

3D Deep Shape Descriptor

Yi Fang¹, Jin Xie¹, Guoxian Dai¹, Meng Wang¹, Fan Zhu¹, Tiantian Xu², Edward Wong²,

¹Department of Electrical and Computer Engineering, New York University Abu Dhabi

²Polytechnic School of Engineering, New York University

Shape descriptor refers to an informative description that provides a 3D object with an identification as a member of some category. The development of an effective and efficient 3D shape descriptor poses several technical challenges, including, in particular, the high data complexity of 3D models and their representations, the structural variations, noise, and incompleteness present in 3D models [1, 2]. To address these challenges, we develop a deep shape descriptor (DeepSD), which includes 1) the heat shape descriptor (HeatSD) that uses the point-based heat kernel signature (HKS) and 2) the combination of the eigen-shape descriptor (ESD) and the fisher-shape descriptor (FSD). Our deep shape descriptor has high discriminative power that tends to maximize the inter-class margin while minimizing the intra-class variance.

DeepSD is mainly constitute of four components, where the illustration on how these four components are mapped onto a deep neural network is illustrated in Figure 1. The first component is a 3D shape database where a large volume of shapes are stored. The second component is shape feature extraction where two features: heat kernel signature (HKS) and heat shape descriptor (HeatSD), are extracted. The third component is a deep neural network for learning deep shape descriptor. A multi-layer deep neural network is used in our method. A collection of HeatSDs are used in the training of principal component analysis (PCA) and linear discriminant analysis (LDA) to generate the Eigen-shape descriptor (ESD) and Fisher-shape descriptor (FSD) respectively. The fourth component is the target value of DNN where pre-computed ESD and FSD are used as target values in the training the DNN.

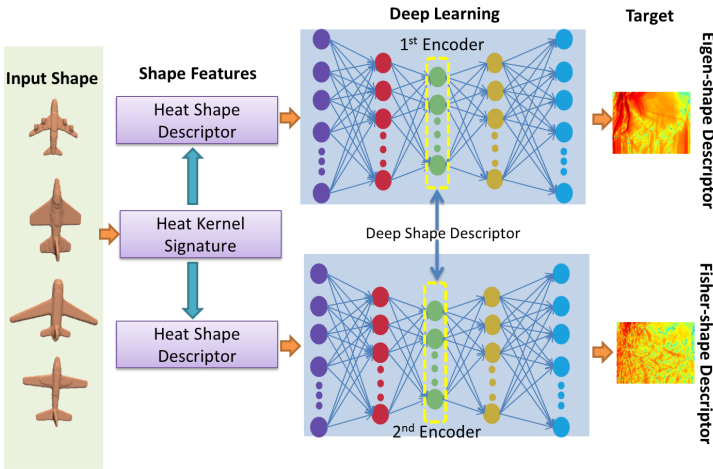


Figure 1: Pipeline of learning deep shape descriptor. Given input shapes, three steps are included along with the pipeline: 1) Heat kernel signatures are extracted for each shape in the database. Heat shape descriptor are computed based on HKS. 2) Heat shape descriptors are fed into two deep neural networks with target values using ESD and FSD, respectively. 3) The deep shape descriptor is formed by concatenating nodes in hidden layer (circled by yellow dash lines).

Heat kernel signature: The heat kernel $H_t(i, j)$ aggregates heat flow through all possible paths between two vertices on the meshed surface, and a heat kernel signature can be defined as:

$$HKS(p) = (H_{t_1}(p, p), H_{t_2}(p, p), \dots, H_{t_n}(p, p)) \quad (1)$$

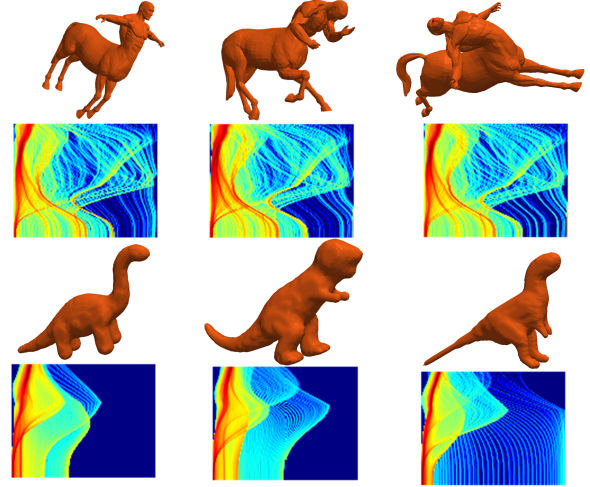


Figure 2: Illustration of heat shape descriptor. (A) illustrates the HeatSD for three centaur models undergone isometric transformation. (B) illustrates the HeatSD for three dinosaur models with moderate structural variations.

where p denotes a point on the surface, $HKS(p)$ denotes the heat kernel signature at point p , $H_t(p, p)$ denotes the heat kernel value at point p , t_n denotes the diffusion time of the n th sample point.

Heat shape descriptor: HeatSD is developed using probability distribution of HKS values at all vertices and at all scales. At each scale, HeatSD is defined based on the probability distribution of HKS at that scale. HeatSD is a multi-scale shape descriptor, thereby it can provide a complete and local-to-global description of 3D shape. Figure 2 displays 3D objects and their corresponding HeatSDs.

Deep shape descriptor: We use the architecture of a many-to-one encoder neural network to develop our encoder for deep shape descriptor. By enforcing the target value to be unique for input HeatSDs from the same group but with structural variations, the deep shape descriptor represented by the neurons in the hidden layer is invariant to within-group structural variations but will discriminate against other groups. We set the target value of the neural network as pre-computed Eigen-shape descriptor and Fisher-shape descriptor for each group. We formulate the objective function of the proposed sparse many-to-one encoder by the square-loss function with sparse constraint on the weights as:

$$\operatorname{argmin}_{W, b} \frac{1}{2} \sum_{i, j} \left\| Y_i - h(x_i^j, W, b) \right\|_2^2 + \frac{\lambda}{2} \|W\|_F^2, \quad (2)$$

where L is the number of layers in the deep neural network, W is the weight matrix of the multiple-hidden-layer neural network, b is the bias matrix of the neural network, x_i^j represents the j -th training sample from the i -th group, $h(x_i^j, W, b)$ in general is a non-linear mapping from the input x_i^j to the output. The parameter λ is the weight of the regularizer, and Y_i is the target value for the i -th group.

[1] Natraj Iyer, Subramaniam Jayanti, Kuiyang Lou, Yagnanarayanan Kalyanaraman, and Karthik Ramani. Three-dimensional shape searching: state-of-the-art review and future trends. *Computer-Aided Design*, 37(5):509–530, 2005.

[2] Johan WH Tangelder and Remco C Veltkamp. A survey of content based 3d shape retrieval methods. *Multimedia tools and applications*, 39(3):441–471, 2008.