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

SHREC 2010: robust feature detection and description benchmark

A. M. Bronstein[†1], M. M. Bronstein[†,1], B. Bustos[6], U. Castellani[†,3] M. Crisani[3,9], B. Falcidieno[4], L. J. Guibas[†,2],

I. Kokkinos[5], V. Murino[3,9], I. Isipiran[6], M. Ovsjanikov[†,8], G. Patané[4], M. Spagnuolo[4], J. Sun[7],

[1]Department of Computer Science, Technion – Israel Institute of Technology
[2]Department of Computer Science, Stanford University
[3]Department of Computer Science, University of Verona
[4]CNR-IMATI Genova
[5]Deparment of Applied Mathematics, École Centrale de Paris
[6]Department of Computer Science – University of Chile
[7]Princeton University
[8]Institute for Computational and Mathematical Engineering, Stanford University
[9]Italian Institute of Technology, Genova

Abstract

Feature-based approaches have recently become very popular in computer vision and image analysis applications, and are becoming a promising direction in shape retrieval. SHREC'10 robust feature detection and description benchmark simulates the feature detection and description stages of feature-based shape retrieval algorithms. The benchmark tests the performance of shape feature detectors and descriptors under a wide variety of transformations. The benchmark allows evaluating how algorithms cope with certain classes of transformations and strength of the transformations that can be dealt with. The present paper is a report of the SHREC'10 robust feature detection and description benchmark results.

Categories and Subject Descriptors (according to ACM CCS): H.3.2 [Information storage and retrieval]: Information Search and Retrieval—Retrieval models I.2.10 [Artificial intelligence]: Vision and Scene Understanding—Shape

1. Introduction

Feature-based approaches have recently become very popular in computer vision and image analysis applications, notably due to the works of Lowe [Low04], Sivic and Zisserman [SZ03], and Mikolajczyk and Schmid [MS05]. In these approaches, an image is described as a collection of local features ("visual words") from a given vocabulary, resulting in a representation referred to as a *bag of features*. The bag of features paradigm relies heavily on the choice of the local feature descriptor that is used to create the visual words. A common evaluation strategy of image feature detection and description algorithms is the stability of the detected features and their invariance to different transformations applied to an image. In shape analysis, feature-based approaches have been introduced more recently and are gaining popularity in shape retrieval applications.

SHREC'10 invariant feature detection and description benchmark simulates the feature detection and description stages of feature-based shape retrieval algorithms. The benchmark tests the performance of shape feature detectors and descriptors under a wide variety of different transformations. The benchmark allows evaluating how algorithms cope with certain classes of transformations and what is the strength of the transformations that can be dealt with.

† Organizer of the SHREC track. All organizers and participants are listed in alphabetical order. For any information about the benchmark, contact mbron@cs.technion.ac.il

2. Data

The dataset used in this benchmark was from the TOSCA shapes [BBK08], available in the public domain. The shapes were represented as triangular meshes with approximately 10,000–50,000 vertices.

The dataset consisted of 3 shapes, with simulated transformations applied to them. For each null shape, transformations were split into 9 classes shown in Figure 1: *isometry* (non-rigid almost inelastic deformations), *topology* (welding of shape vertices resulting in different triangulation), *micro holes* and big *holes*, *global* and *local scaling*, additive Gaussian *noise*, *shot noise*, down-*sampling* (less than 20% of the original points).

In each class, the transformation appeared in five different versions numbered 1–5. In all shape categories except scale and isometry, the version number corresponded to the transformation strength levels: the higher the number, the stronger the transformation (e.g., in noise transformation, the noise variance was proportional to the strength number). For scale transformations, the levels 1–5 corresponded to scaling by the factor of 0.5, 0.875, 1.25, 1.625, and 2. For the isometry class, the numbers do not reflect the strength of the transformation. The total number of transformations per shape was 45, and the total dataset size was 138. The dataset is available at http://tosca.cs.technion.ac.il/book/shrec_feat.html.

3. Evaluation methodology

The evaluation was performed separately for feature detection and feature description algorithms. The participants were asked to provide the following: (i) for each shape Y in the dataset, a set of detected feature points $\mathcal{F}(Y) = \{y_k\}_k$ (typically $|\mathcal{F}(Y)| << |Y|$); (ii) optionally, for each detected point y_k, a descriptor vector $\{\mathbf{f}(y_k)\}_{k=1}^{|\mathcal{F}(Y)|}$; or alternatively, (iii) for each shape Y in the dataset, a dense descriptor $\{\mathbf{f}(y_k)\}_{k=1}^{|Y|}$. The performance was measured by comparing features and feature descriptors computed for transformed shapes and the corresponding null shapes.

