Eurographics Workshop on 3D Object Retrieval (2017)
I. Pratikakis, F. Dupont, and M. Ovsjanikov (Editors)

Shrec'17 Track: Retrieval of surfaces with similar relief patterns

S. Biasotti[1*], E. Moscoso Thompson[1*], M. Aono[5], A. Ben Hamza[9], B. Bustos[7], S. Dong[4], B. Du[4], A. Fehri[3], H. Li[4], F. A. Limberger[8], M. Masoumi[9], M. Rezaei[9], I. Sipiran[6], L. Sun[4], A. Tatsuma[5], S. Velasco Forero[3], R. C. Wilson[8], Y. Wu[4], J. Zhang[4], T. Zhao[4], F. Fornasa[2*] and A. Giachetti[2*]

[1] CNR - IMATI
[2] Dept. of Computer Science, University of Verona
[3] PSL Research University - MINES ParisTech, CMM, Center for Mathematical Morphology
[4] Beijing Technology and Business University
[5] Toyohashi University of Technology
[6] Department of Engineering - Pontifical Catholic University of Peru
[7] Dept. of Computer Science, University of Chile
[8] Dept. of Computer Science, University of York
[9] Concordia University, Montreal, QC, Canada
∗ Track organizers

Abstract

This paper presents the results of the SHREC'17 contest on retrieval of surfaces with similar relief patterns. The proposed task was created in order to verify the possibility of retrieving surface patches with a relief pattern similar to an example from a database of small surface elements. This task, related to many real world applications, requires an effective characterization of local "texture" information not depending on patch size and bending. Retrieval performances of the proposed methods reveal that the problem is not quite easy to solve and, even if some of the proposed methods demonstrate promising results, further research is surely needed to find effective relief pattern characterization techniques for practical applications.

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

1. Introduction

The challenge of this SHREC'17 track is to evaluate the performance of existing retrieval algorithms when a 3D surface is characterized by different relief patterns, such as knitted fabrics. This task is particularly challenging and interesting for practical applications, for example to retrieve artworks' patterns, artists' styles or to classify natural structures like tree barks [OVSP13], rock types or engravings [ZPS*16], etc. A peculiar characteristic of patterns is the fact their style does not depend by the overall structure of the shape rather than it identifies parts and local properties that are independent of the global shape.

The key point in this retrieval task is to find a patch descriptor and a distance measure that do not depend on the patch bending and are robust to different mesh samplings, but only on the relief pattern characterizing its surface, that can have different size and a depth of a few millimeters.

2. Data acquisition and processing

15 different textiles (with well defined and specific relief patterns) have been placed in different poses possibly lying on differently shaped objects and stretched, so that the same surface assumes different space embeddings (examples of different embeddings of the same patch are shown in Figure 1.

Then, their surface has been acquired with an Intel Realsense F200 depth sensor and the 3D System Sense app. [sen]. Exported meshes, without texture, have been processed to remove inconsistencies and 12 patches of approximately $100-200cm^2$ have then been obtained for each textile class. With this procedure 180 models have been created. These models are also referred as *original* surfaces.

To each patch, three processing operations (two adaptive simplifications to 10K and 5K vertices and one re-sampling operation to 15K vertices) have been applied to each scan. All these transformations have been designed to alter the mesh connectivity of the original scans, see Figure 2. These processing operations were automatically done using the ReMESH software [AF06]. Then, all dangling edges and self-intersections have been manually removed, so that the final patches are oriented, made of a single connected component and locally a regular. At the end of these processing operations, the full dataset is made of 720 models (180 original

DOI: 10.2312/3dor.20171058

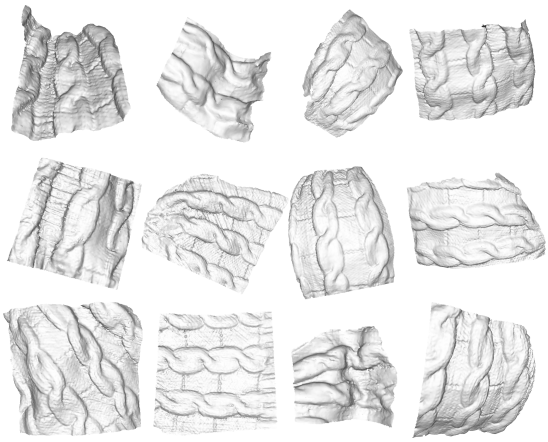

Figure 1: *Renderings of the 12 patches acquired and cropped for a single tissue class.*

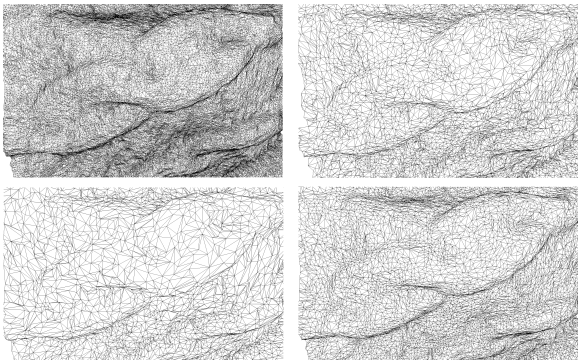

Figure 2: *Detail of an original mesh scan and its three sampling variations.*

surfaces plus 540 tessellation variations). Figure 3 shows a representative patch for each class of patterns.

To test the methods and tune the parameters, the participants were provided with a set of examples. Such a set contains three sample patterns in different poses with four different samplings on the mesh.

3. Evaluation

Participants were asked to send dissimilarity matrices (of size 720x720) for the whole dataset. The challenge is to retrieve the patches that are closest to a given query. To analysis how methods are really independent of the spatial bending and robust to different tessellations, we evaluated the the retrieval performance both on the whole dataset (720 models) and on the subset of 180 scans (the similarity scores over the 180 models are derived from the matrices submitted The retrieval performances of the proposed methods were then evaluated according to the classical measures used in [SMKF04], e.g. Nearest Neighbor (NN), First Tier (FT), Second Tier (ST), e-measure (E) and Discounted Cumulated Gain (DCG). Furthermore, Precision-Recall plots have been analyzed and from the PR curves the Mean Average Precision (MAP), e.g. the average of all precision values computed for each subject in the retrieved list was estimated. As a further evaluation measure, we consider the confusion matrices obtained from the NN classification deduced from the similarity matrices.

