

PMP-Net 通过学习多步的点移动路径实现点云补全

Wen X, Xiang P, Han Z, et al. Pmp-net: Point cloud completion by learning multi-step point moving paths[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021: 7443-7452.

[PDF 下载链接](#)

[原始代码](#)

摘要

点云补全任务的目的是预测不完整的三维形状中的缺失部分。一种广泛使用的策略是从不完整的点云中生成一个完整的点云。然而，点云的无序性削弱了生成高质量三维形状的可能，因为离散点的结构和拓扑细节很难通过仅使用隐式代码的生成过程来捕获。在本文中，我们通过从一个新的角度来重新考虑如何解决上述问题，即将预测表述为一个点云的变形过程。具体来说，我们设计了一个新的神经网络，名为PMP-Net，来模拟地壳运动的方式。通过移动不完整输入的每个点来补全点云，其中点移动路径（Point Moving Paths, PMP）的总距离应该是最短的。因此，PMP-Net根据所有点移动距离这个约束，为每个点预测一个唯一的点移动路径。因此，该网络在点水平上学习了一个严格的和唯一的对应关系。从而提高了预测的完整形状的质量。我们在 Completion3D 和 PCN 数据集上进行了全面的实验，并且证明了我们相对于最先进的点云补全方法存在优势。

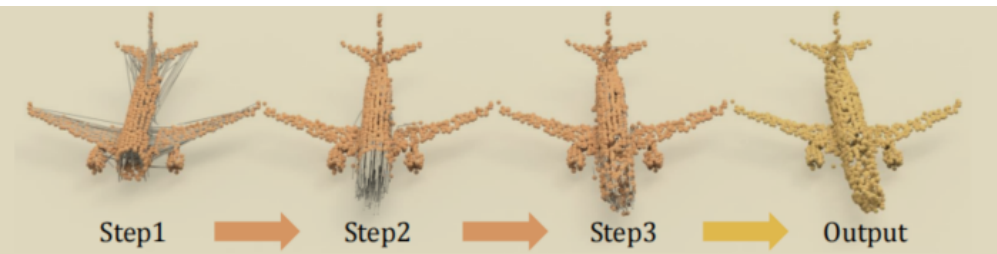


图1：多步骤点云变形的补全。点移动路径用灰色线表示。在每一步中，原始点云使用橙色表示，目标点云使用黄色表示。

Ch01 介绍

点云作为一种广泛使用的三维形状表示方法，可以很容易地通过深度相机或者其他三维扫描设备获得。由于三维扫描设备的观察点或者遮挡的局限性，原始点云通常是稀疏的和不完全的^[41]。因此，对于下游的三维计算机视觉应用，如：分类、分割及其他视觉分析工作，形状的补全或者合并过程通常需要生成三维形状的缺失区域。

在本文中，我们专注于由点云表示的三维对象的补全工作，其中的缺失部分是由于扫描仪视角的自遮挡造成的。以往的方法大多将点云补全表示为点云生成问题，其中通常采用编码器框架从输入的不完整点云中提取潜在编码，并且将提取的潜在编码解码为完整的点云。利益于基于深度学习的点云学习方法，这方面的点云补全方法^{[41][43]}在过去几年中取得了巨大的进展。然而，由于使用深度神

经网络生成点云仍然是一项困难的任务，因为点云的无序性导致生成模型难以捕获离散点之间的结构和拓扑细节^[34]，因此，基于生成模型点云补全质量难以让人满意。

为了解决这个问题，本文提出了一种新型的神经网络（PMP-Net），从一个新的角度来描述点云补全任务。与生成模型不同的是，生成模型是直接预测三维空间中所有点的坐标，而PMP-Net通过学习将原始的三维形状中的点移到目标位置。通过点的移动过程，PMP-Net建立了原始点云与目标点云之间的对应关系，捕获两个点云之间的结构关系和拓扑细节。另一方面，因为两个点云之间将点从原始位置移动到目标位置存在着各种各样的路径，这些路径将使训练过程变得混乱。因此，为了鼓励网络学到一个特定的点移动路径，我们在地动距离（Earth Mover's Distance, EMD）的启发下，提出了有序的点移动路径网络（PMP-Net）和所有点的移动距离（Point Moving Distance, PMD）这样的网络约束。

此外，为了更加准确地预测点移动路径，我们还提出了一种多步骤路径搜索策略，即在多尺度搜索半径下对点移动路径进行连续调优。具体来说，路径搜索以一种非常精确的方式重复了多个步骤（见图2）。每一步都将考虑到前面预测的路径，然后根据前面的路径规划其下一个移动路径，为了记录和聚合点移动路径的历史信息，我们受门控循环单元（GRU）的启发，提出了一种新的循环路径聚合（Recurrent Path Aggregation, RPA）模块。这种模块可以记住和聚合每个点的路线序列，并且将之前的信息与点的当前位置结合起来，以预测下一步移动的方向和长度。通过逐步减少搜索半径，PMP-Net可以持续地为每个点细化一个越来越精确的路径，实现点从不完整的点云的原始位置移动到完整点云的目标位置。总之，我们的工作的主要贡献如下：

- 我们提出了一种新的网络（PMP-Net）用于点云补全任务，网络将每个点从不完整的形状移动到完整的形状，从而实现高质量的点云补全。与以往的点云补全方法相比，PMP-Net通过预测点移动路径，学习得到了“点”级别的对应关系，能够更加详细地了解不完整形状和完整形状之间的更多的结构关系和拓扑细节。
- 通过对使用了地动距离（EMD）约束的网络的正则化，我们提出了一种学习输入点云与输出点云之间特定的点移动路径。因此，网络不会被点移动的多重解所混淆，并且最终预测出原始点云和目标点云之间的有意义的“点”级别的对应关系。
- 我们提出了从粗到细的方式多步骤搜索点移动路径。每个步骤都将通过使用所提出的循环路径聚合（RPA）模块，根据来自先前路径的聚合信息及其当前位置来决定下一步的移动。

Ch02 相关工作

三维重建和表示学习中的深度学习技术促进了三维形状补全的研究，这些研究大致分为两类：

- H1
1. 传统的三维形状补全方法^{[1][33][35][15]}，通常基于手搓特征（如：表面光滑度或者对称轴）来推断缺失的区域；而其他一些方法^{[19][25][30][31]}考虑使用大规模的三维形状数据集来辅助，通过搜索找到相似的补丁来三维形状的不完整区域。

2. 基于深度学习的方法^{[14][16][29][50]}利用其强大的表示学习能力，从不完整的输入形状中提取几何特征，并且根据所提取的特征直接推断出完整的形状。与传统的补全方法相比，这些可以学习的方法不需要预先定义的手工特征，并且可以更好地利用大规模补全数据集中的丰富的形状信息。

