Deep Aggregation of Local 3D Geometric Features for 3D Model Retrieval

Takahiko Furuya
takahikof AT yamanashi.ac.jp

Ryutarou Ohbuchi
ohbuchi AT yamanashi.ac.jp

Integrated Graduate School of Medicine,
Engineering, and Agricultural Sciences,
University of Yamanashi,
Kofu, Japan

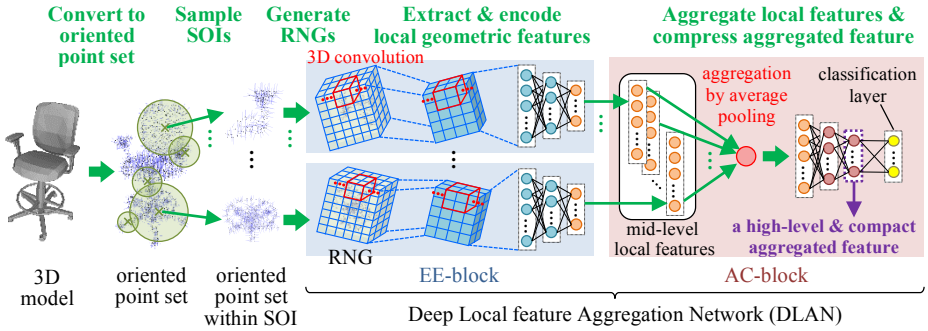

Figure 1: DLAN architecture for extracting a rotation-invariant and salient feature per 3D model.

Aggregation of Local Features (*ALF*) is a well-studied approach for image as well as 3D model retrieval (3DMR). A carefully designed local 3D geometric feature is able to describe detailed local geometry of 3D model, often with invariance to geometric transformations including 3D rotation of local 3D regions. For efficient 3DMR, these local features are aggregated into a feature per 3D model by using Bag-of-Features, Fisher Vector coding, etc.

Recent trend is to use end-to-end 3D Deep Convolutional Neural Network (3D-DCNN) (e.g., [1]) for 3DMR. 3D-DCNNs have often shown accuracies better than methods based on ALF. However, current 3D-DCNN based methods have weaknesses; they lack invariance against 3D rotation and/or they often miss geometrical details as they coarsely quantize shapes into voxels in applying 3D-DCNN.

Our goal is to extract a 3D model feature that is invariant against 3D rotation and more accurate than the existing ALF and 3D-DCNN based approaches. To this end, we combine ALF with 3D-DCNN.

We propose a novel deep neural network for 3DMR called *Deep Local feature Aggregation Network* (*DLAN*) that performs extraction of rotation-invariant 3D local features and their aggregation by using a single deep architecture. A DLAN (Figure 1) first describes local 3D regions of a 3D model by using "mid-level" local features invariant to 3D rotation. The set of local features is aggregated into a rotation-invariant and compact feature vector per 3D model.

Experimental evaluation using three benchmark datasets shows effectiveness of the DLAN. Here, we present results on the ModelNet40 dataset [1]. The proposed DLAN significantly outperforms the state-of-the-arts including 3D-DCNN based [1] and 2D-DCNN based [2][3] 3DMR algorithms.

algorithms	MAP [%]
3D ShapeNets [1]	49.2
MVCNN [2]	79.5
GIFT [3]	81.9
DLAN (proposed)	**85.0**

Table 1: Comparison of retrieval accuracy.

[1] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang and J. Xiao. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling. *Proc. CVPR 2015*, 1912–1920, 2015.

[2] H. Su, S. Maji, E. Kalogerakis, and E. Learned-Miller. Multi-View Convolutional Neural Networks for 3D Shape Recognition. *Proc. ICCV 2015*, 945–953, 2015.

[3] S. Bai, X. Bai, Z. Zhou, Z. Zhang, and L. J. Latecki. GIFT: A Real-time and Scalable 3D Shape Search Engine, *Proc. CVPR 2016*, 5023–5032, 2016.