

Harris 3D: a robust extension of the Harris operator for interest point detection on 3D meshes

Ivan Sipiran · Benjamin Bustos

© Springer-Verlag 2011

Abstract With the increasing amount of 3D data and the ability of capture devices to produce low-cost multimedia data, the capability to select relevant information has become an interesting research field. In 3D objects, the aim is to detect a few salient structures which can be used, instead of the whole object, for applications like object registration, retrieval, and mesh simplification. In this paper, we present an interest points detector for 3D objects based on Harris operator, which has been used with good results in computer vision applications. We propose an adaptive technique to determine the neighborhood of a vertex, over which the Harris response on that vertex is calculated. Our method is robust to several transformations, which can be seen in the high repeatability values obtained using the SHREC feature detection and description benchmark. In addition, we show that Harris 3D outperforms the results obtained by recent effective techniques such as Heat Kernel Signatures.

Keywords 3D interest points detection · Local features · Harris operator

1 Introduction

Many applications have benefited from the wide diffusion of 3D models. Areas such as medicine, engineering, entertainment, and so on are increasingly relying in processes that

involve this kind of information. Coupled with this, an improved ability of capture devices has been observed, allowing to generate low-cost three-dimensional objects and make extensive use of them. In addition, tasks involving 3D models such as mesh analysis and processing are active research areas.

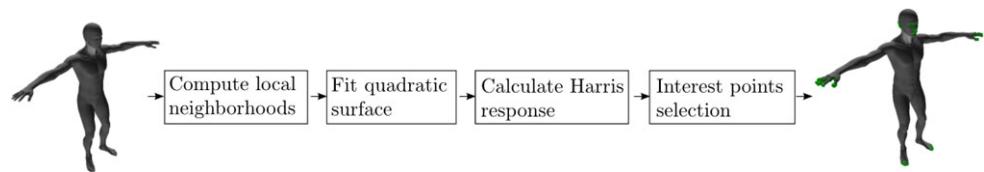
For the same reasons, there is a growing interest in the use of detailed models for better representation. This implies a higher necessity for feature detection tasks. As with images, the better is the resolution of a 3D object, the better the representation of some entity and therefore it is necessary to be able to select distinctive points on a 3D model in order to keep the efficiency in the processes applied on them. Some tasks that benefit from this capability are object registration [5], object retrieval and matching [9], mesh simplification, viewpoint selection [14], and mesh segmentation [11, 25], just to name a few.

The interest point detection on 3D data is a challenging problem for several reasons. First, there is no consensus about the definition of an interest point. A commonly used definition (that we use in this paper) relates the measure of interest with the level of protrusion of outstanding local structures. So, vertices on smooth or nearly planar sections of a surface will have low interest, as opposite to vertices in regions with uncommon local structure. For instance, in a human-shaped model, an interest point detector should select vertices on the face, hands, and feet. Second, the topology in 3D meshes is arbitrary. That is, a vertex can have an arbitrary number of neighboring vertices. This makes the tasks of selecting a local neighborhood around a vertex harder. In addition, this drawback causes that different tessellations can represent the same locality and therefore, an interest point detector should be able to deal with that. Third, without a well-defined topological structure for meshes, the extent of a locality in which a vertex is an interest point is

I. Sipiran (✉) · B. Bustos
PRISMA Research Group, Department of Computer Science,
University of Chile, Santiago, Chile
e-mail: isipiran@dcc.uchile.cl

B. Bustos
e-mail: bebustos@dcc.uchile.cl

Fig. 1 This figure shows the steps we propose to detect interest points in 3D meshes



unknown or difficult to compute. Finally, there is no additional information other than the position of vertices and the connectivity information among them. This fact complicates the process because the level of interest needs to be measured using the available information, which also depends of the topology of the mesh.

In a different scenario, interest point detectors for images have reached an acceptable effectiveness. The reason is that image structure is well-defined and interest points correspond with pixels that represent interesting structures in the scene captured in the image. In that sense, several methods have been proposed for detecting interest points taking into account different scales and transformations. As a result, the range of applications in computer vision that makes use of interest points detection has increased considerably in recent times such as image matching [17], image stitching [2], and human activity recognition [12, 13], just to name a few.

However, despite the success of these techniques in the image domain, trying to adapt them to 3D meshes is not a trivial task. Firstly, the adaptation is not direct because 3D meshes structure is very different from images. Secondly, the transformations required to be robust in 3D domains (isometry, topology, change of tessellations, downsampling, etc.) are also different.

In this paper, we present an effective and efficient extension of the Harris operator for 3D meshes. We chose the Harris operator for several reasons. First, the computation of the operator is an efficient and simple task. This is an important issue if we want use the interest point detection as a preliminary stage of subsequent process such as shape matching, registration or object recognition. Second, Harris-based methods have been effectively used in a number of applications and they have a high effectiveness as reported in the evaluation reports [19, 21]. Finally, an interesting evidence has been found recently, which greatly favors the Harris interest point detection method. Loog and Lauze [16] recently showed an important connection between the Harris operator and the computational visual attention model of visual perception. Roughly speaking, their model proved that interest points detected by the Harris method have low probability of appearing in other locations in the same image. These reasons encouraged us to investigate an effective extension of the Harris operator for 3D meshes, trying to comply with the robustness to the transformations in 3D domain.

Our method is outlined in Fig. 1. Given a 3D mesh, not necessarily manifold, the main process is performed in a

vertex-wise manner. The overall process consists of four steps. Firstly, our algorithm determines a local neighborhood around a vertex. The subsequent tasks are performed over this local neighborhood. Secondly, the neighborhood is processed so that it is prepared for a fitting step. We try to fit a quadratic surface to the set of points. This surface is a good representation of the locality and we consider it as an local image. Thirdly, we propose to calculate derivatives using a smoothing over the surface. We can use these derivatives for computing the Harris response for each vertex. Finally, our method selects the final set of interest points. A preliminary version of our method has been presented in a conference paper [23].

The contributions of this paper are summarized as follows:

- We improve the process for calculating the Harris operator for 3D meshes, making it robust to noise, change of tessellations and other transformations which deform the mesh structure such as local scaling, shot noise, presence of holes, just to name a few.
- We propose a novel method to define the neighborhood size of a vertex, depending on its surrounding structure.
- We give several options to select a few interest points using the information that the Harris operator provides.
- We present a comprehensive experimentation, trying to investigate the effect of different parameters choices and comparing our method with effective methods in the state of the art.

The organization of this paper is as follows. Section 2 presents the related works. Section 3 presents a detailed description of our method. Section 4 presents and discusses the experimental results. Finally, Sect. 5 concludes the paper.

2 Related work

The interest point detection topic emerged in the computer vision community with the aim of reducing the amount of information used in high-level vision tasks. A pioneering work was presented by Harris and Stephens [7], which was the basis for many later works. For readers interested in interest points detectors on images, we recommend the evaluation paper presented by Schmid et al. [21], which contains detailed descriptions and performance evaluation of several proposed methods.

For 3D meshes, several approaches have been proposed, most of which have tried to extend the detectors proposed for images. After the SIFT method proposed by Lowe [17], a number of extensions have been presented which use Difference-of-Gaussians (DoG) as interest point detector. Castellani et al. [3] applied the DoG detector over vertices in scale space obtained with successive decimations of the original shape. Vertices with high response in their DoG operator are selected as interest points. In the same way, Zou et al. [27] proposed to build a geodesic scale space, and subsequently to apply DoG detector on that space for detecting interest points on a surface. Also, Zaharescu et al. [26] assumed that the vertices of a 3D object have associated information such as curvature or photometric properties. Defining a discrete Difference-of-Gaussians operator, the authors applied this operator on the function defined by the associated information over a manifold. This approach showed good results in matching of 3D models sequences.

