Eurographics Workshop on 3D Object Retrieval (2015)
I. Pratikakis, M. Spagnuolo, T. Theoharis, L. Van Gool, and R. Veltkamp (Editors)

SHREC15 Track: Retrieval of non-rigid (textured) shapes using low quality 3D models

Andrea Giachetti[1], Francesco Farina[1], Francesco Fornasa[1], Atsushi Tatsuma[2], Chika Sanada[2], Masaki Aono[2], Silvia Biasotti[3], Andrea Cerri[3], Sungbin Choi[4]

[1]Department of Computer Science University of Verona, Italy - Track organizers
[2]Department of Computer Science and Engineering, Toyohashi University of Technology, Japan
[3]Istituto di Matematica Applicata e Tecnologie Informatiche "E. Magenes", CNR, Italy
[4] Seoul National University, South Korea

Abstract

This paper reports the results of the SHREC 2015 track on retrieval of non-rigid (textured) shapes from low quality 3D models. This track has been organized to test the ability of the algorithms recently proposed by researchers for the retrieval of articulated and textured shapes to deal with real-world deformations and acquisition noise. For this reason we acquired with low cost devices models of plush toys lying on different sides on a platform, with articulated deformations and with different illumination conditions. We obtained in this way three novel and challenging datasets that have been used to organize a contest where the proposed task was the retrieval of istances of the same toy within acquired shapes collections, given a query model. The differences in datasets and tasks were related to the fact that one dataset was built without applying texture to shapes, and the others had texture applied to vertices with two different methods. We evaluated the retrieval results of the proposed techniques using standard evaluation measures: Precision-Recall curve; E-Measure; Discounted Cumulative Gain; Nearest Neighbor, First-Tier (Tier1) and Second-Tier (Tier2), Mean Average Precision. Robustness of methods against texture and shape deformation has also been separately evaluated.

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

1. Introduction

This SHREC contest aims at combining different research objective of past SHREC contests: testing of ability of 3D shape retrieval methods to deal with deformations [LGB*11], testing the ability of the methods to be applied on real-world data [PSR*14], testing the use of models captured with cheap depth sensors [MFP*13], testing the fusion of texture and shape features for retrieval [BCA*14]. To reach this goal, we proposed tasks based on low quality acquisitions (based on registration of point clouds acquired with a depth sensor) of a set of plush toys in different poses and with different illumination conditions. Participants were asked to test untrained shape retrieval performances based on only shape or on shape and texture information.

To allow an easy use of different descriptors proposed in the past SHREC contests, we provided watertight meshes, without texture or with interpolated color values stored as vertex features, saved in ASCII Object File Format.

2. Data acquisition and proposed tasks

We decided to use plush toys to test the robustness of shape retrieval methods against articulated and local deformations, as they can be easily posed in different ways, present a meaningful texture and the surface, acquired with depth sensor may be noisy. We selected a set of 12 toys with approximately similar size (Figure 1) and for each model we acquired 10 3D models. Poses are varied making the toys lying on different sides and moving articulated parts.

3D models have been created by acquiring for each toy 4 different point clouds with a Asus Xtion Live Pro depth sensor, capturing both depth and texture. The registration was made easier by placing the toys over a platform with markers. Creation, printing and recognition of markers was real-

DOI: 10.2312/3dor.20151067

Figure 1: *The 12 different toys placed on the acquisition platform.*

Figure 2: *Example models captured with different poses/illumination conditions*

ized using the ArUco toolkit [GJMSMCMJ14]. Each complete point cloud was finally meshed with the Poisson algorithm, with fixed parameters, in order to create a watertight model, and possibly simplified to remove degenerate triangles or to change resolution. The non-textured dataset *DATA_NOTEX*, in fact, was processed further by differentiating the resolution of meshes with similar poses.

Two textured datasets were then obtained from the reconstructed meshes and from the original textured point clouds by means of two different procedures: in the first, used to create the dataset *DATA_TEX*1, we generated a triangle parametrization and extrapolated vertex color in the whole mesh from this parametrization, in the second procedure, used to create the dataset *DATA_TEX*2, we attributed to each vertex the color of the closest point in the cloud within a threshold distance, leaving white the mesh vertices too far from the closest point.