4. Participants and methods

Ten groups subscribed to the contest but only **six** finally submitted their contributions with (up to three) dissimilarity matrices. A total of **14** runs/matrices have been evaluated and compared, including one created with a simple baseline method.

In the following sections we present all the participants and proposed methods.

4.1. Statistics of Local Features Extracted from Local Binary Pattern Image (LBPI), by Atsushi Tatsuma and Masaki Aono

In order to capture texture features of surface object, authors estimate statistics of local features extracted from the depth-buffer image as a shape descriptor of the surface object. To emphasize the texture of surface object, they convert the depth-buffer image into the Local Binary Pattern (LBP) [OPH94] image.

Figure 4 shows the overview of our feature extraction process. As a preprocessing, authors perform the normalization of the position and scale for a given surface object. They translate the mean of the vertices to the origin, and normalize the size of the surface object to the unit sphere.

After normalization of the surface object, they render a depth-buffer image from the viewpoint defined by the mean of the face-normal vectors. To emphasize the texture of the surface object represented with depth values, they convert the depth-buffer image into the LBP image. And then, they extract local features from the LBP image. For feature detection and description algorithm, the KAZE features [ABD12] are employed.

To obtain the shape descriptor of the surface object, authors calculate the mean, covariance matrix, skewness, and kurtosis of the local features. Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$ be the local features of size N. The mean μ, covariance matrix C, skewness s, and kurtosis k are calculated as follows [Mar70]:

$$\mu = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i,$$

$$C = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}_i-\mu)(\mathbf{x}_i-\mu)^\top,$$

$$s = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\{(\mathbf{x}_i-\mu)^\top C^{-1}(\mathbf{x}_i-\mu)\}^3,$$

$$k = \frac{1}{N}\sum_{i=1}^{N}\{(\mathbf{x}_i-\mu)^\top C^{-1}(\mathbf{x}_i-\mu)\}^2.$$

Since the covariance matrix C lies on the Riemannian manifold of symmetric positive semi-define matrices, the method used maps the covariance matrix onto a point in the Euclidean space by using Pennec et al.'s method [PFA06].

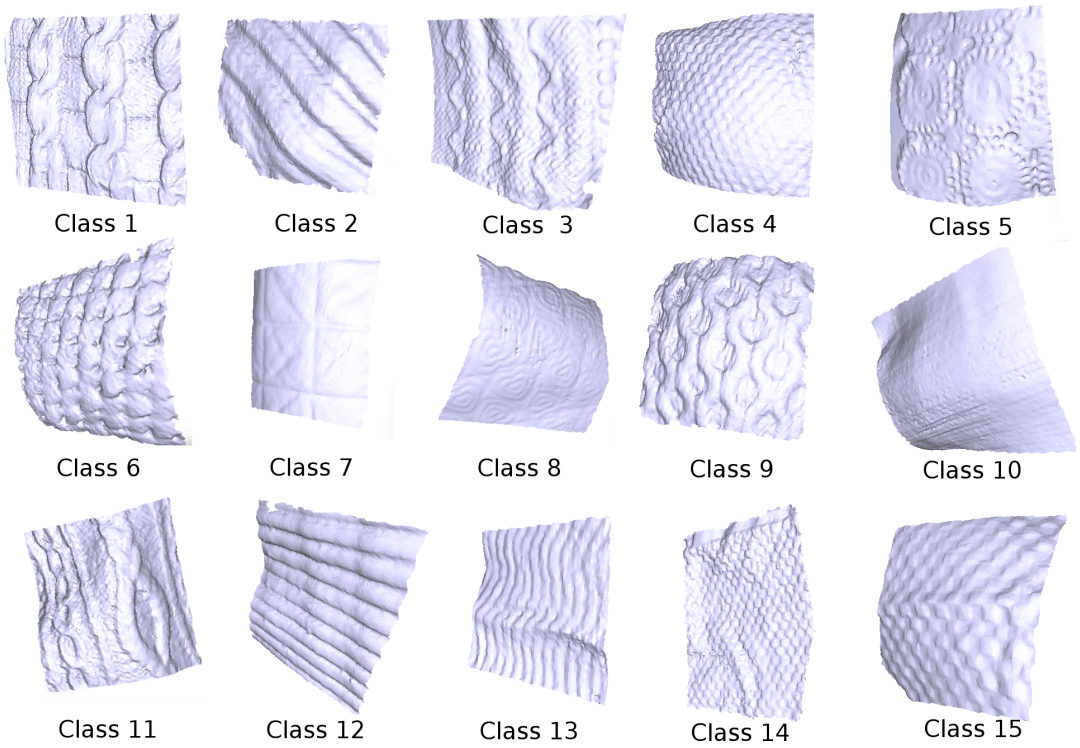

Figure 3: *Example meshes acquired on the 15 tissue classes.*

The final shape descriptor is obtained by concatenating the mean, covariance, skewness, and kurtosis of the local features. The shape descriptor is normalized with the signed square rooting and ℓ_2 normalization [JC12]. For the dissimilarity between two shape descriptors, the Euclidean distance is employed.

4.2. Kinetic Laplace-Beltrami operator, by Frederico A. Limberger and Richard C. Wilson (KLBO-FV-IWKS/KLBO-SV-IWKS)

The key idea of this method is to test a curvature-based Laplace-Beltrami operator (KLBO) to describe the relief patterns of surfaces. After computing the eigendecomposition of the KLBO, the Improved Wave Kernel signature (IWKS) [LW15] is computed and encoded using two different encoding schemes: the Fisher Vector (FV) or the Super Vector (SV). Lastly, differences between encodings are computed using Euclidean distance, after reducing feature dimensionality by computing PCA. In the following, we explain in more details each one of these steps.

