Eurographics Workshop on 3D Object Retrieval (2011)
H. Laga, T. Schreck, A. Ferreira, A. Godil, I. Pratikakis, and R. Veltkamp (Editors)

Evaluation of 3D Interest Point Detection Techniques

Helin Dutagaci, Chun Pan Cheung, and Afzal Godil

National Institute of Standards and Technology (NIST), USA

Abstract

In this paper, we compare the results of five 3D interest point detection techniques to the interest points marked by human subjects. This comparison is used to quantitatively evaluate the interest point detection algorithms. We asked human subjects to look at a number of 3D models, and mark interest points on the models via a web-based interface. We propose a voting-based method to construct ground truth out of humans' selections of interest points. Evaluation measures, namely False Positive and False Negative Errors, are then defined based on the geodesic distance between the interest points detected by a particular algorithm and the human-generated ground truth.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation—Line and curve generation

1. Introduction

Interest points, also referred to as feature points, salient points, or keypoints, are those points which are distinctive in their locality, and are present and stable at all instances of an object, or of its category of objects. Interest point detection is an important processing step involved in various 3D vision algorithms, such as 3D face recognition, registration in medical imaging, tracking, and stereoscopic vision. In the areas of shape characterization, recognition, and retrieval, using salient points is a popular approach since these salient points provide local features that are invariant to rotation, scaling, noise, deformation and many other transformations of 3D objects.

We aimed to develop an experimental setup by which the performance of interest point detection algorithms can be compared quantitatively. A widely used evaluation criterion is "repeatability", which measures the stability of the detected points on a particular object with respect to various transformations that object undergoes, such as deformation, change in resolution, and addition of noise. Our motivation, on the other hand, is to measure the detection and localization success of the algorithms with respect to human-generated ground truth.

Most of the 3D interest point detection algorithms developed in the last decade defined functions summarizing the geometrical content of localities on a 3D model in multiple scales, and selected local extrema of those functions as interest points. This approach is in accordance with the fact that humans respond more to significant local changes on the surface. However, humans also tend to see interest points within "meaningful" parts that are geometrically homogeneous; and they tend to ignore geometrical features that do not characterize the underlying class of the object.

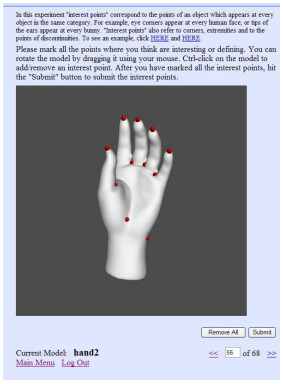


Figure 1: *User interface for marking interest points.*

Contrasting computational approaches for salient point detection to the humans' perception of interest points may lead to new directions of research, especially for object classification. With this motivation, we designed experiments to measure how close the points detected by an algorithm are

DOI: 10.2312/3DOR/3DOR11/057-064

to those considered as interest points marked by human subjects. We developed a web-based application where human subjects were asked to mark the interest points of a model. We collected data from nine subjects for 51 models. We compared five different interest point detection techniques based on "false negative" and "false positive" errors, employing the data provided by human subjects as ground truth.

2. Related work

2.1. Interest point detection techniques

There is considerable work on interest point detection applied to 2D image vision, 3D medical imaging, and range vision. In this paper, we confine our discussion to the interest point techniques that are designed to work on 3D mesh models. Although other higher-level local structures, such as larger regions of interest [GCO06], line-like features [PKG03], and edges are of great importance for shape characterization, we concentrate on techniques that detect isolated points of interest on a 3D mesh model.

In [GMGP05], an invariant descriptor called the integral volume descriptor is calculated at each vertex of the model. It corresponds to the volume of the intersection of the interior of the model with a ball centered at the vertex. This descriptor is closely related to the mean curvature at that point. A histogram of the descriptor is formed and points that correspond to the least populated bins are selected as candidate interest points. By operating on balls with varying radii, the authors employ a multi-scale approach. The points that have persistently rare values at different radii are marked as interest points.

