3D Object Partial Matching Using Panoramic Views

Konstantinos Sfikas[1], Ioannis Pratikakis[1,2], Anestis Koutsoudis[1], Michalis Savelonas[1], and Theoharis Theoharis[3,4]

[1] ATHENA Research and Innovation Center, Xanthi, Greece,
[2] Democritus University of Thrace, Department of Electrical & Computer Engineering, Xanthi, Greece,
[3] Department of Informatics and Telecommunications, University of Athens, Greece
[4] IDI, Norwegian University of Science and Technology (NTNU), Norway

Abstract. In this paper, a methodology for 3D object partial matching and retrieval based on range image queries is presented. The proposed methodology addresses the retrieval of complete 3D objects based on artificially created range image queries which represent partial views. The core methodology relies upon Dense SIFT descriptors computed on panoramic views. Performance evaluation builds upon the standard measures and a challenging 3D pottery dataset originated from the Hampson Archeological Museum collection.

Keywords: 3D object retrieval, partial matching, 3D pottery dataset.

1 Introduction

Generic 3D object retrieval has considerably matured and a number of accurate and robust descriptors have been proposed in the literature [5, 14, 21, 22, 27]. However, there are still many challenges regarding the retrieval of 3D models that originate from specific domains and/or exhibit special characteristics. One such domain is Cultural Heritage (CH) which includes 3D models that are usually deteriorate, due to aging, their shape is altered due to environmental factors and in most cases only incomplete artefacts have been preserved.

The problem with the partial data is that it is not possible to effectively match them against a full 3D model representation, since most of it may be missing. The representation gap makes it difficult to extract a signature that will be, at least partially, similar when presented with a complete 3D model and when presented with a partial query of a similar object.

We have addressed this challenge by using a panoramic view representation that is able to encode the 3D surface characteristics onto a 2D image map, as well as range image representation which can be mapped to the panoramic views through interest points correspondence. For the complete 3D models, we compute a number of panoramic views on axes, which are perpendicular to the faces of a dodecahedron. Each axis defines three panoramic view cylinders

A. Petrosino, L. Maddalena, P. Pala (Eds.): ICIAP 2013 Workshops, LNCS 8158, pp. 169–178, 2013.

(one for the axis itself and two more for any two axes in order to make up an orthonormal basis, along with the first one.). Then, we apply the Dense SIFT (DSIFT) algorithm [9, 18] to the points and calculate the corresponding descriptor. In the same spirit, for the partial objects, we initially compute a range image representation used as the query model for which the DSIFT algorithm is calculated.

The remainder of the paper is structured as follows. In Section 2, recent work on 3D model retrieval based on range image queries is discussed. Section 3 details the proposed method and Section 4 presents experimental results achieved in the course of the method's evaluation. Finally, conclusions are drawn in Section 5.

2 Related Work

Over the past few years, the number of works addressing the problems of partial 3D object retrieval has increased significantly. Although this task still remains non-trivial, very important steps have been made in the field.

Stavropoulos et al. [25] present a retrieval method based on the matching of salient features between the 3D models and query range images. Salient points are extracted from vertices that exhibit local maxima in terms of protrusion mapping for a specific window on the surface of the model. A hierarchical matching scheme based is used for matching. The authors experimented on range images acquired from the SHape REtrieval Contest 2007 (SHREC'07) *Watertight Models* [11] and the Princeton Shape Benchmark (PSB) standard [24] datasets. Chaouch and Verroust-Blondet [3] present a 2D/3D shape descriptor which is based on either silhouette or depth-buffer images. For each 3D model a set of six projections in calculated for both silhouette and depth-buffers. The 2D Fourier transform is then computed on the projection. Furthermore, they compute a relevance index measure which indicates the density of information contained in each 2D view. The same authors in [4] propose a method where a 3D model is projected to the faces of its bounding box, resulting in 6 depth buffers. Each depth buffer is then decomposed into a set of horizontal and vertical depth lines that are converted to state sequences which describe the change in depth at neighboring pixels. Experimentations were conducted on range images artificially acquired from the PSB dataset. Shih et al. [23] proposed the elevation descriptor where six depth buffers (elevations) are computed from the faces of the 3D model bounding box and each buffer is described by a set of concentric circular areas that give the sum of pixel values within the corresponding areas. The models were selected from the standard PSB dataset.

