# 3D Point of Interest Detection via Spectral Irregularity Diffusion

**Ran Song · Yonghuai Liu · Ralph R. Martin · Paul L. Rosin**

**Abstract** This paper presents a method for detecting points of interest on 3D meshes. It comprises two major stages. In the first, we capture saliency in the spectral domain by detecting spectral irregularities of a mesh. Such saliency corresponds to the interesting portions of surface in the spatial domain. In the second stage, to transfer saliency information from the spectral domain to the spatial domain, we rely on spectral irregularity diffusion (SID) based on heat diffusion. SID captures not only the information about neighbourhoods of a given point in a multiscale manner, but also cues related to the global structure of a shape. It thus preserves information about both local and global saliency. We finally extract points of interest by looking for global and local maxima of the saliency map. We demonstrate the advantages of our proposed method using both visual and quantitative comparisons based on a publicly available benchmark.

**Keywords** Mesh saliency · Points of interest · Laplacian · Eigendecomposition

## 1 Introduction

Detection of points of interest on a 3D surface is a fundamental problem in computer vision and graphics, with applications to problems such as shape matching,

Ran Song (✉) · Yonghuai Liu
Department of Computer Science, Aberystwyth University, UK
E-mail: res@aber.ac.uk

Ralph R. Martin · Paul L. Rosin
School of Computer Science & Informatics, Cardiff University, UK

shape-based retrieval, mesh simplification and viewpoint selection. Its widespread applicability has led to a range of definitions in the literature. Usually, points of interest are considered to be points which are (i) locally distinctive, and (ii) likely to attract visual attention from a human. The former is merely related to the local geometry of the surface, and can be easily formulated once distinctiveness has been defined. However, the latter corresponds to the semantic relevance of points, and is hard to describe by a computational model even if this is often consistent to some degree with local geometric distinctiveness [3].

Generally, two different processes influence visual attention. One operates top-down and depends on the task at hand, or the cognitive and affective factors of the observer. The other operates bottom-up, and is driven by features within the input data. The relationship between top-down and bottom-up factors remains controversial. For example, can top-down interest counteract strong visual stimulus signals and shift attention away from bottom-up features? In this article, we focus on bottom-up point of interest detection. In our experiments, we will compare different bottom-up point of interest detectors based on human-generated ground truth data, in the hope that we can fairly measure how much a detector relates to the real interests of human perception.

### 1.1 Related work

Early work on detecting points of interest on 3D meshes mostly relied on measures of local surface geometry. [8] developed a curvature-related descriptor working at multiple levels of surface detail. A histogram of the descriptor was then calculated over the mesh, and points

corresponding to the least populated bins were selected as candidate points of interest. For shape retrieval, [19] proposed a Monte-Carlo strategy to select points on a surface with each point having the same initial probability of being chosen. The assumption behind this method is that the task can be affected by shape tessellations. [24] selected points that contribute to improving retrieval performance by assigning a predicted distinctiveness value to each selected point using a training phase. [29] built to detect points of interest. [28] assumed that the vertices of a 3D object have associated information such as curvature or photometric properties. The authors applied a discrete DoG on the function defined by the associated information.

However, these methods lack an important consideration in human perception: a mechanism take account of global shape information. They thus tend to miss the globally-salient points. For example, in the human-generated ground truth shown in [3,6], a large group of points (the segment centres) which are not salient in their local neighbourhoods are nevertheless marked as points of interest. Here, to analyse and capture the global shape in a mesh, we detect the irregularity of its Laplace spectrum composed of a group of eigenvalues related to the global structure of the mesh.

Since the seminal proposal of mesh saliency [16], a measure of perceptual importance on 3D meshes, recent papers have developed point of interest detectors incorporating this idea. [2] first defined a saliency measure by applying a Gaussian at the vertices, then a scale space was constructed and vertices highly displaced after the filtering were marked as candidate points of interest. [13] defined the geometric energy of a vertex via the eigendecomposition of the Laplace-Beltrami spectrum of a mesh, capturing saliency in the spectral domain. A vertex is selected as a point of interest if it remains as a local maximum of the geometry energy function over several successive frequencies. [26] used the heat kernel signature as a temporal domain restriction of the heat kernel to a manifold; this is also related to eigendecomposition of the Laplace-Beltrami operator. A vertex is selected as point of interest when its signature is a local maximum. Another method also employing the heat kernel for shape analysis is the autodiffusion function [7], which shows nice properties for feature based skeletonisation and segmentation. [20] extracted scale-invariant key points and ranked them using a measure directly related to their repeatability and the distinctiveness of the underlying local descriptor.

In general, saliency-based point of interest detectors use various mechanisms to extract saliency information, and then integrate such information in the process of localising points of interest. Their performances are thus largely dependent on the design of the computational models for saliency detection. In this work, we also concentrate on this vital stage, and give an algorithm for saliency detection together with an analysis.

## 1.2 Our work

Inspired by the spectrum-based methods for detecting 2D image saliency, we first capture potential saliency by investigating local irregularities in the spectrum of the Laplace-Beltrami operator. Then, to construct a saliency map, we use a curvature-weighted heat diffusion function to deliver and reorganise the saliency information into the spatial domain in a multiscale way. The points of interest are finally localised by considering both local and global maxima of the saliency map.

Two main contributions are highlighted in this paper. Firstly, we develop a fully automatic method for detecting sparse points of interest to humans on 3D surface meshes. To achieve this goal, we design a novel mechanism which combines both local geometric cues and global shape information obtained via spectral irregularity diffusion.

Secondly, we provide a theoretical analysis for the proposed saliency detection method based on spectral irregularity. To the best of our knowledge, it is the first time that spectral shape analysis has been used to detect saliency on 3D meshes. In our theoretical analysis, we investigate the intrinsic relationship between points of interest to human perception and the representation of shape information in the spectral domain.

