

東北大學

硕士学位论文开题报告及论文工作计划书

学号：2270874

姓名：付凌坤

导师：王庆

学科类别：☐ 工学 ☒ 全日制专业学位

学科/工程领域：控制工程

所属学院：信息科学与工程学院

研究方向：计算机视觉

拟选题目：基于图像融合的
点云补全研究

选题时间：2023 年 8 月 1 日

东北大学研究生院

2023 年 10 月 14 日

填表说明

- 1、本表应在导师指导下如实填写。
- 2、学生在开题答辩前一周，将该材料交到所在学院、研究所。
- 3、按有关规定，没有完成开题报告的学生不能申请论文答辩。
- 4、全文正文均用小四号宋体，单倍行距，段前段后间距为 0，如果页数不够，可以整页扩页，其他格式要求参见《东北大学硕、博士学位论文格式》。

一、前期工作基础（本节可以整页扩页）

课程学习及选题开题阶段，在导师指导下从事研究工作总结（不少于 2000 字）

1.1 研究工作总结

点云是空间中代表 3D 形状或物体的一组数据点。它可以通过激光雷达、摄像头、三维扫描仪等设备获取。这些点在几何上代表物体的 3D 表面。点云补全任务是指将不完整的点云数据补全成完整的三维模型。然而，在实际应用中，由于各种因素的干扰，采集到的点云数据往往存在缺失、噪声等问题，这就需要对点云进行补全处理。点云补全任务可以应用于很多领域，例如机器人视觉、虚拟现实、医学图像处理等。

一直以来，点云补全任务所面临的核心问题在于如何通过部分点云获得形状完整的高质量 3D 点云，经过前期调研发现，点云补全可以分为传统方法和深度学习方法。在传统方法中，基于几何形状和模板特征匹配是两种主流思想。而经过研究，这类方法只能处理一些点云缺失率很低、结构特征十分明显的缺失点云，而难以处理复杂的形状。与传统方法相比，这些基于深度学习的方法在处理具有不规则结构和几何形状复杂的对象方面表现出了显著的优势。因此本课题拟采用基于深度学习的点云补全方法进行研究。在这之后本人还对点云不完整的原因以及所面临的挑战进行了分析，并系统性的总结了点云补全领域常用的数据集与评价指标。

1.1.1 点云不完整的原因

部分或不完整的点云是指有缺失点的点云。缺失点是指点云中由于各种原因缺失了一部分点。在数据采集过程中，三维激光扫描仪会受到被测物体特征、处理方法、环境等因素的影响，不可避免地导致点的缺失(图 1-1)。

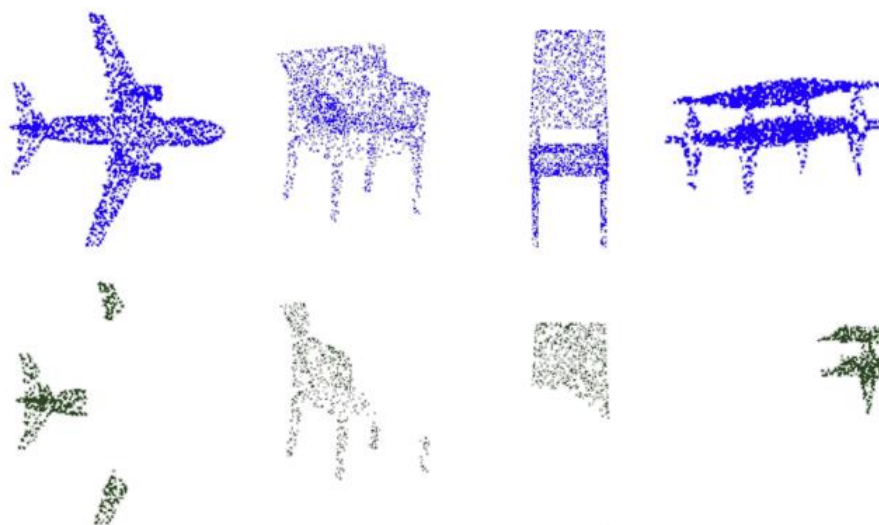


图 1-1 完整点云和缺失 70%点云示意图

主要原因可以归结为镜面反射、信号吸收、外部物体遮挡、自遮挡和盲点（图 1-2）。前两者是由于物体的表面材料，可能会吸收或反射激光雷达信号。后三种主要是由于遮挡，遮挡可以借助物体的其他部分或利用多源数据来完成。而且，三维扫描仪在扫描过程中的稳定性也对扫描质量有特殊的影响。数据采集完成后，点云还需要进行一系列处理，如点云去噪、平滑、配准、融合等。同时，这些操作会显著加剧点云的缺失。这不仅会影响数据完整性，导致拓扑错误，还会影响点云重构、三维模型重建、局部空间信息提取以及后续处理的质量。

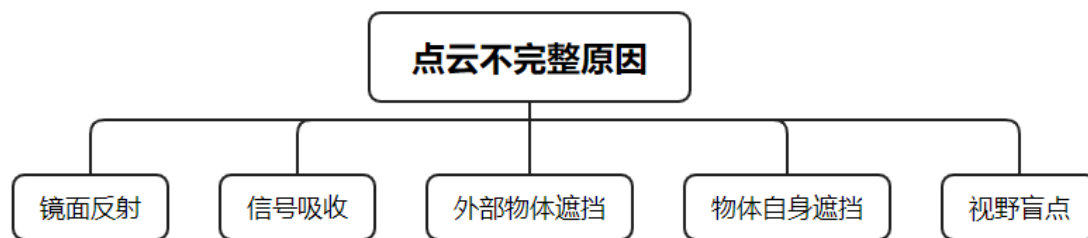


图 1-2 点云不完整的原因

1.1.2 面临的挑战

1) 结构性信息挑战

重建完整的点云具有挑战性，因为点云补全任务所需的结构信息与点云的无序和非结构化特性相矛盾。现实世界中 3D 物体点云可以被视为低层次和高层次的配置，包括表面、语义部分、几何元素等。现有的点云生成框架要么在其设计的解决方案中排除结构，要么假设并执行特定的结构或拓扑来生成 3D 物体的完整点云。因此，学习点云的结构特征对于更好地重建完整的点云非常重要。

2) 细粒度完整形状挑战

在 3D 形状补全方面，应该使用关系结构信息来重建合理的细致完整点云，例如几何对称性、规则排列和表面平滑度。虽然一些研究已经通过迭代细化、全局特征和局部特征的集成、跳过连接、残差连接等充分利用了结构信息，但还需要更多的研究来生成更为细致完整的形状。