Experimenting on the SHREC'09 *Querying with Partial Models* [7] dataset, Daras and Axenopoulos in [6] present a view-based approach for 3D model retrieval. The 3D model is initially pose normalized and a set of binary (silhouette) and range images are extracted from predefined views on a 32-hedron. The set of features computed on the views are the Polar-Fourier transform, Zernike moments and Krawtchouk moments. Each query image is compared to all the extracted views of each model of the dataset. Ohbuchi et al. [20] extract features from 2D range images of the model viewed from uniformly sampled locations on a view sphere. For every range image, a set of multi-scale 2D visual features are computed using the Scale Invariant Feature Transform (SIFT) [18]. Finally, the features are integrated into a histogram using the Bag-of-Features approach [10]. The same authors enhanced their approach by pre-processing the range images, in order to minimize interfere caused by any existing occlusions,as well as by refining the positioning of SIFT interest points, so that higher resolution images are favored [9, 19]. Their works have experimented on and have participated on both corresponding SHREC'09 *Querying with Partial Models* and SHREC'10 *Range Scan Retrieval* [8] contests. Wahl et al. [28] propose a four-dimensional feature that parameterizes the intrinsic geometrical relation of an oriented surface point pair (surflets). For a 3D model a set of surflet pairs is computed over a number of uniformly sampled viewing directions on the surrounding sphere. This work was one of the two contestants of the SHREC'10 *Range Scan Retrieval* track. Finaly, Koutsoudis et al. [16, 17] presented a set of 3D shape descriptors designed for content based retrieval of complete or nearly complete 3D vessels. Their performance evaluation experiments were performed on a dataset that included among others a subset of Virtual Hampson Museum 3D collection [15].

In this work, we propose the use of a 3D model representation that bridges the gap between the 3D model and the range scan.

3 Methodology

The main steps of the proposed methodology for partial 3D object retrieval via range queries are: (i) shape descriptors extraction from each full 3D model of the dataset (off-line), (ii) range image computation of the partial query model, (iii) shape descriptor extraction from the range image query (on-line) and (iv) matching of the query descriptor against the descriptor of each 3D model of the dataset.

In the case of the full 3D models, a number of panoramic views of each model are extracted on viewpoint axes that are defined by a dodecahedron, thus extending the PANORAMA [21] method to multiple viewpoint axes. Each axis defines three panoramic view cylinders (one for the axis itself and two more for any two axes so that, along with the first one to make up an orthonormal basis). To obtain a panoramic view, we project the model to the lateral surface of a cylinder of radius R and height $H = 2R$, centered at the origin with its axis parallel to one of the coordinate axes (see Fig. 1a). We set the value of R to $2 * d_{max}$ where d_{max} is the maximum distance of the model's surface from its centroid. In the following, we parameterize the lateral surface of the cylinder using a set of points $s(\phi, y)$ where $\phi \in [0, 2\pi]$ is the angle in the xy plane, $y \in [0, H]$ and we sample the ϕ and y coordinates at rates B and $2B$, respectively (we set $B = 256$). Thus, we obtain the set of points $s(\phi_u, y_v)$, where $\phi_u = u * 2\pi/(2B)$, $y_v = v * H/(B)$, $u \in [0, 2B - 1]$ and $v \in [0, B - 1]$. These points are shown in Fig. 1b.

The next step is to determine the value at each point $s(\phi_u, y_v)$. The computation is carried out iteratively for $v = 0, 1, ..., B - 1$, each time considering the