For each model we first acquired the shape in a fixed pose under 5 different illumination conditions. Illumination was varied by putting lights in different positions around the models. Then we acquired other 5 models with fixed illumination conditions but changing the model pose, e.g. putting a differend side on the platform, moving limbs or other deformable parts. This results in both articulated deformations and local deformations due to the side in contact with the platform. Renderings of textured models captured in different poses/illumination conditions are shown in Figure 2. Data are now publicly available at the website `http://www.andreagiachetti.it/shrec15`

We finally proposed three different contests for the three datasets created: participants could freely decide to test their methods on all the datasets or on selected ones. The fact that the dataset features make the tasks extremely challenging was confirmed by the fact that of 9 groups registered for the contest, only 3 sent results for the evaluation, including as required matrices and executable codes. A.Tatsuma, C.Sanada and M. Aono from Toyohashi University of Technology sent results obtained on all the three datasets based on Local Binary Pattern and Hue Histograms; S.Biasotti and A.Cerri from IMATI-CNR sent results obtained on all the three datasets using Spherical Harmonics possibly coupled with Hue information, S. Choi from Seoul National University sent results only for the non-textured models, based on Geodesic Distance Matrices. These methods have been compared with different baseline techniques tested by the organizers, e.g. Area+Volume, ShapeDNA and Histograms of Area Projection Transform (HAPT) for non-textured meshes; CIE u-v histograms alone or differently coupled with HAPT for the textured meshes. All the methods are summarized in the following section.

3. Description of methods

3.1. Multiresolution Representation Local Binary Pattern and Hue Histograms, by A. Tatsuma, C. Sanada, and M. Aono

The Multiresolution Representation Local Binary Pattern (MRLBP) histograms has been proposed in [BCA*14] as a view-based 3D shape descriptor. The overview of the approach is illustrated in Figure 3. Authors performed the Multidimensional Scaling method [EK03] to obtain the canonical form of a 3D model as preprocessing. Here, to reduce the computational cost for the calculation of geodesic distances, the number of vertices of the 3D model was down sampled to about 1000 points. The transformed 3D model was enclosed within a unit geodesic sphere after pose normalization by the Point SVD [TA09]. From each vertex of the unit geodesic sphere, authors rendered depth and color buffer images with 256×256 resolution; a total of 38 view-

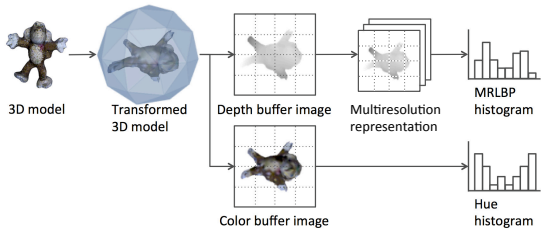

Figure 3: *Overview of the Multiresolution Representation Local Binary Pattern (MRLBP) based approach described in Section 3.1*

points were defined. MRLBP and Hue histograms were used as a feature vector for each view.

To capture shape features on a 3D model, the MRLBP histogram was calculated for each depth buffer image. Multiresolution representations were obtained by applying a Gaussian filter with varying standard deviation parameters to a depth buffer image. In the chosen implementation, authors selected 0.8, 1.4, 2.0, 3.2 and 6.8 for the standard deviation parameters. For each filtered image, an LBP histogram was computed. To incorporate spatial location information, the filtered image was partitioned into 4×4 blocks and the 64-bin LBP histogram was calculated at each block. The MRLBP histogram was generated by merging the LBP histograms of filtered images through the selection of the maximum value of each histogram bin. Moreover, to capture color features on a 3D model, the Hue histogram was computed for each color buffer image. The Hue histogram has invariance to light intensity changes [vdSGS10]. The color buffer image was partitioned into 4×4 blocks and a 6-bin Hue histogram was computed at each block.

The MRLBP and Hue histograms were l_1 normalized and mapped into the Jensen-Shannon kernel space [VZ12]. The final feature vector for a view was obtained by concatenating the mapped MRLBP and Hue histograms for textured meshes. To compare two 3D models, the Hungarian method [Kuh55] was applied to all pair dissimilarities between their feature vectors. The Euclidean distance was used to estimate the dissimilarity between two feature vectors. The method is referenced as $T1$ in Section 5

4. Spherical harmonics, CIELab colour space and persistence by S.Biasotti and A. Cerri

Authors of this submission proposed a shape signature arising from the progressive aggregation of three different, yet complementary ingredients:

- A geometric description. This was obtained with the popular spherical harmonic (SH) descriptor introduced in [KFR03]. Spherical harmonics are scale- and rotation-invariant. Once a shape is normalized and uniformly voxelized, it is associated with an array (512 elements in our implementation) encoding the indices of spherical harmonic values of the shape.
- A purely colourimetric descriptor. CIELab colour coordinates $(L,a,b$ channels) at each point of a triangle mesh were used to build a concatenated colour histogram (in the current implementation, 128 bins for each colour channel). The CIELab colour space was considered here as it well represents how human eyes perceive colours. Indeed, uniform changes of coordinates in the CIELab space correspond to uniform changes in the colour perceived by the human eye. This does not happen with some other colour spaces, for example the RGB space.
- A persistence-based colourimetric descriptor. Colour coordinates (normalized L,a,b channels) are seen as either scalar or multi-variate functions defined over the shape. Here authors exploited the fact that in the CIELab colour space, tones and colours are held separately: the L channel is used to specify the luminosity or the black and white tones, whereas the a channel specifies the colour as either a green or a magenta hue and the b channel specifies the colour as either a blue or a yellow hue. Following [BCGS13], the colourimetric description was finally included in the persistence framework.

