Eurographics Workshop on 3D Object Retrieval (2010)
I. Pratikakis, M. Spagnuolo, T. Theoharis, and R. Veltkamp (Editors)

A Robust 3D Interest Points Detector Based on Harris Operator

Ivan Sipiran[1] and Benjamin Bustos[1]

[1]PRISMA Research Group
Department of Computer Science, University of Chile

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 affine transformations(partially for object rotation) and distortion transformation such as noise addition. Moreover, the distribution of interest points on the surface of an object remains similar in transformed objects, which is a desirable behavior in applications such as shape matching and object registration.

1. Introduction

Many applications have benefited with 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, such is the impact it has caused that almost all processes involving 3D models are active research areas.

For the same reasons, many models have considerably large sizes. 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 efficiency in the processes applied on them. Some tasks that benefit from this capability are object registration [GMGP05], object retrieval and matching [HH09], mesh simplification, viewpoint selection [LVJ05], and mesh segmentation [TVD08, KLT05], just to name a few.

An interest points detection method for 3D objects must have some desirable properties. For example:

- It must be invariant to affine transformations.
- It must be robust to noise, which can be introduced during the capture process.
- It must be robust to different tessellations.

In this paper, we present an efficient interest points detector based on Harris operator defined for images (Figure 1). Our method holds the requirements previously described. 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 and tessellations.
- We propose a novel method to define the neighborhood size of a vertex, depending of its surrounding structure.
- We give several options to select a few interest points using the information that the Harris operator provides.

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, Section 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 [HS88], which

Figure 1: *Examples of interest points detected with our method.*

was the basis for many later works. For readers interested on interest points detectors on images, we recommend the evaluation paper presented by Schmid et al. [SMB00], 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 [Low04], a number of extensions have been presented which use Difference-of-Gaussians(DoG) as interest point detector. Castellani et al. [CCFM08] applied the DoG detector over vertices in scale-space obtained with successive decimations of the original shape. Vertices with high response in its DoG operator are selected as interest points. In the same way, Zou et al. [ZHDQ08] 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. [ZBVH09] assumed that the vertices of an 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.

As a 3D surface property, the Laplace-Beltrami operator has been also used to detect interest points. Hu and Hua [HH09] 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. [SOG09] 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.

Differently, Liu et al. [LZQ06] 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 [SF06] 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 [LVJ05] 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 to 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.

On the other hand, Mian et al. [MBO09] 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 [HS88] 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 [SMB00]. The Harris detector is based on the local auto-correlation function of a signal, which measures the local changes of the signal with patches shifted by a small amount in different directions. The local auto-correlation is defined as:

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

where $I(.,.)$ 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:

$$e(x,y) = [\triangle x \triangle y] \begin{bmatrix} \sum_{x_i,y_i} W.I_x^2 & \sum_{x_i,y_i} W.I_x.I_y \\ \sum_{x_i,y_i} W.I_x.I_y & \sum_{x_i,y_i} W.I_y^2 \end{bmatrix} [\triangle x \triangle y]^T$$
$$= [\triangle x \triangle y] E(x,y) [\triangle x \triangle y]^T \quad (2)$$

where I_x and I_y denotes 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:

$$h(x,y) = det(E) - k.(tr(E))^2 \quad (3)$$

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 [Glo09] 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 to 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 (green circles), the second ring (blue circles), and k-th ring (yellow circles). All these points correspond to the neighborhood $V_k(v)$. The method to calculate k will be explained later in this section.

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. We think that applying PCA is a better choice than 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.