## 2 Mesh saliency via spectral shape analysis

Various methods have detected saliency through spectral analysis of 2D images. One common motivation is the intrinsic relation between the spectral properties of images and the characteristics of the human perceptual system. On one hand, spectral analysis can provide information about global cues or statistics of images which are difficult to extract in the spatial domain; on the other hand, the human perceptual system has the ability to rapidly recognise objects or features without a slow process of scrutinising details. The human perceptual system achieves this by swiftly extracting certain low-frequency saliency information [10,12,18]; this can be computationally modeled via spectral analysis. In this work, we extend such spectral analysis to 3D shapes.

## 2.1 Preliminaries

For a 2D image, it is well-known that the spectral analysis can be performed via the Fourier transform. As summarised in [17], one way to extend the notion of Fourier analysis to the manifold or surface setting, where the signal represents the geometry of a 3D surface instead of the intensity of a 2D image, is to define appropriately a Laplacian operator and rely on its eigenvalues and eigenvectors to perform frequency analysis. The definition of the Laplacian can be generalised to functions defined over a manifold $\mathcal{M}$ with metric $g$, to give the Laplace-Beltrami operator [17]:

$$\Delta = \operatorname{div} \operatorname{grad} = \delta d = \sum_i \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial x_i} \sqrt{|g|} \frac{\partial}{\partial x_i} \qquad (1)$$

where $|g|$ is the determinant of $g$. Since a closed form expression for the eigenvectors of the Laplace-Beltrami operator on an arbitrary manifold is generally unavailable, the eigenvectors are numerically calculated by discretising the Laplace-Beltrami operator. To solve the eigenproblem

$$-\Delta \psi_j = \lambda_j \psi_j \qquad (2)$$

where $\lambda_j$ is the eigenvalue corresponding to the eigenvector $\psi_j$ and the '−' sign is here required for the eigenvalues to be positive, it is typical to discretise the Laplace-Beltrami operator using the cotan formulation [5,23]. Doing so, Eq. (2) is simplified to the following generalised eigenproblem in matrix form:

$$-Q\psi = \lambda B \psi \qquad (3)$$

where the matrix $Q$ is called the stiffness matrix and $B$ is the mass matrix. $\lambda$ represents the diagonal matrix formed by unknown eigenvalues (frequencies) $\lambda_f$ where $f$ is the frequency index. $\psi$ is a matrix whose columns are the corresponding eigenvectors. $Q$ and $B$ can be determined by the finite element method (FEM) [4,22, 17].

## 2.2 A computational saliency model

We define the Laplace spectrum of a mesh as $\mathcal{H}(f) = \{\lambda_f : 1 \leq f \leq m\}$ where $m$ denotes the number of vertices on the mesh. Fig. 1(a) shows the Laplace spectrum of a mesh and Fig. 2 shows plots of the eigenvectors of the Laplace operator for the *Girl* model.

[14] revealed that the human visual system regards features that deviate from the norm as informative and stays sensitive to them. The norm represents the expected, indistinctive and ever-present patterns. Its spectrum, as formulated in [12], is usually a smooth curve. [12] generated a residual spectrum of the log-Fourier
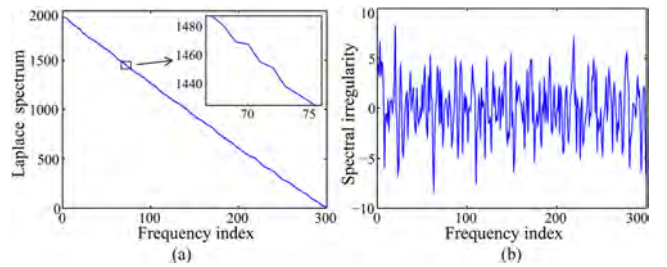


**Fig. 1** (a) The Laplacian spectrum (300 smallest eigenvalues only) of the *Girl* model (see Fig. 2); (b) Laplacian spectral irregularity.

spectrum of an image to approximately represent the features which deviate from the *norm* in the spectral domain. Inspired by this, we calculate spectral irregularities by computing the difference between the Laplacian spectrum and its locally smoothed spectrum.

A simple procedure to accomplish a locally smoothed spectrum is to adopt a discrete Laplacian, denoted by $\delta = \left[ \frac{1}{2} -1 \frac{1}{2} \right]$. The smoothed spectrum is calculated as

$$\mathcal{S}(f) = \mathcal{H}(f) + \frac{1}{2}\delta * \mathcal{H}(f) \qquad (4)$$

*Remark 1* As pointed out in [18], convolution of the amplitude spectrum with a low-pass Gaussian operator provides a saliency detector for 2D images. The effect of applying the 1D discrete Laplace operator over such a monotonic spectrum is similar to Gaussian smoothing since both lead to attenuation of high-frequency content. However, whereas Gaussian smoothing is expressed in terms of squared differences from the mean, Laplacian smoothing is expressed in terms of absolutes difference from the mean. Consequently, the Laplacian has a fatter tail than the Gaussian, so it is less locally-supported and thus a better approximation of the spectrum of the norm.

