计 算 机 研 究 与 发 展
Journal of Computer Research and Development
DOI: 10.7544/issn1000-1239.202330077
61(11): 2870−2896, 2024

基于图的点云研究综述

梁　循　李志莹　蒋洪迅
（中国人民大学信息学院　北京　100872）
（xliang@ruc.edu.cn）

Research on Graph-Based Point Cloud: A Survey

Liang Xun, Li Zhiying, and Jiang Hongxun
（*School of Information*, *Renmin University of China*, *Beijing* 100872）

Abstract　The processing, transmission, and semantic segmentation of point clouds are important analytical tasks in the field of 3D computer vision. Nowadays, the effectiveness of graph neural networks and graph structures in point cloud research has been confirmed, and graph-based point cloud (GPC) research continues to emerge. Therefore, a unified research perspective, framework and methodology need to be formed. This paper systematically sorts out various application scenarios of GPC research, including registration, denoising, compression, representation learning, classification, segmentation, detection and other tasks, summarizes the general framework of GPC research, and proposes a technical route covering the current GPC research. Specifically, this paper gives the hierarchical concept category of GPC research, including low-level data processing, intermediate representation learning, and high-level recognition tasks. GPC models or algorithms in various fields are reviewed, including static and dynamic point cloud processing algorithms, supervised and unsupervised representation learning models, and traditional or machine learning GPC recognition algorithms. The representative achievements and their core ideas are summarized, such as dynamically updating the nearest neighbor graph at each layer of feature space, hierarchical and parameter sharing dynamic point aggregation module, and segmentation accuracy improvement employing graph partitioning as well as graph convolution. The model performances are compared, including OA (overall accuracy), mAcc (mean accuracy) and mIoU (mean intersection over union). Based on the analysis and comparison of existing models and methods, the main challenges faced by GPC are summarized, the corresponding research issues are put forward, and the future research directions are explored. The GPC research framework established in this paper is general and comprehensive, which provides field positioning, technical summary and macro perspective for subsequent researchers to engage in this new cross-field research of GPC. The emergence of point cloud research is the result of the rapid progress of detector hardware technology. The current state of point cloud research indicates some challenges between theory and practical applications, and there are still some key issues to be addressed. However, the development of point cloud research is expected to propel artificial intelligence into a new era.

Key words　point cloud; graph structure; graph-based point cloud (GPC); graph signal processing; spatial-temporal graph; graph neural network

摘　要　点云的处理、传输、语义分割等是3维计算机视觉领域重要的分析任务. 现如今, 图神经网络和图结构在点云研究方面的有效性已被证实, 基于图的点云（graph-based point cloud, GPC）研究不断涌现. 因

收稿日期: 2023-02-16; **修回日期**: 2023-12-26
基金项目: 国家自然科学基金项目(72071203)
This work was supported by the National Natural Science Foundation of China (72071203).
通信作者: 蒋洪迅(jianghx@ruc.edu.cn)

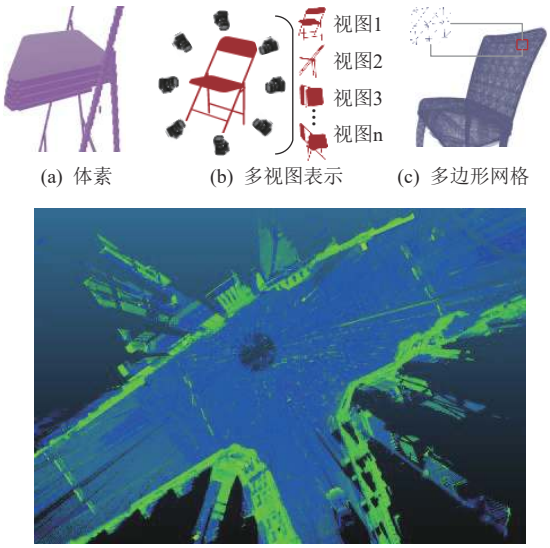
此，一种统一的研究角度、框架和方法论亟待形成. 系统性梳理了 GPC 研究的各种应用场景，包括配准、降噪、压缩、表示学习、分类、分割、检测等任务，概括出 GPC 研究的一般性框架，提出了一条覆盖当前 GPC 全域研究的技术路线. 具体来说，给出了 GPC 研究的分层概念范畴，包括底层数据处理、中层表示学习、高层识别任务；综述了各领域中的 GPC 模型或算法，包括静态和动态点云的处理算法、有监督和无监督的表示学习模型、传统或机器学习的 GPC 识别算法；总结了其中代表性的成果及其核心思想，譬如动态更新每层特征空间对应的最近邻图、分层以及参数共享的动态点聚合模块，结合图划分和图卷积提高分割精度；对比了模型性能，包括总体精度（overall accuracy, OA）、平均精度（mean accuracy, mAcc）、平均交并比（mean intersection over union, mIoU）；在分析比较现有模型和方法的基础上，归纳了 GPC 目前面临的主要挑战，提出相应的研究问题，并展望未来的研究方向. 建立的 GPC 研究框架具有一般性和通用性，为后续研究者从事 GPC 这个新型交叉领域研究提供了领域定位、技术总结及宏观视角. 点云研究的出现，是探测器硬件技术长足进步后应运而生的结果；点云研究的现状表明在理论和实践之间存在一些挑战，一些关键问题还有待解决. 同时，点云研究的发展将推动人工智能进入新的时代.

关键词 点云；图结构；基于图的点云；图信号处理；时空图；图神经网络

中图法分类号 TP391

1 图和点云交叉领域研究总析

3 维视觉研究在虚拟现实、3 维重建[1-3] 和动画[4] 等生产、生活和娱乐消费领域都有着重要意义. 3 维视觉数据的可视化表达，如图 1 所示，包括体素、多视图、多边形网格和 3 维点云等主要形式. 体素指体积元素，即具有固定大小的立方体. 例如著名游戏 Minecraft 就采用体素进行 3 维建模，多个体素像 2 维平面的像素块那样紧密排布来构成 3 维物体或场景[5-7]. 多视图表示指从多个视角对物体拍照得到 2 维图像集合来表示 3 维物体[8-11]. 多边形网格是指使



(a) 体素 (b) 多视图表示 (c) 多边形网格

(d) 3维点云

Fig. 1 Three-dimensional representation

图 1 3 维表示形式

用几何凸多边形共用顶点连成网格形表面来近似真实物体或纹理[12-13]，譬如经典游戏魔兽世界就是采用多边形网格建模 3 维场景. 3 维点云是指具有 *xyz* 三维坐标的点的集合[14-15]，在 3 维空间中随机放置的离散点汇聚在一起，呈现出云团状，故得名点云. 上述 4 种表达形式各有优劣：体素表示方法分辨率低且存储成本较高，难以细腻描绘真实物体或场景；多视图表示由于 2 维图像的局限性，对 3 维视觉的表现能力有限且存储和计算成本高；多边形网格表示常见于动画，依赖于模拟或仿真设计软件，一般无法用传感器直接获取；3 维点云来自于激光雷达（light detection and ranging, LiDAR）传感器的直接输出，传感器通过发射激光到物体表面[16]，对物体表面精细扫描后，从激光往返时间、激光传输方向以及传感器位置计算出反射点的 3 维位置坐标. LiDAR 点云不仅保留了 3 维空间的原始定位信息，还能精确地表征目标的形状[17]，是自动驾驶[18-20] 或机器人[21-22] 等应用场景的首选 3 维数据表示方式. 随着 3 维点云数据获取设备逐渐低成本化，视觉任务逐渐由 2 维转向 3 维，较为常见的数据来源有基于深度相机的方法[23]、基于 LiDAR 的方法[24] 以及基于激光扫描仪的方法[25]. 例如自动驾驶领域，传统的摄像头数据获取方式很容易受到天气、光照等影响，而 LiDAR 即便在光照不足的情况下也获取到清晰的 3 维点云数据. 除了标的物的坐标信息和物体形状，传感器还可能返回反射表面的 RGB 颜色或回波的反射强度等数据，所以某些场景下点云数据的属性会超过 3 维点云数据[26].

除以上介绍的常用 3 维数据表现形式，还应注意

介于2维和3维之间的深度图. 深度图是一种准3维表现形式，它由2维图像表示，其中每个像素点的值代表了该点的深度值. 点云可通过将点坐标投影到2维平面，同时第3维度作为深度值的方式转换为深度图. 相反地，深度图也可通过反投影的方式还原得到对应点云. 深度图仅需要2维像素信息，因此在计算和存储方面具有优势. 然而，深度图的表达能力相对较弱，无法像点云那样准确地表达复杂的3维结构，以及保留物体的几何形状等细节信息. 实际应用往往根据任务要求和数据特点来选择最合适的表达形式.

点云包含的丰富信息及其独特性质，为其数据处理、学习表示和分析识别任务带来了巨大的挑战和困难. 现有研究表明，3维点云数据具备无序性[27]、变换不变性[28]和点互动性[29]3种基本特性，其庞大的数据规模、非结构化的数据性质、传感器产生的噪声干扰等都是挑战3维点云分析研究的主要问题[30-33]. 以PointNet[14]为代表的逐点模型部分解决了这些特性，但缺乏对点云之间的局部非结构化信息的捕获. 基于图的研究方法为解决这一缺陷提供了强有力的支持.

图(graph)作为基本的抽象数据结构之一，天然具有非结构化和不规则性，在处理具有相似特性的非结构化的点云数据上具有极大优势[34]. 很多学者利用网格数据及其对应的图结构(即顶点和连边)提取所需信息，实现了更好地描述3维数据的局部信息. 文献[35−36]通过传播和聚合相邻顶点的特征来获得最终的表示，而文献[37−38]则选择具有固定模式的精细核来观察局部信息. 点云转化为图表示并聚合局部信息的示例如图2所示. 在点云的诸多应用场景，如表示学习[39-40]、分类[41-42]和分割[34,43]等任务上，基于图结构和图技术的方法和模型[44-46]不断涌现，逐渐成为点云研究的重要分支之一. 不过，现有的成果仍然分散在点云研究的各个小方向内，尚未形成统一的研究体系. 况且，很多基础性的方法和关键技术目前仍是开放性问题，譬如如何对3维点云数据快速有效地构建图并应用图技术[47]，怎样针对点云数据的特性将图的强大表示能力充分挖掘[37]，亟待更多交叉性的研究和探索.

1.1 区域趋势分析

本节对基于图的点云(graph-based point cloud, GPC)这一交叉领域现有研究文献进行限定搜索和区域趋势分析. 鉴于WoS(Web of science)能够提供质量更高、更为细致的引文分析，本文选择WoS数据库作为文献数据源，并利用Citespace[48-49]进行分析.

根据应用领域划分，图3展示了GPC论文涉及最多的WoS领域类别，说明了这类成果主要涉及遥感、图像科学、人工智能等技术，并应用在工程电子电气、地理科学等多学科领域. 本文采集这些论文的总结或展望部分的内容，去除“图”“点云”等停用词后进行文本主题挖掘分析. 表1汇总了学者对于GPC研究下一步发展方向的建议和共识，包括多模态融合、弱监督与自监督、邻域搜索、动态点云等. 本文第7节将对这些未来趋势做详细分析.