Within this framework authors submitted one run for the non-textured dataset, based on SH description only, and three for the textured datasets, obtained as follows:

Run 1. (BC1) This was based on a simple geometric description. Each model was represented by the associated SH descriptor. For two models S_1 and S_2, their geometric dissimilarity $d_{geo}(S_1,S_2)$ was measured as the L_1-distance between the corresponding geometric descriptions;

Run 2. (BC2) The geometric description was enriched by adding some colourimetric information. In order to do this, CIELab colour histograms were included in the loop. Beyond the perceptual uniformity property directly inherited by the CIELab colour representation, colour histograms behave well against localized colourimetric noise, even when characterized by large variations in the L, a, b values. Indeed, in this case colour distribution is not altered greatly. For two models S_1 and S_2, their colourimetric dissimilarity $d_{clr}(S_1,S_2)$ was measured as the L_1-distance between the corresponding colour histograms; the final dissimilarity score was given by the normalized weighted sum $\lambda_1 d_{geo}(S_1,S_2) + \lambda_2 d_{clr}(S_1,S_2)$, with $\lambda_1 = 5$ and $\lambda_2 = 1$. The difference in weights is motivated by the intuition that geometry is more relevant than colour in assessing shape (dis)similarity.

Run 3. (BC3) The contribution of CIELab colour histograms was refined by introducing a persistence-based colourimetric description [BCGS13]. Indeed, the stability properties of persistence [CSEH07, CL15] imply robustness against small variations in the L,a,b values. This also holds when colour perturbations are widely spread over the surface model, as in the case of slight illumination changes. Thus, the idea was to use the persistence framework to comple-

ment the robustness of colour histograms against localized colourimetric noise. Precisely, the a, b coordinates were used to jointly define a bivariate function, whereas L was used as a scalar function. In this way, colour and intensity are treated separately. For a model S, the two functions $f_L : S \to \mathbb{R}$ and $f_{a,b} : S \to \mathbb{R}^2$ were considered, the former taking each point $x \in S$ to the L-channel value at x, the latter to the pair given by the a- and the b-channel values at x, respectively. The values of f_L and $f_{a,b}$ were then normalized to range in the interval [0,1]. Last, S is associated with the $0th$ persistence diagram $\mathrm{dgm}(f_L)$ and the $0th$ persistence space $\mathrm{spc}(f_{a,b})$: these descriptors encode the evolution of the connectivity in the sublevel sets of f_L and $f_{a,b}$ in terms of birth and death (i.e. merging) of connected components, see [BCGS13] for more details. For two models S_1 and S_2, their persistence-based distance $d_{pers}(S_1,S_2)$ was the normalized sum of the Hausdorff distance $d_H(\mathrm{dgm}_1(f_L),\mathrm{dgm}_2(f_L))$ between the corresponding persistence diagrams, and the Hausdorff distance $d_H\left(\mathrm{spc}_1(f_{a,b}),\mathrm{spc}_2(f_{a,b})\right)$ between the corresponding persistence spaces. The final dissimilarity score was given by the normalized weighted sum $\lambda_1 d_{geo}(S_1,S_2) + \lambda_2 d_{clr}(S_1,S_2) + \lambda_3 d_{pers}(S_1,S_2)$, with $\lambda_1 = 5$, $\lambda_2 = 2/3$ and $\lambda_3 = 1/3$.

4.1. Geodesic Distance Matrix by S. Choi

The author sent three runs only for the non textured task. The runs are based on simple geodesic distance matrices [SFH*09]. Downsampling was applied to the models to obtain 1.600 points per each mesh. Then geodesic distance matrix were built by calculating all geodesic distances between sampled points. Three variants of the method have been proposed:

Histogram 1 (C1): Each model's geodesic distance matrix is converted into histogram. Model to model dissimilarity was computed using Euclidean distance measure between two histogram.

Histogram 2 (C2): Each model's geodesic distance matrix is converted into histogram. Model to model dissimilarity from the correlation measure between two histograms.

Eigenvalue (C3): Each model's geodesic distance matrix's top 50 eigenvalues were extracted and converted to a vector. Model to model dissimilarity was computed using mean normalized Manhattan distance between two vectors.

4.2. Baseline methods by A.Giachetti

Three runs were proposed as baseline methods both for non-textured and textured tasks.

For non textured shape retrieval we propose three simple methods:

Area and Volume (G1): surface area and volume of the meshes were computed and the two elements vector obtained, normalized mapping the variable ranges in the interval [0 1] was used as shape descriptor. Model to model dissimilarity was estimated with Euclidean distance.