On the other hand, the geometric diffusion theory can be used for detecting interest points on surfaces. The diffusion process reveals interesting characteristics from the intrinsic geometry of a surface which can be exploited to detect outstanding structures. As a 3D surface property related to the diffusion process on a manifold, the Laplace–Beltrami operator has been also used to detect interest points. Hu and Hua [9] defined the geometric energy of a vertex as function of the eigenvalues and eigenvectors of the Laplace–Beltrami spectrum of a given object. Vertices where the energy is a maximum are considered as interest points. In addition, the energy provides the scale where the selected vertices are interesting. The selected interest points were used in a matching task with promising results. On the other hand, Sun et al. [24] defined the Heat Kernel Signature as a temporal domain restriction of the Heat Kernel on a manifold, which is related to the Laplace–Beltrami spectrum. In 3D meshes, each vertex has an associated signature. A vertex is selected as interest point, when for large time values, its signature has a maximum with respect to the neighbor vertices.

Similarly, Zou et al. [28] proposed to build a scale space of the surface geometry via the surface Ricci flow which satisfies a set of desired properties for a multi-scale representation. The authors applied the Ricci flow over a metric of the surface based on edge lengths. The scale space is represented as a matrix with curvature values calculated from the set of diffused metrics. Then, the Laplacian of a vertex is computed using the cotangent schema using the curvature as associated values. A vertex is considered as an interest point if it is an extreme of the Laplacian in the 1-ring neighborhood and neighboring scales.

Also, curvature-based methods have been proposed. Gelfand et al. [5] described an interest point detector based on a new descriptor called the integral volume descriptor. For each vertex, the amount of volume in the intersection

of a ball centered in the vertex and the 3D object describes an interesting local measure. The authors showed that this quantity is closely related to the curvature in the vertex. Vertices with uncommon integral values are selected as interest points. Also using curvature in vertices, Ho and Gibbins [8] suggested a measure called the curvedness measure in order to describe the geometric information in a vertex. The curvedness is calculated from the principal curvatures of a vertex. This measure can be calculated in different scales by selecting different neighborhood sizes which are used to fit quadratic patches over which curvatures are computed. Vertices with extremal values in its curvedness, with respect to neighboring points and scales, are selected as interest points.

Differently, Liu et al. [15] proposed a Monte Carlo strategy to select a random set of points on a surface with each point having the same probability to be chosen. These points were used in partial shape retrieval. The assumption behind this proposal is that the vertices of a shape are samples of the original surface and the tasks that use them can be affected by shape tessellations. Similarly, Shilane and Funkhouser [22] considered random points on a 3D surface, selecting only those points that contribute to improve the retrieval performance. With a training phase, it was possible to assign a predicted distinction value to each selected point in the 3D collection and thus, using that values to assign new ones to points of a new shape.

As another approach, the mesh saliency defined by Lee et al. [14] has proven to be a robust feature to many 3D applications. The process to compute the mesh saliency of a 3D object begins calculating a Gaussian-weighted average of the mean curvature on a surface. Each vertex in an object is thus associated with the difference of such average in different scales, which is the saliency of that vertex. Vertex with the highest saliency can be considered as interest points.

Conformal parameterization has also been used to propose interest points detectors. Methods based on conformal parameterization [10, 20] transform a 3D surface into a 2D parameterization that can be seen as an image. A 3D to 2D mapping is said to be conformal if angles are preserved. Once an image is computed, interest points can be detected on it, and subsequently these are mapped back to the 3D domain.

Finally, Mian et al. [18] related the repeatability of keypoints (extracted from partial views of an object) with a quality measure based upon principal curvatures.

3 Interest points detection

Harris and Stephens [7] proposed an interest points detector for images. Their method is a popular technique due to its strong invariance to rotation, scale, illumination variation,

and image noise [21]. The Harris detector is based on the local autocorrelation function of a signal, which measures the local changes of the signal with patches shifted by a small amount in different directions. The local autocorrelation is defined as:

$$e(x, y) = \sum_{x_i, y_i} W(x_i, y_i) [I(x_i + \Delta x, y_i + \Delta y) - I(x_i, y_i)]^2 \tag{1}$$

where $I(\cdot, \cdot)$ denotes the image function and (x_i, y_i) are the points in the Gaussian function W centered on (x, y) , which defines the neighborhood area in analysis.

Using a Taylor expansion truncated to the first order terms to approximate the shifted image, we obtain:

$$e(x, y) = \mathbf{S} \begin{bmatrix} \sum_{x_i, y_i} W \cdot I_x^2 & \sum_{x_i, y_i} W \cdot I_x \cdot I_y \\ \sum_{x_i, y_i} W \cdot I_x \cdot I_y & \sum_{x_i, y_i} W \cdot I_y^2 \end{bmatrix} \mathbf{S}^T \tag{2}$$

$$= \mathbf{S} \mathbf{E}(x, y) \mathbf{S}^T,$$

where $\mathbf{S} = [\Delta x \ \Delta y]$ is a shift vector, I_x and I_y denote the partial derivatives in x and y , and along with W are evaluated in (x_i, y_i) points.

Harris and Stephens proposed to analyze the eigenvalues of matrix E , which contains enough local information related to the neighborhood structure. In addition, to avoid the expensive eigenvalue calculation, they proposed to assign to each pixel in the image the following value:

$$h(x, y) = \det(E) - k(\text{tr}(E))^2 \tag{3}$$

with k constant.

The Harris operator has been used in many applications in image processing and computer vision by its simplicity and efficiency. However, the problem with 3D data is that the topology is arbitrary and it is not clear how to calculate the derivatives. To cope this problem, Glomb [6] suggested some approaches. We take this work as a basis for proposing a robust interest points detector on 3D meshes.

3.1 Robust Harris operator on 3D meshes

Given a vertex of a 3D object, we are interested in calculating the Harris operator value associated with that point. A 3D object is represented as a set of vertices V and a set of faces F with adjacency information between these entities. In addition, our method is not restricted to manifold meshes.

Let v be the analyzed vertex and $V_k(v)$ the neighborhood considering k rings around v . Figure 2 shows vertex v (black circle), the first ring around v (path formed by green circles), the second ring (path formed by blue circles), and the k th ring (path formed by yellow circles). All these vertices correspond to the neighborhood $V_k(v)$. The method to calculate k will be explained later in this section.

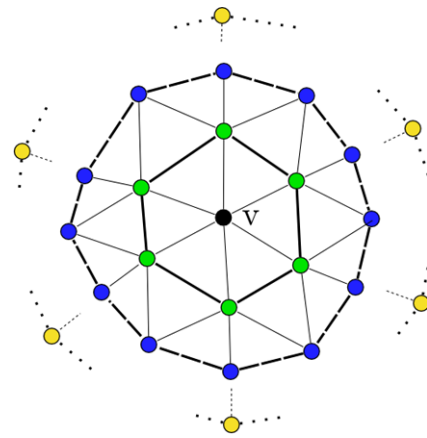


Fig. 2 Point v and its neighbor rings. Firstly, $V_1(v)$ is composed by vertices connected by *strong edges*. Secondly, $V_2(v)$ is composed by vertices up to those connected by *dashed edges*. Finally, $V_k(v)$ is composed by all vertices until those connected by *pointed edges*

We calculate the centroid of $V_k(v)$ and translate the set of points so the centroid is in the origin of the 3D coordinate system. Then, we compute the best fitting plane to the translated points. To do so, we apply Principal Component Analysis to the set of points and we choose the eigenvector with the lowest associated eigenvalue as the normal of the fitting plane. In our opinion, applying PCA is a better choice than the least square fitting because the assumption $z = f(x, y)$ does not have a good behavior when the data do not exhibit such functional characteristic.

