Eurographics Workshop on 3D Object Retrieval (2009)
I. Pratikakis, M. Spagnuolo, T. Theoharis, and R. Veltkamp (Editors)

SHREC'09 Track: Structural Shape Retrieval on Watertight Models

J. Hartveldt[1], M. Spagnuolo[2], A. Axenopoulos[3], S. Biasotti[2], P. Daras[3], H. Dutagaci[4], T. Furuya[6], A. Godil[4], X. Li[4], A. Mademlis[3], S. Marini [2], T. Napoleon[5], R. Ohbuchi[6] and M. Tezuka[6]

[1] Utrecht University (UU), Netherlands
[2] Institute of Applied Mathematics and Information Technologies (IMATI) - Italian National Research Council (CNR), Italy
[3] Informatics Telematics Institute (ITI), Greece
[4] National Institute of Standards and Technology (NIST), USA
[5] TelecomParisTech; Laboratoire Traitement et Communication de l'Information (CNRS-LTCI), France
[6] University of Yamanashi, Japan

Abstract

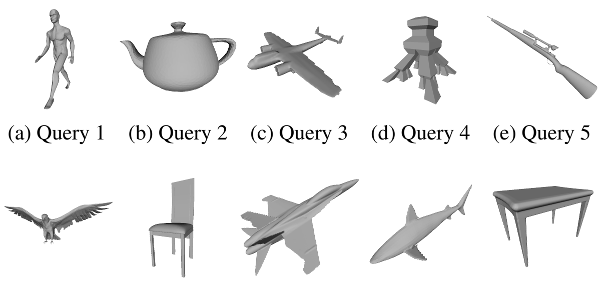
The annual SHape REtrieval Contest (SHREC) measures the performance of 3D model retrieval methods for several different types of models and retrieval purposes. In this contest the structural shape retrieval *track focuses on the retrieval of 3d models which exhibit a relevant similarity in the shape structure. Shape structure is typically characterised by features like protrusions, holes and concavities. It defines relationships in which components of the shape are connected.*

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

1. Data collection and Queries

The models have been gathered and constructed as structurally different. No metadata other than the model's mesh has been included in the shape, since the emphasis of the contest track lies on these structural differences.

The shape repository contains 200 models. The models are classified in 10 main classes. Each main class contains a pair of subclasses with 10 models. The 10 queries chosen for this contest are models from the repository (Figure 1). Matching models from inside a subclass are highly relevant and matching results between models in the two pair class are marginally relevant. Figure 2 provides a snapshot of all models currently used for the contest track.