我们所提出的PMP-Net也属于基于深度学习的方法，基于上述的分类方式可以进一步分类如下：

- **体素辅助的形状补全：**卷积神经网络（CNN）的表示学习能力在二维计算机视觉研究中得到了广泛的应用，近年来有关二维图像内绘制的应用研究持续爆发。从 2D CNN 在图像绘制研究的成功中，可以直观地借鉴其思想到三维形状补全的研究中。近年来，人们提出了几种基于 3D CNN 结构的体积辅助形状补全方法。请注意，我们使用术语“体素辅助”来描述这类方法，是因为三维体素通常不是网络的最终输出。相反，预测得到的体素将被进一步细化，并且转换为其他表示（如：网络^[2]或点云^[42]），以产生更加精细的三维形状。因此，体素更像是一个中间的辅助工具，用于帮助补全网络推断出完整的形状。值得注意的是，3D-EPN^[2]和GRNET^[42]已经提出了从粗到细的方式重建完整的三维体素。它们首先在编码器-解码器框架下使用 3D CNN 预测一个粗糙的完整形状，然后使用从完整形状数据集中选择的类似补丁来细化输出^[2]，或者根据输出体素进一步重建点云的细节^[42]。此外，还有一些研究考虑了纯粹的体素数据用于形状补全任务。例如：Han等人^[4]提出了通过同时推断全局结构和局部几何来预测完整形状的细节，从而直接生成高分辨率的三维体素形状；Stutz等人^[32]提出了一种基于变分自动编码器的方法来在弱监督下补全三维体素数据。尽管 3D CNN 在特征学习方面具有极好的能力，但是计算成本是输入的体素数据的分辨率的立方，使得处理细粒度形状变得困难。
- **基于点云的形状补全。**近年来，由于点云是许多三维扫描设备的直接输出结果，而且点云的存储和处理需要的计算机成本也比体积数据要低许多，因此最近的许多研究都考虑在三维点云上执行补全任务，因此这类方法也越来越受关注。从点云表示学习^{[26][27]}的改进中可以看出，以往的方法（如：TopNet^[34], PCN^[48], SA-Net^[41]）将解决方案作为编码器-解码器框架下的生成模型。他们采用PointNet^[26]等编码器从不完整的提取全局特征，并且使用解码器根据提取的特征推断完整点云。与PCN^[48]相比，TopNet^[34]改进了解码器的结构，以便在一个根树结构^[34]中隐式地建模和生成点云。SA-Net更进一步，通过跳跃-注意机制，保存与传递了不完整形状的几何信息的细节。其他著名的工作（如：RLGAN-Net^[29]和Render4Completion^[13]）专注于对抗性学习框架，以提高生成的完整形状的真实性和一致性。综上所述，上述方法大多是点云补全任务的生成式解决方案，不可避免地存在点云的无序性，这使得使用生成解码器重建详细的类型或者结构变得困难。因此，为了避免预测无序数据的问题，PMP-Net采用了一种不同的方法的来重建完整点云，通过学习从初始输入开始移动所有点，代替从潜在编码中直接生成最终点云。PMP-Net的思想也与三维形状变形^[45]的研究有关，^[45]只考虑了一步变形。然而，在不完整的形状和完整的形状之间的一步变形极具挑战性，这需要在没有任何其他先验信息的情况下，推断缺失区域中完全未知的几何形状。相比之下，我们提出了多步搜索，以鼓励PMP-Net对缺失区域推断出更多几何信息的细节，并且正则化点的移动距离保证多步推理的效率。

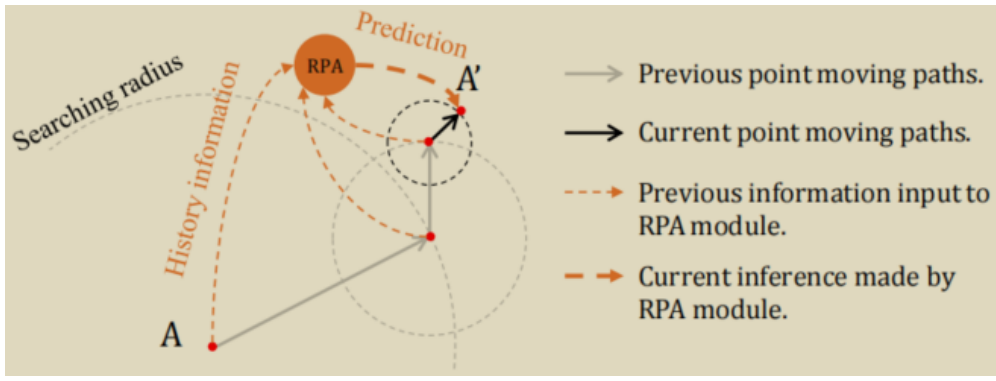


图2：从粗到细的搜索半径下使用多步骤路径搜索。PMP-Net将点A移动到A'需要通过三个步骤，每一步都减小其搜索半径，并且回顾移动历史，以决定下一个需要移动的位置。

Ch03 PMP-Net的架构

3.1. 点位移预测

多步骤架构

H1

提出的PMP-Net的概念如图3(a)所示，在给定输入点云 $P = \{\mathbf{p}_i\}$ 和目标点云

H2

$P' = \{\mathbf{p}'_i\}$ 时，PMP-Net的目标是预测一个位移向量集 $\Delta P = \{\Delta \mathbf{p}_i\}$ ，该向量集将 P 中的每个点通过三步移动到 P' 中的位置上，即： $\{(\mathbf{p}_i + \Delta \mathbf{p}_i)\} = \{\mathbf{p}'_j\}$

H3

，其中第 k 步的位移向量记为 $\Delta \mathbf{p}_i^k$ ，得 $\Delta \mathbf{p}_i = \sum_{k=1}^3 \Delta \mathbf{p}_i^k$ 。对于步骤 k ，网络将第 $k-1$ 步的变形点云 $\{\mathbf{p}_i^{k-1}\} = \{\mathbf{p}_i + \sum_{j=1}^{k-1} \Delta \mathbf{p}_i^j\}$ 作为输入点云，并且依据其计算新的位移向量。因此，预测的形状将一步步地细化，最终产生一个高质量的补全形状。

位移向量预测

H3

在第 k 步时，为了预测每个点的位移向量 $\Delta \mathbf{p}_i^k$ ，我们首先从点云中提取每个点的特征。这种特征是由PointNet++[27]的基本框架从输入的三维形状的全局特征中提取的，然后使用特征传播模块在三维空间中将全局特征传播到每个点，并且最终为每个点 \mathbf{p}_i^k 生成特征 $\mathbf{h}_i^{k,l}$ 。由于我们的实验实现了将特征传播的三个层次应用到了分层次生成每个点的特征（见图3(b)） $\mathbf{h}_i^{k,l}$ ，其中上标 k 表示步骤、下标 l 表示层次。然后每个点的特征 $\mathbf{h}_i^{k,l}$ 与随机噪声向量 \hat{x} 拼接起来，从而给每个点一个很小的扰动从而迫使它们离开原始的位置[45]。然后，最终点特征 $\mathbf{h}_i^{k,3}$ 在第 k 步第3层送入其后的多层感知机（MLP）及一个tanh激活函数，输出一个三维向量作为点 \mathbf{p}_i^k 的位移向量 $\Delta \mathbf{p}_i^k$ ：

$$\Delta \mathbf{p}_i^k = \tanh\left(MLP([\mathbf{h}_i^{k,3} : \hat{x}])\right) \quad (1)$$

