Graph-Based Combinations of Fragment Descriptors for Improved 3D Object Retrieval

Tobias Schreck
Data Analysis and Visualization Group
University of Konstanz, Germany

Maximilian Scherer
Interactive Graphics Systems
Technische Universität Darmstadt, Germany

Michael Walter
Interactive Graphics Systems
Technische Universität Darmstadt, Germany

Benjamin Bustos
KDW+PRISMA
Department of Computer Science, University of Chile, Chile

Sang Min Yoon
Yonsei Institute of Convergence Technology, Yonsei University
Korea

Arjan Kuijper
Fraunhofer Institute for Computer Graphics Research
Darmstadt, Germany

ABSTRACT

3D Object Retrieval is an important field of research with many application possibilities. One of the main goals in this research is the development of discriminative methods for similarity search. The descriptor-based approach to date has seen a lot of research attention, with many different extraction algorithms proposed. In previous work, we have introduced a simple but effective scheme for 3D model retrieval based on a spatially fixed combination of 3D object fragment descriptors. In this work, we propose a novel flexible combination scheme based on finding the best matching fragment descriptors to use in the combination. By an exhaustive experimental evaluation on established benchmark data we show the capability of the new combination scheme to provide improved retrieval effectiveness. The method is proposed as a versatile and inexpensive method to enhance the effectiveness of a given global 3D descriptor approach.

Categories and Subject Descriptors

I.3.7 [**Computer Graphics**]: Three-Dimensional Graphics and Realism; H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval

General Terms

Algorithms, Experimentation

Keywords

3D object retrieval, descriptor combinations, effectiveness

MMSys'12, February 22-24, 2012, Chapel Hill, North Carolina, USA

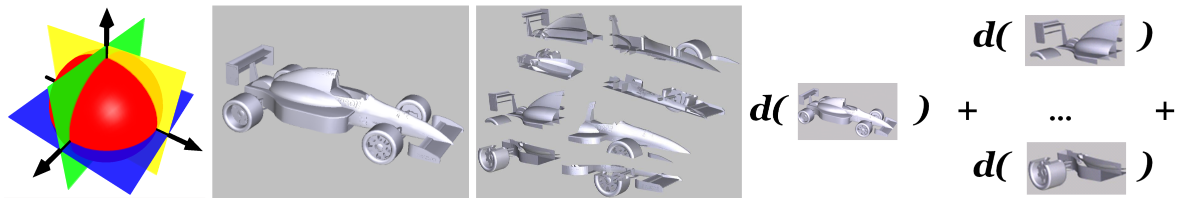
1. INTRODUCTION

3D Object Retrieval is concerned with devising methods for similarity search in databases of 3D objects. To this end, 3D similarity functions are researched which provide content-based access. Together with query-by-example or query-by-sketch, access to and re-usage of existing content in applications such as Computer-Aided Design, Simulation, and Visualization become feasible. Also, analytic applications such as cluster analysis rely on similarity functions. Under the popular descriptor-based approach, descriptors (or signatures) are computed for each 3D object, and a distance function defined on the descriptors is taken as a similarity measure. To date, many different descriptor extraction methods have been proposed, with no single method showing best for any kind of application.

In [4], we proposed a scheme for the generic improvement of given global 3D object descriptors. It heuristically partitions a 3D object into a number of fragments. Descriptors are extracted from the global object and from its fragments, and all are combined to form a joint descriptor. The discrimination performance of the joint descriptor was shown to outperform the performance of a number of baseline description extraction methods in experiments. A major drawback of the original scheme was that it used a *spatially fixed* matching of fragment descriptors in the combination. In this work, we introduce a *spatially flexible*, graph-based matching of fragment descriptors which improves substantially over the original fixed scheme. The scheme is proposed as a simple, yet effective method to boost the retrieval performance of given global 3D descriptors. It furthermore is inexpensive in terms of implementation and runtime complexity required. Given its independence from any particular descriptor implementation, it is orthogonal to descriptor definition and can accommodate additional descriptors to be developed in the future.