Shape DNA (G2): the first 100 eigenvalues of the Laplace Beltrami decomposition were computed and used as descriptor [RWP06]. The original code by M. Reuter was used `http://reuter.mit.edu/software/shapedna/`. Model to model dissimilarity was estimated with Euclidean distance.

Histograms of Area Projection Transform (G3): Histograms of Area Projection Transform have been computed with the method proposed in [GL12]. Considering the size of the objects, a fixed discretization step of 4mm. was used and the multiscale representation was based on 9 radii ranging from $r = 8mm$ to $r = 36mm$. Histograms at different scales (computed on 12 bins) were then concatenated, normalized and compared with the Cityblock distance.

For textured shape retrieval we propose three simple methods based on adding the CIELab color histograms to the shape descriptors. Only a and b components were considered as assumed to be less dependent on global illumination. Three runs were generated for each of the two textured datasets:

Color Only: Cielab ab histogram (G1T): The distance matrix is obtained only using color information from histograms of the CIELab a and b components. Shape dissimilarity is obtained measuring histogram differences with the Chi-square distance.

HAPT + ab histogram matrix combination (G2T): The distance matrix computed as in run G3 of the non textured dataset is combined with the distance matrix obtained as in run G1T of the textured cases.

HAPT + ab histogram feature combination (G3T): Finally a feature fusion approach was applied to the two histograms used in G2T. The two histograms, after normalization of the ranges into the [0 1] interval, were concatenated and then compared with the Cityblock distance.

5. Evaluation

The retrieval performance of the methods were evaluated according to the classical evaluation measures used in [SMKF04], e.g. Nearest Neighbor (NN), First Tier (FT), Second Tier (ST), e-measure (E) and Discounted Cumuated 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 in each relevant object in the retrieved list was estimated.

5.1. Retrieval of non-textured shapes

Table 1 shows the retrieval scores obtained on the full dataset. While it is clear that the performances of Geodesic Distances are quite poor, the other methods provide reasonable results even if the scores are lower than those obtained on other popular datasets. Looking also at the Average Precision-Recall curves, BC1 provides the best results especially for higher recall rates, with T1 and G3 providing close results.

	NN	1-Tier	2-Tier	E	DCG	MAP
C1	0,37	0,21	0,36	0,23	0,52	0,26
C2	0,38	0,21	0,36	0,24	0,52	0,27
C3	0,18	0,12	0,23	0,15	0,43	0,17
T1	**0,88**	0,57	0,67	0,34	**0,82**	0,61
G1	0,40	0,38	0,60	0,34	0,64	0,41
G2	0,76	0,44	0,58	0,33	0,74	0,49
G3	0,81	0,53	0,67	0,34	0,79	0,58
BC1	0,83	**0,60**	**0,71**	**0,37**	**0,82**	**0,63**

Table 1: *Retrieval scores obtained by the different runs on the DATA_NOTEX database. Bold fonts indicate best results.*

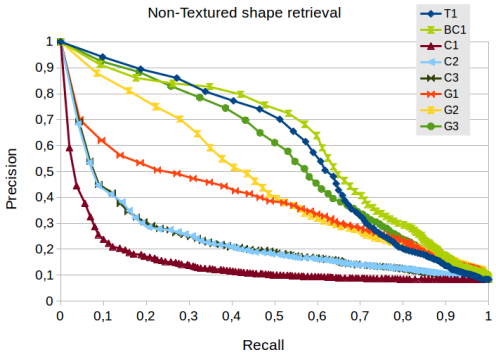

Figure 4: *Averaged Precision vs Recall curves for all the methods tested on the DATA_NOTEX dataset. Bold fonts indicate best results.*

If we analyze the results obtained on different subsets of the data, we can however understand that the "good" methods, behave differently and are better adapted on different model types. In fact, if we select only the subset of models of objects acquired on the same approximate pose, but under different illumination and differently remeshed (Table 2) it is possible to see that the results are quite high as expected and all the methods behave well with BC1 providing the best results, nearly perfect.

On the other hand, if we select only the subset of models of objects acquired changing the pose, but not the illumination conditions and the final resolution (Table 3) it is possible to see that the results are lower as expected and performances of BC1 and G2 are more relevantly decreased than those obtained by T1 and G3. This fact can be interpreted with the fact that MRLBP and MAPT are more robust against nonrigid (and also non-isometric) deformations, but more sensitive to acquisition noise and meshing, than SH. This is reasonable condidering that the SH descriptor is computed directly on voxelized shapes, and is therefore insensitive to meshing problems.

