# Meng Wang

## Personal Information

| | |
|---|---|
| **Status:** | **PH.D. STUDENT** |
| **Program:** | **Computer Science and Engineering** |
| **School:** | **Tandon School of Engineering, New York University** |
| **Period:** | **From 2014-09 to 2020-05** |

## Biography

**I was a Ph.D. student at New York University and advised by Professor YiFang. During my Ph.D. period, I was a research assistant in NYU Multimedia and Visual Computing (MMVC) Lab. I am broadly interested in 3D Computer Vision and Deep Learning. Now I am about to join Xmotors.ai (CA,USA) as a research scientist.**

**Research Project:** Hand-crafted Feature Extraction for Shape Correspondence

## Description

Shape correspondence provides important information for many shape analysis such as shape alignment and registration, texture mapping, information transformation among shapes, shape morphing, statistical shape modeling and so on. The accuracy of shape correspondence algorithm is ultimately determined by the quality and characteristics of the shape signature that captures the geometric essence of points on 3D surface. Based on the diffusion geometry, we proposed a local diffusion map signature (LDMS) based on a novel strategy of projecting heat diffusion distribution on 3D surface into 2D image along the direction determined by gradients of heat diffusion field. we experimentally verify the performance of LDMS. First, we conduct several shape matching test under different situations. Then we conduct some comparisons over other shape matching methods. The proposed method is tested on SCAPE, TOSCA, and Watertight database.

## Method

The proposed LDMS mainly includes three steps. The first step is mapping a 3D surface with heat diffusion distribution. A variety of heat diffusion distribution, such as HKS, SIHKS, and Heat Kernel Matrix, can be chosen for this purpose. HKS is chosen in our development. The second step is generating LDMS for repressive points sampled from 3D shape. The third step is producing a multi-scale LDMS for repressive points sampled from 3D shape. The multi-scale LDMS is produced by concatenating 2D images by projecting HKS based heat diffusion distribution at different scales. In addition, multi-scale LDMS shown provides much more discriminative power compared to local depth projection.

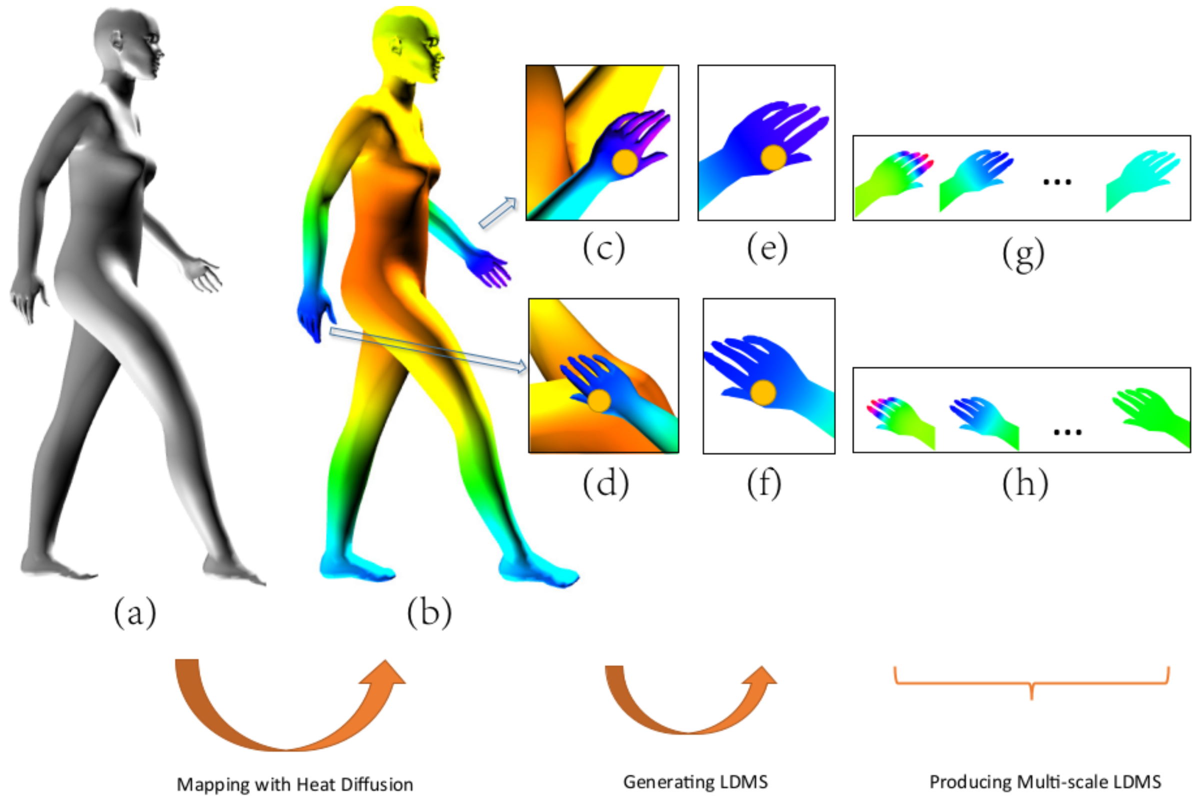


Figure 1: Pipeline of our local feature extraction method.