The set of points is rotated so that the normal of the fitting plane is the z -axis. As we choose the less principal component as normal, the points exhibit a good spread in the XY -plane after rotation and therefore we can only work in XY -plane to calculate the derivatives. As final step before calculating derivatives, we translate the set of points so that the point v is in the origin of the XY -plane. This step will facilitate the further analysis.

To calculate derivatives, we fit a quadratic surface to the set of transformed points. Using least square method, we find a paraboloid of the form:

$$z = f(x, y) = \frac{p_1}{2}x^2 + p_2xy + \frac{p_3}{2}y^2 + p_4x + p_5y + p_6. \tag{4}$$

We chose a quadratic surface with only six terms because it represents a paraboloid. That is, it is the best choice if we need a function of two variables with quadratic terms. Adding more terms implies that it is possible to fit a more complex surface. However, more complex surfaces do not have well-defined derivatives in certain points of the domain. In addition, we needed a simple expression in order to apply the derivatives.

As we are interested in derivatives in the point v , one could directly evaluate the derivatives of $f(x, y)$ in the point

(0, 0), i.e.:

$$f_x = \left. \frac{\partial f(x, y)}{\partial x} \right|_{x=0}, \tag{5}$$

$$f_y = \left. \frac{\partial f(x, y)}{\partial y} \right|_{y=0}. \tag{6}$$

The above expressions should be a good estimate of derivatives. However, these can be influenced by noise. Instead, we propose to apply a Gaussian function as proposed originally by Harris and Stephens [7]. However, a difficulty arises because in the original expression the derivatives are discrete functions and our derivatives are continuous functions. To address this problem, we propose to apply the integration of the derivatives with a continuous Gaussian function as follows:

$$A = \frac{1}{\sqrt{2\pi}\sigma} \int_{R^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \cdot f_x(x, y)^2 dx dy, \tag{7}$$

$$B = \frac{1}{\sqrt{2\pi}\sigma} \int_{R^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \cdot f_y(x, y)^2 dx dy, \tag{8}$$

$$C = \frac{1}{\sqrt{2\pi}\sigma} \int_{R^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \cdot f_x(x, y) \cdot f_y(x, y) dx dy, \tag{9}$$

where σ is a constant, which defines the support of the Gaussian function and the factor $1/\sqrt{2\pi}\sigma$ is a normalization value.

Using calculus, we can reduce the expressions to

$$A = p_4^2 + 2p_1^2 + 2p_2^2, \tag{10}$$

$$B = p_5^2 + 2p_2^2 + 2p_3^2, \tag{11}$$

$$C = p_4p_5 + 2p_1p_2 + 2p_2p_3. \tag{12}$$

Finally, we can formulate the matrix E associated with the point v using the previously calculated values:

$$E = \begin{pmatrix} A & C \\ C & B \end{pmatrix}. \tag{13}$$

The Harris operator value in the point v is calculated as in (3).

3.2 Adaptive neighborhood size

Several approaches can be considered to select the number of rings around a point as neighborhood. If the object tessellation is uniform, i.e., almost all triangles in the object have the same size, we can use a constant number of rings to all points, or use the points contained in a ball of radius r and centered in point v . However, in irregular and complex meshes, these methods do not approximate a neighborhood adequately.

To tackle this problem, we propose an adaptive technique. Our method selects a different neighborhood size depending on the tessellation around a point. Let us consider an object as a graph $G(V', E')$, where $V' = V$ and E' is the set of edges obtained from the adjacency information of the object.

Given a point $v \in V'$, a k -ring around v is the set of points where the length of the shortest path to v is k :

$$ring_k(v) = \{w \in V' \mid |shortest_path(v, w)| = k\}. \tag{14}$$

The distance from a point v to the $ring_k(v)$ is defined as

$$d_{ring}(v, ring_k(v)) = \max_{w \in ring_k(v)} \|v - w\|_2. \tag{15}$$

Finally, we define the neighborhood size of a point v as

$$radius_v = \{k \in N, d_{ring}(v, ring_k(v)) \geq \delta \text{ and } d_{ring}(v, ring_{k-1}(v)) < \delta\}, \tag{16}$$

where δ is a fraction of the diagonal of the object bounding rectangle.

It is important to note that the proposed method always finds a neighborhood to a point, even with complex and irregular tessellations around that point.

3.3 Selecting interest points

With each vertex associated with its Harris operator value, we propose two ways to select the interest points of a given object. Firstly, we preserve the vertices which are local maximum. To do so, we select a vertex v which holds the following condition:

$$h(v) > h(w), \quad \forall w \in ring_1(v). \tag{17}$$

Secondly, we propose two approaches to select the final set of interest points.

- *Select the points with the highest Harris response.* We can pick a constant fraction of interest points depending on the application. In this proposal, we obtain the points with higher saliency and therefore, some portions of the object do not have interest points.
- *Representatives of Interest Points Clusters.* This approach can be used when we want a good distribution of interest points in the object surface. This proposal consists of two steps. First, we sort the pre-selected interest points according to their Harris operator value in decreasing order. Second, we apply Algorithm 1 to cluster the sorted points and select the final set of interest points.

The value of ρ can be considered as a fraction of the diagonal of the object bounding rectangle and it has effect in the number of returned interest points.

Figure 3 shows the result of the two options to select interest points.

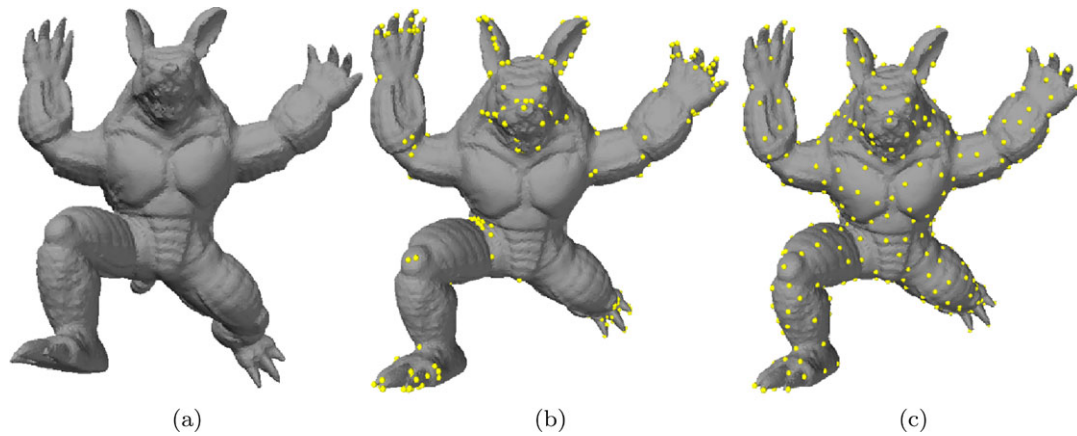


Fig. 3 Selection options. (a) Armadillo model. (b) Selected points with highest Harris response. (c) Selected points by clustering

Algorithm 1 Interest Points Clustering

Require: Set P of pre-selected interest points in decreasing order of Harris operator value

Ensure: Final set of interest points

```

1: Let  $Q$  be a set of points
2:  $Q \leftarrow \emptyset$ 
3: for  $i \leftarrow 1$  to  $|P|$  do
4:   if  $\min_{j \in [1, |Q|]} \|P_i - Q_j\|_2 > \rho$  then
5:      $Q \leftarrow Q \cup \{P_i\}$ 
6:   end if
7: end for
8: Return  $Q$ 

```

4 Experimental evaluation and discussion

In this section, we show the experimental results of our implementation of the Harris 3D method using a standard benchmark. The presentation of results is divided in two parts. First, we experiment with several aspects and parameters of our method. The objective is to investigate the effect of the parameters on the repeatability of interest points. Second, we compare our method with methods in the state of the art, namely the Heat Kernel Signatures proposed by Sun et al. [24] and the Salient Points detection method proposed by Castellani et al. [3].