Feature detection. The quality of the feature detection was measured using the *repeatability* criterion. Assuming for each transformed shape Y in the dataset the groundtruth dense correspondence to the null shape X to be given in the form of pairs of points $\mathcal{C}_0(X,Y) = \{(y'_k, x_k)\}_{k=1}^{|Y|}$ (and same way, $\mathcal{C}_0(Y,X)$), a feature point $y_k \in \mathcal{F}(Y)$ is said to be *repeatable* if a geodesic ball of radius R around the corresponding point $x'_k : (x'_k, y_k) \in \mathcal{C}_0(X,Y)$ contains a feature point $x_j \in \mathcal{F}(X)$. The subset $\mathcal{F}_r(Y) \subseteq \mathcal{F}(Y)$ of repeatable features is given by

$$\mathcal{F}_{R,X}(Y) = \{y_k \in \mathcal{F}(Y) : \mathcal{F}(X) \cap B_R(x'_k) \neq \emptyset, (x'_k, y_k) \in \mathcal{C}_0(X,Y)\},$$

where $B_R(x'_k) = \{x \in X : d_X(x, x'_k) \leq R\}$ and d_X denotes the geodesic distance function in X. The repeatability $\mathrm{rep}(Y,X)$ of $\mathcal{F}(Y)$ in X is defined as the percentage of features from $\mathcal{F}(Y)$ that are repeatable,

$$\mathrm{rep}(Y,X) = \frac{|\mathcal{F}_{R,X}(Y)|}{|\mathcal{F}(Y)|}.$$

For a transformed shape Y and the corresponding null shape X, the overall feature detection quality was measured as $(\mathrm{rep}(Y,X) + \mathrm{rep}(X,Y))/2$. The value of $R = 5$ was used in the benchmark. This radius constitute approximately 1% of the shapes diameter. Features without groundtruth correspondence (e.g. in regions in the null shape corresponding to holes in the transformed shape) were ignored.

Feature description. The quality of feature descriptor was measured as the average normalized L_2 distance between the descriptor vectors in corresponding points,

$$Q(X,Y) = \frac{1}{|\mathcal{F}(X)|} \sum_{k=1}^{|\mathcal{F}(X)|} \frac{\|\mathbf{f}(y_k) - \mathbf{g}(x_j)\|_2}{\|\mathbf{f}(y_k)\|_2 + \|\mathbf{g}(x_j)\|_2},$$

$(x_j, y_k) \in \mathcal{C}(X,Y)$. For sparse feature descriptors $\{\mathbf{f}(y_k)\}_{k=1}^{|\mathcal{F}(Y)|}$, $\{\mathbf{g}(x_j)\}_{j=1}^{|\mathcal{F}(X)|}$ computed on feature points $\mathcal{F}(X)$ and $\mathcal{F}(Y)$, respectively, the correspondence was defined between the closest detected points,

$$\mathcal{C}(X,Y) = \left\{ (\arg \min_{x_j \in \mathcal{F}(X)} d_X(x_j, x'_k), y_k \in \mathcal{F}(Y)) \right\},$$

$(x'_k, y_k) \in \mathcal{C}_0(X,Y)$. For dense descriptors, the groundtruth correspondence $\mathcal{C}_0(X,Y)$ was used as $\mathcal{C}(X,Y)$.

4. Feature detection methods

Three families of feature detection methods were evaluated in this benchmark: heat kernel-based features [SOG09] (denoted HK for notation brevity), 3D Harris features [IB10] (denoted H), and salient points [CCFM08] (hereinafter, SP).

HK1–2: Heat Kernel features

Sun *et al.* [SOG09] proposed a feature detection algorithm based on the *heat kernel signature* (HKS), $\mathbf{h}(x) = (h_{t_1}(x,x), \ldots, h_{t_n}(x,x))$, where $h_t(x,y) \approx \sum_{i=0}^{K} e^{-\lambda_i t} \phi_i(x) \phi_i(y)$ is the heat kernel associated with the positive-semidefinite Laplace-Beltrami operator Δ_X, and λ and ϕ are the eigenvalues and eigenfunctions of Δ_X. The HKS provides a multi-scale notion of curvature, and local maxima of the HKS function for a large time parameter correspond to tips of protrusions that can be used as stable features. At the same time, HKS is also used as a feature descriptor.

In the first setting of this method (denoted in the following as HK1 for brevity), feature points are detected on a given mesh by computing the HKS function $h_t(x,x)$ for a large fixed value of t, and declare a point x a feature point if $h_t(x,x) > h_t(x',x')$ for all x' in a 2-ring neighborhood