## Results

We conducted the experiments for shape matching to evaluate the performance of the proposed LDMS. First we conducted some experiment on SCAPE library to test the performance on non-rigid shapes. We also tested the performance of shape matching on incomplete shapes and shapes under noise. Experimental results show that our method performs best in all compared methods.

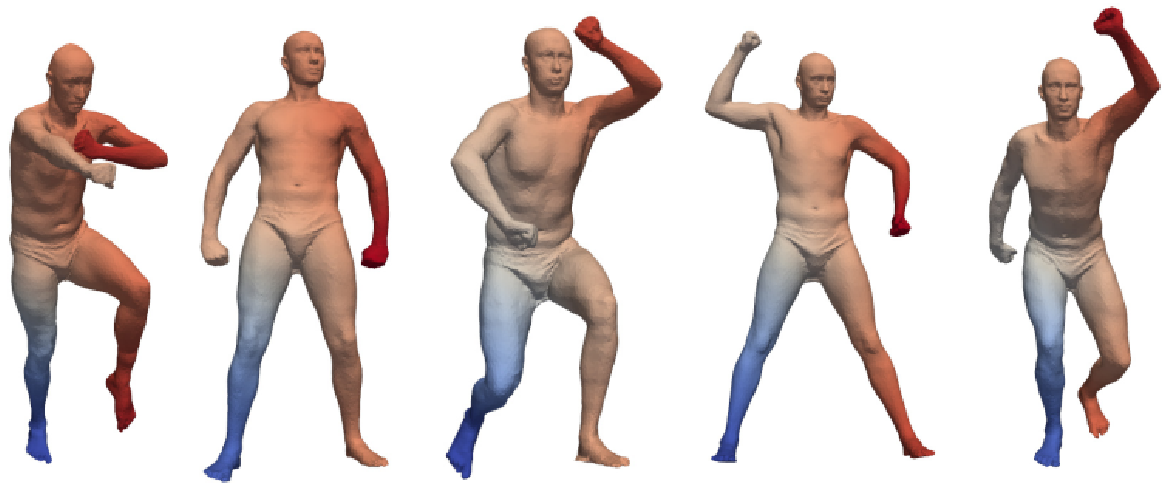

Figure 2: Matching result on SCAPE. Points with same color corresponds to each color.

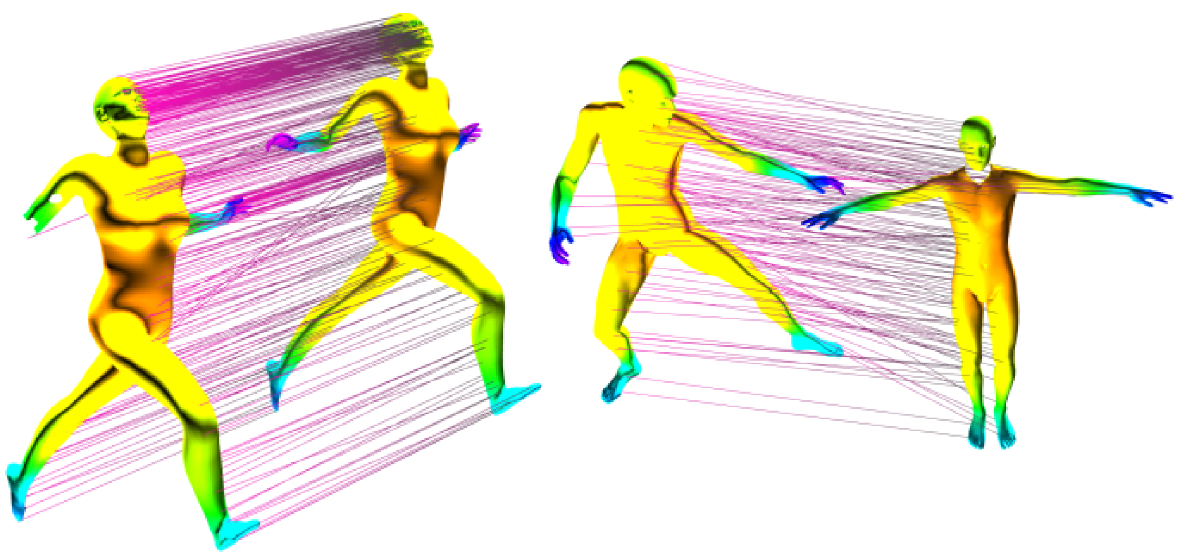

Figure 3: Partial shape matching.