2. RELATED WORK

3D Object Retrieval is an active field of research, concerned with the definition and evaluation of methods for similarity search for 3D objects. Its roots are in Computa-



(a) (b) (c) (d)

Figure 1: Our approach is based on combining global 3D descriptors by concatenation of feature vectors extracted from fragments of the respective model (d). A whole 3D model, such as illustrated in (b), is partitioned by an Octree scheme (a) into eight individual fragments (c). We introduce a substantially improved new scheme for forming descriptor combinations, providing more effectiveness in the description.

tional Geometry and Computer Vision, but also Multimedia Databases and Computer Graphics. Early methods for comparison of 3D shapes include, for example, the Iterative Closest Points method [1], which computes an alignment of two point sets for comparison. Under the descriptor-based approach, descriptive measurements are calculated from 3D objects. Forming descriptors e.g., in form of feature vectors or graphs, these can be used to compare 3D shapes for similarity [14, 3]. *Global* methods extract one descriptor for a given 3D object, while *local* methods extract descriptors for local parts of the model. While global methods support a notion of global object similarity, local methods are capable to support partial similarity and a variety of invariance properties. Local methods can be considered more complex, in that they need to identify the number of local objects parts based on 3D object segmentation [5] or identification of interest points [12, 6] among others. While global methods often can employ structurally simple distance functions, local methods are often based on matching approaches or employ bag-of-feature [2] schemes, to compare sets of local descriptors.

Global descriptor methods typically are simple to implement and therefore, are often preferable in practical applications. However, also current evaluations show that to date, the search for effective global descriptors is not solved [7]. Improving the performance of global descriptors remains a challenging topic, and improvement approaches orthogonal to descriptor definition remain important [4].

3. APPROACH

We next recall the baseline procedure from [4] and present our new improved descriptor combination scheme.

3.1 Spatially Fixed Combination Scheme

Global 3D object descriptors produce one descriptor as the results of some global 3D shape analysis process. A prominent example includes image-based descriptors such as proposed in [17]. These (1) produce a number of reference 2D views of the objects, (2) calculate descriptors for each 2D view, e.g., Fourier or HOG descriptors, and (3) combine the view descriptors in a joint similarity function. However, global extraction techniques may suppress local information. This may result from specifics of the method. E.g., in case of view-based descriptors, occlusion may prevent relevant object detail to enter in the descriptor. Or, in case of sampling-based schemes such as 3D centroid descriptors, relevant object detail may be missed due to sampling artifacts. To overcome such implicit method problems, in [4] we proposed to partition a given 3D object into a number of fragments, and combine descriptors calculated for the whole model and all of its fragments. In particular, descriptors were given as feature vectors, and the combination FV^c was obtained by concatenation of normalized and weighted feature vectors:

$$FV^c = \frac{FV^g}{||FV^g||} \oplus \frac{w}{8}\frac{FV^1}{||FV^1||} \oplus ... \oplus \frac{w}{8}\frac{FV^8}{||FV^8||}, \qquad (1)$$

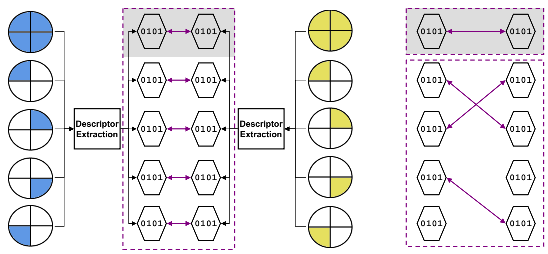
where FV^g is the global descriptor, and FV^n is the descriptor of the nth object fragment ($n = 1, \dots, 8$) from the Octree partitioning of the model. All descriptors in the combination are normalized to unit length; and weight w is used to scale the fragment descriptor importance relative to the global descriptor.

This approach allowed for a more complete description of the objects with respect to the aforementioned potential problems. Figure 1 illustrates the overall process. As every model was normalized for rotation prior to partitioning and descriptor extraction, and as the descriptor concatenation proceeds in order of the fragment indexes, the method yields a spatially fixed 1:1 mapping between object fragments (see Figure 2(a) for a 2D illustration). While the partitioning was done in a simple and heuristic way based on Octree partitioning, the approach managed to increase the discrimination performance, as compared to several baseline descriptors. Retrieval rates were improved up several percentage points, as compared to the original descriptors [4].

3.2 New Combination Scheme Based on Bipartite Graph Matching

A major drawback of the original scheme was that it combined the global and fragment descriptors in a spatially fixed scheme. The original spatially fixed 1:1 mapping may not be the best mapping in every case. In particular, many 3D objects show local symmetries along their main principal directions. While we assume the 1:1 spatially fixed matching is adequate for many classes, we further assume that a more flexible matching scheme is adequate for other models.