其中的冒号‘:’表示拼接操作。

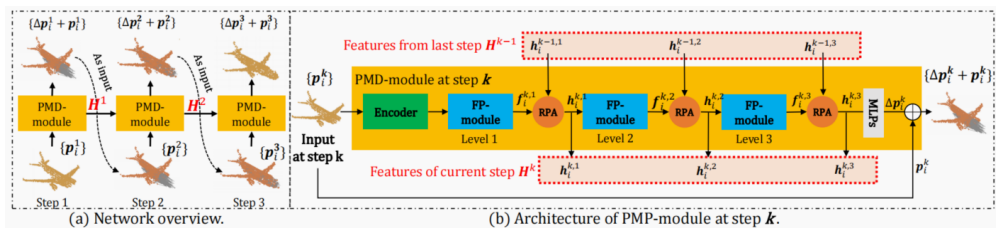


图3：基于从粗到细的搜索半径实现的多步骤路径搜索。PMP-Net通过三步将点 A 移动 A' ，每一步都减少搜索半径，回顾移动历史，决定下一个要移动的位置。

各步骤之间的循环信息流

前面的移动信息对于网络决定当前的移动是至关重要的，因为前面的路径可以用来推断单个点的最终位置。因此，这些信息可以引导网络找到下一步移动的方向和距离，防止网络在点移动路径搜索的多个步骤中改变了最终目的。为了满足这个要求，我们提出了在特征传播模块的每一步和每一层都使用特定RPA单元，这个模块可以记忆先前路径的信息，并且可以用于推导每个点的下个位置。 RPA 模块在 $(k$ 步、 l 层)的输入是第 $i-1$ 层的输出，并将之与第 $k-1$ 步得到的特征 $\mathbf{h}_i^{k,l-1}$ 混合，从而在相同的 l 层输出当前层的特征 $\mathbf{h}_i^{k,l}$ ：

$$\mathbf{h}_i^{k,l} = RPA(\mathbf{f}_i^{k,l-1}, \mathbf{h}_i^{k-1,l}) \quad (2)$$

其中， RPA 模块的细节参见上述。

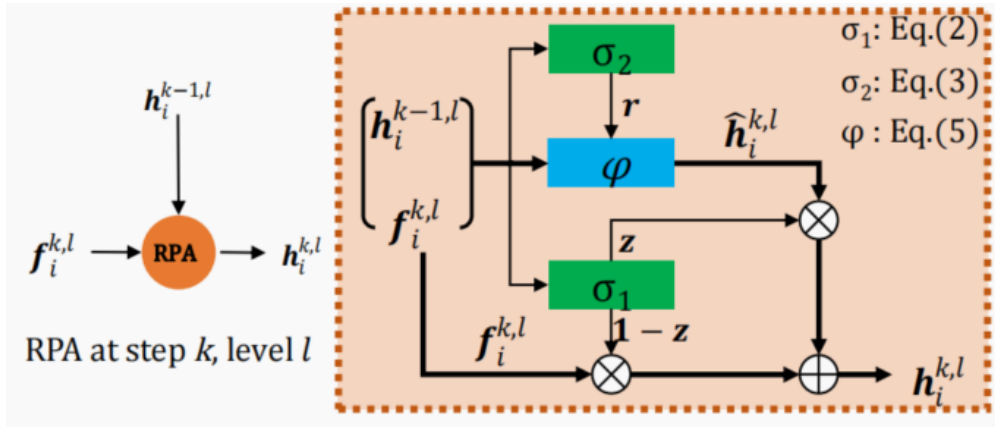


图4：RPA模块在第 k 步、第 l 层的结构细节

3.2. 循环路径聚合

点移动的先前路径可以看作为序列数据，其中包含的每次移动的信息将在此过程中被选择性地记住或者忘记。遵循这个理念，我们利用递归神经网络提供的灵感，模仿了门循环单元（GRU）的行为，计算更新门 z 和重置门 r 从而编码和遗忘信息，这是依据第 $k-1$ 的点特征 $\mathbf{h}_i^{k-1,l}$ 和第 k 步的点特征 $\mathbf{h}_i^{k,l-1}$ 实现的。两个门计算的形式化公式如下：

$$z = \sigma(W_z[\mathbf{f}_i^{k,l-1} : \mathbf{h}_i^{k-1,l}] + \mathbf{b}_z) \quad (3)$$

$$r = \sigma(W_r[\mathbf{f}_i^{k,l-1} : \mathbf{h}_i^{k-1,l}] + \mathbf{b}_r) \quad (4)$$

其中， W_z, W_r 是权重矩阵， $\mathbf{b}_z, \mathbf{b}_r$ 是偏差。 σ 是sigmoid激活函数，用于预测一个0到1之间的值，表示允许通过门的信息的比率。冒号‘:’表示两个特征的拼接。

与标准的GRU不同的是，RPA在计算输出特征 $\mathbf{h}_i^{k,l}$ 时，强调的是保存先前的信息更为重要，因此我们解决了在当前输入信息中保存更多重要性的问题，并且提出了计算输出特征 $\mathbf{h}_i^{k,l}$ 的公式：

$$\mathbf{h}_i^{k,l} = z \odot \hat{\mathbf{h}}_i^{k,l} + (1 - z) \odot \mathbf{f}_i^{k,l-1} \quad (5)$$

其中， $\hat{\mathbf{h}}_i^{k,l}$ 是当前步骤的中间特征。公式包含了过去路径的保留信息，并且是根据当前的输入特征计算得到的，其计算公式为：

$$\hat{\mathbf{h}}_i^{k,l} = \varphi(W_h[r \odot \mathbf{h}_i^{k-1,l} : \mathbf{f}_i^{k,l-1} + \mathbf{b}_h]) \quad (6)$$

其中， φ 在我们的实现中是relu激活函数。

使用 $\hat{\mathbf{h}}_i^{k,l}$ 与 $\mathbf{f}_i^{k,l-1}$ 的融合特征，而不是 $\mathbf{h}_i^{k-1,l}$ 是因为相比RNN中的标准单元，当前点的位置对下次移动的决策影响力更大一些。特别是，当RPA模块需要忽略在当前决策中并不重要的先前信息时，因此等式(5)可以通过简单地将更新门 z 置为零向量就能忘记所有的历史，从而使得RPA模块完全关注当前输入 $\mathbf{f}_i^{k,l-1}$ 里的信息。