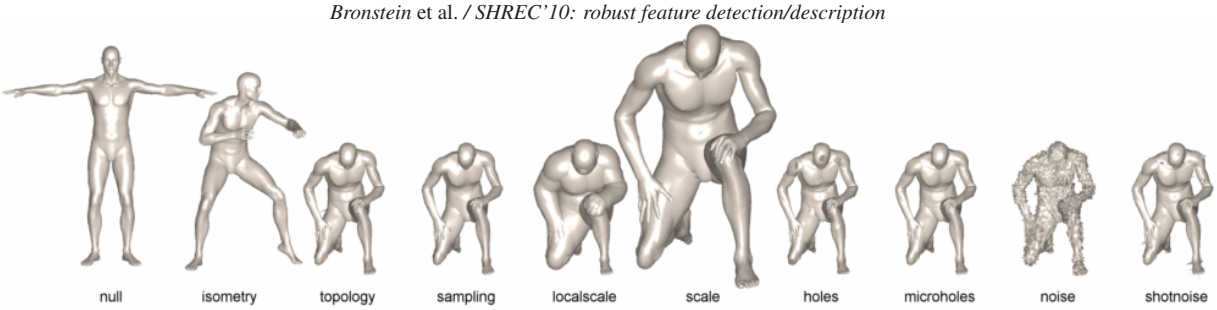

Figure 1: *Transformations of the human shape used in the tests (shown in strength 5, left to right): null, isometry, topology, sampling, local scale, scale, holes, micro holes, noise, shot noise.*

of x. To reduce processing time, each mesh was simplified to have at most 10,000 vertices, non-manifold edges were removed, and each face of the mesh was consistently oriented using the Geomagic software. The discretization of the Laplace-Beltrami operator Δ_X was computed using the MeshLP method of Belkin *et al.* [BSW08]. The HKS was computed based on the eigendecomposition of the discretizing matrix, after normalizing the shape to have total surface area 1. This also allowed to use a uniform value of $t = 0.1$ across all shapes. After detecting feature points on the simplified mesh using the local 2-ring maximum criterion, the detected features were projected onto the original high-resolution mesh.

The second method (HK2) follows the same basic procedure described above, but differs in the final step of selecting feature points. After defining the function $h_t(x,x)$ for $t = 0.1$, persistent homology was used to filter out unstable feature points. For this, the 0 dimensional persistence diagram of this function [ELZ00, CGOS09] was computed. A point was declated a feature, if it is a local maximum of this function and furthermore, if the difference between the death and birth times of the corresponding connected component is above a threshold α (see [ELZ00] for details). A uniform $\alpha = 0.1$ was used for all shapes, which was chosen by manually examining the persistence diagram of one of the null shapes. † This process generally results in a sparser set of features, but that should be more stable to perturbations of the function.

H1–3: Harris 3D features

Isipiran and Bustos [IB10] applied a method similar to Harris and Stephens [HS88] interest points detector for images, referred here as 3D Harris point detector. The problem with 3D data is the arbitrary topology and subsequently, the calculation of derivatives. Isipiran and Bustos [IB10] used [Glo09] as a basis for proposing a robust interest points detector on 3D meshes.

Interest points detection: Let x be a vertex of the shape X and $V_k(x)$ denote the neighborhood of k rings around x. First, the centroid of $V_k(x)$ is calculated, and the set of points is translated so the centroid is in the origin of the 3D coordinate system. Next, a plane is fit to the translated points applying PCA to the set of points and choosing the eigenvector with the lowest associated eigenvalue as the normal of the fitting plane. The set of points is rotated so that the normal of the fitting plane is the z-axis. To calculate derivatives, a quadratic surface $f(u,v)$ is fit to the set of transformed points.

Second, the matrix E associated to the point x

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

is computed, where

$$\begin{aligned} A &= \frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}^2} e^{\frac{-(u^2+v^2)}{2\sigma^2}} . f_u^2(u,v) du dv; \\ B &= \frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}^2} e^{\frac{-(u^2+v^2)}{2\sigma^2}} . f_v^2(u,v) du dv; \\ C &= \frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}^2} e^{\frac{-(u^2+v^2)}{2\sigma^2}} . f_u(u,v) f_v(u,v) du dv, \end{aligned}$$

and σ is a parameter defining the support of the Gaussian function (the setting of σ is considered next). The Harris operator value at point x is calculated as $H(x) = det(E) - 0.04(tr(E))^2$.

Finally, a constant fraction of the total number of vertices with the highest Harris operator response is selected as feature points.

Adaptative neighborhood size: Different neighborhood size is selected depending of the tessellation around a point. Given a point $x \in X$, $R_k(x)$ is the boundary of the k-ring neighborhood $V_k(x)$ around x (the set of points distant from x by k edges). The distance from a point x to $R_k(x)$ is defined as the closest point in the Euclidean sense,

$$d(x, R_k(v)) = \max_{x' \in R_k(x)} \|x - x'\|.$$

† M. Ovsjanikov, J. Sun, and L. J. Guibas would like to thank Primoz Skraba for providing the code for the computation of persistence diagrams.

Finally, the neighborhood size of a point x is defined as

$$\rho(x) = \{k : d(x, R_k(x)) \leq \delta, d(x, R_{k-1}(x)) < \delta\},$$

where δ is a fraction of the diagonal of the object bounding rectangle. It is important to note that the above method always find a neighborhood to a point, even with complex and irregular tessellations around that point.