Lee [LVJ05] also proposes a multi-scale approach based on mesh saliency. Mesh saliency is defined at each vertex as a function of the differences of Gaussian-weighted curvatures at successive scales. High saliency points are marked as interest points. Liu et al. [LLKR07] combine mesh saliency with Morse theory, while Castellani et al. [CCFM08] define another saliency measure. They apply Gaussian filtering directly on the vertex positions rather than the curvature values. Difference-of-Gaussians (DoG) are calculated at various scales, and vertices that are highly displaced after the filtering are marked as interest point candidates. The DoG approach is also used in [ZHDQ08] and [ZBVH09]. In [WNK06], the mesh is filtered with a set of Laplacians of Gaussian (LoG) to construct a pyramid. Points with local minima in both spatial and scale dimensions are declared as interest points.

Walter et al. [WAL08] extend the 2D SUSAN operator to 3D meshes to compute the saliency degree on the vertices. Sipiran and Bustos [SB10] use a 3D extension of the 2D Harris operator, which is based on the local autocorrelation of an image. They detect the local maxima of the Harris response as candidate interest points and then reduce the set either by thresholding or clustering.

Mian et al. [MBO10], define a local, thus invariant coordinate system around a point using the cropped surface surrounding it. They calculate the covariance matrix of the points on the cropped surface and use this covariance matrix to calculate the ratio between the first two principal axes of the surface. Novatnack and Nishino [NN07] parametrize a 3D mesh model onto a 2D plane, and construct a dense surface normal map. They build a scale-space by convolving the normal map with a set of Gaussian filters, and detect corners on each scale separately.

Hu and Hua [HH09] operate on Laplace-Beltrami spectral domain instead of spatial domain. They define the geometry energy on the vertices as a function of eigenfunctions and eigenvalues of the spectrum. A point is selected as an interest point if it remains as a local maximum of the geometry energy function within several successive frequencies. Thus the distinctiveness of an interest point is required to be stable within a portion of the spectrum. As another approach related to Laplace-Beltrami spectral analysis, Sun et al. [SOG09] use heat kernel function of the mesh. A point is chosen as an interest point where this function is a local maximum.

Shilane and Funkhouser [SF06] adopt a very different approach. The interest points are not detected independently for a single 3D object; they are rather picked in comparison with a large dataset of classified objects. The authors aim at determining those points which are specific to the object's category and which lead high retrieval performance. Thus, distinctive interest points are detected via a learning procedure supervised by class information of the models.

Atmosukarto and Shapiro [AS08] also use a learning scheme to determine interest points. They calculate curvature-based descriptors on each vertex of the mesh and use Support Vector Machines (SVM) to train a classifier that distinguishes between salient and non-salient points. The training data consist of manually marked salient and non-salient points on a set of models.

In this paper, we analyze the results of five of these algorithms with respect to the human-generated ground truth: Mesh saliency [LVJ05], salient points defined by Castellani et al. [CCFM08], 3D-Harris [SB10], 3D-SIFT [GW11], and scale-dependent corners [NN07]. We give detailed descriptions of these methods in Section 3.

2.2. Evaluation methods

There are a number of ways authors have used to demonstrate the success of their interest point detection algorithms: 1) Visualization of detected points on sample 3D mesh models; 2) End results of the ultimate task to which the detection algorithm serves, such as recognition or retrieval performance, or accuracy of registration; 3) Repeatability rate.

Repeatability rate is defined as the percentage of the detected points that are common in two different instances of

a scene or an object ([SL03], [BBB*10]). Usually, two detected points on the two instances are considered to be common if one falls within a neighborhood of the other, and the size of the neighborhood is denoted by ε. The repeatability is then referred as ε-repeatability.

SHREC'10 robust feature detection and description benchmark [BBB*10] evaluates and compares 3D salient point detection algorithms. The evaluation is based on the repeatability of the detected points under a variety of transformations. The dataset consists of three models (human, dog, horse) and their transformed versions. Each model has gone under 45 different transformations, such as changes in topology, sampling, scale, and addition of noise and holes. Three feature detection approaches are compared at [BBB*10]: Heat kernel-based signatures [SOG09], Salient points of Castellani et al. [CCFM08], and 3D Harris Features [SB10].

Our schema of evaluation is complementary to the analysis in [BBB*10]. We gather ground truth from human subjects, and base our analysis on humans' judgments about the interest/salient points on generic 3D models.

3. Interest point detection techniques evaluated in this work

3.1. Mesh saliency

Mesh saliency [LVJ05] is based on the local curvature over the surface. The mean curvature at each vertex is weighted by two Gaussian filters, one with scale twice the other. The absolute difference between the weighted curvatures at two scales corresponds to the mesh saliency at that scale pair. The procedure is repeated for a number of different scale pairs, then the total mesh saliency at a vertex is calculated as the sum of mesh saliency values at these scale pairs.