Spectral deviation can now be computed as the spectral irregularity $\mathcal{R}$:

$$\mathcal{R}(f) = \mathcal{H}(f) - \mathcal{S}(f) \qquad (5)$$

*Remark 2* The spectral decomposition of the Laplacian is well-known for facilitating analysis of global structure of shape [17]. However, saliency eventually needs to be localised in the spatial domain to generate a saliency map. Eigenvectors of the Laplacian, as shown in Fig. 2, give no help in localising salient features in the spatial domain. In fact, the use of eigenvectors can suppress the influence of small-scale fluctuations, as shown in the minimisation of the RatioCut model [11] where the eigendecomposition is used for graph partitioning. The spectral irregularity computed via Eq. (5) helps to localise saliency in the spatial domain. Note that in [16, 2], saliency is directly localised in the spatial domain
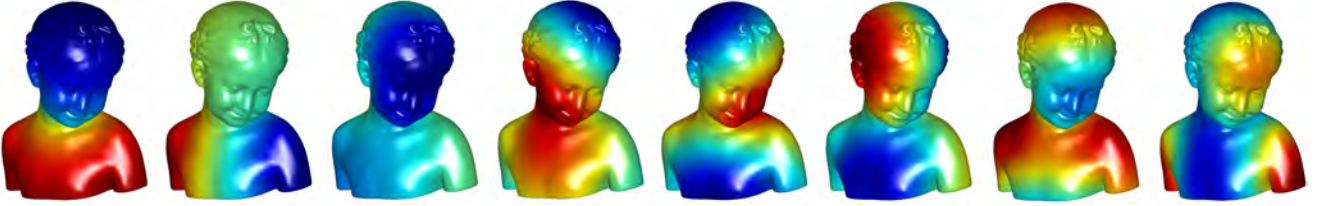
**Fig. 2** Plots of the eigenvectors of the Laplace operator corresponding to the eight smallest nonzero eigenvalues

by calculating a difference between Gaussian-weighted averages.

*Remark 3* The nice localisation property of spectral irregularity can also be interpreted via its relation to wavelets. Eqs. (4) and (5) and the curve shown in Fig. 1(b) are reminiscent of wavelet functions which have powerful localisation properties in both spectral and spatial domains. It is critical for a wavelet function to satisfy the *admissibility* condition since it ensures the wavelet transform can be recovered, leading to desirable properties for various graphics applications. Here we demonstrate that spectral irregularity is *admissible*.

**Proposition 1** . *The proposed spectral irregularity $\mathcal{R}$ suffices the admissibility condition, expressed in discrete manifold space as*

$$\sum_{f=0}^{\infty} \frac{|\Psi_{\mathcal{R}}(f)|^2}{f} < \infty \tag{6}$$

*where $\Psi_{\mathcal{R}}$ is the manifold harmonic transform of $\mathcal{R}$.*

*Proof* . Recall that the zero-frequency manifold harmonic basis $\phi_0 = \frac{1}{\sqrt{A}}$ where $A$ is the total area of the mesh, which is a constant everywhere on the mesh. According to Eqs. (4) and (5), $\mathcal{R}$ has zero mean. Hence, it vanishes at zero frequency in the manifold harmonic transform (MHT)[27],

$$\Psi_{\mathcal{R}}(0) = \langle \mathcal{R}, \phi_0 \rangle = 0, \tag{7}$$

where $\langle x, y \rangle$ denotes the inner product of $x$ and $y$. Also, since $\mathcal{R}$ is compactly-supported, it has limited bandwidth in its MHT. If $f^*$ is its upper frequency, we have

$$\Psi_{\mathcal{R}}(f) = 0, \quad \text{for} \quad f > f^*. \tag{8}$$

Therefore, we have

$$\sum_{f=0}^{\infty} \frac{|\Psi_{\mathcal{R}}(f)|^2}{f} = \sum_{f=1}^{f^*} \frac{|\Psi_{\mathcal{R}}(f)|^2}{f} < \infty, \tag{9}$$

Hence, $\mathcal{R}$ is *admissible*.

## 3 Spectral irregularity diffusion

In this section, we propose a method for transferring saliency captured by spectral irregularity to the spatial domain, based on diffusion.

Gaussian filtering has been widely used for multiscale saliency detection. Given a $d$-dimensional signal $U : \mathbb{R}^d \to \mathbb{R}$, its linear scale-space representation $F : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ is defined as the convolution:

$$F(\cdot, t) = U(\cdot) * g(\cdot, t), \tag{10}$$

where $t$ is a scale parameter and $g(\cdot, t)$ is a Gaussian kernel with standard deviation $\sigma$ determined by the scale parameter: $\sigma = \sqrt{t}$.

The generating equation of $F$ in Eq. (10) is the heat diffusion equation [15], so $F$ can also be obtained as the solution to a diffusion process:

$$\left( \Delta + \frac{\partial}{\partial t} \right) F(p, t) = 0. \tag{11}$$

Given an initial heat diffusion $F_0(p)$, for a mesh $M$ which can be regarded as a compact Riemannian manifold, the solution is expressed as

$$F(p, t) = \int_M h_t(p, q) F_0(q) dq \tag{12}$$

where $dq$ is the volume form at $q \in M$. The minimum function $h_t(p, q)$ that satisfies Eq. (12) is called the heat kernel, and can be regarded as the amount of heat that is transferred from vertex $p$ to $q$ in time $t$ given a unit heat source at $p$. The eigendecomposition of the heat kernel can be expressed as

$$h_t(p, q) = \sum_{f=1}^{m} e^{-\lambda_f t} \psi_f(p) \psi_f(q) \tag{13}$$

where again, $\lambda_f$ and $\psi_f$ are eigenvalues and eigenvectors of the Laplacian operator respectively.

The heat kernel $h_t(p, q)$ has quite a few nice properties [26]. However, computations using it have high complexity as it operates in both spatial and temporal domains. [26] demonstrated that under some mild assumptions, restricting the kernel to only the temporal domain can still preserve all of the information about the intrinsic geometry of the shape, and also retain the major useful properties of the heat kernel. The concise
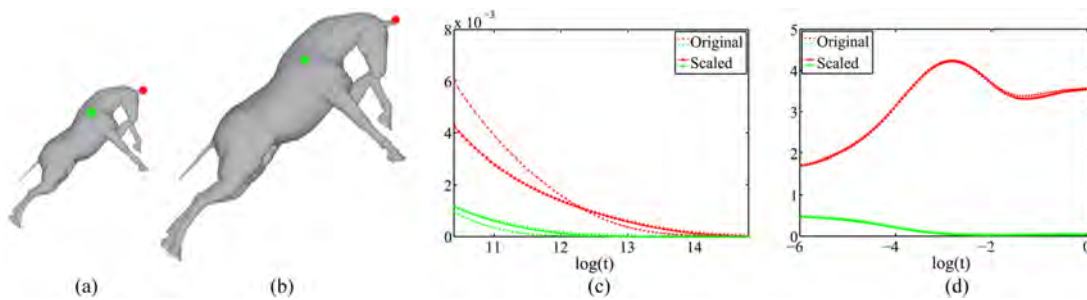
**Fig. 3** (a) The original model; (b) The scaled model where the length of the diagonal of its bounding box is 1.75 times of that of the original model; (c) The unnormalised IKS for the green and the red points on the two models; (d) The normalised IKS for the green and the red points on the two models.

