Neurocomputing, Volume 197, 12 July 2016, Pages 1-13,
DOI: 10.1016/j.neucom.2016.01.012

Region-based Saliency Estimation for 3D Shape Analysis and Understanding

Yitian Zhao[a], Yonghuai Liu[b], Jian Yang[a], Yongjun Wang[c], Baogang Wei[d], , Yifan Zhao[e], and Yongtian Wang[a]

[a] *Beijing Engineering Research Center of Mixed Reality and Advanced Display, School of Optics and Electronics, Beijing Institute of Technology, China*
[b] *Department of Computer Science, Aberystwyth University, UK*
[c] *School of Mathematics and System Sciences, Beihang University, China*
[d] *College of Computer Science and Technology, Zhejiang University, China*
[e] *EPSRC Centre for Innovative Manufacturing in Through-life Engineering Services, Cranfield University, Cranfield, UK*

Abstract

The detection of salient regions is an important pre-processing step for many 3D shape analysis and understanding tasks. This paper proposes a novel method for saliency detection in 3D free form shapes. Firstly, we smooth the surface normals by a bilateral filter. Such a method is capable of smoothing the surfaces and retaining the local details. Secondly, a novel method is proposed for the estimation of the saliency value of each vertex. To this end, two new features are defined: Retinex-based Importance Feature (RIF) and Relative Normal Distance (RND). They are based on the human visual perception characteristics and surface geometry respectively. Since the vertex based method cannot guarantee that the detected salient regions are semantically continuous and complete, we propose to refine such values based on surface patches. The detected saliency is finally used to guide the existing techniques for mesh simplification, interest point detection, and overlapping point cloud registration. The comparative studies based on real data from three publicly accessible databases show that the proposed method usually outperforms five selected state of the art ones both qualitatively and quantitatively for saliency detection and 3D shape analysis and understanding.

Keywords: Saliency, 3D surface, Retinex, local detail, global geometry.

January 21, 2016

Published by Elsevier. This is the Author Accepted Manuscript issued with:
Creative Commons Attribution Non-Commercial No Derivatives License (CC:BY:NC:ND 3.0).
The final published version (version of record) is available online at DOI:DOI: 10.1016/j.neucom.2016.01.012.
Please refer to any applicable publisher terms of use.

1. Introduction

Visual saliency is a predictor of object regions which attract human attention: it indicates their relative importance and is closely related to how we perceive and process visual stimuli. It is under investigation by multiple disciplines, including cognitive psychology [36, 38], neurobiology [7, 26], and computer vision [17, 1]. In computer vision tasks, finding salient regions in the visual field is essential, because it allows computer vision systems to process a flood of visual information and allocate limited resources to relatively small but interesting regions, or to a few interesting objects. Most approaches focus on the extraction of surface regions that are significantly different from their surroundings. The detected salient regions facilitate the understanding of the structure and the search for the regions/components that are particularly important to some applications. For example, the saliency values of vertices are used to determine in a mesh simplification algorithm the order in which they are decimated.

The salient features of a surface typically characterize the surface well and form a basis for a non-global similarity measure among sub-parts of shapes [10]. However, most existing techniques employ purely geometric measures - such as local curvature - or require user input to indicate important areas. For example, Lee et al. [22] defined a *mesh saliency* by using the Gaussian-weighted mean curvatures. Intuitively, the salient regions are not always those with specific curvature profiles. In other words, strictly geometric measures, such as curvature maxima or minima, do not always correlate with perceptual importance.

1.1. Related Work

Most of the existing saliency detection and estimation works are inspired by corresponding work on 2D images [17, 14, 15, 13, 5, 31, 32, 3]. In this section, we will review the most related work that operate 3D data directly.

One of the earliest saliency estimations on 3D surfaces was proposed by Lee et al. [22]: it was inspired by earlier work on saliency detection in 2D images [17]. The authors introduced a novel approach: *mesh saliency*. For each vertex, it defines a function of the differences of Gaussian-weighted mean curvatures at successive scales. The final mesh saliency is computed by adding together the saliency maps at these scales, after applying a non-linear normalization of suppression. Howlett et al. [16] introduced a method for saliency prediction in simplified polygonal models by using eye-tracking

devices. The location of a participant when fixating while viewing a particular model was captured by a high-speed eye-tracking system. In Gal et al. [10], the salient geometric features were constructed by clustering together a set of descriptors that are sufficiently interesting, in the sense that they have a high curvature relative to their surroundings, and a high variance of curvature values. Shilane and Funkhouser [33] proposed a shape-based descriptor for the representation of each region of a given object and used it as a query into a database for the analysis of distinctive regions. The distinctiveness of the descriptor is computed by evaluating a retrieval performance metric: *discounted cumulative gain* (DCG). Feixas et al. [9] presented a unified framework for viewpoint selection via view-based mesh saliency. The view-based mesh saliency is based on the idea how the polygons are perceived from a set of viewpoints. The saliency of a polygon is estimated as Jensen-Shannon (JS) divergence between its visibility probability distribution and those of its neighbors. Leifman et al. [23] proposed a viewpoint selection method based on vertex distinctness. They introduced a diffusion distance-based dissimilarity measure, which models the difference between two spin images as a temperature field, and considers the diffusion process on the field. Wu et al. [39] proposed a method to detect the mesh saliency based on the observation that salient regions are both locally prominent and globally rare. The saliency is obtained by the linear combination of the local contrast and the global rarity.

The methods proposed in [5, 33] require per-category training with an extremely large dataset. In this case, the detected salient regions undesirably change with training data. A single-saliency map is usually generated [27, 22, 33, 5] for information fusion by simply computing the sum or the average of all multi-scale saliency maps, and then uses threshold-based methods to determine whether a point is salient or not. While these methods are fast, they do not make good use of the information embedded in the multi-scale saliency maps [35]. On the other hand, a large number of methods, such as [22, 10, 4], rely heavily on the curvature alone. Their performance may degrade catastrophically due to the high sensitivity of curvature to imaging noise.

Saliency is a relative concept, judged not just against local neighbours, but also global overall shape. Local details are such geometric properties of an object defined by local and neighboring points. Global geometry is such geometric properties of an object that are defined by all the points/segments in the dataset for its description. It is interchangeably used throughout

this paper with global shape and global information without any ambiguity. The limitations of existing methods can be summarized as: (i) they mainly consider immediately neighboring points, overlooking global geometry; and (ii) they confuse the global overall shape with local details.

1.2. Our Work

The proposed saliency estimation method consists of three main steps: bilateral normal filtering, vertex-based saliency estimation (VBS), and region-based saliency estimation (RBS). This is a modified, improved, and extended version of our work [40]. The original idea of the VBS method was proposed in [40], and it has been improved in this paper in the sense that a bilateral filter has been proposed to smooth surface normals and suppress imaging noise, the mean curvature feature has been replaced by the Relative Normal Distance (RND) and it has been validated more extensively using experiments based on real data. Since the VBS method does not always detect complete and meaningful regions, a region-based saliency estimation method is proposed in this paper in order to refine the vertex-based saliency, through considering the votes of different surface regions for estimating their relative contrast and saliency values instead. The final vertex-based saliency is estimated as a ratio between the region-based saliency and the vertex based saliency. In this case, the finally estimated saliency of points considers both local details and also the global geometry.

To validate the proposed method, real range scans and models of various objects of interest from three publicly accessible databases are used. Due to either lack or subjective nature of ground truth, the detected saliency is mainly presented for visual inspection. The usefulness of the detected saliency is then used to guide the existing techniques for such 3D shape analysis and understanding tasks as mesh simplification, interest point detection, and overlapping point cloud registration. Both the qualitative and quantitative experimental results show that the proposed method is more powerful and effective for saliency detection and 3D shape analysis and understanding than the five selected state of the art ones.

The main contributions of this paper are fourfold:

- A bilateral filter has been proposed to smooth the surface normals of faces. Since this filter prevents those with a large difference in the normal vectors from that of the face of interest for participation and is based on the statistics of the Euclidean distance between normals of

neighboring faces, it is more robust to imaging noise and the variation of imaging resolution and scale. The smoothed normals are then used to estimate the surface properties of interest. Experimental results based on real data show that the estimated features are more expressive of the geometry and details in a surface.

- Two new features: Retinex-based Importance Feature (RIF) and Relative Normal Distance (RND) have been proposed in this paper. The RIF simulates the characteristics of human visual systems, and can capture the overall geometry of a shape. The RND is defined by neighboring points and can thus capture the local details of the surface. The experimental results based on real data show that the two features are supplementary and powerful in describing different aspects of the structure in a shape and in detecting salient regions.