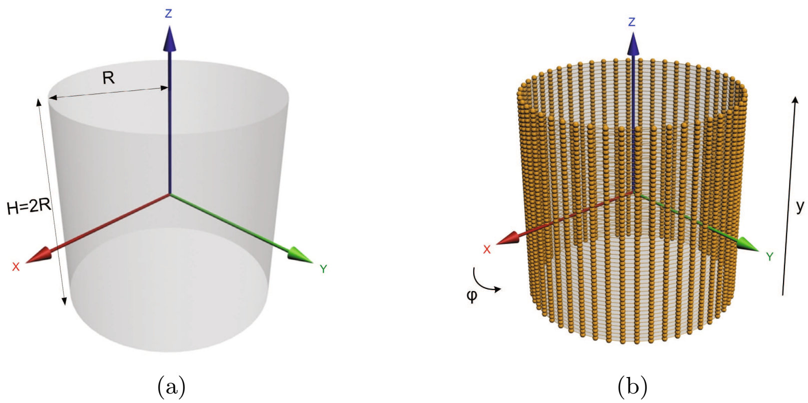


Fig. 1. (a) A projection cylinder for the acquisition of a 3D model's panoramic view and (b) the corresponding discretization of its lateral surface to the set of points $s(\phi_u, y_v)$

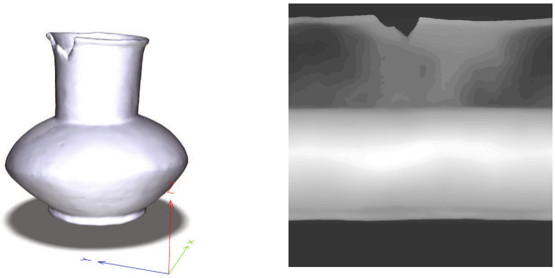
set of coplanar $s(\phi_u, y_v)$ points i.e. a cross section of the cylinder at height y_v and for each cross section we cast rays from its center c_v in the ϕ_u directions. To capture the position of the model surface, for each cross section at height y_v we compute the distances from c_v to the intersections of the model's surface with the rays at each direction ϕ_u.

Let $pos(\phi_u, y_v)$ denote the distance of the furthest from c_v point of intersection between the ray emanating from c_v in the ϕ_u direction and the model's surface; then $s(\phi_u, y_v) = pos(\phi_u, y_v)$. Thus the value of a point $s(\phi_u, y_v)$ lies in the range $[0, R]$, where R denotes the radius of the cylinder.

A cylindrical projection can be viewed as a 2D gray-scale image where pixels correspond to the $s(\phi_u, y_v)$ intersection points in a manner reminiscent of cylindrical texture mapping [26] and their values are mapped to the [0,1] range. In Fig. 2a, we show an example 3D model and in Fig. 2b the unfolded visual representation of its corresponding cylindrical projection $s(\phi_u, y_v)$.

Once the panoramic views have been extracted, the DSIFT descriptor is calculated on the cylindrical depth images. The first step of the DSIFT computation, is the extraction of a number of interest points, for which the DSIFT descriptors are calculated. The original implementation by Lowe, calculates these interest points through the Difference of Gaussians (DoG) method, which is geared towards enhancing the edges and other details present in the image. It has been experimentally found that the calculation of the DSIFT descriptors over the complete image for a large number of randomly selected points $[1, 2, 9, 18]$ (frequently defined as Dense SIFT/ DSIFT, in the literature), instead of selecting a limited number of interest points, yields better results in terms of retrieval accuracy.

In the case of query models, range images are computed by taking the depth buffer projection. The selected size of sampling is 256×256 pixels and the DSIFT descriptor is computed directly on the range image. The histogram produced out of the total number of DSIFT descriptors is stored as the signature.



(a) (b)

Fig. 2. (a) An example 3D model and (b) its corresponding cylindrical projection on the z-axis.

Finally, the descriptors for each interest point of the range image must be matched against the 3D model dataset descriptors. To this end, every DSIFT point of the range image is compared against every DSIFT point of a 3D model's panoramic views, for which the minimum is kept in terms of $L2$ distance. The average of these minimum distances is stored as the final distance of the query and the target model.

4 Evaluation

For the experimental evaluation we have use a dataset related to the cultural heritage domain which is a 3D pottery dataset originated from the Hampson Archeological Museum collection.

The Hampson Archeological Museum collection composes a major source of information on the lives and history of pre-Columbian people of the Mississippi river valley [12]. The Centre of Advanced Spatial Technologies - University of Arkansas worked on the digitisation of numerous artefacts from the Hampson museum collection using a Konica-Minolta Vivid 9i short-range 3D laser scanner. The digitisation was performed at a precision close to 0.2 mm. The 3D digital replicas are covered by the creative common 3.0 license and are offered for online browsing or downloading in both high (>1M facets) and low (<= 25K facets) resolutions [13].