Using this information, the variance of the Gaussian function in the Harris operator is defined at each point as $\sigma(x) = \delta/\rho(x)$. Therefore, each point has a different support for the applied Gaussian window when calculating its operator value, which is consistent with the neighborhood size.

PB1–3: Salient points

The salient points detection method proposed by Castellani *et al.* [CCFM08] and used in [TCF09] is inspired by the research on saliency measure on 2D images. The source mesh is decomposed in multiscale representations, and a saliency measure is defined by combining the results gathered at each scale level. The method consists of the following main steps:

Multiscale representation: The first step consists of applying N Gaussian filters on the shape X, obtaining N multi-dimensional filtering maps $\{F_s\}_{s=1}^{N}$. Gaussian filtering is applied as follows: let $g(x,\sigma)$ the Gaussian operator with standard deviation σ, applied on the vertex $x \in X$. The *neighborhood region* of x, over which the filtering is applied, is built by expanding a n-rings search starting from x, and collecting all those vertices displaced within a distance equal to 2.5σ. The difference-of-Gaussians (DoG) operator is defined as:

$$F_s(x) = g(x,\sigma_s) - g(x,k\sigma_s) \quad (1)$$

where σ_s is the value of the standard deviation associated to scale s, and k is a constant equal to 2. 5 Six scales of filtering have been fixed, corresponding to standard deviation values $\sigma_s \in \{1\varepsilon, 2\varepsilon, \ldots, 6\varepsilon\}$, where ε amounts to 0.1% of the length of the main diagonal located in the bounding box of the model. Note that, as studied by [Low04], fixing a constant factor k for DoG computation provides a close approximation to the scale-normalized Laplacian of Gaussian, which is required for true scale invariance.

3D saliency measure definition: This step aims at obtaining a dense measure of mesh saliency (i.e., associated to each vertex). Note that $F_s(x)$ is a 3D vector measuring how much the vertex x has been moved from its original position after the filtering. In order to reduce such displacement in a scalar quantity, the displacement vector $F_s(x)$ is projected to the normal $n(x)$ at x. This way, the *scale* map X_s is obtained as

$$X_s(x) = \langle n(x), g(x,\sigma_s) - g(v,k\sigma_s)\rangle. \quad (2)$$

Furthermore, this reduces the shrinking effect which arises typically when Gaussian filter is applied to meshes. Each map is normalized by adopting the Itti's approach [IKN98]: normalizing the values in the map to a fixed range $[0,\ldots,R]$; finding the location of global maximum T; finding all the other local maxima and computing their average $\hat{t}$; globally multiplying the map by $(T-\hat{t})^2$ by obtaining the final normalized scale map $\hat{X}_s$. The effect of this normalization is to increase the evidence of the highest peaks.

Feature points detection: The above peaks enhancement procedure is emphasized by introducing an adaptive *inhibition-process* on each normalized scale map. From each vertex $x \in X$, all the values of the scale map $\hat{X}_s$ observed on the neighborhood of x are considered. If the $\hat{X}_s(x)$ is higher than the 85% of the values in its neighborhood, the value is retained, otherwise $\hat{X}_s(x)$ is set to zero. Therefore, the *inhibited* saliency map is obtained by simply adding the contribution of each *inhibited* scale map. Finally, in order to detect salient points a *non-maximum suppression* phase on the *inhibited* saliency map is performed: a point is detected if it is a local maximum and its value is higher than the 30% of the global maximum. Note that, after the inhibition phase, the neighborhood of a point is adaptively defined by expanding the local region while new non-zero points are found.

In this benchmark, three settings were used (denoted SP1–3), differing in the number of feature points selected.

5. Feature description methods

Three families of feature description methods were evaluated in this benchmark: heat kernel signature [SOG09] computed on the feature points detected by the method HK1–3 (referred hereinafter as sparse HKS or SHK); dense heat kernel signature as used in [OBG09, BBOG10], and spin image signatures [JH99] computed on the feature points detected by the method SP1, as used in [TCF09].

SHK1–2: Sparse Heat Kernel Signature

As a feature descriptor, HKS computed at the heat kernel feature points (HK1–2) sampled at 100 values of t was used. In other words, the descriptor of point x is a vector with entries $h_{t_i}(x,x)$ for 100 values of $t_1,\ldots,t_n$. Logarithmic sampling in time proposed by Sun *et al.* [SOG09] was used with $t_i = \log 10/\lambda_{\max},\ldots,4\log 10/\lambda_2$.