heat kernel, or heat kernel signature (HKS) [26], is expressed as

$$H_t(p) = \sum_{f=1}^{m} e^{-\mathcal{H}(f)t} \psi_f^2(p). \tag{14}$$

### 3.1 Irregularity kernel signature

In this work, we propose a novel signature, which we call the *irregularity kernel signature* (IKS)

$$I_t(p) = \sum_{f=1}^{m} e^{-\mathcal{R}(f)t} \psi_f^2(p). \tag{15}$$

It can be proved that the IKS inherits most of the advantages of the HKS including isometric invariance, multiscale properties, stability, etc. Isometric invariance is important for shape matching and shape-based retrieval [26], for example.

**Proposition 2** : Isometric invariance. *If $T: M \to N$ is an isometry between two Riemannian manifolds $M$ and $N$, then $I_t^M(p) = I_t^M(T(p))$, $\forall p \in M$ and $\forall t > 0$.*

The IKS inherits the isometric invariance of the HKS because it is also based on the Laplace-Beltrami operator. The spectrum of the Lalace-Beltrami operator has isometric invariance since it only depends on the gradient and divergence which in turn only depend on the Riemannian structure of the manifold.

However, it inevitably inherits the disadvantages of the HKS, e.g., the sensitivity to scale [1]. Here, we normalise the Laplacian eigenvalues and eigenvectors to achieve scale invariance for the IKS

$$\mathcal{R}'(f) = \frac{A_p \mathcal{R}(f)}{\sum_{p \in M} A_p}, \quad \psi_f'(p) = \sqrt{\frac{A_p}{\sum_{p \in M} A_p}} \psi_f(p) \tag{16}$$

where $A_p$ denotes the area associated with the vertex $p$; we typically normalise the IKS using

$$I_t'(p) = \frac{1}{\sum_{p \in M}(A_p \cdot I_t(p))} I_t(p). \tag{17}$$

Fig. 3 compares the original IKS and the normalised IKS. After the normalisation, the IKS of corresponding points computed at the two different scales are almost identical, which means it achieves scale invariance.

To make the eigendecomposition of the Laplacian tractable, we only compute the 300 smallest eigenvalues and their corresponding eigenvectors in practice. The smallest eigenvalues correspond to the lowest frequencies, conveying information mostly about the global structure of the shape—the IKS is equivalent to a collection of low-pass filters. The trade-off is that we discard high-frequency information corresponding to surface details to efficiently compute the IKS. The emphasis on low frequencies could potentially damage the ability of the IKS to precisely localise features, although saliency detected using spectral irregularity has a nice localisation property. In this work, we simply weight the IKS using Gaussian curvature to help improve feature localisation. Note that, due to the multi-scale property of the IKS, for point of interest detection, we typically set $t$ to a small value causing the IKS to focus more on local details. Weighting by Gaussian curvature further strengthens the localisation property of the detector. To further balance local and global features, and make the process adjustable, we design a simple but fast scheme for point of interest selection.

### 3.2 point of interest selection

The IKS (which we henceforth assume to be weighted by Gaussian curvature) gives each mesh vertex a response informed by both global and local saliency information. To select points of interest based on such a saliency map, we find the *globally-salient points* and the *stable locally-salient points* in turn. A point on a mesh is a globally-salient point if its IKS is (i) maximal over its 2-ring neighbourhood and (ii) greater than $T_{global}I_M$ where $I_M$ is the global maximum of the IKS over all points and $T_{global}$ is a threshold factor. A point

is a stable locally-salient point if its IKS is (i) maximal over its 2-ring neighbourhood and (ii) greater than $T_{local}I_S$ where $I_S$ is the second-largest IKS in its neighbourhood and $T_{local}$ is a thresholding factor.

## 4 Experiments

Our experiments are based on the publicly available benchmark [6] (http://www.itl.nist.gov/iad/vug/sharp/benchmark/3DInterestPoint). In this benchmark, human subjects were asked to mark points of interest on models, giving human-generated ground truth data for use in measuring the performance of 3D point of interest detectors. It also provided results of tests on six algorithms: 3D-Harris [25], 3D-SIFT [9], HKS [26], mesh saliency [16], salient points [2] and SD-corners [21]. Previous evaluation methods usually measured the repeatability rate according to varying factors, such as model deformation, scale change, different modalities, noise, and topological change. Unlike them, [6] used false positive error (FPE), false negative error (FNE) and weighted miss error (WME) to directly measure the key quality of a 3D point of interest detection method: how similar are the points it detects of interest to those selected by human perception?

