Autonomous Robot Lab 自主机器人实验室

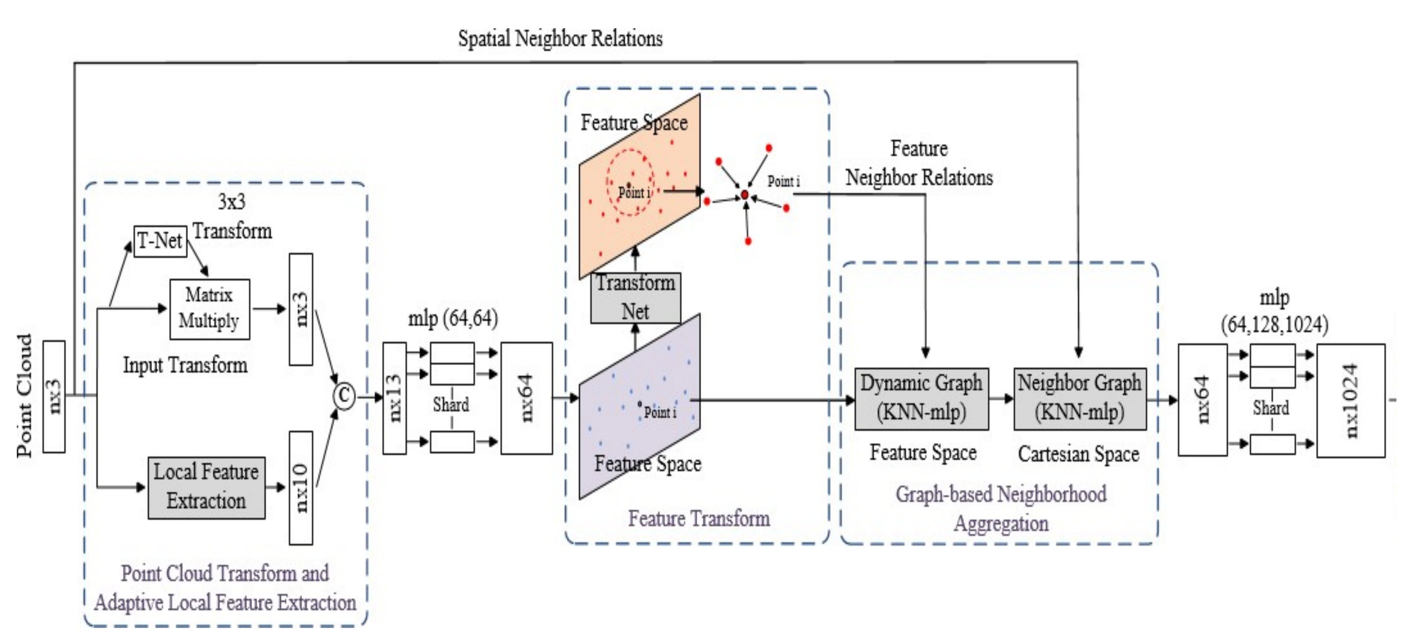
End-to-End 3D Point Cloud Learning for Registration Task Using Virtual Correspondences

