The Method of Sketch-Based 3D Shape Retrieval

Hongzhi Liu

April 23, 2018

1 The Background of Research Method

Recently extensive research efforts have been dedicated to sketch-based 3D shape retrieval. Besides, deep features have been proposed for the retrieval due to the success of deep neural networks in different applications. By reading Professor Xie's thesis, I can learn Wasserstein barycenters of multiple views of 3D shapes in the feature space for sketch-based shape retrieval.

The team mainly employ two deep CNNs to extract the CNN features of sketches and 2D projections from a set of rendered views. And then the barycenters of CNN features of 2D projections can then be computed to characterize 3D shapes. Finally, the proposed approach for sketch-based 3D shape retrieval can be demonstrated effective through experimental results.

2 Proposed Approach

In this section, I will introduce main research theory and method in the thesis. One is Wasserstein barycenters and the other is cross-domain matching.

2.1 Wasserstein barycentric representation method

The Wasserstein distance has been widely used in computer vision. And barycenters p_b of a set of probability distributions p_i can be defined as [1]:

$$p_i = \arg\min_{p_b} \sum \lambda \times D(p_i, p_b) \tag{1}$$

where $D(p_b, p_i)$ is the Wasserstein distance between p_b and p_i and λ is the weight.

Due to the property of the Wasserstein barycenters that can capture the structure of the high-dimensional data well, Professor Xie propose to use the Wasserstein barycenters of projections from multiple views in the feature space to characterize 3D shapes.

2.2 3D shape retrieval methods

In the thesis, I can learn the Wasserstein barycenters of the deep CNN features of 3D shapes for sketch-based 3D shape retrieval. As is mentioned in paper, author employ two AlexNets to extract deep CNN features of 2D projections and sketches, respectively. Once the CNN features from 3D shapes and sketches are obtained, we use fully connected layers to construct a metric network. Fig. 1 illustrates the cross-domain matching framework with two networks, one for sketch network and the other for view network.
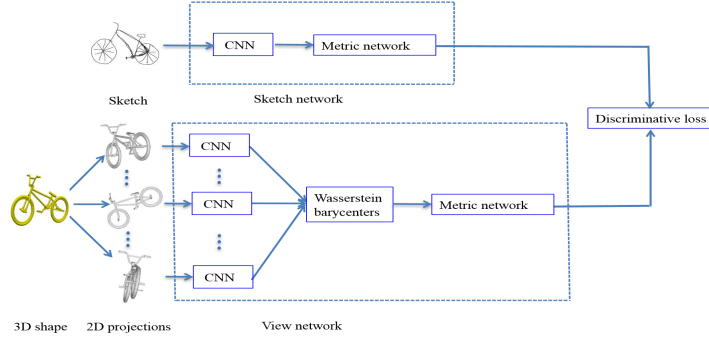

Figure 1: The cross-domain matching framework for sketch-based 3D shape retrieval. By rendering 3D shapes at multiple views, Professor Xie extract deep CNN features of 2D projections. The Wasserstein barycenters of the deep CNN features are computed to represent 3D shapes. With the metric network of fully connected layers, they then formulate a discriminative loss to learn sketch and shape features for cross-domain.

References

[1] Marco Cuturi and Gabriel Peyr. A smoothed dual approach for variational wasserstein problems. *Mathematics*, 90(1):499–505, 2015.