To this end, we introduce a more flexible scheme to the combination process, considering the problem of comparing the fragment descriptors as a bipartite graph matching problem. Specifically, we define an edge between each pair of fragment descriptors of the objects to be compared, weighted by the Manhattan distance between the descriptors of the respective fragments. We apply the *Hungarian method* [10] to compute a matching between the fragment descriptors,



(a) (b)

Figure 2: Combining global and fragment descriptors. (a) Spatially fixed combination employed in [4]). (b) Spatially flexible scheme proposed here, based on bipartite graph matching (three best matching fragments shown in illustration).

which minimizes the total sum of edge weights, in effect minimizing the dissimilarity between the matching solution. Finally, we form the descriptor concatenation according to the alignment of fragments as provided by the solution of the Hungarian algorithm. Figure 2(b) illustrates the matching. We allow two parameters for our combination scheme. Parameter w in Equation 1 (named *local_weight* in Section 4) determines the relative weight between the global and the fragment descriptors in the combined descriptor. Also, we allow to consider a variable number of fragment descriptors in the combination (parameter *parts* in Section 4). In particular, we allow to consider only a number of best matching edges from the bipartite graph matching solution (see Figure 2(b) for an illustration). The idea behind the latter is that two models may be similar even though that some parts of them might not be similar. Both parameters are examined experimentally in Section 4.

4. RESULTS

4.1 Experimental Setup

We experimentally evaluate our approach using three established 3D object retrieval benchmarks. The benchmarks are: The Purdue Engineering Benchmark (ESB, [9]), the Princeton Shape Benchmark Test Partition (PSB, [11]), and the SHREC 2009 New Generic Benchmark (SHREC, [8]). Each of these benchmarks comprises hundreds of 3D mesh models of generic (PSB, SHREC) and engineering (ESB) model typology. Each benchmark provides class labels for the contained objects. We performed retrieval experiments, in which for each benchmark, we use each classified object as a query object. We then calculate average R-Precision (or First Tier) [11] scores as the effectiveness measure to compare our combination schemes.

As to the considered global descriptors, analogous to [4] we consider the following descriptors: Rays with Spherical Harmonics representation (RSH), Silhouette (SIL), Depth Buffer (DBD), and Desire (DSR) descriptors [16, 15]. RSH, SIL, and DBD are established, standard global 3D object descriptors. DSR is a combined descriptor formed from concatenating RSH, SIL, and DBD descriptors. Also analogous to [4], we partition each 3D object by Octree partitioning into 8 fragment parts, after rotation, translation, and scale normalization. We then calculate descriptors for the whole models and all of their fragments.

To compare our combination schemes, we performed exhaustive experiments. For each descriptor and benchmark, we considered spatially fixed and flexible combination schemes. Specifically, we vary two parameters to model the spectrum of possible combination schemes. Parameter $parts = 0, \ldots, 8$ specifies the fragment selection method as follows: $p = 1, \ldots, 8$ allows for 1 up to 8 pairs of matched fragments to contribute to the distance while the rest is ignored (cf. also Section 3.2). $parts = 0$ corresponds to the spatially fixed combination scheme using all 8 fragments and represents the method studied in [4]. Parameter $local\_weight = [0, 1]$ gives the sum of the weights of all fragment descriptors relative to the global descriptor (corresponds to parameter w in Section 3.1). $local\_weight = 0$ corresponds to using only the global descriptor, while $local\_weight = 1$ corresponds to the global and fragment descriptors contributing to the overall distance function at equal weight.

4.2 Experimental Results

We performed exhaustive experiments calculating average R-Precision scores for each descriptor and benchmark, while varying parameters *parts* and *local_weight*. As we want to study two parameters, we chose to visually analyze the results using heatmap displays. In this display, the x-axis corresponds to the parameter *parts*, and the y-axis corresponds to the parameter *local_weight*, the latter sampled at steps of 0.1. We normalize the R-Precision scores over all cells linearly according to the minimum and maximum values occurring, and map the respective values inversely to a black-to-green color gradient. Thereby, shades of more intense green correspond to higher R-Precision scores (or better discrimination power)[1].