Zhijian Qiao, Zhe Liu, Yueling Shen, and Hesheng Wang

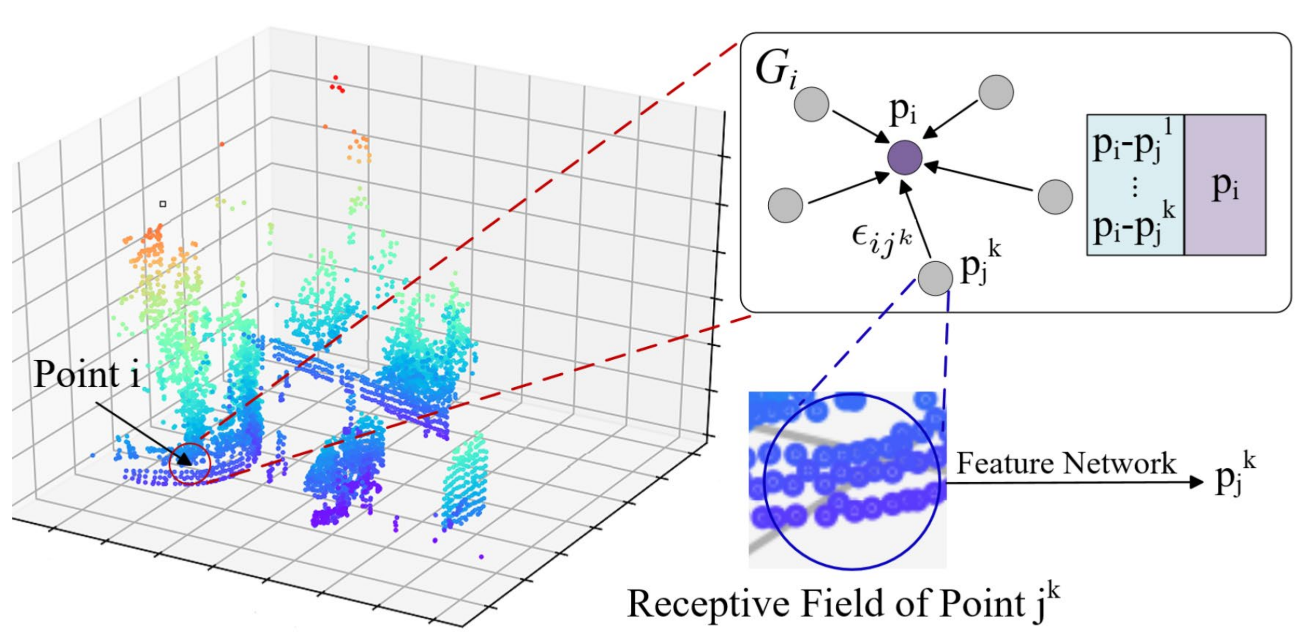
SHANGHAI JIAO TONG UNIVERSITY 1896

For robotics, autonomous driving and other fields, point cloud registration is a key problem. We present a point cloud registration method based on deep learning, which can be used for object, indoor, and large-scale point clouds. This method consists of three main parts: local feature extraction, consistent point cloud indicator, and weighted Procrustes solver. We use GPU to train our model on ModelNet40 and KITTI datasets.

Local Feature Extraction

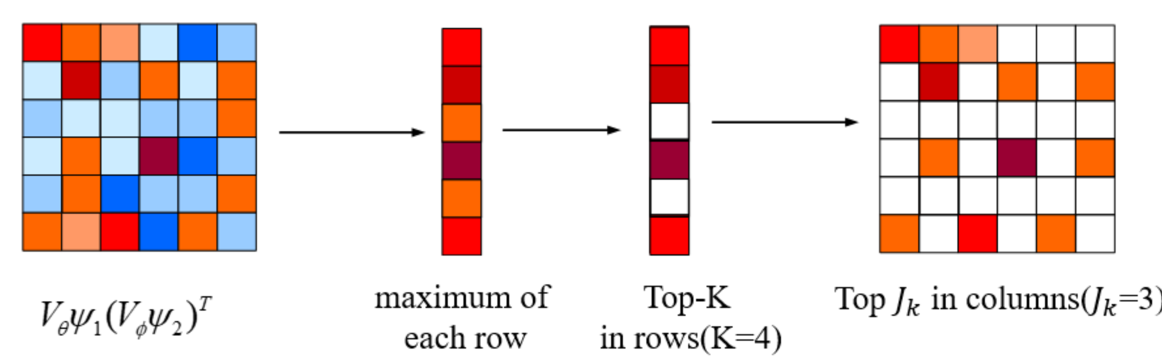


The raw point cloud data is simultaneously input to the Input Transformation Net and the Adaptive Local Feature Extractor, the former aims to ensure the rotational translation invariance of the input point coordinates, and the latter aims to fully consider the statistical local distribution characteristics.



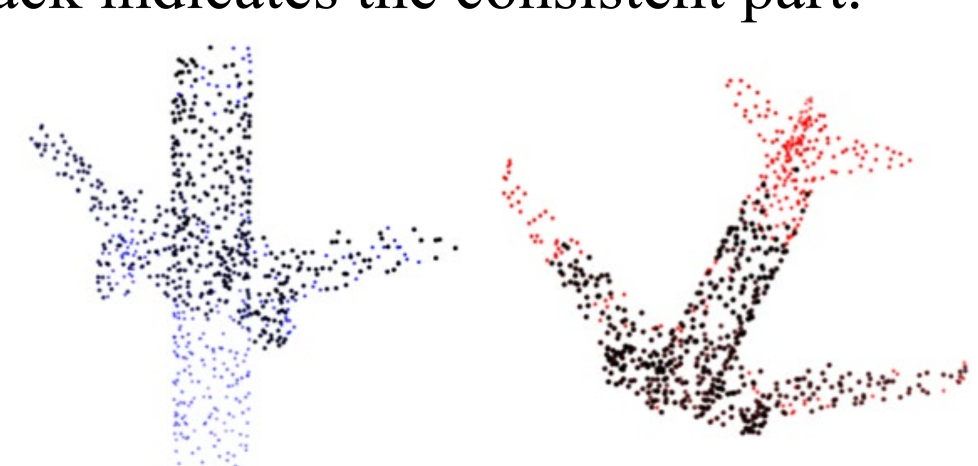
Graph formulation. Note that the receptive field of each point corresponds to a local neighborhood in the original point cloud, since the FN(Feature Network) has introduced the local structure into the feature of each point. Then we utilize GNN for feature aggregations.

