a new *simplification error deviation* for an edge collapse as $\sum_{c\in\mathcal{C}}\Gamma_s(c)|f(c,\mathcal{M})-f(c,\mathcal{M}')|$, where $\mathcal{C}$ is the global camera space, $\Gamma_s(\cdot)$ is the viewpoint saliency, and $f(c,\cdot)$ is calculated from the projected areas of each polygon. The effectiveness of the scheme is demonstrated in Fig. 8.

5.2. Mesh segmentation

Mesh segmentation is a classical problem in computer graphics for cutting up a complex object into simpler sub-objects. It can be interpreted either in a purely geometric sense or in a more semantics-oriented manner. The first case complies with some properties of meshes (e.g., curvature or distance to a fitting plane), while the second case aims at identifying parts that correspond to high-level features.

The mesh decomposition method (called K-means) [85] was applied to the metamorphosis of polyhedral surfaces, which contains four steps: preprocessing, electing the initial representatives of the patches, determining the decomposition, and reelecting the representatives. In the fourth step, the goal is to minimize the following energy function

$$\min\sum_{\mathbf{p}}\sum_{\mathbf{t}\in\mathcal{P}(\mathbf{p})}Dist(\mathbf{p},\mathbf{t}),$$

where $Dist(\mathbf{p},\mathbf{t})$ is the shortest distance from a patch representative $\mathbf{p}$ to a triangular face $\mathbf{t}$ belonging to the path $\mathcal{P}$ represents. To improve decompositions where the human intuition is pertinently matched, Wang et al. [29] revised the above equation by adding a new term in which the saliency similarity between $\mathbf{p}$ and $\mathbf{t}$ is considered. From Fig. 9, it is seen that the boundaries of different segmented parts are closely related to the structures, thus the final segmentation results are improved effectively.

5.3. Mesh resizing

The resizing of 3D meshes is a popular and useful operation that is often regarded as a special kind of shape deformation. By *resizing* we mean scaling or stretching the object along several orthogonal directions or dimensions to fit a new prescribed size [86]. One of the key issues in mesh resizing is that some visually salient features of the underlying model should be preserved as much as possible.

Jia et al. [54] suggested resizing should be distributed non-homogeneously throughout the mesh, protecting salient regions, while stretching others excessively. To do so, the following energy function is to be minimized

$$\min\left[\sum_{\mathbf{e}\in E(\mathcal{M})}\left|\frac{(\mathbf{v}'_{\mathbf{e}_0}-\mathbf{v}'_{\mathbf{e}_1})-w(\mathbf{e})(\mathbf{v}_{\mathbf{e}_0}-\mathbf{v}_{\mathbf{e}_1})}{(\mathbf{v}_{\mathbf{e}_0}-\mathbf{v}_{\mathbf{e}_1})+\alpha}\right|^2+g\right].$$

In the first term, edges of a mesh are stretched by updating positions of vertices and keeping the connectivity unchanged. In order to preserve those edges whose saliency is high, the weighting is a linear combination of uniform and target scale β, that is $w(\mathbf{e})=\Gamma_s(\mathbf{e})+(1-\Gamma_s(\mathbf{e}))\beta$. The second term, i.e.,

Fig. 8. An illustration of saliency-driven mesh simplification with Buddha model ($|\mathrm{T}(\mathcal{M})|=60,866$) [76]. The left and right show the original model and saliency map, respectively. The middle top and middle bottom are close-ups of decimation result ($|\mathrm{T}(\mathcal{M})|=8,999$) with and without considering mesh saliency, respectively. It is seen that the improvements achieved make the most salient regions such as the mouth, eyes, wearing, and hair are preserved better.

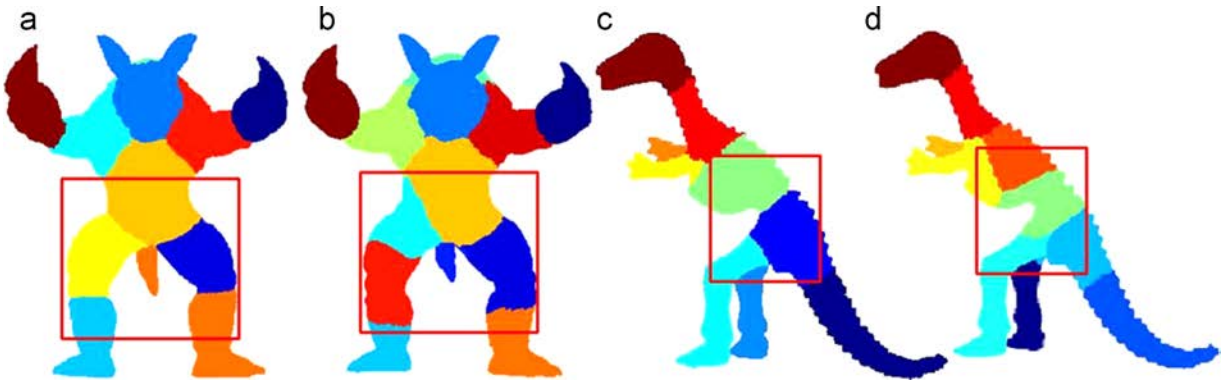

Fig. 9. An illustration of saliency-driven mesh segmentation with Armadillo and Dinosaur models. [76]. (a) and (c) are saliency driven K-means mesh decompositions [29]. (b) and (d) are K-means mesh segmentations [85].

Fig. 10. An illustration of saliency-guided mesh resizing with Lion model. Left: the original model; middle: the saliency guided mesh resizing result [54]; right: the non-homogeneous mesh resizing result [86].

$g = \sum_{\mathbf{v} \in V(\mathcal{M})} (\mathbf{L}\mathbf{v}' - g(\mathbf{v}))^2$, is added to avoid a situation where the face area may change dramatically during the optimization. Via the comparison in Fig. 10, it is seen that saliency-guided mesh resizing achieves more visually appealing results.

5.4. Normal enhancement

The normal enhancement operation is a technique in guiding the shading scheme for 3D shape depiction. To do so, a novel approach was presented by Miao et al. [87], where, by dynamically perturbing the surface normals, the illumination and shading are adjusted to push the influence of visual attention into the graphics rendering pipeline.

This approach depends on the classical Phong local lighting model in which the lighting of vertex $\mathbf{v}$ is determined by ambient lighting, diffuse lighting, and specular lighting. Additionally, it employs the orthogonal coordinate system whose coordinate axes $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ are considered as the halfway unit normal $\mathbf{h}$, $(\mathbf{h} \times \mathbf{l})/(\|\mathbf{h} \times \mathbf{l}\|)$, and $\mathbf{e}_1 \times \mathbf{e}_2$, respectively, where $\mathbf{l}$ is the unit light direction. Thus, given a vertex $\mathbf{v}$ on a mesh $\mathcal{M}$ the perturbation of the normal vector can be represented in the above coordinate system as

$$\nabla \mathbf{n}(\mathbf{v}) = \delta_1 \mathbf{e}_1 + \delta_2 \mathbf{e}_2 + \delta_3 \mathbf{e}_3.$$

First, with the constraint that $\nabla \mathbf{n}(\mathbf{v}) \cdot \mathbf{h}(\mathbf{v}) = 0$, it holds $\delta_1 = 0$. Second, to solve the equation above, a saliency map Γ_s of $\mathcal{M}$ is computed, normalized, and linearly scaled onto the bound limitation of $\delta\mathbf{i} = \nabla \mathbf{i}/\mathbf{i}_s$, where $\mathbf{i}$ is the lighting map estimated by Phong model, and $\mathbf{i}_s$ is the point light source intensity. Finally, the obtained saliency-guided luminance variance $\delta\mathbf{i}$ is applied to calculating δ_3 that in turn assists in computing δ_2. From Fig. 11, it is seen that the salient surface features are effectively enhanced by altering the shade of these regions. For instance, the brim of Bunny's ear and the muscle of Bunny's leg are also improved clearly and attract viewers' attention.

5.5. Volume rendering

Volume rendering is a technique to compute a 2D projection of a three-dimensional data set that allows for a layered representation of three dimensional intricate structures, where a common problem is to find a suitable transfer function that provides a good characterization of those different, even partly overlapping structures in a 3D volume [82].