DHK1–3: Dense Heat Kernel Signature

In [OBG09, BBOG10], a dense version of the HKS descriptor was used. In this approach, the feature detection stage is avoided, and the feature descriptor is computed at all the points of shapes. In the present benchmark, three settings of the dense HKS were used. In DHK1, values of $K = 100$, $n = 6$ were used, and $t_1,\ldots,t_6$ were chosen as $1024, 1351, 1783, 2353, 3104$ and 4096 (these are setting identical to [OBG09]). In DHK2, eigenpairs and the mass matrix obtained by the linear finite elements method (FEM) described in [PSF10] were used to compute the heat kernel signatures. Same settings for K, n and $t_1,\ldots,t_n$ as in SG1

were used. Such a discretization is known to be less sensitive to geometric and topological noise, irregular sampling, and local shape deformations. In DHK3, the scale-invariant heat kernel signature (SI-HKS) [BK10] was used. SI-HKS is

$$\hat{\mathbf{h}}(x) = |\mathcal{F}_\tau \mathrm{diff}_\tau \log(h_{\alpha^{\tau_1}}(x,x),\ldots,h_{\alpha^{\tau_n}}(x,x))|,$$

where diff denotes the finite difference operator and $\mathcal{F}$ is the Fourier transform. Cotangent weights and $K = 100$ first eigenpairs were used to obtain $\mathbf{h}$. Value of $\alpha = 2$ and τ ranging from 1 to 25 with increments of $1/16$ were used. The first six discrete frequencies of the Fourier transform were taken (these are settings identical to [BK10]).

SI: Spin images

Spin images (SI) were introduced in the seminal work of [JH99], to characterize the properties of a 3D object with respect to a single oriented point. Spin image descriptors have been successfully applied in computer vision and in computer graphics for several applications such as object recognition [JH99], partial views registration [ABC04] and 3D object retrieval [ABADB07]. In practice, spin images are computed as the 2D histogram of distances of neighboring points from the normal vector and the tangent plane respectively. Given a vertex x of the shape X, and its normal $n(x)$, the *spin-map* $S_X : \mathbb{R}^3 \to \mathbb{R}^2$ is defined as:

$$S_X(x') \to (\alpha, \beta),$$

where $x' \in \mathcal{N}(x)$ denotes neighbor vertices, $\alpha = \sqrt{\|x'-x\|^2 - \langle n(x), x'-x\rangle^2}$, and $\beta = \langle n(x), x'-x\rangle$. The values α and β are quantized into $n \times m$ bins and their occurrences are accumulated by forming the spin image.

6. Results

Feature detectors. Tables 1–8 show the repeatability of the feature detection approaches compared in this benchmark. Heat kernel-based feature detectors (HK1–2) perform the best on the average (HK1 being slightly superior to HK2 in average repeatability in strength 5). HK2 shows the best results in isometry, micro holes, scale, local scale, and sampling transformation classes, as well as noise till strength 4. HK1 shows the best results in noise of strength 5 class and shot noise. Harris 3D features appear the most robust to topology changes (H2) and holes (H1) of strength 5. Table 9 summarizes the best performing feature description algorithms.

Feature descriptors. Tables 10–12 and 13–15 show the average normalized L_2 error in descriptor vectors at corresponding points for sparse and dense feature description approaches compared in this benchmark, respectively. Heat kernel signatures show the best results among the compared algorithms. Among sparse descriptors (SHK1–2 and SI), the best results in average repeatability and local scale, and sampling classes are achieved by SHK1; in micro holes and scale

	Strength				
Transform.	**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

Table 1: *Repeatability of HK1: heat kernel based feature detection algorithm. Average number of detected points: 23.*

	Strength				
Transform.	**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 2: *Repeatability of HK2: heat kernel based feature detection algorithm. Average number of detected points: 9.*

	Strength				
Transform.	**1**	**≤2**	**≤3**	**≤4**	**≤5**
Isometry	90.47	91.94	91.71	91.88	92.10
Topology	90.33	90.21	89.93	89.97	89.82
Holes	89.59	89.41	89.25	88.82	88.49
Micro holes	90.42	90.40	90.36	90.33	90.31
Scale	92.21	91.61	90.67	89.55	88.19
Local scale	88.08	86.49	83.64	80.99	78.98
Sampling	84.81	84.80	82.37	78.76	70.68
Noise	89.27	87.36	83.20	79.76	74.53
Shot noise	90.73	90.84	89.43	87.94	86.37
Average	89.55	89.23	87.84	86.44	84.38

Table 3: *Repeatability of H1: Harris 3D feature detection algorithm. Average number of detected points: 303.*

	Strength				
Transform.	**1**	**≤2**	**≤3**	**≤4**	**≤5**
Isometry	89.66	91.49	90.99	91.70	91.90
Topology	90.08	90.11	90.02	89.80	89.83
Holes	88.44	88.47	88.41	88.12	87.58
Micro holes	90.12	90.20	90.20	90.10	89.94
Scale	90.58	90.22	90.05	89.61	89.24
Local scale	89.71	89.14	87.38	84.95	82.77
Sampling	88.38	85.05	81.30	76.03	72.03
Noise	88.38	78.44	70.19	65.13	62.21
Shot noise	89.39	88.56	87.38	86.11	84.48
Average	89.42	87.96	86.21	84.62	83.33