(a) Query 1 (b) Query 2 (c) Query 3 (d) Query 4 (e) Query 5

(f) Query 6 (g) Query 7 (h) Query 8 (i) Query 9 (j) Query 10

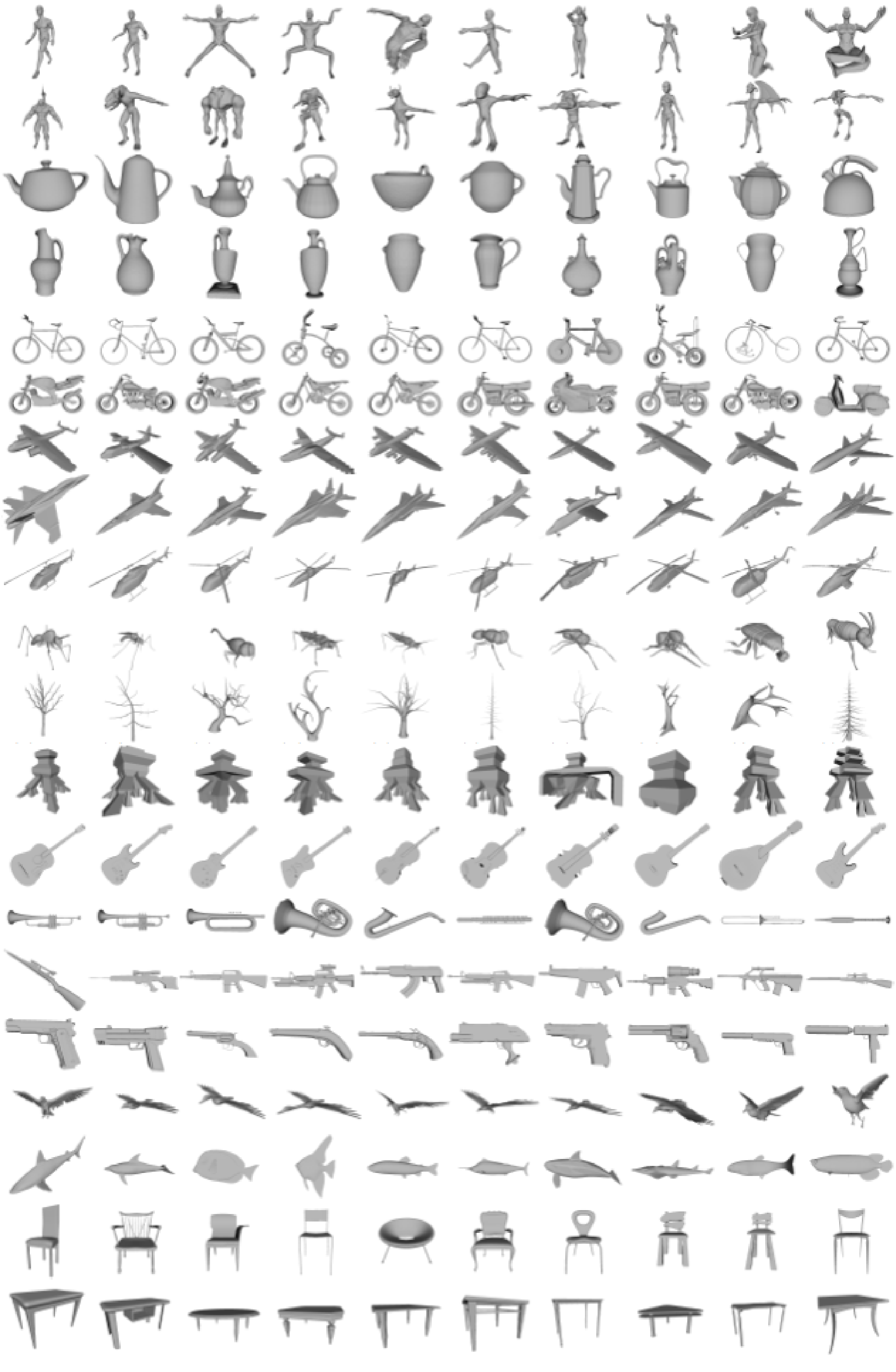
Figure 1: The query set.

2. Participants and Methods

Five out of six registrants have successfully submitted their track results. The results were submitted in the form of ten ranked lists. Each list containing the query item and the item that was returned, together with a computed distance and the resulting rank.

- Silvia Biasotti, Simone Marini (IMATI):
  ERG 1, 2
- Petros Daras, Apostolos Axenopoulos, Athanasios

Figure 2: The complete repository.



Mademlis (Informatics Telematics Institute):
Compact Multi View Descriptor 1
Compound SID CMVD 1, 2, 3
- XiaoLan Li, Afzal Godil, Helin Dutagaci (NIST):
BagOfWords 1
ConcentricBagOfWords 2
- Thibault Napoleon (TelecomParisTech CNRS LTCI):
MCC 1, 2, 3, 4
- Ryutarou Ohbuchi, Takahiko Furuya, Masaki Tezuka (University of Yamanashi):
BF-SIFT 1
MR-SPRH-UDR 1

The following subsections give a brief overview of the methods used in the contest. Please view the papers corresponding to these methods for a more detailed description.

2.1. ERG

The shape description and matching framework from the ERG (Enhanced Reeb Graph) [BM09] based experiments have been originally used to approach the sub-part correspondence problem, i.e., to find similar sub-parts of objects represented as 3D polygonal meshes. In the method, the geometry and the structure of the shapes are coupled in a descriptor that provides a flexible coding, grounded on solid mathematical theories, and that can be adapted to the user's needs and to the context of applications.
The matching framework is a graph-matching technique, which builds the common sub-graphs between the two shapes and highlights the maximal sub-parts whose structure and space distribution are similar.

2.2. Compound SID CMVD

The Compound SID CMVD (Compact Multi View Descriptor and Shape Impact Descriptor) [DMA09] method combines a volume-based approach with a view-based method, producing a compound descriptor. The resulting field is described using both Newton's and General Relativity's laws. The latter is applied to a set of 2D images (multi-views), which are automatically generated from a 3D object, by taking views from uniformly distributed viewpoints. For each image, a set of 2D rotation-invariant shape descriptors is extracted.

2.3. (Concentric) BagOfWords

The first method represents shape as one Bag-of-Words (BW) model and the second method represents shapes with Concentric Bag of Words (CBW) model [LGD09]. Based on the first method, each shape is recorded as a global histogram, while based on the latter one; each shape is recorded as several concentric histograms. In both of these methods Spin Images are used as local shape descriptor.
The bag-of-word was originally used for text retrieval. Recently attracting a lot of interest in the computer vision field. It has been applied to image classification, video classification, 3D shape retrieval and analysis and so on. They first use the original bag-of-words (BW) representation, and then introduce the expanded procedure of the Concentric Bag-of-Words (CBW) method.