### 4.1 Evaluation methods

[6] denotes the set of points of interest detected by an algorithm on model $M$ as $\mathcal{A}$ and the set of ground truth points of interest as $\mathcal{G}(n,\sigma)$ where $n$ and $\sigma$ are two controlling parameter, to which the final set of ground truth points of interest are highly sensitive. In our experiments, we fix $n = 2$ and $\sigma = 0.03$. For a point of interest $g$ in set $\mathcal{G}$, a geodesic neighbourhood of radius $r$ is defined as $\mathcal{C}_r(g) = \{p \in M | d(g,p) \le r\}$ where $d(g,p)$ is the geodesic distance between points $g$ and $p$. The parameter $r$ controls the localisation error tolerance. A point $g$ is considered to be 'correctly detected' if a detected point $a \in \mathcal{A}$ exists in $\mathcal{C}_r(g)$, such that $a$ is not closer to any other points in $\mathcal{G}$. Denoting the number of correctly detected points in $\mathcal{G}$ as $N_C$, the FNE at localisation error tolerance $r$ is

$$FNE(r) = 1 - \frac{N_C}{N_G} \tag{18}$$

where $N_G$ is the number of points in $\mathcal{G}$.

Each correctly detected point $g \in \mathcal{G}$ corresponds to a unique $a$, the closest point to $g$ among the points in $\mathcal{A}$. All points in $\mathcal{A}$ without a correspondence in $\mathcal{G}$ are declared as false positives. Then, the FPE at localisation error tolerance $r$ is

$$FPE(r) = \frac{N_A - N_C}{N_A} \tag{19}$$

where $N_A$ is the number of detected points of interest.

To incorporate the saliency of a point of interest in the evaluation, [6] also proposed the WME. Assume that within a geodesic neighbourhood of radius $r$ around the ground truth point $g_i \in \mathcal{G}$, $n_i$ human subjects have marked an point of interest. Then the WME is defined as

$$WME(r) = 1 - \frac{1}{\sum_{i=1}^{N_G} n_i} \sum_{i=1}^{N_G} n_i \delta_i \tag{20}$$

where $\delta_i = 1$ if $g_i$ is detected and 0 otherwise.

An algorithm gets a low WME if it manages to detect a point that is frequently voted for by human subjects. Thus it measures the ability of an algorithm to detect the most salient points. In contrast, FNE and FPE treat all ground truth points of interest equally.

### 4.2 Experimental results and analysis

Fig. 4 visualizes over 8 models: the ground truth, the results of our proposed spectral irregularity diffusion method (SID), and the 6 competing methods from the benchmark. The models were selected to represent different classes of objects.

**Global saliency**: 3D-Harris, mesh saliency, salient points and SD-corners respond to regions where there is a strong local geometric response; i.e. they do not locate points of interest in flat regions. It is usually a challenge for bottom-up point of interest detectors to capture globally-salient points as they are mostly driven by local geometrical cues. However, the ground truth data show that human perception has the ability to capture points of global interest in flat regions. In this regard, 3D-SIFT, HKS and SID perform better since they not only capture points with a strong local geometric response, but also detect some points of high global saliency in flat regions. 3D-SIFT sometimes locates points of interest in geometrically insignificant regions since it works on a coarse voxel structure. SID works well as it captures saliency in the spectral domain, which considers the global structure of the object.