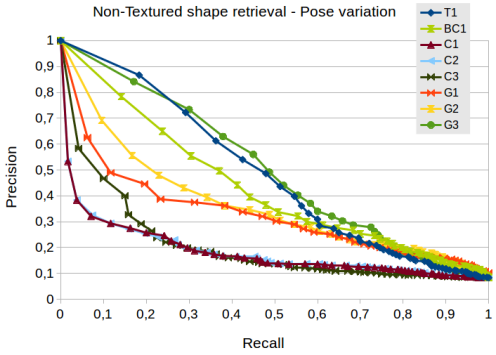

Figure 5: *Averaged Precision vs Recall curves obtained on a the subset of models with relevant shape deformations.*

5.2. Retrieval of textured shapes

Surprisingly, the addition of texture information seems to help the retrieval less than expected. In fact, looking at the scores obtained on the two textured datasets proposed, it is evident that the retrieval performances are only marginally improved or even not improved at all. This is evident looking at the runs of Biasotti and Cerri (BC1,BC2,BC3), proposing as first run a matrix obtained only from the shape descriptor. The other two runs obtain results only minimally better than the first. The same considerations hold, however, for the method of Tatsuma et al. (T1): here the scores reported in Table 4 and 5 are even worse than those obtained on the non-textured dataset (Table 1), apart for the NN. A slightly more evident improvement is obtained adding color information to the MAPT (G2T,G3T) descriptor, even if also in this case only for NN.

Results for the two datasets textured in different ways are more or less the same, slightly better for the set with white regions where texture is invisible (*DATA_TEX2*). This can be seen comparing Tables 4 and 5. This is reasonable, as in both cases a wrong information is in any case used to fill the

	NN	1-Tier	2-Tier	E	DCG	MAP
C1	0,55	0,36	0,51	0,20	0,62	0,47
C2	0,55	0,38	0,54	0,21	0,63	0,48
C3	0,20	0,13	0,19	0,12	0,38	0,19
T1	0,98	0,91	0,94	**0,22**	0,97	0,89
G1	0,33	0,29	0,58	0,21	0,59	0,45
G2	0,92	0,76	0,82	**0,22**	0,88	0,77
G3	0,97	0,86	0,93	**0,22**	0,94	0,86
BC1	**1,00**	**0,95**	**0,97**	**0,22**	**0,99**	**0,92**

Table 2: *Results of the best methods on models of object without pose changes but with different illumination conditions and mesh postprocessing. Bold fonts indicate best results.*

non visible part. Note that the results on the two datasets are also the same when only the color information is used for retrieval as in runs G1T.

Note that color information is not negligible, as the run G1T, using only color histogram information are acceptable, considering the acquisition noise and varying illumination. The fact that the combined use of color and shape does not give tangible advantages, demonstrates that there is room for a large room for improvements in the colour and shape descriptors. However, considering the different approaches proposed and the similar results, it seems that an effective feature fusion is harder than it may be expected.

The average Precision vs Recall plots (Figure 6) reveal that, as in the nontextured case, the use of SH descriptors gives some advantages, especially at higher recall rates.

Also in the textured case, if we separate the data acquired in different poses from the repeated acquisitions with similar poses and varied illumination, the ranking of the three best methods is completely different. Table 6 and Table 7 show the results obtained on the fixed pose-varying illumination subset and on the varied pose-fixed illumination subset. It is clear that the methods based on SH are more robust against the effect of illumination variations on the acquisition even not exploiting color information, while MAPT and MRLBP appear more robust against large deformations, as revealed also by the specific PR plot represented in Figure 7. As expected, the color-only descriptor (G1T) performs very well when no illumination changes are created.

To understand better the behavior of the best three methods, we plotted the sort of confusion matrices for the runs T1, BC3 and G3T performed on *DATA_TEX*1 that are represented in Figure 8, Figure 9 and Figure 10. In these plots models are ordered by class and for each query model in a row, the columns represent the first 9 retrieved models with color code depending on the retrieval order (black corresponds to the nearest neighbor, slightly lighter gray to the second retrieval, etc.). Looking at the matrices obtained, that ideally should be block matrices with white squares everywhere except for 10x10 squares along the principal diagonal (that is also white), it is possible to see that false positives are

	NN	1-Tier	2-Tier	E	DCG	MAP
C1	0,07	0,12	0,26	0,17	0,38	0,20
C2	0,07	0,12	0,27	0,16	0,38	0,20
C3	0,17	0,16	0,25	0,14	0,42	0,23
T1	**0,73**	0,43	0,56	0,19	0,67	0,52
G1	0,25	0,23	0,47	**0,21**	0,52	0,36
G2	0,38	0,29	0,49	**0,21**	0,56	0,40
G3	0,69	**0,45**	**0,61**	0,20	**0,69**	**0,53**
BC1	0,57	0,37	0,51	0,20	0,62	0,47

Table 3: *Results obtained on models of object with pose changes and constant illumination conditions. Bold fonts indicate best results.*