In order to draw visual attention to user-specified regions in a direct volume rendering environment, Kim and Varshney [83]

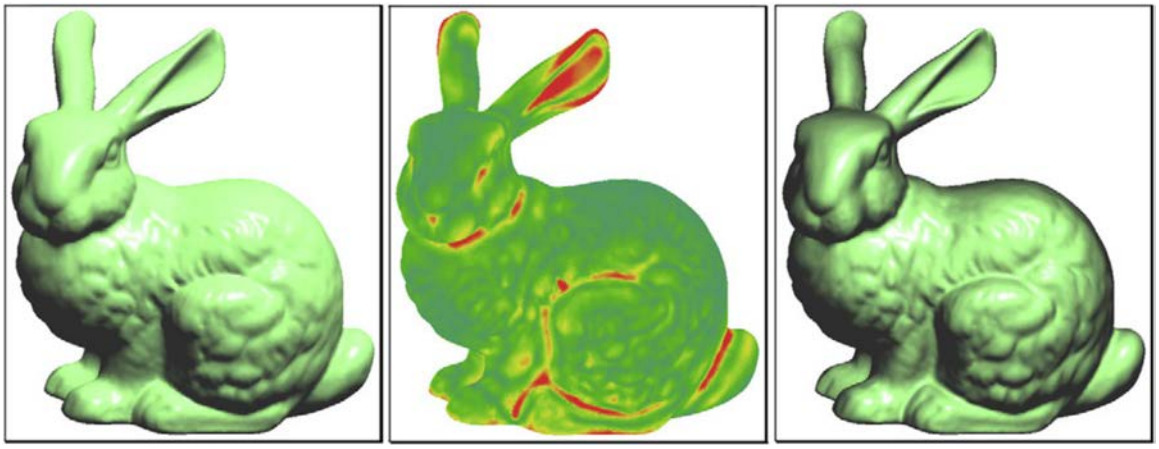

Fig. 11. An illustration of saliency-guided normal enhancement technique for 3D shape depiction [87]. Left: the original Stanford Bunny; middle: saliency map; right: the expressive rendering result.

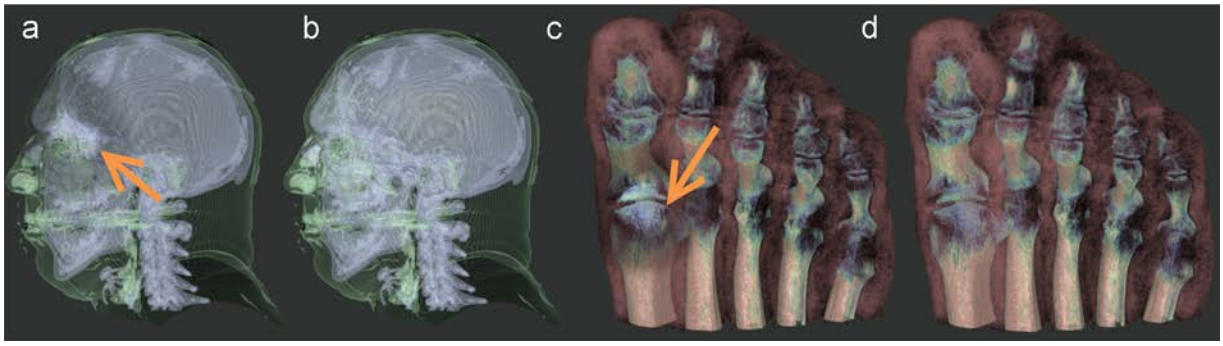

Fig. 12. An illustration of saliency-guided volume rendering direct volume rendering (DVR) technique [83]. (b) and (d) show the traditional volume visualization. (a) and (c) show the results of applying the saliency-guided enhancement operator to specific regions indicated by the arrows.

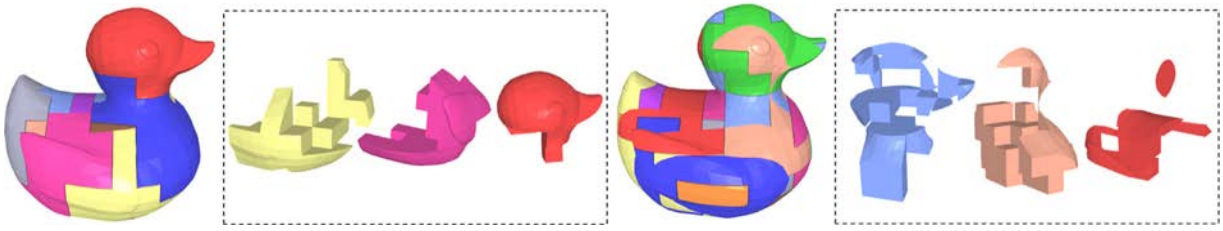

Fig. 13. An illustration of 3D printing with (the left Duck) and without (the right Duck) saliency guidance [88].

presented a new visualization enhancement operator that is inspired by the center-surround mechanism of visual saliency. The starting point for this approach is the generation of a saliency field Γ_s that defines a value of saliency for every voxel. This assignment could be based upon user specification (manual painting), eye-tracking data, or feature computation. Then, the known saliency map assists in computing a virtual emphasis field $\mathbf{B}$, i.e., $\mathbf{B}=\mathbf{A}^{-1}\Gamma_s$. The matrix $\mathbf{A}$ is with respect to $a_{ij}=\alpha G_w((\mathbf{v}_i-\mathbf{v}_j),\sigma)-\beta G_w(\mathbf{v}_i-\mathbf{v}_j,2\sigma)$, where $\alpha=3/4$, $\beta=1/4$, $G(\cdot,\cdot)$ is the Gaussian kernel, σ is the scale, and $\mathbf{v}_i$ and $\mathbf{v}_j$ are voxels in a neighborhood. Once we have computed the emphasis field it is used to modulate the various visualization parameters. Fig. 12 displays the results where the saliency-guide emphasis field is exploited to modulate its brightness of every voxels. It can facilitate the users' understanding by guiding their attention to regions and objects selected by a domain expert. The work showed that the saliency enhancement operator is more effective at eliciting viewer attention than the traditional Gaussian enhancement operator.

5.6. 3D printing

To address an intrinsic limitation of 3D printing that a single object cannot be fabricated if it is larger than the working volume of a 3D printer, Song et al. [84] proposed to connect the printed 3D parts by 3D interlocking, instead of using connectors, glue, or skew. When building the saliency connection graph for boundary voxels to later avoid putting cutting seams (between interlocking 3D parts) on salient object features, they first applied the method in [15] to compute the saliency value at every mesh vertex and then estimated the saliency value of each boundary voxel by averaging the saliency values of all mesh vertices it contains. Therefore, a strong saliency connection was built only when the two related voxels had large saliency values (see Fig. 13).

Another saliency-guided 3D printing work was carried out by Wang et al. [88], who presented an adaptive slicing method for reducing the manufacturing time of 3D printers while preserving the visual quality of printing results, with which the printing time

is saved by 30–40%. They applied the method in [7] to compute the saliency of an original input model first, and then the visual quality of an ith layer l_i is measured by a metric as follows

$$\Gamma_s(l_i) = \Gamma_s(\mathbf{v} \in l_i) \cdot h_i \cdot \cos\,\theta_i,$$

where Γ_s is the saliency map, h_i is the thickness, and $\theta_i = \min_{\mathbf{v} \in L_i} \theta(\mathbf{v})$ in which $\theta(\mathbf{v})$ is the angle between the norm of a surface point $\mathbf{v}$ and the Z-direction.

5.7. Scan integration

Automatic 3D model reconstruction from multiple registered range images plays a key role in robot vision, reverse engineering, and 3D printing. Mainly, it contains four steps: registration, segmentation, integration, and triangulation. The third step of *scan integration* refers to how to integrate the registered range images captured from different viewpoints into a surface representation of the same object. In general, it is a rather challenging task, because the registration errors can be accumulated from one view to another, leading the overlapping areas, segmented in the second step, to deviate away from each other.