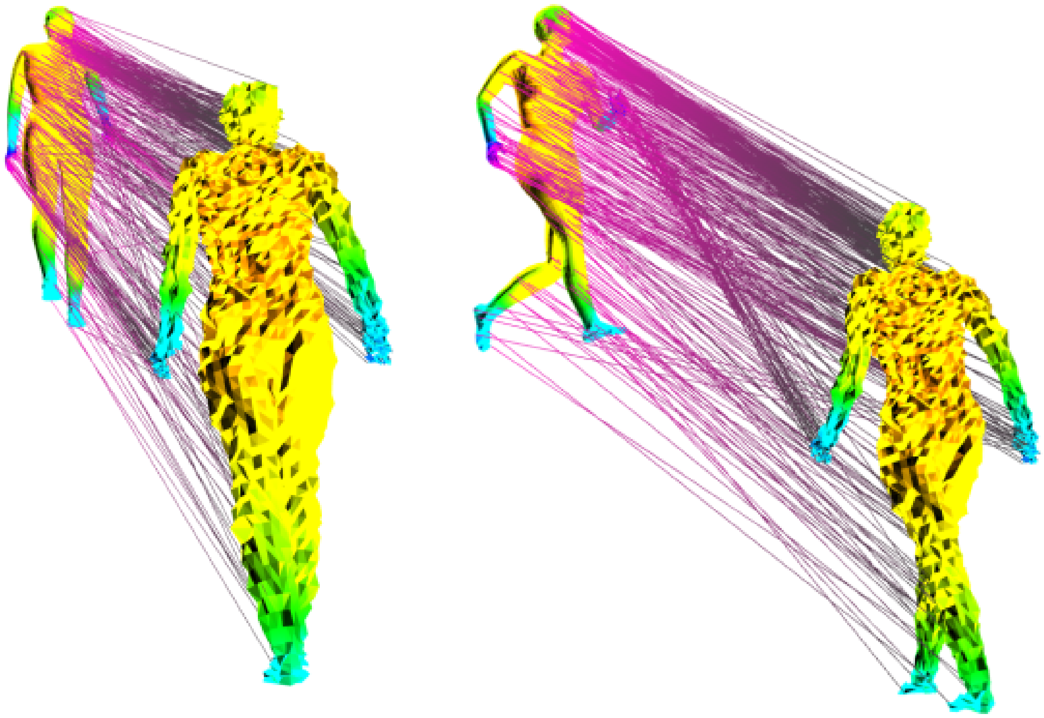

Figure 4: Noise shape matching.

| Methods | Average Geodesic Error |
|---|---|
| LDMS | **0.039** |
| Blended | 0.081 |
| Mobius Voting | 0.162 |
| Best Conformal | 0.113 |
| GMDS | 0.276 |
| HKM 1 corr | 0.272 |
| HKM 2 corr | 0.256 |

Figure 5: Result average geodesic error on SCAPE library.

**Research Project:** Learning-Based 3D Geometric Feature Extraction

## Description

Handcrafted features are only applicable to special situations and normally have a bad performance with large shape deformations. Learning based methods have much more representative power for large shape deformations. We first propose to learn category-specific dictionaries to encode subtle visual differences among different categories, and a shared dictionary to encode common patterns among all the categories to gain both representative power and discriminative power. With the LDA constraint on the coefficients of dictionary learning, the new learnt coefficients have small within-class scatter and big between-class scatter.

## Method

(i) With the sparse coding scheme, we construct a structured dictionary to represent 3D shapes, including the shared common sub-dictionary and class-specified sub-dictionaries. (ii) Based on the constructed dictionary, we impose the LDA constraint to representation coefficients to learn a novel 3D shape descriptor that is insensitive to different deformations.

