Retrieval of 3-Dimensional Rigid and Non-Rigid Objects

Konstantinos Sfikas*

National and Kapodistrian University of Athens,
Department of Informatics and Telecommunications,
ksfikas@di.uoa.gr

Abstract. This dissertation focuses on the problem of 3D object retrieval from large datasets in a near realtime manner. In order to address this task we focus on three major subproblems of the field: (i) pose normalization of rigid 3D models with applications to 3D object retrieval, (ii) non-rigid 3D object description and (iii) search over rigid 3D object datasets based on 2D image queries. Regarding the first of the three subproblems, 3D model pose normalization, two main novel pose normalization methods are presented, based on: (i) 3D Reflective Object Symmetry (ROSy) and (ii) 2D Reflective Object Symmetry computed on Panoramic Views (SymPan/SymPan+). Considering the second subproblem, a non-rigid 3D object retrieval methodology, based on the properties of conformal geometry and graph-based topological information (ConTopo++) has been developed. Furthermore, a string matching strategy for the comparison of graphs that describe 3D objects, is proposed. Regarding the third subproblem a 3D object retrieval method, based on 2D range image queries that represent partial views of real 3D objects, is presented. The complete 3D objects of the database are described by a set of panoramic views and a Bag-of-Visual-Words model is built using SIFT features extracted from them. The methodologies developed and described in this dissertation are evaluated in terms of retrieval accuracy and demonstrated using both quantitative and qualitative measures via an extensive consistent evaluation against state-of-the-art methods on standard datasets.

Keywords: 3D Objects, Rotation Normalization, Shape Modelling, Partial Matching, Range Images

1 Introduction

Information, commonly refers to a useful portion of data located among a collection of related entities. Recent advances in storage technologies and the widespread use of the Internet, have resulted in a vast increase of the amount of data stored in and distributed from large databases. Any attempt for manual annotation and information extraction is almost impossible, therefore rendering the need for an automated procedure, mandatory.

* Dissertation Advisor: Theoharis Theoharis, Professor

The process of extracting useful information from large amounts of data, in an automated manner and based on an example or descriptive query, is called information retrieval. Common types of information that can benefit from such a retrieval process are: textual, visual, audio and video data and most recently, 3D and 4D (3D over time) data.

In recent years, through the creation of inexpensive 3D scanners and the simplification of 3D modelling software, a large volume of 3D data has been created and stored in corresponding scientific and industrial/commercial repositories. Furthermore, 3D data can be processed in various, application dependent, ways and occasionally be combined with data of other types and modalities (e.g. textual annotation and/or thumbnails of 3D models). These data types can further be used as queries for the retrieval of 3D objects.

Some example applications that exploit the properties of 3D models and could greatly benefit from a retrieval process follow: in medicine large diagnostic 3D data are compared and researched in order to assist the process of making medical decisions. In biometrics a person's 3D facial model is searched over corresponding databases for identification purposes. Game development utilizes retrieval and reusability of 3D models in order to minimize production times and reduce the size of the final product. Other example application areas include engineering and archaeology. It can therefore be easily deduced, that 3D object retrieval is a key process, although in general it is complex and highly depended on the application.

2 Framework and problem statement

3D object retrieval applications can be classified into two major categories: interclass and intra-class retrieval. Inter-class retrieval focuses on a generic domain of 3D objects and aims at finding the closest match among a set of 3D models that belong to a broad range of different classes. In this case, there is usually no prior knowledge regarding the characteristics or the nature of the 3D objects. Intra-class retrieval targets a specific 3D object domain (e.g. 3D faces, non-rigid 3D models, human action models, engineering models etc), where a match is sought between 3D models that belong to the same class but have their special characteristics defined differently. Intra-class 3D object retrieval methods usually exploit domain knowledge and shape characteristics of the 3D models, in order to attain higher performance.

For both categories, the generic framework of a 3D object retrieval system can be outlined as follows: preprocessing, pose normalization, shape descriptor extraction, feature matching.

At the first step of the 3D object retrieval pipeline, 3D models are preprocessed. In this step, the 3D models are cleaned up of any inconsistencies present due to the digitization process, i.e. double or reversed faces, structural gaps, etc. This step is highly dependent on the method/equipment used for the creation of the 3D models and may differ greatly from one application to another.