**Local saliency**: Although 3D-SIFT works well with respect to marking points of global interest, it does not localise the finer points of interest as well as the other algorithms (see the tentacle tips of the *Octopus*, the nose of the *Human* and the feet and the ear tip of the *Horse*). This undesirable behaviour of missing some locally salient points is caused by its coarse voxelisation scheme which nonetheless benefits its ability to detect globally-salient points. Typically, it is difficult for a detector to well capture both global and local saliency.
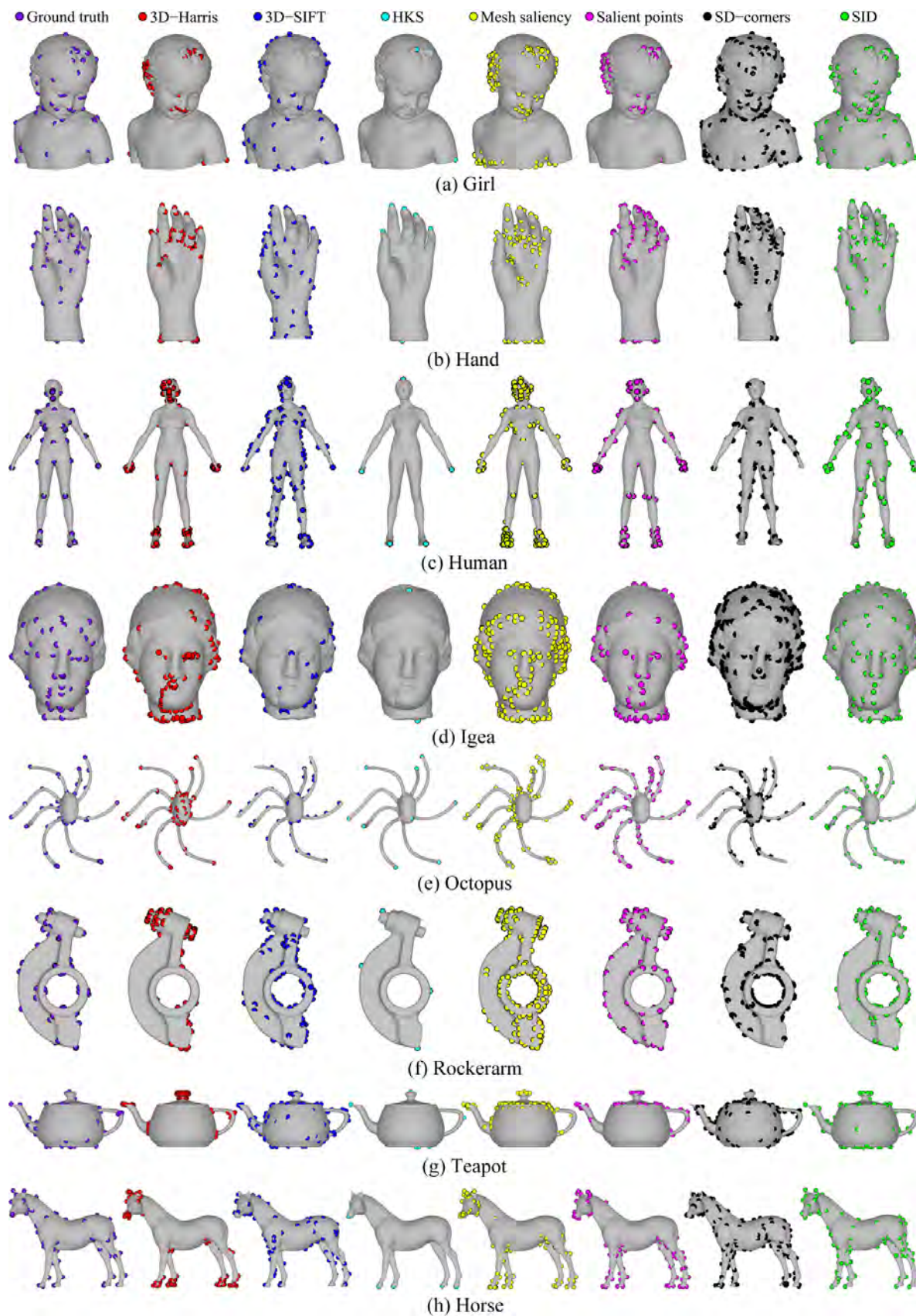
**Fig. 4** Ground truth obtained by setting $\sigma = 0.03$ and $n = 2$ (first column); points of interest detected by the algorithms: 3D-Harris [25] (second column), 3D-SIFT [9] (third column), HKS [26] (fourth column), Mesh saliency [16] (fifth column), Salient points [2] (sixth column), SD-corners [21] (seventh column) and the proposed SID method (eighth column).
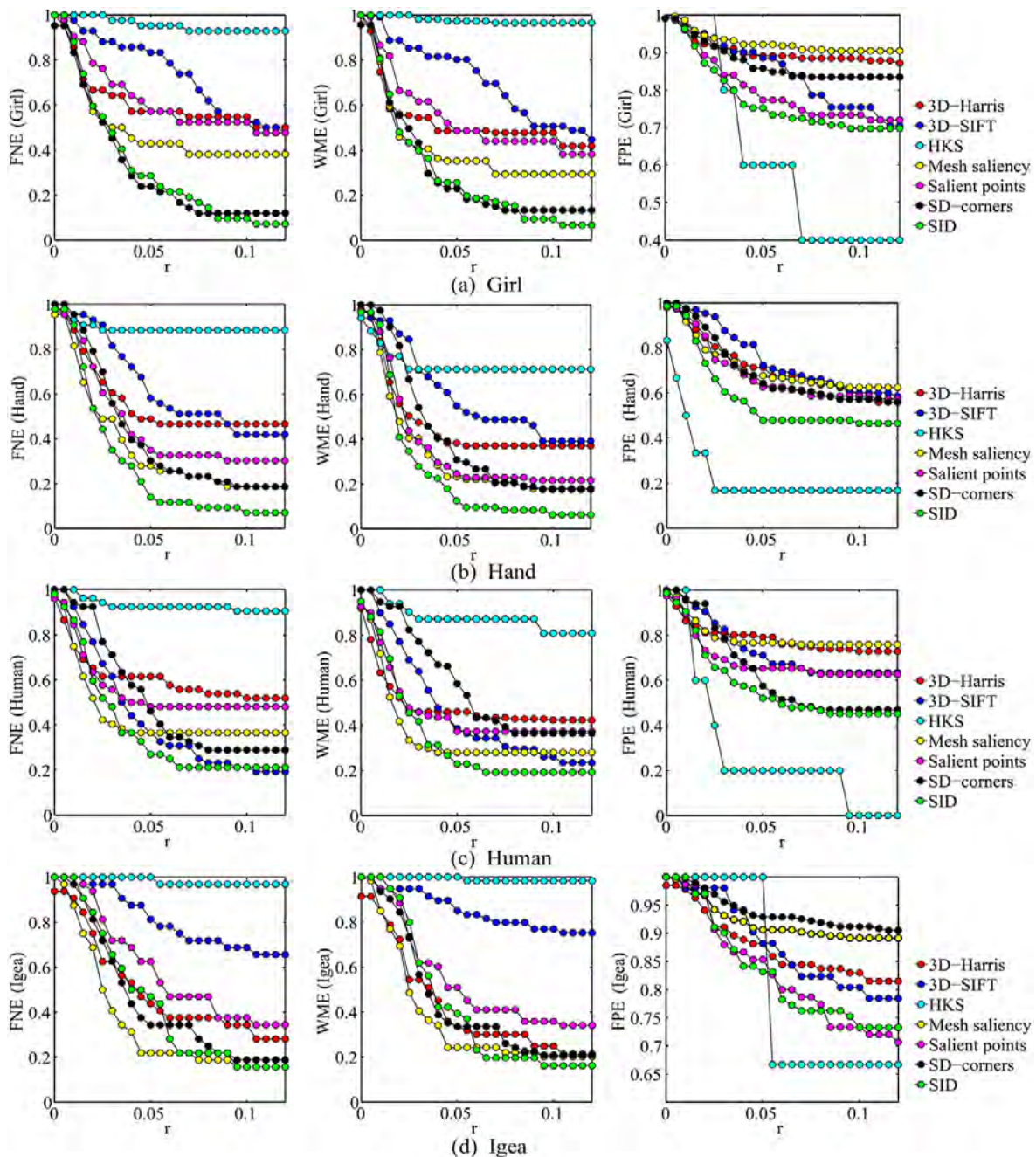
**Fig. 5** FNE, WME, and FPE graphs of various algorithms for the test models as rendered in Fig. 4.

SID performs significantly better than 3D-SIFT at localising such finer points of interest.

**Number of points of interest**: All methods, except for HKS which returns few points of interest, tended to mark more points of interest than human users. For the IKS used in SID, we set $t = 0.05$, the same setting for the implementation of HKS in the benchmark. It can be seen that mesh saliency and SD-corners pick up points of interest from almost all singularities while other methods including SID provide fewer points of interest.