Recently, Song et al. [7] showed how mesh saliency is able to significantly improve scan integration (see Fig. 14). The core idea is that given two registered 3D scans, $\mathcal{M}_1$ and $\mathcal{M}_2$, they are initialized and partitioned into salient and nonsalient regions with respect to the pre-computed saliency maps of the scans. In this way, it redistributes registration errors within each scan in such a way as to ensure that the salient regions suffer less from registration errors. The integration scheme for nonsalient regions is similar to the one presented in [89]. Nonetheless, for the salient regions where scan integration does not rely on point normals, a different scheme is put in place by employing the iterative closest point (ICP) algorithm to reposition points in salient regions. With applying two different schemes to the saliency-driven segments, the integrated point cloud $S_{integrated} = S_{non-salient} \cup S_{salient}$ is achieved. If there are further scans, it next combines $S_{integrated}$ integrated with the next input scan $\mathcal{M}_3$. After all input scan images have been integrated through this procedure, a single integrated point set is obtained, and the scan integration procedure is done.

6. Evaluation of mesh saliency detection methods

Table 1 summarizes feasible methods for evaluating mesh saliency, to which some comments are given. In general, among them the application-guided assessment is superior, and the reasons are four-fold. First, unlike the means of qualitative evaluation, it does not require (pseudo-)ground truth. Second, it does compare various methods directly and quantitatively, preventing an assessment from subjective judgement. Third, it is self-adapted, as different applications have different scales and metrics usually. Lastly, it is consistent with the acknowledgement that the true effectiveness of a saliency detection method depends on the applications [13].

6.1. Qualitative evaluation

Unfortunately, there is no real-ground truth for qualitative evaluation in the context of mesh saliency. The reason is that a saliency map is not as objective as surface analysis metrics such as area and curvature (see more discussions in Section 7.1). As a result, the majority of mesh saliency approaches, such as [58,69], opted to compare their results with the previous. Such an option, however, is problematic when the results to be compared are not accessible, as the instances the prior work give are always limited. Even if the references are available, it is not a straightforward task of comparing saliencies by the readers themselves, as different people might have different ideas of what it means to be important (e.g., functional, structural, social, and visual) [90]. Carrying out user studies can attenuate the issues aforementioned [74], however it is difficult to prevent an assessment from subjective judgement; also, the data output is less reusable.

A better way to go is taking advantage of *pseudo-ground truth*, which essentially consists of eye-fixation data obtained by physical eye-tracking or laborious annotations. For instance, an open benchmark http://www.itl.nist.gov/iad/vug/sharp/benchmark/3DInterestPoint of human-generate ground truth was presented via a web-based subjective experiment [8]. There, the user is shown the 3D models, one at a time. Given the UI, each person is free to rotate an object in 3D and then asked to mark the "interest points", which are supposed to correspond to the points that appear at every object in the same category. Considering that the human judgment is subjective by nature, eventually, some consensuses among the users are reached in order to merge all the marked points into a final set of pseudo-ground truth. Another open benchmark http://points.cs.princeton.edu, based on *Schelling's focal points*, was introduced by Chen et al. [90], who, rather than asking people "click on important points" [34,35,25,8], requested participants "select points they think will be selected by others".

Now, both benchmarks are reusable for qualitative evaluation (see Fig. 15) and quantitative evaluation (see Section 6.2). Nevertheless, one drawback is that the scheme of the datasets is designed for evaluating local contrast based mesh saliency. Hence, in the future it will be rewarding to provide other pseudo-ground truth for assessing global contrast based mesh saliency.

6.2. Quantitative evaluation

The aforesaid means of qualitative evaluation are sort of empirical evaluation. Meanwhile, there are criteria, which treat the saliency detection task as a binary decision problem, for evaluating new algorithms of saliency quantitatively. The term of *decision* is referred to as the gold standard or ground truth judgment of relevance. In practice, with pseudo-ground truth such as Schelling points [90] a value normalized needs to be chosen as a

Fig. 14. An illustration of how the task of scan integration is significantly improved by the saliency-guided method in [7] (the right most). The first and the second are respectively one the input scans and the computed saliency map. The third to the seventh are results of the state-of-the-art. It is seen that more geometric details are preserved by considering the factor of mesh saliency.

Table 1
A table of feasible mesh saliency evaluation methods, where the second column tells the type of an evaluation method that is either qualitative or quantitative; the third column shows if pseudo-ground truth (p-GT) is required by the method; and the fourth column suggests how costly the method is in general.

Evaluation method	Type	p-GT	Cost	Comments
Saliency vs. saliency	qlty	No	Low	It is impractical when the references to be compared are not accessible
User study	qlty	No	High	It is effective, while the generated data is less reusable
Saliency vs. GT	qlty	Yes	Low	The comparisons are convincing; however, to date ground truth for evaluating global contrast based mesh saliency is not available
PR, ROC, AUC, CC, and F-measures	qnty	Yes	Low	These measures have not yet been applied to known mesh saliency methods, although they are highly cited in the study of visual saliency
FNE, FPE, and WME	qnty	Yes	Low	These measures are easy to compute and commonly used
Application-driven assessment	qnty	No	Medium	It outperforms the others, as the true effectiveness of a saliency detection method depends on the applications

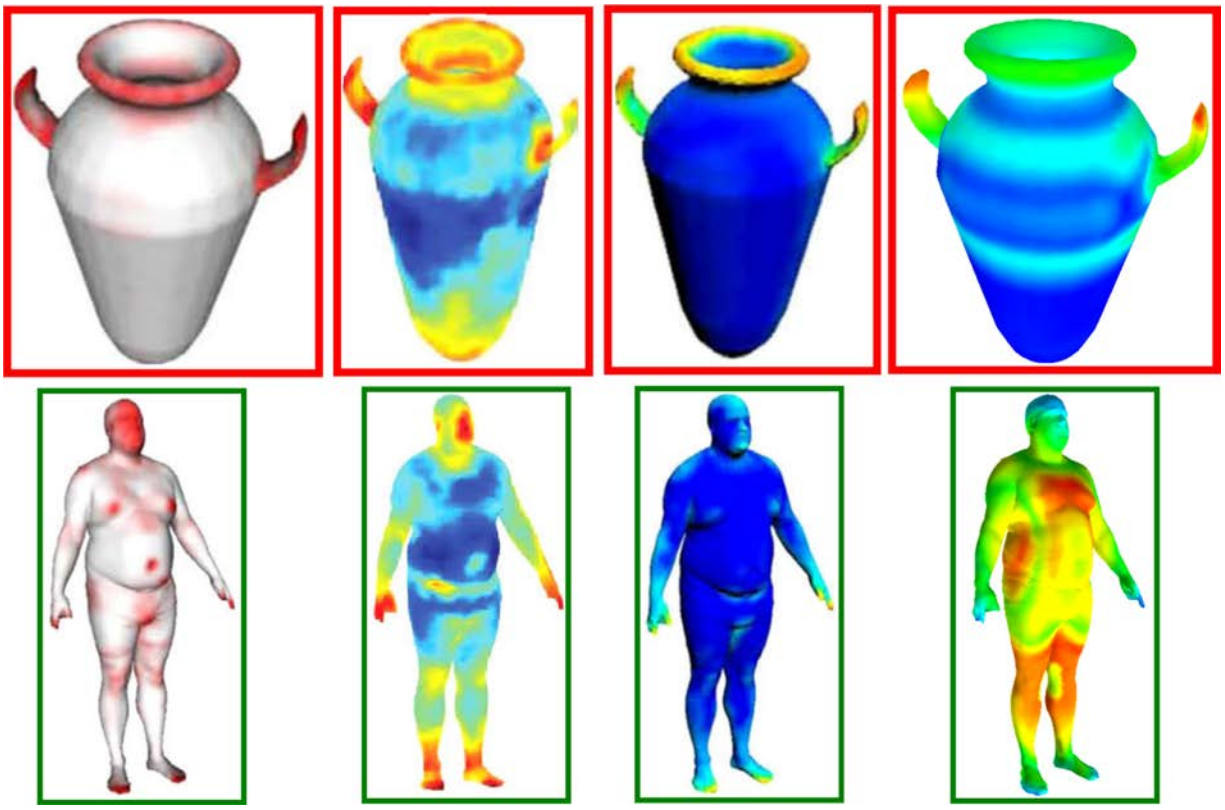

Fig. 15. An illustration of qualitative evaluation of salient regions determined by pseudo-ground truth [90], spectral mesh saliency [7], ranking unsalient patches [14], and DOG of surface curvature [15], from left to right. With the reference of *Schelling points*, it is convincing that the method in [7] achieves poorer results than the one in [14] and the one in [15] with the jar model in red box, while it outperforms the other two with the human model in green box. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

threshold in advance, such that the relevant space is separated from the irrelevant. Thus, the charts of indicators such as *precision–recall* (PR), *receiver operator characteristic* (ROC), area under ROC (a.k.a., AUC), correlation coefficients (CC), and *F-measures* can be drawn to evaluate the effectiveness, i.e., the accuracy of decisions made by different methods. But, in the survey we are aware that these metrics are never used by the known mesh saliency methods, even though they are highly cited in visual saliency study [13,91,92].