Figure 3 shows the results for the SHREC data set and the DBD descriptor in an annotated heatmap display. The first column (marked by a thin orange frame) shows the results for $parts = 0$, which is the baseline spatially fixed scheme combining all 8 fragment descriptors with the global descriptor. We see that in this scheme, R-precision is maximum for the smallest *local_weight* at 10%, and then decreases, as *local_weight* increases. Considering the columns to the right (large orange frame), we see that the flexible combination scheme manages to outperform the spatially fixed combination scheme, for a substantial area in the parameter space. This can be visually seen by comparing the green shades of the cells in the first column (fixed scheme) with the shades in the columns to the right. In particular, for the parameter range between 4 and 7 fragments allowed in the bipartite graph solution, and *local_weight* between 20% and 40%, we see maximum relative R-Precision results (yellow frame in the figure).

Table 1 shows a numeric comparison of the best R-Precision rates obtained for the two combination schemes for the DBD descriptor, and all three benchmark data sets. The spatially fixed scheme provides 42.8%, 34.7%, and 35.4% R-PRecision for the SHREC, PSB, and ESB data sets, at parameter *local_weight* of 10%, 10%, and 20%, respectively. The flexible combination scheme provides 44.9%, 35.8%, and 36.5% R-Precision at parameters *local_weight* of 20%, 10%, 20%,

[1]We recommend viewing the diagrams on a color printout of sufficient color contrast or on a monitor display, for best perception of shading differences.

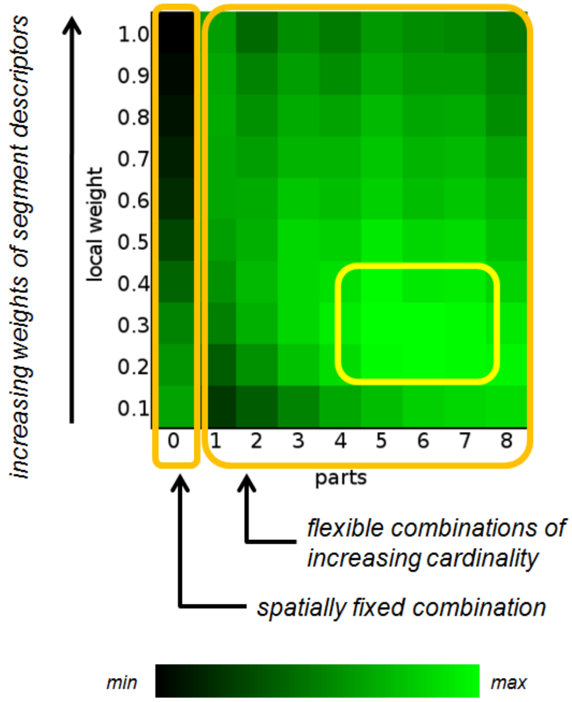


Figure 3: R-Precision result heatmap obtained for the DBD descriptor on the SHREC benchmark. The flexible combination scheme improves over the spatially fixed combination scheme, for a substantial range in the parameter space. This result can be seen by visually comparing the two schemes via the green cell shades in the orange frames (more green intensity indicates higher R-Precision). The parameter region of maximum R-Precision is shown by the yellow frame.

and *parts* of 6, 6, 7, respectively. This implies a relative increase in R-Precision of the flexible scheme over the spatially fixed scheme between 3% and 5%, for these parameter ranges.

We also observe similar rates of improvement for the other tested descriptors RSH, SIL, and DBD (cf. Tables 2, 3, 4). Figure 4 visualizes the results in an array of heatmap diagrams, where rows correspond to data sets, and columns correspond to descriptors. For almost all cases, a rather large and stable region in parameter space exists, for which the R-Precision results of the spatially fixed combination scheme are outperformed. This parameter space region is consistently located around low to medium values of *local_weight*, and medium to high values of *parts*.