- A region-based saliency (RBS) estimation method has been proposed. The RBS method takes all the surface regions into account and each of them contributes to the estimation of the saliency of another, the RBS method is more robust to imaging noise and variation in imaging resolution and scale for the detection of semantically continuous and meaningful salient regions. Experimental results based on real data show that the proposed RBS method is successful in combining both the global shape and local details and outperforms qualitatively three selected state of the art ones for the detection of salient regions in the shapes with varied geometry and imaging noise.

- The detected saliency by the proposed method has been applied to guide the existing techniques for such 3D shape analysis and understanding tasks as surface simplification, interest point detection, and overlapping point cloud registration. Experimental results based on real data demonstrate that the RBS method is quantitatively more powerful and effective for 3D surface analysis and understanding than three selected state of the art ones.

The remainder of this paper is organized as follows. Our three-step saliency estimation approach is described in detail in Sections 2, 3, and 4 respectively. Then we experimentally demonstrate its robustness, effectiveness, and efficiency for saliency estimation in Section 5. The application of the detected saliency to guide the existing techniques for different tasks is

demonstrated in Section 6. Finally, we draw conclusions and indicate future work in Section 7.

2. Bilateral Normal Filter

In this section, a novel bilateral filter is proposed to smooth surface normals and suppress the imaging noise with minimal damage to the geometric features of the object. Bilateral filter [37] is a non-linear filter, where the weight considers two components in the spatial domain as well as in the intensity domain. The larger the distance in the spatial domain and the difference in the intensity domain, the less impact a neighbouring pixel will impose on the pixel of interest. The weight of each component is computed using a Gaussian function. It has been extended to filter the meshes in [8] due to its nonlinear, feature-preserving characteristics. Recently, the bilateral filter was extended in [22, 41] to smooth the normals of faces of meshes.

Let $\mathbf{n_i}$ be a unit normal of a triangular face i, and $\mathbf{c_i}$ be its centroid, the bilateral normal filter is usually defined as:

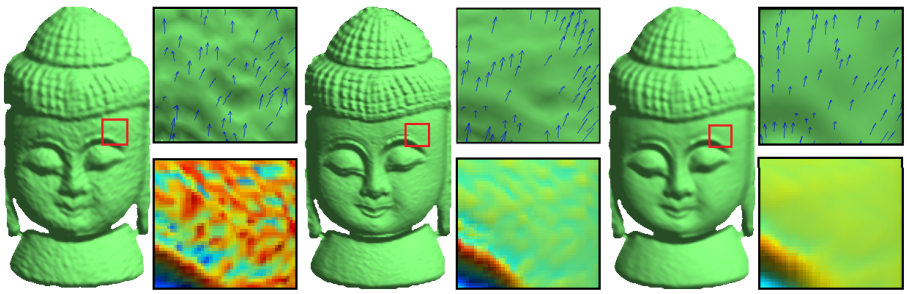
$$\overline{\mathbf{n}}_\mathbf{i} = \frac{\sum\limits_{j \in \mathbf{F}_i} W_c(\|\mathbf{c_i} - \mathbf{c_j}\|) W_s(\|\mathbf{n_i} - \mathbf{n_j}\|) \mathbf{n_j}}{\sum\limits_{j \in \mathbf{F}_i} W_c(\|\mathbf{c_i} - \mathbf{c_j}\|) W_s(\|\mathbf{n_i} - \mathbf{n_j}\|)} \tag{1}$$

where $\mathbf{F}_i$ is the 1-ring neighboring faces of i, the spatial weight function: $W_c = \exp\{-\frac{\|\mathbf{c_i}-\mathbf{c_j}\|^2}{2\sigma_c^2}\}$, and the feature preserving weight function σ_s: $W_s = \exp\{-\frac{\|\mathbf{n_i}-\mathbf{n_j}\|^2}{2\sigma_s^2}\}$ are the standard Gaussian function in terms of the geometric distances between the centroids and the unit normals of two faces, with the standard deviations σ_c and σ_s respectively. Since the absolute distance between normal vectors tends to capture the coarse and large variation and ignore the fine and small variation in the surface normal, a new bilateral normal filtering method is proposed in this section.

Given a face with a unit normal $\mathbf{n_i}$ and a centroid $\mathbf{c_i}$, the bilateral filtered normal $\overline{\mathbf{n}}_\mathbf{i}$ of the face is defined as:

$$\overline{\mathbf{n}}_\mathbf{i} = \frac{\sum\limits_{j \in \mathbf{F}_i} W_c(\|\mathbf{c_i} - \mathbf{c_j}\|) W_s(\mathbf{n_i}, \mathbf{n_j}) \mathbf{n_j}}{\sum\limits_{j \in \mathbf{F}_i} W_c(\|\mathbf{c_i} - \mathbf{c_j}\|) W_s(\mathbf{n_i}, \mathbf{n_j})} \tag{2}$$

where W_s is defined as the standard Gaussian function in terms of the relative normal distance $\mathrm{RND}(\mathbf{n_i}, \mathbf{n_j})$ instead:



(a) (b) (c)

Figure 1: The effect of smoothing by the proposed method on the scan *buddha*. (a) Noise corrupted scan with Gaussian white noise, standard deviation=0.2; (b) Original scan; (c) Smoothed version of (a). The regions in the red rectangle are also illustrated in the snapshots, showing that two surface features: surface normal (top), and mean curvature (bottom) are stable.

$$W_s(\mathbf{n_i}, \mathbf{n_j}) = \begin{cases} 0 & \text{if } (\mathbf{n_i} - \mathbf{n_j}) \cdot \mathbf{n_i} \geq \text{RND}(\mathbf{n_i}, \mathbf{n_j}) \\ \frac{1}{\sqrt{4\pi}} \exp\{-\frac{\text{RND}(\mathbf{n_i}, \mathbf{n_j})^2}{2}\} & \text{otherwise.} \end{cases} \tag{3}$$

The relative normal distance $\text{RND}(\mathbf{n_i}, \mathbf{n_j})$ between $\mathbf{n_i}$ and $\mathbf{n_j}$ is defined as:

$$\text{RND}(\mathbf{n_i}, \mathbf{n_j}) = \frac{\|\mathbf{n_i} - \mathbf{n_j}\|}{\text{ave}_{\mathbf{k} \subset \mathbf{F}_i}(\|\mathbf{n_i} - \mathbf{n_k}\|)} \tag{4}$$

where $\text{ave}_{\mathbf{k} \subset \mathbf{F}_i}(\|\mathbf{n_i} - \mathbf{n_k}\|)$ is the average Euclidean distance between $\mathbf{n_i}$ and other neighboring normals $\mathbf{n_k}$.

The relative distance is used in case the distribution of the data is not uniform. For two sets of points with a similar neighboring relationship but different densities, the Euclidean distances between corresponding points differ dramatically from each other, but the relative distances are in general similar. This is an advantage of the relative distance over the absolute distance in resisting the change of imaging resolution, occlusion, and viewpoint. We truncate the normal vectors if their differences are greater than the relative average distance of the normal vectors. Thus, large noisy normal vectors are excluded by this filter, leading the proposed relative normal distance based filter to be less sensitive to high levels of noise. After the face normals have been smoothed, the normal vector of a vertex is estimated as the weighted

average of the normals of the triangles sharing it with weights defined as the areas of these triangles, leading such properties as mean curvature and shape index to be accurately and reliably estimated.

Figure 1 shows the smoothing results for the scan *buddha* that was corrupted by Gaussian white noise. The quality of the smoothing is illustrated by different surface features in the snapshots of Figure 1: both the surface normal and mean curvature show that the proposed method is not only capable to preserve such local details as eyebrows, but also has the ability to smooth such flat regions as cheeks.

3. Vertex-based Saliency Estimation

The original vertex-based saliency was proposed in [40], we refer this method as OVBS in this paper. The OVBS diffuses the shape index with the surface curvatures, and builds a center surround operator for saliency detection. However, the combination of the shape index and surface curvatures as the individual feature channels have a drawback: they may contain a great deal of redundant information and thus may not be expressive to characterize saliency, since each of these channels is based on the principal curvatures. In order to address the drawbacks of the OVBS method, in this section, we thus improve it mainly based on two novel features: *Relative Normal Distance* proposed in the last section, and *Retinex-based Importance Feature* to be described below in this section.