4.1 The data set

For our experiments, we use the SHREC robust feature detection and description benchmark [1], which is available in the public domain. This data set is composed of triangular meshes with approximately 10,000–50,000 vertices.

The collection consists of three basic shapes (null shapes) from which a set of transformations have been applied. For each null shape, nine transformations were used: isome-

try (non-rigid transformation), topology, big holes and micro holes, local and global scaling, additive Gaussian noise, shot noise, and downsampling (less than 20% of the original points).

Each transformation is performed in five versions. In all transformations, except isometry and scale, each version represents a strength level, so high levels correspond with stronger transformations. The scales used in scaling transformation were 0.5, 0.875, 1.25, 1.625, and 2. For the isometry transformation, each version reflects a non-rigid transformation of the null shape. Therefore, each null shape has 45 transformed shapes and there are 138 shapes in the whole collection.

4.2 Evaluation methodology

An interest point detection method returns a set of detected points $F(Y) = \{y_k\}_k$ for each shape Y (typically, $|F(Y)| \ll |Y|$). The performance is measured by comparing the interest points computed for transformed shapes and the corresponding null shape.

The quality of the interest points detection was measured using the repeatability criterion. Assuming for each transformed shape Y in the data set the ground-truth dense correspondence to the null shape X to be given in the form of pairs of points $C_0(X, Y) = \{(y'_k, x_k)\}_{k=1}^{|Y|}$ (and same way, $C_0(Y, X)$), an interest point $y_k \in F(Y)$ is said to be repeatable if a geodesic ball of radius R around the corresponding point $x'_k : (x'_k, y_k) \in C_0(X, Y)$ contains an interest point $x_j \in F(X)$. The subset $F_r(Y) \subseteq F(Y)$ of repeatable interest points is given by

$$F_{R,X}(Y) = \{y_k \in F(Y) : F(X) \cap B_R(x'_k) \neq \emptyset, (x'_k, y_k) \in C_0(X, Y)\}, \quad (18)$$

where $B_R(x'_k) = \{x \in X : \text{geod}(x, x'_k) \leq R\}$ and geod denotes the geodesic distance function in X . The repeatability

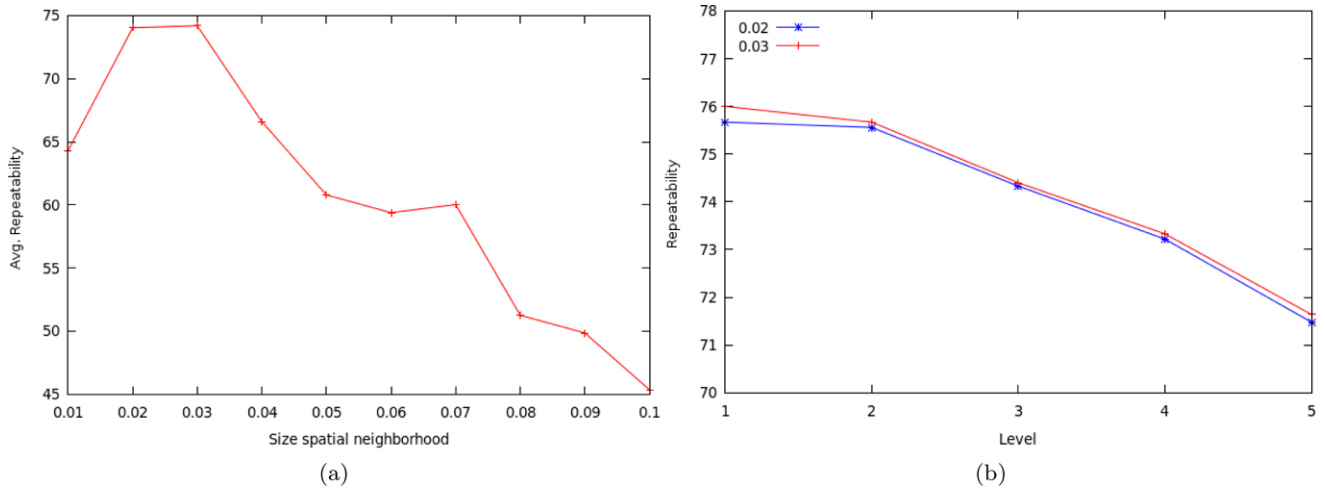


Fig. 4 (a) Repeatability values obtained by selecting spatial neighborhoods. The radii of the ball was chosen as a fraction of the diagonal of the bounding box of the object. Results show average repeatability for

the range [0.01, 0.1] for the radii. (b) Average repeatability for spatial neighborhoods with radii 0.02 and 0.03, respectively. The comparison was done by levels of transformation

$rep(Y, X)$ of $F(Y)$ in X is defined as the percentage of interest points from $F(Y)$ that are repeatable:

$$rep(Y, X) = \frac{|F_{R,X}(Y)|}{|F(Y)|} \tag{19}$$

For a transformed shape Y and the corresponding null shape X , the overall feature interest points detection quality was measured as $(rep(Y, X) + rep(X, Y))/2$. The value of $R = 5$ was used in the benchmark. This radius constitutes approximately 1% of the shapes diameter. Interest points without ground-truth correspondence (e.g. in regions in the null shape corresponding to holes in the transformed shape) were ignored.

4.3 Analysis of parameter values

In this section, we present the experimental results of our method. We experimented with several aspects and investigated the effect of the parameters on the repeatability of our proposal. Specifically, we evaluated the following aspects:

- *The type and size of local neighborhood.* We tested three options: spatial neighborhood, adaptive neighborhood and ring neighborhood. We are interested in evaluating the effectiveness of each option.
- *The parameter K .* We tested with different values for this parameter in order to investigate its effect in the calculation of Harris response.
- *The type of interest point selection method.* We only evaluated the method that selects the points with higher response. We intended to figure out the effect of the number of selected interest points in the effectiveness of our algorithm.

It is important to point out that for all experiments we used a basic configuration of parameters and only the analyzed parameter was changed. Our basic configuration consisted of adaptive local neighborhood with $\delta = 0.01$, $K = 0.04$, and selection of 1% of the number of vertices with higher Harris response as final interest points set. All graphs presented in our results were adequately scaled for a better visualization.

4.3.1 Local neighborhood

In order to assess the importance of the determination of the local neighborhood around an analyzed vertex, we performed three experiments: spatial, adaptive and ring neighborhoods.

Spatial neighborhoods Our first experiment consisted in evaluating the repeatability of our method when local neighborhoods were determined as the set of vertices lying inside of a ball centered in the analyzed vertex. We varied the radius of the ball and calculated the average repeatability for all transformations and for all levels. The radii were taken as fraction of the diagonal of the bounding box of the object. Figure 4a shows the results of this experiment.

We varied the fraction of the diagonal in the range [0.01, 0.1]. The highest values of repeatability were obtained for 0.02 and 0.03. In order to find the best value, we plot the repeatability curves for these values with respect to the level of transformation (see Fig. 4b). Both curves are very similar, with a slight advantage for $fraction = 0.03$.