After basic preprocessing, *Pose Normalization* ensures that the geometric properties of the 3D models are defined in a uniform manner. The diversity of 3D object acquisition sources implies that 3D objects which may even be part of the same dataset, have their geometrical properties arbitrarily defined. Therefore, before any kind of processing is carried out, it must be ensured that the 3D objects have been normalized in terms of position, scaling and rotation (Fig 1 shows an example rotation normalization). Pose normalization of 3D objects is a common preprocessing step in various computer graphics applications [2, 22, 23, 27]. Visualization, broken fragment reconstruction, biometrics and 3D object retrieval are only a few examples of applications that benefit from a pose normalization procedure. To achieve pose normalization, for every 3D object, a corresponding set of normalization transformations in 3D space must be defined.

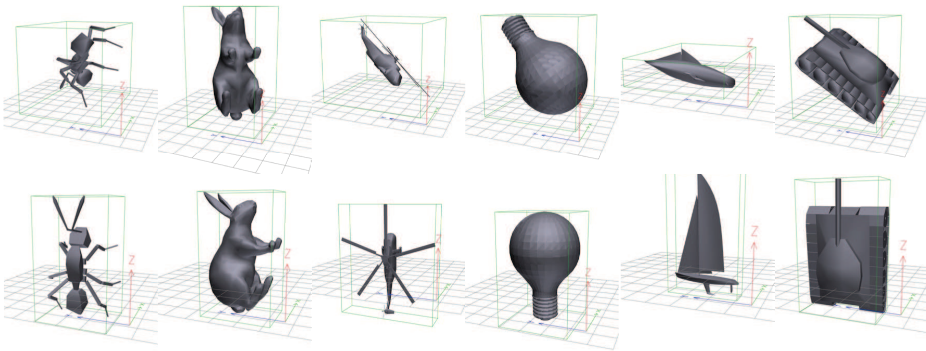

Fig. 1: Examples of non-aligned objects (top-row) and the corresponding rotation normalizations (bottom row).

The main step of a 3D object retrieval system is the computation of a feature set for each 3D model. In this step, the structural and/or other special characteristics of a 3D object are modelled and a shape descriptor that faithfully encodes the shape of the 3D model, in an efficient manner, is created. Feature selection is tightly connected to the corresponding application and can vary greatly for each 3D object retrieval system (e.g. intra-class retrieval exploits features that are more distinguishing within a specific domain, whereas inter-class retrieval uses more generic characteristics).

Finally, each 3D object's shape descriptor is used as a signature during the matching procedure. At this step, the signatures of the 3D models, stored in the database, are compared to the corresponding signatures of the query 3D model(s), using a specified metric. The selected metric is also dependent on both the features selected and the corresponding application. Finally, the response of the 3D object retrieval system is the set of 3D object(s) that correspond to the closest match(es) of the given user query.

3 Contributions

This dissertation has made the following research contributions in the field of 3D object retrieval: two new 3D model pose normalization methods, a non-rigid 3D object retrieval methodology and a 3D object retrieval algorithm, based on range image queries. In detail, the contributions of this dissertation are the following:

3.1 ROSy Pose Normalization Method

A general purpose global pose normalization method, based on 3D object reflective symmetry.

In the ROSy method, the problem of pose normalization is described through the Surface-Oriented Minimum Bounding Box (SoMBB), a modified version of the Axis-Aligned Bounding Box (AABB) which is commonly used in collision detection techniques [24, 8].

The motivation behind the proposed method is to minimize the SoMBB of a 3D object so that the latter becomes aligned with its SoMBB and consecutively with the principal axes of space. Furthermore, to ensure that the 3D object's large planar areas are also in alignment with the principal planes of space, it is required that the average normal to the object's large planar areas become parallel to the box's face normals (Fig. 2).

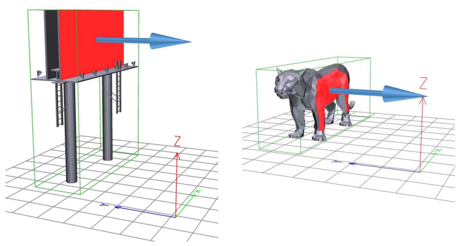

Fig. 2: 3D objects enclosed in their SoMBBs.