3.1. Retinex-based Importance Feature

Land and McCann [20] first proposed the Retinex theory in 1986. Since then it has been influential in the field of computer vision, especially in the sense that is adapted to remove unfavourable illumination effects from images to improve their quality and contrast. We refer the reader to [20, 18] for more details. When the Retinex theory is applied to 2D images, it focuses on reflectance and illumination. It is our argument in this paper that the equivalence of reflectance can be used to enhance 3D shape information: global shape and local geometrical details, and the output of such enhancement is called *Retinex-based Importance Feature (RIF)*. After the generation of RIF, the 3D shape, component or surface can be represented more faithfully to the original. The important details or regions in which the geometry or topology differs significantly from their neighbors will be highlighted, in order to improve their saliency for subsequent processing.

Our work was inspired by the Retinex theory and is based on a bilateral filter for the enhancement and optimization of the shape index [19]. The shape index of a point u is defined as:

$$K = \frac{2}{\pi} \arctan \frac{k_2 + k_1}{k_2 - k_1} \tag{5}$$

where k_1 and k_2 are the maximum and minimum curvatures of point u and $K \in [-1, 1]$. Unlike the mean curvature and the Gaussian curvature, the shape index is invariant to scalings of the shape.

The bilateral filter in our case is written as:

$$L(u) = N^{-1}(u) \int_M K(\ell) g(\ell, u) s(\ell, u) d\ell, \tag{6}$$

where M is the 2-ring neighboring vertices of u: $\ell \in M$, $N(u)$ is a normalization factor: $N(u) = \int_M g(\ell, u) s(\ell, u) d\ell$, where the geometric measurement g and similarity function s are Gaussian functions of the Euclidean distance between their arguments:

$$g(\ell, u) = \exp\{-\frac{1}{2}(\frac{d(\ell, u)}{\sigma_d})^2\}, \tag{7}$$

$$s(\ell, u) = \exp\{-\frac{1}{2}(\frac{d(K(\ell), K(u))}{\sigma_r})^2\}, \tag{8}$$

where σ_d and σ_r control how distant vertices and dissimilar features will impact the filtering of the shape index of the current vertex. The larger the parameters σ_d and σ_r, the more contribution they make. Both σ_d and σ_r are empirically chosen as 0.3 in our work. $d(\ell, u)$ denotes the Euclidean distance between ℓ and u: $d(\ell, u) = \|\ell - u\|$, and $d(K(\ell), K(u))$ measures the absolute difference between two shape index values $K(\ell)$ and $K(u)$: $d(K(\ell), K(u)) = |K(\ell) - K(u)|$.

The bilateral filtering replaces the shape index value at vertex u with an average of similar and nearby shape index values. In the smooth regions, shape index values in a small neighborhood are similar to each other, and the filtered shape index will not change significantly. Therefore, the bilateral filtering averages away small, weakly correlated differences of shape index. The normalization term N ensures that the weights add up to one for all shape index values. As a result, the filter replaces the large shape index at

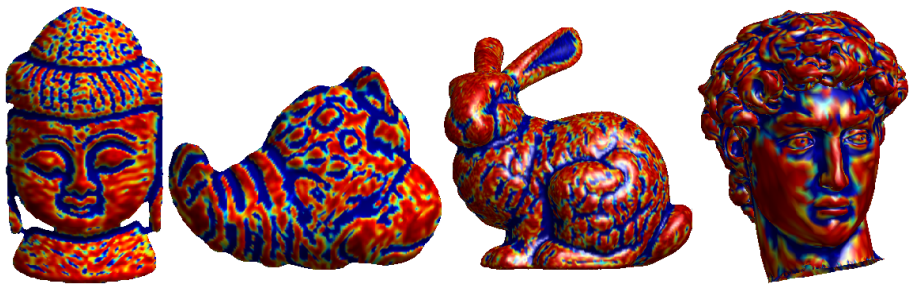


Figure 2: Example RIF of 3D models. It is easier to identify the local details of the given shapes due to clear representation of planar, convex and concave regions. From left to right: *buddha*, *lobster* , *bunny*, and *david*.

the centre by an average of the large shape index value in its vicinity, and vice versa.

Let K be the shape index values for the given 3D surface and let u be one of its vertices. By taking the difference between the logarithms of the input $K(u)$ and the bilateral-filtered shape index $L(u)$, the RIF is defined as:

$$\mathrm{RIF}(u) = log(K(u) + 1) - log(L(u) + 1). \tag{9}$$

From the definition, it can be seen that RIF is a relative measure and thus is expected to resist to the change in imaging resolution, noise and viewpoint. In line with the assumptions of the Retinex theory, the reflectance RIF is restricted to be in the range $\mathrm{RIF} \in [0, 1]$. Thus, we normalized it into the range $[0, 1]$.

Figure 2 illustrates the RIF maps of different 3D models. Overall, the results demonstrate that the global shape and local details have been more clearly represented. Areas such as the nose, mouth and eyes of the *buddha* and *david*, the bumpy area of the *lobster*, and the hair texture of the *bunny* can be easily recognized.

3.2. Vertex-based Saliency Estimation

Denote u and v as two vertices of a given mesh, whose coordinates are $u = (x_u, y_u, z_u)$ and $v = (x_v, y_v, z_v)$. The saliency value of u due to v in the RIF and RND channels are defined by the function $A(\mathrm{RIF}, u, v)$ and $A(\mathrm{RND}, u, v)$ respectively, where,

$$A(F,u,v) = \frac{\|F_u - F_v\|}{\sqrt{(x_u - x_v)^2 + (y_u - y_v)^2 + (z_u - z_v)^2}} \tag{10}$$

where F is either of the geometric features, RIF and RND, of a point. The numerator calculates the feature difference of points, while the denominator calculates their Euclidean distance. In this case, the more similar the geometric features of and the larger the Euclidean distance between points u and v, the less salient the point u is relative to the point v.

After all the saliency values of a point have been determined using different features, they are fused [40] together as a total and final saliency value $\text{Saliency}_{\text{VBS}}(u,v)$:

$$\text{Saliency}_{\text{VBS}}(u,v) = \sqrt{A(\text{RIF},u,v)^2 + A(\text{RND},u,v)^2}. \tag{11}$$

As a feature, RND describes the relative change in surface orientation, and is invariant to the relative change in the density of points and the scale of the objects. RIF simulates the characteristics of human visual perception of shapes, and enables different features and the global structure more discriminative. Thus, the proposed vertex-based method is expected to produce good results for detection of saliency on 3D shapes.

The proposed VBS method has a computational complexity of $O(n)$ in the estimation of shape index and relative normal distance of each vertex, $O(n)$ in bilateral filtering, and $O(n)$ in the estimation of vertex saliency. Thus, it has an overall linear computational complexity, $O(n)$, in the number n of vertices in the mesh.

4. Region-based Saliency Estimation

Although the VBS method described in the last section is able to capture the regions in which the most important features or components lie, the experiments still revealed that the method has a drawback: the detected salient regions in some cases are not continuous - in other words, it is hard to identify the meaningful regions completely. To overcome this drawback, in this section we propose a region based method to refine the estimated vertex saliency values. This method treats patches or regions instead as units for saliency estimation through considering how similar different patches are and letting them vote to each other for the determination of their relative contrast and saliency.

In this work, the mesh segmentation technique is embedded into the framework of the proposed region-based saliency (RBS) estimation method. For a recent review and comparison of the existing mesh segmentation techniques, we refer the reader to [2] and [30] for details. The method in [21] is employed here to segment a given surface into a number of regions according to the detected vertex based saliency values. The main reason for the selection of this method in [21] is that its code is freely available to download online. More importantly, this algorithm has a region merging process for the avoidance of over-segmentation, which is a non-trivial adaptation of an image processing method, taking into account common perimeters of the regions.

After the 3D surface has been segmented into several regions, we estimate the saliency value of each region. The saliency value of a region r_k considers not only its average vertex-based saliency value, but also how the vertex-based saliency values distribute in each region and how such distributions from different regions are similar:

$$\text{Saliency}_{\text{RBS}}(r_k) = \sum_{r_k \neq r_i} w(r_i) D_r(r_k, r_i) \tag{12}$$

where $w(r_i)$ is the weight of region r_i. Here we use the mean of the saliency values of vertices in segment r_i as $w(r_i)$. $D_r(r_k, r_i)$ is the similarity of the vertex-based saliency values between two regions r_k and r_i:

$$D_r(r_k, r_i) = \sum_{j=1}^{n_k} \sum_{l=1}^{n_i} f(c_{j,k}) f(c_{l,i}) D(c_{j,k}, c_{l,i}) \tag{13}$$

where n_k denotes the number of vertex-based saliency values in region r_k, $c_{j,k}$ indicates the j^{th} saliency value among all n_k relative saliency values in the region r_k, $D(c_{j,k}, c_{l,i})$=$|c_{j,k} - c_{l,i}|$, and $f(c_{j,k})$ is the frequency of the saliency value $c_{j,k}$, estimated by building a histogram over all $c_{j,k}$ values of vertices in the selected segment r_k. The number of bins in the histogram was set as 20 in this paper.