3.3. 唯一路径的优化搜索

最小化移动距离

H2 如图5所示，点云的无序性允许在输入形状变形为目标形状时出现多个解，并且在变形后的形状上的直接约束（如：倒角距离）和其基准数据无法保证输入点集与目标点集之间建立的对应关系的唯一性。否则，网络就会被点移动的多个解所混淆，这可能导致无法捕捉不完整形状和补全后的形状之间的拓扑细节与结构关系。
H3 为了在输入点云和目标点云之间建立唯一的和有意义的点对应关系，我们受到地动距离[²⁸]的启发，提出了训练PMP-Net在源点云和目标点云之间，基于所有点的移动距离约束，学习路径规划 ϕ 。具体来说，给定源点云 $\hat{X} = \{\hat{x}_i | i = 1, 2, 3, \dots, N\}$ 和目标点云 $X = \{x_i | i = 1, 2, 3, \dots, N\}$ ，我们遵循EMD约束来学习 ϕ ：

$$\mathcal{L}(\hat{X}, X) = \min_{\phi: \hat{X} \rightarrow X} \frac{1}{|\hat{X}|} \|\hat{x} - \phi(\hat{X})\| \quad (7)$$

在等式(7)中， ϕ 被看作双向函数，用于最小化 \hat{X} 与 X 之间的对应点的平均距离。

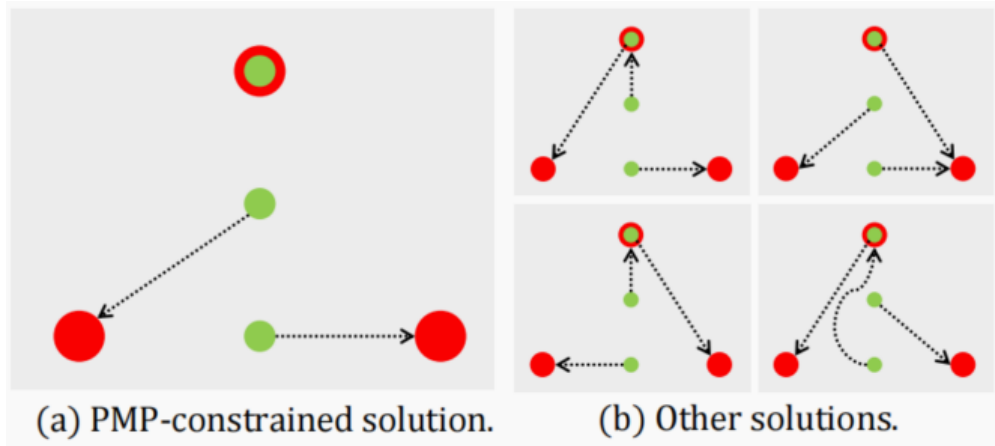


图5：将输入点云（绿色）变形为目标点云（红色）时多个解的情况描述。PMD约束保证了输入点云与目标点云之间的点对应关系的唯一性（见图5a），并且过滤了移动点的各种冗余解（见图5b）

根据等式(7)可知，双射函数 ϕ 是通过网络最小化移动距离时获得的，移动距离是从输入形状到目标形状移动点时产生的。然而，即使输入点云与目标点云之间的对应关系是唯一的，实现两个点云的对应关系的移动路径并不唯一（见图6）。因此，为了鼓励网络学习得到一个最优的点移动路径，我们选择最小化点移动距离损失函数（ \mathcal{L}_{PMD} ），即所有位移向量 $\{\Delta \mathbf{p}_i^k\}$ 的和，这些位移向量是由PMP-Net的三个步骤产生的。点移动距离损失函数的公式如下：

$$\mathcal{L} = \sum_k \sum_i \|\Delta \mathbf{p}_i^k\|_2 \quad (8)$$

等式(8)比EMD的约束更加严格。它不仅要求所有点的整体位移达到最短距离，而且还限制了每一步中点的移动路径必须最短。因此，在每一步中，都会鼓励网络按照之前的方向搜索新的路径（见图6），这将减少冗余的移动决策，提高搜索效率。

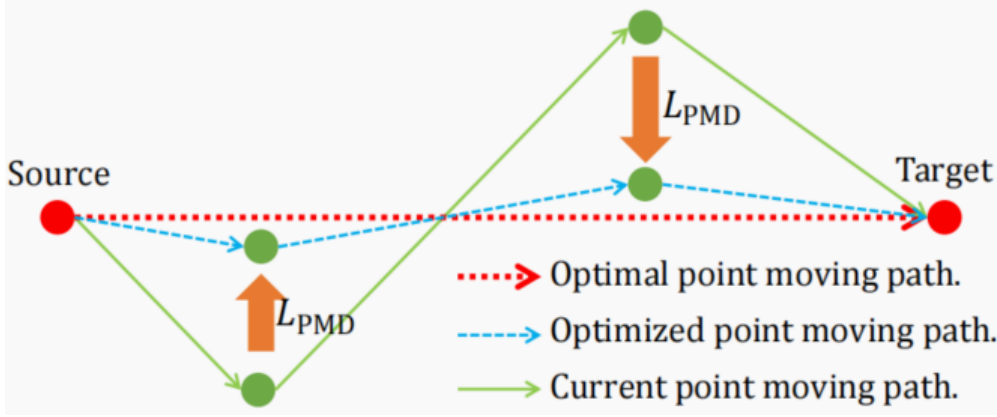


图6: \mathcal{L}_{PMD} 的有效性的描述。通过最小化点移动距离，鼓励网络从源到目标的过程中学习得到更加一致的路径，从而减少每一步的冗余搜索，提高效率

多尺度搜索半径

PMP-Net以从粗到细的方式搜索点移动路径，并且每一步都以10的幂来减少移动点的最大步进，即第 k 步位移 $\Delta \mathbf{p}_i^k$ 在公式(1)的计算机结果限制为 $10^{-k+1} \Delta \mathbf{p}_i^k$ ，这使得网络在训练期间能够更快地收敛。同时，减少的搜索范围将保证网络在下一步不会推翻其在上一步中做出的决定，特别是对于远程移动。因此，可以防止网络在路径搜索过程中做出的冗余决策。

3.4. 训练损失函数

通过倒角距离（Chamfer Distance, CD）和地动距离（Earth Mover Distance, EMD）将完整的基准点云正则化为变形后的形状。遵循等式(7)中的符号，倒角距离定义如下：

$$\mathcal{L}_{CD}(X, \hat{X}) = \sum_{x \in X} \min_{\hat{x} \in \hat{X}} \|x - \hat{x}\| + \sum_{\hat{x} \in \hat{X}} \min_{x \in X} \|\hat{x} - x\| \quad (9)$$