根据作者地域划分，图4列举了论文发表数量排名前10的国家，显然中国和美国是目前GPC研究领域的中坚力量. 从发表时间趋势看，图5展示了过去20年来中美德3国每年的论文发表数量. 显然，美国和德国GPC研究起步较早，而中国启动虽晚但增速较快，目前论文总量上已超越了德国和美国. 中国GPC研究经历了后起、狂追和超越3个阶段，在点云

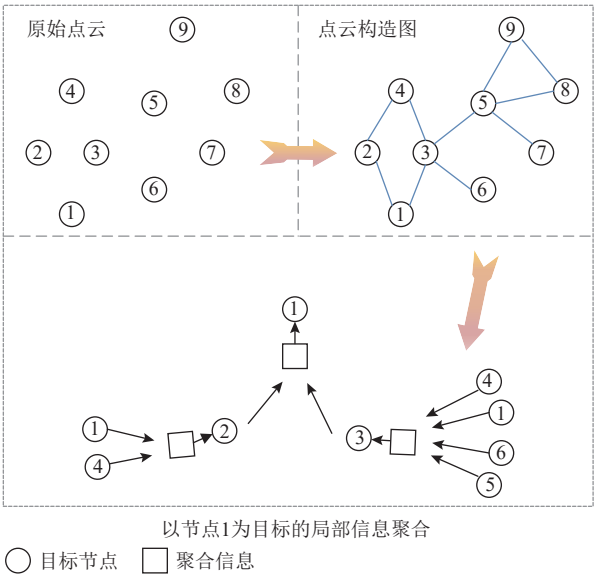


Fig. 2 Graph construction and local information aggregation of point cloud

图2 点云图构造及局部信息聚合



Fig. 3 Block plot of the top 5 categories published in cross-cutting studies

图3 交叉领域研究中发表数量前5的类别块图

Table 1 Summary of Top 8 Future Development Directions

表 1 前 8 项未来发展方向汇总

排序	方向	权重	排序	方向	权重
1	多模态融合	42	5	动态	12
2	目标检测	31	6	无监督	10
3	图像	20	7	邻域	9
4	弱监督	15	8	自监督	9

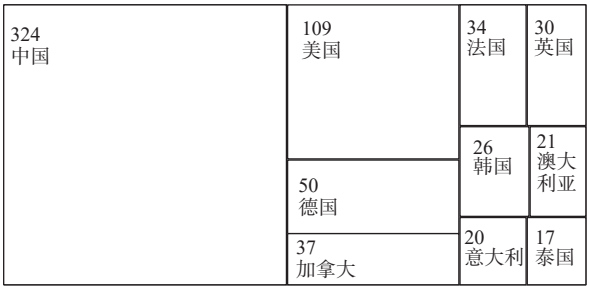


Fig. 4 Block plot of the top 10 countries for the number of papers published in cross-cutting studies

图 4 交叉领域研究中发表论文数量前 10 的国家块图

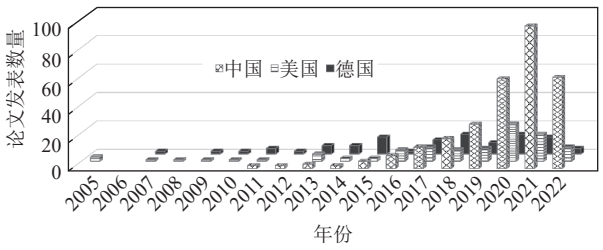


Fig. 5 Comparison of the number of publications in China, America and Germany by year

图 5 中美德国家逐年发表论文数量对比

和图的交叉领域成为有影响力的大国.

除作者外，还有一些机构在点云研究领域做出了重要贡献，推动了图在点云中应用的不断发展，如世界著名的研究机构—— 马克斯·普朗克研究所(Max Planck Institute, MPI)，在计算机视觉、图像处理和点云分析等领域有着卓越的研究成果. 此外，卡内基梅隆大学(Carnegie Mellon University, CMU)作为一所享有盛誉的研究型大学，在 3 维视觉和人工智能领域有着深厚的学术积淀. 该机构的研究人员积极参与点云与图交叉领域的研究，推动了该领域的发展. 斯坦福大学(Stanford University)在计算机视觉和机器学习等领域拥有世界领先的研究实力，其研究人员在点云与图交叉领域的研究方面也取得了令人瞩目的成果.

1.2 主题层次分析

根据文献的关键词共引分析，本文得到了如表 2 所示的图领域和点云领域的高频术语表. 显然，点云和图交叉的 GPC 研究，以解决分割、分类等识别任务居多，采用的技术主要包括图信号处理、图卷积网络等模型，构造图的方法主要通过最近邻计算、权重图等方式. 表 2 的术语词频还表明，相比于点云领域的比较聚合的术语关键词，图领域的关键词比较离散、术语同义不同词、单个关键词频次偏低，说明学者们在图技术的关键词使用上截至目前尚未达成共识.

Table 2 Top 7 High-Frequency Terms in the Graph and Cloud Domains

表 2 前 7 项图-点云领域高频术语

图领域		点云领域	
术语	词频	术语	词频
graph signal processing	12	segmentation	47
graph convolutional (network)	10	classification	18
graph neural network	6	recognition	14
attention mechanism	5	compression	13
neighborhood graph/ nearest neighboring graph	3	object detection	10
weighted graph	3	surface reconstruction	9
graph optimization	2	registration	4

图 6 展示了文献标题共引网络的聚类分析结果. 我们去除了 point cloud, graph, deep learning 等停用词后，得到的聚类结果包括目标检测(簇#4)、分类(簇#1)、分割(簇#7)、表示学习(簇#0)、配准(簇#3)、压缩(簇#6)这些大类别. 此外，簇#14 是作为前述簇的强相关图技术，表明点云各任务与图匹配和图神经网络技术具有强相关和交叉性.

图 7 呈现了点云各任务之间的上下游关系. 点云底层处理任务(low-level tasks)包括配准、去噪、压缩和修复. 它们是获取原始点云数据后首要解决的问题，在点云各个处理任务中处于基石地位；中层任务(intermediate tasks)为表示学习抽取点云特征嵌入，其

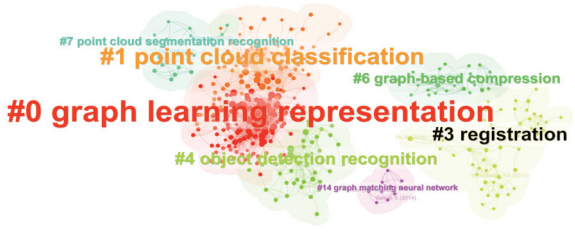


Fig. 6 Visualization of literature co-citation clustering results by title

图 6 按标题文献共引聚类结果可视化

输出可进一步作用于下游点云任务；点云处理高层任务(high-level tasks)主要包括分类、分割、目标检测等识别任务，体现更具体的应用目的. 本文后续部分将按照该领域实体顺序展开，如图 8 所示.

GPC 是点云和图领域交叉的新涌现研究方向，早期的一些研究将其归类为基于点集的研究方法之一，这种分类不够准确；近年来少量点云综述文献[50−51]将 GPC 的某些算法作为独立小类专门阐述，但通常仅仅覆盖了点云研究的一些少量处理任务. 本文试图从全层次、多粒度处理领域上概括 GPC 的现有研究. 据我们所知，这是将 GPC 作为一种跨领域相对独立的研究分支，囊括了点云表示学习、分类、分割、目标检测、压缩、配准等多任务领域的最为全面的 GPC 方法综述，试图总结出一种通用性的 GPC 研究框架，为点云和图跨领域研究者提供所需的研究基础和研究思路参考，促进更多的跨领域合作，推动前述 GPC 开放性问题得到突破性解决.

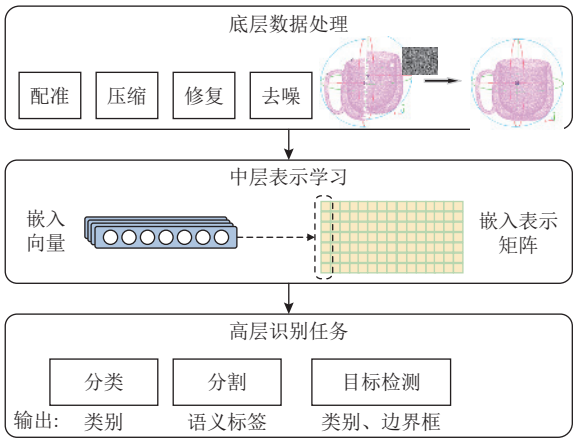


Fig. 7 Domain entities and their associated order

图 7 领域实体及其关联顺序

2 理论和概念基础

点云是定义在 3 维空间上的一组无序点的集合，其形式化表示为 $P=\{p_1,p_2,\cdots,p_n\}\subseteq\mathbb{R}^3$，其中 n 为点云的大小，即点的数量. 任何节点除了包含 3 维坐标外，还可能包含诸如颜色(RGB)、深度(depth)、曲率(curvature)、法线(normal)、光强度(intensity)等衍生信息. 根据获取途径的不同，点云数据可分为静态点云、动态点云. 静态点云指物体和扫描设备均为静止状态下所获取的数据集，可用来储存静态的建筑物、文化遗产、地理信息等；动态点云指物体处于运动状态而扫描设备为静止状态下获取的数据集，如人物动态行为的点云视频流[52-53]，或设备运动状态下扫描获取的数据集，如载有 LiDAR 的汽车行驶过程中动态扫描的含时间戳的道路景观点云视频流[54].

图 9 呈现了点云数据的排列不变性、变换不变性、内部交互性 3 种基本特性. 1)排列不变性[27]. 点云是 3 维空间中一系列点的集合，在空间上点的排列顺序不影响它对物体特征信息的表示，如图 9(a)所示，无论点云中输入数据点的排列顺序如何变化，算法或模型输出的特征提取结果都应该相同. 2)变换不变性[28]. 点云是 3 维物体的表现形式，相同的点云集合在空间中经过一系列刚性变换，如旋转或平移，坐标发生变化后仍然表示的是同一个数据，即模型应该对于无论何种坐标系下的相同点云数据，都能做出相同的判断和相同的处理结果. 图 9(b)的物体来自公开数据集 PCPNet[55-56]中的杯子点云，为方便读者观察，本文在可视化结果上手动标注了黑色匹配线，以突出旋转后的各部位不变性. 3)内部交互性[29]. 由于点云中每个数据点都由一个 3 维空间坐标