Candidate interest points are picked from the local maxima of the total mesh saliency function. A vertex is marked as a local maximum if its total mesh saliency is higher than all its neighboring vertices. Then the candidate points with a saliency measure higher than a threshold are selected as final interest points. We set the threshold as the average of the total mesh saliency over the local maxima.

3.2. Salient points

Castellani et al. [CCFM08] also adopted a multi-resolution approach and defined another measure of saliency on the 3D mesh model. In their approach, instead of filtering the curvature values, they filter the 3D locations of the vertices via Gaussians, and they base their saliency measure on the amount of displacement of the vertices from those of the original mesh.

For each scale, two Gaussian filters, one with twice the scale of the other, are applied on the mesh vertices. The difference between the two filtered models corresponds to the DoG map (F_s) at that particular scale. F_s at each vertex is a 3D vector measuring the displacement of that vertex within twice the scale s. F_s is projected onto the normal of the vertex to obtain a scalar quantity, which is then referred to as "scale-map".

The scale-map is normalized to a fixed range of values, and an inhibition process is applied to enhance the peaks of the map. Then, a non-maximum suppression step is implemented to detect interest points. A local maximum with an inhibited saliency value higher than the 30% of the global maximum is assigned as an interest point.

We refer to this method as "Salient points" to be consistent with the terminology in [CCFM08] and [BBB*10]. In this study, we used the "Mesh Tool" program available at the authors' web site [Mes].

3.3. 3D-Harris

Sipiran and Bustos [SB10] extended the 2D corner detection method of Harris and Stephens [HS88] to 3D mesh models. We used the code provided by the authors, and here, we give a brief description of their method.

First, a neighborhood of k rings around each vertex x is selected, and this set of points is translated so that its centroid coincides with the origin of the local coordinate frame. Then, a plane is fit to the point set using PCA. The eigenvector with the lowest eigenvalue corresponds to the normal of the fitting plane. The point set is rotated so that the normal coincides with the z-axis of the local coordinate system. A quadratic surface is fit to the transformed surface patch, and its derivatives are calculated to construct a 2×2 matrix E, elements of which involve the integration of Gaussian-weighted derivatives in two directions. This matrix contains local geometric information, and it indicates a significant change within the surface patch if both of its eigenvalues are high. The Harris operator value at the vertex is calculated as $H(x) = det(E) - 0.04(tr(E))^2$. Then, a constant fraction of the total number of vertices with the highest Harris response are selected as final interest points.

Since, setting a constant number of rings k will vary the actual neighborhood size significantly depending on the tessellation around a vertex, the authors use an adaptive scheme to determine k for each vertex. They limited the minimum distance from a vertex to the boundary of its neighborhood with a fraction of the diagonal of the object bounding rectangle, and set k for that vertex accordingly.

3.4. 3D-SIFT

The 3D-SIFT technique [GW11] described here operates on 3D voxel space; therefore it involves a voxelization step. After voxelization, in parallel to the SIFT approach in [Low04], a scale space is constructed by applying 3D Gaussian filters with increasingly large scales to the voxelized model. If the

voxelized model is denoted by a binary function $M(x,y,z)$, then each layer of the scale space is represented by its convolution with a 3D Gaussian function. Then, the Difference of Gaussian (DoG) for each level is computed by subtracting the original model from the scaled model at the corresponding level.

The extrema points are detected by searching the DoG space in both spatial and scale dimensions. The extrema points which are located on the surface are declared as interest points. Notice that these interest points are located on a voxel grid. Their locations are mapped back to the 3D space where the original mesh was defined, and the closest vertices are marked as final interest points.

3.5. Scale-dependent corners

Novatnack and Nishino [NN07] also built a scale-space representation of the model; however they analyzed the scales independent of each other to detect scale-dependent corners. We will refer to their method as "SD-corners" method.

As the first step, the vertices of a mesh model are unwrapped onto a 2D plane through embedding. Out of the 2D embedding, a "distortion map" is computed. The distortion map encodes the relative change in the model edge lengths after they have been mapped from the 3D surface onto the 2D plane. Then, the surface normals at the embedded vertices are interpolated to obtain a dense and regular "normal map" over the 2D plane. This 2D vector field (normal map) is filtered with Gaussians of varying scales to obtain a scale-space representation. The Gaussian kernels are modified using the distortion map so that the distance between two points on 2D can be corrected to match the geodesic distance on 3D. Then, first and second order partial derivatives of the normal map are calculated at each scale.