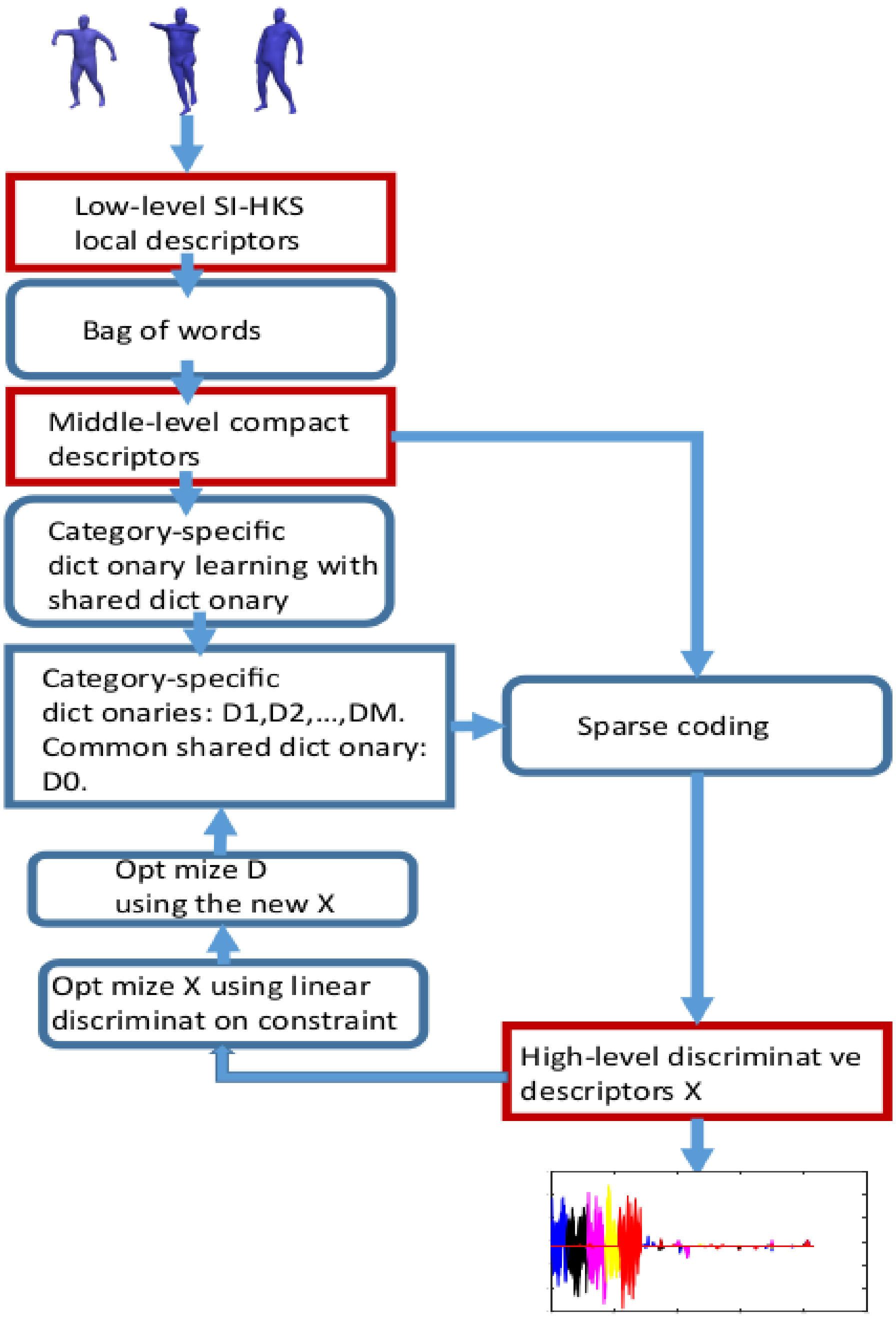


Figure 6: Flowchart of LDDL. The rectangle boxes represent input and output, the rounded rectangle boxes represent processing unit.

# Results

We conducted our experiments on SHREC10 (ShapeGoogle), SHREC 14 human body dataset to evaluate and compare the shape retrieval performance of our LDDL. we can see the LDDL method can achieve best result in synthetic dataset when evaluated using mean average precision. We can see that the LDDL get great performance in retrieval as indicated in the precisionrecall curve. And in the more challenging scanned dataset, although the proposed LDDL method only gets the second best result, our method trains much faster than the supervised sparse coding based method.

| Methods | VQ | UDL | SDL | LDDL |
|---|---|---|---|---|
| Isometry | 0.988 | 0.977 | 0.994 | **0.998** |
| Topology | **1.000** | **1.000** | **1.000** | **1.000** |
| Isometry+topology | 0.933 | 0.934 | 0.956 | **0.971** |
| Partiality | 0.947 | 0.948 | 0.951 | **0.988** |
| Triangulation | 0.954 | 0.950 | 0.955 | **0.965** |

Figure 7: Comparison of different retrieval methods in mean average precision on ShapeGoogle. 3-fold cross-validation is used.