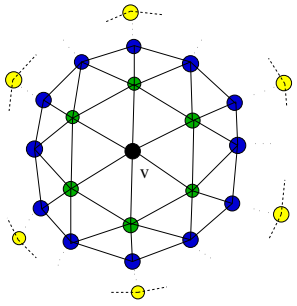

Figure 2: *Point v and its neighbor rings. Firstly, $V_1(v)$ is composed by green vertices. Secondly, $V_2(v)$ is composed by blue and green vertices. Finally, $V_k(v)$ is composed by all vertices until the yellow vertices.*

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:

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

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.:

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

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

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 [HS88]. 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:

$$A = \frac{1}{\sqrt{2\pi}\sigma} \int_{R^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}} . f_x(x,y)^2 dxdy \quad (7)$$

$$B = \frac{1}{\sqrt{2\pi}\sigma}\int_{R^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}}.f_y(x,y)^2 dxdy \quad (8)$$

$$C = \frac{1}{\sqrt{2\pi}\sigma}\int_{R^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}}.f_x(x,y).f_y(x,y) dxdy \quad (9)$$

where σ is a parameter, which defines the support of the gaussian function. The setting of σ will be considered later in this section.

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

$$E = \begin{pmatrix} A & C \\ C & B \end{pmatrix} \quad (10)$$

The Harris operator value in the point v is calculated as in Equation 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, 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 of 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:

$$ring_k(v) = \{w \in V' || shortest\_path(v,w)| = k\} \quad (11)$$

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

$$d_{ring}(v, ring_k(v)) = max_{w \in ring_k(v)} \|v - w\|_2 \quad (12)$$

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

$$radius_v = \{k \in N, d_{ring}(v, ring_k(v)) \geq \delta \text{ and } d_{ring}(v, ring_{k-1}(v)) < \delta\} \quad (13)$$

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

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

In addition, as we provide an approximate extent to each neighborhood, we can use this information to consistently apply the Gaussian function when calculating the Harris operator value. The extent of the Gaussian is controlled by the parameter σ, which we define as:

$$\sigma_v = \frac{\delta}{radius_v} \quad (14)$$

Therefore, each point have a different support for the applied Gaussian window when calculating its operator value and it is consistent with the neighborhood size as well.

3.3. Selecting Interest Points

With each vertex associated to 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:

$$h(v) > h(w), \forall w \in ring_1(v) \quad (15)$$

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 of the application. In this proposal, we obtain the points with higher saliency and therefore, some portions of the object does 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 step. First, we sort the pre-selected interest points according to its Harris operator value in decreasing order. Second, we apply the following algorithm to cluster the sorted points and select the final set of interest points.

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

```
Let Q be a set of points
Q ← ∅
for i ← 1 to |P| do
    if min_{j∈[1,|Q|]} ||P_i − Q_j||_2 > ρ then
        Q ← Q ∪ {P_i}
    end if
end for
Return Q
```

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

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

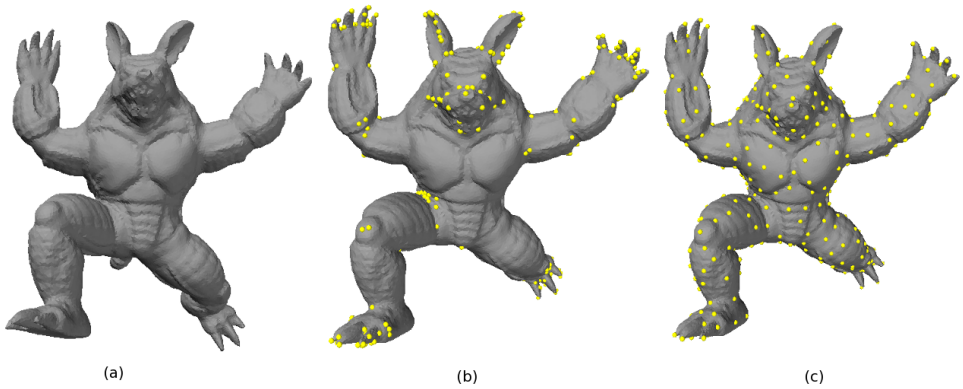

Figure 3: *Selection options. (a) Object from the collection. (b) Selected points with highest Harris response. (c) Selected points by clustering.*