The KLBO is a method which computes the underlying spectral components of 3D meshes using a curvature-based kinetic term, which removes the influence of shape's articulations. By defining the Lagrangian of the dynamics over the object's surface using classical field theory (1), it is possible to weight the physical field using a smooth positive kinetic density. The Lagrangian density of the system is given by:

$$\mathcal{L}(\phi, \bigtriangledown\phi, \dot{\phi}, \mathbf{x}, t) = T - V \tag{1}$$

where T is the kinetic energy and V is the potential energy, and ϕ represents a field defined over the surface of the object. By weighting T in the Euler-Lagrange equation, which is given by the action of the system ($S(\mathcal{L}) = \int \mathcal{L} d\mathbf{x} dt$) it is possible to weight the kinetic energy differently across the field. After computing all respective derivations, it leads to the following generalized eigenproblem:

$$W\phi = \lambda A K \phi \tag{2}$$

where K is a diagonal matrix where the diagonal element K_{ii} is the respective mean curvature H_i of the vertex v_i. This formulation is similar to the already known generalized eigenproblem ($W\phi = \lambda A\phi$), where W is also given by

$$W(i,j) = \begin{cases} \frac{(\cot\alpha_{ij} + \cot\beta_{ij})}{2} & \text{if } (i,j) \in E, \\ \sum_{k \neq i} W(i,k) & \text{if } i = j. \end{cases} \tag{3}$$

and A is a diagonal matrix where A_{ii} represents the area of the Voronoi diagram around v_i. Differently from the standard approach, the proposed formulation is less variant to non-rigid motions since articulated regions are downweighted by the curvature-based kinetic term. Therefore, the knitted fabrics or other materials do not need to be bended in a specific way.

After computing the shape spectrum (λ, ϕ) using (2), participants compute the IWKS which is a spectral descriptor based on the Wave Kernel Signature [ASC11]. To encode the local spectral descriptors they use two methods: the Fisher Vector (KLBO-FV-IWKS run) and the Super Vector (KLBO-SV-IWKS run). For more details about them and how to encode local features using these

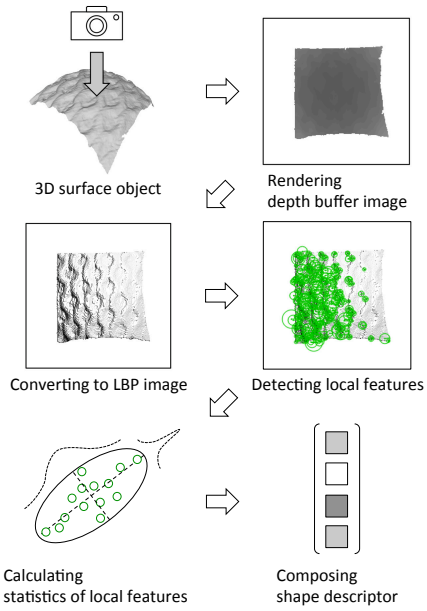

Figure 4: *Overview of feature extraction process based on statistics of local features.*

methods we refer to reader to [LW15]. Distances between signatures are computed using Euclidean distance, but first, feature dimensionality is reduced to 50 by computing PCA of the encodings.

Parameters and Performances: Authors computed the first 300 eigenvalues and eigenfunctions of the KLBO using the cotangent scheme method with the kinetic-energy modifications. Then, they computed the IWKS with 100 frequencies using *iwksvar* = 5. For computing the dictionary, they used the first 29 models of the database to create GMMs with 38 components for each signature frequency. All experiments were performed in Matlab in a Windows PC (Intel Core i7 3.4GHz, 8GB RAM). In the beginning, all meshes were resampled to 10K faces. It takes, in average, 7 seconds to resample and compute the every model signature. The time for computing all IWKS descriptors, including the KLBO eigendecomposition, was 82 minutes. The times for computing either the FV or SV encoding for all signatures was 105 seconds and 410 seconds, respectively.

4.3. Signature Quadratic Form by Ivan Siprian and Benjamin Bustos (SQFD)

This method consists of applying the Signature Quadratic Form Distance (SQFD) [BUS09] along with intrinsic spectral descriptors. On the one hand, the SQFD distance is a suitable and effective alternative to compare 3D objects represented as a collection of local features [SLBS16]. On the other hand, spectral features have proven to be robust against several transformations while keeping discriminative geometric information. This proposal combines these two methods in order to represent and assess the similarity between relief patterns.

Let M be a 3D shape. First a set of local descriptors is computed on M. The computation of spectral descriptors depends, in turn, on the computation of the Laplace-Beltrami operaror and its eigendecomposition. As the complexity of this decomposition is directly related to the number of vertices of the input mesh, input models are simplified to have approximately 20,000 vertices by using the QSlim algorithm [GH97].

A local descriptor for each vertex is thus estimated in the simplified mesh. We denote the set of local descriptors of the mesh M as F_M. The approach to use the SQFD distance establishes that we need to compute a more compact representation called *signature*. Let us suppose the existence of a local clustering on F_M that groups similar local descriptors such that the number of clusters is n and $F_M = C_1 \bigcap C_2 \bigcap \ldots C_n$. The signature is defined as $S^M = \{(c_i^M, w_i^M), i = 1, \ldots, n\}$, where $c_i^M = \frac{\sum_{d \in C_i} d}{|C_i|}$ and $w_i^M = \frac{|C_i|}{|F_M|}$. Each element in the signature contains the average descriptor in the cluster (c_i^M) and a weight (w_i^M) to quantify how representative is the cluster in the collection of local descriptors.

The local clustering is a key ingredient of the computation of the signatures. Participants use an adaptive clustering method that searches groups of descriptors using two distance penalties. The method uses an intra-cluster penalty λ that accounts for the maximum distance between descriptors in the same cluster. Also, the method uses an inter-cluster penalty β that accounts for the minimum distance between centroids of different clusters. In addition, the clustering method only preserves clusters with a number of descriptors greater than a parameter N_m. More details can be found in [SLBS16].

Given two objects M and N, and their respective signatures S^M and S^N, the Signature Quadratic Form Distance is defined as

$$SQFD(S^M, S^N) = \sqrt{(w^M | -w^N) \cdot A_{sim} \cdot (w^M | -w^N)^T} \quad (4)$$

where $(w^M | w^N)$ denotes the concatenation of two weight vectors. The matrix A_{sim} is a block similarity matrix that stores the correlation coefficients between clusters. To transform a distance between cluster centroids to a correlation coefficient, it is necessary to apply a similarity function and the Gaussian similarity function

$$sim(c_i, c_j) = \exp(-\alpha d^2(c_i, c_j)). \quad (5)$$

was chosen. Furthermore, it is necessary to choose the value of parameter α and the ground distance for descriptors. In all the proposed experiments, the choice were $\alpha = 0.9$ and L^2 as ground distance. More details about the computation of signatures and the SQFD distance can be found in [SLBS16].

Experimental Settings

Authors proposed three runs using different configurations of the method. Here we describe the parameters used in each run

- **SQFD(WKS).** Participants used the normalized Wave Kernel Signature [ASC11] as local descriptor. The parameters for local clustering were $\lambda = 0.2$, $\beta = 0.4$, $N_m = 30$.