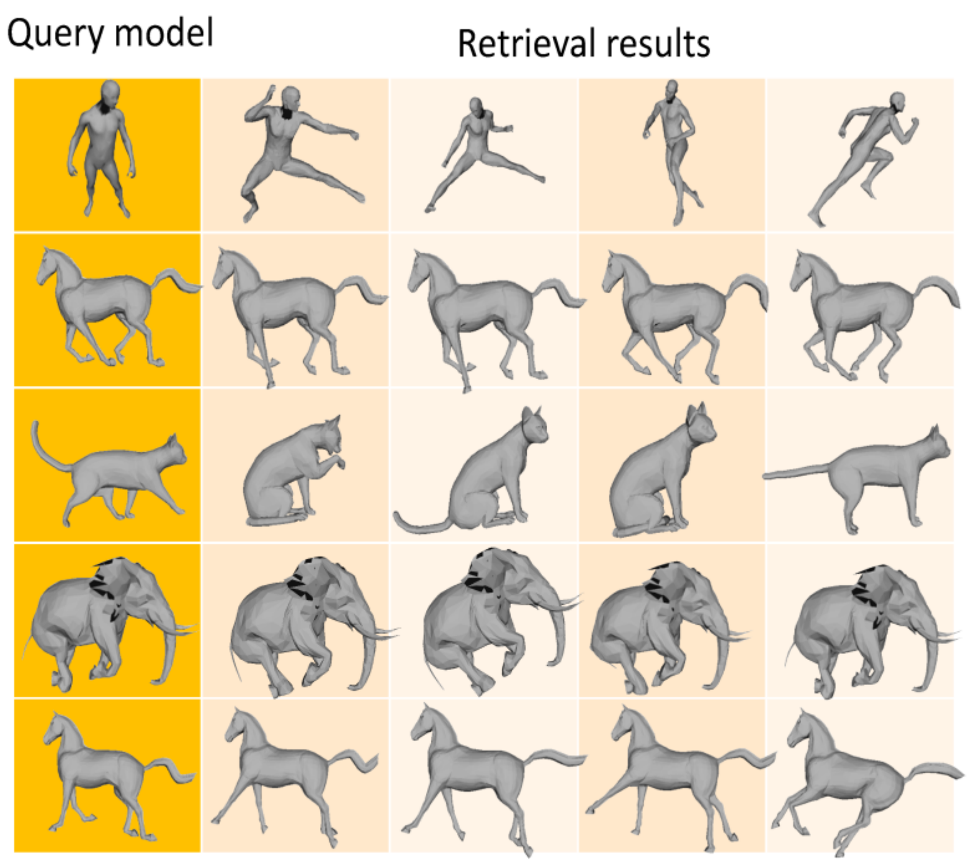


Figure 8: Shape retrieval result on ShapeGoogle dataset. Top is the precision-recall curve for shapes under both isometry and topology deformations. Bottom are first 4 retrieval results.

| Methods | Synthetic | Scanned |
|---|---|---|
| ISPM ([80]) | 0.902 | 0.258 |
| DBN ([104]) | 0.842 | 0.304 |
| R-BiHDM ([150]) | 0.642 | 0.640 |
| HAPT ([49]) | 0.817 | 0.637 |
| VQ ([15]) | 0.813 | 0.514 |
| UDL ([87]) | 0.842 | 0.523 |
| SDL ([87]) | 0.954 | **0.791** |
| LDDL | **0.972** | 0.732 |

Figure 9: Comparison of different retrieval methods in mean averaged precision on SHREC 14 Human bodies. Bold numbers are the best result of listed, underlined numbers are the second best result.

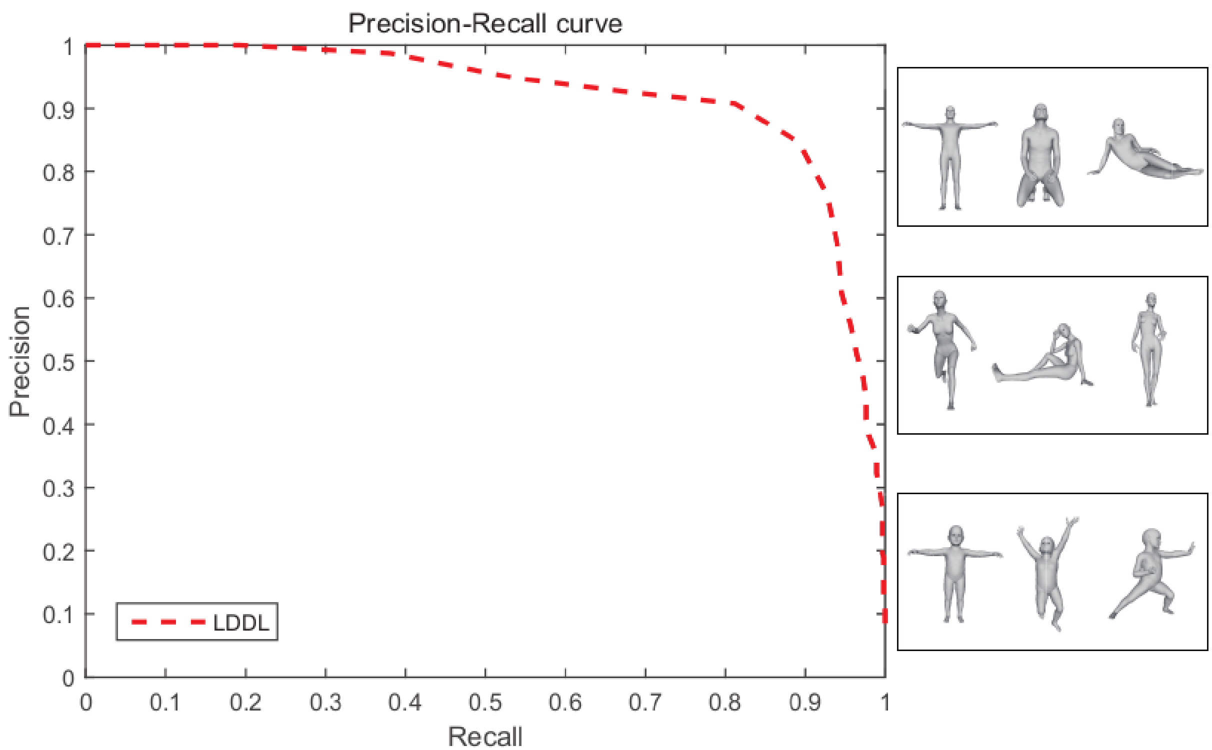


Figure 10: Precision-recall curve of SHREC14 synthetic set. Right are some shapes of the dataset.