Initially, the axis-aligned minimum bounding box of a rigid 3D model is modified by requiring that the 3D model is also in *minimum angular difference* with respect to the normals to the faces of its bounding box.

To estimate the modified axis-aligned bounding box, a set of predefined principal planes of symmetry is used and the corresponding symmetric models are computed. Then, a combined spatial and angular distance, between the 3D model and its symmetric model, is calculated.

By minimizing this combined distance, through a set of rotations in space, the 3D model fits inside its modified axis-aligned bounding box and alignment with the coordinate system is achieved. [18]

3.2 SymPan+ Pose Normalization Method

A pose normalization method, based on panoramic views and reflective symmetry, is presented.

The motivation for the proposed method is that the use of reflective symmetry as a feature for pose normalization and 3D object retrieval seems to enhance the results [11], as most of the 3D objects exhibit symmetrical properties to some degree. These properties tend to be distinct between different classes and similar between objects of the same class, therefore enhancing the discrimination achieved by other commonly used characteristics, such as the spatial distribution and/or surface orientation of the 3D models. Qualitative and experimental investigation in 3D data-sets has led us to the observation that most objects possess at least a single plane of symmetry. Our approach is thus guided by this observation.

Initially, the surface of a 3D model is projected onto the lateral surface of a circumscribed cylinder, aligned with the primary principal axis of space. Based on this cylindrical projection, a *normals' deviation map* is computed.

Through an iterative procedure, the symmetry plane of the 3D model is parallelized with the axis of the projection cylinder, thus computing the first principal axis of the 3D model. This is achieved by rotating the 3D model and computing reflective symmetry scores on panoramic view images (Fig 3).

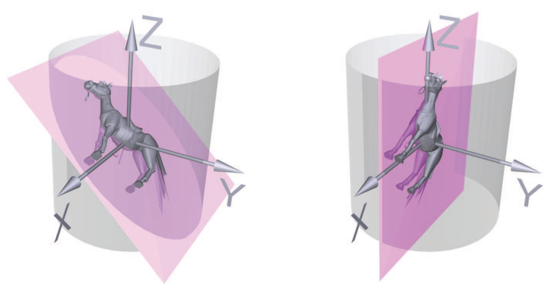

Fig. 3: Aligning the symmetry plane normal with the XY plane

The other principal axes of the 3D model are then estimated by computing the variance of the 3D model's panoramic views. [17, 21]

3.3 ConTopo++ Non-Rigid 3D Object Retrieval

Combining the properties of conformal geometry and graph-based topological information, a non-rigid 3D object retrieval methodology is proposed, which is both robust and efficient in terms of retrieval accuracy and computation speed. While graph-based methods are robust to nonrigid object deformations, they require intensive computation which can be reduced by the use of appropriate representations, addressed through geometry-based methods. In this respect, a 3D object retrieval methodology, which combines the above advantages in a unified manner, is presented.

Initially, we define a graph, that captures the topological structure of an arbitrary 3D mesh. Each node of the graph represents a unique connected component, while each edge of the graph describes the relation between adjacent connected components. Each connected component is composed of 3D mesh faces that have the same label and are also pathwise-connected.

In this work we have used discrete conformal factors [1] as a labeling criterion due to their ability to identify the protrusive parts in a mesh. The faces of the 3D mesh are partitioned based on a linear multi-thresholding of the values of the discrete conformal factor, thus splitting the mesh into a set of connected components (see Fig. 4).

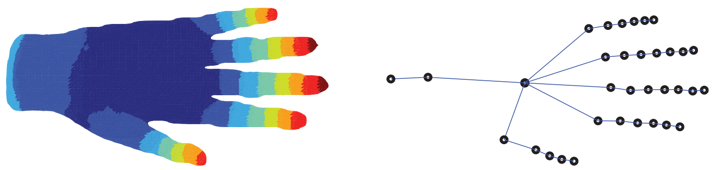

Fig. 4: Illustration of an eight-level quantized mesh and the corresponding graph.