完整的损失函数定义如下：

$$\mathcal{L} = \sum_k \mathcal{L}_{CD}(P^k, P') + \mathcal{L}_{PMD} \quad (10)$$

其中， P^k 和 P' 分别表示第 k 步的输出点云和目标点云。请注意，找到最优 ϕ 的计算代价是非常高昂的。在实验中，我们遵循了[48]的简化算法来估计 ϕ 的逼近。

Ch04 实验

4.1. Completion3D 数据集评估

数据集

我们在广泛应用的三维点云补全基准（如：Completion3D^[34]）上评估了PMP-Net。Completion3D数据集是从ShapeNet数据集上推导出来的大规模三维对象数据集。部分三维形状是通过将部分视图中的2.5D深度图像反向投影到三维空间中生成的。我们遵循Completion3D中的配置分割训练/验证/测试集，以便与其他方法进行公平的比较。

评价指标

根据之前的研究^{[34][42][41][48]}，我们使用每个点的L1/L2倒角距离（CD）作为评价度量，其中的L1版本是等式(9)中CD的平均点数，L2版本是等式(9)中CD的L2范数（用于取代L1范数）

定量比较

PMP-Net与其他最先进的点云补全方法的比较结果（见表1）说明PMP-Net在所有类别的平均倒角距离都取得了最好的成绩。与第二好的补全方法GRNet比较平均倒角距离，PMP-Net在8个类别中有6个都取得了更好的效果，证明了PMP-Net在不同形状类别中具有更好的泛化能力。正如我们在第2节讨论的那样，GRNet^[42]是一种体素辅助形状补全方法，其中点云与三维体素之间的转换，以及3D CNN在体素数据上的应用都是一个耗时的过程。在表1中的其他方法（如：SA-Net^[41]）都是典型的完全基于点云的生成类补全方法，并且在这些方法上对PMP-Net进行非平凡的改进证明了在点云补全任务中基于变形的解的有效性。

表1：在Completion3D数据集上点云补全中每个点的L2倒角距离 $\times 10^4$ 。（越低越好）

Methods	Average	Plane	Cabinet	Car	Chair	Lamp	Couch	Table	Watercraft
FoldingNet [43]	19.07	12.83	23.01	14.88	25.69	21.79	21.31	20.71	11.51
PCN [48]	18.22	9.79	22.70	12.43	25.14	22.72	20.26	20.27	11.73
PointSetVoting [49]	18.18	6.88	21.18	15.78	22.54	18.78	28.39	19.96	11.16
AtlasNet [3]	17.77	10.36	23.40	13.40	24.16	20.24	20.82	17.52	11.62
SoftPoolNet [37]	16.15	5.81	24.53	11.35	23.63	18.54	20.34	16.89	7.14
TopNet [34]	14.25	7.32	18.77	12.88	19.82	14.60	16.29	14.89	8.82
SA-Net [41]	11.22	5.27	14.45	7.78	13.67	13.53	14.22	11.75	8.84
GRNet [42]	10.64	6.13	16.90	8.27	12.23	10.22	14.93	10.08	5.86
PMP-Net(Ours)	9.23	3.99	14.70	8.55	10.21	9.27	12.43	8.51	5.77

定性比较

在图7(a)中，我们在Completion3D数据集上将PMP-Net与其他补全方法进行了可视化地比较，从中我们可以发现PMP-Net在不同的输入类别上预测了更加准确的完整形状，而其他的方法可能在某些形状上输出一些失败的案例。例如，图7(a)中输入丢失了整个桌面的“桌子”后，GRNet与SA-Net几乎无法预测出完整的桌面，而其他方法FoldingNet^[43]、PCN^[48]和TopNet^[34]试图修复桌面，但是无法预测出良好的形状。此外，基于变形的PMP-Net相对于其他生成类方法的优势可以通过图7(a)中的案例“船”得到很好地证明。特别是与PCN和TopNet相比，这两类生成类方法虽然成功地学习到输入的“船”的整体结构，但是无法重建输入的“船”的剩余部分。另一方面，基于变形的PMP-Net可以通过在特定区域中移动少量的点来完成补全，从而保留了输入的形状，并且保证了输入形状的整体结构不变。在图7(b)中，我们可视化了PMP-Net的更多补全结果。

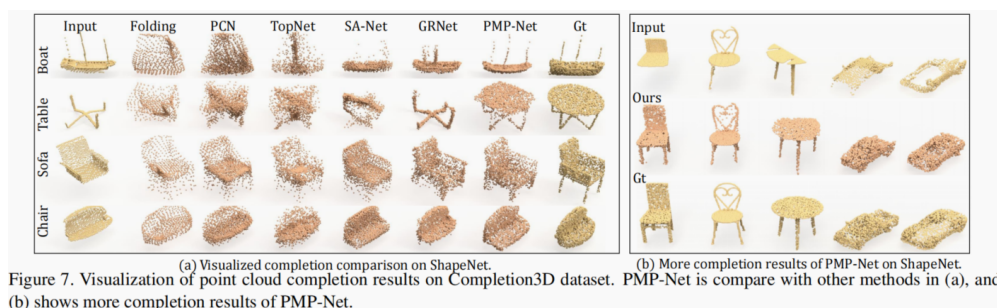


图7：在Completion3D数据集上点云补全结果的可视化。(a)PMP-Net与其他方法的比较；(b)PMP-Net的更多补全结果

对ScanNet中“椅子”的定性比较

H3 为了评估PMP-Net在点云补全任务上的泛化能力，我们在Completion3D数据集预训练了PMP-Net模型，评估了其在ScanNet中“椅子”实例上的调优前的性能，并且与GRNet进行了对比（GRNet是表1中第二好的方法，也在Completion3D数据集上进行了预训练）。比较的可视化结果（见图8）说明PMP-Net相比GRNet在补全形状上噪声更少，这得益于其点移动方法。虽然Completion3D与ScanNet在数据分布上的差异使得网络不可避免地被混淆，但是基于移动点的PMP-Net可以选择保留对象的残留部分的那些点在它们自己的位置上从而保持了更好的形状，而基于生成的GRNet则必须为对象残留部分和缺失部分预测新的点。