**Research Project:** Encoder-free 3D Geometric Feature Learning

## Description

The encoder-decoder based learning method can extract 3D geometric features well, but there are several limitations. To eliminate the explicit designs of encoder and to alleviate the dependency of corresponding pairs of complete shape and partial shape, we propose a weakly supervised approach to 3D shape completion problem which does not require corresponding pairs of complete shape and partial shape. Specifically, the 3D shape completion problem can be transform to a shape deformation problem if we can deform a zero-shot shape to partial observed point cloud which requires no template for a specific category. Targeting on this goal, we use an implicit way to extract the deep shape feature by directly optimizing the latent space vector using unsupervised loss.

## Method

We introduce a shape descriptor learning method that require no explicit encoder. For a given observation $p$, we adopt a randomly initialized latent space vector $z$. In order to strengthen the representative ability of latent space vector, the value of $z$ can be vary in any range. However, the infinite latent space range resulting in a slow optimization and it make the loss hard to converge. We set a proper range of latent space to compensate those above issues. Similar with DeepSDF, we assume a multivariate-Gaussian distribution over the latent space, therefore, we design the value of $z$ to lie in a unit sphere. Without using a encoder, we directly use backpropagation to optimize $z$ by using the loss produced from the prediction of shape morphing function.

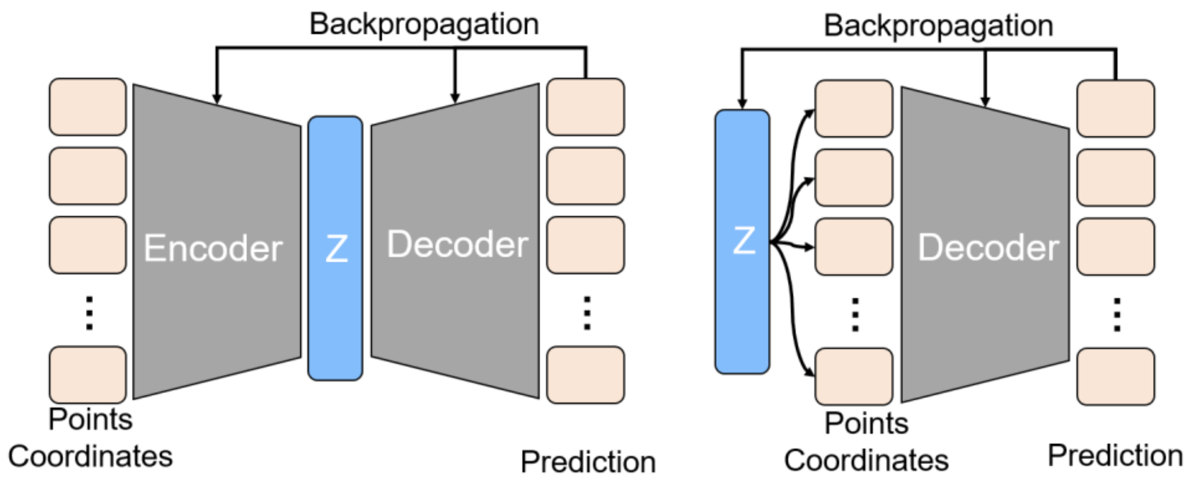


Figure 11: An auto-encoder architecture is presented on the left part. Different from the auto-encoder, the auto-decoder can encode the shape feature without a explicit encoder. A random latent space vector is optimized by the backpropaga- tion of the neural network to represent shape descriptor.

# Results

We first conduct 3D car complete experiment on ShapeNet and KITTI benchmark which is of great interesting in autonomous driving field. To validate our model can approach 3D shape completion on various categories, we use ModelNet to evaluate our method. We perform experiments on several datasets and various shapes to show the ability to predict a future motion and its correspondences. We adopt a human pose SCAPE benchmark dataset and we also use TOSCA benchmark dataset.

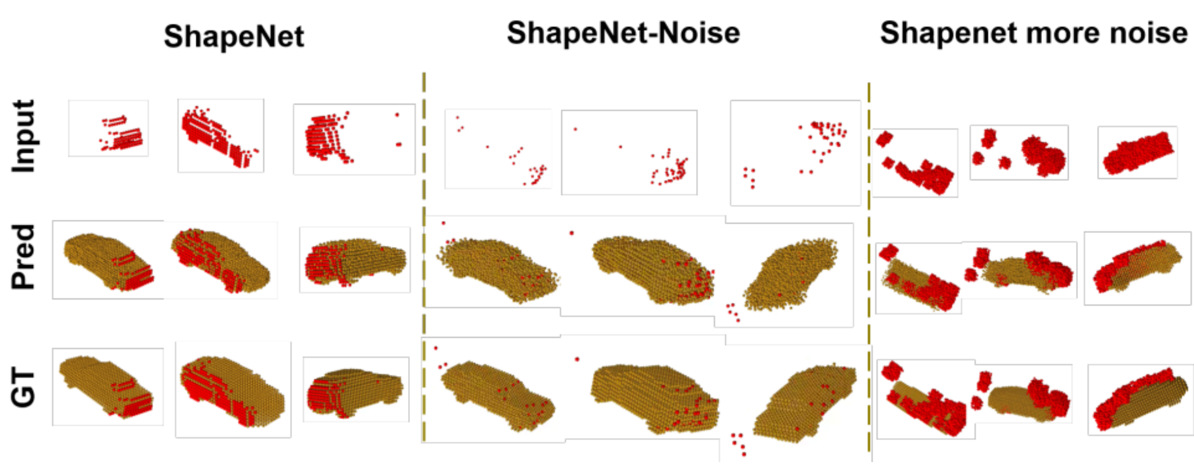


Figure 12: Qualitative Result on Shapenet.