4. Experimental Evaluation and Discussion

To test the affine transformation invariance and the robustness to noise, we made several experiments. The criterion to evaluate the experiments was the repeatability of the interest points. Given an object O and a transformation function T, which can be a translation, scaling, rotation, or noise addition, $T(O)$ is a transformed object. In addition, P_O represents the set of interest points extracted from O. Thus, the repeatability is defined as:

$$R_{O,T(O)} = \frac{|P_O \cap P_{T(O)}|}{|P_O|} \tag{16}$$

For all described experiments, we used the values $k = 0.04$ and $\delta = 0.025$ in our method. In addition, we used the highest Harris response points as interest points, where we select the 1% from the size of set V in each object. With respect to the used data, we pick 20 models from different collections available on the Web. Figure 4 shows some used models. Table 1 shows the detailed list of objects used in our experiments.

Figure 4: *Some objects used in our experiments.*

To evaluate the affine transformation invariance, each object was rotated with 10 randomly chosen angles, arbitrarily in the three coordinate axes. Thus, we calculated the repeatability between the original model and each transformed object, obtaining an average for each object in our collection. Finally, we calculate the mean of average repeatabilities of each object. For rotation, we obtained a repeatability of 0.8745.

Collection	Object / File
SHREC 09 - Partial Shape Retrieval	D00018.off D00048.off D00096.off D00245.off D00290.off D00341.off D00364.off D00405.off D00482.off D00597.off D00621.off D00644.off D00708.off D00724.off D00745.off D00772.off
Mesh Segmentation Benchmark [CGF09]	281.off 284.off 285.off
Stanford 3D Scanning Repository	asian dragon

Table 1: *Objects used in our experiments.*

We did not get a total rotation invariance because the process to calculate the Harris operator relies on a good quadratic surface fitting. In addition, the surface fitting relies on the distribution of points in the XY-plane and the problem is that PCA allows us to approximate the normal of a best fitting plane to the points, however the direction

of the normal of that plane is arbitrary. Therefore, different quadratic surfaces could be fitted to the same point in different orientations, affecting the Harris response computation.

In the same fashion, to test scale invariance, we proceeded to select random scales in the interval $[0.5, 2.0]$ and following the same described process above, we obtain a repeatability of 1.0. This good result is obtained because the neighborhood sizes are relatives to the object size, without affecting the subsequent processes.

In the case of noise addition, we added up a random offset to each vertex of an object in arbitrary directions. We considered different offset values to evaluate the effect of the noise amount in the interest points detection, with values in the interval $[10^{-4}, 10^{-3}]$. As in affine transformation experiments, we following the same methodology to calculate the repeatability in each offset. Figure 5 shows an example of interest points detection in noisy objects and Figure 6 shows the effect of offset amount in the repeatability.

This experiment also allowed us to test the proposed adaptive neighborhood estimation. As we selected arbitrary directions when noise was added up, the local tessellations around a vertex changed considerably. In this case, our proposal estimated good neighborhoods to mitigate the noise.

As we expected, the noise affects the Harris operator calculation. However, although the repeatability decreases when more noise is added up, the distribution of detected interest points on the surface remains, as shown in Figure 5. This is a important issue because, even in presence of noise, the interest points can be used trustly in tasks such as object registration and shape matching.

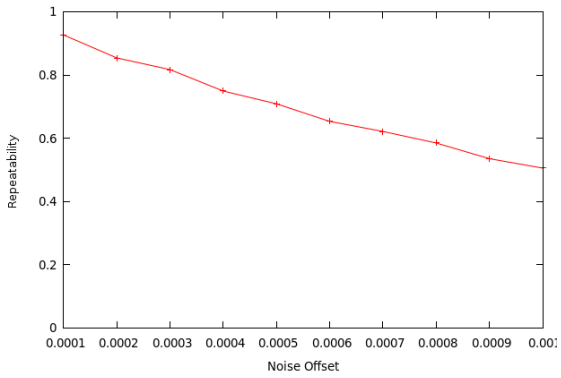

Figure 6: *Effect of noise in repeatability.*

In addition, we present some results applied to models with different level of detail in Figure 7. To simplify the models, we use the method proposed by Garland [GH97]. Obviously, the fewer are the vertices in a model, fewer interest points are detected. However, the distribution of interest points on the model surface is surprisingly similar.