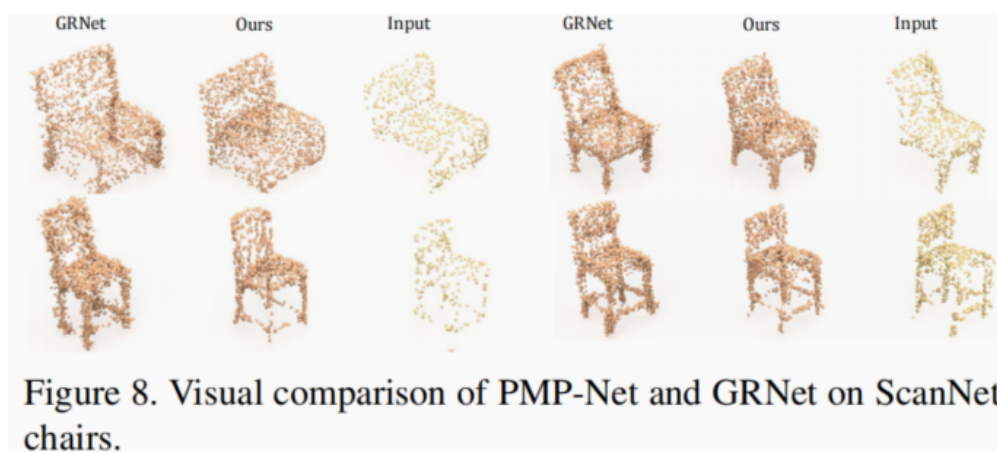


图8：PMP-Net和GRNet在SccanNet的椅子数据的可视化比较

4.2. 应用到稠密点云补全

H2 我们证明了PMP-Net在稀疏点云上的学习可以直接应用到稠密点云补全中。具体来说，我们在2048个点的稀疏形状上保持对PMP-Net的训练，并且通过在PCN数据集^[48]上预测16384个点稠密完整形状来提示其泛化能力。因为部分形状拥有不同的点数，所以我们对大于2048个点的形状下采样，对小于2048个点的形状上采样（随机拷贝），使它们都符合2048个点数要求。因为PMP-Net学习的是移动点而不是生成点，因此输入的不完整点云与目标的完整点云的点的数目必须相同。为了预测16384个点的完整形状，我们在测试阶段为每个形状的预测都重复了8次，每次只移动2048个点。请注意，PMP-Net仍然是在具有2048个点的稀疏点云上进行训练，这些点云都是从PCN数据集的稠密点云中采样出来的。

表2中的对比说明，在PCN数据集上，PMP-Net的性能与最先进的方法^[36]相比是差不多的（排名第二）。Wang等人^[36]的结果引自其原始论文，而其他方法的结果引自^[42]。请注意，在稠密点云补全上，大多数基于生成的方法（如：CDN^[36]和GRNet^[42]）设计了一个从粗到细的生成过程，从而获得更好的性

能。而PMP-Net则基于2048个点通过简单地重复点移动过程就可以得到任意数量的稠密点云，并且仍然可以获得与对应方法相当的结果。

4.3. 模型分析

在本节中，我们分析了PMP-Net中不同部分的影响。默认情况下，除了被分析的部分外，我们在所有的实验中都使用相同的网络配置。为了方便起见，所有的研究通常都是在Completion3D数据集的验证集上进行的，数据分为四类（即：飞机、汽车、椅子和桌子）。

RPA模块与PMP损失函数的分析

通过使用PMP-Net中的其他单元替换RPA模块来分析该模块的有效性，通过从PMP-Net中移除PMP损失函数来分析该函数的有效性。具体来说，我们开发了新课程不同的变体来进行比较：(1)NoPath：从网络中移除RPA模块；(2)Add：在网络中使用逐元素添加层来替换RPA模块；(3)RNN：使用RNN替换RPA模块；(4)LSTM：使用LSTM替换RPA模块；(5)GRU：使用GRU替换RPA模块；和(6)w/o PMP：从网络中移除PMP损失，在训练中仅使用倒角距离。

形状补全的结果（见表3）可以发现基线RPA模块获得了最好的性能，而最差的结果是Add变体，通过与NoPath变体相比可知，不加处理地直接使用历史信息（上一步的路径）会导致网络性能退货。RPA基线与GRU变体的比较证明了我们设计的RPA模块的有效性，它比GRU单元更加能够考虑当前步骤的信息，将会有助于网络更加精准地做出点移动的决策。

通过与w/o PMP基线变体的比较，我们发现PMP损失显著地提升了网络的性能，这与我们的观点保持一致，即点移动路径应该被正则化，以便更好地捕捉三维形状的结构和拓扑细节。

表3：RPA模块和PMP损失的分析（基线使用星号“*”标注）

Table 3. Analysis of RPA and PMP loss (baseline marked by “*”).					
Unit.	avg.	plane	chair	car	table
NoPath	11.95	3.55	8.30	16.15	19.79
Add	12.23	3.32	8.10	16.47	21.05
RNN	12.12	3.55	8.14	16.19	20.58
LSTM	11.99	3.79	8.09	15.37	20.72
GRU	11.87	3.44	7.85	15.44	20.72
w/o PMP	13.66	4.26	8.21	16.69	25.19
baseline*	11.58	3.42	7.87	15.88	19.15

多步骤路径搜索

通过表4，我们分析了不同步骤对点云变形的影响。具体来说，我们确定了每一个步长之间的搜索半径与10的比值，并且评估了PMP-Net在不同步长配置下的性能。例如：将step=4的搜索半径分别设置为 $10^0, 10^{-1}, 10^{-2}, 10^{-3}$ ，将step=2的搜索半径分别设置为 $10^0, 10^{-1}$ 。从表4中，通过比较第1步、第2步和第3步的不同结果，我们可以发现通过多步骤变形点云有效地提升了补全的性能。另一方面，通过比较第3步、第4步的不同结果，说明了多步骤路径搜索的性能也存在极限，因为过多的步骤可能会导致路径搜索中的信息冗余。

表4：不同步骤的影响（基线标注为星号“*”）

Table 4. The effect of different steps (baseline marked by “*”).					
Steps.	avg.	plane	chair	car	table
1	12.26	3.71	8.27	15.59	21.48
2	11.90	3.47	7.95	15.66	20.53
3*	11.58	3.42	7.87	15.88	19.15
4	11.67	3.39	7.91	15.89	19.48

搜索半径的分析

默认情况下，我们将每一步搜索半径减小为前一步的1/10。在表5中，我们分析了搜索半径的不同配置，并且评估了它们对PMP-Net性能的影响。此外，我们还测试了两种不同的策略来执行点移动路径搜索，即固定搜索半径策略（[1.0, 1.0, 10]）和递减缩小策略（[1.0, 0.5, 0.25]）。基线结果是PMP-Net的默认配置（[1.0, 0.1, 0.01]）。表5说明PMP-Net在固定搜索半径策略上得到最差的结果，证明了递减缩小策略的有效性。对比两个递减缩小策略可以发现比例越大，模型的性能越好，因为大比例可以避免网络撤销前一步做出的决策。我们还注意到，当下降比例变大后PMP-Net将会逼近表4中step=1的网络行为，这可能会反向影响形状补全的性能。

表5：搜索半径的影响（基线标准为星号“*”）

Table 5. The effect of searching radius (baseline marked by “*”).					
Radius.	Avg.	Plane	Chair	Car	Table
[1.0, 1.0, 1.0]	12.01	3.61	8.22	16.44	19.79
[1.0, 0.5, 0.25]	11.77	3.36	8.01	15.92	19.79
[1.0, 0.1, 0.01]*	11.58	3.42	7.87	15.88	19.15

