Fusing Multiple Features for Shape-based 3D Model Retrieval

Takahiko Furuya
g13dm003@yamanashi.ac.jp

Ryutarou Ohbuchi
ohbuchi@yamanashi.ac.jp

Graduate School of Medicine and Engineering,
University of Yamanashi,
Yamanashi, Japan

Abstract

Fusing multiple features is a promising approach for accurate shape-based 3D Model Retrieval (3DMR). Most of the previous algorithms either simply concatenate feature vectors or sum similarities derived from features. However, ranking results due to these methods may not be optimal as they don't exploit distributions, i.e., manifold structures, of multiple features. This paper proposes a novel 3DMR algorithm that effectively and efficiently fuses multiple features. The proposed algorithm employs a Multi-Feature Anchor Manifold (MFAM) that approximates multiple manifolds of heterogeneous features with small number of "anchor" features. Given a query, ranks of 3D models are computed efficiently by diffusing relevance on the MFAM. Distance metrics of heterogeneous features are fused during the diffusion for better ranking. Experiments show that our proposed algorithm is more accurate and much faster than 3DMR algorithms we have compared against.

Proposed algorithm

For accurate and efficient 3DMR, we propose *3D model retrieval by Visual Feature Fusion (3DVFF)* algorithm that fuses multiple visual features of 3D models via an unsupervised distance metric fusion algorithm called *Multi-Feature Anchor Manifold Ranking (MFAMR)*. Figure 1 shows an overview of the proposed algorithm. The 3DVFF algorithm first extracts two visual features SV-DSIFT and LL-MO1SIFT from each 3D model in a database. The SV-DSIFT aggregates local visual features extracted from multiple viewpoints by Super Vector (SV) coding [1], and shows high accuracy for models having global deformation and/or articulation. The LL-MO1SIFT aggregates per-view global image features by using Locality-constrained Linear (LL) coding [2], and shows high accuracy for rigid models. For each feature, to reduce computational cost, a manifold of all the features is approximated by a manifold of anchors. Then, the two anchor manifold graphs are fused into a MFAM graph [3]. Ranking of the 3D models in the database for a given query is efficiently computed by relevance diffusion from the query to the 3D models over the MFAM. The two heterogeneous manifolds, one for the SV-DSIFT and the other for the LL-MO1SIFT, are fused during the relevance diffusion over the MFAM to yield a fused distance metric.
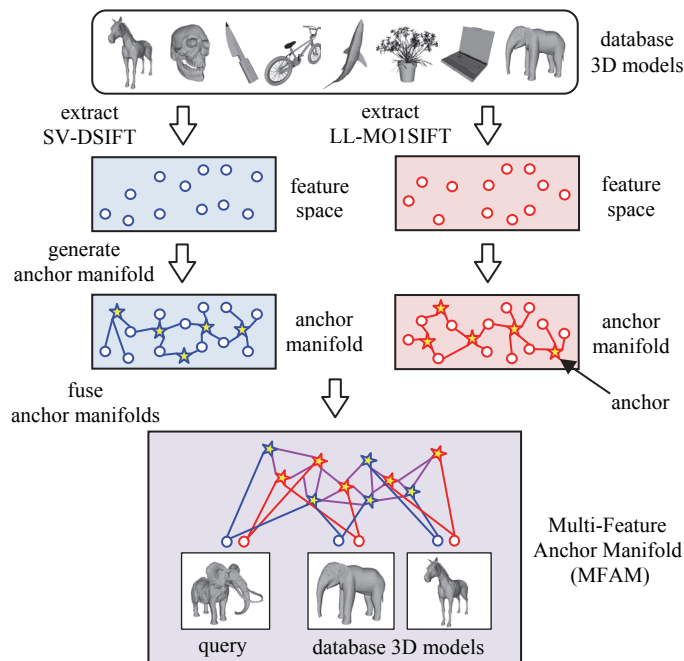

Figure 1: Overview of the proposed algorithm.