- **SQFD(HKS).** The normalized Heat Kernel Signature [SOG09] was used as local descriptor. The parameters for local clustering were $\lambda = 0.2$, $\beta = 0.4$, $N_m = 30$.
- **SQFD(SIHKS).** The Scale-invariant Heat Kernel Signature [BK10] was used as local descriptor. The parameters for local clustering were $\lambda = 0.01$, $\beta = 0.02$, $N_m = 20$.

4.4. Covariance descriptor by morphological analysis of curvature estimations by Santiago Velasco-Forero and Amin Fehri

The proposed method computes an image covariance descriptor from morphological transformation of local estimation of curvature for a given 3-dimensional mesh $\mathcal{S}$. A scheme to illustrated the description of the method is given in Fig. 5. Four main components are used to compute this descriptor:

1. First, the local principal curvatures, κ_1, κ_2 and the *Gaussian curvature*, $\kappa_G = \kappa_1, \kappa_2$ are computed [CSM03, ACSD*03]. The estimation of κ_1 and κ_2 is performed by the largest and smallest eigenvalues of a smoothed estimation of local normal sections [Pey11]. Authors fixed 11 as smoothing parameter.
2. Second, curvature values on the three-dimensional surface are projected to a flat surface (two dimensional). Accordingly, the *boundary* of the mesh, i.e. the set of vertices that are only referenced by a single triangle in the mesh was found. This set is denoted by $\texttt{Boundary}(\mathcal{S})$. Then, a support vector machine (SVM) was trained in a regression problem from the 3D coordinate of $\texttt{Boundary}(\mathcal{S})$ to a discrete square of value from $[0,256] \times [0,256]$, i.e. $T : \mathbb{R}^3 \mapsto \mathbb{Z}^2$. In this way it is possible to represent the curvature information $(\kappa_1, \kappa_2, \kappa_G)$ of $\mathcal{S}$ by three images.
3. At this point, it is possible to take advantage of image processing texture descriptors to characterize the original 3D surface. Participants used mathematical morphology operators in this task [Soi13]. For binary or grey-scale images, they are simple in the sense that they usually have an intuitive interpretation. Erosion $\varepsilon_{\texttt{SE}}(\mathbf{I})$ (or minimum operator in a neighborhood) shrinks bright objects, whereas dilation $\delta_{\texttt{SE}}(\mathbf{I})$ (or maximum operator in a neighborhood) expands bright objects at the boundary. The size and shape of the effect is controlled by the structuring element SE (neighborhood). The morphological *opening* $\gamma_{\texttt{SE}}(\cdot)$ is an idempotent transformation defined by composition of erosion and dilation, i.e. $\gamma_{\texttt{SE}}(\mathbf{I}) = \delta_{\texttt{SE}}(\varepsilon_{\texttt{SE}}(\mathbf{I}))$. As the goal is to characterize the local texture of the curvature, rather than finding an optimal fixed value of size of SE, authors focus on a range of values for the scale parameter. Thus, they produce 32 transformations of each curvature image by considering: 1) $\gamma_{\texttt{SE}_\circ}$ operator by considering $\texttt{SE}_\circ$ as disks with radius from 1 to 16 2) The maximum of $\gamma_{\texttt{SE}_0}, \gamma_{\texttt{SE}_{45}}, \gamma_{\texttt{SE}_{90}}$ where $\texttt{SE}_\texttt{x}$ a discrete line of angle x and radius from 1 to 32 by 2. In conclusion, they have 96 images from each 3-dimensional mesh $\mathcal{S}$.
4. Finally, the covariance matrix of 96 images is computed, to produce a descriptor as a square matrix of size 96×96 in the sense of [TPM06].
5. The similarity between two meshes is calculated via their representation as covariance matrices by using the following measures:

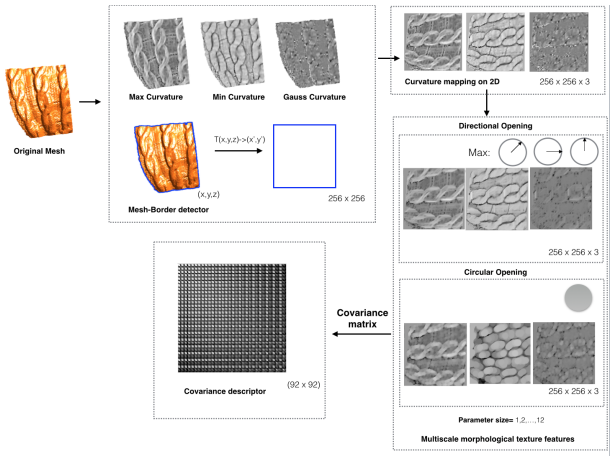

Figure 5: *Proposed method includes three local curvature estimations, a projection from $T : \mathbb{R}^3 \mapsto \mathbb{Z}^2$ and a set of morphological transformation by opening. The final descriptor is covariance matrix of size 96×96. Details are included in Section 4.4.*

   a. The *cosine similarity (CMC-1)* between the upper triangular matrix with diagonal, i.e., for two covariance matrices,Σ_i and Σ_j, $D_1(\Sigma_i, \Sigma_j) := \frac{\texttt{triu}(\Sigma_i)^T \texttt{triu}(\Sigma_i)}{||\texttt{triu}(\Sigma_i)|| ||\texttt{triu}(\Sigma_i)||}$
   b. The *von Neuman divergence (CMC-2)* for two covariance matrices [DT08], Σ_i and Σ_j, $D_2(\Sigma_i, \Sigma_j) := \text{tr}(\Sigma_1 \log \Sigma_i - \Sigma_i \log \Sigma_j - \Sigma_i + \Sigma_j)$
   c. The *Euclidean distance (CMC-3)* between the upper triangular matrix with diagonal, i.e., for two covariance matrices, Σ_i and Σ_j, $D_3(\Sigma_i, \Sigma_j) := ||\texttt{triu}(\Sigma_i) - \texttt{triu}(\Sigma_i)||$

4.5. Interior Dihedral Angle Histogram (IDAH) and HOG of the Geometry images (HOG-GI), by Li Sun,Yujuan Wu, Junjie Zhang, Bowen Du, Tianyu Zhao, Shuilong Dong and Haisheng Li

Interior Dihedral Angle Histogram (IDAH) The models' convex/concave properties and local geometrical features can be well described by the interior dihedral angle of each edge of the mesh. Authors therefore propose a new statistical feature for 3D shape retrieval called Interior Dihedral Angle Histogram (IDAH). The process can be described as follows:

Firstly, they calculate all the interior dihedral angles of the model surface. Then, the distribution histogram is calculated in different intervals. Finally, they adopt the Manhattan distance between histograms to describe the similarity of models. Besides, they set two interval values, bin=130:10:230 and bin=105:3:255 (Unit of Measure: Degree), and the dimension of the feature vector is 11 and 51 respectively.

