PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation

论文首先所描述的知识是三维基本的数据类型，查阅资料后有如下几种：

1.多视角（multi-view）：通过多视角二维图片组合为三维物体，此方法将传统CNN应用于多张二维视角的图片，特征被view pooling procedure聚合起来形成三维物体；

2.体素（volumetric）：通过将物体表现为空间中的体素进行类似于二维的三维卷积（例如，卷积核大小为5x5x5），是规律化的并且易于类比二维的，但同时因为多了一个维度出来，时间和空间复杂度都非常高，目前已经不是主流的方法了。（论文中有所体现，主要的缺点是表现转换时候会出现大量无效的数据，增加其人工操作）

3.点云（point clouds）：直接将三维点云抛入网络进行训练，数据量小。主要任务有分类、分割以及大场景下语义分割；（论文中所描述的方法处理的数据类型）

4.非欧式（manifold，graph）：在流形或图的结构上进行卷积，三维点云可以表现为mesh结构，可以通过点对之间临接关系表现为图的结构。（目前还没有涉及）

第二涉及到数据转化的不变性。序列问题，作者借鉴的是NLP过程中的处理方法，见论文（O. Vinyals, S. Bengio, and M. Kudlur. Order matters: Sequence to sequence for sets. arXiv preprint arXiv:1511.06391, 2015. 2, 4, 7）



整个网络的结构图，第一个所需要解决的问题是怎么处理点云数据的对称性和无序性。



论文构建了一个函数g，达到对称的效果，用上图的转化网络结构来表示。这个网络结构为什么能够起到这种效果未知。从中学习的网络结构有借鉴意义。

论文的特点在于其对称性变换，在实验中对比的数据集用的是不同的表达方式。

MLP多层感知机层就能处理该问题，其中没有类似于2D的卷积处理。