1.1.3 数据集

关于点云补全，数据集可以分为两种类型：人造数据集和真实世界数据集。最常研究的四个数据集如下：

1) PCN：从 PCN 派生的计算机辅助设计（CAD）数据集，总共包含来自 8 个特定类别的 30974 个 3D 模型。包括在基于真实点云表面上均匀采样的 16384 个点。

2) KITTI：该数据集是通过 Velodyne 激光扫描仪收集的。最初是为了评估立体匹配的性能而设计的，其中包含 22 个立体序列的 LiDAR 点云。收集了部分车辆以评估点云补全方法在缺乏基于真值的实际扫描中的性能。

3) ModelNet40：一个包含全面的 3D CAD 模型的数据集，其中包括 40 个类别和 13356 个模型。

4) Completion3D：一个在线平台，用于基于 ShapeNet 数据集的一个子集评估形状补全方法。需要注意的是，一些成对的点云可能面临无法匹配比例的问题，这可以通过提出的单侧 CD 损失来解决。

除了上述数据集之外，最近提出了 Shapenet 34/55 和 MVP 数据集，以增加物体的种类和数量、不同的视角以及尽可能接近真实世界物体的不同程度的缺陷。

表 1-1 点云补全常用数据集

数据集	年份	分类数量	来源	分辨率
PCN	2015	8	CAD	2048/4096/8192/16384
KITTI	2012	8	RGB&LiDAR	与网络有关
ModelNet40	2015	40	CAD	2048/16384
Completion3D	2019	8	CAD	1024/2048/16384

1.1.4 评价指标

对于点云补全, *Chamfer Distance*(CD)、*Earth Mover's Distance*(EMD)和*F - Score*是最常用的性能评估标准。 CD 试图找到两组点之间的最小距离, EMD 评估点云的重建质量。 $F - Score$ 评估物体表面之间距离。

1) *Chamfer Distance*(CD)

$$CD(S_1, S_2) = \frac{1}{|S_1|} \sum_{x \in S_1} \min_{y \in S_2} \|x - y\|_2 + \frac{1}{|S_2|} \sum_{y \in S_2} \min_{x \in S_1} \|y - x\|_2$$

其中, CD 代表输出点云 S_1 中的点到完整点云 S_2 中的点的平均最近距离之和, 以及 S_2 中的点到 S_1 中的点的平均最近距离之和。 CD 有两个变体: $CD - T(CD - \ell_1)$ 和 $CD - P(CD - \ell_2)$ 。两个点云 S_1 和 S_2 之间 $CD - T$ 和 $CD - P$ 的定义如下:

$$\begin{aligned} \mathcal{L}_{S_1, S_2} &= \frac{1}{|S_1|} \sum_{x \in S_1} \min_{y \in S_2} \|y - x\|_2^2 \\ \mathcal{L}_{S_2, S_1} &= \frac{1}{|S_2|} \sum_{y \in S_2} \min_{x \in S_1} \|x - y\|_2^2 \\ \mathcal{L}_{CD-T}(S_1, S_2) &= \mathcal{L}_{S_1, S_2} + \mathcal{L}_{S_2, S_1} \\ \mathcal{L}_{CD-P}(S_1, S_2) &= \frac{(\sqrt{\mathcal{L}_{S_1, S_2}} + \sqrt{\mathcal{L}_{S_2, S_1}})}{2} \end{aligned}$$

2) *Earth Mover's Distance*(EMD)

EMD 的目标是找到一个双射函数 $\phi: S_1 \rightarrow S_2$, 以最小化部分点和完整点之间对应点的平均距离。与 CD 不同的是, S_1 和 S_2 的大小需要相同。

$$EMD(S_1, S_2) = \min_{\phi: S_1 \rightarrow S_2} \frac{1}{|S_1|} \sum_{x \in S_1} \|x - \phi(x)\|_2$$

3) *F - Score*

$F - Score$ 被视为精确率和召回率的调和平均值。精确率计算在一定距离范围内重建点的百分比, 代表重建的准确性。另一方面, 召回率计算在一定距离范围内基于真实值上的点的百分比, 代表重建的完整性。距离阈值 d 可以用来控制 $F - Score$ 的严格程度。 $F - Score$ 可以评估正确重建的点或表面面积的百分比, 可以定义如下:

$$F - Score(d) = \frac{2P(d)R(d)}{P(d) + R(d)}$$

其中, $P(d)$ 和 $R(d)$ 分别表示距离阈值 d 的精确度和召回率。

$$P(d) = \frac{1}{|S_1|} \sum_{r \in S_1} \left[\min_{t \in S_1} \|t - r\| < d \right]$$

$$R(d) = \frac{1}{|S_2|} \sum_{t \in S_2} \left[\min_{r \in S_1} \|t - r\| < d \right]$$

其中 S_1 是正在评估的重建点集, S_2 是真实值。 $|S_1|$ 和 $|S_2|$ 分别是 S_1 和 S_2 中点的数量。

二、选题依据（本节可以整页扩页）

课题背景、选题依据、课题研究目的、理论意义和应用价值（工学硕士）/工程背景和实用价值（专业学位硕士）（不少于 1000 字）

2.1 课题背景

随着激光雷达、激光或 RGB-D 扫描仪等 3D 扫描设备的普及，点云变得更容易获取，目前在机器人、自动驾驶、3D 建模和制造等领域引发了大量研究。然而，由于遮挡、反射、透明度以及设备分辨率和角度的限制，这些设备直接采集的原始点云主要是稀疏的、不完整的。这限制了点云数据的应用，如 3D 场景识别和理解工作、无人驾驶、机器人视觉等。因此，点云补全是一项非常必要的工作，它会利用先验知识从不完整的点云补全已完成的点云。从部分观测中生成完整的点云对于推动下游应用至关重要。

点云补全早期尝试通过体素定位和 3D 卷积将 2D 补全任务中成熟的方法转移到 3D 点云。然而，由于空间分辨率的提高，这些方法的计算成本很高。随着 PointNet 和 PointNet++ 的巨大成功，直接处理三维坐标已成为基于点云的三维分析的主流。近年来，许多其他方法如基于点的、基于卷积的、基于折叠的、基于视图的、基于生成模型的和基于变换器的方法如雨后春笋般涌现，并取得了显著成果（图 2-1）。