Table 4: *Repeatability of H2: Harris 3D feature detection algorithm. Average number of detected points: 303.*

	Strength				
Transform.	**1**	**≤2**	**≤3**	**≤4**	**≤5**
Isometry	81.97	85.23	84.32	85.59	85.67
Topology	82.05	82.21	82.46	82.20	82.06
Holes	80.30	80.58	81.07	81.24	80.91
Micro holes	82.19	82.50	82.52	82.28	82.19
Scale	84.09	83.65	82.45	80.84	79.43
Local scale	81.72	79.00	75.07	71.60	68.23
Sampling	77.36	71.78	66.51	60.50	57.06
Noise	80.25	63.55	55.08	51.69	48.93
Shot noise	81.62	79.73	77.18	74.44	71.25
Average	81.28	78.69	76.30	74.49	72.86

Table 5: *Repeatability of H3: Harris 3D feature detection algorithm. Average number of detected points: 151.*

	Strength				
Transform.	**1**	**≤2**	**≤3**	**≤4**	**≤5**
Isometry	64.48	70.10	71.73	73.05	73.27
Topology	65.52	64.13	63.05	62.32	62.60
Holes	64.80	64.56	64.28	63.49	62.85
Micro holes	50.84	45.54	41.29	37.80	35.00
Scale	73.33	70.06	67.53	65.18	63.50
Local scale	65.48	60.99	59.82	57.81	54.09
Sampling	66.15	66.41	65.80	63.45	60.35
Noise	70.39	65.19	62.17	58.76	56.61
Shot noise	63.44	51.07	42.07	36.10	33.91
Average	64.94	62.00	59.75	57.55	55.80

Table 6: *Repeatability of SP1: salient points feature detection algorithm. Average number of detected points: 88.*

	Strength				
Transform.	**1**	**≤2**	**≤3**	**≤4**	**≤5**
Isometry	79.01	83.50	83.90	84.33	84.79
Topology	78.10	77.26	77.00	77.18	77.35
Holes	79.55	79.34	78.64	77.68	76.50
Micro holes	59.07	53.36	49.60	46.42	43.74
Scale	84.68	82.36	80.77	78.98	77.42
Local scale	78.91	77.68	76.37	74.39	71.98
Sampling	76.39	75.58	74.15	72.02	69.09
Noise	83.45	80.36	77.73	75.41	73.08
Shot noise	77.78	73.31	66.06	62.25	59.68
Average	77.44	75.86	73.80	72.07	70.40

Table 7: *Repeatability of SP2: salient points feature detection algorithm. Average number of detected points: 205.*

	Strength				
Transform.	**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

Table 8: *Repeatability of SP3: salient points feature detection algorithm. Average number of detected points: 409.*

	Strength				
Transform.	**≤ 1**	**≤ 2**	**≤ 3**	**≤ 4**	**≤ 5**
Isometry	HK2	HK2	HK2	HK2	HK2
Topology	HK1	HK1	HK1	HK1	H2
Holes	HK1	HK1	H1	H1	H1
Micro holes	HK2	HK2	HK2	HK2	HK2
Scale	HK2	HK2	HK2	HK2	HK2
Local scale	HK1	HK2	HK2	HK2	HK2
Sampling	HK2	HK2	HK2	HK2	HK2
Noise	HK2	HK2	HK2	HK2	HK1
Shot noise	HK2	HK1	HK1	HK1	HK1
Average	HK2	HK1	HK1	HK1	HK1

Table 9: *Winning feature detection algorithms across transformation classes and strengths. H1–3=Harris 3D features, HK1–2=heat kernel features.*

	Strength				
Transform.	**1**	**≤2**	**≤3**	**≤4**	**≤5**
Isometry	0.05	0.04	0.04	0.04	0.04
Topology	0.05	0.06	0.12	0.14	0.19
Holes	0.07	0.07	0.07	0.08	0.09
Micro holes	0.05	0.05	0.06	0.06	0.06
Scale	0.05	0.05	0.05	0.05	0.05
Local scale	0.07	0.09	0.10	0.12	0.14
Sampling	0.06	0.06	0.06	0.08	0.13
Noise	0.08	0.09	0.11	0.12	0.13
Shot noise	0.05	0.06	0.10	0.16	0.25
Average	0.06	0.06	0.08	0.09	0.12

Table 10: *Robustness of SHK1: heat kernel signature feature description algorithm based on featured detected by HK1 (average L_2 distance between descriptors at corresponding points). Average number of points: 23.*