The authors define two types of geometric corners: Points that have high curvature isotropically, and points that have high curvature in at least two distinct tangential directions. They compute the Gram Matrix of first order partial derivatives of the normal map at each point. If the maximum eigenvalue of the Gram Matrix is high at a point then the point is considered to have a high corner response. Some of these candidate corners may reside on edges rather than on corners, so they are eliminated using second order derivatives of the normal map.

The corners are separately detected at each scale, and then corners at different scales are unified into a single set. Finally, these corners detected in 2D domain are mapped back onto the surface of the 3D model.

4. Subjective experiments

4.1. The 3D dataset

The 3D object dataset used in our experiments consists of 51 triangular meshes. Some models are standard models that are widely used in 3D shape research, such as Armadillo, David's head, Utah teapot, Bunny. We chose some of the models from The Stanford 3D Scanning Repository [Sta] and some others from the SHREC2007 watertight model database [Wat].

4.2. User interface for collecting ground truth

We created a web-page where users can login using an alias and participate in the experiments [Int]. Figure 1 shows a snapshot of the user interface. The user is shown the 51 3D models, one at a time. The user is free to rotate the object in 3D. She is asked to mark the interest points on the 3D model, then to click on the submit button to proceed to the next model.

Up to now, 11 participants have marked interest points on all of the 51 models. With a quick inspection, we discarded data from two users since they seemed to arbitrarily mark the interest points. We constructed the ground truth and evaluated interest point detection techniques based on the points marked by these nine subjects.

5. Ground truth

Human judgment of interest points is subjective by nature. Figure 2 shows two models marked by three different subjects. Some people tend to elaborately mark points on the smallest details (subject a in Figure 2), while others choose far less points (subject b in Figure 2). There are also different choices in locating the interest points; for example, subjects don't mark interest points on the same exact location around the smooth corners of the chair in Figure 2.

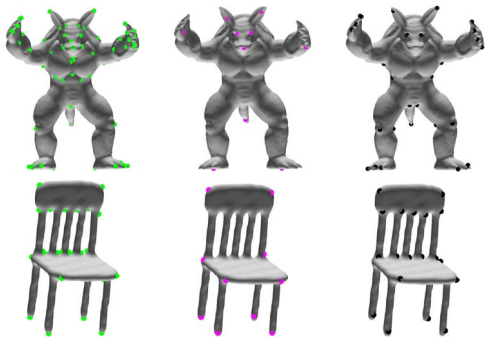


Figure 2: *Models marked by three different subjects.*

We look for some consensus among the users in order to merge all the marked points into a final set of ground truth interest points. It is also necessary to reject outliers and incorrectly marked points, and discard small variations of localization. We have two criteria while constructing the ground truth: The radius of an interest region, and the number of users n that marked a point within that interest region.

We set the radius of interest region as σd_M, where d_M stands for the model diameter; i.e. the largest Euclidean distance between all pairs of vertices of model M. We group all the interest points (marked by distinct subjects) whose geodesic distances to each other are less than $2\sigma d_M$. If the number of points in the group is less than n we discard that group. Otherwise, we select a representative among the group, and set it as a ground truth interest point. The point with the minimum sum of geodesic distances to the other points in the group is selected as the representative. Notice that, two groups can be overlapping, i.e. can have vertices in common. If the distance between two representatives turn out to be less than $2\sigma d_M$, the representative with the smaller number of group points is discarded from the ground truth interest point set.

Figure 3 shows a model with the interest points gathered from nine subjects (on the left), and the final ground truth points (on the right). We zoomed two sample groups, and circled their corresponding representatives on the right figure.

We will denote the set of ground truth points obtained with the parameters n and σ as $\mathcal{G}_M(n,\sigma)$ for a particular model M. These two parameters highly determine the final set of ground truth interest points. With high n, we have less number of ground truth points, since not all users choose small details as interest points (Figure 2). As σ increases, we expect to have more ground truth points, since we accept more variation on localization of the points marked by the subjects. However as σ further increases the region it defines tend to include distinct interest regions, thus close interest points marked on distinct structures start to merge. In Section 7, we give the average number of ground truth points on our dataset with varying n and σ.