2.4. MCC

The MMC method [Nap09] presents a new approach for each 3D shape retrieval stages (alignment, descriptors extraction and similarity measure). First, they introduce a new model alignment to capture the pose of the 3D object. The smallest enclosing sphere and a pose (computed with various principal components analysis) selection based are used on the minimal visual hull to solve this problem. After this alignment process, they extract a robust and compact signature from a derived representation as a set of images. This method is based on a multi-view approach that keeps 3D model coherence by considering simultaneously a set of 2D images in specific view directions. The various silhouettes of a model are strongly correlated, a selected set of them help to better discriminate one model among others. The extracted signature is based on a multi-scale representation. They keep a histogram of this information as a global geometric feature and the whole information as a local geometric feature. Finally, the matching process is performed with a pruning approach based on the advantages of the local and global signature of the object. This framework allows them to control the performance of the method accordingly to the computation time.

2.5. BF-SIFT and MR-SPRH-UDR

The bag-of-local visual feature approach BF-SIFT and the unsupervised dimension reduction approach MR-SPRH-UDR [TFO09] have been used without combining multiple methods. Instead, they chose to test the basic performance. The BF-SIFT uses range images of size 256 x 256 taken from 42 equally spaced view orientations about the 3D model. For each 3d model there is an average of 2,057 features in the database. The MR-SPRH-UDR uses the Locally Linear Embedding to learn the mapping from the 625D input space of the SPRH feature to output a 400D subspace. The mapping is learned from a total of 5,000 SPRH features. The contest track database was considered too small for this learning purpose and therefore the LLE was trained using 5,000 models randomly selected from the union of Princeton Shape Benchmark database and the National Taiwan University database. The learned mapping is approximated by using RBF network with Gaussian kernel. They used the cosine similarity measure as a distance computation among the dimension-reduced features.

3. Performance Measures

The performance of the shape retrieval is measured using 1^{st} and 2^{nd} tier *precision* and *recall*, finally presented as the *F-Measure*. This measure gives a nice overview of the complete retrieval performance and the method, especially since F-Measure is only high when both precision and recall are high as well. The average precision and recall values of the ten queries are used as input for the computation of this F-Measure. Since the values have been calculated for two Tiers this results in two F-Measures per method run. Finally the two F-Measures are compared to their maximum to create a percentage. The average of the two F-Measure percentages determined the position in the ranked list.

Precision is defined as the ratio of number of relevant retrieved objects to the number of retrieved objects. Objects retrieved from a highly relevant class are rewarded 1 point. Objects retrieved from a marginally relevant class are rewarded a score of 0.5 points. Thus, the maximum precision for the first tier (10 retrieved objects) is $10/10 = 1$, and the maximum precision for the second tier (20 retrieved objects) is $15/20 = 3/4$.

Recall is defined as the ratio of number of relevant retrieved objects to the total number of relevant objects. Thus, the maximum recall for the first tier is $10/15 = 2/3$, and for the second tier $15/15 = 1$.

F-Measure is separately calculated for the 1^{st} and 2^{nd} tier. The F-Measure will only be high if both precision and recall are high:

$$\text{F-Measure} = \frac{2*RecallAvg*PrecisionAvg}{RecallAvg+PrecisionAvg}.$$

The final **percentage** is calculated based on the two F-Measures, where n is the number of tiers:

$$\text{RankPercentage} = \frac{\sum(\frac{FM}{MaxFM}*100)}{n}.$$

4. Results and Discussion

In the first image below (Figure 3) both the 1st and 2nd Tier F-Measure values have been plotted in a graph where the top line represents the maximum value. An absolute representation where the maximum values are stated in the top left corner for respectively 1st and 2nd tier. The size of the bars are therefore not related to the relative value the methods have for tier 1 and 2, but to a percentage. This percentage represents a ratio of the achieved F-Measure compared to it's maximum possible value. Since the tiers have their own maximum values the lengths are on a different scale.

To compute the final ranked list, the two percentages have been averaged into a single percentage value. All values are ordered in Table 1. Note that the maximum values for the Tier 1 and Tier 2 F-measures are respectively 0.8 and 0.86 as displayed in the top-left corner of Figure 3.

The averages used to build the F-Measures, as stated in section 3, are listed in table 2.