Experiments and results

To evaluate accuracy and efficiency of the proposed 3DVFF algorithm, we use two benchmarks; the *Princeton Shape Benchmark* (*PSB*) [4] and the SH14LC [5] (Results for other benchmarks are presented in a full paper). Figure 2 shows examples of 3D models for the benchmarks. For the PSB, we use 400 anchors for SV-DSIFT and 500 anchors for LL-MO1SIFT to approximate structure of multi-feature manifold. For the SH14LC, we use 2,000 anchors for SV-DSIFT and 2,500 anchors for LL-MO1SIFT. We use *Mean Average Precision* (*MAP*) [%] for quantitative evaluation of retrieval accuracy.

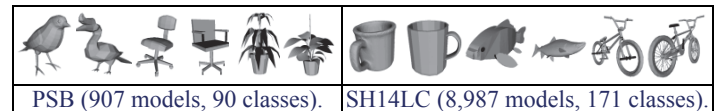| PSB (907 models, 90 classes). | SH14LC (8,987 models, 171 classes). |

Figure 2: Examples of 3D models for the benchmarks.

Efficiency: Table 1 shows computation time per query. Feature extraction is accelerated by using a GPU and multi-core CPUs. MFAMR is computed on a single thread. Our 3DVFF using MFAMR is much faster than previous feature fusion method MR-early, which sums multiple feature similarities to generate a manifold graph and performs naive Manifold Ranking [6] at every query. The 3DVFF takes less than 3 seconds per query for the SH14LC having 8,987 3D models.

algorithms	Feature extraction	Ranking	Total
MR-early	2.575	34.779	37.354
3DVFF (proposed)	2.575	0.031	**2.606**

Table 1: Computation time per query for the SH14LC benchmark.

Accuracy: Table 2 compares retrieval accuracy of the proposed 3DVFF algorithm with other state-of-the-art 3D model retrieval algorithms. For both the PSB and the SH14LC, our proposed 3DVFF algorithm achieved the highest retrieval accuracy among the algorithms listed in Table 2. For the SH14LC benchmark, MAP score of our 3DVFF is 3% higher than LCDR-DBSVC, the best performing algorithm among the SH14LC track entries [5].

algorithms	PSB	SH14LC
MR-D1SIFT [6]	62.9	46.4
LCDR-DBSVC [5]		54.1
SV-DSIFT	63.4	46.4
LL-MO1SIFT	55.3	39.9
3DVFF (proposed)	**72.6**	**57.2**

Table 2: Comparison of retrieval accuracy (MAP [%]).

[1] X. Zhou, K. Yu, T. Zhang, and T.S. Huang. Image Classification using Super-Vector Coding of Local Image Descriptors, *Proc. ECCV 2010*:141–154, 2010.

[2] J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong. Locality-constrained Linear Coding for Image Classification, *Proc. CVPR 2010*:3360–3367, 2010.

[3] S. Kim and S. Choi. Multi-view anchor graph hashing, *Proc. ICASSP 2013*:3123–3127, 2013.

[4] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser. The Princeton Shape Benchmark, *Proc. SMI 2004*:167–178, 2004.

[5] B. Li, Y. Lu, C. Li, A. Godil, T. Schreck, M. Aono, Q. Chen, N. K. Chowdhury, B. Fang, T. Furuya, H. Johan, R. Kosaka, H. Koyanagi, R. Ohbuchi, and A. Tatsuma. Large Scale Comprehensive 3D Shape Retrieval, *Proc. EG 3DOR 2014*:131–140, 2014.

[6] D. Zhou, J. Weston, A. Gretton, O. Bousquet, and B. Schölkopf. Ranking on Data Manifolds, *Proc. NIPS 2004*, 2004.

[7] R. Ohbuchi and T. Furuya. Distance metric learning and feature combination for shape-based 3D model retrieval, *Proc. 3DOR 2010*:63–68, 2010.