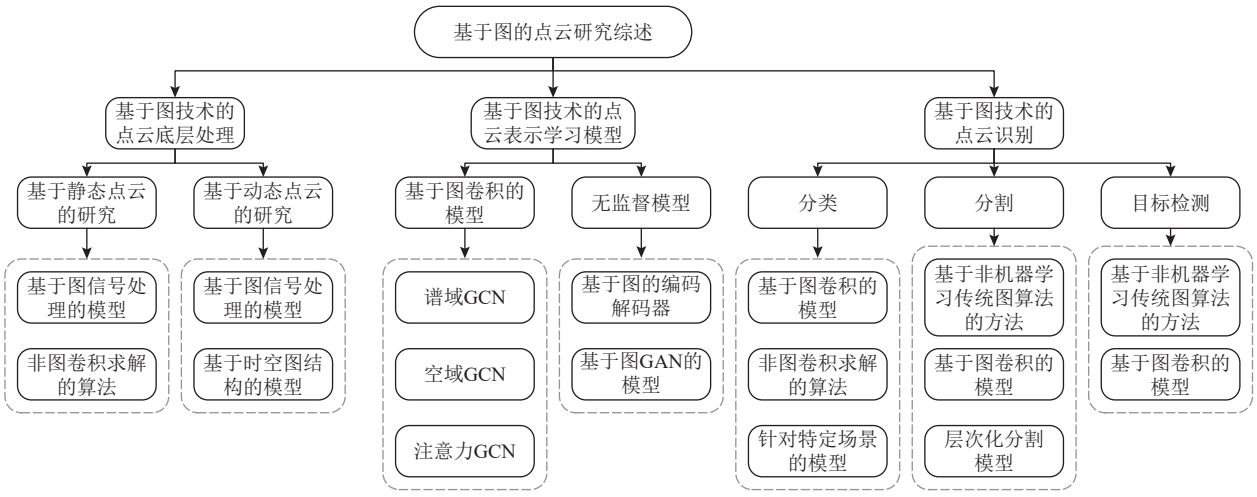
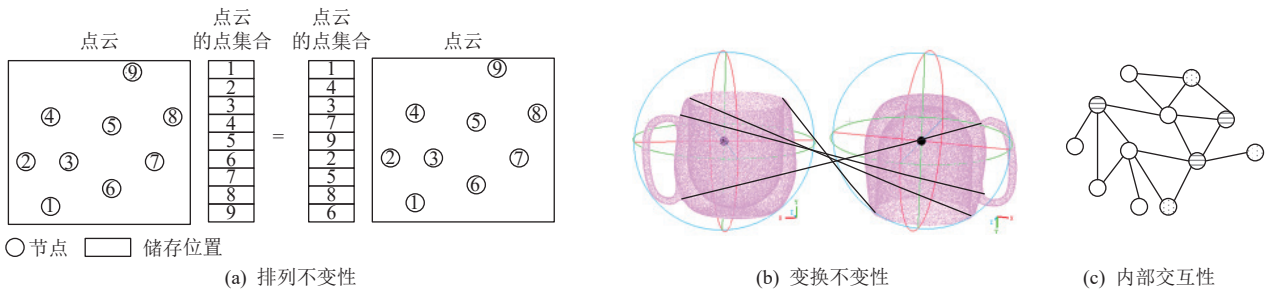


Fig. 8 Classification of graph-based point cloud studies

图 8 基于图的点云研究分类



(a) 排列不变性 (b) 变换不变性 (c) 内部交互性

Fig. 9 Three characteristics of point clouds

图 9 点云的 3 种特性

表示，如图 9(c) 所示，各点之间存在一定的空间距离关系，各点之间不是孤立存在的，相邻的点可以形成一个有意义的子集，反映物体的局部特征，即数据中各点之间具有一定的内部交互关系.

点云是一种无层次、不规则的数据形式，为其处理和分析任务带来了很大的挑战. 当前对点云数据的处理方式多是将其构造为规则的网格型数据，所耗代价昂贵[57-58]；或者利用体素化技术处理，导致不可避免的低分辨率问题[14]. 直接针对点集合处理的方法势必会弱化 3 维点云数据间的拓扑连接关系[34,59]，从而影响数据的特征表达效果. 如何直接高效处理点云数据，是研究者须面对的关键挑战.

2.1 点云的图结构

图是一种非结构化的抽象数据结构，由节点(node)和连接节点的边(edge)组成，形式化定义为 $G=(V,E)$，其中 $V=\{v_1,v_2,\cdots,v_n\}$代表节点集合，$E=\{e_1,e_2,\cdots,e_m\}$代表边集合. 节点又称为顶点(vertex)，是对实体属性特征的抽象，设每个节点存在 F 维特征，则图的节点特征矩阵记为 $\boldsymbol{X}^{N\times F}$，其中 N 表示节点个数；边又称为连接(connection)，是对实体拓扑依赖关系的抽象，图的边可以表达为邻接矩阵 $\boldsymbol{A}^{N\times N}$. 图有很多种类，根据边是否带有方向，可划分为有向图和无向图；根据连边是否含有权重，可分为带权图和无权图；根据图中所有节点所代表的实体是否属于同一类型，可分为同质图或异质图.

点云的图结构化是将 1 个点云物体或场景视作 1 个图，物体或场景中的每个点代表图中的 1 个节点. 点云中每个点的坐标及其他属性特征作为图中节点的特征，为点云中的点构造出边集合，即可生成点云的图结构 $G=(P,E)$. 构造边的方法可以通过对每个点计算其 k 个最近邻居，即 k 近邻(k-nearest neighbors, k-NN)，将点和邻居进行连接. 对 3 维空间上点云构造出图结构之后，便可利用相关图技术解决后续任务.

2.2 点云的图技术

GPC 研究所涉及的图技术主要包括图分割、图匹配、图优化、图神经网络等. 图分割将一个大图划分成若干个子图，使得子图之间的割边权值和最大，属于 NP 难问题，可用于社交网络的聚类或社区发现，主要思想是一个好的分割结果能够使得子图内部连接数量最大化，而不同子图之间的连接数量最小化. 经典算法主要有点划分法、边划分法和超图划分法等[60]. 图匹配是图形学和计算机视觉领域中的基本问题，即在 2 个或多个图之间寻找顶点对应关系，是 NP 难的离散组合优化问题，经典算法主要有图同构算法、最大公共子图检测算法、概率松弛方法等[61]. 图优化是用图的方式描述常规优化问题的技术，顶点作为待优化的变量，而边作为误差项. 例如 SLAM 领域中一个移动的机器人某时刻的位姿(pose)可视为图中的一个顶点，而位姿之间的关系构成一条边，通过设置合适的边的约束，优化问题则转化为调整机器人的位姿(顶点)来尽可能满足边的约束，使得误差最小，最终得到优化后的位姿地图. 目前较为流行的优化方法有 g2o, Ceres Solver, GTSAM, SESync 等[62].

图神经网络(graph neural network, GNN)[63] 是指用于图结构数据上的模式提取和特征学习的神经网络模型的总称. GNN 可用来完成图上的聚类、分类、分割等任务，其主要分支有频谱域和空域上的图卷积网络(graph convolution network, GCN)、图编码器(graph auto-encoders, GAE)、图生成网络(graph generative network)、图注意力网络(graph attention network, GAT)和图嵌入(graph / network embedding)等. GCN[64] 主要基于卷积神经网络(convolutional neural network, CNN)，分为谱域和空域 2 种. 谱域图卷积借助图信号处理的相关概念，将节点特征作为图顶点的信号，利用邻接矩阵和图傅里叶变换[65](graph Fourier transform, GFT)在谱域上对图信号进行卷积，经典的谱域图卷积算子包括 Chebyshev 卷积算子[66] 和 CayleyNets[67]. 空域图卷积在空间拓扑结构上定义卷积操作，经典的空域图卷积算子包括 GraphSAGE[68] 和 FastGCN[69]. 图编码器是将输入信息编码再解码，以获得隐含信息

表征的人工神经网络模型, 包括 MGAE[70] 和 SIG-VAE[71]. 图生成网络属于一种生成对抗网络(generative adversarial network, GAN), 它借用 GAN 学习图的特征分布构建新的图结构数据的 GNN 模型, 包括 GraphGAN[72] 和 GraphSGAN[73]. 图注意力机制的核心思想是为中心节点的邻居节点分配权重因子, 使得 GNN 只关注特定邻居节点, 而不是给所有邻居分配相同的注意力, 主要模型有 GAM[74] 和 ChebyGIN[75]. 图嵌入又称网络嵌入或网络表示学习, 旨在将网络中每个节点嵌入映射到低维向量空间, 获得内在拓扑信息的节点的特征表示, 包括 DeepWalk[76], Node2vec[77], CANE[78], SepNE[79] 等.

2.3 基于图的点云方法与其他方法的比较

虽然本文主要关注 GPC 方法, 但将 GPC 与其他在点云分析中常用的方法, 如 PointNet 和 Transformer 进行对比具有重要意义. 在本节中, 我们旨在突显 GPC 的独特优势以及它相较于其他方法的优越性.

PointNet 提出了直接处理无序点云的思想, 无需使用体素化或基于网格的表示. 它在点云分类和分割等各种任务中取得了令人瞩目的成果. 然而, PointNet 存在一些局限性. 作为一种全局特征提取方法, 它可能无法充分捕捉复杂场景中的局部几何结构, 限制了其处理复杂细节的能力. 此外, PointNet 在排列不变性方面较为脆弱, 点的顺序微小的变化会导致不同的特征表示.

Transformer 架构以其在自然语言处理任务中的成功而闻名, 也被应用于点云分析. 通过使用自注意机制, Transformer 可以有效地建模点之间的长期依赖关系. 然而, 由于输入数据的高维性, 将 Transformer 应用于点云时常常导致高计算开销, 使其在实时应用或大规模点云数据集上的应用不太实际.

相比之下, GPC 方法为点云分析提供了灵活而有效的框架. 通过将点云表示为图, 这些方法本质上捕捉了点之间的局部和全局几何关系. 这种特性使它们能够更高效地保留细粒度的结构, 并从点云中提取有意义的特征. GPC 方法可以解决 PointNet 中的排列不变性问题. 通过使用图卷积或消息传递技术, 这些方法可以从相邻点聚合信息, 从而获得鲁棒且排列不变的特征表示. 此外, GPC 提供了灵活性, 可以将其他信息(如法线、颜色或语义标签)作为图的边缘或节点属性进行嵌入. 这种丰富的表示增强了基于图的模型的区分能力, 使它们能够有效处理各种点云分析任务.

目前, GPC 在点云分割、目标检测和形状分类等任务中展现出更高的准确性. 此外, GPC 方法往往具有更好的泛化性能, 能够处理不完整或带有噪声的点云数据, 使其特别适用于应对复杂真实场景. 虽然 PointNet 和 Transformer 在点云分析领域做出了重要贡献, 但 GPC 在保留细粒度结构、处理排列不变性以及融合附加信息等方面具有独特优势. GPC 的灵活性和优越性使其成为各类点云分析任务中的有力选择.

3 基于图的点云底层处理

LiDAR 扫描过程会遇到各种状况, 譬如遮挡产生盲区、设备抖动导致噪声以及对大物体或大环境多次扫描形成不同坐标系的点云数据. 点云数据这些缺陷需要系统对不同坐标系的点云进行底层处理[80-83]. 当点云规模过于庞大时还需压缩数据以利于储存和传输[84].

点云配准[80,85] 是指对在不同方位和角度下测得的 3 维点云进行刚体变换, 使原本不同的坐标系变得对齐和统一. 以面向城市建筑物场景建模的地面激光扫描(terrestrial laser scanning, TLS)技术为例, 由于城市中多栋建筑物和其他地面物体排布复杂、互相遮挡, 限制了 TLS 设备的扫描视角. TLS 往往需要多视角扫描才能完整覆盖城市建筑物场景. 获得的多角度点云数据处于不同的局部坐标系内, 需要经过配准才能得到统一坐标的完整场景. 鉴于大规模场景构成复杂多变, 使得 TLS 或 LiDAR 获得的点云数据包含大量噪声和不完整结构, 点云数据量庞大、局部密度变化明显. 如何提高大规模复杂多变场景下的点云自动配准精度和效率是目前亟待解决的问题.

点云去噪[86] 是指过滤点云数据中的干扰噪声, 即原始数据中的散列点、孤立点、异常扰动点等. 噪声来自于物理测量误差或扫描场景中的无关物体干扰. 相对于静态点云, 动态点云上的去噪和修复均面临着时间和空间域的不规则采样模式所造成的巨大挑战. 目前已有研究[87] 借助时空图结构及其上的优化模型来处理动态点云的时间和空间上的对应关系, 取得了突出的成效.