HOG of the Geometry images Based on paper [SBR16], authors converted the model into a *geometry image*. Firstly, they use authalic parametrization to parametrize a 3D model over a spherical domain. Then, the spherical parametrization is mapped onto an octahedron. Finally, they cut the octahedron along its edges to output

a flat and regular geometry image. They further use the VLFeat function vl_hog to extract the HOG features of geometry images to evaluate similarity between models.

4.6. Geodesic Multi-resolution (GMR) Approach by Majid Masoumi, Mahsa Rezaei, A. Ben Hamza

Authors use a geodesic multi-resolution descriptor [MLH16] by incorporating the vertex area into the definition of spectral graph wavelet [LB13] in a bid to capture more geometric information and, hence, further improve its discriminative ability. Moreover, Mexican hat wavelet has been utilized as a generating kernel, which considers all frequencies equally-important overall as opposed to the cubic spline kernel [LB13]. Furthermore, in order to capture the spatial relations between features, the proposed geodesic multi-resolution descriptor is weighted by geodesic exponential kernel. While the approach focuses primarily on 3D object retrieval, this framework is fairly general and can be used to address a variety of shape analysis problems, including segmentation and classification [MH17].

The geodesic multi-resolution framework consists of four main steps. The first step is to represent each 3D shape in the dataset $\mathcal{D}$ by a spectral graph wavelet signature matrix $\mathbf{S}$, where $\mathbf{S} = (\mathbf{s}_1, \ldots, \mathbf{s}_m) \in \mathbb{R}^{p \times m}$, and $\mathbf{s}_i$ is the p-dimensional local descriptor at vertex i and m is the number of mesh vertices. In the second step , the area-weighted spectral graph wavelet signatures $\mathbf{s}_i$ are mapped to high-dimensional mid-level feature vectors using the soft-assignment coding step of the BoF model, resulting in a $k \times m$ matrix $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_m)$ whose columns are the k-dimensional mid-level feature codes. In the third step, the $k \times k$ geodesic multi-resolution matrix $\mathbf{F}$ is computed using the mid-level feature codes matrix and a geodesic exponential kernel as follows:

$$\mathbf{F} = \mathbf{U}\mathbf{K}\mathbf{U}^{\mathsf{T}}, \tag{6}$$

where $\mathbf{U}$ is a $k \times m$ matrix of geodesic multi-resolution codes (i.e. mid-level features), and $\mathbf{K} = (\kappa_{ij})$ is an $m \times m$ geodesic exponential kernel matrix whose elements are given by

$$\kappa_{ij} = \exp\left(-\frac{d_{ij}^2}{\varepsilon}\right), \tag{7}$$

with d_{ij} denoting the geodesic distance between any pair of mesh vertices $\mathbf{v}_i$ and $\mathbf{v}_j$, and ε is a positive, carefully chosen parameter that determines the width of the kernel. Then $\mathbf{F}$ is reshaped into a k^2-dimensional global descriptor $\mathbf{x}_i$. In the fourth step, the geodesic multi-resolution vectors $\mathbf{x}_i$ of all n shapes in the dataset are arranged into a $k^2 \times n$ data matrix $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$. Finally, a query $\mathbf{x}$ is compared to all data points in $\mathbf{X}$ using ℓ_1-distance to find the most relevant shapes to the query. The lower the value of this distance is, the more similar the shapes are.

4.7. Baseline method: Curvature Histograms (CH)

As a baseline of the benchmark we have implemented the histogram of the minimal curvature. In our settings, for every surface, we computed the value of the minimal curvature on each vertex using an implementation of the method [CSM03]. Then, an histogram of 128 bins is kept as signature of the that surface, see Figure 6. As distance between two surfaces we adopt the Manhattan distance between the corresponding histograms. The histograms of curvature used as baseline are not multi-resolution neither store any adjacency of relationship between the elements of the surface.

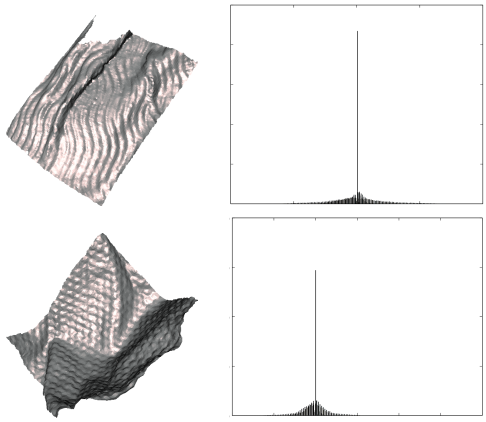

Figure 6: *Two surfaces (left) and the corresponding histograms of minimal curvature (right).*

5. Results

Retrieval scores obtained by the different runs of the proposed methods are shown in Table 1. It is possible to see that globally the performances appear rather poor. Sufficiently good results on first retrieval (NN) should not mislead, as they are clearly biased by the inclusion in the dataset of four different meshes representing the same acquired patch. If we estimate the same scores on a subset of the data with single instances of each patch (180 patches with 12 differently shaped patches for each class), this effect is removed (Table 5). The method that seems more effective in retrieving similar patterns independently on the background shape is therefore not the one with top performances on the whole dataset (KLBO-FV-IWKS), but CMC-2 that actually increases many retrieval scores on the single-instances dataset. This could mean that the method is not robust against remeshing.

KLBO-FV-IWKS demonstrates a sufficient ability in pattern retrieval, but, on the other hand, seems to include in the similarity estimation information on the global shape, not only on the relieved pattern. Note that even if a method correct large scale bending of a shape, a signature may depend on the global area and contour of the shape while in our task it is important to characterize statistically the relief pattern.

Other methods seem mostly to fail both in retrieve patches with similar shape and similar pattern.