The measures of *false negative error* (FNE), *false positive error* (FPE), and *weighted miss error* (WME) are prevalent, by contrast. They were firstly defined in [8] and later used in [56,93]. Let G be the set of ground truth points, and A be the set of points detected by an algorithm against a model $\mathcal{M}$. A point $\mathbf{p} \in G$ is considered to be "correctly detected", if there exists a detected point $\mathbf{q} \in A$ such that $\mathbf{q}$ is close to $\mathbf{p}$ but not closer to any other points in G. Therefore, it can be formulated that $FNE = 1 - N_C/|G|$ and $FPE = (1 - N_C/|A|$, where N_C is the number of correctly detected points, and $|\cdot|$ returns the size of a set. In order to take the prominence of $\mathbf{p}$ into account, WME assumes that $\mathbf{p}$ is marked by n subjects within a parameterized geodesic neighborhood. Therefore, $WME = 1 - \sum_{i=1}^{|G|} n_i \delta_i / \sum_{i=1}^{|G|} n_i$, where if $\mathbf{p}$ is correctly detected by the algorithm δ_i returns 1; otherwise, it returns 0.

In mesh saliency it is more popular that the quantitative evaluation is conducted through the *application-guided assessment* using saliency as a weight map (see Section 5). Taking the application of mesh simplification discussed in Section 5.1 as an example, a better mesh saliency detection method is the one that induces less error between the original mesh and the simplified. To do so, the measure of the *root mean square error* (RMSE) is widely adopted. It is either image-based [94], which renders both models from vertices of a surrounding dodecahedron using flat shading then calculates the squared sum of pixel-wise intensity difference between output "images", or geometry-based, which

refers to the well-known Metro method [95]. The measure called MESH error [96] is common as well, which improves the Metro method by using the Hausdorff distance. Although neither RMSE nor MESH measures how well salient areas are preserved, they are useful tools when comparing the results of two different ways of simplifying meshes based on saliency [7].

Taking the application of scan integration discussed in Section 5.7 as another example, the integration error introduced in [89] measures the average squared Euclidean distance between the integrated points and their closest points in the input range scans.

7. Discussion: Current trends and outlook

A current trend in mesh saliency detection reflects that global geometric (or topological) data [69,74,67,7,14] of graphics are more effective than local cues or their combinations [15,39,42] in achieving perceptually meaningful saliencies. Another trend is that, besides conventional surface curvatures, recently new thoughts about the HVS have been explored and incorporated into mesh saliency detection. The motion of vertices [41], global rarity [74], visibility [69], and temporal variations of geometry [75] are a few examples.

Next, we present four key insights into mesh saliency detection, which are expected to be helpful for researchers in the future study.

(i) First, of course mesh saliency is an intrinsic attribute of objects and can benefit various applications, while there are many open problems: *is mesh saliency pose-aware*? *how to evaluate the saliency of a region being self-occluded*? *do shape contexts change mesh saliency*? *which properties of mesh saliency in the first-stage are preserved in the second-stage, and which are not*? Answering them is rather challenging. Not only is mesh saliency selectively subjective in fact but also there are no hints and comments given by prior work about them. A test case, illustrated in Fig. 16, is carried out by computing a saliency map of the same model but in different poses. In order to evaluate the map, the values are scaled into a range of $[0, 1]$ first, and then a histogram is constructed for the model with 100 uniform bins spanning the range scaled. Consequently, the task of comparing 3D shapes can be converted to computing distances between their corresponding histograms under the ℓ_1 norm. The experiment results indicate that mesh saliency should be pose variant, also it is certain that this simple test is far away a satisfactory reply to those problems.

(ii) Second, as stated by Sipiran and Bustos in [32], the detection of robust components will be the next step in the search of reliable local structures of 3D models. Compared to small-scale salient elements, large-scale salient components are more structure-preserved thus better adaptable to the fidelity of HVS. Moreover, it is reminiscent of the last open problem aforementioned.

(iii) Third, since multimedia data plays a key role in human-centered visual computing field, there will be more attentions being paid to the study of saliency. Visual saliency, which mainly targets 2D images, is a topic that is closely related to mesh saliency. Even if 3D data requires its own processing and analysis methods, basic knowledge about the similarities and differences between other topics and the topic itself would be helpful for making progresses in the future. For instance, visual saliency is recently exploited to guide scene classification in image processing [97], while such an idea has never been attempted in 3D cases.

(iv) Fourth, along with the emergence of low cost 3D acquisition devices such as Microsoft Kinect, the amount of range image data is growing quickly. Such a current trend has reflected the flourishment of research towards the development of 3D object recognition systems [11], in which the concern of self-occlusion in mesh saliency turns out to be an inter-occlusion problem in *scene saliency*. Also, treating mesh saliency maps as surface features (or shape descriptors) offers a new insight into object recognition potentially.

7.1. Salient points vs. interest points vs. critical points

In mesh saliency detection, it is important to be clear about a concern – "*are salient points equivalent to interest points and critical points? if no, what is their correlation?*". Indeed, the three types are highly relevant to the features of 3D models, as all take account of the fact that the human vision's tendency is drawn to differences. Thereby, owing to their efficiency of visual persuasion in classical geometric analysis and processing, these topics have now been well studied and widely used in computer graphics.

However, as interpreted by Fig. 17, a better answer to the concern is "not exactly!". The interest points, being referred to keypoints [9] as well, are locations that are distinctive in their locality, stable, and present at all instances of an object or of its category of objects [8,98]. In general, they are detected by defining functions summarizing the shape content of localities on a 3D model in multiple scales, such that the local extrema of those functions are selected as interest points. Notice that almost (or all) intrinsic geometric and topological information of shapes can be helpful for interest point detection, while only the part that is dependent on the HVS are suitable for mesh saliency detection. The reason is that, besides the deviations due to geometry and topology, the eye gaze of the HVS is more prone to visual uniqueness, unpredictability, rarity, or surprise. To summarize, salient points are required to be perception-dependent, while interest points are not.

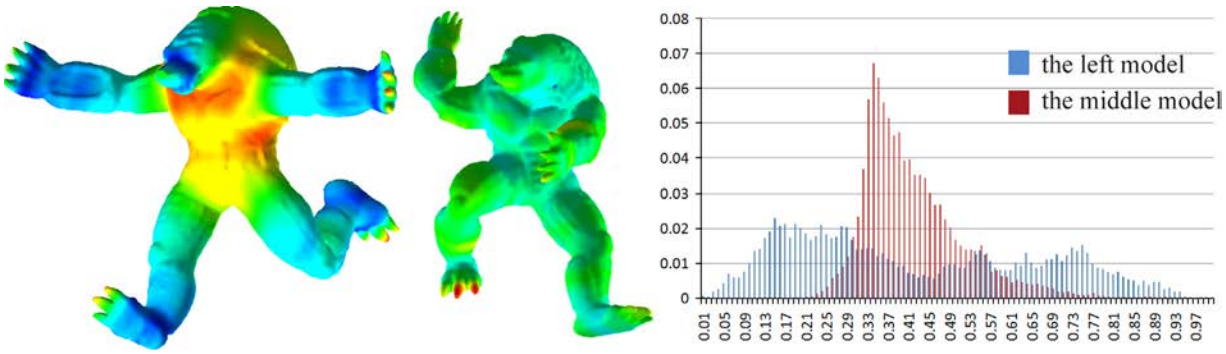

Fig. 16. The left and the middle are the same model but in different poses. The right diagram shows that the distance between histograms of saliency map [15] is 1.194 under the ℓ_1 norm, indicating that mesh saliency is pose variant.