	Strength				
Transform.	**1**	**≤2**	**≤3**	**≤4**	**≤5**
Isometry	0.04	0.03	0.04	0.04	0.04
Topology	0.04	0.06	0.11	0.13	0.18
Holes	0.06	0.07	0.08	0.08	0.09
Micro holes	0.04	0.04	0.05	0.05	0.05
Scale	0.04	0.04	0.04	0.04	0.04
Local scale	0.07	0.08	0.10	0.13	0.16
Sampling	0.05	0.05	0.05	0.07	0.14
Noise	0.08	0.09	0.11	0.12	0.13
Shot noise	0.05	0.08	0.15	0.24	0.31
Average	0.05	0.06	0.08	0.10	0.13

Table 11: *Robustness of SHK2: heat kernel signature feature description algorithm based on featured detected by HK2 (average L_2 distance between descriptors at corresponding points). Average number of points: 9.*

the best results are of SHK2; in isometry, holes, and noise classes, SHK1 and SHK2 perform the same; and spin image (SI) feature descriptor performed the best in topology and shot noise classes. Among dense descriptors (DHK1–3), DHK1 and DHK2 show equal average performance, with FEM-based descriptor (DHK2) being slightly better in the topology, local scale, sampling, and noise classes; the scale-invariant heat kernel signatures (DHK3) perform the best in the scale class. Overall, the top performing descriptor is DHK2, and summarized in Table 16.

7. Conclusions

Among the compared feature detection algorithms, heat kernel-based methods show the highest overall repeatability. These methods also perform the best in most of the transformation classes, excepting topology and holes, where 3D Harris feature detector shows the best results. Among

	Strength				
Transform.	**1**	**≤2**	**≤3**	**≤4**	**≤5**
Isometry	0.12	0.10	0.10	0.10	0.10
Topology	0.11	0.11	0.11	0.11	0.11
Holes	0.12	0.12	0.12	0.12	0.12
Micro holes	0.15	0.15	0.16	0.16	0.16
Scale	0.18	0.15	0.15	0.15	0.15
Local scale	0.12	0.13	0.14	0.15	0.17
Sampling	0.13	0.13	0.13	0.13	0.15
Noise	0.13	0.15	0.17	0.19	0.20
Shot noise	0.11	0.13	0.16	0.17	0.18
Average	0.13	0.13	0.14	0.14	0.15

Table 12: *Robustness of SI feature description algorithm based on features detected by SP2 (average L_2 distance between descriptors at corresponding points). Average number of points: 205.*

	Strength				
Transform.	**1**	**≤2**	**≤3**	**≤4**	**≤5**
Isometry	0.02	0.01	0.01	0.01	0.01
Topology	0.02	0.02	0.02	0.02	0.03
Holes	0.02	0.02	0.02	0.03	0.03
Micro holes	0.01	0.02	0.02	0.02	0.02
Scale	0.25	0.15	0.13	0.14	0.16
Local scale	0.02	0.04	0.05	0.08	0.11
Sampling	0.02	0.02	0.02	0.02	0.03
Noise	0.02	0.06	0.09	0.13	0.16
Shot noise	0.02	0.02	0.02	0.02	0.02
Average	0.04	0.04	0.04	0.05	0.06

Table 13: *Robustness of DHK1: dense heat kernel signature feature description algorithm using cotangent weight discretization (average L_2 distance between descriptors at corresponding points).*

	Strength				
Transform.	**1**	**≤2**	**≤3**	**≤4**	**≤5**
Isometry	0.01	0.01	0.01	0.01	0.01
Topology	0.02	0.02	0.02	0.02	0.02
Holes	0.02	0.02	0.02	0.03	0.03
Micro holes	0.01	0.01	0.01	0.01	0.02
Scale	0.25	0.15	0.13	0.14	0.16
Local scale	0.02	0.03	0.05	0.07	0.10
Sampling	0.02	0.02	0.02	0.02	0.02
Noise	0.03	0.06	0.09	0.12	0.15
Shot noise	0.01	0.01	0.02	0.02	0.02
Average	0.04	0.04	0.04	0.05	0.06

Table 14: *Robustness of DHK2: dense heat kernel signature feature description algorithm using FEM discretization (average L_2 distance between descriptors at corresponding points).*

	Strength				
Transform.	**1**	**≤2**	**≤3**	**≤4**	**≤5**
Isometry	0.09	0.07	0.07	0.07	0.07
Topology	0.10	0.10	0.11	0.11	0.11
Holes	0.09	0.10	0.12	0.13	0.14
Micro holes	0.09	0.09	0.09	0.09	0.09
Scale	0.12	0.10	0.10	0.10	0.09
Local scale	0.10	0.13	0.15	0.19	0.22
Sampling	0.09	0.09	0.10	0.10	0.12
Noise	0.11	0.16	0.19	0.21	0.24
Shot noise	0.09	0.09	0.09	0.09	0.10
Average	0.10	0.10	0.11	0.12	0.13

Table 15: *Robustness of DHK3: dense scale-invariant heat kernel signature feature description algorithm using cotangent weight discretization (average L_2 distance between descriptors at corresponding points).*