To better understand the behavior of the methods, it is interesting to investigate the performances of the methods for the different classes of patterns. Figures 7 and 8 depict the Precision-Recall curves for the whole dataset of 720 patches and the original set of 180 models, respectively. Details of the Precision-Recall curves of some specific classes are shown in the Figures 9, 10 and 11.

The most interesting fact is that the performances of the meth-

	NN	1-Tier	2-Tier	e	DCG	mAP
CH	0.196	0.089	0.160	0.077	0.500	0.122
LBPI	0.828	0.248	0.400	0.232	0.697	0.283
IDAH-1	0.390	0.157	0.254	0.145	0.578	0.174
IDAH-2	0.306	0.141	0.244	0.127	0.559	0.163
GI HOG	0.686	0.107	0.176	0.102	0.561	0.131
SQFD-SIHKS	0.168	0.127	0.239	0.106	0.532	0.158
SQFD-HKS	0.536	0.117	0.192	0.110	0.558	0.139
SQFD-WKS	0.510	0.112	0.191	0.102	0.549	0.136
CMC-1	0.718	0.258	0.372	0.247	0.673	0.260
CMC-2	0.763	0.272	0.389	0.261	0.686	0.271
CMC-3	0.647	0.219	0.323	0.208	0.639	0.218
KLBO-FV-IWKS	**0.986**	**0.333**	**0.449**	**0.332**	**0.759**	**0.339**
KLBO-SV-IWKS	0.978	0.287	0.409	0.283	0.732	0.296
GMR	0.079	0.066	0.128	0.054	0.474	0.113

Table 1: *Retrieval scores of the different runs proposed on the full set of surface patches. NN values are biased by the remeshing of the same patches.*

	NN	1-Tier	2-Tier	e	DCG	mAP
CH	0.206	0.132	0.222	0.145	0.437	0.165
LBPI	0.339	0.207	0.353	0.237	0.518	0.250
IDAH-1	0.272	0.163	0.262	0.175	0.480	0.201
IDAH-2	0.339	0.182	0.271	0.181	0.503	0.215
GI HOG	0.089	0.069	0.130	0.097	0.373	0.118
SQFD-SIHKS	0.167	0.111	0.213	0.157	0.423	0.163
SQFD-HKS	0.106	0.066	0.137	0.102	0.376	0.123
SQFD-WKS	0.067	0.073	0.148	0.108	0.380	0.129
CMC-1	0.600	0.342	0.461	0.274	0.641	0.371
CMC-2	**0.633**	**0.363**	**0.494**	**0.293**	**0.662**	**0.390**
CMC-3	0.533	0.281	0.394	0.242	0.596	0.308
KLBO-FV-IWKS	0.522	0.295	0.412	0.247	0.603	0.307
KLBO-SV-IWKS	0.489	0.249	0.375	0.235	0.570	0.273
GMR	0.061	0.061	0.119	0.089	0.364	0.112

Table 2: *Retrieval scores on the partial dataset with single instances of each patch.*

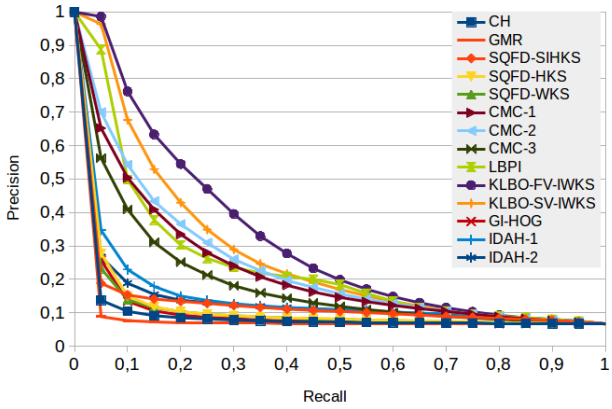

Figure 7: *Precision vs. Recall plot for all the proposed runs on the full dataset.*

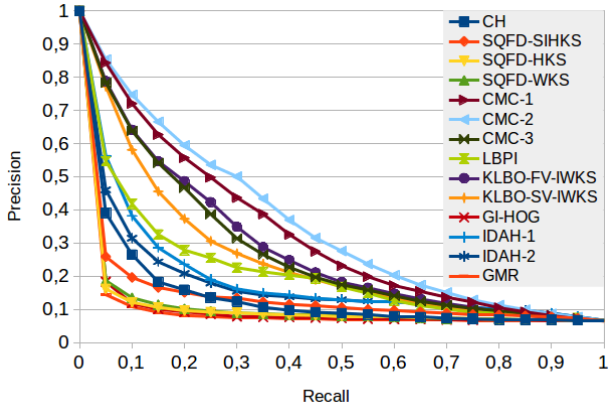

Figure 8: *Precision vs. Recall plot for all the proposed runs on the single patch instances dataset. Overall results are similar, but CMC behaves now better than KLBO and NN performances are decreased.*

ods are not similarly ranked for all the classes, as methods that are not successful on selected classes are often good on others. Figure 11 shows the Precision vs Recall plot for class 9 showing that for this pattern (regularly spaced motifs of about 1-2 cm of size) the method giving the best performances on most classes (CMC-2) is not effective, while LBPI gives sufficiently good results. This fact shows that methods may be differently tuned for pattern size and depth and a combination of them could be largely improve the retrieval scores.

Given a dissimilarity matrix, the associated confusion matrix CM is a squared matrix whose order is equal to the number of classes (15) in the dataset. For a row i in CM, the element $CM(i,j)$ gives the number of items which have been misclassified, resulting as elements of class j rather then elements of class i. Thus, the classification matrix CM of an ideal classification system should be a diagonal matrix, such that the element $CM(i,i)$ equals the number of items belonging to the ith class. The confusion matrices shown in Figure 12 give an overview of the correct/incorrect classification achieved when the NN is used as a classifier. These matrices are estimated on the 180 patches original dataset without remeshings, so they are not biased by global shape matching. The perfect matrix should contain only a diagonal of yellow elements while the others should be green. What is interesting is that some classes (like 4 and 10) are challenging for all methods while others are quite well kept by several techniques (for instance 1, 4, 9, 15). CMC-2 is also able to provide for some classes high precision values at high recall values. Classes with larger and more relieved patterns seems to be more easily characterized and this is probably related to the noisy real-world acquisition.