Consistent Point Cloud Indicator



Based on the features extracted in the previous module, a self-attention mechanism is introduced for enhancing the static structure information and weakening the parts which are difficult to match in the large-scale point cloud. Considering the requirements of finding correspondences between two point clouds in the registration task, our method employs cross attention module. Corresponding information is aggregated into the embedding, which promotes the efficiency of finding corresponding parts. Since the two point clouds are only partially corresponded, we present a Top-K operation to find K best corresponding points from source point cloud and for each one, we find J best corresponding candidate points in target point cloud.

The visualization is as follows. The left is source point cloud while the right is the target. The black indicates the consistent part.

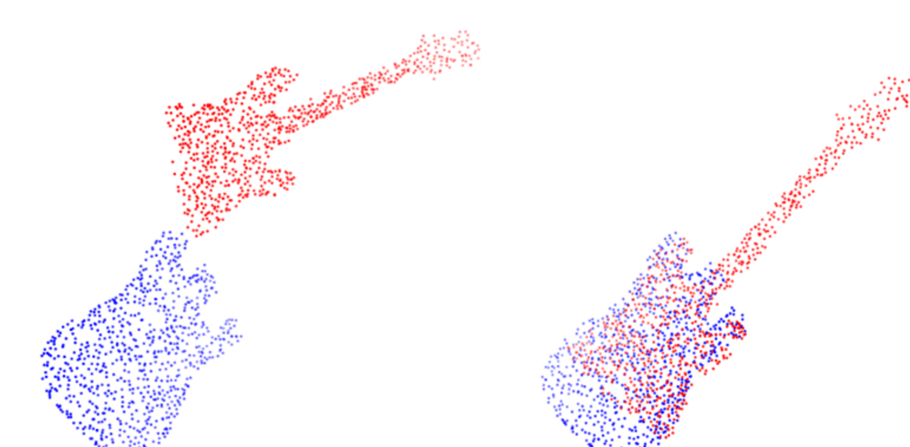


Weighted Procrustes Solver

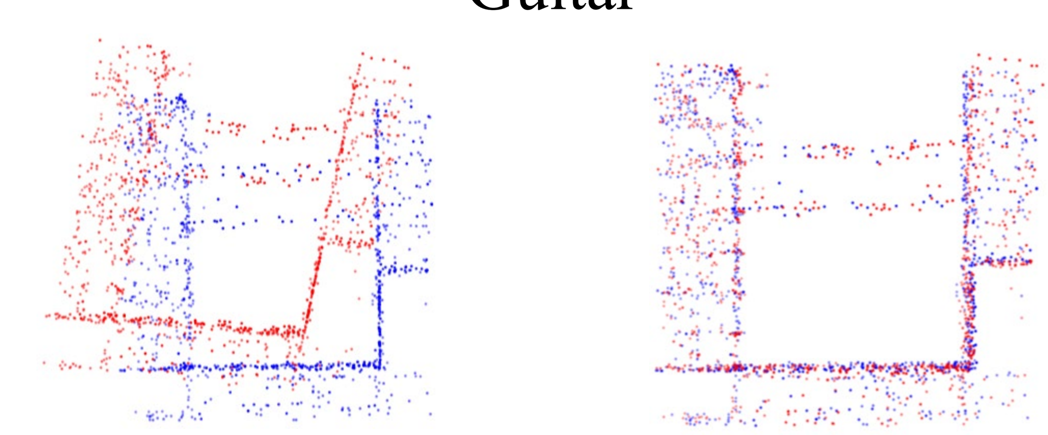
The inlier likelihood estimated by the attention mechanism provides a weight for each correspondence. The original Procrustes method minimizes the mean squared error between corresponding points $\frac{1}{N}\sum_{(i,j)\in\mathcal{M}}\|\mathbf{x}_i-\mathbf{y}_j\|^2$. In contrast, we minimize a weighted mean squared error. This change allows us to We can reduce the influence of unreliable points on the pose estimation to increase the accuracy. Formally, Weighted Procrustes analysis minimizes:

$$\begin{aligned} e^2 &= e^2(R,\mathbf{t};\mathbf{w},X,Y) \\ &= \sum_{(i,j)\in\mathcal{M}} \tilde{w}_{(i,j)}(\mathbf{y}_j-(R\mathbf{x}_i+\mathbf{t}))^2 \\ &= \operatorname{Tr}\left((Y-RX-\mathbf{t}\mathbf{1}^T)W(Y-RX-\mathbf{t}\mathbf{1}^T)^T\right) \end{aligned}$$

All experiments are conducted on a desktop computer with two Nvidia GTX 1080Ti GPUs. The results of point cloud registration are as follow.



Guitar



Real scene