The *critical points*, nevertheless, are related to the topology of manifolds and well defined by Morse theory [99]: given a smooth, real-valued function $f : \mathcal{M} \to \mathbb{R}$ defined on $\mathcal{M}$, the critical points of f are exactly those where the gradient vanishes. In general, a critical point can be further classified as the *ridge*, the *valley*, or the *umbilical*. Ridge points and valley points, whose extremality coefficients degenerate, identify places where a surface bends sharply, thus they are perceptually salient [100]. In [63], critical points were used for complementing mesh saliency detection. However, compared to geometry of 3D shapes, the uses of topology in mesh saliency detection are still relatively less.

7.2. Visual saliency vs. mesh saliency

There has been a substantial body of literature on visual saliency (see [101] for a review). Unlike images and videos where color and temporal coherence are the most important attributes, meshes esteem geometry and topology of shapes as the most important contributors. Thus, the nature of the problem is different, so are the solutions to saliency. Since mesh saliency was derived from visual saliency detection [15], essentially both tasks share a great common interest in defining saliency maps representing the level of saliency for visual attentions.

Taking the global contrast based saliency detection as an example, Shen and Wu [92] suggested that salient object detection is not a pure low-level, bottom-up process, but higher-level knowledge is nontrivial even for task-independent image saliency. So an image is represented as a low-rank matrix plus sparse noises in a certain feature space, where the non-salient regions (or background) can be explained by the low-rank matrix, and the salient regions are indicated by the sparse noises. In the feature space, each vector consists of 53 features that are stacked vertically with conventional local low-level information of RGB color with global high-level guidance of steerable pyramids and Gabor filters. After clustering, by staking all f_i, i.e., the representative feature vector of the ith segment, into a matrix, which is further decomposed into two parts, the low-rank matrix and the sparse matrix. Apparently, this visual saliency detection work, along with other work [102] based on low-rank and sparse analysis, strongly inspired the recent work of Wang et al. [29] for mesh saliency detection.

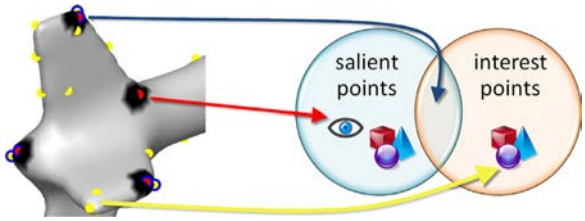

Fig. 17. An illustration of the correlation between *interest points* and *salient points*. In the left image, the red dots stand for points indicated as human-generated ground truth by collecting eye fixation data, and the yellow dots stand for points marked by the interest point detection algorithm presented in [8]. Suppose that the ground truth amounts to salient points, as diagrammed in the right image, the two types of interest points and salient points do have overlaps but are not equal to each other. The lines in blue, red, and yellow represent regions that are true positive, false negative, and false positive, respectively. This illustration verifies the correlation that *interest points* $\approx$ *salient points*. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

Except that, there exist works in which visual saliency can accommodate geometric processing directly. For example, as demonstrated in Fig. 18, given a parameterized mesh and a texture map the visual perceptual attributes of the texture are first computed using a visual discrimination metric, then the precomputation of visual saliencies is then used to guide the distribution of samples to the surface mesh. During mesh remeshing, few samples are distributed to texture areas with strong visual masking properties and more samples to texture areas with weaker visual masking properties.

Meanwhile, we note that applying mesh saliency to image processing is a problem that has not yet been attempted, whereas investigating the problem can potentially benefit many tasks, such as texture synthesis. In [103], the algorithm called *feature-aligned shape texturing* extracts some salient feature curves of a 3D shape as prior knowledge of the sequent process of texture analysis and synthesis, where periodic texels are oriented and aligned with the direction of the salient curves on the 3D shape. One limitation is that the texel structure of source patterns is in the form of stripe or block with regular boundaries, to which those linear shape features are catered. Nonetheless, if fractal textures are used, the algorithm may fail to deal with a case where the structure turns to be irregular. In the case, the tool of mesh saliency, detecting regional saliencies of the shape, can be a better choice to obtain prior knowledge, such that, given a salient region and an unsalient region, different schemes of texture synthesis are employed to determine the location and connectivity of the texels.

7.3. Object recognition vs. mesh saliency

This subsection mentions another active research area called *object recognition*, which is relevant to mesh saliency, due to the correlation between salient points and interest points as explained in Section 7.1. Generally speaking, the goal of an object recognition method, which consists of three main steps: feature detection, feature matching, and feature verification, is to correctly identify objects present in a range image, and determine their poses (i.e., positions and orientations) [104]. By contrast, mesh saliency aims to detect regions and points of a range image, and determine their saliencies (i.e., distinctness, rarity, and surprise). Though their

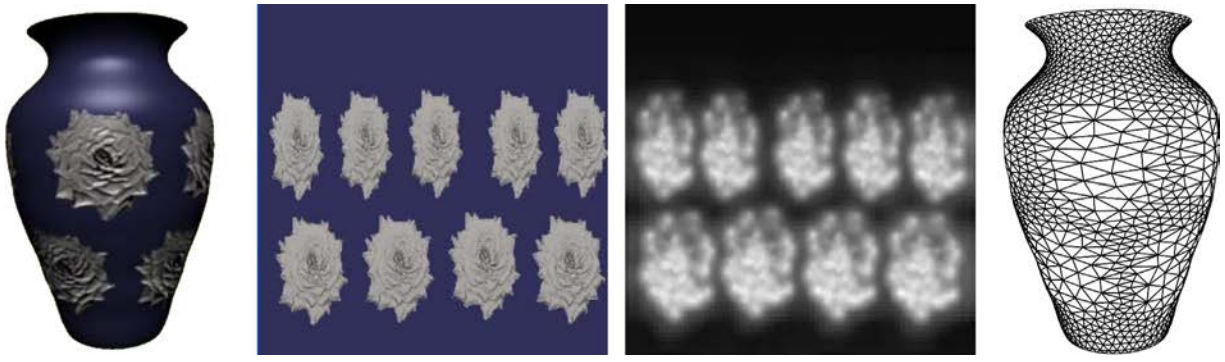

Fig. 18. An illustration of the case where visual saliency is applied to mesh remeshing [79]. Left: the original textured vase model; middle: the texture and its computed visual saliencies; right: the saliency-guided remeshed vase model. The original vase has 7171 vertices, and the remeshed vase has 2000 vertices.

Table 2
The state-of-the-art work of mesh saliency detection that is ordered by year. Note that the eleventh column only lists applications the work itself tested. The set of applications includes simplification (a), segmentation (b), good view selection (c), shape matching and retrieval (d), alignment (e), watermarking (f), feature evaluation (g), scan integration (h), illustration (i), medical imaging analysis (j), sampling (k), resizing (l), and smoothing (m).