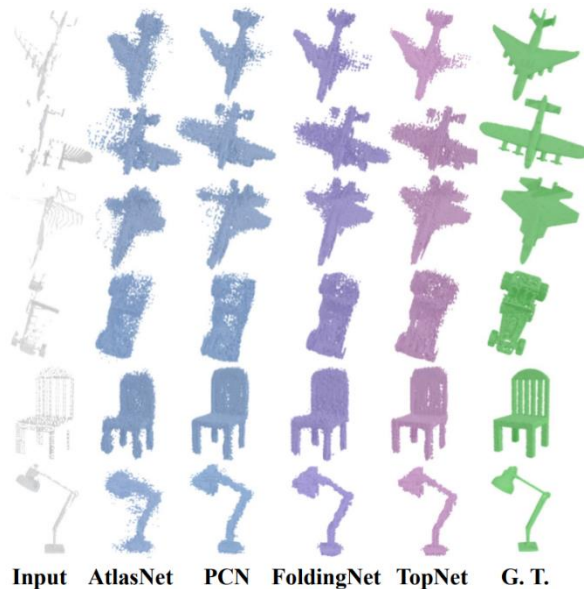


图 2-1 常用点云补全算法结果示意图

2.2 选题依据

目前，点云补全可以分为传统方法和深度学习方法。在传统方法中，基于几何先验规则和模板特征匹配是两种主流思想。与传统方法相比，这些基于深度学习的方法在处理具有不规则结构和几何形状的对象方面表现出了显著的优势。深度学习方法可以分为单模态和多模态点云补全方法。而单模态的补全方法由于缺少额外信息，在点云特征提取方面仍存在不足。因此，采用图像引导的方法来改进提升点云特征的提取。由于点云的无序和非结构化性质，无法直接使用纹理属性提取图像等特征。而在点云补全中使用图像信息来表示点云信息可以提高补全精度。

2.3 课题研究目的

为了更好地将点云应用于三维重建、机器人视觉、自动驾驶、虚拟现实等领域。通过点云补全技术，可以使得不完整的点云数据变得更加完整，从而提高三维重建的精度和效率；也可以为机器人视觉和自动驾驶等领域提供更加准确的环境感知和障碍物检测；同时，

还能为智能制造领域提供完整的高精度点云模型以提高工业生产效率。此外，点云补全还可以为虚拟现实技术提供更加真实的三维场景。因此，点云补全的研究目的是为了提高现有技术的性能和应用范围，为各个领域的发展提供更好的支持。

2.4 工程背景和实用价值

2.4.1 工程背景

随着机械化制造和数字化制造的发展，机器人已由点对点应用发展至离线编程轨迹应用。针对各工作场景，对机器人进行离线编程轨迹是如今各大主流机器人厂商的优势区域。此类编程方式在针对刚性对象的加工中完成度较好。但对压铸件、橡胶件、超大型工件等变形量无法有效控制的加工对象进行加工时，离线编程轨迹的方法将不再适用。针对此类痛点，使用 3D 视觉传感器进行实物 3D 建模，完成机器人在线轨迹生成是重要的解决办法，然而此类技术在当前各大机器人厂商与集成商中尚未得到充分的研究。其主要原因为基于 3D 视觉的机器人在线轨迹生成方法存在多学科交叉，需突破多个关键技术，如针对复杂干扰情况的有效成像技术、高精度的成像系统标定、高精度的点云拼接、高精度的点云补全、高精度手眼标定等，需要掌握机器视觉、机器人控制等多个学科的专业知识。因此，当前基于 3D 机器视觉的机器人工艺轨迹在线生成技术研究尚未完善且该技术存在大量蓝海场景，具有较高的研究价值。

由于待测的工件样式多变，具有不同的形状、材质、颜色和反射性质等，如何合理的进行视觉传感器选型与设计是首先需要解决的问题；此外由于单位位置的视觉传感器视场有限，为获得工件的完整数字模型，通常需要在多个位置采集工件的点云子块再进行拼接。但是获取到的单位位置的点云可能存在残缺或空洞（图 2-2），如何根据当前已有点云特征自适应的完成点云补全同样需要进行研究，点云补全结果的好坏直接影响着工业生产的效率以及工业产品质量的好坏。



图 2-2 工业环境下实际采样到的手机边框点云，两端存在残缺，中间存在空洞

2.4.2 实用价值

1) 工业制造：针对正在研究的 3D 视觉工业平台，点云补全可以帮助工业制造业在生产过程中获取更准确、更完整的产品三维模型，提高生产效率和产品质量。

2) 建筑设计：点云补全可以在建筑设计过程中生成完整的三维建筑模型，帮助建筑师更好地理解设计和规划。

3) 3D 打印：点云补全可以生成完整的三维模型，从而帮助 3D 打印行业更准确地打印出产品。

4) 虚拟现实和增强现实：点云补全可以生成完整的三维模型，从而帮助虚拟现实和增强现实应用更加真实和逼真。

5) 医学影像：点云补全可以帮助医学影像学家更好地理解患者的病情和病变部位，从而提供更准确的诊断和治疗方案。

6) 汽车和航空工业：点云补全可用于汽车和航空工业中的 CAD 建模和虚拟样机。

7) 机器人导航：点云补全可用于机器人导航，帮助机器人识别环境并规划路径。

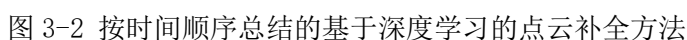
8) 3D 目标检测：点云补全可以通过对点云数据进行插值、重建等操作，将缺失部分进行填补，得到完整的点云数据。这样一来，就可以更准确地进行 3D 目标检测，提高目标检测的精度和可靠性。

综上所述，点云补全任务的实用价值在于为下游任务提供更加详细和准确的数据输入，通过进一步研究和发展点云补全，可以推动这些领域的创新和发展。

国内外研究现状、发展动态描述（不少于 1000 字）；所阅文献的查阅范围及手段，附参考文献（不少于 10 篇，其中近 3 年文献不少于 5 篇，英文文献不少于 3 篇，全部按照标准格式列出，并在文中顺序标注）

```
graph TD; A[点云补全] --> B[基于点的方法]; A --> C[基于视图的方法]; A --> D[基于卷积的方法]; A --> E[基于折叠的方法]; A --> F[基于Transformer的方法]; A --> G[基于生成式模型的方法]; A --> H[其他方法]; G --> I[基于GAN的方法]; G --> J[基于VAE的方法];
```

如图 3-1 所示, 基于点云补全和生成所采用的网络结构, 现有的架构可分为基于点、基于视图、基于卷积、基于图形、基于变换器、基于生成模型和其他方法。所有的方法总结在图 3-2 中。大多数工作都是混合方法, 根据它们所提出的亮点, 它们可能属于几种方法。在这里我们主要研究点云补全的开创性工作以及与视图相关的点云补全方法。