There are two important issues that we should remark. First, the average repeatability decreases with large neighborhoods. This is because large neighborhoods cannot be

Fig. 5 Average repeatability for adaptive neighborhoods with several δ values

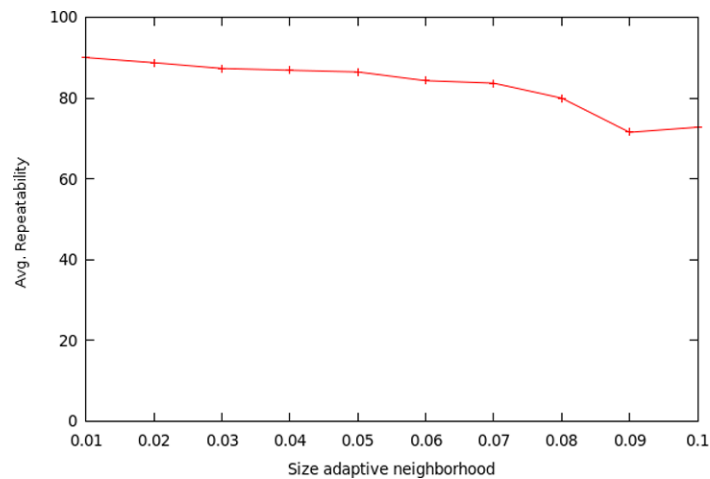


Table 1 Repeatability of our method using spatial neighborhood with $fraction = 0.03$. Average number of detected points: 303

Transform.	Strength				
	1	≤ 2	≤ 3	≤ 4	≤ 5
Isometry	75.77	79.89	78.01	79.45	79.85
Topology	76.25	76.39	76.21	76.19	76.18
Holes	75.56	75.52	75.36	75.34	74.98
Micro holes	76.19	76.17	76.10	75.98	75.92
Scale	82.58	80.01	77.38	74.79	72.29
Local scale	74.33	72.49	69.34	65.46	62.66
Sampling	73.21	71.06	69.35	66.75	59.51
Noise	73.81	73.14	71.78	69.94	67.48
Shot noise	76.32	76.36	76.04	76.09	75.90
Average	76.00	75.67	74.40	73.33	71.64

well fitted by a quadratic surface, and therefore the calculation of derivatives and the Harris response are not robust. Second, as it can be seen in Fig. 4b, repeatability always decreases with stronger levels of transformations (except in those where the level does not represent the strength of the transformation). Nevertheless, this behavior was expected.

Table 1 shows the repeatability for each transformation and each level for $fraction = 0.03$. Although there are transformations where, on average, the repeatability is greater than 75% (for example, isometry, topology, holes, micro holes, scale and shot noise), spatial neighborhoods have some problems. On the one hand, vertices belonging to a spatial neighborhood do not necessarily belong to the same local surface around the analyzed vertex. Moreover, the set of vertices could not belong to the same connected component. Having neighborhoods with vertices not belonging to the local surface, the fitting task is not robust, damaging the overall process and therefore the effectiveness. On the other hand, spatial neighborhoods do not ensure a good balance

Table 2 Repeatability of our method using adaptive neighborhoods with $\delta = 0.01$. Average number of detected points: 303

Transform.	Strength				
	1	≤ 2	≤ 3	≤ 4	≤ 5
Isometry	93.59	95.14	94.43	95.05	95.18
Topology	93.48	93.43	93.25	93.16	93.15
Holes	92.08	92.04	91.60	91.25	90.67
Micro holes	93.59	93.59	93.56	93.55	93.51
Scale	94.29	93.80	93.42	93.04	92.60
Local scale	93.49	92.89	91.27	88.81	86.55
Sampling	92.23	90.40	87.83	84.52	77.98
Noise	92.33	81.83	72.51	66.75	63.44
Shot noise	93.54	92.60	91.27	89.93	88.20
Average	93.18	91.75	89.90	88.45	86.81

around the analyzed vertex, which is a problem if the models have poor triangulations.

Adaptive neighborhoods Our second experiment consisted in varying the parameter δ with the adaptive local neighborhood approach. Figure 5 shows the average repeatability obtained in this experiment. Table 2 shows all values of repeatability obtained by the best configuration ($\delta = 0.01$). We tested values smaller than 0.01 for δ ; however, the improvement was not significant.

Similarly to the spatial neighborhoods, the average repeatability decreases as the extent of the neighborhoods increases. Large neighborhoods cause an inadequate fitting, affecting the effectiveness of the method. Results show an improvement of more than 15% in almost all entries with respect to those obtained by spatial neighborhoods. This percentage represents approximately 45 repeatable interest points detected in difference, compared to the spatial neighborhood effectiveness. However, although the results improve significantly, there is still visible a rapid fall in the re-

peatability values for noise transformation, even below those obtained by spatial neighborhoods.

From Table 2 we can note some interesting issues. On the one hand, from the transformations where the level represents strength, some of them present only a slight fall of average repeatability between the minimum and maximum levels: topology (0.33%), holes (1.41%) and micro holes (0.08%). On the other hand, from stronger level, the worst performance is obtained by sampling and noise transformations.

The behavior with the sampling transformation is expected because our method relies on selecting a fraction of the number of vertices as interest points. So the number of interest points detected in level 1 is much larger than in level 5. Clearly, the number of repeatable interest points in this transformation cannot be larger than the number of interest points in high levels. Therefore, the repeatability decreases considerably with downsampling. Also, with respect to the noise, the level of distortion of the meshes in stronger levels of this transformation is high, causing a considerable deformation in the shapes. Surely, local neighborhoods are affected considerably and the process of fitting is not robust. We argue that the decrement with respect to spatial neighborhoods is due to the use of geodesic distances when collecting the rings around the analyzed vertex. As the noise is applied in the direction of the normal of each vertex, the same parameter δ in different levels of transformations can determine neighborhoods of considerable different sizes. Obviously, for the reasons explained before, the repeatability of our method is affected.

Ring neighborhoods The third experiment evaluated the repeatability with different numbers of rings selected as neighborhood. Figure 6 shows the result for this experiment. Table 3 shows all repeatability values.

From Fig. 6 we can observe the same effect on the size of the neighborhood. With larger neighborhoods, average repeatability begins to decrease systematically. The interesting thing about this experiment is to note that a very small neighborhood gave the best results. One reason could be that, with small neighborhoods, the fitting step is robust and, thus, derivatives are well calculated. Nevertheless, small neighborhoods are unstable in presence of noise or another distortion transformation. Still, repeatability values for stronger levels of noise improve up to 13% with respect to spatial neighborhoods and up to 15% with respect to adaptive neighborhoods.

4.3.2 Parameter K

The parameter K is used in (3) to calculate the Harris response for a given vertex. This parameter needs to be tuned experimentally. We varied the parameter in the range

Table 3 Repeatability of our algorithm using one ring neighborhood. Average number of detected points: 303

Transform.	Strength				
	1	≤ 2	≤ 3	≤ 4	≤ 5
Isometry	95.86	96.74	96.51	96.78	96.79
Topology	95.92	95.92	95.84	95.75	95.70
Holes	93.60	93.77	93.74	93.83	93.48
Micro holes	95.86	95.86	95.88	95.89	95.93
Scale	96.76	96.54	96.24	95.93	95.39
Local scale	95.92	94.97	93.49	91.35	89.01
Sampling	94.69	93.55	92.19	89.05	80.26
Noise	92.97	92.07	90.82	89.71	88.54
Shot noise	95.99	95.57	94.90	93.62	92.41
Average	95.28	95.00	94.40	93.55	91.95

[0.04, 0.1] and tested the average repeatability. Figure 7 shows the results. In our implementation, the best result was for $K = 0.07$. However, the improvement was not significant with respect to our default value $K = 0.04$ (approximately 0.34%, representing almost 2 repeatable interest points of difference).