Year	Sly	Type		Scale			Signal			Connectivity	Applications	Performance
		LC	*GC*	*SS*	*MS*	*GS*	*PF*	*CF*	*SF*			
2005	Lee et al. [15]	*			*		*			Mesh	a, c	Fast
2006	Lavoué et al. [37]	*		*			*			Mesh	f, g	Fast
	Gal and Cohen-Or [39]	*		*				*		Mesh	d, e	Slow
2007	Liu et al. [46]	*			*		*			Mesh	g	Fast
	Shilane and Funkhouser [58]	*		*				*		Mesh	a, c, d	Slow
2008	Zou et al. [9]	*			*		*			Mesh	d	Slow
	Atmosukarto and Shapiro [40]	*		*				*		Mesh	d	Slow
	Castellani et al. [42]	*			*		*			Point set	h	Fast
	Ohbuchi et al. [52]	*			*		*			Mesh	d	Slow
	Feixas et al. [24]	*				*	*			Mesh	c	Slow
	Zaharescu et al. [48]	*			*		*			Mesh	d	Fast
2009	Zou et al. [59]	*				*	*			Mesh	a, d, g	Fast
	Sun et al. [62]	*				*	*			Mesh	d	Fast
	Hu et al. [64]	*				*			*	Mesh	d	Fast
	Miao and Feng [41]	*		*			*			Mesh	i	Fast
2010	Pratikakis et al. [43]	*		*			*			Mesh	d	Slow
	Ruggeri et al. [65]	*				*			*	Mesh	d	Fast
2011	Godil and Wagan [51]	*			*		*			Mesh	d	Slow
	Benjamin et al. [68]	*				*	*			Mesh	b	Fast
2012	Hou and Qin [53]	*			*				*	Mesh/point set	g	Fast
	Leifman et al. [69]	*				*		*		Mesh	c	Fast
	Zhao and Liu [72]	*				*		*		Mesh	a	Slow
	Song et al. [77]		*		*		*			Mesh	a	Fast
	Ivan and Benjamin [32]		*	*				*		Mesh	–	Fast
	Hadi et al. [47]	*			*		*			Mesh	d	Fast
2013	Doraiswamy et al. [63]	*				*	*			Mesh	–	Fast
	Song et al. [67]	*			*				*	Mesh	–	Fast
	Wu et al. [74]	*				*		*		Mesh	a, k	Slow
2014	Song et al. [7]	*				*			*	Mesh	a, b, h	Fast
	Jia et al. [54]	*			*			*		Mesh	l	Fast
	Miao et al. [55]	*			*		*			Mesh	a	Fast
	Bulbul et al. [75]	*				*		*		Mesh	–	Slow
	Wang et al. [29]		*			*			*	Mesh	b, m	Fast
2015	Tao et al. [14]	*				*		*		Mesh	–	Slow
	Tasse et al. [56]	*				*		*		Point set	–	Fast

goals are different, likewise keypoints widely used in the application of object recognition (see the latest survey in [11]), salient points have been found uses in extracting surface features and devoted to matching similarities between two or more shapes [58,9,40,52,59].

7.4. Scene saliency vs. mesh saliency

In the context of image processing, visual saliency detects salient objects within a 2D image representing a shot of a real scene. Often, in computer graphics a scene is the form of 3D scanned point clouds. As a result, in conjunction with the rapid technology advancement of data acquisition and computing power, *scene saliency* becomes a more promising research area, but defining a distinctness metric that suits point sets is challenging since the data designated to work is not only noise-prone but also large-scale. By far, little attention have been paid to overcoming the challenges.

A remarkable progress was made by Shtrom et al. [93], who presented an efficient means in a multi-level manner. For a point $\mathbf{p}$, a histogram (e.g., 33 bins) of the quantized angular variations between $\mathbf{p}$ and its neighbors is constructed, with which a slightly-modified point descriptor called *fast point feature histogram* is computed. Then, given two points $\mathbf{p}_i$ and $\mathbf{p}_j$, their low-level dissimilarity $d(\mathbf{p}_i, \mathbf{p}_j)$ is proportional to the difference measured under χ^2 and inverse proportional to the distance measured under ℓ_2. Due to the lack of connectivity between points, the neighborhood of $\mathbf{p}_i$ is specified as $d_{\chi^2}(\mathbf{p}_i, \mathbf{p}_j) \le d_{\min}$, where $d_{\min}$ is tuned. The final saliency map is comprised of three components: low-level distinctness, point association, and high-level distinctness.

Unlike the above method [93], it was argued that in larger scene saliencies are units of various classes [105], with a naive assumption that computing object-based saliency in laser data is formulated as searching for a composite of geometric features (GFs). First, the seeds of line-type and plane-type GFs are detected via a decision tree; next, two procedures of region growing and fragment merging are conducted. In order to describe the obtained combinations properly and compactly, a graphical object representation is used, such that all potential objects in the range image are represented by graphs. Consequently, the object-based saliency computation is converted into a graph matching problem. To do so, graphs of objects of interest are first learned via manually chosen object samples composed of laser points. Because a limited set of GFs is used, the method runs rather efficiently, even if the scenes designated are complex.

Another object-based method was presented in [106]. Differing from the aforesaid method [93] where a few classes are predefined, it conducts a strategy of *planar region extraction* iteratively before proceeding saliency evaluation, until 95% of the point cloud is processed, such that all potentially distinct objects are singled out. After that, two steps of Euclidean clustering and saliency ranking are executed. However, we are aware that the method tested in a controlled indoor environment, if not impractical, is hard to work in real scenes.

8. Conclusions

The survey provided the first and comprehensive reference resource of mesh saliency detection as well as its applications, whose structure was depicted in Fig. 2 in Section 2.

Overall, local contrast based mesh saliency detection methods, elaborated in Section 3, may be sensitive to noises and other transformations. On the other hand, although local saliencies can represent visually interesting information of 3D objects at fine scales, often they are too “dispersed”, even the structures of 3D models are somewhat “damaged”. Therefore, the global contrast based methods, elaborated in Section 4, are capable of finding larger and more interesting structures. It is expected that more attentions will be paid to developing GC methods to overcome the shortcomings LC methods have. Additional, rather than introducing the large body of applications enumerated in Table 2, Section 5 chose six typical scenarios where mesh saliency plays a key role, including mesh simplification, mesh segmentation, mesh resizing, normal enhancement, volume rendering, and 3D printing.

To shed light on current trends and outlook of mesh saliency detection for future study, in Section 7 three insights into the topic were presented, and some open problems were raised. Unlike images and videos where color and coherence are the most important attributes, meshes esteem geometry and topology of shapes as key contributors. As a result, the nature of the task is different, so are the solutions to saliency. Furthermore, salient points differ from interest points and critical points, as mesh saliency metrics are required to be perception-dependent (see Fig. 17). Even if 3D data requires its own processing and analysis methods, basic knowledge about the similarities and differences between other topics and the topic itself would be helpful for making progresses in the future.

Acknowledgements

This work was sponsored by China Scholarship Council (No. 201206230015). Also, the authors were supported in part by National NSFC project (61222206), National NSFC key program project (No. 61133009) and National 973 Program of China (No. 2011CB302203).