基于点的方法通常使用多层感知器 (MLP) 独立地对每个点进行建模。由于点云的变换不变性, 全局特征通过对称函数 (如最大池化) 进行聚合。然而, 整个点云中的几何信息和相关性仍未完全考虑。作为处理特征的常用方法, 本节仅回顾主要使用基于点的网络的方法。最初由 PointNet [1] 开创, 一些作品使用 MLP 处理和恢复点云, 因为它具有简明和非可忽略的表示能力 [2], [3]–[5]。PointNet++ [6] 和 TopNet [7] 融入了分层结构以考虑几何信息。PointNet++ 提出了两个集合抽象层, 智能地聚合多级信息, 而 TopNet 则提出了一种新的解码器, 可以生成没有假定任何特定结构或拓扑的结构化点云。受 PointNet 和 PointNet++ 的启发, Yu 等人 [8] 提出了 PU-Net, 通过基于亚像素卷积层的特征缩放学习多尺度特征。缩放恢复方法对提取的特征进行 1×1 内核的卷积。然后, 提取的特征被分解并重构成一组上采样点的簇。联合损失函数用于均匀地分布生成的点云在潜在表面上。然而, PU-Net 主要设计用于从稀疏点云生成单个更密集的点云, 而不是执行点云补全。它无法填补大的孔和缺失的部分,

也无法在点云的重采样部分添加有意义的点。为了减轻 MLP 带来的结构损失，提出了 AtlasNet[9]和 MSN[10]，通过评估一组参数化表面元素重构完整输出，从而可以生成完整的点云。具体而言，AtlasNet[9]采用了 2D 点的附加输入，并将其应用于在表面上产生单个点。因此，输出是平面的连续图像。可以多次重复这种方法，以从众多表面元素的组合中重构出 3D 形状。为了避免结构损失，MSN [10]引入了基于变形的解码器，可以将单位正方形变形为一组聚合到粗点云中的表面元素。PCN 驱动方法首次，Hebert 等人[11]提出了一种基于学习的形状补全方法，Point Completion Network (PCN)。与现有方法不同，PCN 直接处理原始点云，不需要任何关于基础形状（如对称性）或注释的结构假设（如语义类）。它具有允许生成细粒度补全并保持少量参数的解码器设计。通过与 PCN 和点卷积相结合，Xu 等人[12]设计了 FinerPCN，以局部信息为考虑因素，以粗到细的方式生成完整和精细的点云。之后，Zhang 等人[13]提出了骨架桥接点补全网络（SK-PCN）。SK-PCN 具有预测以学习全局信息的 3D 骨架。然后，使用骨架点的位移补全表面。在 MSPCN 中，Xiao 等人[14]使用一对上采样模块重构细粒度输出，并监督每个阶段以生成具有更多信息和有益中间输出的输出。此外，他们提出了一种通过将所选点与最大池化和体积下采样点相结合来识别关键集（MVCS）的方法。此 MVCS 可以考虑重要特征和整体形状。通过将部分点云视为同一分布的不同样本，设计了基于结构检索的点云补全网络（SRPCN）[15]。SRPCN 使用 k-means 聚类来提取结构点并将它们分散到分布中，然后使用 KL 散度作为度量来在数据库中找到最佳匹配输入的完整结构点云。为了获得可以捕捉全局结构和局部几何细节的良好特征表示，提出了 FSNet [16]，通过从局部区域学习多个潜在模式，适应地将点特征聚合成 2D 结构特征图。然后，将 FSNet 集成到粗到细的管道中以进行点云补全。

3.1.2 基于视图的点云补全方法

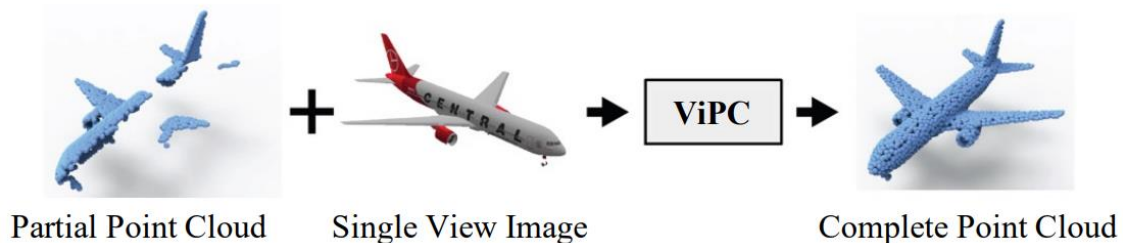


图 3-3 ViPC: 使用来自其他单视图图像的补充信息补全部分点云的方法

通过图像模式的优点，解决点云完整性的关键挑战是有效地整合由不完整的姿态和区域细节以及单视图图像的全局形状信息带来的特征（图 3-3）。作为一个传感器融合网络，Zhang 等人[17]提出了 ViPC，这是一种视图引导的架构。ViPC 从额外的单视图图像中检索缺失的全局结构信息。ViPC 的主要贡献在于“动态偏移预测器”，它可以改进粗糙的输出。Zwicker 等人[18]提出了一种多视角一致的推理方法，以加强基于视图的三维形状补全中的几何一致性。并且定义了一种多视角一致性损失算法，用于推理优化，可以在没有基于真实监督的情况下实现。此外，ME-PCN[19]利用深度扫描使网络对形状边界敏感，从而使 ME-PCN 能够恢复细粒度的表面细节，同时保持一致的局部拓扑。为了在多个局部观察派生自同一对象的帮助下估计三维规范形状的 6-DoF 姿态，Gu 等人[5]提出了一种弱监督方法来解决这个问题。在训练过程中，网络使用多视角几何约束来共同优化规范形状和姿态，可以在单个局部输入的条件下推导出完整的结果。多视角补全网络（MVCN）[20]结合了 GAN 和多视角信息以增强点云补全的性能。其主要特点和缺点如下：

- 与其他方法不同，基于视图的方法的输入是图像，可以是 RGB-D 图像或深度图像。
- 由于可以从这些图像中获取不同的信息，因此性能在很大程度上取决于视角和视图数量。