We further incorporate spatial information by applying a spatial weighting term into Equation 12 to increase the impacts of closer regions and decrease those of more distant regions. Specifically, for any region r_k, the spatially weighted region-based saliency is defined as:

$$\text{Saliency}_{\text{RBS}}(r_k) = \sum_{r_k \neq r_i} \exp\left(\frac{-D_s(r_k, r_i)}{\sigma_s^2}\right) w(r_i) D_r(r_k, r_i) \tag{14}$$

where $D_s(r_k, r_i)$ is the spatial distance between segments r_k and r_i, defined as the Euclidean distance between their centroids. σ_s controls the strength of the spatial weighting. Larger values of σ_s increase the spatial weighting of the distant segments and vice versa. In our implementation, we set $\sigma_s^2 = 0.4$ empirically.

It can be seen that each segment (region) has only one saliency value, leading to mosaic-like results. For more accurate saliency estimation, we propose to combine the saliency values at a region and any point inside respectively. The final saliency value $S(u)$ of a point u is therefore estimated as: $S(u) = \frac{\text{Saliency}_{\text{RBS}}(r)}{\text{Saliency}_{\text{VBS}}(u)}$, where $\text{Saliency}_{\text{VBS}}(u)$ is the saliency value at point u and $\text{Saliency}_{\text{RBS}}(r)$ is the saliency value of segment r in which point u lies. The saliency values are finally normalized into the range of $[0, 1]$ for the convenience of further saliency-guided applications.

The proposed RBS method has a computational complexity of $O(n)$ in segmentation, $O(N^2)$ in the estimation of region based saliency, and $O(n)$ in the estimation of vertex saliency, where N is the number of regions segmented in the mesh. Since N is usually significantly smaller than n, the proposed RBS method still has an overall linear complexity, $O(n)$, in the number n of vertices in the mesh.

5. Experimental Results and Analysis

In this section, we validate the proposed RBS method for the detection of saliency in 3D shapes using three publicly accessible databases: Ohio State University Range Image Database (OSURID)[1], the Stuttgart Range Image Database (SRID) [2], and the Stanford 3D Scanning Repository (S3DSR) [3]. While the bunny and david datasets were from S3DSR, the dragon dataset was from SRID, and all the remaining datasets were from OSURID. For the sake of comparative studies, three state-of-the-art methods were selected: the original VBS (OVBS) method [40], Mesh Saliency (MESA) [22], and Distinctive Region (DIRE) [33]. All the experiments described in this paper were implemented on a PC with 3.1Ghz Intel core i5 system with 8GB RAM. Since the ground truth is usually either missing or limited and subjective, the detected saliency results are presented mainly for visual comparison. To

[1] `http://sampl.eng.ohio-state.edu/~sampl/database.htm`

[2] `http://range.informatik.uni-stuttgart.de/htdocs/html/`

[3] `http://graphics.stanford.edu/data/3scanrep/`

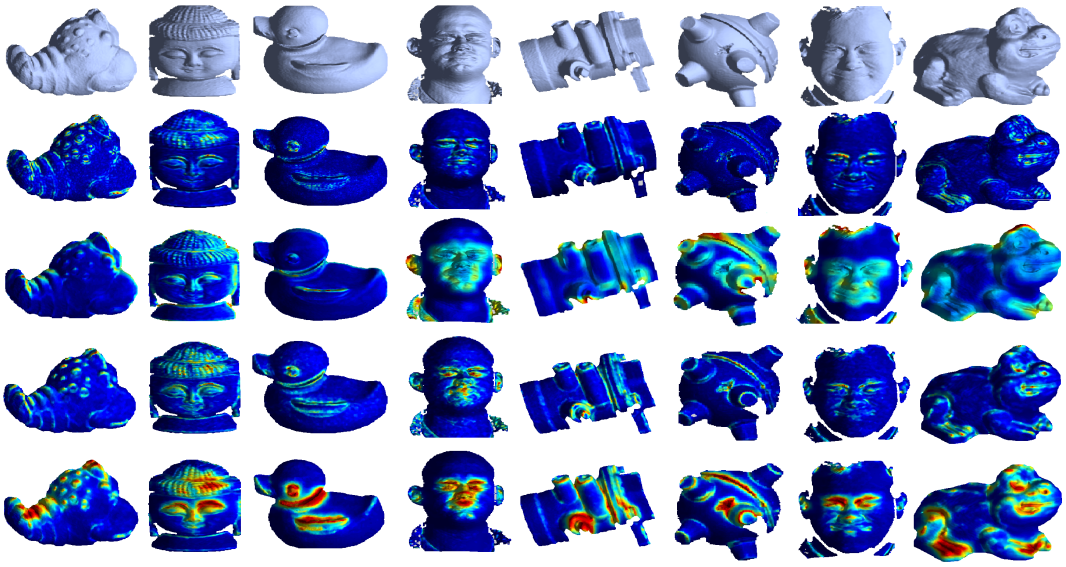


Figure 3: Salient regions detected by different methods in different real range scans. From top to bottom: the original, MESA, DIRE, OVBS, and RBS. From left to right: scans *lobster*, *buddha*, *duck*, *rick-face*, *valve*, *frame*, *pat-face*, and *frog*.

investigate whether the proposed method is robust to noise, we also add Gaussian white noise and compare the saliency detected before and after noise corruption. The details are presented in the following two subsections.

5.1. Saliency detection

Figure 3 shows saliency detected by different methods in 8 real range scans. The MESA method captured the visually salient features and detected the large curvature regions of the *lobster*, *buddha*, *duck* and *frog* surfaces as salient, but failed to describe such salient and important regions as the face and pipe orifices in the *rick-face*, *valve*, *frame* and *pat-face* scans. This is because the original models have a low resolution of 200 by 200 and contain significant noise, and the definition of saliency as a function of the mean curvature is sensitive to such low resolution and heavy imaging noise.

The DIRE method produced relatively better results than the MESA method. However, for some cases, such as the scan *rick-face*, it over-detected the salient regions: not only the eyes, nose and mouth, but also the forehead and some parts of the cheek. The hair region in the *buddha* scan was not completely detected, while such high curvature areas as the openings of the

pipe orifices of the valve and frame and such boundaries as the feet of the frog were detected as salient regions as well. This is because it is difficult for the DIRE method to represent surface regions and the definition of the distinctive regions varies with the database for their performance measurement.

In the case of the OVBS method, the salient regions are distributed usually in the geometrically complicated area, but the located salient regions are not always continuous. For instance, it detected the nose and eyes incompletely in the *buddha* scan. It hardly detected the nose, eyes and mouth areas as semantically meaningful and complete parts of the scan *pat-face*: because the shape index and the mean curvature used by this method are redundant as all the function of the same principal curvatures and thus less expressive of saliency and the centre-surround operator is not powerful enough to capture all the global shape information.

The last row of Figure 3 shows the results of the RBS method. This method addresses the drawbacks of all of the above methods: salient regions are falsely or partially detected, or missed; and over-sensitive to noise or boundaries. The detected salient regions are expanded, and usually correspond to the components of the object of interest. The "warm color" covers the most important regions (visually and geometrically) of the surfaces. This is because both the RIF and RND are expressive features for representing saliency and the region based voting is powerful for capturing the global geometry. The proposed RBS method strikes an appropriate balance between describing both local details and global geometry.

5.2. Effects of Noise

To demonstrate the robustness of our method, random Gaussian white noise with a standard deviation of $\sigma = 0.3$ was added to the original real range scans, and then both the RBS and OVBS methods were run again. The results of the detected saliency are shown in Figure 4. It can be seen that both the RBS and OVBS methods can still distinguish salient and non-salient regions in the presence of a considerable amount of noise. While the detected salient regions by the OBVS method scatter over the whole noise corrupted scans, as expected, those by the RBS method remain at the similar regions. These remarkable results show that the voting from different segments in the proposed RBS method is robust to imaging noise, and is powerful and stable in detecting salient regions from surfaces with varied complexities of geometry and magnitudes of imaging noise.