Fig. 5 gives FNE, WME and FPE graphs with respect to localisation error tolerance $r$. As more points of interest are captured, more false positives are normally detected, although that usually corresponds to a lower FNE. Similarly, if an algorithm tends to mark fewer points of interest, this results in a lower FPE, at the cost of a higher FNE. An ideal method should keep both FNE and FPE low. Also, since the WME considers the semantic importance of each ground truth point, it is desirable that a method should have a low WME.
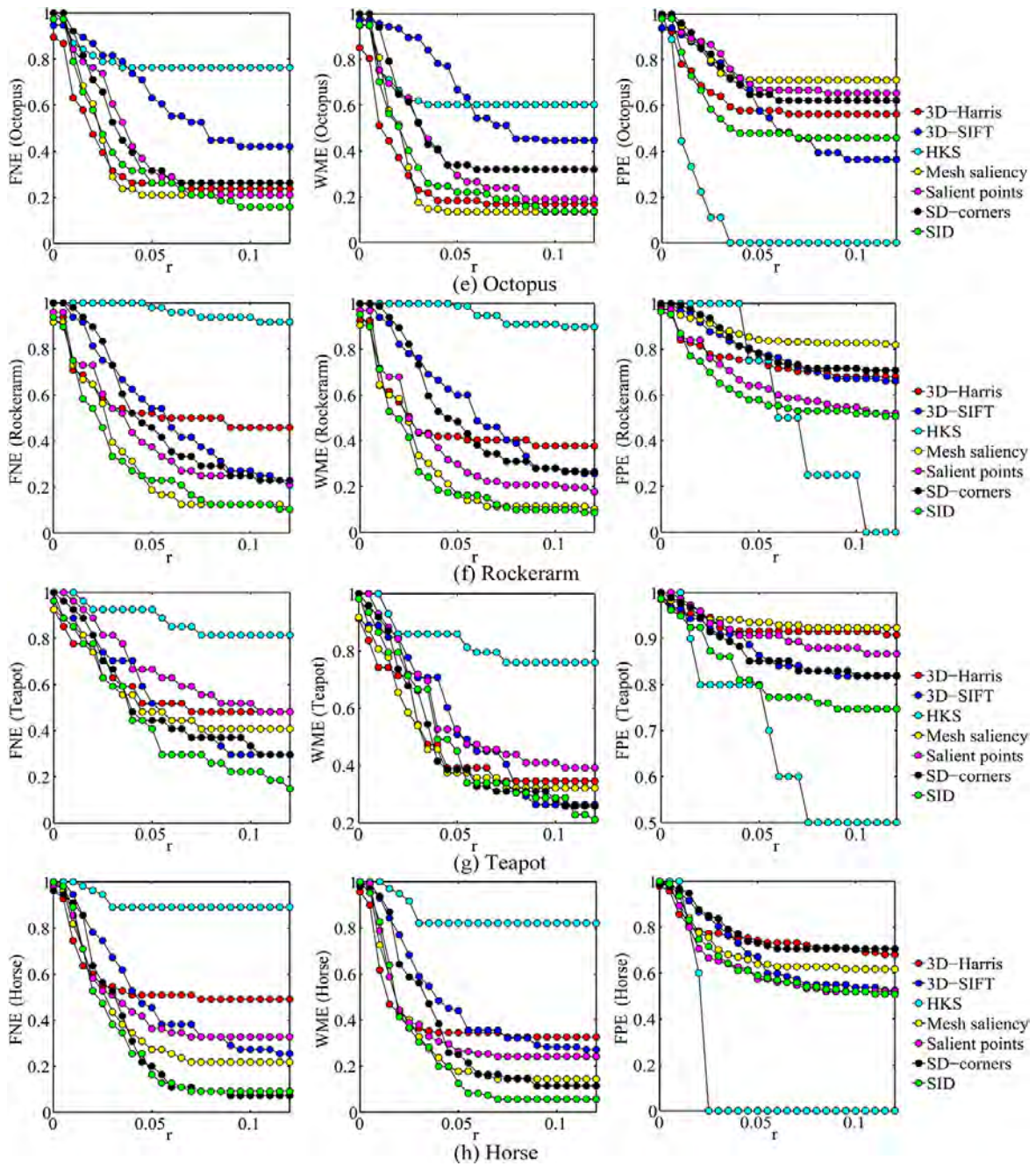
**Fig. 5** (Continued)

**FNE**: Mesh saliency, SD-corners and SID localise the points of interest well, as demonstrated by their low FNE. SID stably achieves the lowest FNE on almost all test models. In most graphs, SID has a similar FNE to other methods at the starting point, but its FNE drops very quickly with respect to $r$. A rapid drop in FNE means that the method finds the points of interest with a low localisation error. In most cases, the FNE curve for 3D-Harris drops slowly, which indicates that the detected points of interest are poorly localised. 3D-SIFT does not perform as well as other methods in terms of FNE, since the coarse voxel structure does not allow good localisation of points of interest.

**WME**: SID returns the lowest WME in most cases, and SID's WME curve is usually the one which drops fastest in each graph. 3D-SIFT does not perform well in quite a few cases. The good performance of SID as measured by WME means that it is less likely to miss the most salient points (i.e. points marked by many human subjects). This also demonstrates that the proposed saliency detection scheme is effective. Mesh saliency also performs well in most cases because it also has

a mechanism for considering saliency in its computational model.

**FPE**: SID achieves the lowest FPE in most cases. In sharp contrast, although mesh saliency and SD-corners perform well in terms of FNE and WME, they also have high FPE, as they have a tendency to return an excessive number of points of interest, including 'false' ones.