3.1.3 基于 GAN 的方法

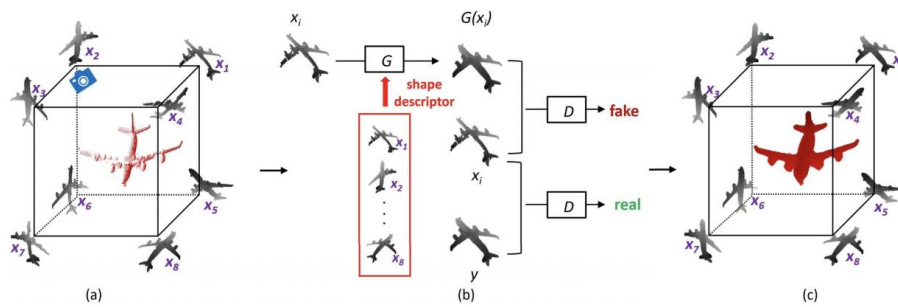


图 3-4 生成式多视图表示的图示

多视图 GAN：同一 3D 模型的不同视角共享一些可以探索的公共信息，包括从不同角度看到的全局和区域信息。Zwicker 等人[20]提出了一个多视图补全网络（MVCN）（图 3-4），它利用来自 3D 形状所有视图的信息来辅助补全每个单独的视图。受益于多视图呈现和具有条件 GAN 的网络架构，MVCN 增强了 3D 补全的性能。Liu 等人[21]试图将三维点云生成问题转化为二维几何图像生成问题，并引入了对抗 VAE 来优化通过将对抗学习与 VAE 相结合提出的 GIG。虽然很容易独立地创建 3D 形状的深度图，但有两个缺点。首先，它们不鼓励来自同一 3D 对象的深度图之间的一致性，影响通过反投影完成的深度图获得的 3D 对象的准确性。其次，它们无法补全深度图并使用来自同一 3D 对象的其他深度图的信息。补全单个深度图的准确性有限。

3.2 所查阅文献的范围及手段

通过企业及企业实验室的局域网，访问如 CVPR、ICCV、IROS、ICRA 等高质量期刊会议官网；通过百度学术查找本领域的高被引论文；或者登陆校内图书馆网站，下载国内外大量的电子版文献进行集中阅读。在查阅论文时使用的具体办法是，首先确定几个关键词搜索相关领域论文，然后在知网、百度学术以及顶会期刊上找到优质的综述性文章，根据不同的细分方向和研究难点，选出代表性论文进行阅读，再针对其中遇到的特别的算法和技术问题，细化关键词进行进一步查阅。所查阅文献的范围如表 3-1 所示：

表 3-1 文献查阅范围

数据库名称	时间跨度
中国学术期刊全文数据库	2000-2023
中国优秀博硕士论文数据库	2000-2023
中国学术会议论文数据库	2000-2023
中国学位论文数据库	2000-2023
外文检索数据库	2000-2023
SCI（科学引文索引）	2000-2023
Ei Village（工程索引网络版）	2000-2023
Google 学术	2017-2023
百度学术	2017-2023

参考文献：

[1] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, “Pointnet: Deep learning on point sets for 3d classification and

- segmentation,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 652–660.
- [2] X. Yu, Y. Rao, Z. Wang, Z. Liu, J. Lu, and J. Zhou, “Pointnet: Diverse point cloud completion with geometry-aware transformers,” in Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 12 498–12 507.
- [3] R. Li, X. Li, C.-W. Fu, D. Cohen-Or, and P.-A. Heng, “Pu-gan: a point cloud upsampling adversarial network,” in Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 7203–7212. [42] X. Wang, M. H. Ang Jr, and G. H. Lee, “Cascaded refinement network for point cloud completion,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 790–799.
- [4] X. Wang, M. H. Ang Jr, and G. H. Lee, “Cascaded refinement network for point cloud completion,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 790–799.
- [5] J. Gu, W. Ma, S. Manivasagam, W. Zeng, Z. Wang, Y. Xiong, H. Su, and R. Urtasun, “Weakly-supervised 3d shape completion in the wild,” in Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part V 16. Springer, 2020, pp. 283–299.
- [6] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, “Pointnet++: Deep hierarchical feature learning on point sets in a metric space,” arXiv preprint arXiv:1706.02413, 2017.
- [7] L. P. Tchammi, V. Kosaraju, H. Rezatofighi, I. Reid, and S. Savarese, “Topnet: Structural point cloud decoder,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 383–392.
- [8] L. Yu, X. Li, C. Fu, D. Cohen-Or, and P. A. Heng, “Pu-net: Point cloud upsampling network,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 2790–2799.
- [9] T. Groueix, M. Fisher, V. G. Kim, B. C. Russell, and M. Aubry, “A papier-mâché approach to learning 3d surface generation,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 216–224.
- [10] M. Liu, L. Sheng, S. Yang, J. Shao, and S. Hu, “Morphing and sampling network for dense point cloud completion,” in Proceedings of the AAAI conference on artificial intelligence, vol. 34, no. 07, 2020, pp. 11 596–11 603.
- [11] W. Yuan, T. Khot, D. Held, C. Mertz, and M. Hebert, “Pcn: Point completion network,” in 2018 International Conference on 3D Vision (3DV). IEEE, 2018, pp. 728–737.
- [12] Y. Chang, C. Jung, and Y. Xu, “Finerpcn: High fidelity point cloud completion network using pointwise convolution,” Neurocomputing, vol. 460, pp. 266–276, 2021.
- [13] Y. Nie, Y. Lin, X. Han, S. Guo, J. Chang, S. Cui, and J. Zhang, “Skeleton-bridged point completion: From global inference to local adjustment,” arXiv preprint arXiv:2010.07428, 2020.
- [14] W. Zhang, C. Long, Q. Yan, A. L. Chow, and C. Xiao, “Multi-stage point completion network with critical set supervision,” Computer Aided Geometric Design, vol. 82, p. 101925, 2020.
- [15] K. Zhang, X. Yang, Y. Wu, and C. Jin, “Srpcn: Structure retrieval based point completion network,” arXiv preprint arXiv:2202.02669, 2022.
- [16] Z. Su, H. Huang, C. Ma, H. Huang, and R. Hu, “Point cloud completion on structured feature map with feedback network,” arXiv preprint arXiv:2202.08583, 2022.
- [17] X. Zhang, Y. Feng, S. Li, C. Zou, H. Wan, X. Zhao, Y. Guo, and Y. Gao, “View-guided point cloud completion,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 15 890–15 899.
- [18] T. Hu, Z. Han, and M. Zwicker, “3d shape completion with multiview consistent inference,” in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, no. 07, 2020, pp. 10 997–11 004.
- [19] B. Gong, Y. Nie, Y. Lin, X. Han, and Y. Yu, “Me-pcn: Point completion conditioned on mask emptiness,” in Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 12 488–12 497.

[20] T. Hu, Z. Han, A. Shrivastava, and M. Zwicker, “Render4completion: Synthesizing multi-view depth maps for 3d shape completion,” in Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, 2019, pp. 0–0.

[21] L. Wang, Y. Huang, P. Tao, Y. Hou, and Y. Liu, “Learning geometryimage representation for 3d point cloud generation,” arXiv preprint arXiv:2011.14289, 2020.