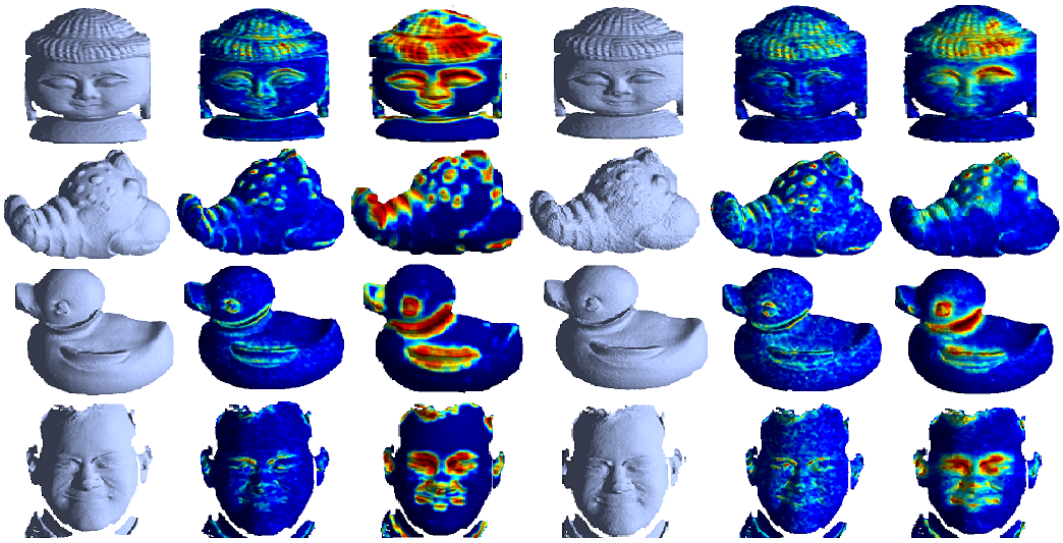


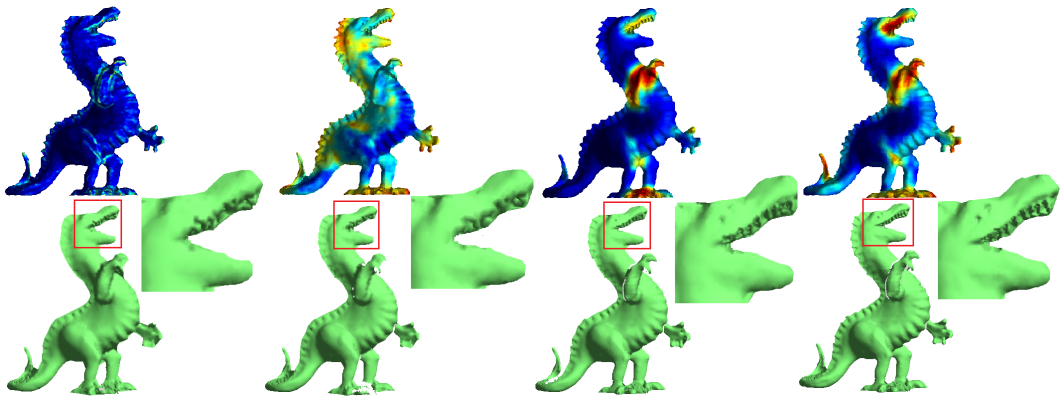
Figure 4: Saliency detected by the OVBS and RBS methods, with data corrupted by random Gaussian white noise with a standard deviation of $\sigma = 0.3$. From left to right: the original surface; OVBS; RBS; the noise corrupted surface; OVBS over the noise-corrupted surface; RBS over the noise-corrupted surface.

6. Saliency-guided 3D shape analysis and understanding

The detected saliency of 3D shapes is of broad interest since it can potentially improve the existing techniques for 3D shape analysis and understanding. In this section, the detected saliency by our method is applied to guide three tasks: interest point detection, surface simplification, and overlapping point cloud registration as detailed below.

6.1. Surface Simplification

To incorporate saliency information into surface simplification, in this section, we weight the quadric errors of edges [11] by the saliency values estimated by the proposed RBS method. The quality of a surface simplification method is measured by the Root Mean Square Error (RMSE) and Metro error [6] between the original and the simplified mesh. For the sake of comparative study, three existing state of the art techniques were selected again: the OBVS method [40], Mesh Saliency (MESA) [22], and Distinctive Region (DIRE) [33].

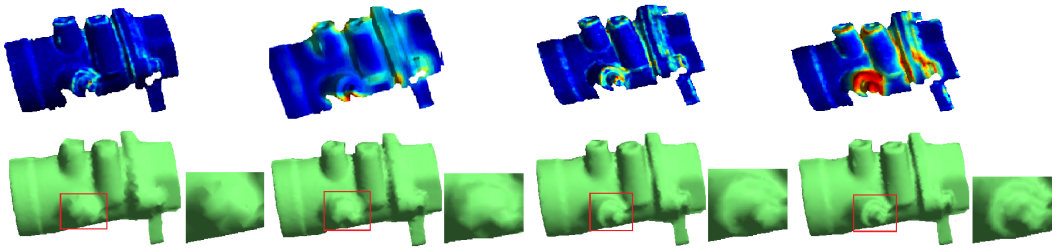


(a) (b) (c) (d)

Figure 5: Examples of simplification results guided by different methods on model *dragon* with a simplification rate of 95%. The snapshots illustrate the local details after simplification. The original model has 28730 vertices, and the simplified model contains 1436 vertices. (a) MESA; (b) DIRE; (c) OVBS; (d) RBS.

Figures 5 and 6 illustrate the detected saliency and saliency-guided surface simplification results, where the simplification rate was 95% for all the methods. It can be seen that both the MESA and DIRE methods smooth the teeth and mouth corners of the model dragon and the cylinder shapes of the pipe orifices of the model valve significantly. In contrast, both the OBVS and RBS methods achieved visually more pleasing results, not only retaining the model's global shape, but also preserving such significant local features as the teeth, mouth corners and the spine of the *dragon* and the pipe orifices of the *valve*.

The outstanding performances of the proposed method have been verified in Figure 7. All the surfaces were tested using three different simplification rates: 50%, 80% and 95%. As the simplification rate increases, the RMSE and Metro errors increase for all the methods, as expected. On careful scrutiny of all the cases, it can be seen that the RBS method always achieves the best simplification results than the alternatives, decreasing the RMSE and Metro errors of the existing methods by as much as 41% and 38% respectively. These remarkable results show that the simplified surfaces guided by our RBS method are better approximations of the original surfaces than those guided by the other methods, even though the imaging noise, the simplification rates and the geometry of the models vary significantly.



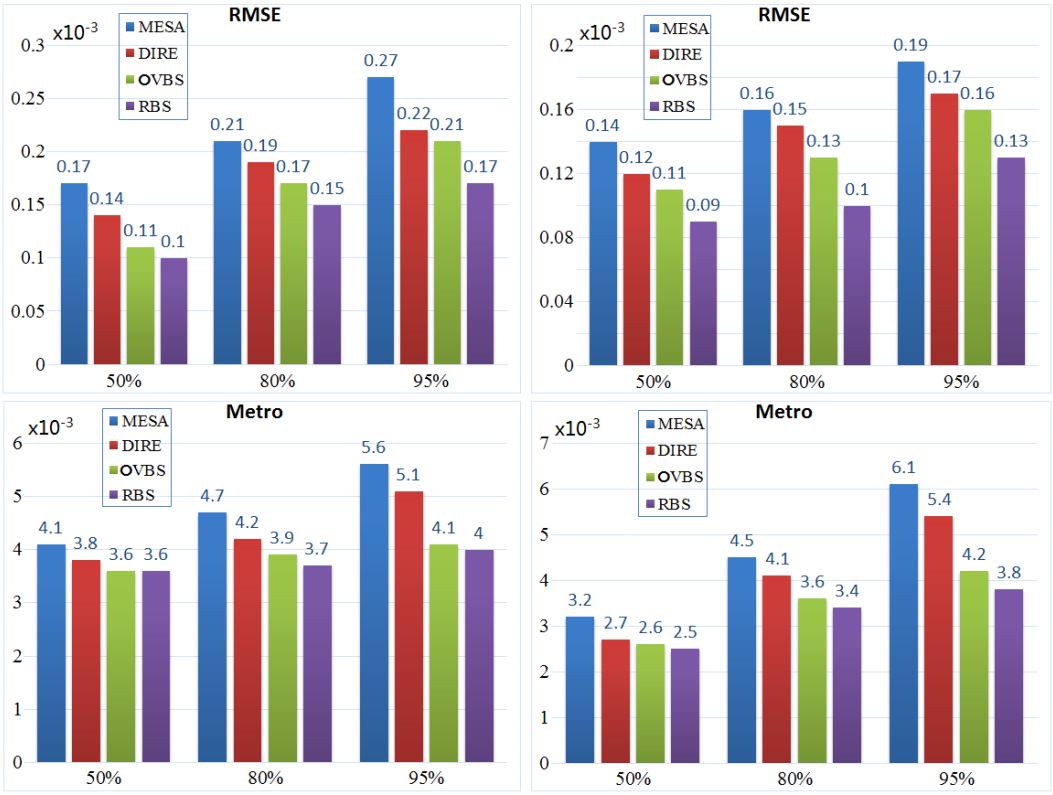
(a) (b) (c) (d)

Figure 6: Examples of simplification results guided by different methods on model *valve* with a simplification rate of 95% . The snapshots illustrate the local details after simplification. The original model has 12787 vertices, and the simplified model contains 640 vertices. (a) MESA; (b) DIRE; (c) OVBS; (d) RBS.