点云修复[82] 旨在修补点云数据中的孔洞, 以利于后续的渲染或更高层次任务的解析. 特别是对于动态点云来说, 移动中的设备使得动态点云相较于静态点云存在更多数据缺失导致的孔洞. 近年来随着深度传感器和激光扫描技术的成熟, 动态点云的获取越来越便利, 大大推广了动态点云在自动驾驶、增强和虚拟现实等运动视觉领域中的应用. 不过, 动态点云的修复[88] 作为预处理级别的底层任务, 受重

视程度远小于高层识别任务，始终处于被忽视地位. 况且不规则的空间和时间采样模式也给动态点云的修复带来了极大的困难.

3维点云压缩[84]的重要性日益凸显. 随着采集设备精度的提高，点云的数据规模与日俱增，这对点云的存储和传输提出了更大的挑战. 基于图的点云压缩技术逐渐兴起，特别是对于动态点云的压缩方面，基于动态图结构或时空图结构的相关模型陆续涌现[52,89]. 点云的无损压缩技术算法的性能衡量主要从压缩后的比特流大小、压缩过程的时间复杂度2个方面去考虑. 压缩后比特流大小是指编码后输出的比特数，时间复杂度是指点云编码以及解码全过程所耗费的时间. 对于有损压缩，除以上2方面外，还需要衡量点云的失真程度. 失真包括几何失真和属性失真2部分，即编解码对应点的几何距离损失和属性数值损失. 对应点是指原始点与重建后寻找到的最近点或者重建点在原始点云中的最近点，而最近点的寻找是通过构建 k-d 树（k-dimensional tree, k-d tree）结构来完成[84].

点云底层数据是点云分析和处理的基础，该领域常用特定的点云处理工具包来操作和处理点云数据. 一些代表性的点云处理工具包简介如下：

1）Open3D. 支持包括 pts, ply, pcd, stl 等常见点云文件格式，提供丰富的点云数据结构、滤波、配准、重建和可视化功能，简单易用、功能强大，使其成为研究者和从业者广泛使用的工具.

2）PCL（点云库）. 拥有大量的点云滤波、分割、配准、特征提取等算法，支持多种点云数据表示和输入输出格式，并提供了 C++, Python 等编程接口，为点云处理提供了全面的支持.

3）PyVista. 基于 Python，提供强大的3维数据可视化功能，可用于加载、处理和可视化大规模点云数据，有助于研究者更直观地理解点云数据.

4）LAStools. 专为 LiDAR 数据开发，支持大规模 LiDAR 点云数据的压缩、滤波、分类和特征提取等操作，广泛用于地理信息系统和地图制图.

这些点云处理工具包集成了丰富的工具和方法，可以极大地简化点云数据处理的流程，并加速点云相关算法的实现和研究.

3.1 静态点云处理

图信号处理对点云底层处理模型提供了强有力的理论和技术支持. 对于静态点云的特征捕获问题，Zeng 等人[86]将基于图像块的低维流形模型扩展到静态点云中的表面块上，采用了基于块的图拉普拉斯正则化算子近似计算连续域的流形维数，提出了一种新的离散块距离度量方法来量化2个相同大小的表面块之间的相似性，基于 k-NN 在相似块之间的构造图结构，采用图拉普拉斯正则化算子加速计算迭代. 同样基于图信号处理理论，一种基于谱域特征的滤波算法[90]利用加权图来表示 LiDAR 点云，将机载 LiDAR 数据作为图节点的不规则信号，从而将谱域图信号整合到滤波过程中，根据谱特征来过滤地面点和非地面点.

针对具体场景，通过构建图结构进行匹配或信息变换的算法同样行之有效. 针对 TLS 获取的大跨度、高分辨率复杂岩体表面点云的配准问题，Wang 等人[91]提出基于 n 点完全图的配准算法，采用分步策略来降低计算中所涉及的点数：首先对原始点云中所有点计算高斯曲率，过滤掉曲率较小的点，得到初步筛选出的关键点；其次聚类这些关键点并计算出每个簇的质心；最后根据每个质心及其 $n-1$ 个近邻构建出 n 点完全图作为局部特征描述算子. 将2个点云生成的描述子的对应点对计算相似度系数，然后根据系数值排序匹配点对，由此实现点云的2视配准. 一种新的3维点云颜色属性压缩算法[84]基于图变换来关联属性，对预编码的几何信息构建底层图结构来预测编码点，并利用图变换对预测残差进行变换.

3.2 动态点云处理

本文在第2节介绍了动态点云的获取方式和数据内容. 由于近年来3维动态点云在运动3维物体研究中的重要地位，对蕴含庞大数据量的3维动态点云的处理成为了新的研究前沿. 动态点云由一系列静态点云组成，每个静态点云由1组离散点组成，用3维坐标表示几何信息. 动态点云的每一帧指的是某一时刻瞬间捕获的点云，其每一帧包含的点数可能不相同，每帧的点之间没有随时间的明确关联. 3维动态点云包含了时间顺序，能够连贯地表示物体移动过程，可广泛应用于3维动画、3维游戏、3维虚拟现实技术、自主车辆导航等领域. 然而，除数据量大、点空间分布不规则这些静态点云研究面临的问题以外，3维动态点云每帧的点数随时间变化而变化，给点云的压缩编码带来了更严峻的挑战.

动态点云压缩可预先对物体进行运动估计和运动补偿，以促进动态点云的有效压缩. Thanou 等人[52]针对3维动态点云几何和颜色属性的压缩，利用点云序列中的时间相关性，提出了运动估计和补偿算法. 他们利用 k-NN 构造连通图，坐标和颜色属性作为节点的图信号，通过 GFT 将节点属性在谱域表示. 而且提出一个重要假设，即因为3维动态点云中连

续时刻的帧是相关的,则连续时刻的图信号也是相关的.因此,可通过比较相邻时刻图顶点上的信号来处理时间相关性,将点云序列表示为一组图,则动态点云上的运动估计问题随之等价为动态图中的特征匹配问题.

动态点云由于其时间和空间的不规则采样模式,去噪任务面临着更大的处理困难.Hu 等人[87]考虑到底层表面(流形)的时间一致性,基于时空图来表示动态点云.他们首先利用最远点采样法(farthest point sampling, FPS)采样块中心点,并搜索中心点的 k-NN 构成一个局部表面块,然后搜索邻接块构造图结构.在时间域上,他们假设图算子与黎曼流形在泛函上有着离散对应关系,通过定义流形间距离及其离散对应关系,从而度量时间域上表面块之间基于变分的内在距离.在空间域上,他们基于时序距离构建相应表面块之间和相邻块的点之间的时空图连通性.利用初始图表示,他们将动态点云去噪定义为点云和底层图表示的联合优化,同时通过空间平滑和时间一致性进行正则化.

动态点云修复方面,时空图相关技术举足轻重.基于可学习时空图表示,Fu 等人[88]提出一种动态点云修复算法,在时序对应相邻帧搜索相同的底层表面,通过时序图中边的权值表示它们之间的关系.该算法利用 2 阶帧间一致性和帧内自相似性,将动态点云修复问题定义为点云与底层时空图的联合优化问题,并通过时序边权重一致性和空间域平滑性对其进行正则化,并对此设计了一种交替最小化优化算法求解.利用时空图结构,Xu 等人[89]对动态点云的属性编码问题提出了针对性的压缩算法.该算法通过 k-NN 构造时空图,利用 GFT 得到谱域信号表示,然后结合高斯马尔可夫随机场模型设计了一种最优内预测方法,进而使用精细运动估计预测变换编码来压缩点云的属性特征.

3.3 点云底层处理的挑战和机遇

总体来看,静态点云的处理算法主要利用图结构编码点的几何信息以及构建属性特征之间的关联.对于动态点云,每帧的点云所含点数不同给底层处理算法带来了更大的困难,一种处理途径是将不规则点数通过体素化下采样并构建图结构来转化为图的特征匹配问题;另一种则通过时空关系构建时空图,利用时空图上的预测优化模型来解决运动估计问题.现阶段,静态点云和动态点云的底层处理仍较多地依靠对点云的体素化来应对海量点的问题[92].然而,体素化虽然能够大幅度降低点的规模,但不可避免地造成了离散化损失.此外,对点云的体素化表示仍旧耗费大量的存储空间.

目前,点云各任务处理领域的数据集都存在不平衡性,即数据集涵盖的目标和场景都有一定偏向,不足以均衡覆盖多种场景和目标.此外,遥感领域可用的点云数据集往往具有大量噪声,数据完整度受限.数据集的质量极大地影响了点云高层次识别任务,如分类、分割和目标检测的最终精度、鲁棒性和实践场景.然而,目前点云识别研究往往忽视了底层处理,使得现有的点云分类和分割方法在遥感领域数据上的进步受到极大限制.结合底层处理对大规模复杂场景点云数据进行高层识别任务,是未来 GPC 发展的重要趋势.例如,大规模场景点云常常存在数据缺失导致的孔洞,极大地影响了点云语义分割算法的精度.如何有效结合点云修复方法进行分割,仍处于探索阶段,是进一步提高点云分割算法性能未来发展的关键.

4 基于图的点云表示学习

点云表示学习又称为点云表征学习或点云特征学习,是点云数据中层处理的核心任务,它是贯通点云底层数据处理和高层识别任务的关键渠道.点云表示学习旨在抽取无序点云数据的通用表征,以便用于下游处理任务,如分类、分割、检测等识别任务.早期的此类算法也称为特征工程或特征提取,依赖于经典的统计分析方法[93-95].在深度学习兴起后,基于深度神经网络学习点云特征的嵌入模型或算法成为点云表示学习的特指[14,36,96].高质量的点云嵌入有助于提高下游任务的精度,且有利于迁移至多个不同任务场景下使用.

鉴于图结构能充分保留点云数据信息,图表示学习模型及算法为点云表示学习提供了有力支持.基于 GNN 的点云表征学习,最早可追溯到用图结构表示的不规则几何数据的特征学习[97-98].构造图结构来提取点云的局部特征是早期 GPC 特征描述的关键.例如,在处理 3 维点云时,Chen 等人[47]使用图过滤器和基于图的重采样策略设计特征描述算子,来选择信息最丰富的 3 维节点.近年来,深度学习在图结构上的学习取得了重要突破,图卷积、图生成对抗、图编码器等多种 GNN 模型闪耀登场,图结构在表示点云的非结构化信息上的强大潜力再一次受到重视.越来越多的点云表示学习模型开始采用 GNN 提取点云特征.表 3 列出了本领域的代表性模型,其出现

Table 3 Point Cloud Representation Learning Models

表 3 点云表示学习模型

模型	来源	核心思想	图技术	训练	数据集	下游任务	复杂度	推理时间/ms
LocalSpecGCN[40]	ECCV 2018	谱域 GCN 捕获局部信息及递归聚类池化	图卷积	有监督	③⑥⑦	分类、部件分割、语义分割		
DGCNN[58]	TOG 2019	动态更新每层特征空间对应的最近邻图	图卷积	有监督	③⑤⑥	分类、部件分割、语义分割	$O(KN)$	27.2
DeepGCNs[36]	ICCV 2019	引入残差、密集连接和扩张卷积的深层 GCN	图卷积	有监督	⑤	语义分割		
Hassani 等人[99]	ICCV 2019	多尺度多任务无监督编码	图编码	无监督	③⑥	分类、部件分割		
Chen 等人[100]	TIP 2020	用图拓扑作辅助信息设计解码器	图编码	无监督	①②③④	重建、可视化、转移分类		
Grid-GCN[37]	CVPR 2020	基于网格空间提高 GCN 的查询和上下文聚合速度	注意力	有监督	③④⑤⑦	分类、语义分割		26.2
Valsesia 等人[101]	TMM 2021	从生成器特征构造最近邻图来执行 GCN	图生成	无监督	①	上采样、分类		
Graph-PBN[102]	GM 2022	PointNet 和 GCN 并行分支结构	注意力	有监督	③⑤⑥	分类、部件分割、语义分割	$O(KN)$	

注：数据集：①ShapeNet②ShapeNetCore③ModelNet40④ModelNet10⑤S3DIS⑥ShapeNet Parts⑦ScanNet；$O(KN)$ 表示模型为 N 个中心点获取 K 个邻居点的时间复杂度，推理时间为模型在 ModelNet40 数据集上的一个 epoch 内所需的前向传播时间.