	NN	1-Tier	2-Tier	E	DCG	MAP
T1	0,89	0,44	0,53	0,27	0,75	0,49
BC1	0,83	**0,60**	**0,72**	**0,37**	0,82	0,63
BC2	0,83	**0,60**	**0,72**	**0,37**	0,82	0,63
BC3	0,85	**0,60**	**0,72**	**0,37**	**0,83**	**0,64**
G1T	0,73	0,31	0,39	0,20	0,64	0,33
G2T	0,88	0,39	0,47	0,24	0,72	0,44
G3T	**0,91**	0,52	0,65	0,34	0,81	0,58

Table 4: *Results obtained on the DATA_TEX*1 *dataset. Bold fonts indicate best results.*

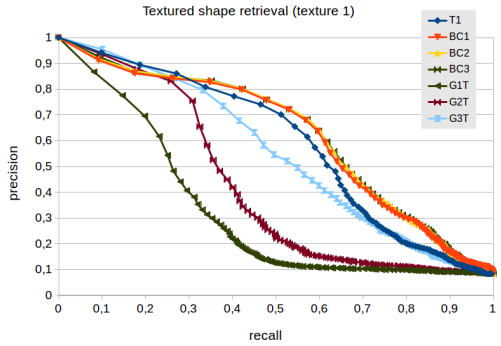

Figure 6: *Averaged Precision vs Recall curves for all the methods tested on the DATA_TEX*1 *dataset.*

quite different for the different methods and are largely dependent on classes and on subsets (illumination variations vs pose variations). In fact, for all the classes, represented in the figure with the plush toy icon, the set of 10 model is also ordered putting the 5 models with illumination variations first, then the models with deformations. It is possible to see that, for example, T1 and G3T are quite good in retrieving models from the same class subset while rarely retrieve in the first 9 spots models belonging to the subset of the same class different from the query one. This is revealed by the fact that the block matrix structures seem actually made by 5x5 blocks instead of 10x10 blocks. This is the reason why the methods have lower scores than BC3 for higher recall rates.

	NN	1-Tier	2-Tier	E	DCG	MAP
T1	0,89	0,46	0,55	0,28	0,77	0,51
BC1	0,83	0,60	**0,72**	**0,37**	0,82	0,63
BC2	0,85	0,60	**0,72**	**0,37**	**0,83**	**0,64**
BC3	0,85	**0,60**	**0,73**	**0,37**	**0,83**	**0,64**
G1T	0,73	0,31	0,39	0,20	0,64	0,34
G2T	0,88	0,39	0,47	0,24	0,72	0,44
G3T	**0,91**	0,50	0,65	0,33	0,80	0,56

Table 5: *Retrieval scores obtained on the DATA_TEX*2 *dataset. Bold fonts indicate best results.*

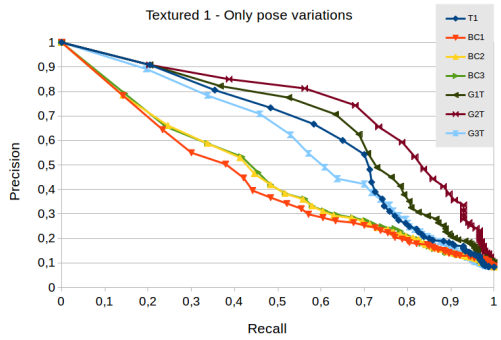

Figure 7: *Averaged Precision vs Recall curves for all the methods tested on a subset of DATA_TEX1 with large pose variations and no illumination changes*

6. Discussion

We created a novel dataset (distributed in different versions) and proposed a SHREC 2015 contest to verify the robustness of methods for the retrieval of deformable and optionally textured 3D shapes. The dataset has several challenging features:

- It has been acquired with current low cost depth sensor technology
- Objects are subject to articulated and non-isometric deformations and can change their topology
- Real and relevant changes of illumination conditions are represented
- Meshes are closed but meshed with different resolutions

From the analysis of the results submitted it is possible to understand interesting facts. First of all, the best methods proposed can deal with geometrical and topological noise and deformations, but each one has different and specific weaknesses. Color and shape information fusion did not perform well. Retrieval performance also depends strongly on class types. These facts may suggest that to apply successfully 3D shape retrieval methods on noisy textured data possibly acquired on deformable objects, research work is surely needed to find optimal ways to merge different descriptors. Specific learning techniques could be also tested in order to adapt the algorithms to the specific classes of interest. We hope that the research efforts needed to overcome the limitations of the current techniques will benefit from the availability of our datasets and of their possible future extensions.