5. Conclusions and Future Work

We have developed a robust interest points detector for three-dimensional objects based on Harris operator. We have shown how incorporating a Gaussian function when calculating the derivatives involved in Harris response, in addition to adaptive neighborhood determination, can enhance the results of detection and make the process robust against noise and tessellation variations in the neighborhood of a vertex. Also, we have shown in this paper that the proposed method can keep the distribution of interest points as visually similar in presence of distortions such as noise addition. This is an important issue in applications such as object registration and shape matching. In addition, we proposed two options to select interest points from the Harris response in each vertex, which can used depending of applications requirements.

In the future, we would like to extend the method to support levels of detail in the objects representations. Clearly, we do not claim robustness against the resolution of an object. It should be an important direction for further research due to its applicability in shape matching. Also, we plan to research the potential applications of our method in tasks such as partial shape matching, global and partial object registration, mesh simplification, and mesh segmentation.

References

[CCFM08] CASTELLANI U., CRISTANI M., FANTONI S., MURINO V.: Sparse points matching by combining 3D mesh saliency with statistical descriptors. *Computer Graphics Forum 27*, 2 (2008), 643–652.

[CGF09] CHEN X., GOLOVINSKIY A., , FUNKHOUSER T.: A benchmark for 3D mesh segmentation. *ACM Transactions on Graphics 28*, 3 (2009).

[GH97] GARLAND M., HECKBERT P. S.: Surface simplification using quadric error metrics. In *Proc. International Conference and Exhibition on Computer Graphics and Interactive Techniques SIGGRAPH '97* (New York, NY, USA, 1997), ACM Press/Addison-Wesley Publishing Co., pp. 209–216.

[Glo09] GLOMB P.: Detection of interest points on 3D data: Extending the harris operator. In *Computer Recognition Systems 3*, vol. 57 of *Advances in Soft Computing*. Springer Berlin / Heidelberg, May 2009, pp. 103–111.

[GMGP05] GELFAND N., MITRA N. J., GUIBAS L. J., POTTMANN H.: Robust global registration. In *Proc. Eurographics Symposium on Geometry Processing SGP '05* (Aire-la-Ville, Switzerland, Switzerland, 2005), Eurographics Association, p. 197.

[HH09] HU J., HUA J.: Salient spectral geometric features for shape matching and retrieval. *Visual Computer 25*, 5-7 (2009), 667–675.

[HS88] HARRIS C., STEPHENS M.: A combined corner and edge detection. In *Proc. of The Fourth Alvey Vision Conference* (1988), pp. 147–151.

[KLT05] KATZ S., LEIFMAN G., TAL A.: Mesh segmentation using feature point and core extraction. *Visual Computer 21*, 8 (2005), 649–658.

[Low04] LOWE D. G.: Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision 60*, 2 (2004), 91–110.