4.3 Discussion of Results and Extensions

Based on our experimental evaluation, we argue that the flexible combination scheme can provide better R-Precision than the spatially fixed combination scheme. The improvement comes at the price of an additional parameter (*parts*). However, based on the experimental results, a rather large and stable region in parameter space exists for which improvements can be realized. A simple rule-of-thumb for this parameter, as derived from the analysis of results in Figure 4, can be used in practical implementations. Note that improvements of even a few percentage points in retrieval precision are important, as we are considering the problem of improving the effectiveness of the retrieval here. Unlike efficiency, effectiveness cannot be addressed by scaling-up the system hardware/software, but improvements to the core search algorithms need to be found. Our approach involves the Hungarian method to solve a bipartite graph matching problem. This algorithm has runtime complexity of $O(n^3)$ in the number of nodes to match. Given that we consider a small number of fragments (eight), this runtime complexity is deemed acceptable. We have yet not explored the implications of using a higher number of fragments. We note that for more fragments, possibly also more efficient matching schemes, e.g., relying on approximation, could be necessary. Detailed runtime consideration was beyond the scope of this study and is left for future work. We have tested our scheme on four different 3D descriptors. While to date, a wealth of other descriptors is available, the tested descriptors resemble a set of robust and standard methods. Further methods should be considered for generalizability. The improved combination approach can be recommended as a simple-to-implement scheme that is orthogonal to the definition or global 3D object descriptors.

We see several promising extension possibilities. It would be interesting to further analyze the reason for the improvements brought about by the flexible matching scheme. In particular, if we could identify cases for which model classes which fragments typically match, we might be able to incorporate heuristic rules to automatically guide the matching

Method (DBD)	SHREC	PSB	ESB
Global only	39.4%	33.2%	33.7%
Best spatially fixed	42.8%	34.7%	35.4%
Best spatially flexible	44.9%	35.8%	36.5%
Improvement wrt. global	14.0%	7.8%	8.3%
Improvement wrt. fixed	4.9%	3.2%	3.1%

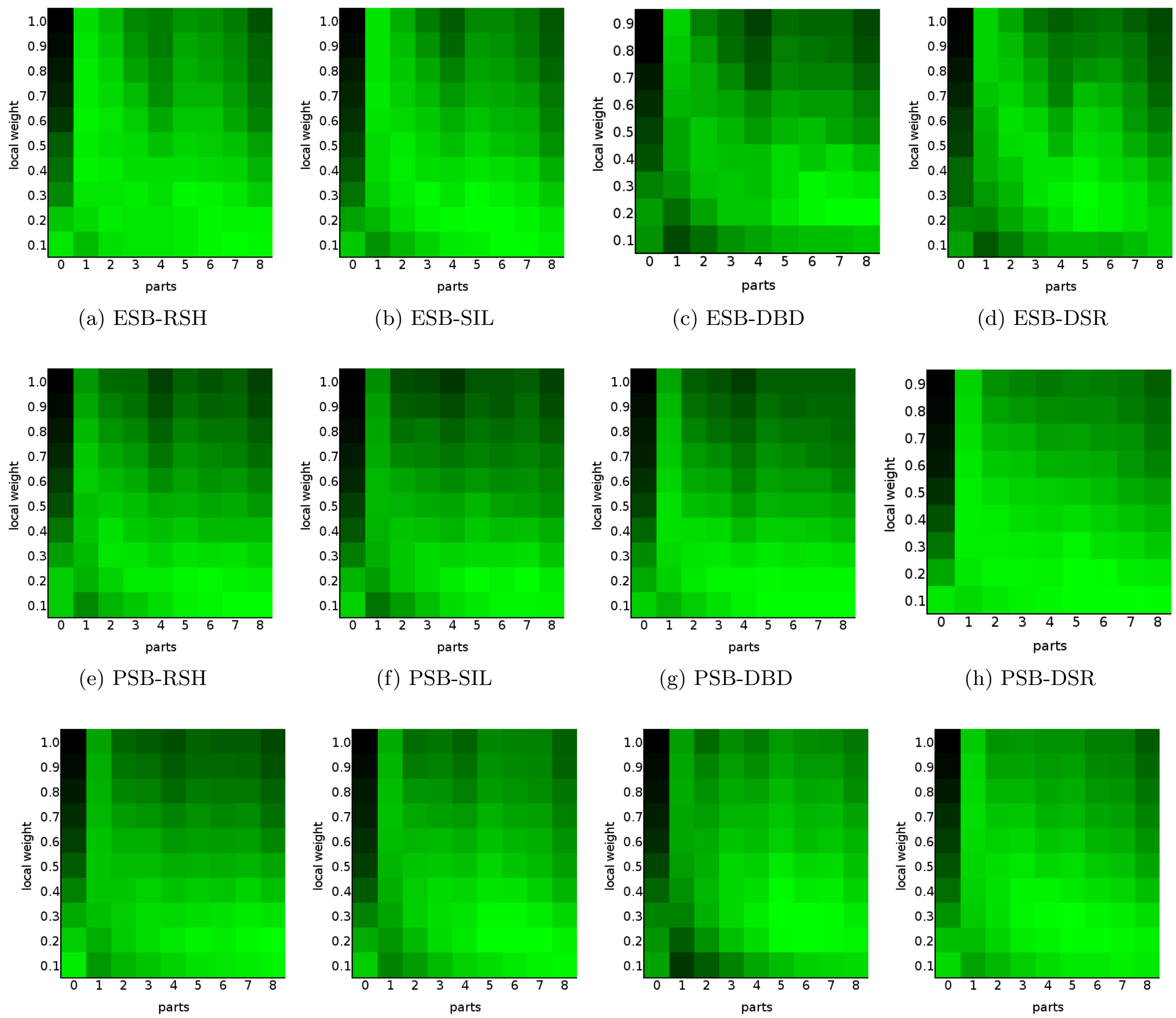
Table 1: Summary of the best R-Precision results obtained for the DBD descriptor and the three considered benchmarks. The new combination scheme provides robust effectiveness improvements over the spatially fixed scheme.