Mesh matching compares both geometrical and topological features as a measure of similarity between two 3D meshes in a unified manner. During matching, the topological equivalence between the graphs of two 3D meshes is examined and enhanced by node-to-node comparison of geometrical features.

The matching procedure is based on string matching. Each ordered path, of graph nodes, that extends from the *core* partition of the 3D mesh down to each of its articulations is considered a *string*. Furthermore, besides the ordered connectivity of the string (graph) nodes, a number of features are also attached to them, which are used for the geometrical matching. [19]

3.4 3D Object Retrieval based on 2D Range Image Queries

A 3D object retrieval method, based on range image queries that represent partial views of real 3D objects, is presented.

The motivation behind the proposed method, is to use a 2D image in order to query a database of 3D objects and bridge the representation gap between the two in an efficient manner.

The complete 3D models of the database are described by a set of panoramic views and a Bag-of-Visual-Words model is built using SIFT features extracted from them. To address the problem of partial matching, a spatial histogram computation scheme, on the panoramic views, that represents local information by taking into account spatial context, is suggested.

Furthermore, a number of optimization techniques are applied throughout the process, for enhancing the retrieval performance. [20]

4 Experimental Results

The experimental evaluation is based on the Precision-Recall curves and five quantitative measures: Nearest Neighbor (NN), First Tier (FT), Second Tier (ST), E-measure (E) and Discounted Cumulative Gain (DCG) [9, 22] for the classes of each corresponding dataset.

4.1 ROSy and SymPan+ Pose Normalization Methods

For the evaluation of the ROSy pose normalization method, we have chosen a state-of-the-art 3D object retrieval methodology, by Papadakis et al. [14], as the evaluation vehicle.

Papadakis' 3D object retrieval system, in its original form, uses a combination of the CPCA and NPCA algorithms to achieve pose normalization of a 3D model. ROSy itself has similar performance to CPCA and NPCA. However, the combination of the three pose normalization methods (namely ROSy+) gives a significant boost to the discriminative power of the retrieval process, outperforming the original hybrid (CPCA, NPCA) approach.

Similar to the way that the ROSy+ system has been used for the quantitative evaluation of the ROSy pose normalization method, for SymPan+ we have chosen the PANORAMA state-of-the-art 3D object retrieval system, by Papadakis et al. [15] as the evaluation vehicle. The proposed method replaces the NPCA pose normalization method in the existing hybrid scheme.

The direct effect of the proposed alignment methods can be evaluated by comparing against the original 3D object retrieval methods' performance. In terms of object retrieval performance, we compared against DLA [3], GSMD+SHD+R [12], Lightfield [4], SH-GEDT [10] and DESIRE [25].

In Fig. 5, using the experimental results given in [18, 17], we illustrate the P-R scores for the test subset of the PSB dataset, for the proposed pose normalization methods.

ROSy+ is able to achieve an average performance gain of about 3% over the original hybrid approach (mean value over the quantitative measures used). Furthermore, it is clear that ROSy+ performs better than state-of-the-art methods by an average of 2% - 5%. SymPan+ improves the discriminative power of the PANORAMA 3D object retrieval system by an average of 7% over the original approach. Furthermore, the SymPan+ method exhibits improved performance over ROSy+ by an average of 2 - 3%.

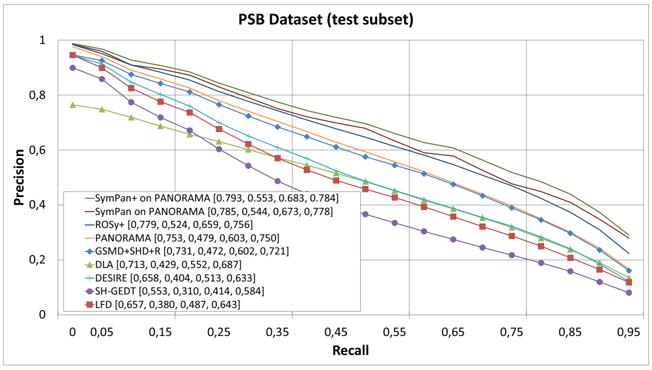

Fig. 5: Precision-Recall plot for the Princeton Shape Benchmark test dataset. SymPan, SymPan+ 3D model pose normalization methods on PANORAMA retrieval results are compared against state-of-the-art 3D object retrieval techniques.