As a testbed for content based retrieval and partial matching experiments, we have used 384 models of low resolution, that were downloaded from the website of the museum along with associated metadata information, as a testbed for content based retrieval and partial matching experiments. Initially the models were classified by the museum into six general classes (Bottle, Bowl, Jar, Effigy, Lithics and Others). As the current classification did not ensure similarities based on shape within a given class, we performed an extended shape-oriented classification. We initially organised the models into thirteen classes of different

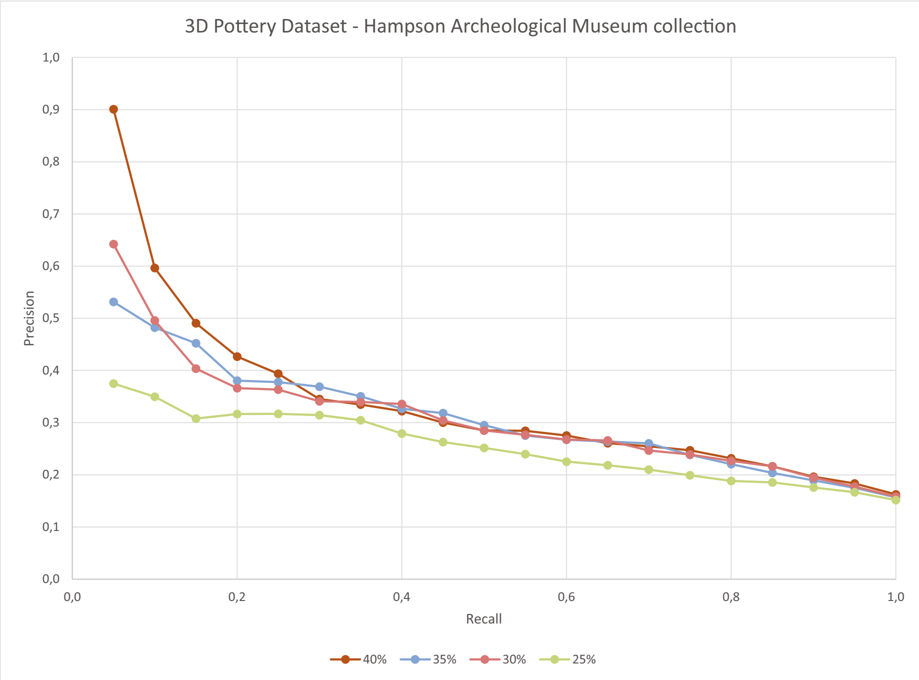


Fig. 3. Average P-R scores for the pottery dataset originating from the Hampson Archeological Museum collection. Illustrated is the performance of the presented method, obtained using queries with reduced surface (40% - 25%) with respect to the surface of the original complete 3D models.

populations (Bottles, Bowls 1 - 4, Figurines, Masks, Pipes, Tools, Tripod-Base Vessels, Conjoined Vessels, Twin Piped Vessels and Others).In the sequel five of these classes (all Bottle and Bowls classes) were further divided into 15 subclasses resulting in a total of 23 distinct classes.

Since this dataset does not contain any partial 3D object that can be used as query, we artificially created a set of 20 partial queries by slicing and cap filling an amount of complete 3D objects, originating from those classes that are densely populated. The partial queries comprise objects with a reduced surface compared to the original 3D object by a factor which ranges from 40% to 25% with a step of 5%.

Our experimental evaluation is based on Precision-Recall (P-R) plots and five quantitative measures: Nearest Neighbor (NN), First Tier (FT), Second Tier (ST), E-measure (E) and Discounted Cumulative Gain (DCG) [24] for the classes of the corresponding datasets. For every query model that belongs to a class C, recall denotes the percentage of models of class C that are retrieved and precision denotes the proportion of retrieved models that belong to class C

over the total number of retrieved models. The maximum score is 100% for both quantities. Nearest Neighbor (NN) indicates the percentage of queries where the closest match belongs to the query class. First Tier (FT) and Second Tier (ST) statistics, measure the recall value for the $(D-1)$ and $2(D-1)$ closest matches respectively, where D is the cardinality of the query class. E-measure combines precision and recall metrics into a single number and the DCG statistic weights correct results near the front of the list more than correct results later in the ranked list under the assumption that a user is less likely to consider elements near the end of the list [24].