的时间线形式如图 10 所示.

4.1 基于图卷积的模型

图卷积操作在提取点云局部特征方面极具成效，是目前 GPC 表示学习的研究热点. Wang 等人[40] 借助图谱域上图信号的处理方法，设计了在谱域提取点云局部信息的卷积算子 LocalSpecGCN. 该模型对每个点计算 k-NN 并组成局部邻域点集 $V=\{x_1,x_2,\cdots,x_k\}$，计算逐点对距离（坐标之间的欧氏距离或特征空间中的距离）来构建权重为 $W_{ij}=dist\left(x_i,x_j\right)$ 的边集合 $E\subseteq V\times V$，由此构建局部表示图 $G=(V,E,W)$，并定义谱域上的图卷积为

$$\tilde{h}=\tilde{\boldsymbol{X}}\odot\tilde{g}, \tag{1}$$

其中 $\tilde{\boldsymbol{X}}$ 代表输入点集特征矩阵经 GFT 后的形式，$\tilde{g}$ 代表图傅里叶域中的滤波器，$\tilde{h}$ 代表滤波后的输出，$\odot$ 表示哈达马积（Hadamard product）. 经过逆傅里叶变换后得到原始空间域中的卷积输出 h 后，组合点集特征抽象操作定义为

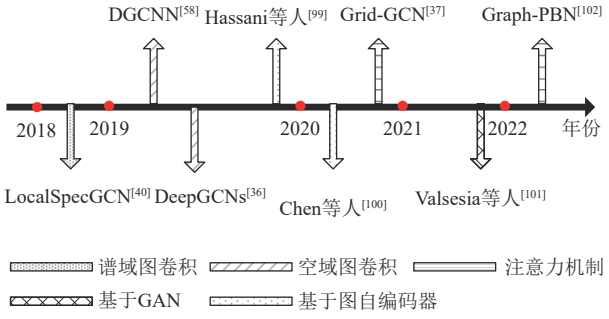


Fig. 10 Chronological overview of point cloud representation learning models

图 10 点云表示学习模型的时序总览

$$f(x_1,x_2,\cdots,x_k)=\oplus(h_1,h_2,\cdots,h_k), \tag{2}$$

其中 h_i 是在点 x_i 的卷积输出 $h(x_1,x_2,\cdots,x_k)$，$\oplus$ 为基于分层聚类和簇内池化提出的新型集合激活函数，本质上是一种 k-NN 图池化算法，旨在总结来自 k-NN 图的信息.

空域图卷积也越来越多地用于点云表示学习中. 谱域图卷积虽然有充实的图谱理论和卷积定理作为理论基础，但必须将数据由空域转换到谱域做处理. 而空域图卷积不依靠图谱卷积理论，直接在空间上定义卷积操作，灵活性更强. 由于面临的具体问题不同，空域图卷积思想在点云任务处理上发展出多种延伸和变体. Wang 等人[58] 提出动态图 CNN 架构（dynamic graph CNN, DGCNN），在每一层生成的特征空间中使用最近邻重新计算图来学习语义分组点. 该模型将初始图构造为特征矩阵 $\boldsymbol{X}\subseteq\mathbb{R}^F$ 的 k-NN 图 $G=(V,E)$，该图包含自环，即每个节点也指向自己. 由此边特征定义为 $e_{ij}=h_\Theta\left(x_i,x_j\right)$，其中 $h_\Theta:\mathbb{R}^F\times\mathbb{R}^F\to\mathbb{R}^{F'}$ 是具有一组可学习参数 Θ 的非线性函数. 由此该模型设计了一种新的卷积操作 EdgeConv，定义为

$$x_i'=\blacksquare_{j:(i,j)\in E}h_\Theta\left(x_i,x_j\right), \tag{3}$$

其中 ■ 选择适合的聚合函数，譬如求和操作，而 $h_\Theta()$ 选用非对称边缘函数，定义为

$$h_\Theta\left(x_i,x_j\right)=\tilde{h}_\Theta\left(x_i,x_j-x_i\right), \tag{4}$$

这使得在卷积的过程中能够将中心点的全局特征 x_i 与局部邻域特征 x_j-x_i 相结合，更好地捕获点云的全局与局部信息. 在神经网络的每一层，都获得一个更新的图 $G^l=\left(V^l,E^l\right)$，其中第 l 层的边为每个点 x_i^l 在该层特征空间内的 k-NN 相连组成的集合. 这样，在保

持排列不变性的同时, 该模型不仅可以学习如何提取局部几何特征, 还可以学习特征空间中对点的不同分组. 然而, DGCNN[58] 无法避免梯度消失的问题, 即所提取的节点深度特征与邻居节点特征趋于同质化, 不能提供有价值的边缘向量. 为缓解 DGCNN[58] 的这一缺陷, LDGCNN[103](linked dynamic graph CNN)对各动态图之间的层次特征构建链接, 使用 k-NN 搜索节点邻域并构建局部有向图结构, 然后通过图卷积聚合局部特征.

当考虑更深层的 GCN 时, GCN 中动态变化的邻域有助于缓解过度平滑问题并有效扩大感受野. DeepGCNs 模型[36] 通过每层特征空间中的扩张 k-NN 函数重新计算顶点之间的边, 以进一步增加感受野. 在该模型中, 通过聚合上一层相邻顶点特征得到的当前层中心点的表示定义为

$$h_{v_{l+1}}=\varphi\left(h_{v_l},\rho\left(\left\{h_{u_l}|u_l\in N(v_l)\right\},h_{v_l},W_\rho\right),W_\varphi\right), \tag{5}$$

其中 ρ 是顶点特征聚合函数, φ 是顶点特征更新函数, h_{v_l} 和 $h_{v_{l+1}}$ 分别是第 l 层和 $l+1$ 层的顶点特征. $N(v_l)$ 是 v 在第 l 层的相邻顶点的集合, h_{u_l} 是由 W_ρ 参数化的那些相邻顶点的特征. W_φ 包含这些函数的可学习参数. 为简单起见且不失一般性, 该模型使用 $\rho=\max(h_{u_l}-h_{v_l}|u_l\in N(v_l))$, 即无可学习参数的最大池化顶点特征聚合器来池化邻域内特征, 将具有批标准化(batch normalization)和 RELU(rectified linear unit)激活函数的多层感知机(multi-layer perception, MLP)作为顶点特征更新器 φ.

为训练更深层次的 GCN, 残差、密集连接和扩张卷积被引入, 分别如式(6)(7)所示. 残差映射 $\mathcal{F}$ 学习将一个图作为输入, 并为下一层输出一个残差图表示 G_{l+1}^{res}, W_l 是第 l 层的可学习参数集, 由此通过拟合映射 $\mathcal{F}$ 来学习底层映射 $\mathcal{H}$. 对于密集融合, 算子 $\mathcal{T}$ 是一个顶点级联函数, 它将输入图 G_0 与所有中间 GCN 层输出密集融合, 使得 G_{l+1} 包含来自先前层的所有 GCN 转换.

$$G_{l+1}=H(G_l,W_l)=\mathcal{F}(G_l,W_l)+G_l=G_{l+1}^{\text{res}}+G_l, \tag{6}$$

$$\begin{aligned}G_{l+1}=&\mathcal{H}(G_l,W_l)=\mathcal{T}(\mathcal{F}(G_l,W_l),G_l)=\\&\mathcal{T}(\mathcal{F}(G_l,W_l),\cdots,\mathcal{F}(G_0,W_0),G_0).\end{aligned} \tag{7}$$

最后, 对于扩张卷积, 该模型使用扩张 k-NN, 通过跳过每 d 个邻居来返回顶点 v 在 $k\times d$ 邻域内的 k-NN, 构成顶点 v 的 d-扩张邻域 $N^{(d)}(v)=\{u_1,u_{1+d},u_{1+2d},\cdots,u_{1+(k-1)d}\}$, d 称为扩张率. 该模型在每个 GCN 层之后找到扩张的邻居并构造一个扩张图. 以上框架共同构建了一个深度达 56 层的图卷积神经网络, 在点云语义分割上取得了比浅层网络更好的性能.

为提高 GCN 模型在点云学习和训练上的计算效率, 针对图结构上的采样查询方法, 如 FPS 和 k-NN 邻居点查询方法对模型训练效率的局限, Grid-GCN[37] 方法结合栅格(grid), 提出了栅格空间上的结构化查询策略以及新的 GCN 聚合模块. 与搜索空间是所有点的传统 k-NN 不同, 该模型提出的 k-NN 只需要在上下文点之间进行搜索, 使得查询速度大大加快. 在查询到局部点群后, 该模型设计了 GCA 模块来聚合节点特征. GCA 模块首先构造一个局部图 $G=(V,E)$, 其中 V 由点群中心节点及其 k 个群成员节点组成, E 由每个节点连接到中心节点构成. GCA 模块的定义为

$$\tilde{f}_{c,i}=e(X_i,f_i)\times\mathcal{M}(f_i), \tag{8}$$

$$\tilde{f}_c=\mathcal{A}\left(\left\{\tilde{f}_{c,i}\right\},i\in 1,2,\cdots,k\right), \tag{9}$$

其中 $\tilde{f}_{c,i}$ 是节点的贡献, X_i 是节点的 xyz 坐标. $\mathcal{M}$ 是 MLP, e 是边注意力函数, $\mathcal{A}$ 是聚合函数. 由此, GCA 模块将节点的特征 f_i 投影到 $\tilde{f}_i$, 根据节点与中心的边关系计算出 $\tilde{f}_i$ 的贡献并将所有这些特征聚合为中心的特征 $\tilde{f}_c$.

注意力机制被用来优化基于图卷积技术的 GPC 表示模型的性能, 使网络聚合过程聚焦于相对重要的邻域信息. Grid-GCN[37] 设计一种基于注意力的 GCA 模块来提取栅格空间邻域的上下文特征, 用覆盖权重和栅格上下文池化定义边注意力函数为 $e=mlp(mlp_{\text{geo}}(X_c,X_i,w_i),mlp_{\text{sem}}(f_{\text{cxt}},f_i))$, 同时捕获上下文语义特征以及地理位置信息. Cui 等人[104] 提出了用于点云分析的几何注意力动态 GCN, 将几何注意力边缘卷积模块引入 DGCNN 的 EdgeConv 中, 以捕捉点云的内在特征似然, 提高网络对几何表示的学习能力. 基于 GCN 的强大表征能力, Graph-PBN 模型[102] 利用注意力机制设计了一种新的卷积算子 EAGConv, 用于收集局部的几何上下文 k-NN 信息.