Method (DSR)	SHREC	PSB	ESB
Global only	43.8 %	29.4%	29.8%
Best spatially fixed	45.3 %	30.1%	31.4%
Best spatially flexible	45.9 %	30.7%	32.4%
Improvement wrt. global	4.8%	4.4%	8.4%
Improvement wrt. fixed	1.3%	2.0%	2.9%

Table 2: Results obtained for the DSR descriptor.

Method (SIL)	SHREC	PSB	ESB
Global only	36.1 %	23.9%	27.8%
Best spatially fixed	37.7 %	25.6%	29.0%
Best spatially flexible	38.4 %	26.1%	30.1%
Improvement wrt. global	6.3%	9.2%	8.3%
Improvement wrt. fixed	1.9%	2.0%	3.8%

Table 3: Results obtained for the SIL descriptor.



(a) ESB-RSH (b) ESB-SIL (c) ESB-DBD (d) ESB-DSR

(e) PSB-RSH (f) PSB-SIL (g) PSB-DBD (h) PSB-DSR

(i) SHREC-RSH (j) SHREC-SIL (k) SHREC-DBD (l) SHREC-DSR

Figure 4: R-Precision heatmap diagrams obtained for all benchmarks (rows) and descriptors (columns). For almost all cases, the flexible combination scheme improves over the fixed combination scheme, and for a substantial area in the parameter space.

Method (RSH)	SHREC	PSB	ESB
Global only	32.4 %	20.2%	26.6%
Best spatially fixed	34.3 %	21.4%	27.3%
Best spatially flexible	34.7 %	21.8%	28.0%
Improvement wrt. global	7.1%	7.9%	5.3%
Improvement wrt. fixed	1.2%	1.9%	2.6%

Table 4: Results obtained for the RSH descriptor.

process. Learning-based approaches could be useful to this end. Our scheme is currently based on a simple Octree-based partitioning scheme. Data-dependent partitioning schemes, e.g., based on local interest point analysis [12, 6], could further improve the method and should be experimentally compared to the proposed method. Also, the output of local interest point analysis could serve to find locally-adaptive weighing schemes (currently, *local_weight* is set uniformly for all fragment indices and query objects). Our distance function violates the triangle inequality, and thereby, is not a metric. Thus, many established indexing techniques for metric spaces are not applicable, but non-metric approaches [13] need to be considered. Finally, applicability of our approach to further multimedia object types with spatial reference such as images, would be interesting to consider in the future.

5. CONCLUSIONS

3D object retrieval using global descriptors remains a challenging research problem. We proposed a scheme to generically improve the retrieval precision of global 3D descriptors based on a new combination scheme. The scheme forms combinations of global and fragment descriptors, allowing for flexible fragment matching based on solving a bipartite graph problem. By exhaustive systematic experiments on several benchmark data sets and descriptors, we showed that the new matching scheme provides retrieval precision rates exceeding those of using only the global descriptors. And in particular, the new scheme improves over an existing scheme based on spatially fixed fragment combinations. The method is proposed as an inexpensive and practical approach to improve the retrieval performance of 3D object retrieval based on global 3D descriptors. A range of interesting future work possibilities exists as discussed.

