SHREC'08 Entry: Semi-Supervised Learning for Semantic 3D Model Retrieval

Akihiro Yamamoto, Masaki Tezuka, Toshiya Shimizu, Ryutarou Ohbuchi

University of Yamanashi[†]

ABSTRACT

A shape feature by itself is not sufficient for effective 3D model retrieval. Long-lasting semantics shared by a community as well as a short-lived intention of a user determines the similarity of 3D models. In this paper, we describe a method of shape-based 3D model retrieval that employs off-line, semi-supervised learning of multiple classes in the database to capture long-lasting, shared semantic knowledge. The method performs two learning based dimension reductions, first one to accommodate distribution of features in the feature space and the second one to accommodate the semantic knowledge embodied in a set of user-defined semantic labels. We evaluate the method by using the SHREC'08 3D Generic and CAD Models Track.

KEYWORDS: Content-based retrieval, manifold learning, multi-scale feature.

INDEX TERMS: H.3.3 [Information Search and Retrieval]: Information filtering. I.3.5 [Computational Geometry and Object Modeling]: Surface based 3D shape models. I.4.8 [Scene Analysis]: Object recognition.

1 INTRODUCTION

Shape-based retrieval of 3D models could benefit significantly from supervised learning of commonly shared semantic classes or an intention of a user. Overall shape of bananas and dolphins are similar. In such a case, a supervised learning scheme would help a lot in distinguishing semantic classes.

An intention is often captured by using *Relevance feedback* (*RF*), which learns, supervised, an intention of a user through interactive user feedback. The RF approach captures an intention, a semantics that lasts a short period and shared by a few. A drawback of the RF is that every retrieval session must start anew with a low retrieval performance.

To exploit long-lived, commonly shared knowledge, or *semantics*, one may employ an *off-line supervised learning* of *multiple semantic classes*. Supervised learning of multiple semantic classes can be difficult, due to small number of labeled examples per class and mutually conflicting constraints from multiple classes.

We have previously proposed a semi-supervised dimension reduction (SSDR) approach [3] to learn multiple semantic classes off-line, and successfully improved retrieval performance of the *Princeton Shape Benchmark* (*PSB*) [5], a collection of "generic" and diverse 3D models. In this paper, we apply the same approach to two *SHREC'08 Tracks*.

[†]4-3-11 Takeda, Kofu-shi, 400-8511, Japan.
yama.akihiroAT jp.fujitsu.com, g06mk009AT yamanashi.ac.jp
t04kf026AT yamanashi.ac.jp, ohbuchiAT yamanashi.ac.jp

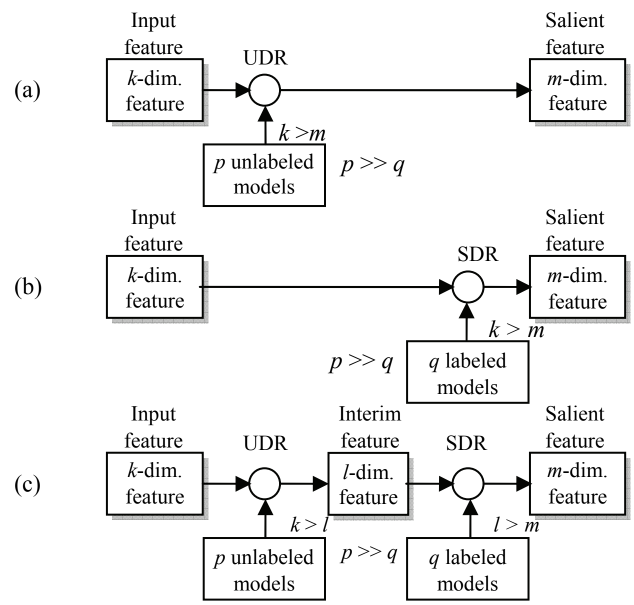

Figure1. *Unsupervised* (*UDR*) (a), *Supervised* (*SDR*) (b), and the *Semi-Supervised* (*SSDR*) (c) dimension reduction methods.

2 METHOD

The SSDR method [3] first generates a 6-level set of morphological *Multi-Resolution* (*MR*) models from each model in order to compare features at finer as well as coarser (i.e., nearly convex hull) scales. Following steps, the feature extraction, dimension reduction, and distance computation, are all performed at each resolution level before the distances from all *MR*-levels are fused into an overall distance (Figure 2).