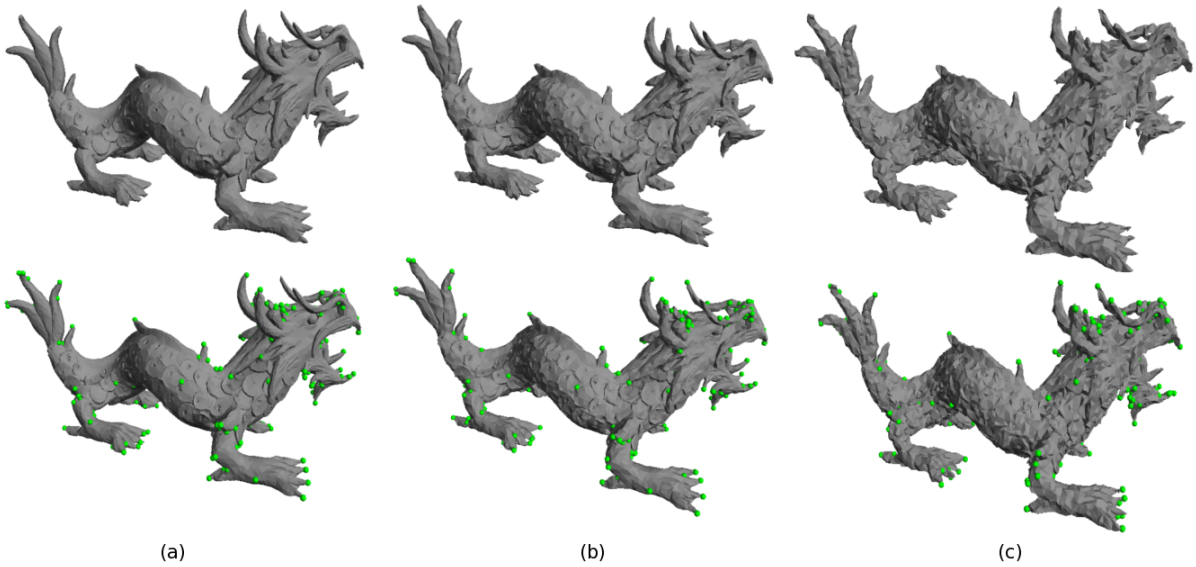

Figure 5: *Example of results on noisy objects. (a) Object without noise. (b) Object with 0.001 noise offset. (c) Object with 0.002 noise offset.*

Figure 7: *Interest points detected with different levels of detail. First column: original objects. Second column: objects with 75% of vertices. Third column: objects with 50% of vertices. Fourth column: objects with 25% of vertices.*

[LVJ05] LEE C. H., VARSHNEY A., JACOBS D. W.: Mesh saliency. In *Proc. International Conference and Exhibition on Computer Graphics and Interactive Techniques SIGGRAPH '05* (New York, NY, USA, 2005), ACM, pp. 659–666.

[LZQ06] LIU Y., ZHA H., QIN H.: Shape topics: A compact representation and new algorithms for 3D partial shape retrieval. In *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR '06* (Washington, DC, USA, 2006), IEEE Computer Society, pp. 2025–2032.

[MBO09] 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, Special Issue on 3D Object Retrieval (2009).

[SF06] SHILANE P., FUNKHOUSER T.: Selecting distinctive 3D shape descriptors for similarity retrieval. In *Proc. IEEE International Conference on Shape Modeling and Applications SMI '06* (Washington, DC, USA, 2006), IEEE Computer Society, p. 18.

[SMB00] SCHMID C., MOHR R., BAUCKHAGE C.: Evaluation of interest point detectors. *International Journal of Computer Vision 37*, 2 (2000), 151–172.

[SOG09] SUN J., OVSJANIKOV M., GUIBAS L. J.: A concise and provably informative multi-scale signature based on heat diffusion. *Computer Graphics Forum 28*, 5 (2009), 1383–1392.

[TVD08] TIERNY J., VANDEBORRE J.-P., DAOUDI M.: Enhancing 3D mesh topological skeletons with discrete contour constrictions. *Visual Computer 24*, 3 (2008), 155–172.

[ZBVH09] ZAHARESCU A., BOYER E., VARANASI K., HORAUD R. P.: Surface feature detection and description with applications to mesh matching. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition CVPR '09* (Miami Beach, Florida, June 2009).

[ZHDQ08] ZOU G., HUA J., DONG M., QIN H.: Surface matching with salient keypoints in geodesic scale space. *Computer Animation and Virtual Worlds 19*, 3-4 (2008), 399–410.