四、研究内容（本节可以整页扩页）

1. 研究构想与思路、主要研究内容及拟解决的关键问题（不少于 1000 字）

4.1 研究构想与思路

本课题针对基于视图的点云补全展开研究，主要有三个阶段。首先通过阅读国内外相关的文献，为算法的研究打下基础。算法的研究方向分为图像的全局和局部特征提取、部分点云的特征提取、图像与点云的模态融合等研究。

在对这三部分算法进行基本的研究过后，针对其中具体存在的问题提出解决办法，然后对算法分别进行实验，对实验结果进行分析，并找到问题，通过查阅资料、阅读文献对算法进行进一步的改进，直到达到预期的指标。在算法测试完成后，将三个部分进行融合，并重复之前的实验再进行改进，最终对整个课题内容进行总结。

4.2 主要研究内容

1) 基于深度学习的图像全局和局部特征提取

为了解决图像模态方面的问题，本课题拟采用基于卷积的神经网络的方法。基于卷积的神经网络能够有效地提取图像的全局特征和局部特征，不仅要关注图像的局部特征，同时也需要关注图像局部特征与局部特征之间的联系。通过使用基于卷积的神经网络，能够提取丰富的图像信息，同时考虑网络的抗噪能力和不同物体的泛化能力。获取准确的图像特征，将为后续的研究提供更加可靠的数据支持。

2) 点云数据的特征提取

由于现实中采集的点云数据往往存在一定的噪声和缺失，因此需要对缺失点云进行处理，使其能够被用于特征提取。针对部分缺失点云的特点，需要设计特征提取算法，以提高特征的鲁棒性和准确性。获取的特征包括但不限于点云的姿态和形状等信息。

3) 图像和点云融合网络的设计

获取到图像和点云数据后，并对数据进行预处理，以便进行后续的融合处理。在对图像数据进行特征提取后，还需要对点云数据进行特征提取，还需要对它们进行特征匹配，以便将两种模态的数据对齐。为了对图像和点云特征进行融合，本课题拟采用基于图像和点云多模态融合的方式进行研究，而模态融合网络的好坏直接决定了最终点云补全效果的好坏。对于融合完成后的点云，还需要考虑其边界情况，对边界部分进行细化，以进一步提高补全质量。最后需要对融合结果进行评估，以验证多模态融合网络的有效性和准确性。

4.3 拟解决的关键问题

1) 基于卷积的图像特征提取

将图像应用到点云补全任务中可以提供更好的视觉效果，因为图像数据通常更加容易获取和处理并且可以提供更多的形状和纹理信息，使得补全后的点云更加精准。并且由于图像数据通常比点云数据小得多，因此使用图像数据可以加快点云补全任务的计算速度。然而，将图像特征与局部点云的特征进行融合仍然是一个具有挑战性的问题。因为图像可能会存在噪声干扰，所以在提取特征的同时对网络的抗干扰能力进行提升也是一项非常重要的任务。

2) 对部分缺失点云特征提取

部分缺失点云的特征提取对于点云的补全准确性是非常重要的。通过对部分缺失点云进行特征提取，可以提取出点云的局部几何特征、法向量、曲率等信息，从而更加准确地重建点云模型。同时，在点云补全任务中，需要处理大量的点云数据。通过对部分缺失点云进行特征提取，可以减少数据处理的时间和计算量，从而提高数据处理的效率。因此对部分缺失点

云特征提取完备与否对补全效果好坏也存在较大影响。

3) 图像和点云特征的融合

图像和点云都包含了不同的信息。图像可以提供颜色、纹理、表面形状等信息，而点云则提供了三维空间中物体的几何形状和位置信息。将这两种信息融合起来可以更好地描述场景和物体的特征，从而提高点云补全的准确性和鲁棒性。

其次，图像和点云特征融合可以帮助更好地处理遮挡和缺失的情况。在某些情况下，点云可能存在缺失或遮挡的情况，这会影响点云补全的效果。通过利用图像信息，算法可以更好地推断被遮挡或缺失的区域，从而更准确地进行点云补全。图像可以提供更为准确的物体边界信息，从而帮助算法更快地确定物体的形状和位置，同时还可以对点云细节进行优化。这可以大大提高点云补全的效率，从而更快地完成任务。

综上所述，如何将图像和点云特征进行高效的融合，也是需要解决的关键问题之一。

2. 拟采取的研究方法、技术路线、实施方案及可行性分析

4.4 拟采用的研究方法

本课题的研究方法主要是基于图像融合的点云补全的研究，通过实验分析来不断调整和改进。通过阅读国内外有关点云补全方向的文献，对其中的优缺点和整个发展脉络进行梳理，选择并构造适合本课题的框架和模型。本课题采用的研究流程如下：

研究流程具体如下：前期调研 → 阅读有关文献（包括图像特征提取、点云特征提取等）→ 理论分析 → 提出相关科学问题 → 网络框架设计 → 功能模块设计 → 关键问题攻克 → 实验结果验证 → 数据整理、论文撰写。本文主要研究方法如下：

1) 理论研究部分：首先，对现有的基于视图的点云补全进行综述调研；其次，筛选出与研究方向相近的文章进行学习，了解网络的结构、适用场景和优缺点等；最后，找到适合自己研究的框架，并能够基于此框架进行改进和创新。

2) 算法研究部分：图像和点云特征的融合是研究的重点，它直接关系到最终生成点云质量的好坏；图像的特征提取和点云的特征提取的研究也十分关键，他们也影响了网络的整体性能。因此，相关算法的研究至关重要。

3) 实验部分：框架建立好之后，需要通过和 SOTA 的点云补全方法比较，分析本研究的不足，并不断调整和改进。

4.5 技术路线

针对基于图像的点云补全任务，结合两方面进行考虑，分别是图像的特征提取和点云特征提取、配准，将他们进行高效地融合，完成网络的搭建后根据实验结果对其进行调整，最后根据所获得的数据进行结果的分析 and 对比，得出最终结论。本课题研究方案采取的技术路线如图 4-1 所示：