4.3.3 Interest point selection

In Sect. 3.3, we proposed two options for the final selection of interest points. The clustering approach is interesting from the point of view of applications requiring points well distributed over the whole surface. Nevertheless, as the process is based on grouping vertices with high responses, this step is not necessarily robust according to the repeatability criterion. Therefore, we did not consider this approach in our evaluation.

On the other hand, selection of vertices with higher response is interesting because the number of selected vertices is an important issue in applications. For example, in shape matching, we might be interested in only a few points, as the efficiency of matching is closely related to the number of points to be matched. So, we evaluated the effect of reducing the number of selected vertices.

Our method selects a fraction of the number of vertices as interest points, so the smaller is the fraction, the fewer interest points are selected. We varied the fraction in the range [0.001, 0.01]. Figure 8 shows our results.

Clearly, the average repeatability increases as the number of interest points is increased. This trend is maintained for values of fractions larger than 0.01. An important aspect to note is that by reducing the number of interest points in half, with respect to the value of fraction 0.01, the average repeatability decreases approximately in 6.34%. This is an advantage because if we need to apply subsequent processes,

Fig. 6 Average repeatability with ring neighborhoods sizes taken from the range [1, 10]

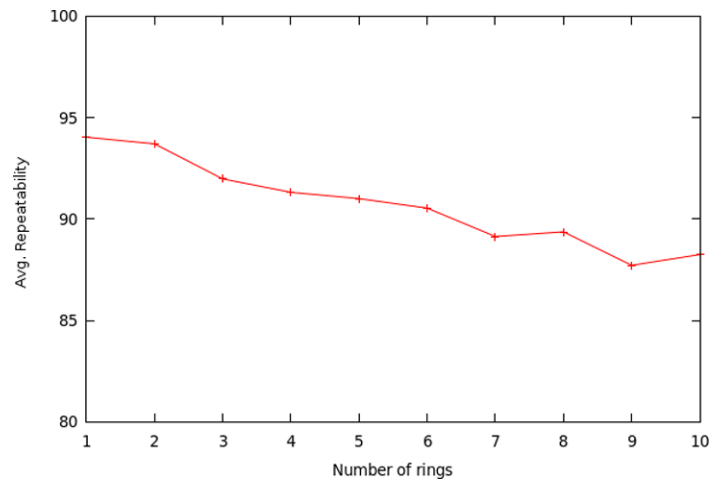


Fig. 7 Effect of varying K in average repeatability

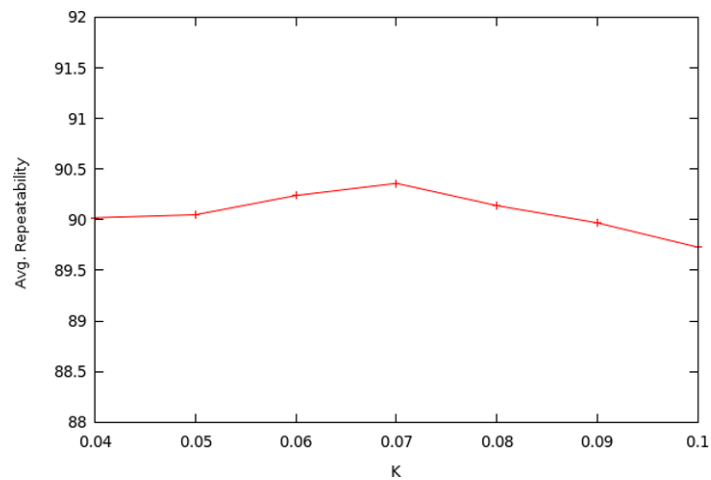
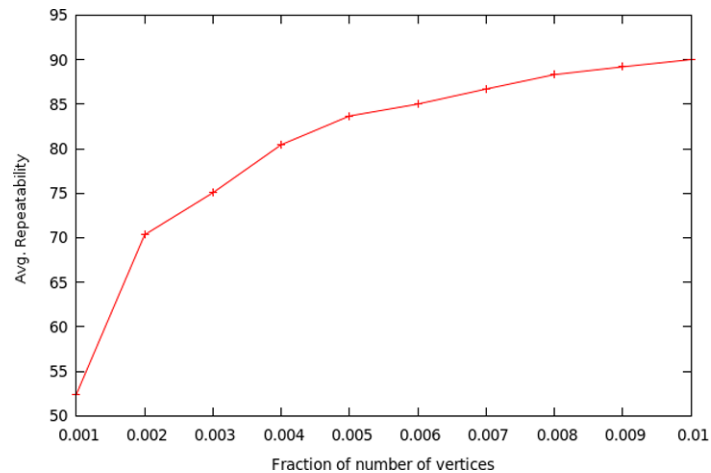


Fig. 8 Effect of reducing the number of interest points in average repeatability



we can opt for reducing the number of interest points selected, at expense of slightly reducing the robustness of the delivered points. The number of interest points will finally depend on the application and the trade-off between robustness and efficiency.

As an additional test, we combine the parameter values with the best effectiveness, obtaining a slight improvement. Table 4 shows the repeatability values for all transformations in all levels. Compared to Table 3, on average results improve for all levels.

Table 4 Repeatability of our algorithm using combination of values with the best effectiveness. Average number of detected points: 303

Transform.	Strength				
	1	≤2	≤3	≤4	≤5
Isometry	96.01	96.73	96.26	96.60	96.62
Topology	96.01	95.97	95.82	95.73	95.71
Holes	94.62	94.43	94.10	94.01	93.81
Micro holes	96.01	96.01	95.98	95.96	95.95
Scale	97.06	96.89	96.28	95.62	94.94
Local scale	96.24	94.96	93.40	91.26	88.84
Sampling	95.31	93.62	92.08	89.13	80.42
Noise	93.09	92.58	91.59	90.33	88.79
Shot noise	96.03	95.66	95.00	93.83	92.79
Average	95.60	95.21	94.50	93.61	91.99

4.4 Comparison with other methods

In order to compare our method with the state of the art, we selected two recent methods for detecting interest points on 3D meshes: Heat Kernel Signatures [24] and Salient Points [3]. Next, we specify the configuration used for these methods:

Heat kernel signature (HKS) As this method relies on the eigendecomposition of the Laplace–Beltrami operator on the mesh, we needed to simplify the meshes to approximately 10,000 vertices (we used Garland’s method [4]). Interest points were computed on the simplified meshes and mapped back to the original mesh by using the nearest neighbor vertex. We used the value $t = 0.1$ from the total area of the surface to evaluate the Heat Kernel Signature and 2-ring neighborhood in order to detect interest points. We present three variations:

- HKS1: No pre-processing step was used. We implemented this variant based on the original HKS implementation.¹
- HKS2: The Geomagic software was used for removing non-manifold edges, and faces were consistently oriented. Results for this variant were taken from the original report [1].
- HKS3: Filtering using persistent homology was used to discard unstable feature points. Results were also taken from [1].

Salient points (SP) We chose the best performance of this method from the SHREC feature detection benchmark.