References

[1] Berger J. Ways of seeing: based on the BBC television series. Penguin Books. Reprint edition, 1990.
[2] Berger J. About looking, Vintage. Reprint edition, 1992.
[3] Tomita H, Ohbayashi M, Nakahara K, Hasegawa I, Miyashita Y. Top-down signal from prefrontal cortex in executive control of memory retrieval. Nature 1999;401(6754):699–703.
[4] Buschman TJ, Miller EK. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. Science 2007;315(5820):1860–62.
[5] Zanto TP, Rubens MT, Thangavel A, Gazzaley A. Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. Nat Neurosci 2011;14(5):656–61.
[6] Katsuki F, Constantinidis C. Early involvement of prefrontal cortex in visual bottom-up attention. Nat Neurosci 2012;15(8):1160–6.
[7] Song R, Liu Y, Martin RR, Rosin PL. Mesh saliency via spectral processing. ACM Trans Graph 2014;33(1):6:1–17.
[8] Dutagaci H, Cheung CP, Godil A. Evaluation of 3D interest point detection techniques via human-generated ground truth. Vis Comput 2012;28(9):901–17.
[9] Zou G, Hua J, Dong M, Qin H. Surface matching with salient keypoints in geodesic scale space. Comput Anim Virtual Worlds 2008;19(3–4):399–410.
[10] Rusinkiewicz S. Estimating curvatures and their derivatives on triangle meshes. In: 2nd international symposium on 3DPVT; 2004. p. 486–93.
[11] Guo Y, Bennamoun M, Sohel F, Lu M, Wan J. 3D object recognition in cluttered scenes with local surface features: a survey. IEEE Trans Pattern Anal Mach Intell 2014;36(11):2270–87.
[12] Creusot C, Pears N, Austin J. A machine-learning approach to keypoint detection and landmarking on 3D meshes. Int J Comput Vis 2013;102(1–3):146–79.
[13] Cheng M-M, Zhang G-X, Mitra N, Huang X, Hu S-M. Global contrast based salient region detection. In: IEEE CVPR; 2011. p. 409–16.
[14] Tao P, Cao J, Li S, Liu X, Liu L. Mesh saliency via ranking unsalient patches in a descriptor space. Comput Graph 2015;46:264–74.
[15] Lee CH, Varshney A, Jacobs DW. Mesh saliency. In: SIGGRAPH'05; 2005. p. 659–66.
[16] Koch C, Poggio T. Predicting the visual world: silence is golden. Nat Neurosci 1999;2(1):9–10.
[17] Garland M, Heckbert PS. Surface simplification using quadric error metrics. In: Proceedings of the 24th annual conference on computer graphics and interactive techniques. New York: ACM Press/Addison-Wesley Publishing Co.; 1997. p. 209–16.
[18] Reuter M, Wolter F-E, Peinecke N. Laplace–Beltrami spectra as ‘Shape-DNA’ of surfaces and solids. Comput Aided Des 2006;38(4):342–66.
[19] Meyer M, Desbrun M, Schröder P, Barr AH. Discrete differential-geometry operators for triangulated 2-manifolds. In: Visualization and mathematics, vol. III. New York: Springer; 2003. p. 35–57.
[20] Grill SK, Malach R. The human visual cortex. Annu Rev Neurosci 2004;27(1):649–77.
[21] Bruce N, Tsotsos J. Saliency based on information maximization. In: Advances in neural information processing systems. Cambridge, MA: MIT Press; 2005. p. 155–62.
[22] Striemer CL, Chouinard PA, Goodale MA, de Ribaupierre S. Overlapping neural circuits for visual attention and eye movements in the human cerebellum. Neuropsychologia 2015;69:9–21.
[23] Walther D, Koch C. Modeling attention to salient proto-objects. Neural Netw 2006;19(9):1395–407.
[24] Feixas M, Sbert M, González F. A unified information-theoretic framework for viewpoint selection and mesh saliency. ACM Trans Appl Percept 2009;6(1):1:1–23.
[25] Kim Y, Varshney A, Jacobs DW, Guimbretière F. Mesh saliency and human eye fixations. ACM Trans Appl Percept 2010;7(2):12:1–13.
[26] Itti L, Koch C. A saliency-based search mechanism for overt and covert shifts of visual attention. Vis Res 2000;40(10):1489–506.
[27] Fu X, Cai L, Liu Y, Jia J, Chen W, Yi Z, et al. A computational cognition model of perception, memory, and judgment. Sci China Inf Sci 2014;57(3):1–15.
[28] Liu R, Cao J, Lin Z, Shan S. Adaptive partial differential equation learning for visual saliency detection. In: 2014 IEEE conference on computer vision and pattern recognition (CVPR); 2014. p. 3866–73.
[29] Wang S, Li N, Li S, Luo Z, Su Z, Qin H. Multi-scale mesh saliency based on low-rank and sparse analysis in shape feature space. Comput Aided Geom Des 2015;35–36:206–14.
[30] Goferman S, Zelnik-Manor L, Tal A. Context-aware saliency detection. IEEE Trans Pattern Anal Mach Intell 2012;34(10):1915–26.
[31] Litman R, Bronstein AM, Bronstein MM. Diffusion-geometric maximally stable component detection in deformable shapes. Comput Graph 2011;35(3):549–60.
[32] Sipiran I, Bustos B. Key-components: detection of salient regions on 3D meshes. Vis Comput 2013;29(12):1319–32.
[33] Dryden IL, Mardia KV. Statistical shape analysis, vol. 4. Chichester: Wiley; 1998.
[34] Howlett S, Hamill J, O'Sullivan C. Predicting and evaluating saliency for simplified polygonal models. ACM Trans Appl Percept 2005;2(3).
[35] McDonnell R, Larkin M, Hernández B, Rudomin I, O'Sullivan C. Eye-catching crowds: saliency based selective variation. In: SIGGRAPH'09; 2009. P. 55:1–10.
[36] Koch C, Ullman S. Shifts in selective visual attention: towards the underlying neural circuitry. In: Vaina L, editor. Matters of intelligence, vol. 188; 1987. p. 115–41.
[37] Lavoué G, Gelasca ED, Dupont F, Baskurt A, Ebrahimi T. Perceptually driven 3D distance metrics with application to watermarking. In: SPIE optics+photonics. Bellingham: International Society for Optics and Photonics; 2006. p. 63120L–63132L.
[38] Corsini M, Gelasca ED, Ebrahimi T, Barni M. Watermarked 3-D mesh quality assessment. IEEE Trans Multimed 2007;9(2):247–56.
[39] Gal R, Cohen-Or D. Salient geometric features for partial shape matching and similarity. ACM Trans Graph 2006;25(1):130–50.
[40] Atmosukarto I, Shapiro LG. A salient-point signature for 3D object retrieval. In: Proceedings of the 1st ACM international conference on multimedia information retrieval, MIR'08; 2008. p. 208–15.
[41] Miao Y, Feng J. Perceptual-saliency extremum lines for 3D shape illustration. Vis Comput 2010;26(6–8):433–43.
[42] Castellani U, Cristani M, Fantoni S, Murino V. Sparse points matching by combining 3D mesh saliency with statistical descriptors. Comput Graph Forum 2008;27(2):643–52.
[43] Pratikakis I, Spagnuolo M, Theoharis T, Veltkamp R. A robust 3D interest points detector based on harris operator. In: Eurographics workshop on 3D object retrieval, vol. 1; 2010.
[44] Harris C, Stephens M. A combined corner and edge detector. In: Alvey vision conference, vol. 15; 1988. p. 50.
[45] Sipiran I, Bustos B. Harris 3d: a robust extension of the Harris operator for interest point detection on 3D meshes. Vis Comput 2011;27(11):963–76.
[46] Shen Liu Y, Liu M, Kihara D, Ramani K. Salient critical points for meshes. In: Symposium on solid modeling and applications; 2007. p. 277–82.