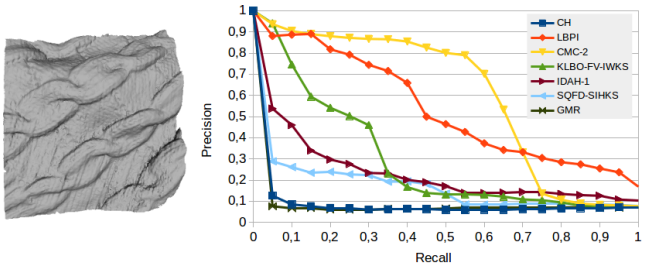

Figure 9: *Precision vs. Recall plot for the best runs of each participants on class 1 objects, characterized by the pattern on the left. Best methods (CMC-2 and LBPI) are able to retrieve similar patterns given the examples.*

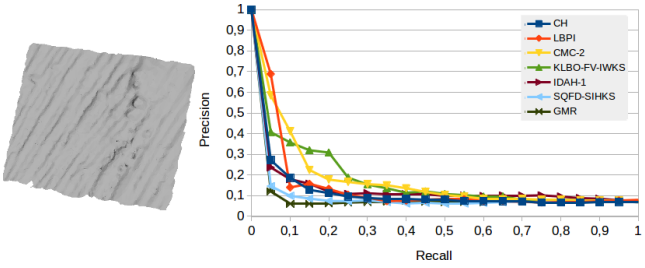

Figure 10: *Precision vs. Recall plot for the best runs of each participants on class 4 objects, characterized by the pattern on the left. No method is successful in retrieving patches of the same class.*

6. Discussion

The retrieval of surface patches with different spatial embeddings but with same relief patterns given an example is a practical task with several potential applications. One particular application that is interesting for track organizers is, for example, artworks' analysis, where the aim is to characterize materials (studied in the Scan4Reco project [sca]) and styles or repeated decorations (incision on beard and hair, helmets, etc) on Cultural Heritage artifacts (such as those analyzed in the GRAVITATE project [GRA]).

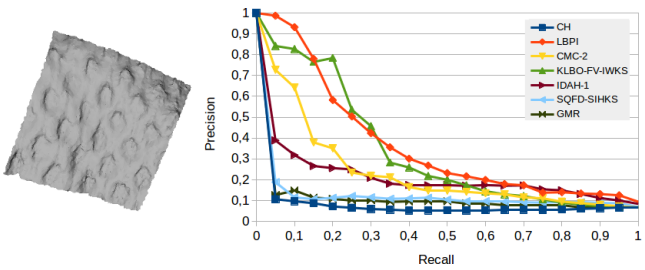

Figure 11: *Precision vs. Recall plot for the best runs of each participants on class 9 objects, characterized by the pattern on the left.*

Several authors recently proposed methods to characterize these patterns, however, the lack of labeled datasets was a limit in evaluating them. For this reason we proposed a novel dataset realized bending differently relieved tissues and acquiring patches with a depth sensor and organized a specific Shape Retrieval contest. This dataset includes patterns of different spatial frequency and resulted more challenging than expected, also due to the fact that it was created with a noisy real world capture of depth data. However, as humans can distinguish and assign the pattern visually on renderings, it is expected that algorithms that can retrieve correctly the same class instances given an example could be developed.

The simple curvature-based baseline method as well as more sophisticated global spectral approaches proposed by participants showed quite poor results. A few methods demonstrated more promising results, especially for selected relief patterns. The Kinetic weighting of Laplace-Beltrami operator proposed by Limberger&Wilson seems effective in the removal of surface bending, even if the descriptor appears to be still influenced by the patch shape. Velasco-Forero&Fehri as well as Tatsuma&Aono showed that exploiting image processing methods for 2D texture may be a viable solution, but that methods should be improved for practical applications.

Acknowledgments This work was partially supported by the Scan4Reco project funded by European Union's Horizon 2020 Framework Programme for Research and Innovation under grant agreement no 665091 and the GRAVITATE project, European project H2020 REFLECTIVE, grant agreement n. 665155, (2015-2018). A. Tatsuma and M. Aono were supported by Kayamori Foundation of Informational Science Advancement, Toukai Foundation for Technology, and JSPS KAKENHI Grant Numbers 26280038, 15K15992.

References

[ABD12] ALCANTARILLA P. F., BARTOLI A., DAVISON A. J.: Kaze features. In *Proc. of the 12th European Conference on Computer Vision* (2012), vol. 4 of *ECCV'12*, pp. 214–227. 2

[ACSD*03] ALLIEZ P., COHEN-STEINER D., DEVILLERS O., LÉVY B., DESBRUN M.: Anisotropic polygonal remeshing. *ACM Trans. Graph. 22*, 3 (July 2003), 485–493. 5

[AF06] ATTENE M., FALCIDIENO B.: Remesh: An interactive environment to edit and repair triangle meshes. In *Proc. SMI'06* (2006), IEEE Computer Society, p. 41. 1

[ASC11] AUBRY M., SCHLICKEWEI U., CREMERS D.: The wave kernel signature: A quantum mechanical approach to shape analysis. In *Computer Vision Workshops (ICCV Workshops), IEEE International Conference on* (Nov 2011), pp. 1626–1633. 3, 4

[BK10] BRONSTEIN M., KOKKINOS I.: Scale-invariant heat kernel signatures for non-rigid shape recognition. In *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.* (June 2010), pp. 1704–1711. 5

[BUS09] BEECKS C., UYSAL M. S., SEIDL T.: Signature quadratic form distances for content-based similarity. In *Proc. ACM Int. Conf. on Multimedia* (New York, NY, USA, 2009), MM '09, ACM, pp. 697–700. 4

[CSM03] COHEN-STEINER D., MORVAN J.-M.: Restricted delaunay triangulations and normal cycle. In *Proceedings of the Nineteenth Annual Symposium on Computational Geometry* (New York, NY, USA, 2003), SCG '03, ACM, pp. 312–321. 5, 6