6.2. Interest Point Detection

In this section, we demonstrate the usefulness of the estimated saliency for the detection of interest points on 3D surfaces. In order to select interest points from the detected salient points with fairly even distribution, so that they describe not just salient features but also the overall geometry of salient regions, the voxelization is employed in our method. Voxelization is concerned with partitioning salient points into a set of voxels. Let the salient points be enclosed by a bounding box, and it can be divided into grids. Normally, the size of each block and the dimensions of the grid need to be specified by the user. In this paper, we need a high enough resolution to make sure that important details are not lost. In this case, as we only aim to detect the interest points from salient points rather than from the whole model, only such salient regions whose saliency values are larger than the average of the saliency values of all the points in the shape need to be voxelized. We defined the size of the voxel as $4mm \times 4mm \times 4mm$ in the experiments described below.

The larger the variation of the saliency values of the salient points inside a voxel, the larger the entropy, the more details the voxel contains. Thus the entropy can be used to guide the point sampling. The entropy of each voxel is estimated from a histogram with 20 bins of the finally estimated saliency values of salient points inside. The interest point selection criterion is as follows: let N_{vox} be the number of voxels after voxelization of the salient points, and let all the voxels be divided into two categories with equal



(a) (b)

Figure 7: RMSE and Metro errors of the mesh simplification guided by different methods over models with simplification rates of 50%, 80%, and 95% respectively. (a) dragon; (b) valve.

number of voxels by their entropy values: E_1 and E_2, where E_1 indicates the half of voxels with lower entropy values and E_2 indicates the half of voxels with higher entropy values. P_{req} denotes the total number of points to be sampled: then the numbers N_{E_1} and N_{E_2} of points to be selected from each voxel in the E_1 and E_2 sets, respectively, are:

$$N_{E_1} = \frac{0.5P_{req}}{N_{vox}}, N_{E_2} = \frac{1.5P_{req}}{N_{vox}}. \tag{15}$$

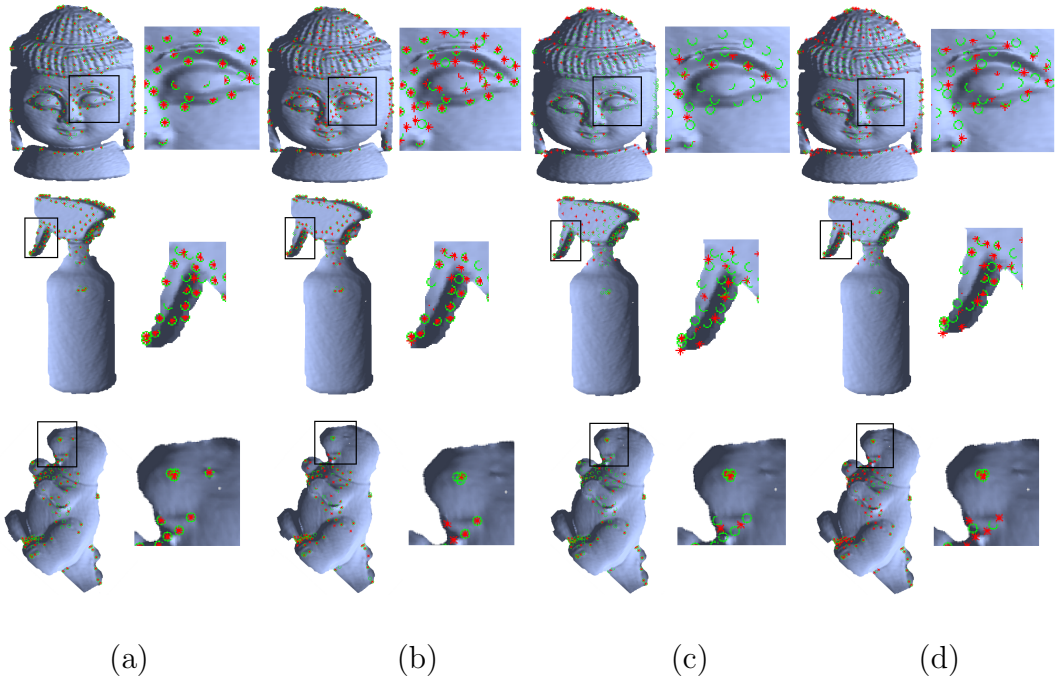
Clearly, more points are selected from the voxel with higher entropy (E_2) than the voxel with lower entropy (E_1). In our experiments described below, the ratio of N_{E_1} and N_{E_2} is set as 1:3, which provides the best performance

during the repeatability evaluation of the detected interest points.

For a comparative study, the saliency detected by the OVBS and RBS methods was used again to guide the interest point detection, and two state-of-the-art 3D interest point detectors, 3D-Harris [34] and 3D-SIFT [12], were also selected. 1% of points were sampled from the original shape as interest points, in order to make a fair comparison. The experimental results are presented in Figure 8. Green circles depict the interest points detected from the reference scans, and the red stars depict the points selected from the partially overlapping scans with the reference ones after some transformations. They show that each of the methods has the capability to detect interest points from the areas with a high density of local details. The interest points guided by the OVBS and RBS methods remain consistently detected, such as the mouth and eyes region of the *buddha*, the handle areas of the *bottle*, and the face of the *dinosaur*. However, 3D-Harris and 3D-SIFT failed to detect a large set of repeatable points as shown in Figure 8 (c) and (d) respectively.

To validate the robustness of our saliency-guided interest point detection method, and make a comparison to the competitors, the repeatability rate [29, 24] was used for quantitative evaluation. Figure 9 shows the repeatability rates of the interest points detected in different 3D scans subject to varied underlying transformations. Most detectors show excellent tolerance with small rotations of the underlying transformations, such as 20°. For the scan *buddha*, RBS achieves a repeatability rate of around 0.90 in Figure 9(a), it increases that of the OVBS, 3D-SIFT and 3D-Harris methods by as much as 6%, 20% and 29% respectively. As expected, with the rotation angle of the underlying transformation increasing, the repeatability rate falls, since the larger the rotation angle of the underlying transformation, the smaller the number of the overlapping points.

With the change of the rotation angle of the underlying transformation from 108° to 72°, the repeatability rates of the detected interest points drop dramatically by 32% and 19% over the scans of the bottle for the 3D-Harris and 3D-SIFT methods respectively. This is because it is difficult for 3D-Harris to select such parameters as the number of rings around each vertex, and for it to estimate corner response. As a result, the rate at which the detected interest points were false-positive increases. 3D-SIFT sometimes detects interest points in insignificant regions. Moreover, while it seeks to avoid detecting interest points from the boundaries and edges, this strategy did not help to identify the interest points, as in our test datasets, some salient points are located at boundaries and edges. While the proposed RBS



(a) (b) (c) (d)

Figure 8: The interest points detected on the overlapping scans buddha0-20, bottle144-180, and dinosaur36-72. All the detected interest points by different methods are represented over the first scans for convenient comparison. Green circles depict the interest points detected from the scans of buddha0, bottle144, and dinosaur36. Red stars depict the interest points detected from the scans of buddha20, bottle180, and dinosaur72. (a) RBS; (b) OVBS; (c) 3D-Harris; (d) 3D-SIFT.

method achieves relatively stable performances on the data captured from different viewpoints, it increases the repeatability rate of the OVBS method by as much as 9%. This is because it includes such ingredients that are supplementary and powerful: relative normal distance is expressive in describing local details, RIF is capable of distinguishing different local features and global geometry, and region based voting takes into account the global shape information. Consequently, they balance well the description of the interest points on the scale of not only the local details, but also the overall global geometry.