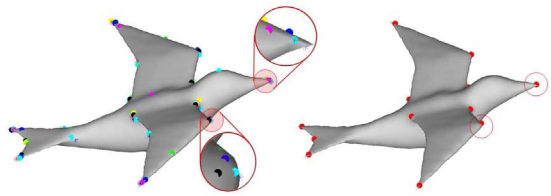


Figure 3: *Ground truth generation. On the left, different markers correspond to the points selected by different users. On the right, resulting ground truth interest points are indicated with red dots ($n = 4$, $\sigma = 0.03$)*

6. Evaluation method

Previous evaluation methods for 3D salient point detectors measured the repeatability rate according to varying factors, such as deformation of the model, scale change, different modalities, noise, and topological changes [LSKW05, BBB*10]. We perform our evaluation on a single instance of a model with respect to human generated ground truth, and use false positive and false negative errors as performance measures.

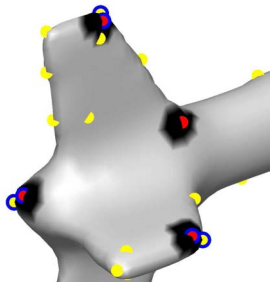


Figure 4: *The red dots indicate ground truth interest points. Yellow dots are the points marked by an interest point detection algorithm. The paired interest points are enclosed by blue circles. The red dot not enclosed by a blue circle is a false negative. The unenclosed yellow dots are false positives. The black regions correspond to the points with geodesic distance to a ground truth point, less than r*

For simplicity, let us denote the set of ground truth points $\mathcal{G}_M(n,\sigma)$ as $\mathcal{G}$, and the set of interest points detected by an algorithm on model M as $\mathcal{A}$. For an interest point g in set $\mathcal{G}$ we define a geodesic neighborhood of radius r:

$$\mathcal{C}_r(g) = \{p \in M | d(g,p) \leq r\}$$

where $d(g,p)$ corresponds to the geodesic distance between points g and p. The parameter r controls the localization error tolerance. A point g is considered to be "correctly detected" if there exists a detected point $a \in \mathcal{A}$ in $\mathcal{C}_r(g)$, and 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, we define the false negative error rate at localization error tolerance r as:

$$FNE(r) = 1 - \frac{N_C}{N_G}, \quad (1)$$

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

The rate of false positives of an interest point detection algorithm is another measure of its relevancy to human perception of interest points. The algorithm is not supposed to find points on regions that are not of interest to humans. To calculate the false positive error rate we proceed as follows: Each correctly detected point in $g \in \mathcal{G}$ 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 number of false positives, denoted as N_F is equal to

$$N_F = N_A - N_C, \quad (2)$$

where N_A is the number of detected interest points by the algorithm. The false positive error rate at localization error

tolerance r is then,

$$FPE(r) = \frac{N_F}{N_A}. \quad (3)$$

Note that our definition of false positive error rate is different than the conventional one, where the number of false positives is normalized by the number of all true negatives; i.e. the number of vertices that are not true interest points. Since this number depends on the tessellation of the mesh model, we prefer to normalize the number of false positives with the number of interest points the algorithm produces.

Figure 4 demonstrates sample false negatives and positives at the tail of an airplane model. The red dots indicate ground truth interest points. Yellow dots are the points marked by an interest point detection algorithm. Corresponding pairs of ground truth points and algorithm-detected points are enclosed by blue circles (correct detection). The red dot not enclosed by a blue circle is a false negative. The unenclosed yellow dots are false positives. The black regions correspond to the points with geodesic distance to a ground truth point, less than r

7. Results

As mentioned in Section 4, we collected data from nine subjects via our web-based interface, and constructed the ground truth as described in Section 5. The number of subjects may be increased to achieve a more reliable ground truth, however marking interest points is an intensive, time-consuming task.