Acknowledgments

This work was partially funded by Fondecyt (Chile) Project 1110111. It was also partially supported by the MKE (The Ministry of Knowledge Economy), Korea, under the "IT ConsilienceCreative Program" support program supervised by the NIPA (National IT Industry Promotion Agency) (NIPA-2010-C1515-1001-0001).

6. REFERENCES

[1] P. Besl and N. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:239–256, 1992.

[2] A. M. Bronstein, M. M. Bronstein, L. J. Guibas, and M. Ovsjanikov. Shape google: Geometric words and expressions for invariant shape retrieval. *ACM Trans. Graph.*, 30:1:1–1:20, February 2011.

[3] B. Bustos, D. Keim, D. Saupe, and T. Schreck. Content-based 3D object retrieval. *IEEE Computer Graphics and Applications, Special Issue on 3D Documents*, 27 (4):22–27, 2007.

[4] B. Bustos, T. Schreck, M. Walter, J. Barrios, M. Schaefer, and D. Keim. Improving 3D similarity search by enhancing and combining 3D descriptors. *Springer Multimedia Tools and Applications*, Online First, January 05, 2011.

[5] X. Chen, A. Golovinskiy, , and T. Funkhouser. A benchmark for 3D mesh segmentation. *ACM Transactions on Graphics*, 28(3), Aug. 2009.

[6] H. Dutagaci, C. P. Cheung, and A. Godil. Evaluation of 3d interest point detection techniques. In *Proc. EG Workshop on 3D Object Retrieval*, pages 57–64, 2011.

[7] H. Dutagaci, A. Godil, P. Daras, A. Axenopoulos, G. C. Litos, S. Manolopoulou, K. Goto, T. Yanagimachi, Y. Kurita, S. Kawamura, T. Furuya, and R. Ohbuchi. Shrec '11 track: Generic shape retrieval. In *Proc. EG Workshop on 3D Object Retrieval*, pages 65–69, 2011.

[8] A. Godil, H. Dutagaci, C. Akguel, A. Axenopoulos, B. Bustos, M. Chaouch, P. Daras, T. Furuya, S. Kreft, Z. Lian, T. Napoléon, A. Mademlis, R. Ohbuchi, P. L. Rosin, B. Sankur, T. Schreck, X. Sun, M. Tezuka, A. Verroust-Blondet, M. Walter, and Y. Yemez. SHREC'09 Track: Generic shape retrieval. In *Proc. EG Workshop on 3D Object Retrieval*, pages 61–68. Eurographics Association, 2009.

[9] S. Jayanti, Y. Kalyanaraman, N. Iyer, and K. Ramani. Developing an engineering shape benchmark for CAD models. *Computer-Aided Design*, 38(9):939–953, 2006.

[10] H. W. Kuhn. The Hungarian method for the assignment problem. *Naval Research Logistic Quarterly*, 2:83–97, 1955.

[11] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser. The princeton shape benchmark. In *Proc. Int. Conference on Shape Modeling and Applications*, pages 167–178. IEEE CS Press, 2004.

[12] I. Sipiran and B. Bustos. Harris 3D: a robust extension of the harris operator for interest point detection on 3d meshes. *The Visual Computer*, 2011. Online first.

[13] T. Skopal and B. Bustos. On nonmetric similarity search problems in complex domains. *ACM Computing Surveys*. To appear.

[14] J. W. H. Tangelder and R. C. Veltkamp. A survey of content based 3D shape retrieval methods. *Multimedia Tools and Applications*, 39(3):441–471, 2008.

[15] D. Vranic. *3D Model Retrieval*. PhD thesis, University of Leipzig, Germany, 2004.

[16] D. Vranic. DESIRE: a composite 3D-shape descriptor. In *Proc. IEEE Int. Conference on Multimedia and Expo*, pages 962–965., Los Alamitos, CA, USA, 2005. IEEE.

[17] S. Yoon, M. Scherer, T. Schreck, and A. Kuijper. Sketch-based 3D model retrieval using diffusion tensor fields of suggestive contours. In *Proc. ACM Multimedia*, pages 193–200. ACM, 2010.