4.2 无监督模型

图自编码器、图生成对抗网络等无监督 GNN 模型逐渐成为在无监督任务场景下点云表示学习研究的一片新的蓝海. 为解决点云的不规则结构问题, FoldingNet[105] 架构设计了一个基于图的编码器和一个基于折叠的解码器. 该编码器是 MLP 和基于图的最大池化层的串联, 且根据输入点集的坐标构造 k-NN 图. 令图的邻接矩阵为 $\boldsymbol{A}$, 图编码层的输入矩阵为 $\boldsymbol{X}$, 那么输出矩阵是 $\boldsymbol{Y}=\boldsymbol{A}_{\max}(\boldsymbol{X})\boldsymbol{K}$, 其中 $\boldsymbol{K}$ 是特征映射矩阵, 且矩阵 $\boldsymbol{A}_{\max}(\boldsymbol{X})$ 的第 i 行第 j 列的输入是

$$(\boldsymbol{A}_{\max}(\boldsymbol{X}))_{ij}=RELU\left(\max_{k\in N(i)}x_{kj}\right), \tag{10}$$

其中局部池化操作本质上基于图结构捕获局部邻域特征. 面向非监督多任务情况下的特征提取与表示学习, Hassani 等人[99]设计了基于多尺度图的编码器, 其总框架由 1 个编码器和 3 个特定任务的解码器组成. 编码器通过在 3 个邻域半径上应用图卷积来提取 3 个中间特征, 所用图卷积定义为

$$z_i=\sum_{p_k\in N(p_i)}h_\Theta\left([p_i\|p_k-p_i]\right), \tag{11}$$

其中z_i是点p_i基于其邻居贡献的学习特征, $p_k\in N(p_i)$是欧几里得空间中离p_i最近的k个点, h_Θ是由θ参数化的非线性函数, $\|$是连接算子, 且对h_Θ使用共享 MLP. 同时使用p_i和p_k-p_i的原因是对每个点的全局信息(p_i)和局部交互(p_k-p_i)进行编码. Chen 等人[100]在早期工作 FoldingNet[105]的基础上, 设计基于图的解码器, 包括图拓扑推理模块和图滤波模块, 利用可学习的图拓扑来推送码字以保持代表性特征.

GAN 在图像生成对抗方面取得巨大成效, 使之成为目前最流行的无监督生成模型, 采用图结构表示点云并结合 GAN 的点云表示模型随之涌现. Achlioptas 等人[96]对一些生成点云的 GAN 架构进行了初步研究. 点云上的生成架构需要解决输入输出对应问题, 即输入顶点和输出顶点之间缺乏规范顺序. 在 3 维网格生成中, 通常假设网格是给定的, 其顶点是规范有序的. 生成问题仅仅是确定网格顶点的嵌入. 不过, 其生成器网络是由密集连接的层组成的, 无法生成局部的特征表示. 考虑到局部特征表示的生成有助于捕获有用的先验知识的数据分布, Valsesia 等人[106]研究如何将图卷积生成模型用于无监督问题, 给出了在生成器输出未知图的条件下图卷积的定义, 提出的生成框架能够输出近似点云的图嵌入表示. 同样针对点云上局部结构捕获和表示问题, Valsesia 等人[101]提出了一种使用图卷积层生成 3 维点云的 GAN 模型. 该网络模型从生成器特征构造最近邻图来实现图卷积操作, 以此捕获局部化特征并输出几何图形的图嵌入向量.

4.3 点云表示学习的挑战和机遇

表示学习广泛用于 2 维图像的分析和理解任务中, 在传统计算机视觉领域的研究已经颇具成效[58]. 然而, 点云的表示学习更具挑战性, 因为它是 3 维空间中不规则分布的点的无序集合[101]. 一方面, 点云是不规则分布在几何空间上的, 应用传统的卷积操作就需要将点云数据转换为体素, 再使用基于栅格的方法来处理和学习 3 维空间信息, 这将不可避免地导致离散化损失[36]. 基于点云数据构建图网络, 并使用图卷积操作捕获点云的邻域特征, 才能最大程度保留原始点云的所有信息. 另一方面, LiDAR 的扫描密度非常大, 通常生成的点云数据的点数量级均以万甚至以千万计, 其存储成本以及后续计算成本都随之激增. 如何有效地保留丰富信息的同时, 快速提取出优质的点云表示向量, 是点云表示学习模型面临的另一挑战. 截至目前, 点云表示学习没有专门的测试数据集和评价指标. 验证模型学习到的点云表示是否有效, 往往通过其在多个下游任务上的表现来体现. 此外, 模型是否能够易于内置到其他神经网络模型中, 也是衡量一个点云表示模型是否具有良好的可扩展性的重要因素.

5 基于图的点云识别

点云识别任务包括分类、分割、目标检测等. 3 维点云分类属于粗粒度识别任务, 只需要识别点云数据中的判别特征, 给出正确的目标类别或者物体形状分类. 3 维点云分割属于细粒度、点级识别任务[43], 根据点云的属性将其分为多个子区域, 且每个相邻区域的特征各不相同, 而区域内部点云数据具有相同的特征[107]. 根据所分割的对象的不同, 可分为一个室内外场景或单一完整物体. 根据分割任务的不同可分为场景语义分割[43]、部件分割[108]和实例分割[109]. 3 维点云数据可描述物体在空间中的位置、形状、大小和其他属性, 所含信息远比 2 维平面图像多. 3 维点云数据宏观上具有物体形状多样、稀疏度不均匀、采样不规则等特点[110], 微观上具有无序性[27], 点云分割和分类的困难性成为研究者的关注热点.

3 维点云目标检测的目的类似于 2 维图像领域的经典任务, 旨在给出某个物体的边界框(bounding box)和类别[111]. 区别在于输入从 2 维图像变成了 3 维点云, 输出的边界框参数数量从 4 个增加至 6 个[112]. 点云数据通过 LiDAR 扫描得到, 是目前 3 维目标检测的重要数据. 与传统 2 维图像相比, 用 3 维点云做检测和跟踪有传输快、存储少、信息保留丰富的优点. 一方面, LiDAR 极为精准的距离测度使得生成的点云数据充分保留 3 维空间中物体的准确几何信息, 有利于准确评估距离和方位. 另一方面, 点云数据相比 2 维图像受光照天气等环境因素的影响程度更小, 在车载、机载设备识别检测时更能够保证准确度, 为自动驾驶和地理测度等工业需要提供了可靠保障.

然而, 如何充分挖掘点云保留的信息仍然是一个开放性问题, 3 维点云的目标检测给自动驾驶、自主导航以及虚拟现实领域带来了新的发展和挑战[13,110]. 点云与图类似, 都属于不规则分布结构. 将规则图像上的卷积操作应用到点云数据, 从而对点云数据进行深度学习, 始终是当前 GPC 研究的焦点.

5.1 基于图的点云分类

点云分类包括对单个物体形状分类和从室内外场景中判定物体类别. 与图像的分类处理类似, 3 维点云分类通常借助一系列卷积操作提取局部细节特征, 通过一个全局池化层提取全局语义特征, 最终给出分类的预测结果[107]. 衡量单目标类别的识别精度包含查准率和精确率 2 个指标, 若存在多个目标类别, 需要平均精度(mean accuracy, mAcc)指标, 即每个类的每个精度的平均值(预测的每个类的精度总和与类的数量的比值), 以及总体精度(overall accuracy, OA)指标, 即正确预测的项目数与要预测的项目总数之比. 虽然 3 维点云数据的分类方法研究已经具有很长的历史, GPC 相关分类研究相对较新, 主要通过图结构表示点云数据, 利用图卷积算子聚合以及提取特征信息. 表 4 列出了代表性模型的核心信息以及在 ModelNet40 和 ModelNet10 上的表现, 采用的评价指标为 OA 和 mAcc. 所列模型的时间事件展现在图 11 中.

鉴于图结构与点云都具有不规则性, 图卷积捕获点云局部特征具有很大优势, 利于点云形状的分类识别. Simonovsky 等人[35] 将图卷积引入到点云数据的分类中, 在空域上定义了利用边标签的局部图邻域卷积算子——边条件卷积(edge-conditioned convolution, ECC), 设计出一个用于图分类的神经网络模型. 该研究不拘泥于传统的图网络测试数据集, 而是将同样不规则的点云数据视为图, 来验证提出的卷积模型对不规则信息的学习能力. 他们在欧几里得空间中直接构建点云的图结构, 将点云分类转化为网络上的图分类任务, 首次提出了不需要体素化而直接使用卷积操作的点云处理方案. DPAM[115] (dynamic points agglomeration module)是一种轻量级的基于图卷积的动态点聚集模块, 将点的相似度图与特征矩阵作为模块的输入, 通过 GNN 传递信息, 实现了语义空间内点的动态凝聚. Guo 等人[117] 设计了基于图的局部特征提取器, 对全局和局部特征分配相应的权重以实现自适应特征融合. 为减少模型对样本标注的依赖, Fan 等人[118] 提出基于图的 One-Shot 点云识别模型, 包括编码器、边条件 GCN 和查询模块.

谱域图卷积同样是 GPC 分类的重要技术. Zhang 等人[38] 提出了 PointGCN 模型, 构造基于高斯核函数的 k-NN 有权图, 定义谱域上切比雪夫多项式作为卷积滤波核, 为点云数据设计了 2 种不同感受野的池化层, 将其与局部图卷积相结合, 以捕获点云的局部与全局特征. 为实现几何变换不变性, 3DTI-Net[113] 在欧氏空间内计算 k-NN 并构造图的邻接距离矩阵, 并转换至谱域的随机游走归一化拉普拉斯矩阵进行卷积.

除了采用图卷积来捕获点云特征外, 还可以将图结构只作为信息传递的局部环境. KCNet 模型[59] 计算各点的 k-NN 图, 以此作为局部几何结构计算核相关响应, 然后在 k-NN 图上重复聚合特征作为图最大池化得到局部高维特征. 为严格确保旋转不变性, ClusterNet 模型[114] 在 k-NN 图上使用 RRI 算子来获取点在邻域下的相对位置模式, 然后通过 EdgeConv 模块将子簇的特征在 k-NN 邻域上聚合获得超簇特征. 单铉洋等人[116] 利用 k-NN 计算中心点的邻居, 构建局部结构之后采用 CNN 进行信息聚合.

Table 4 Comparison of Point Cloud Shape Classification Model on ModelNet10/40 Dataset

表 4 点云形状分类模型在 ModelNet10/40 数据集上的对比

模型	来源	核心思想	图技术	ModelNet40		ModelNet10		参数量×10^6	推理时间/ms
				OA/%	mAcc/%	OA/%	mAcc/%		
ECC[35]	CVPR 2017	在各自边标签上控制每个过滤权重	图卷积		82.4		89.3		
PointGCN[38]	ICASSP 2018	局部化卷积与不同类型的池化结合	图卷积	89.51	86.05	91.91	91.57		
KCNet[59]	CVPR 2018	内核关联和图池化提高 PointNet 局部化性能	k-NN 图、图池化	91.0		94.4		0.9	12.0
3DTI-Net[113]	Arxiv 2018	变换不变性特征编码器	图卷积	91.7				2.6	60
ClusterNet[114]	CVPR 2019	引入分层聚类获得点局部旋转不变性特征表示	k-NN 图	87.1					
DPAM[115]	ICCV 2019	分层以及参数共享的动态点聚合模块	图卷积	91.9	89.9	94.6	94.3	29.5	36.6
RFNet[116]	自动化学报	局部邻域结构和 CNN	k-NN	94.0	91.4			2.36	

注: 推理时间为模型在 ModelNet40 数据集上的 1 个 epoch 内所需的前向传播时间.