	NN	1-Tier	2-Tier	E	DCG	MAP
T1	0,98	0,89	0,96	**0,22**	0,96	0,89
BC1	**1,00**	0,95	0,97	**0,22**	**0,99**	**0,92**
BC2	**1,00**	**0,97**	**0,98**	**0,22**	**0,99**	**0,92**
BC3	**1,00**	0,96	0,98	**0,22**	**0,99**	**0,92**
G1T	0,65	0,41	0,54	0,19	0,67	0,50
G2T	0,92	0,78	0,87	0,21	0,90	0,80
G3T	0,97	0,91	0,94	**0,22**	0,95	0,88

Table 6: *Results obtained on a 60 models subset of DATA_TEX1 with no large pose variations and relevant illumination changes. Bold fonts indicate be results.*

	NN	1-Tier	2-Tier	E	DCG	MAP
T1	0,82	0,58	0,72	0,21	0,76	0,67
BC1	0,57	0,38	0,52	0,20	0,62	0,47
BC2	0,57	0,41	0,56	0,20	0,64	0,49
BC3	0,58	0,42	0,56	0,20	0,64	0,49
G1T	0,82	**0,63**	**0,76**	**0,21**	**0,80**	**0,70**
G2T	**0,83**	0,60	0,75	**0,21**	**0,80**	0,68
G3T	0,79	0,53	0,70	0,21	0,75	0,60

Table 7: *Results obtained on a 60 models subset of DATA_TEX1 with large pose variations and no illumination changes. Bold fonts indicate best results.*

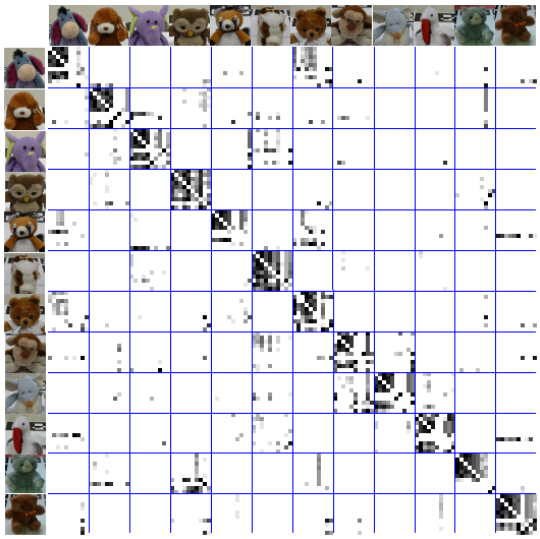

Figure 8: *The matrix represent the first 9 retrieved models (columns) for each query shape (rows) given the distance matrix of the run BC3. Plush toy icons correspond to the class of the index ranges. Darker colors correspond to earlier retrieval.*

References

[BCA*14] BIASOTTI S., CERRI A., ABDELRAHMAN M., AONO M., HAMZA A. B., EL-MELEGY M. T., FARAG A. A., GARRO V., GIACHETTI A., GIORGI D., GODIL A., LI C., LIU Y., MARTONO H. Y., SANADA C., TATSUMA A., VELASCO-FORERO S., XU C.: SHREC'14 track: Retrieval and classification on textured 3D models. In *Eurographics Workshop on 3D Object Retrieval, Strasbourg, France, 2014. Proceedings* (2014), pp. 111–120. doi:10.2312/3dor.20141057. 1, 2

[BCGS13] BIASOTTI S., CERRI A., GIORGI D., SPAGNUOLO

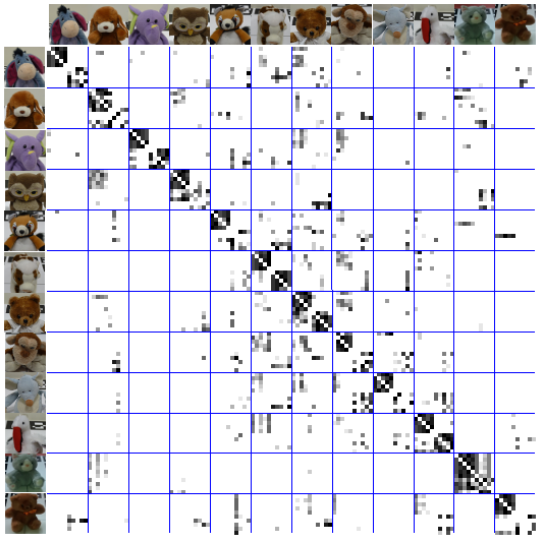

Figure 9: *The matrix represent the first 9 retrieved models (columns) for each query shape (rows) given the distance matrix of the run T1. Plush toy icons correspond to the class of the index ranges. Darker colors correspond to earlier retrieval.*

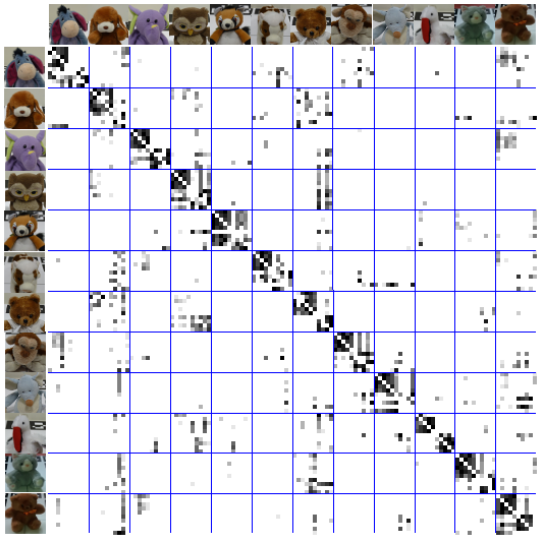

Figure 10: *The matrix represent the first 9 retrieved models (columns) for each query shape (rows) given the distance matrix of the run G3T. Plush toy icons correspond to the class of the index ranges. Darker colors correspond to earlier retrieval.*