The first and second columns of Figure 5 show interest points marked by all the nine subjects (with different markers) and the ground truth points (red dots), respectively. In the figure, the ground truth points correspond to the case $\sigma = 0.03$ and $n = 2$. The last five columns show the points from the five interest point detection algorithms. We can list some of our observations on Figure 5 as follows:

- Marking interest points, especially on detailed objects, is an intensive task. For example, for the Armadillo model in Figure 5-b, some subjects marked only one of two symmetrical interest points. Yet, some subjects fail to well-localize the interest points, for example at the corners of the chair model. These errors are resolved by the ground-truth building process to some extent.
- Some subjects marked points on flat or smooth regions, just to define the object. An example is the green dot on the center of the seat of the chair.
- The algorithms tend to mark more interest points then the human users, especially the mesh saliency and SD-corners methods. The number can be reduced via adjusting the parameters of the algorithms, however that adjustment may cause missing of important interest points.
- All subjects seem to respond to the extremities, regardless of the scale of the local perturbation. Back of the camel and tip of the cactus correspond to a larger scale, while the toes and fingers of the Armadillo model are of fine scale. This necessitates an interest point algorithm to be able to operate on a wide range of scales. The 3D-SIFT algorithm detected an interest point on large scales (shoulder of the head model, back of the camel, tip of the cactus), due to its coarse voxelization strategy. However, it didn't well-localize the finer interest points as the other three algorithms did (ear tips and fingers of the Armadillo model).
- While other algorithms select points that remain salient across scales, SD-corners method gathers interest points from all the scales. The result is a large number of interest points representing both small (facial features of the girl model) and large details (tip of the cactus).
- Most of the subjects marked the facial features of the head model in detail, while putting just one or two representative points on the more detailed hair. The ground truth points of the head model includes very few points from the hair. On the other hand, the interest points generated by the algorithms populate on the hair.
- The algorithms find interest points along the edges of the chair model due to the high curvature in one direction, while there are no ground truth points along those edges.

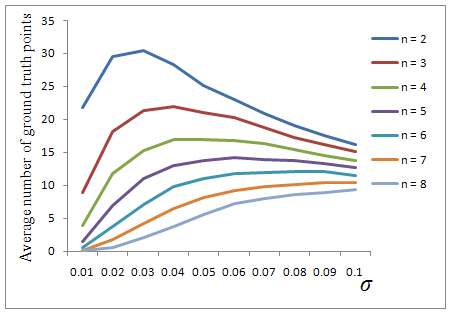


Figure 6: *Average number of ground truth interest points per model with varying σ and n.*

Figure 6 shows the average number of ground truth interest points on a model with varying σ and n. The average is taken over the 51 models we have in our dataset. The case with $n = 1$ creates unreliable ground truth, so we exclude that case from our analysis. With increasing n, i.e. the number of subjects that vote for an interest point, we have less ground truth points. As the localization tolerance, σ, increases to 0.05 we have an increase in the number of ground truth points; then it decreases due to the merging of nearby groups of marked interest points. This effect is especially strong for $n = 2$.

We have calculated False Negative and False Positive errors of the algorithms using ground truth point sets generated with various σ and n. However, here, we only report results with $\sigma = 0.03$ and $n = 2$. Since we have few subjects, there is a small chance that two users will accidentally mark an uninteresting point on the 3D surface. Figure 7 shows the