4.2 ConTopo++ Non-Rigid 3D Object Retrieval Method

In the sequel, we compare the proposed non-rigid 3D object retrieval method ConTopo++ against other state-of-the-art methods on standard datasets.

In Fig. 6 we illustrate the P-R scores of the proposed method against the published results of the SHREC'10 *Non-rigid 3D Models* dataset. It is clear that the proposed method outperforms the track contestants, even though the published results were already of high performance.

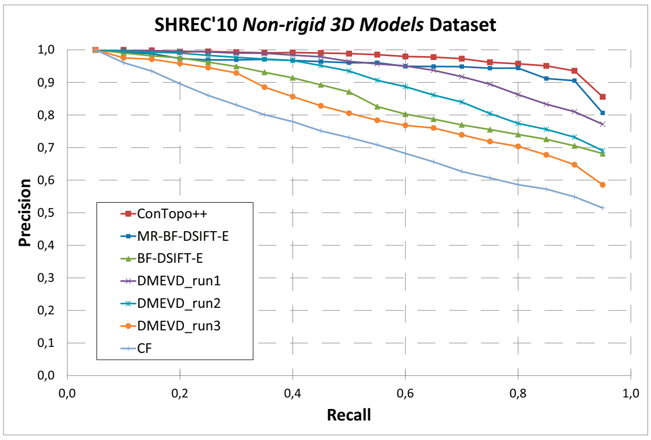

Fig. 6: Comparative results based on the average P-R scores for the SHREC'10 *Non-rigid 3D Models* dataset.

4.3 3D Object Retrieval Based on 2D Range Image Queries

The datasets that we used for the experimental evaluation of our proposed 3D object retrieval, based on 2D image query method are the following: (i) SHREC'09 *Querying with Partial Models* [6] and (ii) SHREC'10 *Range Scan Retrieval* [7]. We compared against existing results of the participating contestants.

More specifically, on the SHREC'09 *Querying with Partial Models* we compared against the variations of CMVD (Compact MultiView Descriptor) by Daras and Axenopoulos [5] and the BF-SIFT and BF-GridSIFT methods by Furuya and Ohbuchi. The P-R scores of Fig. 7 illustrate that the proposed method achieves superior performance compared to the variations of the CMVD, as well as both the BF-SIFT and the BF-GridSIFT retrieval methods.

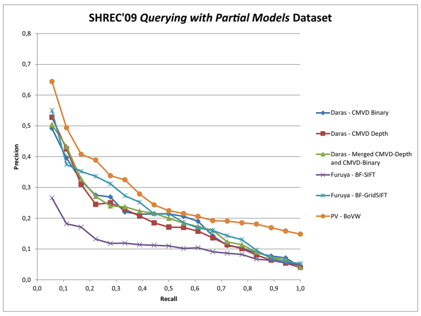

Fig. 7: Comparative results based on the average P-R scores for the SHREC'09 *Querying with Partial Models* dataset.

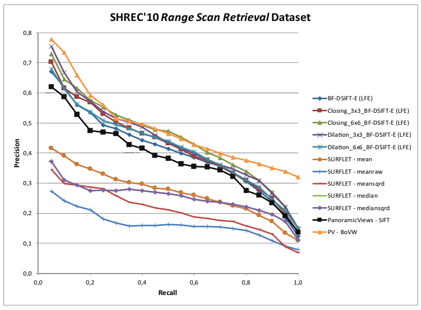

Fig. 8: Comparative results based on the average P-R scores for the SHREC'10 *Range Scan Retrieval* dataset.

On the SHREC'10 *Range Scan Retrieval* dataset we compared against the variations of the BF-DSIFT-E method proposed by Ohbuchi and Furuya [13] and the variations of the SURFLET method proposed by Hillebrand et al. [26]. The P-R scores of Fig. 8, illustrate that the proposed method outperforms the track contestants.

5 Conclusions

To address the problems of 3D model pose normalization, 3D object retrieval with applications to rigid and non-rigid models, as well as image based 3D object retrieval, four novel methodologies have been developed.