The participating methods mostly perform in the same ranking order, for the different queries, as in the final results list. The 'Walking Human' query is the only big exception and it might be interesting to include a larger part of the queries with life forms. It is also interesting to see that a table as a query returns more true positives than a chair as a query on the same ground truth. This might indicate that a directional relation also could be taken into account. There is also more space for partial matching within the participating methods. The military rifle query confirms this assumption. It might be nice to include more partial matching tests in a future repository.

5. Conclusion

The structural shape retrieval contest track of SHREC 2009 can be considered a success. Five out of six participants submitted results for various settings of their methods. The methods are based on very diverse descriptors but show competetive scores. Overall the submissions perform well and a future track might take the difficulty of watertight model queries to a higher level.

Acknowledgements

This work was partially supported by the FP7 project FOCUS-K3D 2007-214993.

References

[AF06] ATTENE M., FALCIDIENO B.: Remesh: An interactive environment to edit and repair triangle meshes. *Shape Modeling and Applications, International Conference on 0* (2006), 41.

[BM09] BIASOTTI S., MARINI S.: Extended reeb graph description for structural comparison. Participant of SHREC09 Structural Shape Retrieval Contest track, jan 2009.

[DMA09] DARAS P., MADEMLIS A., AXENOPOULOS A.: A compound descriptor for 3d object retrieval. Participant of SHREC09 Structural Shape Retrieval Contest track, jan 2009.

[HS09] HARTVELDT J., SPAGNUOLO M.: Shrec 2009 structural shape retrieval. http://shrec.ge.imati.cnr.it/shrec09_structural.html, 2009.

[LGD09] LI X., GODIL A., DUTAGACI H.: 3d shape retrieval based on bag of words method. Participant of SHREC09 Structural Shape Retrieval Contest track, feb 2009.

[Nap09] NAPOLÉON T.: Multi-view 3d retrieval using multi-scale contour representation. Participant of SHREC09 Structural Shape Retrieval Contest track, jan 2009.

[TFO09] TEZUKA M., FURUYA T., OHBUCHI R.: Description of bf-sift and mr-sprh-udr strucutral shape retrieval methods. Participant of SHREC09 Structural Shape Retrieval Contest track, feb 2009.

[VtH09] VELTKAMP R., TER HAAR F.: Shrec 2009 shape retrieval contest. http://www.aimatshape.net/event/SHREC/, 2009.

Method and Score			Tier 1		Tier 2	
Rank	Total %	Methodname	F-Measure (%)		F-Measure (%)	
1	74.67 %	MCC 3	0.648	81.0 %	0.586	68.3 %
2	73.75 %	Compound_SID_CMVD 3	0.620	77.5 %	0.600	70.0 %
3	72.42 %	Compound_SID_CMVD 2	0.612	76.5 %	0.586	68.3 %
4	69.17 %	Compound_SID_CMVD 1	0.592	74.0 %	0.551	64.3 %
4	69.17 %	MCC 2	0.592	74.0 %	0.551	64.3 %
4	69.17 %	BF-SIFT 1	0.592	74.0 %	0.551	64.3 %
7	68.00 %	MR-SPRH-UDR 1	0.576	72.0 %	0.549	64.0 %
8	65.92 %	MCC 4	0.572	71.5 %	0.517	60.3 %
9	65.67 %	Compact Multi View Descriptor 1	0.552	69.0 %	0.534	62.3 %
10	64.58 %	MCC 1	0.548	68.5 %	0.520	60.7 %
11	57.42 %	ERG 2	0.492	61.6 %	0.457	53.3 %
12	52.33 %	ERG 1	0.448	56.0 %	0.417	48.7 %
13	25.83 %	BagOfWords 1	0.232	29.0 %	0.194	22.7 %
14	22.83 %	ConcentricBagOfWords 2	0.200	25.0 %	0.177	20.7 %

Table 1: The contest ranked results list build on F-Measures.

Method and Score			Precision		Recall	
Rank	Total %	Methodname	Tier 1	Tier 2	Tier 1	Tier 2
1	74.67 %	MCC 3	0.810	0.510	0.540	0.683
2	73.75 %	Compound_SID_CMVD 3	0.775	0.525	0.517	0.700
3	72.42 %	Compound_SID_CMVD 2	0.765	0.513	0.510	0.683
4	69.17 %	Compound_SID_CMVD 1	0.740	0.483	0.493	0.643
4	69.17 %	MCC 2	0.740	0.483	0.493	0.643
4	69.17 %	BF-SIFT 1	0.740	0.483	0.493	0.643
7	68.00 %	MR-SPRH-UDR 1	0.720	0.480	0.480	0.640
8	65.92 %	MCC 4	0.715	0.453	0.477	0.603
9	65.67 %	Compact Multi View Descriptor 1	0.690	0.468	0.460	0.623
10	64.58 %	MCC 1	0.685	0.455	0.457	0.607
11	57.42 %	ERG 2	0.615	0.400	0.410	0.533
12	52.33 %	ERG 1	0.560	0.365	0.373	0.487
13	25.83 %	BagOfWords 1	0.290	0.170	0.193	0.227
14	22.83 %	ConcentricBagOfWords 2	0.250	0.155	0.167	0.207

Table 2: The contest ranked results list and the different Precision and Recall values.