Computationally, our RBS and previously proposed OVBS-guided interest point detectors are much more computationally efficient than the 3D-Harris and 3D-SIFT methods. While the former took usually less than 30 seconds on average per scan from the OSURID database, the latter took

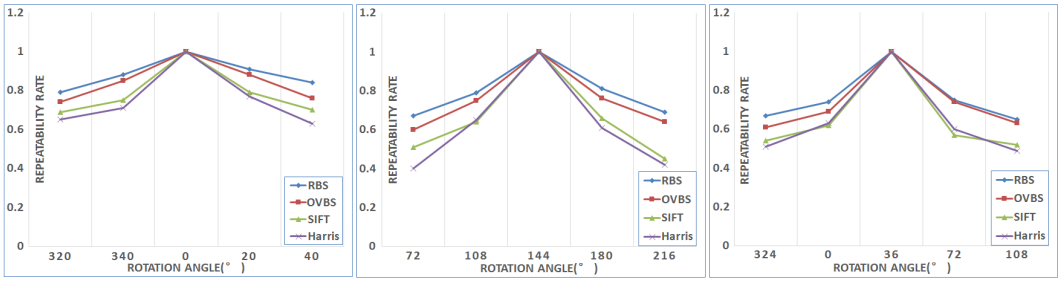


Figure 9: Repeatability rate of the interest points guided by different methods in different scans/models. From left to right: *buddha*, *bottle* and *dinosaur*.

usually more than 1 minute, due to such time intensive operations as local surface patch fitting and corner response estimation in the 3D-Harris method and the computation of the difference of Gaussian (DOG) at multiple scales and detection of extremal responses of DOG in both the spatial and scale spaces in the 3D-SIFT method.

6.3. Overlapping Point Cloud Registration

In this section, the usefulness of our detected interest points is further demonstrated by their application as proxy for the overlapping point cloud registration. The interest points detected by 3D-SIFT and 3D-Harris were also used for the same purpose for a comparative study, so that we can have a better idea whether the detected interest points can well pose the registration problem and which method can detect the most repeatable points for that purpose.

The registration algorithm, fractional iterative closest point (FICP) [28], was selected for the registration of the detected interest points due to its easy implementation and high accuracy. The performance of the registration algorithm was measured by the rotation angle of the finally estimated underlying transformation, average and standard deviation of errors in millimeters of reciprocal correspondences (RCs) [25] and the computational time in seconds. All the overlapping point clouds were subject to relatively small transformations, so that their underlying transformations can be uniformly initialized with the pure translational transformations defined by the difference of centroids of the detected interest points from the two overlapping point clouds. Since the same algorithm was used for registration, the difference in registration results will come uniquely from the detected interest

points: the more accurate the registration results, the more repeatable and informative the detected interest points.

Figure 10: Registration results using the FICP algorithm of the interest points detected from different scans using different algorithms. Columns, left to right: RBS, OVBS, 3D-SIFT, and 3D-Harris. Rows, top to bottom: valve0-10, bottle144-180, buddha0-20, and dinosaur36-72.

The complete sets of points in the original point clouds were also applied for registration by FICP and it was referred to as R-FICP in this section. It provides a performance baseline, allowing evaluation of the extent to which the selected interest points accelerate registration and affect its accuracy. To this end, the overlapping valve0-10, bottle144-180, buddha0-20 and dinosaur36-72 point clouds were selected from the OSURID database for

the experiments.

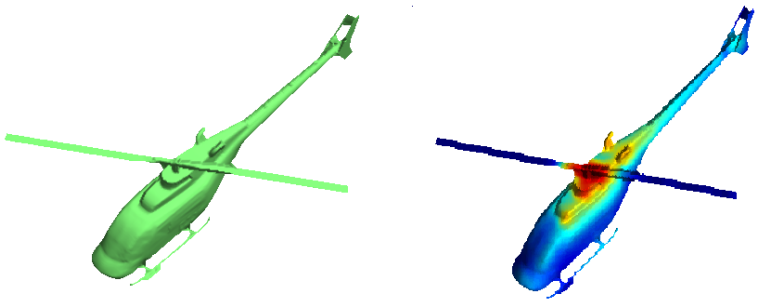
The experimental results are presented in Figure 10 and Table 1. For better visualization, the estimated underlying transformation was applied to the first point cloud, so that the transformed point cloud in *yellow* can be compared with the second reference point cloud in *green*. The interest points detected by 3D-Harris failed to pose the problem for the registration of the overlapping value0-10 point clouds, those detected by either 3D-SIFT or 3D-Harris also failed to act as proxy for the registration of the overlapping buddha0-20 and dinosaur36-72 point clouds, since the estimated underlying transformations failed to bring the scans valve0, buddha0, and dinosaur36 into the best possible alignment with the reference scans valve10, buddha20, and dinosaur72 respectively in 3D space. In sharp contrast, the proposed RBS method successfully identified repeatable interest points, yielding accurate estimations of the underlying transformations that bring all the four pairs of overlapping point clouds into accurate alignment with a large amount of interpenetration between each other.

These superior results have been verified by Table 1. The interest points detected by either the OVBS or RBS method largely overlap after transformation and the resulting estimated rotation angles of the underlying transformations are close to the ground truth. In contrast, those detected by either 3D-SIFT or 3D-Harris are largely displaced, leading the estimated rotation angles of the underlying transformations to have a relative error as large as 83%. While the proposed RBS and VBS methods increase the average error of the R-FICP algorithm slightly by 6% and 10% due to point sampling, both the 3D-SIFT and 3D-Harris methods increase it significantly by as much as 42% and 77% respectively. These results are consistent with those described in the last section. Both 3D-SIFT and 3D-Harris are sensitive to imaging noise and detected less repeatable and useful interest points.

Computationally, all the methods required similar time for the registration of the detected interest points using the FICP algorithm, since the same number of interest points were detected from different scans or models. The detected interest points guided by the saliency estimated by the proposed RBS method has been shown to be more useful in the context of efficient registration of overlapping 3D free form shapes. While it did not sacrifice registration accuracy, it enabled the registration task to be up to 10 times faster than using unsampled complete points.

Table 1: The average e_μ and standard deviation e_δ of registration errors in millimeters of RCs, expected and estimated rotation angles θ and $\widehat{\theta}$ in degrees of the underlying transformation, and the time T in seconds used for the registration of the detected interest points using the FICP algorithm.

Point clouds	Algorithm	$e_\mu(mm)$	$e_\delta(mm)$	$\theta(^\circ)$	$\widehat{\theta}(^\circ)$	$T(s)$
valve0-10	R-FICP	0.39	0.21	10	10.11	45
	RBS	0.42	0.23		10.41	4
	OVBS	0.44	0.26		10.61	4
	3D-SIFT	0.45	0.28		11.22	5
	3D-Harris	1.13	0.92		18.35	5
bottle144-180	R-FICP	0.47	0.38	36	35.68	32
	RBS	0.52	0.41		35.19	3
	OVBS	0.56	0.44		35.08	3
	3D-SIFT	0.89	0.64		30.96	3
	3D-Harris	0.91	0.66		28.81	3
buddha0-20	R-FICP	0.81	0.50	20	19.93	63
	RBS	0.84	0.55		19.45	4
	OVBS	0.86	0.57		19.21	4
	3D-SIFT	0.92	0.60		26.78	5
	3D-Harris	0.91	0.61		9.64	5
dinosaur36-72	R-FICP	0.63	0.85	36	34.14	34
	RBS	0.66	0.89		34.61	3
	OVBS	0.68	0.90		34.79	3
	3D-SIFT	1.01	0.90		25.76	3
	3D-Harris	1.12	0.97		26.17	3



(a) Helicopter (b) RBS

Figure 11: Limitation of the proposed method: some high level saliency was not captured.

7. Conclusion

This paper has proposed a novel technique for detecting the salient regions from different 3D scans or surfaces by combining surface smoothing, definition of expressive features, feature enhancement, and segments based voting. It also demonstrated how the detected saliency can both visually and quantitatively improve the results of several 3D surface analysis and understanding tasks, such as the simplification of surfaces, detection of interest points and registration of overlapping point clouds. Experimental results based on real range scans and models from three publicly accessible databases show that our proposed RBS method could be a more useful tool both qualitatively and quantitatively for the processing and understanding of 3D shapes than five selected state of the art ones.