To extract a feature, a model is converted from a surface model to a point set model via *quasi-Monte-Carlo* sampling of the surfaces. Then the *Surflet Pair Relations Histogram* (*SPRH*) algorithm [7] is applied to compute a k=54=625D histogram. The method then performs Unsupervised Dimension Reduction (UDR) to map the input k-D feature onto a subspace of l-D adapted to the distribution of the features. This step makes the next step, the supervised learning, easier. The interim feature is then processed by *Supervised Dimension Reduction* (*SDR)* algorithm trained by using a labeled training set to produce a salient feature. The dimension m of the salient feature after the *SDR* step is typically has a much smaller than the dimension k of the original feature. Distances computed at six MR levels are combined into one by simply summing them.

We compare the retrieval performance among four methods; (1) (multiresolution version of the) original SPRH feature, (2) UDR processed feature (Figure 1a), (3) SDR processed feature (Figure 1b), and (4) SSDR processed feature (Figure 1c). To compute distance, the original SPRH used the *Kullback-Leibler* divergence, while the UDR, SDR, and SSDR processed features used *Cosine* measure. We used the *Locally Linear Embedding* (LLE) [4] for the UDR, and *Supervised Locality Preserving Projections* (*SLPP*) [1] for the SDR.

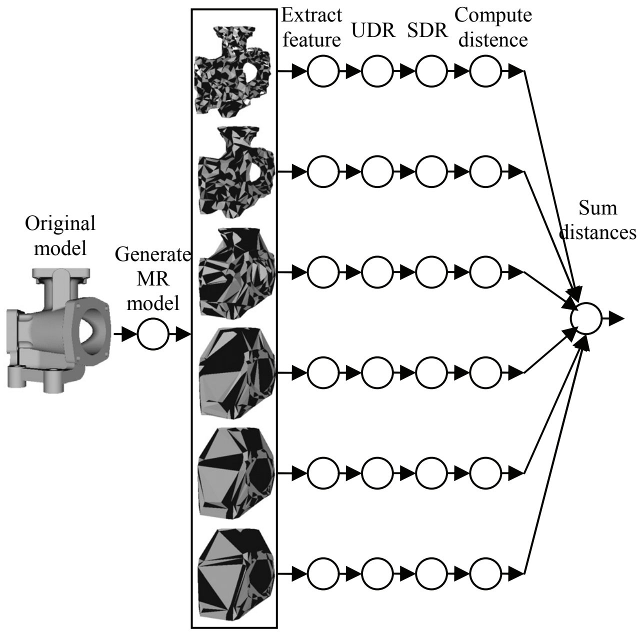

Figure 2. A set of morphological multiresolution models are constructed from the original model for shape comparison. Feature extraction, dimension reduction, and distance computation are performed at each resolution level. Distance from multiple resolution levels are integrated into an overall distance by summing.

3 GENERIC MODELS TRACK

We used our method in the *SHREC'08 Generic Models Track* (*GMT*) for which we briefly summarize the training data and experimental results.

3.1 TRAINING DATA

We trained the UDR algorithm by using the union of the PSB and the National Taiwan University shape benchmark down-sampled to 5,000 models. In the SHREC 2008 GMT, the query set identical to that of SHREC 2006 query set (30 models) is called *Query set 1* (Q1), and the new query set (30 models) is called *Query set 2* (Q2). The SDR algorithm is trained by using the classes of the Q1 only, a set of classes identical to the SHREC 2006 ground truth set classes.

Various parameters for learning algorithms are set to maximize the retrieval performance based on a preliminary set of experiments.

3.2 EXPERIMENTS AND RESULTS

For the Generic Models Track, Table 1 and Figure 2 show excerpts of performance evaluation results. In the case of Q1, the FT=34.93% of the original feature went up to FT=45.4% after the UDR. The SDR applied directly to the original feature produced the FT=38.49%, a figure marginally better than the original, but worse than that of the UDR. The proposed SSDR method yielded the best retrieval performance FT=58.33% for Q1, 23% improvement over the original, MR-SPRH feature.

The performance figures of the method for the Q2 are markedly worse than those for the Q1. In the SHREC'08 GMT, the SSDR method ranked 1st place for the Q1. However, the rank dropped to 3rd place for the Q2. This drop in performance is explained by the fact that the SSDR method is trained by using Q1, but not Q2. Still, the performance of the SSDR method for Q2 is significantly better than that of the UDR method, suggesting that the learning done in the SDR stage for the Q1 classes do generalize to the classes in Q2.