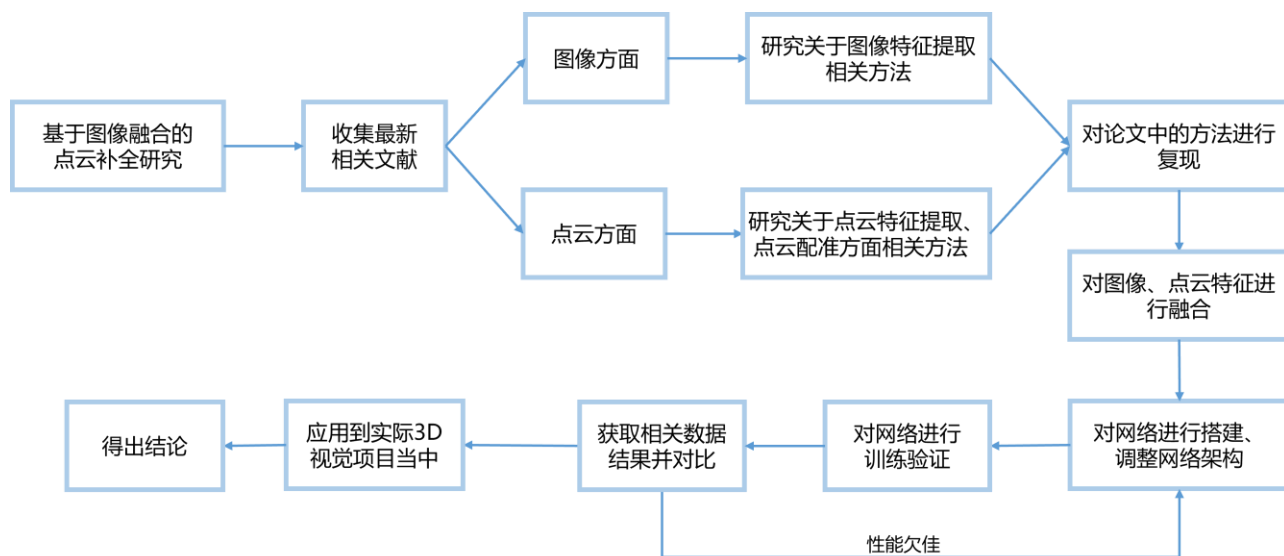


图 4-1 技术路线图

4.6 实施方案

1) 图像特征提取方案

由于提供了图像模态的信息，解决点云的完整性的关键是有效地提取图像中带来的特征。目前的一些视图引导的架构从额外的单视图图像中检索缺失的全局结构信息，再融入到缺失点云中，最后再进行细化，这一方法的弊端是忽略了图像中所蕴含的局部特征信息，这样会导致最终补全的点云会出现边缘不平滑的问题。

为了解决这个问题，在图像特征提取阶段，本课题拟采用基于卷积神经网络的特征提取方式(图 4-2)。基于卷积神经网络的图像特征提取是一种广泛应用于计算机视觉领域的技术。它的主要思想是通过卷积层、池化层和全连接层等一系列神经网络层来逐层提取图像的特征信息。

卷积层通过对输入图像进行卷积操作，提取图像的局部特征。通过滑动滤波器的方式，卷积层可以提取出图像的不同特征，比如边缘、纹理等。池化层通过对输入特征图进行下采样。减少特征图的大小，从而降低计算复杂度，同时也可以防止过拟合。全连接层是将卷积层和池化层中提取的特征信息进行整合的层。它将所有的特征图展开成一个向量，并将向量输入到一个全连接的神经网络中进行分类或回归等任务。

在整个卷积神经网络中，每个卷积层、池化层和全连接层都可以看作是一个特征提取器，通过逐层提取图像的特征信息，最终得到一个高维的特征向量。基于卷积神经网络能较好地提取到图像中蕴含的丰富的全局和局部特征，为后续的融合打下良好的基础。

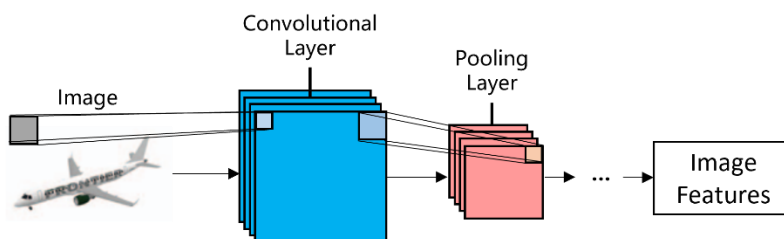


图 4-2 图像特征提取示意图

2) 点云特征提取方案

PointNet 和 PointNet++网络的提出为点云特征提取指明了方向。PointNet 是一种用于点云分类、分割和语义分割任务的神经网络模型。其核心思想是将每个点的特征向量进行最大池化，得到整个点云的全局特征向量，然后通过全连接层进行分类或分割。然而，PointNet 存在一个问题，即无法处理具有复杂结构和局部特征的点云。这对于点云补全任务是存在缺陷的。

鉴于此，本课题拟采用基于 PointNet++网络的特征提取方式(图 4-3)。PointNet++通过构建一系列的点云子集和局部特征集，逐层聚合这些子集和特征，最终得到整个点云的全局特征向量。这种逐层聚合的机制可以在保留全局信息的同时，提取更丰富的局部特征。PointNet++在处理具有复杂结构和局部特征的点云时，具有更好的性能。因此利用 PointNet++网络对点云形状、位姿等特征进行提取具有较高的可行性。

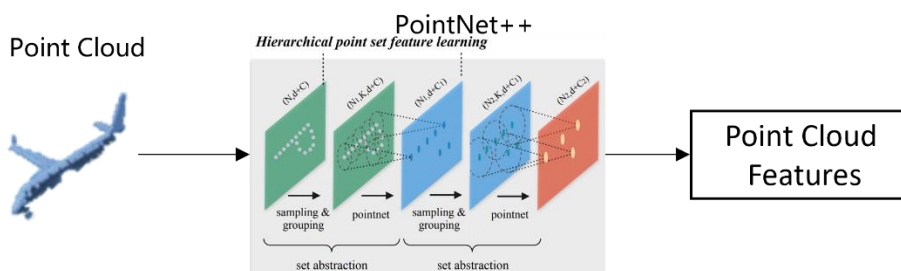


图 4-3 点云特征提取示意图

3) 特征融合方案

针对特征融合可采用多模态的融合方式，多模态融合可分为以下三种：首先是前融合(数据级融合)，它指通过空间对齐直接融合不同模态的原始传感器数据。其次是深度融合(特征级融合)，它指通过级联或者元素相乘在特征空间中融合跨模态数据。最后是后融合(目标级融合)，它指将各模态模型的预测结果进行融合，做出最终决策。由于在本课题中主要任务是

对图像和点云特征进行融合，因此工作主要围绕深度融合进行展开。

本课题拟采用基于 EPNet 进行图像和点云的特征融合。EPNet 的全称是“Embedding Projection Network”，它的核心思想是将不同模态的信息转化为同一维度的向量表示，然后通过投影和融合操作，将它们融合成一个综合的表示。EPNet 包含两个部分：Embedding Network 和 Projection Network。Embedding Network 用于将不同模态的信息转化为向量表示，它可以使用不同的神经网络结构来处理不同类型的数据。Projection Network 则用于将这些向量进行投影和融合，它可以使用多种方法来实现，比如简单的加权求和、门控机制等。EPNet 网络的两个部分也与本课题的设计思路相符，先分别对点云和图像进行特征提取，再进行融合。通过基于 EPNet 进行的特征融合方法完成最终的点云补全任务。