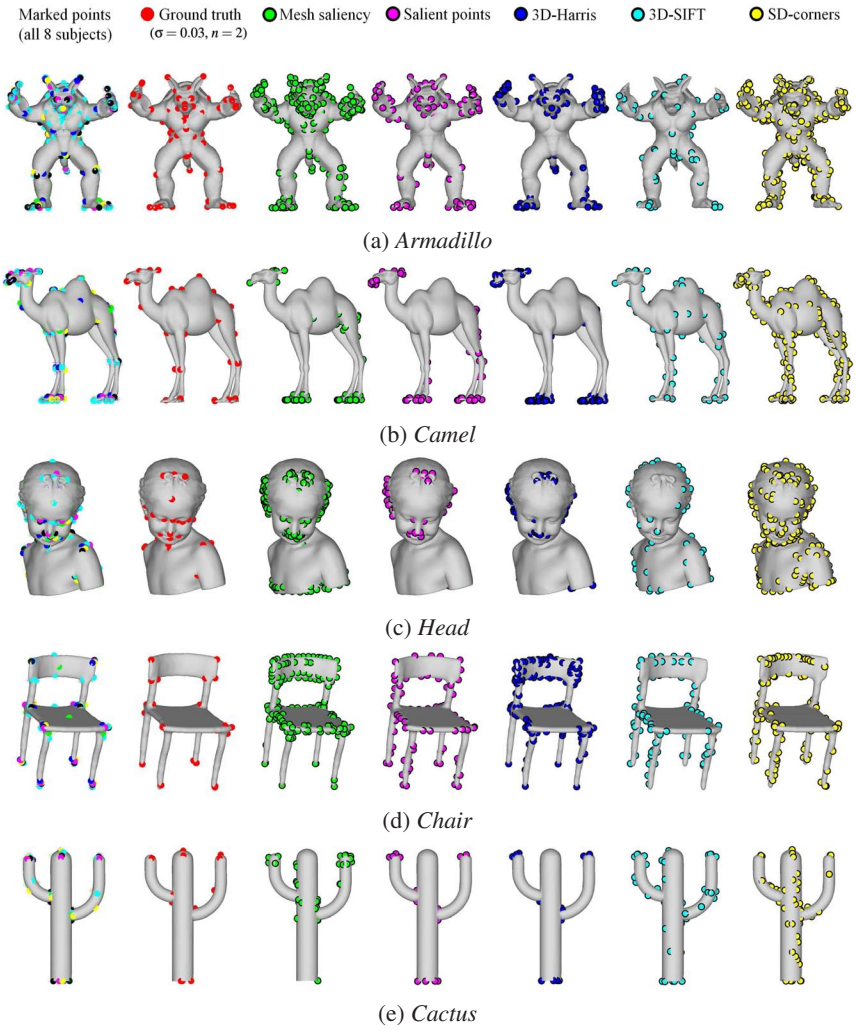
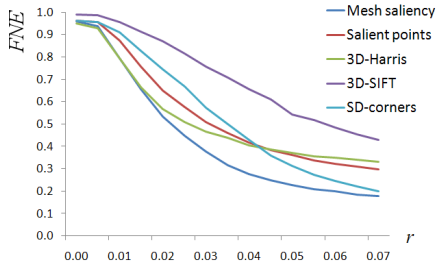


Figure 5: *Interest points marked by all nine human subjects (first column), ground truth obtained by setting* $\sigma = 0.03$ *and* $n = 2$ *(second column), interest points detected by the algorithms: Mesh saliency (third column), Salient points (fourth column), 3D-Harris (fifth column), 3D-SIFT (sixth column), and SD-corners (seventh column).*

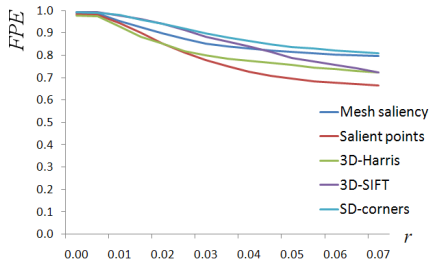
False Negative and False Positive errors of the five interest point detection algorithms with respect to r. The errors are averaged over all the models in the dataset. The False Negative error drops more quickly with the Mesh saliency method and it gives the least error for all values of r. The SD-corners method catches the low False Negative Error levels of Mesh saliency method at about $r = 0.07$, which indicates a high detection rate with a low localization accuracy. We guess, the reason behind the low localization of SD-corners method is due to the unwrapping of 3D model onto a 2D plane. 3D-SIFT doesn't perform as well as other algorithms, since the coarse voxel structure doesn't allow good localization of interest points on the mesh model. Mesh saliency method well localizes the interest points with a cost of a high False Positive error compared to the Salient points and 3D-Harris methods. The SD-corners also have a large False Positive error due to the large number of points it detected.

8. Conclusion

We designed experiments to compare automatic salient point detection algorithms with humans' selections of interest points. We developed a web-based application where human subjects marked interest points on 3D models. We compared five different interest point detection techniques based on False Negative and False Positive errors, employing the human-provided data as ground truth. The 3D model dataset, the ground truth data, and the evaluation code are available at our website [Ben].



(a) *Average False Negative Error*



(b) *Average False Positive Error*

Figure 7: *Evaluation of five algorithms. Ground truth obtained with* $\sigma = 0.03$ *and* $n = 2$.

9. Acknowledgments

This project was funded by the Shape Metrology IMS. We would like to thank Daniela Giorgi and AIM@SHAPE for the models from the Watertight Track of SHREC 2007, and Stanford Computer Graphics Laboratory for the models from The Stanford 3D Scanning Repository. We would like to thank Ivan Sipiran, Benjamin Bustos, Umberto Castellani, Ko Nishino, and Prabin Bariya for sharing their interest point detection codes.