The method is computationally efficient during retrieval, as all the costly learning is done during the preprocessing stage, off-line. Furthermore, the cost of comparison, which is dominant for a large database, is reduced due to dimension reduction. For the six level MR, the dimension of the original, 625×6=3,750, is reduced to 40×6=240 with according reduction in computational cost.

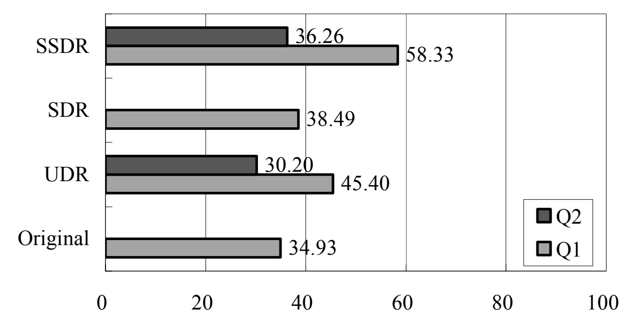

Figure 2. Comparison of the performance across methods and query sets using Mean First Tier (Highly Relevant).

Table 1. Retrieval performances of the Semi-Supervised Dimension Reduction (SSDR)-based method measured using the SHREC 2008 Generic Models Track.

Query set	Method	UDR Dim. *l*	SDR Dim. *m*	Distance	Retrieval performance						
					MAP-HR	MAP-R	AFT-HR	AFT-R	DAR	MNCG @25	MNDCG @25
1	Original	-	-	KLD	0.3761	0.3552	34.93%	32.84%	0.4631	0.4519	0.5101
	UDR	350	-	Cosine	0.4687	0.4278	45.40%	42.18%	0.5413	0.5567	0.6049
	SDR	-	30	Cosine	0.4040	0.3893	38.49%	37.33%	0.4667	0.4886	0.5159
	SSDR	400	50	Cosine	0.6309	0.5846	58.33%	53.63%	0.6514	0.6578	0.6879
	Makadia (run2)	-	-	-	0.4869	0.4364	44.77%	40.55%	0.5499	0.5498	0.5906
2	UDR	350	-	Cosine	0.2818	0.2363	30.20%	26.14%	0.3616	0.3768	0.4120
	SSDR	400	50	Cosine	0.4081	0.3603	36.26%	34.05%	0.4480	0.4612	0.4681

MAP: Mean Average Precision (HR-Highly relevant, R-Relevant)
AFT: Mean First Tier (HR-Highly relevant, R-Relevant)
DAR: Mean Dynamic Average Recall
NCG@25: Mean Normalized Cumulated Gain @25
NDCG@25: Mean Normalized Discounted Cumulated Gain @25

4 CAD Models Track

4.1 Training Data

For the *SHREC'08 CAD Models Track*, we trained the UDR algorithm by using "generic" 3D models; we used 4,000 models quasi-randomly sampled from the union of the PSB and the National Taiwan University shape benchmark. We must use generic models since the ESB with its 866 models is not large enough to train the UDR. The SDR algorithm is trained by using the ESB models and classes.

4.2 Experiments and Results

The results for this track are summarized in Figure 3. Here, we use *Mean First Tier* (FT), among others, as the numerical performance measure. The UDR improved the FT=36.38% of the original feature to 42.75%. This is the method we used to enter the SHREC 2007 CMT [6], but with a slightly different parameter set. The SDR applied directly to the original feature resulted in FT=57.72%, a significant 20% gain over the original. In contrast, in case of the PSB, the SDR-only method did not give significant performance gain. The ESB classes may be easier to learn than the PSB. Finally, the SSDR yielded the best score, FT=78.16%, adding 40% to the original.

Figure 4 shows the change of performance in FT as we vary the number of the salient feature dimension m after the SDR. The graph plots curves for multiple interim feature dimensions l={100, 200, 300, 400, 500} after the UDR. The SSDR overtook SDR-only feature (625D) if the interim feature dimension l>300.

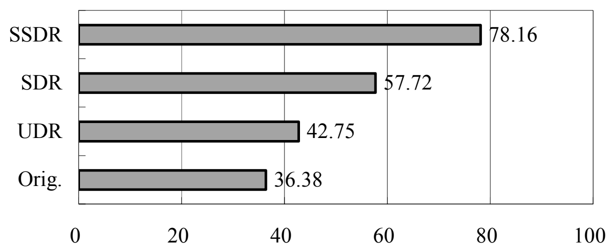

Figure 3. Comparison of the performance using the Mean First Tier.