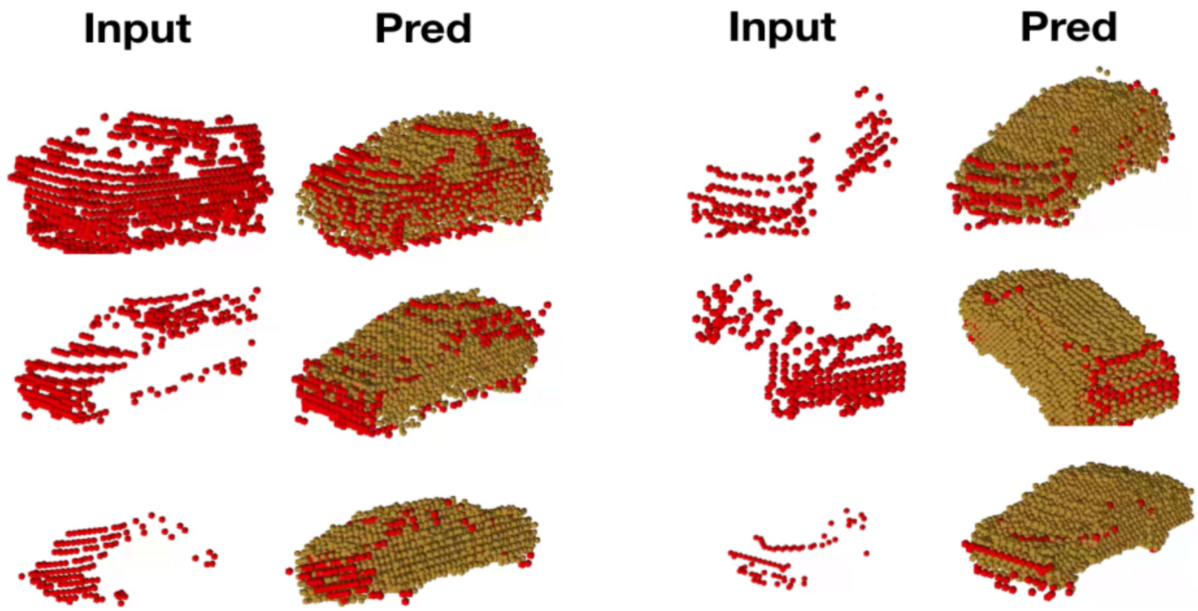


Figure 13: Qualitative Result on KITTI.