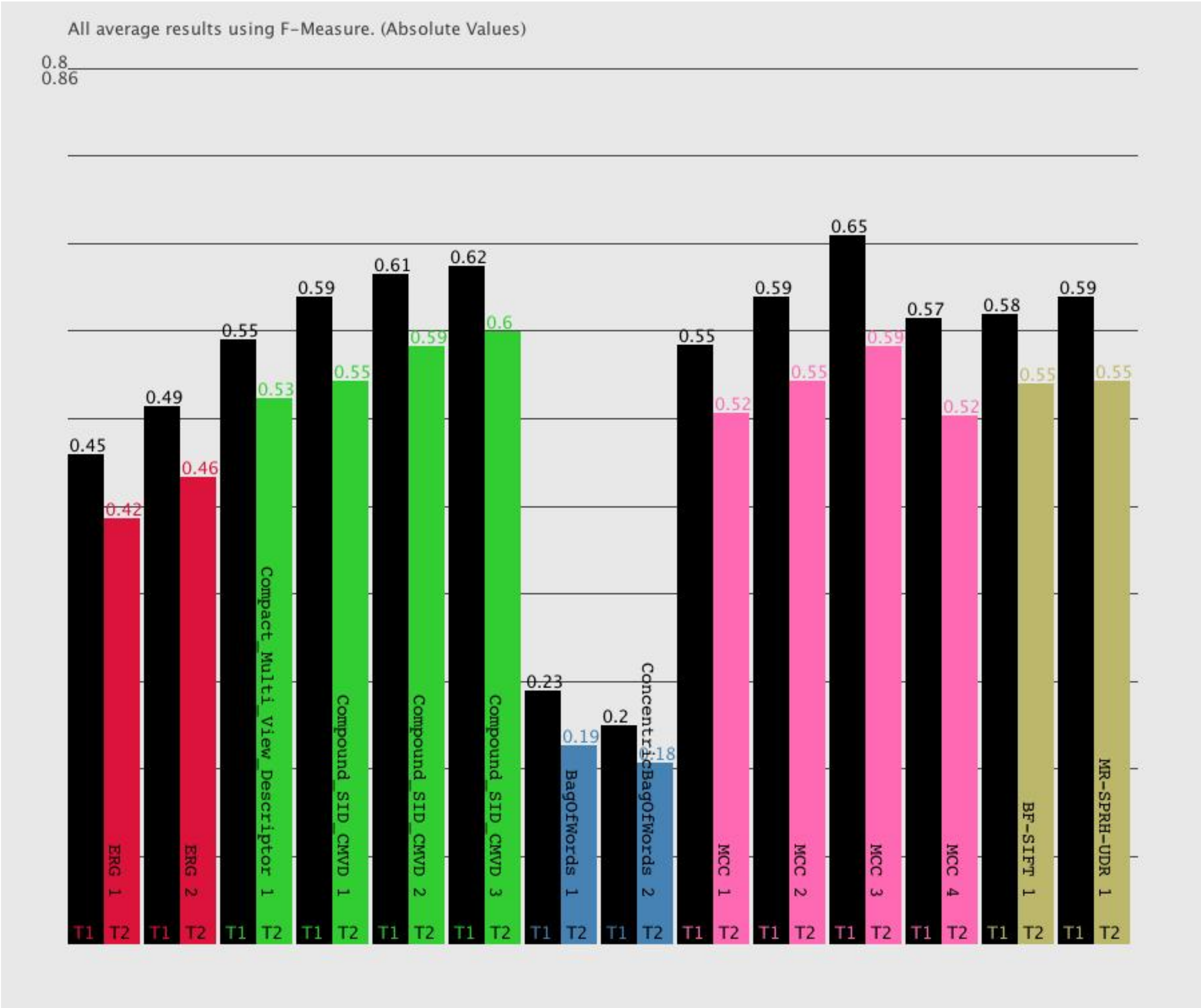


Figure 3: Absolute representation of the contest track performance.