In the field of 3D model pose normalization two main novel methods, based on the reflective symmetry properties of 3D objects, have been proposed. All the proposed methods are able to produce high quality alignments of 3D objects, regardless of their originating class or morphology. These alignments are both stable and consistent.

To address the problem of non-rigid 3D object retrieval the ConTopo++ descriptor has been proposed. This non-rigid 3D object retrieval methodology is able to achieve high levels of retrieval accuracy and outperform many of the competing descriptors at a low computational cost. Fig. 9 illustrates some retrieval samples from the SHREC'10 *Non-rigid 3D Models* dataset.

Fig. 9: Sample queries from the SHREC'10 *Non-rigid 3D Models* dataset. First column indicates the query model and results are illustrated in ranking order. The thumbnails have been taken from the SHREC'10 *Non-rigid 3D Models* dataset.

In the field of image-based 3D object retrieval, we proposed a spatial histograms strategy in a Bag-of-Visual-Words context that fits the information present in panoramic views of 3D objects to the task of partial matching. This improved 3D object retrieval methodology, was evaluated on the SHREC'09 *Querying with Partial Models* and SHREC'10 *Range Scan Retrieval* tracks against the corresponding state-of-the-art 3D object retrieval methodologies. In every case, the proposed method outperforms competing descriptors.

The described methodologies have proven to be robust in terms of retrieval accuracy and outperformed previous state-of-the-art methods in the corresponding evaluation tests. These tests were conducted on publicly available datasets.

Bibliography

[1] Mirela Ben-Chen and Craig Gotsman. Characterizing shape using conformal factors. In Perantonis et al. [16], pages 1–8.

[2] Benjamin Bustos, Daniel A. Keim, Dietmar Saupe, Tobias Schreck, and Dejan V. Vranic. An experimental comparison of feature-based 3D retrieval methods. In *3DPVT*, pages 215–222. IEEE Computer Society, 2004.

[3] Mohamed Chaouch and Anne Verroust-Blondet. Alignment of 3D models. *Graphical Models*, 71(2):63–76, 2009.

[4] Ding-Yun Chen, Xiao-Pei Tian, Yu-Te Shen, and Ming Ouhyoung. On visual similarity based 3D model retrieval. *Comput. Graph. Forum*, 22(3):223–232, 2003.

[5] Petros Daras and Apostolos Axenopoulos. A compact multi-view descriptor for 3D object retrieval. In Stefanos D. Kollias and Yannis S. Avrithis, editors, *CBMI*, pages 115–119. IEEE Computer Society, 2009.

[6] Helin Dutagaci, Afzal Godil, Apostolos Axenopoulos, Petros Daras, Takahiko Furuya, and Ryutarou Ohbuchi. SHREC'09 track: Querying with partial models. In Michela Spagnuolo, Ioannis Pratikakis, Remco C. Veltkamp, and Theoharis Theoharis, editors, *3DOR*, pages 69–76. Eurographics Association, 2009.

[7] Helin Dutagaci, Afzal Godil, Chun Pan Cheung, Takahiko Furuya, Ulrich Hillenbrand, and Ryutarou Ohbuchi. SHREC'10 track: Range scan retrieval. In Mohamed Daoudi, Tobias Schreck, Michela Spagnuolo, Ioannis Pratikakis, Remco C. Veltkamp, and Theoharis Theoharis, editors, *3DOR*, pages 109–115. Eurographics Association, 2010.

[8] Jeffrey Goldsmith and John Salmon. Automatic creation of object hierarchies for ray tracing. *IEEE Comput. Graph. Appl.*, 7(5):14–20, 1987.

[9] Kalervo Järvelin and Jaana Kekäläinen. Cumulated gain-based evaluation of IR techniques. *ACM Trans. Inf. Syst.*, 20(4):422–446, 2002.

[10] Michael M. Kazhdan, Thomas A. Funkhouser, and Szymon Rusinkiewicz. Rotation invariant spherical harmonic representation of 3D shape descriptors. In Leif Kobbelt, Peter Schröder, and Hugues Hoppe, editors, *Symposium on Geometry Processing*, volume 43 of *ACM International Conference Proceeding Series*, pages 156–164. Eurographics Association, 2003.