The limitation of the current work is that it sometimes fails to detect the regions that are relatively similar to their neighbors, but salient to human perception, such as the propeller of the helicopter in Figure 11. The propeller is an important and informative clue to help human to recognize whether the machine is an airborne craft or vessel. Therefore, our future work will focus on how to incorporate such high level semantic cues into saliency detection.

It is worth noting that this paper focuses on the saliency analysis from objects subject to rigid transformations. In this case, the scale is not an issue since the interpoint distances in the objects will not be affected by such transformations. Future research will investigate the effect of the neighbourhood size on the interest point detection, and the usage of the detected saliency for other real-world applications such as 3D shape matching and retrieval.

Acknowledgments

We would like to express our sincere thanks to both the anonymous editor and reviewers for their constructive comments that have significantly improved the quality and readability of the paper. The funding from the National Basic Research Program of China (2013CB328806), the Key Projects in the National Science & Technology Pillar Program (2013BAI01B01), the National Hi-Tech Research and Development Program (2013AA013703), and the Beijing Institute of Technology Research Fund Program for Young Scholars are highly appreciated.

References

[1] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk. Frequency-tuned salient region detection. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, pages 1597–1604, 2009.

[2] M. Attene, S. Katz, M Mortara, and G. Patane. Mesh segmentationa comparative study. In *Proceedings of the Shape Modeling International*, pages 14–25, 2006.

[3] S. Ban, Y. Jang, and M. Lee. Affective saliency map considering psychological distance. *Neurocomputing*, 74(11):1916–1925, 2011.

[4] U. Castellani, M. Cristani, S. Fantoni, and V. Murino. Sparse points matching by combining 3d mesh saliency with statistical descriptors. *Computer Graphics Forum*, 27(2):643–652, 2008.

[5] M. Cheng, G. Zhang, N. Mitra, X. Huang, and S. Hu. Global contrast based salient region detection. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, pages 409–416, 2011.

[6] P. Cignoni, C. Rocchini, and R. Scopogno. Metro: measuring error on simplified surfaces. *Computer Graphics forum*, 17(2):167–174, 1998.

[7] R. Desimone and J. Duncan. Neural mechanisms of selective visual attention. *Annual review of neuroscience*, 19(6):247–248, 2009.

[8] H. Fan, Y. Yu, and Q. Peng. Robust feature-preserving mesh denoising based on consistent subneighborhoods. *IEEE Transactions on Visualization and Computer Graphics*, 16(2):312–324, 2010.

[9] M. Feixas, M. Sbert, and F. Gonzalez. A unified information-theoretic framework for viewpoint selection and mesh saliency. *ACM Transaction on Applied Perception*, 6(1):95–106, 2009.

[10] R. Gal and Cohen-Or. Salient geometric features for partial shape matching and similarity. *ACM Transactions on Graphics*, 25(1):130–150, 2006.

[11] M. Garland and P. Heckbert. Surface simplification using quadric error metrics. In *Proceedings of the SIGGRAPH*, pages 209–221, 1997.

[12] A. Godil and A.I. Wagan. Salient local 3d features for 3d shape retrieval. In *Proceedings of the 3D Image Processing (3DIP) and Applications II.*, pages 93–102, 2011.

[13] S. Goferman, L. Zelnik-Manor, and A. Tal. Context-aware saliency detection. *IEEE Transacation on Pattern Analysis and Machine. Intelligence*, 34(10):1915–1926, October 2012.

[14] J. Harel, C. Koch, and P. Perona. Graph-based visual saliency. *Advances in neural information processing systems*, 19:545–560, 2007.

[15] X. Hou and L. Zhang. Saliency detection: A spectral residual approach. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2007.

[16] S. Howlett, J. Hamill, and C. O'Sullivan. Predicting and evaluating saliency for simplified polygonal models. *ACM Transacation on Applied Perception*, 2(3):286–308, July 2005.

[17] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis,. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 20:1254–1259, 1998.

[18] D. Jobson, Z. Rahman, and G. Woodell. A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transaction on Image Processing*, 6(7):965–976, 1997.

[19] J. Koenderink and A. van Doorn. Surface shape and curvature scales. *Image and Vision Computer*, 10(8):557–565, October 1992.

[20] E. Land. Recent advances in retinex theory. *Vision Research*, 26(1):7–21, 1986.

[21] G. Lavoue, G. Dupont, and A. Baskurt. New cad mesh segmentation method a, based on curvature tensor analysis. *Computer Aided Design*, 37(10):975–987, 2005.

[22] C. Lee, A. Varshney, and D. Jacobs. Mesh saliency. *ACM Transactions on Graphics*, 24(3):659–666, 2005.

[23] G. Leifman. Surface regions of interest for viewpoint selection. In *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition*, pages 414–421, 2012.

[24] Z. Lian, A. Godil, T. Fabry, T. Furuya, J. Hermans, R. Ohbuchi, C. Shu, D. Smeets, P. Suetens, D. Vandermeulen, and S. Wuhrer. Shrec'10 track: non-rigid 3d shape retrieval. In *Proceedings of the 3rd Eurographics conference on 3D Object Retrieval*, pages 101–108, 2010.

[25] Y. Liu, R. Martin, L. Domincis, and B. Li. Using retinex for point selection in 3d shape registration. *Pattern Recognition*, 47(6):2126–2142, 2014.

[26] S. Mannan, C. Kennard, , and Husain. M. The role of visual salience in directing eye movements in visual object agnosia. *Current biology*, 18(1):193–222, 1995.

[27] M. Pauly, R. Keiser, and M. Gross. Multi-scale feature extraction on point-sampled surfaces. *Computer Graphics Forum*, 22(10-12):281–289, 2003.

[28] J. Phillips, R. Liu, and C. Tomasi. Outlier robust icp for minimizing fractional rmsd. In *Proceedings of 3DIM*, page 427434, 2007.

[29] C. Schmid, R. Mohr, and C. Bauckhage. Evaluation of interest point detectors. *International Journal on Comput. Vision*, 37(2):151–172, June 2000.

[30] A. Shamir. A formalization of boundary mesh segmentation. In *Proceeding of the second International Symposium on 3D Data Processing, Visualization, and Transmission*, pages 82–89, 2004.

[31] L. Shao and M. Brady. Specific object retrieval based on salient regions. *Pattern Recognition*, 39:1932–1948, 2006.

[32] L. Shao and M. Kadir, T. Brady. Geometric and photometric invariant distinctive regions detection. *Pattern Recognition*, 39:1932–1948, 2006.

[33] P. Shilane and T. Funkhouser. Selecting distinctive 3d shape descriptors for similarity retrieval. In *Proceeding of the Shape modelling international*, pages 18–27, 2006.

[34] I. Sipiran and B. Bustos. Harris 3d: a robust extension of the harris operator for interest point detection on 3d meshes. *The Visial Computer*, 27(11):963–976, November 2011.

[35] R. Song, Y. Liu, Y. Zhao, R. Martin, and P. Rosin. Conditional random field-based saliency detection. In *Proceeding of the International Conference on Image Processing*, pages 637–640, 2012.

[36] H. Teuber. Physiological psychology. *Annual Review of Psychology*, 6(1):247–248, 1955.

[37] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Proceedings of the International Conference on Computer Vision*, pages 839–845, 1998.

[38] A. M. Triesman and G. Gelade. A feature-integration theory of attention. cognitive psychology. *Annual Review of Psychology*, 12(1):97–136, 1980.

[39] J. Wu, X. Shen, W. Zhu, and L. Liu. Mesh saliency with global rarity. *Graphical Models*, 75:255–264, 2013.

[40] Y. Zhao, Y. Liu, R. Song, and M. Zhang. Extended non-local means filter for surface saliency detection. In *Proceeding of the International Conference on Image Processing*, pages 633–636, 2012.

[41] Y. Zheng, H. Fu, O. Au, and C. Tai. Bilateral normal filtering for mesh denoising. *IEEE Transactions on Visualization and Computer Graphics*, 17(10):1521–1530, October 2011.

Cranfield University

CERES https://dspace.lib.cranfield.ac.uk

School of Aerospace, Transport and Manufacturing (SATM) Staff publications (SATM)

2016-02-01

Region-based saliency estimation for 3D shape analysis and understanding

Zhao, Yitian

Elsevier

Yitian Zhao, Yonghuai Liu, Yongjun Wang, Baogang Wei, Jian Yang, Yifan Zhao, Yongtian Wang, Region-based saliency estimation for 3D shape analysis and understanding, Neurocomputing, Volume 197, 12 July 2016, pp1-13

http://dx.doi.org/10.1016/j.neucom.2016.01.012

Downloaded from Cranfield Library Services E-Repository