基于不同半径时点移动路径的可视化分析

在图9中，我们基于表4中搜索路径的不同策略可视化了搜索过程。通过分析第一步的变形输出，我们可以发现拥有一个从粗到细的搜索策略时，PMP-Net可以在早期的步骤中学习预测更好的形状，在图9(a)中第一步的输出相比图9(b)和图9(c)已经非常完整且干净。此外，一个在早期的步骤中更加全面的形状预测也会使网络更聚集于点云的更好的结构细节，这个结论可以比较图9中的第3步得出，其中图9(a)中红色矩形内的区域比其他两个子图好得多。

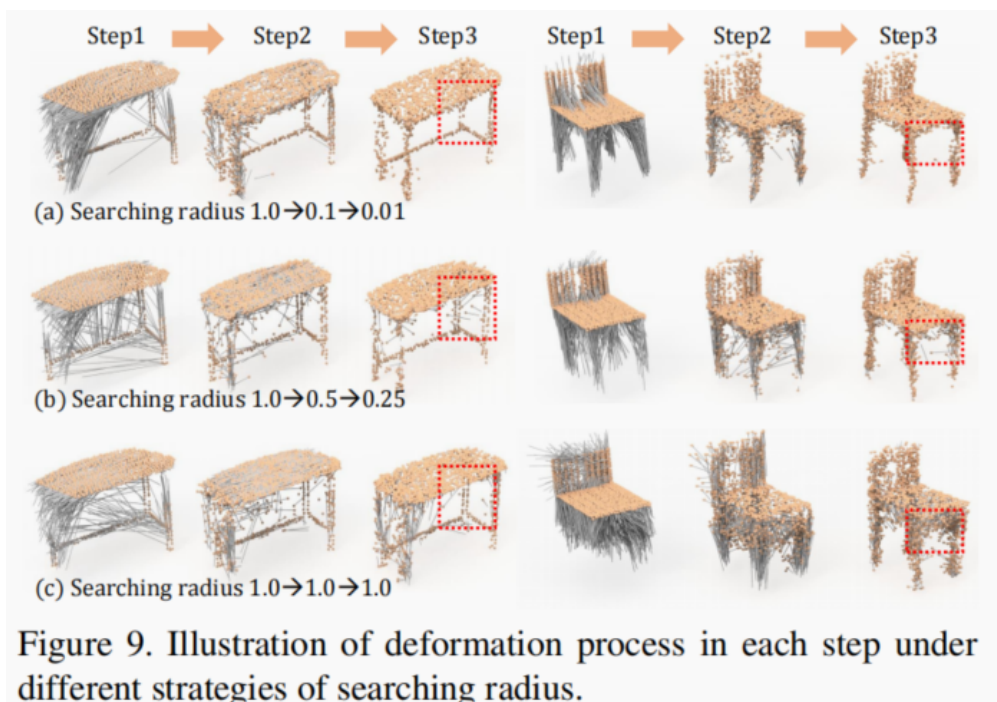


图9：基于搜索半径的不同策略每一步的变形过程展示。

Ch05 结论

H1 本文提出了一种新的用于点云补全的多步骤形状变形网络PMP-Net。通过多步骤从源点云移动点到目标点云，可以更加一致地改善预测形状的结构和拓扑细节，在不完整与完整形状建立点级别的形状对应关系。在实验中，我们与其他方法的比较展示了PMP-Net在Completion3D基准测试和PCN数据集上的优越性。

附录A 配置细节

H1 我们使用PointNet++的单尺度分组（Single Scale Grouping）版本及其特征传播模块作为PMP-Net的基础框架。每个部分的细节分别描述在表6与表7中。

在表6中，`#Points`表示下采样后点的数目；`Radius`表示查询球的半径；`#Sample`表示每个中心点采样邻近点的数目，`MLP`表示编码器的每一层的多层感知机（MLP）的输出通道的数目。

表6：编码器的结构细节

Table 6. The detailed structure of encoder.				
Level	#Points	Radius	#Sample	MLPs
1	512	0.2	32	[64, 64, 128]
2	128	0.4	32	[128, 128, 256]
3	-	-	-	[256, 512, 1024]

表7：特征传播模块的结构细节

Table 7. The detailed architecture of feature propagation module.

Level	1	2	3
MLPs	[256, 256]	[256, 128]	[128, 128, 128]

训练细节

我们使用AdamOptimizer训练PMP-Net，其初始学习率为 10^{-3} ，并且每20个周期指数级衰减0.5。训练过程使用单个的NVIDIA GTX2080Ti GPU完成，其batch_size=24。PMP-Net经过150个周期收敛到PCN和Completion3D数据集。

H2 我们将Completion3D中输入的所有训练形状乘以0.9以避免超出tanh激活函数。

附录B 更多实验

B.1. 噪声向量的维度

H1 论文的等式(1)中的噪声向量用于推动点离开其原始位置。在本节中，我们将分析噪声的维度与标准方差，从而提示它们对点的潜在影响。因为无论是噪声向量的维度还是标准方差，当它们减小到0后，网络就不会再有干扰。另一方面，向量的

H2 维度或者标准方差越大，对网络的干扰越大。在表8中，我们首先分析了噪声向量维度的影响。通过比较0维与其他维度结果的比较，我们可以得出噪声向量引起的扰动对学习点的变形具有重要意义。通过分析不同长度的噪声向量的性能，我们可以发现，与噪声向量的存在相比，噪声向量本身的长度对性能的影响相对较小。

表8：噪声向量的影响（基线标注为星号“*”）

Table 8. The effect of noise dimension (baseline marked by “*”).

Dim.	avg.	plane	chair	car	table
0	14.56	4.39	10.48	19.01	24.33
8	11.85	3.28	7.95	15.65	20.50
16	11.68	3.44	7.86	15.22	20.19
32*	11.58	3.42	7.87	15.88	19.15
64	11.58	3.14	7.96	16.01	19.17

B.2. 噪声分布的标准方差

在表9中，我们展示了PMP-Net中噪声向量在不同的标准方差下的补全结果。与向量维度的分析一样，我们可以得出结论，标准方差越大，则扰动越大，有助于网络获得更好的补全性能。当标准方差达到一定的阈值后，其变化对于网络的影响

H2 就会变弱。

表9：标准方差的影响（基线标注为星号“*”）

Table 9. The effect of standard deviation (baseline marked by “*”).