Tables 5, 6 and 7 show the repeatability obtained by the variants of the Heat Kernel Signature method. Table 8 shows

Table 5 Repeatability of HKS1 feature detection algorithm. Average number of detected points: 35

Transform.	Strength				
	1	≤2	≤3	≤4	≤5
Isometry	83.20	87.39	88.44	87.29	87.21
Topology	79.73	81.27	81.49	81.08	80.94
Holes	80.15	77.57	75.86	73.35	71.16
Micro holes	81.02	80.76	80.68	80.43	80.50
Scale	79.99	79.78	79.90	80.18	80.36
Local scale	80.38	80.65	78.84	75.55	72.99
Sampling	82.70	82.23	82.66	82.04	78.99
Noise	75.80	74.55	72.37	69.34	68.23
Shot noise	79.96	81.14	81.15	80.07	78.77
Average	80.33	80.59	80.15	78.82	77.68

Table 6 Repeatability of HKS2 feature detection algorithm. Average number of detected points: 23

Transform.	Strength				
	1	≤2	≤3	≤4	≤5
Isometry	98.08	98.72	98.01	97.88	98.04
Topology	97.44	96.10	92.26	91.22	88.64
Holes	91.48	90.60	86.78	83.73	81.86
Micro holes	98.08	96.69	96.00	95.52	94.87
Scale	99.36	99.36	98.50	97.90	97.68
Local scale	98.08	94.83	90.09	83.05	78.31
Sampling	97.05	97.88	97.39	96.27	92.35
Noise	95.30	92.78	91.67	89.24	87.62
Shot noise	98.08	96.22	93.39	90.45	87.32
Average	96.99	95.91	93.79	91.70	89.63

the results of the Salient Points method. All comparison were done with respect to our best results shown in Table 4. Our method is represented by H3D.

As can be seen, our method outperformed the HKS1 variant without pre-processing and the Salient Points method. The benefit is consistent in each entry of the table, which is an important result regarding the relevance of the techniques compared. With respect to HKS2 and HKS3, these variants present significant improvements. However, they need well-formed shapes in order to work properly, which affects their efficiency considerably.

Table 9 presents the best method for each entry in the repeatability table. It can be seen that Harris 3D method outperforms the rest in stronger levels (4–5). The Heat Kernel based method is predominant in isometry, scale and sampling transformations. In the isometry transformation, the effectiveness of this method (HKS3 variant) is perfect with an average repeatability of 100%. A similar scenario takes

¹<http://www.geomtop.org/sunjian/software/hks.html>.

Table 7 Repeatability of HKS3 feature detection algorithm. Average number of detected points: 23

Transform.	Strength				
	1	≤ 2	≤ 3	≤ 4	≤ 5
Isometry	100.00	100.00	100.00	100.00	100.00
Topology	94.44	90.38	87.45	88.70	85.76
Holes	80.54	79.00	75.25	72.10	69.99
Micro holes	100.00	100.00	98.15	96.58	95.64
Scale	100.00	100.00	100.00	98.61	97.78
Local scale	97.44	96.79	93.02	87.25	82.90
Sampling	100.00	100.00	100.00	100.00	96.20
Noise	100.00	95.19	93.16	89.37	85.77
Shot noise	100.00	95.30	90.03	82.10	74.38
Average	96.94	95.19	93.01	90.52	87.60

Table 8 Repeatability of SP algorithm. Average number of detected points: 409

Transform.	Strength				
	1	≤ 2	≤ 3	≤ 4	≤ 5
Isometry	86.17	87.42	87.24	87.76	88.15
Topology	86.18	85.63	85.58	85.56	85.56
Holes	85.72	85.10	84.34	83.56	82.58
Micro holes	68.52	62.27	57.96	54.75	51.99
Scale	89.80	88.28	86.82	85.14	83.70
Local scale	85.73	84.97	84.48	83.33	82.12
Sampling	85.02	83.15	82.21	79.94	77.61
Noise	87.31	85.43	83.28	81.36	79.40
Shot noise	85.95	84.42	82.77	81.76	81.23
Average	84.49	82.96	81.63	80.35	79.15

place in the scale transformation, where for small scales (levels 1–2–3) the average repeatability is 100%. For the sampling transformation, the average repeatability is near to perfect too.

The reason for the good performance of the Heat Kernel based method in the aforementioned transformations is its intrinsic definition. This property allows it to appropriately define a characteristic shape based on the spectrum of the Laplace–Beltrami operator, and therefore it is robust against isometry and rigid transformations such as scaling. On the other hand, it is also robust to different samplings of the input mesh, as the interest points are selected for being points with large values of Heat Kernel Signatures in large times. That is to say, the interest points selection process is robust to different tessellations.

Differently, in transformations that deform the local structure of shapes, Harris 3D obtained the best results. Firstly, there is a total predominance of our method with

Table 9 Methods with the best performance by transformations and strengths. HKS—Heat Kernel Signature. H3D—Harris 3D

Transform.	Strength				
	1	≤ 2	≤ 3	≤ 4	≤ 5
Isometry	HKS3	HKS3	HKS3	HKS3	HKS3
Topology	HKS2	HKS2	H3D	H3D	H3D
Holes	H3D	H3D	H3D	H3D	H3D
Micro holes	HKS3	HKS3	HKS3	HKS3	H3D
Scale	HKS3	HKS3	HKS3	HKS3	HKS3
Local scale	HKS2	HKS3	H3D	H3D	H3D
Sampling	HKS3	HKS3	HKS3	HKS3	HKS3
Noise	HKS3	HKS3	HKS3	H3D	H3D
Shot noise	HKS3	HKS2	H3D	H3D	H3D
Average	HKS3	HKS3	HKS3	H3D	H3D

respect to the holes transformation. Secondly, our method outperforms the rest in most levels (3–5) for topology, local scale and shot noise transformations. Finally, it is interesting that our method is also the best in stronger levels for micro holes and noise transformations. Averages Table 9 represent the majority in each level, so we can observe the predominance of Harris 3D in levels 4 and 5.

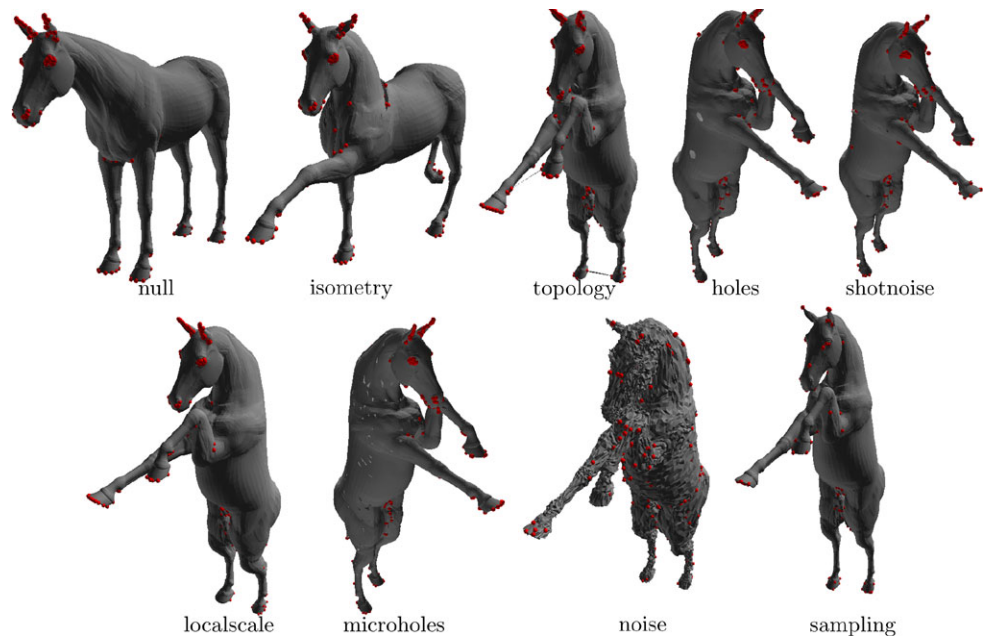
Figure 9 shows some examples of interest points over a class of shape of the SHREC feature detection benchmark using Harris 3D.