[47] Fadaifard H, Wolberg G, Haralick R. Multiscale 3D feature extraction and matching with an application to 3D face recognition. Graph Models 2013;75(4):157–76.
[48] Zaharescu A, Boyer E, Varanasi K, Horaud R. Surface feature detection and description with applications to mesh matching. In: IEEE conference on computer vision and pattern recognition, 2009. CVPR 2009; 2009. p. 373–80.
[49] Zaharescu A, Boyer E, Horaud R. Keypoints and local descriptors of scalar functions on 2d manifolds. Int J Comput Vis 2012;100(1):78–98.
[50] Lowe DG. Distinctive image features from scale-invariant keypoints. Int J Comput Vis 2004;60(2):91–110.
[51] Godil A, Wagan AI. Salient local 3D features for 3D shape retrieval. In: IS&T/SPIE electronic imaging; 2011. p. 78640S–78648S.
[52] Ohbuchi R, Osada K, Furuya T, Banno T. Salient local visual features for shape-based 3D model retrieval. In: IEEE international conference on shape modeling and applications, 2008. SMI 2008; 2008. p. 93–102.
[53] Hou T, Qin H. Admissible diffusion wavelets and their applications in space-frequency processing. IEEE Trans Vis Comput Graph 2012;19(1):3–15.
[54] Jia S, Zhang C, Li X, Zhou Y. Mesh resizing based on hierarchical saliency detection. Graph Models 2014;76(5):355–62.
[55] Miao YW, Hu FX, Chen MY, Liu Z, Shou HH. Visual salience guided feature-aware shape simplification. J Zhejiang Univ Sci C 2014;15(9):744–53.
[56] Tasse FP, Kosinka J, Dodgson N. Cluster-based point set saliency. In: Proceedings of the IEEE international conference on computer vision; 2015. p. 163–71.
[57] Shilane P, Funkhouser T. Selecting distinctive 3D shape descriptors for similarity retrieval. In: IEEE international conference on shape modeling and applications, 2006. SMI 2006; 2006. p. 18–18.
[58] Shilane P, Funkhouser T. Distinctive regions of 3D surfaces. ACM Trans Graph 2007;26(2).
[59] Zou G, Hua J, Lai Z, Gu X, Dong M. Intrinsic geometric scale space by shape diffusion. IEEE Trans Vis Comput Graph 2009;15(6):1193–200.
[60] Hamilton RS. The Ricci flow on surfaces. Contemp Math 1988;71(1):237–61.
[61] Jin M, Kim J, Luo F, Gu X. Discrete surface Ricci flow. IEEE Trans Vis Comput Graph 2008;14(5):1030–43.
[62] Sun J, Ovsjanikov M, Guibas L. A concise and provably informative multi-scale signature based on heat diffusion. In: Computer graphics forum, vol. 28. Wiley Online Library; 2009. p. 1383–92.
[63] Doraiswamy H, Shivashankar N, Natarajan V, Wang Y. Topological saliency. Comput Graph 2013;37(7):787–99.
[64] Hu J, Hua J. Salient spectral geometric features for shape matching and retrieval. Vis Comput 2009;25(5–7):667–75.
[65] Ruggeri MR, Patanè G, Spagnuolo M, Saupe D. Spectral-driven isometry-invariant matching of 3D shapes. Int J Comput Vis 2010;89(2–3):248–65.
[66] Banchoff T, et al. Critical points and curvature for embedded polyhedra. J Differ Geom 1967;1(245–256):34.
[67] Song R, Liu Y, Martin R, Rosin P. 3D point of interest detection via spectral irregularity diffusion. Vis Comput 2013;29(6–8):695–705.
[68] Benjamin W, Polk AW, Vishwanathan S, Ramani K. Heat walk: robust salient segmentation of non-rigid shapes. Comput Graph Forum 2011;30:2097–106.
[69] Leifman G, Shtrom E, Tal A. Surface regions of interest for viewpoint selection. In: 2012 IEEE conference on computer vision and pattern recognition (CVPR); 2012. p. 414–21.
[70] Johnson AE, Hebert M. Using spin images for efficient object recognition in cluttered 3D scenes. IEEE Trans Pattern Anal Mach Intell 1999;21(5):433–49.
[71] Zhao Y, Liu Y, Song R, Zhang M. A saliency detection based method for 3D surface simplification. In: 2012 IEEE international conference on acoustics, speech and signal processing (ICASSP). Kyoto: IEEE; 2012. p. 889–92.
[72] Zhao Y, Liu Y. Patch based saliency detection method for 3D surface simplification. In: 2012 21st international conference on pattern recognition (ICPR); 2012. p. 845–8.
[73] Wolfe JM. Guided search 2.0 a revised model of visual search. Psychon Bull Rev 1994;1(2):202–38.
[74] Wu J, Shen X, Zhu W, Liu L. Mesh saliency with global rarity. Graph Models 2013;75(5):255–64.
[75] Bulbul A, Arpa S, Capin T. A clustering-based method to estimate saliency in 3D animated meshes. Comput Graph 2014;43:11–20.
[76] Castelló P, Chover M, Sbert M, Feixas M. Reducing complexity in polygonal meshes with view-based saliency. Comput Aided Geom Des 2014;31(6):279–93.
[77] Song R, Liu Y, Zhao Y, Martin RR, Rosin PL. Conditional random field-based mesh saliency. In: 2012 19th IEEE international conference on image processing (ICIP); 2012. p. 637–40.
[78] Mao ZH, Ma LZ, Zhao MX, Li Z. Feature-preserving mesh denoising based on contextual discontinuities. J Zhejiang Univ Sci A 2006;7(9):1603–8.
[79] Qu L, Meyer GW. Perceptually driven interactive geometry remeshing. In: Proceedings of the 2006 symposium on interactive 3D graphics and games, I3D'06; 2006. p. 199–206.
[80] Cimen G, Bulbul A, Ozguc B, Capin T. Perceptual caricaturization of 3D models. In: Computer and information sciences, vol. III. London: Springer; 2013. p. 201–7.
[81] Li Y, Liu Y, Wang Y, Wu Z, Yang Y. 3D facial mesh detection using geometric saliency of surface. In: International conference on multimedia computing and systems/International conference on multimedia and expo; 2011. p. 1–4.
[82] Jänicke H, Chen M. A salience-based quality metric for visualization. In: Computer graphics forum, vol. 29. Wiley Online Library; 2010. p. 1183–92.
[83] Kim Y, Varshney A. Saliency-guided enhancement for volume visualization. IEEE Trans Vis Comput Graph 2006;12(5):925–32.
[84] Song P, Fu Z, Liu L, Fu C-W. Printing 3D objects with interlocking parts. Comput Aided Geom Des 2015;35–36:137–48.
[85] Shlafman S, Tal A, Katz S. Metamorphosis of polyhedral surfaces using decomposition. In: Computer graphics forum, vol. 21. Wiley Online Library; 2002. p. 219–28.
[86] Kraevoy V, Sheffer A, Shamir A, Cohen-Or D. Non-homogeneous resizing of complex models. New York: ACM Trans Graph 2008;27(article no. 111).
[87] Miao Y, Feng J, Pajarola R. Visual saliency guided normal enhancement technique for 3D shape depiction. Comput Graph 2011;35(3):706–12.
[88] W. Wang, H. Chao, J. Tong, Z. Yang, X. Tong, H. Li, et al. Saliency-preserving slicing optimization for effective 3D printing. Comput Graph Forum 2015;34:148–60.
[89] Zhou H, Liu Y, Li L, Wei B. A clustering approach to free form surface reconstruction from multi-view range images. Image Vis Comput 2009;27(6):725–47.
[90] Chen X, Saparov A, Pang B, Funkhouser T. Schelling points on 3D surface meshes. ACM Trans Graph 2012;31(4):29:1–12.
[91] Lang C, Liu G, Yu J, Yan S. Saliency detection by multitask sparsity pursuit. IEEE Trans Image Process: A Publication of the IEEE Signal Processing Society 2012;21(3):1327–38.
[92] Shen X, Wu Y. A unified approach to salient object detection via low rank matrix recovery. In: 2012 IEEE conference on computer vision and pattern recognition (CVPR). Providence: IEEE; 2012. p. 853–60.
[93] Shtrom E, Leifman G, Tal A. Saliency detection in large point sets. In: Proceedings of the 2013 IEEE international conference on computer vision; 2013. p. 3591–98.
[94] Lindstrom P, Turk G. Image-driven simplification. ACM Trans Graph 2000;19(3):204–41.
[95] Cignoni P, Rocchini C, Scopigno R. Metro: measuring error on simplified surfaces. Comput Graph Forum 1996;17:167–74.
[96] Aspert N, Santa-Cruz D, Ebrahimi T. Mesh: measuring errors between surfaces using the hausdorff distance. In: Proceedings of 2002 IEEE international conference on multimedia and expo, 2002. ICME'02, vol. 1; 2002. p. 705–8.
[97] Zhang F, Du B, Zhang L. Saliency-guided unsupervised feature learning for scene classification. IEEE Trans Geosci Remote Sens 2015;53(4):2175–84.
[98] Tombari F, Salti S, Di Stefano L. Performance evaluation of 3D keypoint detectors. Int J Comput Vis 2013;102(1–3):198–220.
[99] Milnor JW. Morse theory. no. 51. Princeton: Princeton University Press; 1963.
[100] Yoshizawa S, Belyaev A, Seidel H-P. Fast and robust detection of crest lines on meshes. In: Proceedings of the 2005 ACM symposium on solid and physical modeling. Cambridge: ACM; 2005. p. 227–32.
[101] Borji A, Itti L. State-of-the-art in visual attention modeling. IEEE Trans Pattern Anal Mach Intell 2013;35(1):185–207.
[102] Yan J, Zhu M, Liu H, Liu Y. Visual saliency detection via sparsity pursuit. IEEE Signal Process Lett 2010;17(8):739–42.
[103] Xu K, Cohen-Or D, Ju T, Liu L, Zhang H, Zhou S, et al. Feature-aligned shape texturing. ACM Trans Graph 2009;28(5):108.
[104] Bariya P, Nishino K. Scale-hierarchical 3D object recognition in cluttered scenes. In: 2013 IEEE conference on computer vision and pattern recognition; 2010. p. 1657–64.
[105] Zhao Y, He M, Zhao H, Davoine F, Zha H. Computing object-based saliency in urban scenes using laser sensing. In: 2012 IEEE international conference on robotics and automation (ICRA). Saint Paul: IEEE; 2012. p. 4436–43.
[106] Bhatia S, Chalup SK. Segmenting salient objects in 3D point clouds of indoor scenes using geodesic distances. J Signal Inf Process 2013;04(03):102–8.