[11] Michael M. Kazhdan, Thomas A. Funkhouser, and Szymon Rusinkiewicz. Symmetry descriptors and 3D shape matching. In Jean-Daniel Boissonnat and Pierre Alliez, editors, *Symposium on Geometry Processing*, volume 71 of *ACM International Conference Proceeding Series*, pages 115–123. Eurographics Association, 2004.

[12] Zhouhui Lian, Paul L. Rosin, and Xianfang Sun. Rectilinearity of 3D meshes. *International Journal of Computer Vision*, 89(2-3):130–151, 2010.

[13] Ryutarou Ohbuchi and Takahiko Furuya. Scale-weighted dense bag of visual features for 3D model retrieval from a partial view 3D model. In *Computer*

Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on, pages 63 – 70, 2009.

[14] Panagiotis Papadakis, Ioannis Pratikakis, Theoharis Theoharis, Georgios Passalis, and Stavros J. Perantonis. 3D object retrieval using an efficient and compact hybrid shape descriptor. In Perantonis et al. [16], pages 9–16.

[15] Panagiotis Papadakis, Ioannis Pratikakis, Theoharis Theoharis, and Stavros J. Perantonis. Panorama: A 3d shape descriptor based on panoramic views for unsupervised 3d object retrieval. *International Journal of Computer Vision*, 89(2-3):177–192, 2010.

[16] Stavros J. Perantonis, Nickolas S. Sapidis, Michela Spagnuolo, and Daniel Thalmann, editors. *Eurographics Workshop on 3D Object Retrieval, 3DOR 2008, Crete, Greece, 2008. Proceedings*. Eurographics Association, 2008.

[17] Konstantinos Sfikas, Ioannis Pratikakis, and Theoharis Theoharis. SymPan: 3D Model Pose Normalization via Panoramic Views and Reflective Symmetry. In Umberto Castellani, Tobias Schreck, Silvia Biasotti, Ioannis Pratikakis, Afzal Godil, and Remco C. Veltkamp, editors, *3DOR*, pages 41–48. Eurographics Association, 2013.

[18] Konstantinos Sfikas, Theoharis Theoharis, and Ioannis Pratikakis. ROSy+: 3D object pose normalization based on PCA and reflective object symmetry with application in 3D object retrieval. *International Journal of Computer Vision*, 91(3):262–279, 2011.

[19] Konstantinos Sfikas, Theoharis Theoharis, and Ioannis Pratikakis. Non-rigid 3D object retrieval using topological information guided by conformal factors. *The Visual Computer*, 28(9):943–955, 2012.

[20] Konstantinos Sfikas, Theoharis Theoharis, and Ioannis Pratikakis. 3D object retrieval via range image queries in a Bag-of-Visual-Words context. *The Visual Computer*, 29(12):1351–1361, 2013.

[21] Konstantinos Sfikas, Theoharis Theoharis, and Ioannis Pratikakis. Pose normalization of 3D models via reflective symmetry on panoramic views. submitted to, 2014.

[22] Philip Shilane, Patrick Min, Michael M. Kazhdan, and Thomas A. Funkhouser. The princeton shape benchmark. In *SMI*, pages 167–178. IEEE Computer Society, 2004.

[23] Johan W. H. Tangelder and Remco C. Veltkamp. A survey of content based 3D shape retrieval methods. *Multimedia Tools Appl.*, 39(3):441–471, 2008.

[24] Gino van den Bergen. Efficient collision detection of complex deformable models using AABB trees. *J. Graph. Tools*, 2(4):1–13, 1998.

[25] Dejan V. Vranic. DESIRE: a composite 3D-shape descriptor. In *ICME*, pages 962–965. IEEE, 2005.

[26] Eric Wahl, Ulrich Hillenbrand, and Gerd Hirzinger. Surflet-pair-relation histograms: A statistical 3D-shape representation for rapid classification. In *3DIM*, pages 474–482. IEEE Computer Society, 2003.

[27] Titus B. Zaharia and Françoise J. Prêteux. 3D versus 2D/3D shape descriptors: a comparative study. In Edward R. Dougherty, Jaakko Astola, and Karen O. Egiazarian, editors, *Image Processing: Algorithms and Systems*, volume 5289 of *SPIE Proceedings*, pages 47–58. SPIE, 2004.