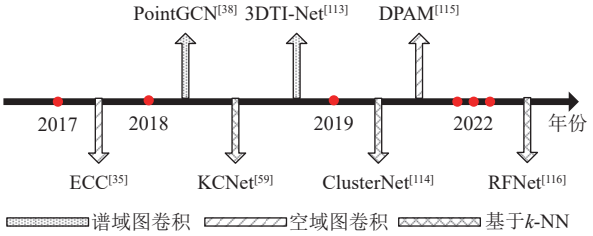


Fig. 11 Chronological overview of point cloud classification models

图 11 点云分类模型的时序总览

针对具体应用场景的分类模型设计同样重要. 基于机载激光扫描场景得到的点云数据的语义分类任务，由于该场景中的物体在垂直方向上重叠，茂密树林和山地等复杂区域空间密度不均匀，数据涉及区域面积广阔导致扫描点数量成千甚至上亿，导致对该场景点云数据的分类存在很大困难. 基于此，Chen 等人[46]提出了一个简单的全卷积网络模型，在 DGCNN 的基础上，将多尺度局部区域内的所有点对连接起来，全面提取上下文信息. 此外，针对在 DGCNN 网络中模型的邻居点之间没有相关性的问题，Chen 等人设计了 Webconv 模块，将局部区域内的点紧密地连接在一起.

分辨率受限或物体被遮挡情况下高效准确地捕获重要物体，一直是点云分类任务上的挑战之一. 而且风的扰动所产生的噪声和杂波也阻碍了提取目标物体的判别特征. Li 等人[119]提出了一种分步模型来对大规模城市环境点云数据中的建筑、树木、行人和汽车进行分类：首先为每个 3 维点构造一组轻量级描述符，有效表征每一类物体的形状和颜色；分别采用全局和局部优化算法细化每个点类别的标注结果，其中全局优化主要依靠图分割技术，使标签具有空间局部连续性以及全局可优化，而局部优化则基于弱监督的目标决策树以提高分类精度.

5.2 基于图的点云分割

与分类只需输出判定类别不同，点云语义分割需要对一个场景进行全面的理解[43]，寻求为每个 3 维点赋予一个语义分类标签. 表 5 列出了代表性模型的核心信息以及在 ShapeNet part, S3DIS Area5, Semantic3D reduced-8, ScanNet 数据集上的表现，采用的评价指标有 OA, mAcc 和平均交并比（mean intersection over union, mIoU），其中交并比（IoU）直观地衡量语义分割的真实值和预测值之间的交集与并集之比，而 mIoU 是每个类上的 IoU 的平均值.

GPC 分割方法起源较早，采用非机器学习的传统图算法对点云进行初始粗粒度分割[124-126]. 为去除冗余背景点，Pan 等人[126]提出一种基于图分割的多视图重构点云自动分割算法，通过连接每个点到它的 k-NN 构造加权图，分割目标和背景点初始 2 值后，迭代执行图分割优化方法更新图结构. 针对道路场景中的遮挡和重叠问题，Luo 等人[127]构建基于局部块的匹配图结构，结合马尔可夫随机场将有标注的点云道路场景中的类别标签传输到无标注的点云道路场景中. Xu 等人[45]基于图分割算法将连通的相邻八叉树体素聚合到一个全连通的局部亲和图（affinity graph）中，以此从建筑工地点云中提取线、平面和圆柱等几何基本体. 过分割利用以颜色信息为主的点云 3 维信息生成点的聚类，记为超点. 文献 [128] 考虑了不同的方法来构造连通图、分配边权值以及定义

Table 5 Comparison of Point Cloud Segmentation Models on the Main Semantic Segmentation Datasets

表 5 点云分割模型在主要的语义分割数据集上的对比

模型	来源	核心思想	图技术	ShapeNet part	S3DIS Area5		Semantic3D reduced-8		ScanNet		参数量×10^6	推理时间/ms
				mAcc/%	OA/%	mIoU/%	OA/%	mIoU/%	OA/%	mIoU/%		
SPG[34]	CVPR 2018	利用超点图和图卷积学习点的语义标签	图卷积		86.38	58.04	94	73.2				
SSP+SPG[39]	CVPR 2019	结合图划分和图卷积提高分割精度	图卷积		87.9	61.7						
GACNet[43]	CVPR 2019	可学习核作为图注意提取局部特征	图卷积		87.79	62.85	91.9	70.8				
HPEIN[120]	ICCV 2019	层次化点边交互网络	FPS，k-NN		87.18	61.85				61.8		
FGCN[121]	CVPRW 2020	端到端 3 层局部 GCN	图卷积	63.22		52.17		62.40				
TGNet[122]	TGARS 2020	泰勒高斯几何关系捕捉	图卷积		88.5	57.8					8.54	
PGCNet[123]	TVC 2020	表面块加速卷积以及层次型聚合特征	图卷积		86.24	63.85			83.9			
HAPGN[108]	TMM 2021	多头注意力聚合与分层提取特征	图卷积、注意力	89.3	85.8							154.6

注：推理时间为模型在 ModelNet40 数据集上的 1 个 epoch 内所需的前向传播时间.

颜色和几何信息上的合并准则,提出一种过分割算法用于提取超点.

利用图卷积捕获和学习点云局部特征,实现对点云的语义标签预测成为近年来点云分割研究的重要趋势之一. Landrieu 等人[34]提出 SPG 模型,通过有向属性图构造点云的超点图来语义分割. SPG 模型将大尺度点云分割问题划分为几何同质分割、超点嵌入、上下文分割 3 个子问题,分别用来计算超点图、降采样和利用门控 GCN 产生点的语义标签. RGCNN[129](regularized graph convolutional neural network)基于点坐标和法线构造图结构,设计了高阶切比雪夫 3 层图卷积神经网络,在损失函数中添加图信号平滑先验来正则化网络的每一层. 为了自适应捕获动态图的结构,RGCNN 在网络的每一层更新图拉普拉斯矩阵. 为了改进划分步骤,Landrieu 等人[39]提出了一种有监督的超点过分割模型,通过构造邻接图将点云过分割问题形式化为图划分问题,引入了基于图结构的对比损失来提高超点边界的对比度. GACNet[43](graph attention convolution network)利用邻居节点构造图结构,将可学习核作为图注意应用于图卷积,提出了一种用于点云语义分割的端到端 GCN. FGCN(feature-based graph convolutional network)模型[121]通过高斯核加权 k-NN 构造图结构,利用切比雪夫多项式定义谱域图卷积来捕捉点的相对位置的局部特征. 针对空域卷积核对高维几何关系以及邻域坐标的捕获能力差,最大值或平均的聚合函数忽略了局部近邻的不同贡献度,导致点云分割性能不佳的问题,TGNet[122](Taylor Gaussian mixture model network)结合泰勒函数和高斯混合模型,设计了一个新的端到端几何 GCN 来提高空域图卷积对点云分割的精度. 为处理大规模点云,PGCNet[123](patch graph convolutional network)模型基于表面块下采样原始点云,动态构造边并应用图卷积来捕获和聚合特征. PointVGG[44]使用 k-NN 构造图结构和边特征,通过逐渐放大感受野来获得可观的局部几何信息,在图结构上学习并聚合点特征.

在层次化分割模型中,图结构起着重要作用. Jiang 等人[120]提出了一种层次化点-边交互网络模型 HPEIN(hierarchical point-edge interaction network),引入编码解码器迭代地更新点云图. 该模型通过 FPS 下采样点云,根据 k-NN 构造边集和初始图结构,设计边分支模块迭代提取边特征,以获取更高维的语义信息. 针对移动 LiDAR 产生的点云重叠区域的分割问题,Xu 等人[130]提出了一种新的最优分层聚类算法,利用 2 部图中的匹配来组合区域,通过基于点的图中的最小代价完美匹配来求解最优聚类组合. 为捕获中心点的各邻域点差异、利用点的空间关系来学习全局特征,Chen 等人[108]提出 HAPGN(hierarchical attentive pooling graph network)模型,利用门控图注意网络分配不同权值增强局部特征,以及利用分层图池模块充分捕获点的空间布局中的层次性信息.

5.3 基于图的点云目标检测

目标检测任务是更细粒度的识别任务,不仅要给出样本中目标的类别,还要给出目标的所在位置,即边界框,仅仅衡量类别预测的准确性是不够的. 为了衡量位置预测的准确性,常采用 mIoU 指标衡量所有类别下的总体准确率,其中利用 IoU 计算预测边界框与真实边界框的交集区域与并集区域之比. 预测时通常为 IoU 设置一个重叠阈值,若 IoU 值超过阈值,则视为位置预测正确,反之则认为位置预测错误.

非机器学习的传统图算法是早期 GPC 目标检测算法主要的理论和技术途径. Golovinskiy 等人[131]基于图分割算法提出了一个多类目标检测框架,对点云的图表示进行图分割,使用 SVM 标注特征向量,从而识别汽车和交通灯等物体,但这种方法需要手动定位. 此外还有图匹配检测定位目标的算法[132-134],以及通过图匹配和路径跟随算法求解关键点图的形状特征匹配算法[135].

深度学习在 3 维目标检测领域近年来成为主流. GNN 引爆非结构化信息深度学习的热潮,GCN 也随之成为 GPC 目标检测的研究热点. Shi 等人[112]基于图卷积提出了 Point-GNN 模型,通过连接节点及其固定半径内邻居点构造图结构,设计了一个具有自动配准机制的 GNN,从而预测图中每个顶点所属的对象的类别和形状,通过框合并和评分操作精确地组合来自多个顶点的检测. 该模型可以在一次拍摄中检测多个目标,表明了 GNN 在点云目标检测上的潜力. Chen 等人[136]基于图卷积提出了一个层级图网络 HGNet(hierarchical graph network)模型,采用 FPS 和 k-NN 实现类似于 FPN 模型的多级语义特征金字塔,利用 k-NN 构造图结构并设计了一种形状注意力图卷积来捕获目标的多层次局部形状语义信息,有效聚合生成的提案区,利用全连通图对提案区推理得到边界框. Zhang 等人[137]提出了一个 2 阶段方法,分别针对点云遮挡物体稀疏和局部点云问题,设计了点云补全和 AMS-GNN(attention based multi-scale graph convolution)两个模块. 基于图卷积设计的 AMS-GNN 模型对每个点及其 k-NN 之间构建连边,从而将提案

点云编码为图结构，通过局部-全局注意机制和基于多尺度图的上下文聚合全面捕捉点之间的关系，以应对点云目标检测中特有的同类实例离散分布差异导致的图节点所学特征对图分辨率和连接关系敏感的问题.

将其他机器学习模型与 GNN 模型融合是提高点云目标检测性能的有效途径. Yin 等人[138]针对多帧点云视频上的目标检测，提出一种新的 GNN 模型 GMPNet(grid message passing network)来编码短期点云. 该模型将每个栅格(即分组的点)视为 1 个节点，并与相邻的栅格构造 k-NN 图，迭代地挖掘相邻帧中不同栅格之间的关系，从而获取运动线索. 经过 GMPNet 提取和聚合短期数据的特征后，长期聚合模块为 ConvGRU 引入注意记忆机制，分别设计时间和空间转换器注意，从而挖掘长期点云的时空相干性. 最后，检测头接收增强的记忆特征并产生检测结果. 受关系推理和 GNN 在各领域的成功启发，Feng 等人[139]提出用 3 维目标-目标关系图结构显式建模提案之间存在的关系信息，以增强 3 维提案特征. 关系图的构造同时考虑了候选边界框的外观与位置 2 种关系，利用图卷积进行关系推理，回归出更精确的 3 维边界框，利用多图和监督图策略来提高关系图的性能. 此外，他们还探讨了外观和位置关系对关系图的不同影响，以及监督图和多图模式对 3 维关系推理的影响.