In Figure 3 we illustrate the average P-R scores for the presented 3D model retrieval method using the artificially created partial queries of the complete pottery dataset. Results are presented for various amounts of partiality. Table 1 shows the corresponding five quantitative measures. Figure 4 illustrates query examples and the corresponding list of the retrieved 3D models.

Our experimental results show that the proposed method is able to handle quite well the problem of partial to complete matching and illustrates stability in its performance with respect to the percentage of partiality. Even at the high partiality levels where only one quarter of the surface of original model is available, the results are promising.

Table 1. Five quantitative measures for the presented 3D object retrieval methodology, using partial queries on the pottery dataset. All measures are normalized.

Method	NN	FT	ST	E	DCG
Proposed Method (25%)	0.23	0.227	0.388	0.185	0.587
Proposed Method (30%)	0.428	0.289	0.495	0.228	0.655
Proposed Method (35%)	0.619	0.372	0.536	0.327	0.713
Proposed Method (40%)	0.857	0.288	0.508	0.237	0.683

Fig. 4. Example retrieval results from the pottery dataset. At each row, a partial query (column 1) and a ranked list of the retrieved 3D objects (columns 2 - 8) are shown.

The proposed method was tested on a Core2Quad 2.5 GHz system, with 6 GB of RAM, running Matlab R2012b. The system was developed in a hybrid Matlab/C++/OpenGL architecture. The average descriptor extraction time for an 100,000 faces 3D model is about 5 seconds.

5 Conclusions

We have presented a method for partial matching based on DSIFT and panoramic views. This method is able to encode the continuous characteristics of the (partial) 3D models into a 2D representation, thus preserving model structure. The performance of the method is evaluated on a pottery dataset originated from the Hampson Archeological Museum collection of historical artefacts. We have shown that using the proposed methodology, we can attain retrieval results which are not severely affected by the reduction in 3D model surface. This work sets a baseline for methodologies addressing partial 3D object retrieval for cultural heritage datasets.

Acknowledgments. The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n ° 600533 PRESIOUS.

References

1. Bosch, A., Zisserman, A., Muñoz, X.: Scene classification via pLSA. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3954, pp. 517–530. Springer, Heidelberg (2006)
2. Bosch, A., Zisserman, A., Muñoz, X.: Image classification using random forests and ferns. In: ICCV 2007, pp. 1–8. IEEE (2007)
3. Chaouch, M., Verroust-Blondet, A.: Enhanced 2D/3D approaches based on relevance index for 3D-shape retrieval. In: SMI, p. 36. IEEE Computer Society (2006)
4. Chaouch, M., Verroust-Blondet, A.: A new descriptor for 2D depth image indexing and 3D model retrieval. In: IEEE International Conference on Image Processing, ICIP 2007, vol. 6, pp. 373–376 (2007)
5. Chen, D.Y., Tian, X.P., Shen, Y.T., Ouhyoung, M.: On visual similarity based 3D model retrieval. Comput. Graph. Forum, 223–232 (2003)
6. Daras, P., Axenopoulos, A.: A compact multi-view descriptor for 3D object retrieval. In: Proceedings of the 2009 Seventh International Workshop on Content-Based Multimedia Indexing, CBMI 2009, pp. 115–119. IEEE Computer Society, Washington, DC (2009)
7. Dutagaci, H., Godil, A., Axenopoulos, A., Daras, P., Furuya, T., Ohbuchi, R.: SHREC'09 track: Querying with partial models. In: Spagnuolo, M., Pratikakis, I., Veltkamp, R., Theoharis, T. (eds.) Eurographics Workshop on 3D Object Retrieval, pp. 69–76. Eurographics Association, Munich (2009)
8. Dutagaci, H., Godil, A., Cheung, C.P., Furuya, T., Hillenbrand, U., Ohbuchi, R.: SHREC'10 track: Range scan retrieval. In: Daoudi, M., Schreck, T. (eds.) Eurographics Workshop on 3D Object Retrieval, pp. 109–115. Eurographics Association, Norrköping (2010)