5 Conclusions

We presented a robust interest point detector for 3D meshes. This task is of importance due to the ability of reducing the amount of information needed in subsequent processes. Our algorithm effectively adapts the well-known Harris corner detection for images in order to be used for 3D meshes. Our proposal has showed to be effective, obtaining high repeatability values using the SHREC feature detection and description benchmark.

Our method is robust for several reasons. First, the use of a Gaussian function to smooth the derivatives surface contributes to effectively mitigate the effect of local deformations introduced by noise, holes, change of tessellations, and so on. Second, our proposal of adaptive neighborhoods improves considerably the alternative of spatial neighborhoods. In addition, from the results obtained by using ring neighborhoods, our experiments confirm the fact that balanced neighborhoods favor the overall process. This is because, with a balanced set of points, the approximation of derivatives in the analyzed point is better. Finally, along with the task of final selection of interest points, Harris 3D proposes several alternatives for its effective use, making an interesting alternative in applications such as shape matching and registration, just to name a few.

Fig. 9 Shapes with interest points. Interest points are represented with small *red balls*



Furthermore, our performed experiments suggest that our method could be used in extreme conditions with good results. This is an important advantage with respect to the state of the art, since it allows us to deal with several kinds of shapes and expand the spectrum of possible applications in the future.

Acknowledgements We thank Michael Bronstein for his extremely useful help with the SHREC feature detection and description benchmark. This project has been partially funded by CONICYT (Chile), through the Doctoral Scholarship, and FONDECYT (Chile) Project 1110111.

References

- Bronstein, A., Bronstein, M., Bustos, B., Castellani, U., Crisani, M., Falcidieno, B., Guibas, L.J., Kokkinos, I., Murino, V., Sipiran, I., Ovsjanikov, M., Patane, G., Spagnuolo, M., Sun, J.: SHREC 2010: Robust feature detection and description benchmark. In: Proc. Eurographics Workshop on 3D Object Retrieval, pp. 79–86. Eurographics Association, Aire-la-Ville (2010)
- Brown, M., Lowe, D.G.: Automatic panoramic image stitching using invariant features. *Int. J. Comput. Vis.* **74**(1), 59–73 (2007)
- Castellani, U., Cristani, M., Fantoni, S., Murino, V.: Sparse points matching by combining 3D mesh saliency with statistical descriptors. *Comput. Graph. Forum* **27**(2), 643–652 (2008)
- Garland, M., Heckbert, P.S.: Surface simplification using quadric error metrics. In: Proc. Int. Conf. and Exhib. on Comput. Graph. and Interact. Tech. SIGGRAPH '97, pp. 209–216. ACM Press/Addison-Wesley, New York (1997)
- Gelfand, N., Mitra, N.J., Guibas, L.J., Pottmann, H.: Robust global registration. In: Proc. Eurographics Symposium on Geometry Processing, p. 197. Eurographics Association, Aire-la-Ville (2005)
- Glomb, P.: Detection of interest points on 3D data: Extending the Harris operator. In: Computer Recognition Systems 3. Advances in Soft Computing, vol. 57, pp. 103–111. Springer, Berlin (2009)
- Harris, C., Stephens, M.: A combined corner and edge detection. In: Proc. of The Fourth Alvey Vision Conference, pp. 147–151 (1988)
- Ho, H., Gibbins, D.: Curvature-based approach for multi-scale feature extraction from 3D meshes and unstructured point clouds. *IET Comput. Vis.* **3**(4), 201 (2009)
- Hu, J., Hua, J.: Salient spectral geometric features for shape matching and retrieval. *Vis. Comput.* **25**(5–7), 667–675 (2009)
- Hua, J., Lai, Z., Dong, M., Gu, X., Qin, H.: Geodesic distance-weighted shape vector image diffusion. *IEEE Trans. Vis. Comput. Graph.* **14**(6), 1643–1650 (2008)
- Katz, S., Leifman, G., Tal, A.: Mesh segmentation using feature point and core extraction. *Vis. Comput.* **21**(8), 649–658 (2005)
- Laptev, I.: On space-time interest points. *Int. J. Comput. Vis.* **64**(2–3), 107–123 (2005)
- Laptev, I., Pérez, P.: Retrieving actions in movies. In: Int. Conf. in Comput. Vis, pp. 1–8 (2007)
- Lee, C.H., Varshney, A., Jacobs, D.W.: Mesh saliency. In: Proc. Int. Conf. and Exhib. on Comput. Graph. and Interact. Tech. SIGGRAPH '05, pp. 659–666. ACM, New York (2005)
- Liu, Y., Zha, H., Qin, H.: Shape topics: A compact representation and new algorithms for 3D partial shape retrieval. In: Proc. IEEE Conf. on Comput. Vis. and Pattern Recognit. CVPR '06, pp. 2025–2032. IEEE Computer Society, Washington (2006)
- Loog, M., Lauze, F.: The improbability of Harris interest points. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(6), 1141–1147 (2010)
- Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **60**(2), 91–110 (2004)
- Mian, A., Bennamoun, M., Owens, R.: On the repeatability and quality of keypoints for local feature-based 3D object retrieval from cluttered scenes. *Int. J. Comput. Vis., Special Issue on 3D Object Retrieval* (2009)
- Mikolajczyk, K.: Scale & affine invariant interest point detectors. *Int. J. Comput. Vis.* **60**(1), 63–86 (2004)
- Novatnack, J., Nishino, K.: Scale-dependent 3D geometric features. In: Proc. Int. Conf. on Comput. Vis., pp. 1–8. IEEE, New York (2007)
- Schmid, C., Mohr, R., Bauckhage, C.: Evaluation of interest point detectors. *Int. J. Comput. Vis.* **37**(2), 151–172 (2000)

22. Shilane, P., Funkhouser, T.: Selecting distinctive 3D shape descriptors for similarity retrieval. In: Proc. IEEE Int. Conf. on Shape Model. and Appl. SMI '06, Washington, DC, USA, p. 18 (2006)
23. Sipiran, I., Bustos, B.: A robust 3D interest points detector based on Harris operator. In: Proc. Eurographics Workshop on 3D Object Retrieval, pp. 7–14. Eurographics Association, Aire-la-Ville (2010)
24. Sun, J., Ovsjanikov, M., Guibas, L.J.: A concise and provably informative multi-scale signature based on heat diffusion. *Comput. Graph. Forum* **28**(5), 1383–1392 (2009)
25. Tierny, J., Vandeborre, J.P., Daoudi, M.: Enhancing 3D mesh topological skeletons with discrete contour constrictions. *Vis. Comput.* **24**(3), 155–172 (2008)
26. Zaharescu, A., Boyer, E., Varanasi, K., Horaud, R.P.: Surface feature detection and description with applications to mesh matching. In: Proc. IEEE Conf. on Comput. Vis. and Pattern Recognit. CVPR '09, Miami Beach, Florida (2009)
27. 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)
28. Zou, G., Hua, J., Lai, Z., Gu, X., Dong, M.: Intrinsic geometric scale space by shape diffusion. *IEEE Trans. Vis. Comput. Graph.* **15**(6), 1193–1200 (2009)



Ivan Sipiran is a Ph.D. candidate at the Department of Computer Science, University of Chile. He is a research assistant of the PRISMA Research Group. His research interests include 3D object retrieval, geometry processing and computer vision.



Benjamin Bustos is an Assistant Professor at the Department of Computer Science, University of Chile. He is head of the PRISMA Research Group. He leads research projects in the domains of multimedia retrieval, multimedia databases, video copy detection, sketch-based image retrieval, and image processing. His research interests include similarity search, multimedia information retrieval, 3D object retrieval, (non)-metric indexing, and pattern recognition. Benjamin Bustos obtained doctoral degree in natural sciences from the University of Konstanz, Germany, in 2006.