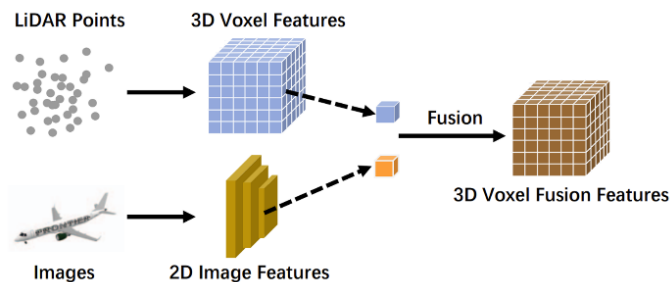


图 4-4 特征融合示意图

4.7 可行性分析

4.7.1 工作基础

1) 调研和综述

前期工作对点云补全的发展历史和研究现状做了详细的调研和综述，明确本课题的研究目的：提出一个基于图像融合的点云补全方法；学习了目前主流的研究方法，并对其优缺点进行总结；在目前常用的点云补全数据集上，对前人的研究进行了实验，通过分析实验结果，找到了研究的突破口。通过对点云补全的文献调研，了解了现有的各种算法及其性能上的比较。

2) 软硬件基础

实验过程中用到的 PointNet、PointNet++ 等网络都已开源，且有完善的学习教程可供指导；目前课题所需要的数据集已经整理好，且拥有 3D 线结构光扫描相机可供实验；实验室配备小型工作站，为实验的顺利进行提供了硬件保障。

3) 个人知识和能力基础

本人已进行过图像处理、模式识别、机器学习、深度学习等相关知识的系统学习，为本课题算法的理解与研究提供了扎实的理论知识基础；对当前点云补全方向的发展有清晰的认识，对深度学习网络有一定了解，可以对课题进行更深入的研究；此外，本人掌握 VSCode、Pycharm、Visual Studio 等开发工具和 Python、C++ 语言以及深度学习框架的实践应用，并且积累了一定的数学理论知识，便于理解复杂的算法和模型。有能力完成网络的设计和改进。

4.7.2 相关技术成熟度

本课题中涉及到的图像、点云特征提取和图像、点云模态融合均有成熟的解决算法，基于图像的点云补全方法尚处于研究阶段。总之，本课题涉及到的相关技术均有成熟或半成熟的解决方案，为课题算法的初步研究提供了可行的选择空间。

4.7.3 技术路线与研究方案

本课题对算法的理论说明、方案设计、网络训练和结果评价均进行了详细的说明，部分模块还设计了具体的方案，说明了本课题具有细致的技术路线和合理的研究方案。前期已对部分研究方案进行试验验证，说明本课题提出的解决方案具有合理性和可行性。

综上所述，本课题中涉及的各项内容均有坚实的理论基础，已有的数据和成果为本课题的开展提供了可对比结果和借鉴依据，基于此制定的研究方法、技术路线具有完全的理论延续性。因此，本课题在理论、技术、软硬件和实践等方面均具有可行性。

五、预期研究成果（本节可以整页扩页）

对所研究的成果进行阐述，同时要对与前文研究内容的相关性及与前人（他人）研究成果的差异性进行描述

5.1 本课题的预期研究成果

1) 功能上：根据整体设计路线，实现图像的特征提取、点云的特征提取、图像与点云模态间的融合，以此实现一套抗噪能力强、对工业生产环境鲁棒的点云补全方法。

2) 性能上：在上述列举的数据集上与 SOTA 方法相比，性能优于 SOTA 方法。运用到实际工业场景中，能准确地补全缺失点云。

3) 可量化成果：根据研究课题发表论文一篇。

5.2 与前文研究内容的相关性

预期的研究成果是在研究内容的基础上完成的。

1) 通过对图像的局部特征和全局特征进行提取，以获取图像所包含的丰富信息，为后续整体点云的精准补全提供前提条件。

2) 通过对部分缺失点云的特征提取，以获取点云中所包含的位姿、形状等信息，为后续整体点云的精准补全提供前提条件。

3) 引入图像和点云的模态融合，将提取到的图像、点云特征融合，实现点云的高精度、精确的补全，解决实际提取到的点云缺失的问题。

5.3 与前人研究成果的差异性

现有的大部分方法都基于点的补全，使用多层感知器（MLP）独立地对每个点进行建模。由于点云的变换不变性，全局特征通过对称函数（如最大池化）进行聚合。然而，整个点云中的几何信息和相关性仍未完全考虑，这必然会影响到最终补全的点云的质量。与之前的研究相比，本课题提出的方法引入了图像融合方案。通过将图像特征与点云特征融合以及图像边缘特征的提取，能够更好地获取完整的点云。

本课题的方法也考虑了图像的边缘特征。传统的基于视图的点云补全方法仅针对图像全局特征进行提取，所补全的点云会出现边缘粗糙的问题。然而，相对于图像的全局特征，图像的局部特征也是不应忽视的。通过提取并融合图像的边缘特征能够进一步提高点云补全的质量。

六、研究条件（本节不允许扩页）

1. 所需实验手段、研究条件和实验条件

（如果没有，可以空白）

6.1.1 所需实验手段

基础知识学习和算法的理论研究；
深度学习网络框架搭建的研究。

6.1.2 研究条件和实验条件

- 1) 工作站一台
- 2) PC 机一台

2. 所需经费，包含经费来源、开支预算（工程设备、材料须填写名称、规格、数量）

（如果没有，可以空白）


七、工作计划（本节不允许扩页）

序号	阶段及内容	工作量估计 (时数)		起止日期	阶段研究成果
1	数据集整理及实验所需软件环境的配置	240		2023/10/01-2023/10/31	整理好实验所需数据集及实验环境的配置
2	研究图像的特征提取并实验	480		2023/11/01-2023/12/31	获取高效的图像特征提取方法
3	研究点云特征提取并进行实验对比	720		2024/01/01-2024/03/31	获取丰富的点云特征提取方法
4	研究图像与点云间的模态融合并实验	480		2024/04/01-2024/05/31	完成图像与点云间的模态融合
5	对各部分研究成果进行整合并实验	480		2024/06/01-2024/07/31	完成点云补全任务网络框架的搭建
6	与 SOTA 方法进行实验对比，分析算法不足及改进	480		2024/08/01-2024/09/30	优化整体的网络框架
7	获取整体实验数据，撰写小论文	720		2024/10/01-2024/11/30	完成小论文撰写工作
8	撰写毕业论文	1200		2024/12/01-2025/04/30	完成毕业论文撰写工作
		合计	4800		

八、导师意见

基于上述内容及学生前期研究情况，导师是否同意其开题

该课题解决问题阐述清晰，研究内容设置合理，同意开题。

导师签字（盖章）：

2023 年 10 月 13 日