9. Furuya, T., Ohbuchi, R.: Dense sampling and fast encoding for 3D model retrieval using bag-of-visual features. In: Proceedings of the ACM International Conference on Image and Video Retrieval, CIVR 2009, pp. 26:1–26:8. ACM, New York (2009)
10. Geurts, P., Ernst, D., Wehenkel, L.: Extremely randomized trees. Machine Learning 36(1), 3–42 (2006)
11. Giorgi, D., Biasotti, S., Paraboschi, L.: SHape REtrieval contest 2007: Watertight models track (2007)
12. Hampson: The hampson archaeological museum state park, http://www.arkansasstateparks.com/hampsonmuseum/#.UahgmdJ8lRg (accessed on May 31, 2013)
13. Hampson: The virtual hampson museum project, http://hampson.cast.uark.edu (accessed on May 31, 2013)
14. Kazhdan, M., Funkhouser, T., Rusinkiewicz, S.: Rotation invariant spherical harmonic representation of 3D shape descriptors. In: Proceedings of the 2003 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing, SGP 2003, pp. 156–164. Eurographics Association, Aire-la-Ville (2003)
15. Koutsoudis: A 3D pottery content based retrieval benchmark dataset, http://www.ipet.gr/~akoutsou/benchmark (accessed on May 31, 2013)
16. Koutsoudis, A., Chamzas, C.: 3D pottery shape matching using depth map images. Journal of Cultural Heritage 12(2), 128–133 (2011)
17. Koutsoudis, A., Pavlidis, G., Liami, V., Tsiafakis, D., Chamzas, C.: 3D pottery content-based retrieval based on pose normalisation and segmentation. Journal of Cultural Heritage 11(3), 329–338 (2010)
18. Lowe, D.G.: Object recognition from local scale-invariant features. In: Proceedings of the International Conference on Computer Vision, ICCV 1999, vol. 2, pp. 1150–1157. IEEE Computer Society, Washington, DC (1999)
19. Ohbuchi, R., Furuya, T.: Scale-weighted dense bag of visual features for 3D model retrieval from a partial view 3D model. In: 2009 IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops), September 27-October 4, pp. 63–70 (2009)
20. Ohbuchi, R., Osada, K., Furuya, T., Banno, T.: Salient local visual features for shape-based 3D model retrieval. In: Shape Modeling International, pp. 93–102. IEEE (2008)
21. Papadakis, P., Pratikakis, I., Theoharis, T., Perantonis, S.J.: 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)
22. Sfikas, K., Pratikakis, I., Theoharis, T.: ConTopo: Non-rigid 3D object retrieval using topological information guided by conformal factors. In: Laga, H., Schreck, T., Ferreira, A., Godil, A., Pratikakis, I., Veltkamp, R.C. (eds.) 3DOR, pp. 25–32. Eurographics Association (2011)
23. Shih, J.L., Lee, C.H., Wang, J.T.: A new 3D model retrieval approach based on the elevation descriptor. Pattern Recognition 40(1), 283–295 (2007)
24. Shilane, P., Min, P., Kazhdan, M., Funkhouser, T.: The princeton shape benchmark. In: Proceedings of the Shape Modeling International 2004, SMI 2004, pp. 167–178. IEEE Computer Society, Washington, DC (2004)
25. Stavropoulos, T.G., Moschonas, P., Moustakas, K., Tzovaras, D., Strintzis, M.G.: 3D model search and retrieval from range images using salient features. IEEE Transactions on Multimedia 12(7), 692–704 (2010)

26. Theoharis, T., Papaioannou, G., Platis, N., Patrikalakis, N.M.: Graphics and Visualization: Principles & Algorithms. A. K. Peters, Ltd., Natick (2007)
27. Vranic, D.V.: DESIRE: A composite 3D-shape descriptor. In: ICME, pp. 962–965. IEEE (2005)
28. Wahl, E., Hillenbrand, U., Hirzinger, G.: Surflet-pair-relation histograms: A statistical 3D-shape representation for rapid classification. In: 3DIM, pp. 474–481. IEEE Computer Society (2003)