**HKS**: We single HKS out since it exhibits a completely different behaviour to the other 6 methods. It detects very few points on the models, most of which correspond to tips of extremities of the models (see Fig. 4). Despite having a low FPE, HKS gives the highest FNE among all other methods since it detects far fewer points than human subjects usually mark.

## 5 Conclusions

We have proposed a novel point of interest detector for 3D meshes. In our method, saliency is captured in terms of the irregularity of the Laplacian spectrum of the mesh Laplacian. Incorporating this irregularity into a diffusion model leads to the IKS which inherits the desirable properties of HKS. Diffusion allows saliency information represented as spectral irregularity in the spectral domain to be transferred to the spatial domain as a per-vertex property. Experiments demonstrate that detected salient points are largely of human interest.

We use Gaussian curvature to weight local geometry to help improve localisation of points of interest. It is possible that other advanced local feature descriptors could be of benefit. A remaining open question is how to balance global and local saliency—locally-salient points are not always globally salient. Further understanding is needed of the corresponding mechanisms within the human perceptual system.

## References

1. Bronstein, M., Kokkinos, I.: Scale-invariant heat kernel signatures for non-rigid shape recognition. In: Proc. CVPR, pp. 1704–1711 (2010)
2. Castellani, U., Cristani, M., Fantoni, S., Murino, V.: Sparse points matching by combining 3d mesh saliency with statistical descriptors. In: Proc. Eurographics, pp. 643–652 (2008)
3. Chen, X., Saparov, A., Pang, B., Funkhouser, T.: Schelling points on 3D surface meshes. In: Proc. SIG-GRAPH (2012)
4. Chung, M., Taylor, J.: Diffusion smoothing on brain surface via finite element method. In: IEEE International Symposium on Biomedical Imaging, pp. 432–435 (2004)
5. Dong, S., Bremer, P.T., Garland, M., Pascucci, V., Hart, J.C.: Spectral surface quadrangulation. In: Proc. SIG-GRAPH, pp. 1057–1066 (2006)
6. Dutagaci, H., Cheung, C., Godil, A.: Evaluation of 3d interest point detection techniques via human-generated ground truth. The Visual Computer **28**, 901–917 (2012)
7. Gebal, K., Bærentzen, J.A., Aanæs, H., Larsen, R.: Shape analysis using the auto diffusion function. Computer Graphics Forum **28**(5), 1405–1413 (2009)
8. Gelfand, N., Mitra, N., Guibas, L., Pottmann, H.: Robust global registration. In: Proc. SGP, pp. 197–206 (2005)
9. Godil, A., Wagan, A.: Salient local 3d features for 3d shape retrieval. In: Proc. SPIE (2011)
10. Guo, C., Zhang, L.: A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression. IEEE Transactions on Image Processing **19**(1), 185–198 (2010)
11. Hagen, L., Kahng, A.: New spectral methods for ratio cut partitioning and clustering. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems **11**(9), 1074–1085 (1992)
12. Hou, X., Zhang, L.: Saliency detection: A spectral residual approach. In: Proc. CVPR, pp. 1–8 (2007)
13. Hu, J., Hua, J.: Salient spectral geometric features for shape matching and retrieval. The Visual Computer **25**(5), 667–675 (2009)
14. Koch, C., Poggio, T.: Predicting the visual world: silence is golden. Nature Neuroscience **2**, 9–10 (1999)
15. Koenderink, J.: The structure of images. Biological cybernetics **50**(5), 363–370 (1984)
16. Lee, C., Varshney, A., Jacobs, D.: Mesh saliency. In: Proc. SIGGRAPH (2005)
17. Lévy, B., Zhang, H.R.: Spectral mesh processing. In: SIG-GRAPH Courses, pp. 8:1–8:312 (2010)
18. Li, J., Levine, M.D., An, X., Xu, X., He, H.: Visual saliency based on scale-space analysis in the frequency domain. IEEE Transactions on Pattern Analysis and Machine Intelligence **99**(PrePrints) (2012)
19. Liu, Y., Zha, H., Qin, H.: Shape topics: A compact representation and new algorithms for 3d partial shape retrieval. In: Proc. CVPR, vol. 2, pp. 2025–2032 (2006)
20. Mian, A., Bennamoun, M., Owens, R.: On the repeatability and quality of keypoints for local feature-based 3d object retrieval from cluttered scenes. International Journal of Computer Vision **89**(2), 348–361 (2010)
21. Novatnack, J., Nishino, K.: Scale-dependent 3d geometric features. In: Proc. ICCV, pp. 1–8 (2007)
22. Pinkall, U., Polthier, K.: Computing discrete minimal surfaces and their conjugates. Experimental mathematics **2**(1), 15–36 (1993)
23. Seo, S., Chung, M., Vorperian, H.: Heat kernel smoothing using laplace-beltrami eigenfunctions. In: Proc. MICCAI, pp. 505–512 (2010)
24. Shilane, P., Funkhouser, T.: Selecting distinctive 3d shape descriptors for similarity retrieval. In: Proc. SMI (2006)
25. Sipiran, I., Bustos, B.: Harris 3d: a robust extension of the harris operator for interest point detection on 3d meshes. The Visual Computer **27**(11), 963–976 (2011)
26. Sun, J., Ovsjanikov, M., Guibas, L.: A concise and provably informative multi-scale signature based on heat diffusion. In: Proc. SGP, pp. 1383–1392 (2009)
27. Vallet, B., Lévy, B.: Spectral geometry processing with manifold harmonics. In: Computer Graphics Forum, vol. 27, pp. 251–260 (2008)
28. Zaharescu, A., Boyer, E., Varanasi, K., Horaud, R.: Surface feature detection and description with applications to mesh matching. In: Proc. CVPR, pp. 373–380 (2009)
29. Zou, G., Hua, J., Dong, M., Qin, H.: Surface matching with salient keypoints in geodesic scale space. Comput. Animat. Virtual Worlds **19**(3-4), 399–410 (2008)