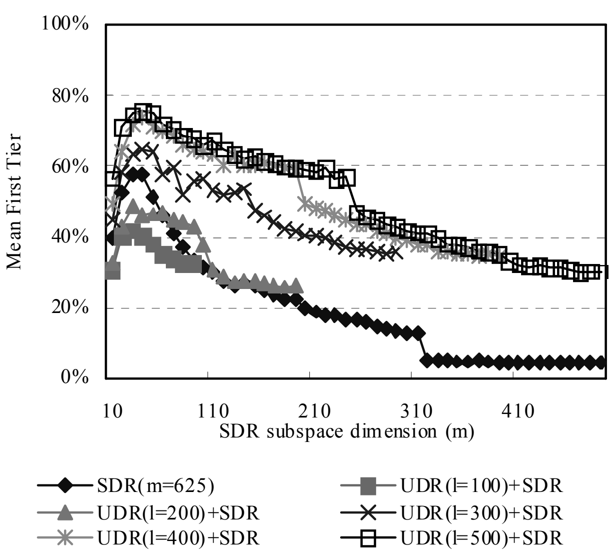

Figure 4. SDR subspace dimension m versus the retrieval performance in Mean First Tier [%] for several UDR subspace dimensions l.

Observation of per-query results seems to suggest the following reasons for the good result; (1) the MR approach that exploits features at multiple scales, (2) the SSDR that learned to use the SPRH feature effectively, and (3) the SPRH feature that detects mutual orientation of surfaces about 3 axes parameterized by distance. In the CMT, the SSDR obtained FT=100% for the majority of the queries.

Some of the CMT queries seem quite difficult even for a non-mechanical engineer human being to do. For example, Query 43, Query 45 (the method has FT=100% on these two) and Query 3 (The method has FT=0% on this) were both difficult for my students.

The method is computationally efficient. Costly learning is done as a preprocessing. Once the UDR and SDR are trained, a dimension reduction is simply a matrix-vector multiplication of a l-dimensional (for the UDR) or m-dimensional (for the SDR) vector. The dimension reduction makes the search through the database much faster. Per resolution level, the dimension of the SSDR processed feature m=40, which is 1/15 of the original feature having k=625. If all six MR levels are combined, the dimension is reduced from 625×6=3,750 to 40×6=240.

5 Conclusion

We have used a *Semi-Supervised Dimension Reduction* (SSDR) method that exploits a large set of unlabeled models as well as a small set of labeled models for an effective dimension reduction [3]. We combined the method with the SPRH feature and the multiresolution (MR) feature extraction approach.

Within the Generic Models Track, the Mean First Tier of 35% for the original MR SPRH increased to 58% after the SSDR. As expected, the method performed very well for the set classes it is trained for, the Q1, but not very well for the Q2.

Within the CAD Models Track, the Mean First Tier of 36% for the original MR SPRH increased to 78% after the SSDR to win the SHREC'08 CAD Model Track with a large (>30% in First Tier) margin. It appears that a task like the CAD Model Track benefit a lot from supervised learning.

References

[1] X. He, P. Niyogi, Locality Preserving Projections, *Advances in Neural Information Processing Systems*, 16, Vancouver, Canada, (2003)

[2] R. Ohbuchi, T. Takei, Shape-Similarity Comparison of 3D Shapes Using Alpha Shapes, *Proc. Pacific Graphics* 2003, pp. 293-302, (2003).

[3] R. Ohbuchi, A. Yamamoto, J. Kobayashi, Learning semantic categories for 3D Model Retrieval, *Proc. ACM MIR* 2007, pp. 31-40, (2007).

[4] S.T. Roweis, L.K. Saul, Nonlinear Dimensionality Reduction by Locally Linear Embedding, *Science*, **290**(5500), pp. 2323-2326, (2000)

[5] P. Shilane, P. Min, M. Kazhdan, T. Funkhouser, The Princeton Shape Benchmark, *Proc. SMI '04*, pp. 167-178, (2004).

[6] R.C. Veltkamp, F.B. ter Haar, SHREC 2007 3D Shape Retrieval Contest, Dept of Info and Comp. Sci., Utrecht Univ., *Tech. Report* UU-CS-2007-015, (2007)

[7] E. Wahl, U. Hillenbrand, G. Hirzinger, Surflet-Pair-Relation Histograms: A Statistical 3D-Shape Representation for Rapid Classification, *Proc. 3DIM 2003*, pp. 474-481, (2003).