References

[AS08] ATMOSUKARTO I., SHAPIRO L. G.: A salient-point signature for 3d object retrieval. In *ACM International Conference on Multimedia Information Retrieval (MIR'08)* (2008), ACM, pp. 208–215.

[BBB*10] BRONSTEIN A., BRONSTEIN M., BUSTOS B., CASTELLANI U., CRISANI M., FALCIDIENO B., GUIBAS L., KOKKINOS I., MURINO V., SIPIRAN I., OVSJANIKOVY M., PATANÈ G., SPAGNUOLO M., SUN J.: SHREC 2010: Robust feature detection and description benchmark. In *Eurographics Workshop on 3D Object Retrieval (3DOR'10)* (2010), pp. 79–86.

[Ben] http://www.itl.nist.gov/iad/vug/sharp/benchmark/3DInterestPoint .

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

[GCO06] GAL R., COHEN-OR D.: Salient geometric features for partial shape matching and similarity. *ACM Trans. Graph. 25*, 1 (2006), 130–150.

[GMGP05] GELFAND N., MITRA N. J., GUIBAS L. J., POTTMANN H.: Robust global registration. In *Symposium on Geometry Processing* (2005), pp. 197–206.

[GW11] GODIL A., WAGAN A. I.: Salient local 3D features for 3D shape retrieval. In *3D Image Processing (3DIP) and Applications II, SPIE* (2011).

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

[HS88] HARRIS C., STEPHENS M.: A combined corner and edge detector. In *4th Alvey Vision Conference* (1988), pp. 147–151.

[Int] http://control.nist.gov/sharp/view/ .

[LLKR07] LIU Y.-S., LIU M., KIHARA D., RAMANI K.: Salient critical points for meshes. In *ACM Symposium on Solid and Physical Modeling (SPM'07)* (2007), pp. 277–282.

[Low04] LOWE D. G.: Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision 60*, 2 (2004), 91–110.

[LSKW05] LLOYD B. A., SZEKELY G., KIKINIS R., WARFIELD S. K.: Comparison of salient point detection methods for 3D medical images. *The Insight Journal - 2005 MICCAI Open-Source Workshop* (2005).

[LVJ05] LEE C. H., VARSHNEY A., JACOBS D. W.: Mesh saliency. In *ACM SIGGRAPH 2005* (2005), pp. 659–666.

[MBO10] 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. Vision 89*, 2-3 (2010), 348–361.

[Mes] http://profs.sci.univr.it/ castella/research.html/.

[NN07] NOVATNACK J., NISHINO K.: Scale-dependent 3D geometric features. In *ICCV* (2007), pp. 1–8.

[PKG03] PAULY M., KEISER R., GROSS M. H.: Multi-scale feature extraction on point-sampled surfaces. *Comput. Graph. Forum 22*, 3 (2003), 281–290.

[SB10] SIPIRAN I., BUSTOS B.: A robust 3D interest points detector based on harris operator. In *Eurographics 2010 Workshop on 3D Object Retrieval (3DOR'10)* (2010), pp. 7–14.

[SF06] SHILANE P., FUNKHOUSER T.: Selecting distinctive 3D shape descriptors for similarity retrieval. In *Shape Modeling International* (June 2006).

[SL03] SEBE N., LEW M. S.: Comparing salient point detectors. *Pattern Recogn. Lett. 24*, 1-3 (January 2003), 89–96.

[SOG09] SUN J., OVSJANIKOV M., GUIBAS L.: A concise and provably informative multi-scale signature based on heat diffusion. In *Eurographics Symposium on Geometry Processing (SGP)* (2009), pp. 1383–1392.

[Sta] http: www.graphics.stanford.edu/data/3Dscanrep/ .

[WAL08] WALTER N., AUBRETON O., LALIGANT O.: Salient point characterization for low resolution meshes. In *ICIP* (2008), pp. 1512–1515.

[Wat] http://watertight.ge.imati.cnr.it/ .

[WNK06] WESSEL R., NOVOTNI M., KLEIN R.: Correspondences between salient points on 3D shapes. In *Vision, Modeling, and Visualization (VMV'06)* (2006), pp. 365–372.

[ZBVH09] ZAHARESCU A., BOYER E., VARANASI K., HORAUD R. P.: Surface feature detection and description with applications to mesh matching. In *CVPR* (2009).

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