	Strength				
Transform.	**≤ 1**	**≤ 2**	**≤ 3**	**≤ 4**	**≤ 5**
Isometry	DHK2	DHK2	DHK2	DHK2	DHK2
Topology	DHK2	DHK2	DHK2	DHK2	DHK2
Holes	DHK2	DHK2	DHK2	DHK2	DHK1
Micro holes	DHK2	DHK2	DHK2	DHK2	DHK2
Scale	SHK2	SHK2	SHK2	SHK2	SHK2
Local scale	DHK2	DHK2	DHK2	DHK2	DHK2
Sampling	DHK2	DHK2	DHK2	DHK2	DHK2
Noise	DHK1	DHK1	DHK2	SHK1	SHK2
Shot noise	DHK2	DHK2	DHK2	DHK2	DHK2
Average	DHK2	DHK2	DHK2	DHK2	DHK2

Table 16: *Winning feature description algorithms across transformation classes and strengths. DHK1=dense heat kernel signatures using cotangent weight discretization, SHK2=sparse heat kernel signatures based on features detected by HK2.*

the compared feature description algorithms, the best results were achieve by heat kernel-based methods.

A more detailed version of this report presenting additional details and experiments will be published separately.

References

[ABADB07] ASSFALG J., BERTINI M., A. DEL BIMBO P. P.: Content-based retrieval of 3d objects using spin image signatures. *IEEE Transactions on Multimedia 9*, 3 (2007), 589–599. 5

[ABC04] ANDREETTO M., BRUSCO N., CORTELAZZO G. M.: Automatic 3-d modeling of textured cultural heritage objects. *IEEE Transactions on Image Processing 13*, 3 (2004), 335–369. 5

[BBK08] BRONSTEIN A. M., BRONSTEIN M. M., KIMMEL R.: *Numerical geometry of non-rigid shapes*. Springer, 2008. 2

[BBOG10] BRONSTEIN A. M., BRONSTEIN M. M., OVSJANIKOV M., GUIBAS L. J.: ShapeGoogle: geometric words and expressions for invariant shape retrieval. *TOG (in review)* (2010). 4

[BK10] BRONSTEIN M. M., KOKKINOS I.: Scale-invariant heat kernel signatures for non-rigid shape recognition. In *Proc. CVPR* (2010). 5

[BSW08] BELKIN M., SUN J., WANG Y.: Discrete Laplace operator on meshed surfaces. In *Proc. SCG* (2008), pp. 278–287. 3

[CCFM08] CASTELLANI U., CRISTANI M., FANTONI S., MURINO V.: Sparse points matching by combining 3D mesh saliency with statistical descriptors. *Computer Graphics Forum 27* (2008), 643–652. 2, 4

[CGOS09] CHAZAL F., GUIBAS L. J., OUDOT S. Y., SKRABA P.: Analysis of scalar fields over point cloud data. In *Proc. SODA* (2009), pp. 1021–1030. 3

[ELZ00] EDELSBRUNNER H., LETSCHER D., ZOMORODIAN A.: Topological persistence and simplification. In *Proc. IEEE Foundations of Computer Science* (2000), pp. 454–463. 3

[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. 3

[HS88] HARRIS C., STEPHENS M.: A combined corner and edge detection. In *Proc. Fourth Alvey Vision Conference* (1988), pp. 147–151. 3

[IB10] ISIPIRAN I., BUSTOS B.: Robust 3D Harris operator. submitted, 2010. 2, 3

[IKN98] ITTI L., KOCH C., NIEBUR E.: A model of saliency-based visual attention for rapid scene analysis. *Trans. PAMI 20*, 11 (1998). 4

[JH99] JOHNSON A., HEBERT M.: Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence 21*, 5 (1999), 433–449. 4, 5

[Low04] LOWE D.: Distinctive image features from scale-invariant keypoints. *IJCV 60*, 2 (2004), 91–110. 1, 4

[MS05] MIKOLAJCZYK K., SCHMID C.: A performance evaluation of local descriptors. *Trans. PAMI* (2005), 1615–1630. 1

[OBG09] OVSJANIKOV M. BRONSTEIN A. M., BRONSTEIN M. M., GUIBAS L. J.: ShapeGoogle: a computer vision approach to invariant shape retrieval. *Proc. NORDIA* (2009). 4

[PSF10] PATANÉ G., SPAGNUOLO M., FALCIDIENO B.: Multi-scale feature spaces for shape analysis and processing. In *Proc. Shape Modeling International (SMI)* (2010). to appear. 4

[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). 2, 4

[SZ03] SIVIC J., ZISSERMAN A.: Video Google: A text retrieval approach to object matching in videos. In *Proc. ICCV* (2003), vol. 2, pp. 1470–1477. 1

[TCF09] TOLDO R., CASTELLANI U., FUSIELLO A.: Visual vocabulary signature for 3D object retrieval and partial matching. In *Proc. Eurographics Workshop on 3D Object Retrieval* (2009). 4