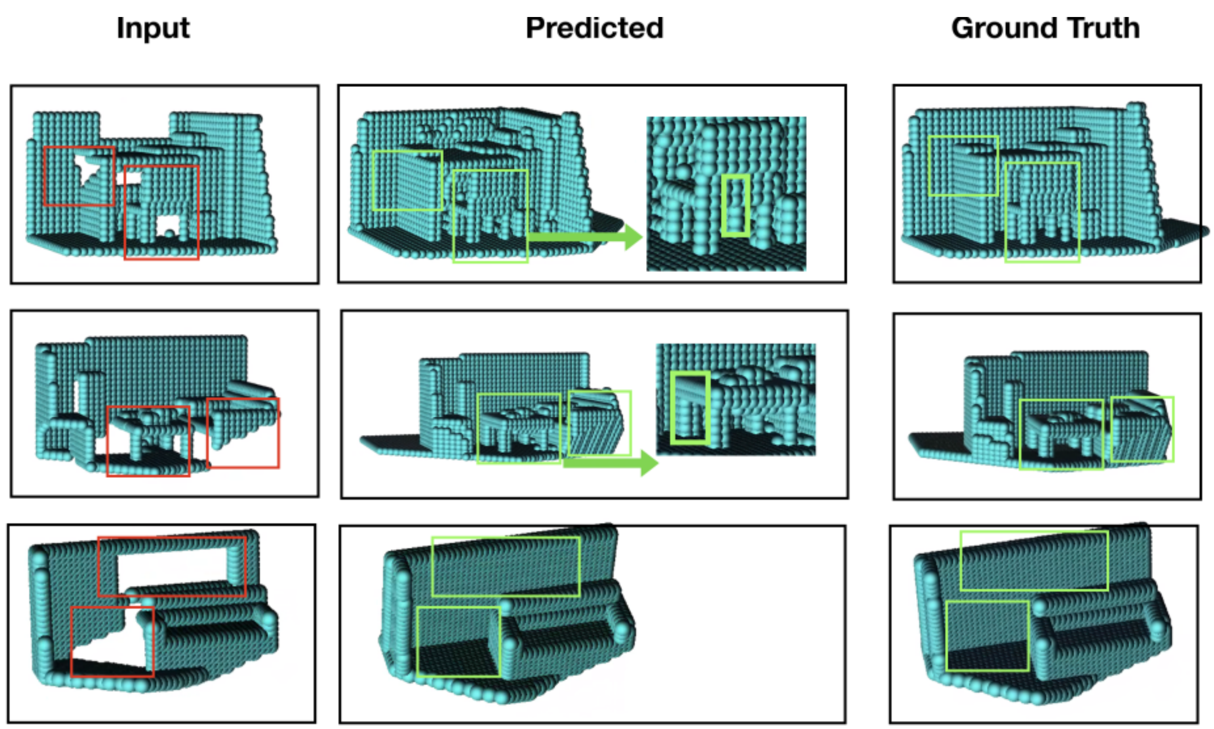


Figure 14: Qualitative result for weakly supervised scene completion. Our method can complete both the overall structure and small objects of the scene.

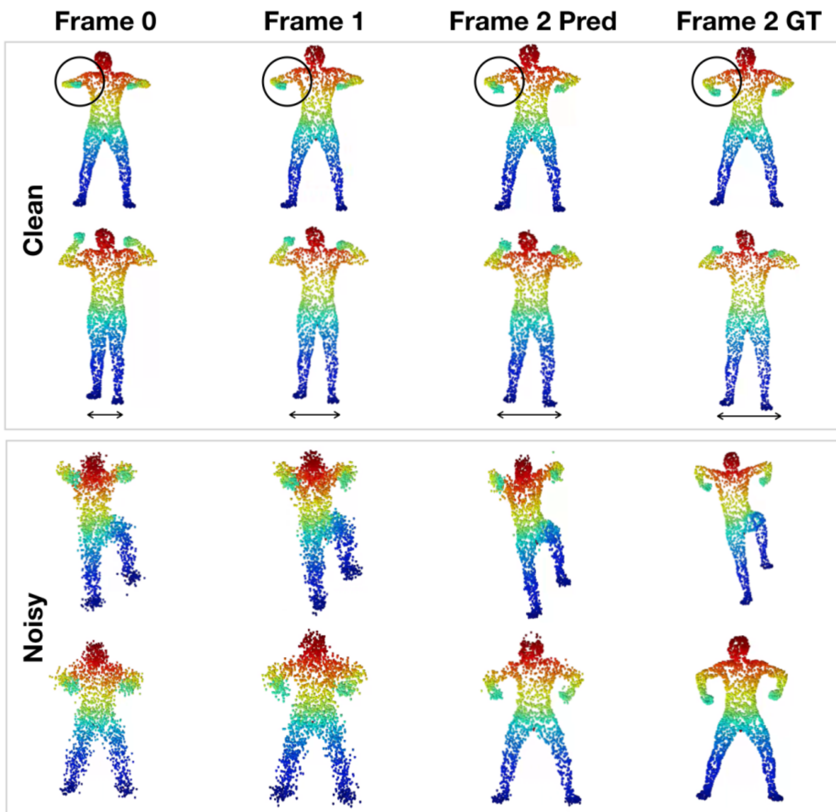


Figure 15: Qualitative Result for partial and hole.

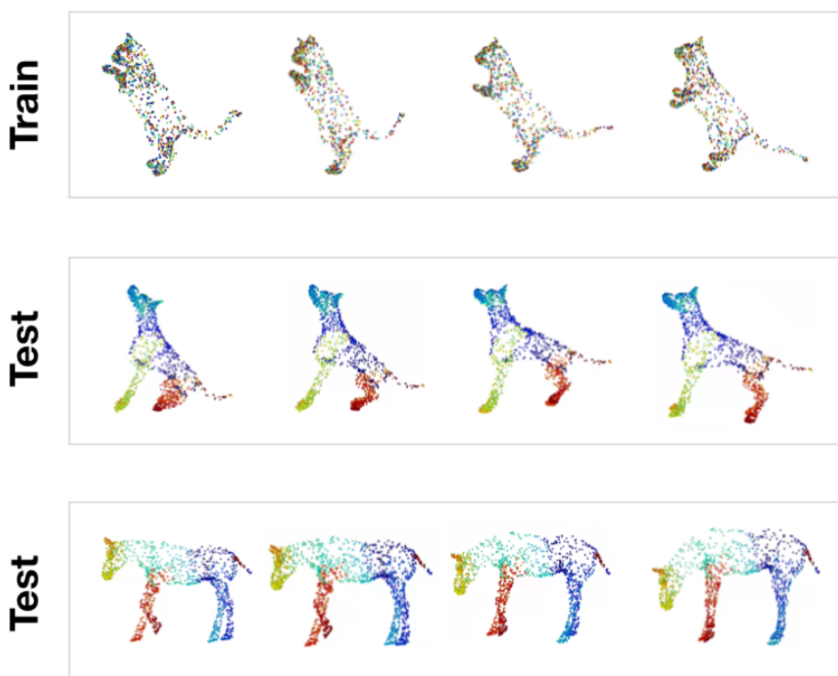


Figure 16: Cross-category Motion Prediction. We train our model in some categories and test on other categories.