Stddev.	Avg.	Plane	Chair	Car	Table
10^{-2}	11.89	3.32	8.15	16.42	19.58
10^{-1}	11.56	3.58	7.78	15.47	19.41
1.0*	11.58	3.42	7.87	15.88	19.15
10	11.62	3.35	7.88	15.29	19.95

B.3. 多步骤搜索的可视化分析

我们在图10中可视化了不同搜索步骤配置下的点变形过程。将最上面一排的4步搜索与其他三种配置进行比较会发现其椅背上的空白究竟形状更加清晰（使用矩形突出显示），证明了多步骤搜索持续优化形状的有效性。

H2

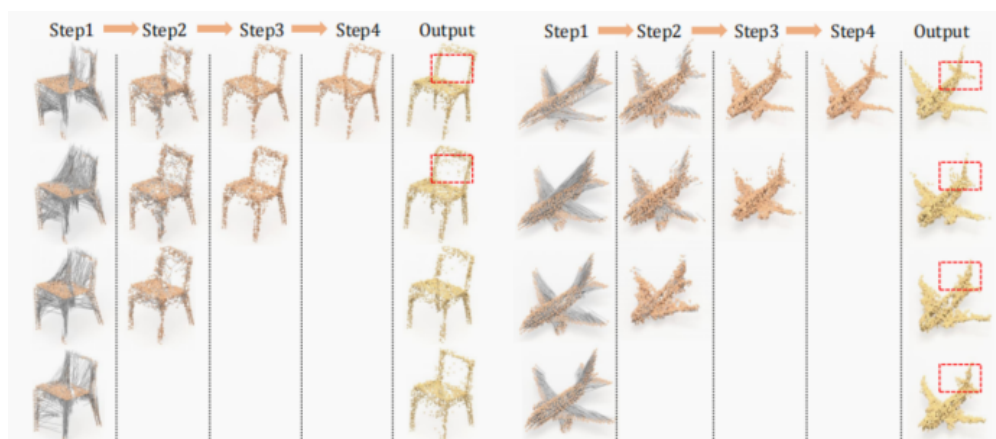


Figure 10. Illustration of multi-step searching under different searching steps. The first row is 4-step completion, and the second row is 3-step completion, and so on.

图10：基于不同搜索步骤的多步骤搜索描述。第一行为4步补全，第二行为3步补全，其他类推。

B.4. 可视化PCN数据集上的补全结果

在图11和图12中，我们补充了在PCN数据集上每个类别的形状补全结果。对于每个类别，第一行是输入的不完整形状，第二行是预测的完整形状，第三行是基准数据。

H2

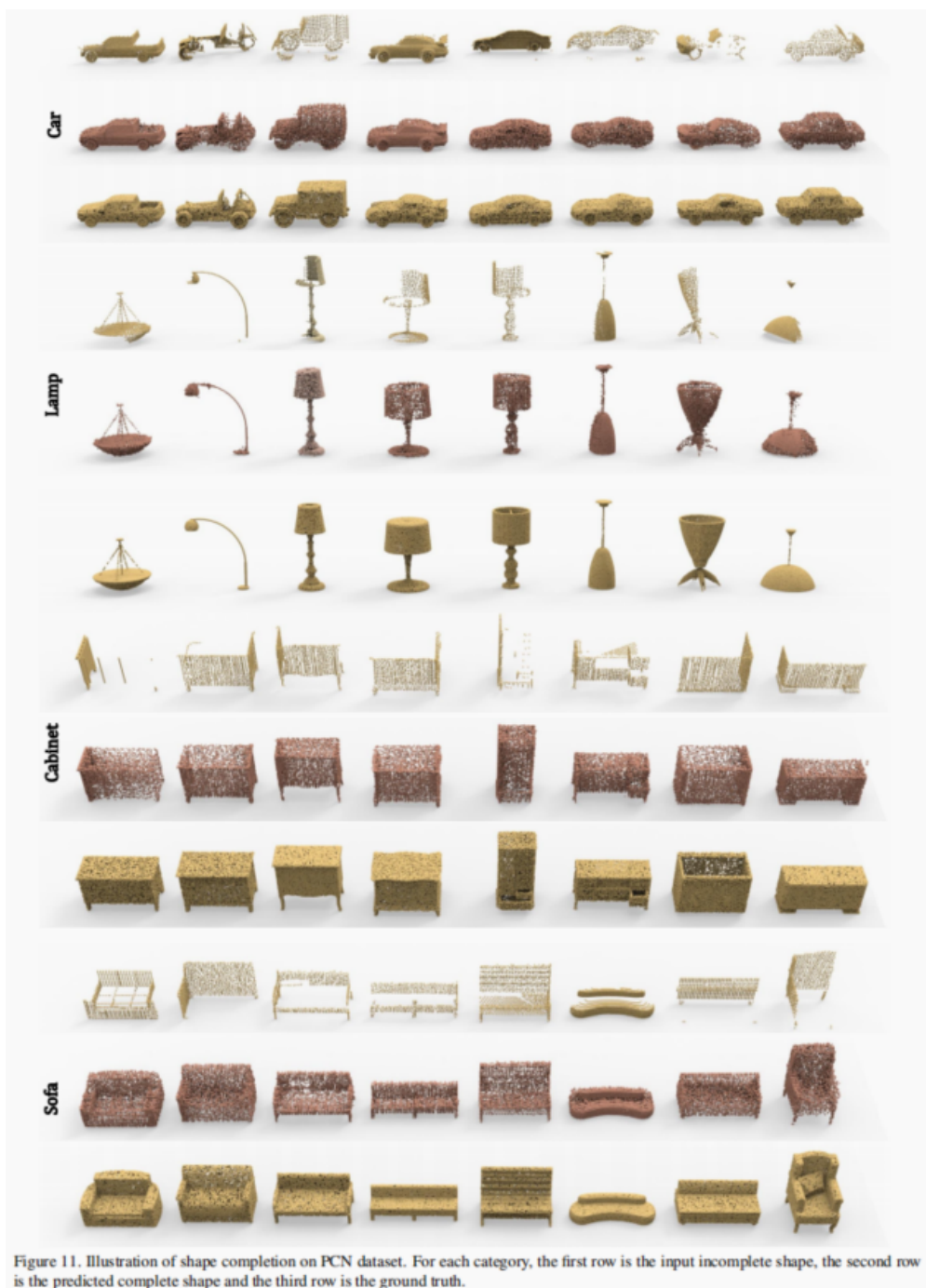


Figure 11. Illustration of shape completion on PCN dataset. For each category, the first row is the input incomplete shape, the second row is the predicted complete shape and the third row is the ground truth.

图11：PCN数据集上形状补全的描述。对于每个类别，第一行是输入的不完整形状，第二行是预测的完整形状，第三行是基准数据。



Figure 12. Illustration of shape completion on PCN dataset. For each category, the first row is the input incomplete shape, the second row is the predicted complete shape and the third row is the ground truth.

图12: PCN数据集上形状补全的描述。对于每个类别, 第一行是输入的不完整形状, 第二行是预测的完整形状, 第三行是基准数据。