M.: PHOG: Photometric and geometric functions for textured shape retrieval. *Comput. Graph. Forum 32*, 5 (2013), 13–22. 3, 4

[CL15] CERRI A., LANDI C.: Hausdorff stability of persistence spaces. *Foundations of Computational Mathematics* (2015), 1–25. doi:10.1007/s10208-015-9244-1. 3

[CSEH07] COHEN-STEINER D., EDELSBRUNNER H., HARER J.: Stability of persistence diagrams. *Discr. Comput. Geom. 37*, 1 (2007), 103–120. 3

[EK03] ELAD A., KIMMEL R.: On bending invariant signatures for surfaces. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 25*, 10 (Oct 2003), 1285–1295. doi:10.1109/TPAMI.2003.1233902. 2

[GJMSMCMJ14] GARRIDO-JURADO S., MUNOZ-SALINAS R., MADRID-CUEVAS F., MARIN-JIMENEZ M.: Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition 47*, 6 (2014), 2280 – 2292. doi:http://dx.doi.org/10.1016/j.patcog.2014.01.005. 2

[GL12] GIACHETTI A., LOVATO C.: Radial symmetry detection and shape characterization with the multiscale area projection transform. In *Computer Graphics Forum* (2012), vol. 31, Wiley Online Library, pp. 1669–1678. 4

[KFR03] KAZHDAN M., FUNKHOUSER T., RUSINKIEWICZ S.: Rotation invariant spherical harmonic representation of 3d shape descriptors. In *Proceedings of the First Eurographics/ACM SIGGRAPH Symposium on Geometry Processing* (Aire-la-Ville, Switzerland, Switzerland, 2003), SGP '03, Eurographics Association, pp. 156–164. 3

[Kuh55] KUHN H. W.: The hungarian method for the assignment problem. *Naval Research Logistics Quarterly 2* (1955), 83–97. 3

[LGB*11] LIAN Z., GODIL A., BUSTOS B., DAOUDI M., HERMANS J., KAWAMURA S., KURITA Y., LAVOUÉ G., VAN NGUYEN H., OHBUCHI R., ET AL.: Shrec'11 track: Shape retrieval on non-rigid 3d watertight meshes. *3DOR 11* (2011), 79–88. 1

[MFP*13] MACHADO J., FERREIRA A., PASCOAL P. B., ABDELRAHMAN M., AONO M., EL-MELEGY M., FARAG A., JOHAN H., LI B., LU Y., ET AL.: Shrec'13 track: retrieval of objects captured with low-cost depth-sensing cameras. In *Proceedings of the Sixth Eurographics Workshop on 3D Object Retrieval* (2013), Eurographics Association, pp. 65–71. 1

[PSR*14] PICKUP D., SUN X., ROSIN P., MARTIN R., CHENG Z., LIAN Z., AONO M., BEN HAMZA A., BRONSTEIN A., BRONSTEIN M., ET AL.: ShrecâĂŹ14 track: Shape retrieval of non-rigid 3d human models. *Proc. 3DOR 4*, 7 (2014), 8. 1

[RWP06] REUTER M., WOLTER F.-E., PEINECKE N.: Laplace–beltrami spectra as âĂŸshape-dnaâĂŹof surfaces and solids. *Computer-Aided Design 38*, 4 (2006), 342–366. 4

[SFH*09] SMEETS D., FABRY T., HERMANS J., VANDERMEULEN D., SUETENS P.: Isometric deformation modelling for object recognition. In *Computer Analysis of Images and Patterns* (2009), Springer, pp. 757–765. 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. 4

[TA09] TATSUMA A., AONO M.: Multi-fourier spectra descriptor and augmentation with spectral clustering for 3D shape retrieval. *The Visual Computer 25*, 8 (2009), 785–804. 2

[vdSGS10] VAN DE SANDE K. E., GEVERS T., SNOEK C. G.: Evaluating color descriptors for object and scene recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 32*, 9 (Sept 2010), 1582–1596. doi:10.1109/TPAMI.2009.154. 3

[VZ12] VEDALDI A., ZISSERMAN A.: Efficient additive kernels via explicit feature maps. *IEEE Transactions on Pattern Analysis and Machine Intelligence 34*, 3 (March 2012), 480–492. 3