[DT08] DHILLON I. S., TROPP J. A.: Matrix nearness problems with

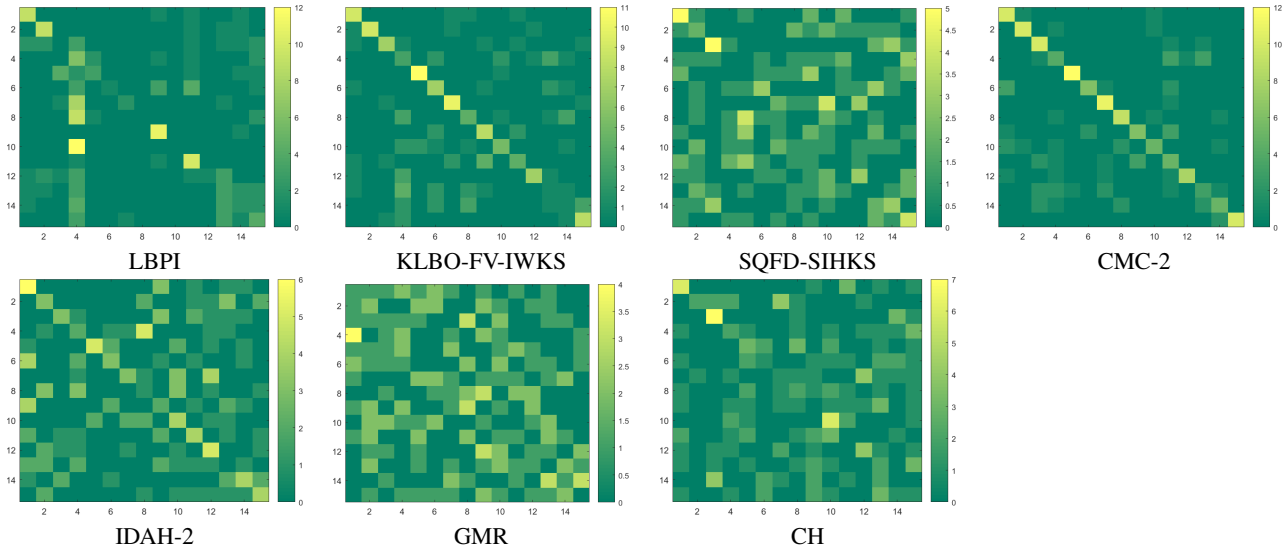

Figure 12: *Confusion matrices for the single instances of each patch (180 models). Colors range from green (0) to yellow (12).*

bregman divergences. *SIAM Journal on Matrix Analysis and Applications 29*, 4 (2008), 1120–1146. 5

[GH97] GARLAND M., HECKBERT P. S.: Surface simplification using quadric error metrics. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques* (New York, NY, USA, 1997), SIGGRAPH '97, ACM Press/Addison-Wesley Publishing Co., pp. 209–216. 4

[GRA] GRAVITATE: Discovering relationships between artefacts using 3D and semantic data. EU H2020 REFLECTIVE project. 8

[JC12] JÉGOU H., CHUM O.: Negative evidences and co-occurences in image retrieval: The benefit of PCA and whitening. In *Proc. of the 12th European Conference on Computer Vision* (2012), vol. 2 of *ECCV'12*, pp. 774–787. 3

[LB13] LI C., BEN HAMZA A.: A multiresolution descriptor for deformable 3D shape retrieval. *The Visual Computer 29* (2013), 513–524. 6

[LW15] LIMBERGER F. A., WILSON R. C.: Feature encoding of spectral signatures for 3D non-rigid shape retrieval. In *Proceedings of the British Machine Vision Conference (BMVC)* (September 2015), BMVA Press, pp. 56.1–56.13. 3, 4

[Mar70] MARDIA K. V.: Measures of multivariate skewness and kurtosis with applications. *Biometrika 57*, 3 (1970), 519–530. 2

[MH17] MASOUMI M., HAMZA A. B.: Spectral shape classification: A deep learning approach. *Journal of Visual Communication and Image Representation 43* (2017), 198–211. 6

[MLH16] MASOUMI M., LI C., HAMZA A. B.: A spectral graph wavelet approach for nonrigid 3d shape retrieval. *Pattern Recognition Letters 83* (2016), 339–348. 6

[OPH94] OJALA T., PIETIKAINEN M., HARWOOD D.: Performance evaluation of texture measures with classification based on kullback discrimination of distributions. In *Proc. of the 12th International Conference on Pattern Recognition* (Oct 1994), vol. 1 of *ICPR'94*, pp. 582–585. 2

[OVSP13] OTHMANI A., VOON L. F. L. Y., STOLZ C., PIBOULE A.: Single tree species classification from terrestrial laser scanning data for forest inventory. *Pattern Recognition Letters 34*, 16 (2013), 2144–2150. 1

[Pey11] PEYRÉ G.: The numerical tours of signal processing. *Computing in Science and Engg. 13*, 4 (July 2011), 94–97. 5

[PFA06] PENNEC X., FILLARD P., AYACHE N.: A riemannian framework for tensor computing. *International Journal of Computer Vision 66*, 1 (2006), 41–66. 2

[SBR16] SINHA A., BAI J., RAMANI K.: *Deep Learning 3D Shape Surfaces Using Geometry Images*. Springer International Publishing, Cham, 2016, pp. 223–240. 5

[sca] Scan4reco: Multimodal scanning of cultural heritage assets for their multilayered digitization and preventive conservation via spatiotemporal 4d reconstruction and 3d printing. 8

[sen] 3d systems sense for intel realsense. 1

[SLBS16] SIPIRAN I., LOKOČ J., BUSTOS B., SKOPAL T.: Scalable 3d shape retrieval using local features and the signature quadratic form distance. *The Visual Computer* (2016), 1–15. 4

[SMKF04] SHILANE P., MIN P., KAZHDAN M., FUNKHOUSER T.: The princeton shape benchmark. In *Shape modeling applications, 2004. Proceedings* (2004), IEEE, pp. 167–178. 2

[SOG09] SUN J., OVSJANIKOV M., GUIBAS L. J.: A Concise and Provably Informative Multi-Scale Signature Based on Heat Diffusion. *Comput. Graph. Forum 28*, 5 (2009). 5

[Soi13] SOILLE P.: *Morphological Image Analysis: Principles and Applications*, 2 ed. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2013. 5

[TPM06] TUZEL O., PORIKLI F., MEER P.: Region covariance: A fast descriptor for detection and classification. In *Proceedings of the 9th European Conference on Computer Vision - Volume Part II* (Berlin, Heidelberg, 2006), ECCV'06, Springer-Verlag, pp. 589–600. 5

[ZPS*16] ZEPPELZAUER M., POIER G., SEIDL M., REINBACHER C., SCHULTER S., BREITENEDER C., BISCHOF H.: Interactive 3d segmentation of rock-art by enhanced depth maps and gradient preserving regularization. *J. Comput. Cult. Herit. 9*, 4 (Sept. 2016), 19:1–19:30. 1