5.4 点云识别公开数据集

表 6 汇总了当前点云识别的常用数据集. 点云分类常用 2015 年普林斯顿实验室提出的 ModelNet 系列数据集，总共包含 662 类物体 127 915 个 3 维 CAD 模型. 由于数据集较大，又分出 2 个标注完整的子数据集 ModelNet10 和 ModelNet40，分别包含 10 个和 40 个物体类别. 由于点云分类属于粗粒度的识别任务，而分割则更为精细，所以服务于分割的 CAD 点云数据集也可以应用于点云分类任务.

表 6 中也列出了 7 个常用的分割数据集. 分割数据集又可以划分为语义分割数据集、实例分割数据集和部件分割数据集. 用于部件分割的数据集有 ShapeNet, ShapeNet part, S3DIS; SceneNN 可用于实例分割; SceneNet, ScanNet, Semantic3D 可用于语义分割. 其中 ShapeNet 和 ShapeNet part 的每个 3 维 CAD 模型都表示一个完整的目标物体，物体的各个不同部件具有不同的语义标签. S3DIS 是大规模的室内场景数据集，于 2017 年由斯坦福大学通过 Matterport 扫描仪扫描获得，各个室内场景的不同部分，如天花板、地板等标注了不同的语义标签. SceneNN 是 RGB-D 数据集，共有 100 个真实环境(包括卧室、厨房、客厅以及办公室等)的室内场景，场景中的所有实例各自具有不同的标签，SceneNet 包含 5 个类别共 57 个室内场景数据. ScanNet 是 RGB-D 数据集，包含 20 类共 1 513 个室内场景. Semantic3D 则是室外场景数据集，包含了村庄、街道、教堂等 8 个不同类型的城市户外场景，场景中每个点除坐标外还包括 RGB 值和强度值作为点属性.

Table 6 Common Datasets Used in Point Cloud Recognition

表 6 点云识别常用数据集

类型	数据集	场景	任务	类别数	年份	样本数
分类	ModelNet10[140-141]	CAD	分类	10	2015	
	ModelNet40[140-141]	CAD		40	2015	12 311
	ModelNet[140-141]	CAD		662	2015	127 915
分割	ShapeNet[142-143]	CAD	部件	55	2015	513 000
	ShapeNet part[142-143]	CAD	部件	16	2015	16 881
	S3DIS[144-145]	室内	部件	11	2017	271
	SceneNN[146-147]	室内	实例	4	2016	100
	SceneNet[148-149]	室内	语义	5	2016	57
	ScanNet[150-151]	室内	语义	20	2017	1 513
	Semantic3D[152-153]	室外	语义	8	2017	
目标检测	KITTI[18, 154]	室外	检测	8	2012	
	nuScenes[155-156]	室外		23	2019	40 000
	Waymo[157-158]	室外		4	2018	
	ONCE[159-160]	室外		5	2021	1×10^6

表 6 还列出了 4 个主要的室外目标检测数据集，分别是 KITTI, nuScenes, Waymo, ONCE 数据集. KITTI 数据集由德国卡尔斯鲁厄理工学院和丰田美国技术研究院联合创办，不仅是目前最常用的目标检测数据集，还是自动驾驶、语义分割等任务常用的基准数据集. nuScenes 数据集由 Motional 团队开发，包含有丰富的目标标注标签，是 KITTI 目标标注总量的 7 倍以上. Waymo 数据集由谷歌公司旗下的无人驾驶公司 Waymo 于 2019 年发布，用于目标检测的大型自动驾驶数据集，包含夜晚、黄昏、雨天和晴天等多种驾驶环境. ONCE 数据集由华为诺亚方舟实验室发布，是在中国驾驶 144 h 采集的车载数据集，包含了多种天气环境，标注了汽车、行人、公交车、卡车等多种目标类别.

5.5 点云识别的挑战与机遇

点云识别是点云研究中的高层处理任务，是点云表示学习重要的下游任务. 与点云表示学习面对

的问题类似，点云识别同样面临海量数据点、非结构化信息难以聚合等挑战. 具体来说，点云分类需要在粗粒度计算全局类别标签，因此，GPC 方法构造高效的局部信息聚合结构，实现快速有效提取全局特征. 其次，点云分割需要在细粒度计算每个点的标签，因此，GPC 方法利用超图、层次聚类以及 GCN 等细粒度耦合，有效提高计算效率、局部非结构化信息提取以及避免梯度消失和向量表示同质化. 最后，点云的目标检测需要在位置和类别上均取得较高的精度，因此，GPC 方法利用图分割以及 GCN 等高效划分和聚合策略，可有效提取局部位置信息以及物体类别信息. 然而，现有的 GPC 分类、分割和目标检测算法在构造图结构、定义卷积操作或聚合策略方面仍主要依赖 *k*-NN 等邻域结构，抑制了图策略的性能. 此外，GPC 目标检测在位置识别精度上仍有很大发展空间，特别是基于图信号处理方法的多目标检测.

6 GPC 研究框架

图 12 综合了当前 GPC 研究的主要技术路线，代表了基于图的点云全领域研究通用的框架. 具体来说，包括以下步骤：首先基于原始点云构造图结构[59,99,108,114,117]，然后将原领域问题部分或整体转化或分解为单个或多个图分析问题，最后利用图技术求解问题[52,87-88]. 构造图结构往往是利用图结构的性质能够有效获取点云的几何结构特征，特别是局部几何关系和上下文信息[36,106].

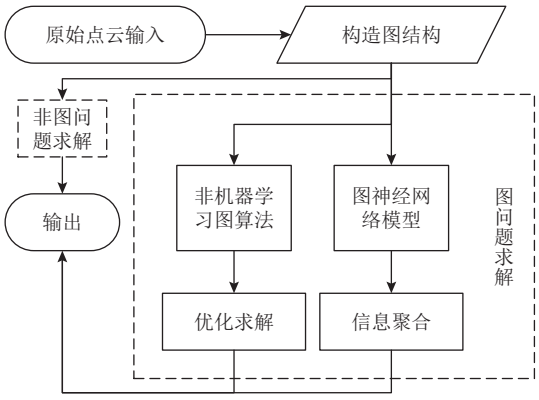


Fig. 12 Diagram of the GPC research framework

图 12 GPC 研究框架图

6.1 点云的图表示构造

当前主流图表示构造算法通常采用 *k*-d 树[84] 或 *k*-NN [59,99,114] 算法来计算每个中心点的邻居节点以构造边集合，形成图结构. 点云如若密度很高，每一个点都作为中心计算其邻居来构造边，产生的图其存储成本和后续计算成本都较为高昂，所以这种构造方法仅适用于单个物体分类的 CAD 数据集. 对于大型室内外场景数据集上的分割、目标检测等任务，可先通过简单分割提取出多个基本形状，将每个基本形状用 1 个节点表示，构造边来表示 2 个形状之间的关联，得到整个场景的拓扑图结构[128]. 此外，也可利用 GAN 模型生成图结构，或者设置 1 个训练模块来迭代地推断出 3 维点之间的不规则关系，以此学习到 1 个拓扑图[100-101]. 其他的图构造方法还包括构建点云级别的关系图. 在这种方法中，属于同一物体类别(如桌子)的多个点云作为图中的节点，根据点云之间的距离远近构建连边. 图中每个点代表了现实中的一个桌子，而图中的边代表的是 2 个桌子之间存在相似关系. 这种方法适用于 Few-Shot 或 One-Shot 点云查询或检索等任务[118]. 此外，利用 FPS 对点云进行点采样，以这些采样点为中心点构造局部 *k*-NN 图，可以降低计算成本.

对于原始点云的深度学习网络模型，随着逐层提取点云的深层次特征，3 维点之间的关系也非固定而是随着特征空间逐层变化[46,58,103]. 在每层特征空间中都重新计算 3 维点的邻居并动态更新图结构，更有利于保留点云丰富的内在特征. 不过，当输入点云在网络层数越深时，特征空间中点间关系的内在特征的丢失比原始几何形状信息的丢失更严重. 对此，可通过将原始几何形状信息作为先验知识，并引入几何注意权重来同时保留几何层面与特征层面的信息[104].

6.2 点云的图问题求解

将原始点云构造为图结构之后，不同任务要求被转化为图问题来求解. 求解方式包括图卷积以及非机器学习的图算法 2 大类. 自 2014 年 Bruna 等人[64] 提出图卷积 GNN 算法后，引爆研究热潮. 图卷积操作天然具有捕获节点之间不规则关系的能力，自 2017 年 ECC[35] 提出以来越来越多地应用于学习点云构造的拓扑图上的内在特征. 点云表示学习任务往往使用各类图卷积核聚合图中的邻域节点信息，提取点云的局部特征信息[40,102,104,106]. 对于分类、分割和识别任务，可先学习表示向量，再执行类别判定以及标签预测等识别任务，或者直接通过端到端的 GNN 模型完成特征提取和预测.

传统的非机器学习的图算法，譬如图分割技术，鉴于其在图计算领域不可忽视的卓越表现，使其几乎成为点云分割优化过程中提高性能的必备模块[124-126]. 对于场景点云上的分割、检测等任务，还需要对基本

形状进行模式匹配，包括线段、边缘等[132-134]. 在构造了图结构之后，通过在2部图中寻找最小代价的完美匹配来优化. 类似地，对于点云的配准，通过构造图为特征描述算子，将2个点云的描述符进行匹配得到点对，从而实现对齐. 此外，动态点云的每一帧可构造为1个拓扑图，将动态点云转换为1组图的序列，将位置和颜色等节点属性作为图顶点上的信号. 以压缩为例，可将连续点云帧之间的运动估计转化为连续图之间的特征匹配问题，根据运动估计改善动态点云的压缩性能[52].

6.3 点云的非图问题求解

将点云构造成图结构之后，也可以仅利用结构特征求解问题，并非一定要转化为图问题来完成点云分析任务. 对于捕获局部结构特征，图结构本身就具有巨大优势，直接利用图结构参与CNN操作、池化操作、传统机器学习模型或随机概率模型的计算. 对于动态点云上的修复、降噪、压缩等任务，可在原始动态点云上构造时空图[87-89]，能够很好地保留时序和点空间坐标信息，便于进一步捕捉动态点云帧到帧之间的时间和位置联系. 通常在时空图结构上应用预测模型或通过定义帧间流形距离来捕获点云相邻帧之间的时间对应关系. 此外，还可以不构造成图，而是直接通过*k*-NN等计算邻域，以局部邻域形式直接参与计算，如作为卷积神经网络的输入[116,161]，也可以归类为点云的非图问题求解的技术路线之一.

6.4 图与非图问题求解的优劣比较

本文所提出的GPC框架中，图问题求解方法和非图问题求解方法在解决不同任务时具有各自的优势和不足. 如果将一个任务建模为图问题求解，则便于利用图结构进行信息传递和处理，适用于处理点云中的拓扑结构和局部关系. 对于点云分割和语义分割等任务，GCN和GAT等图神经网络方法在处理局部和全局特征上表现出色，因此图问题求解在这些任务中具有优势.

然而，非图问题求解方法也不可忽视. 在点云目标检测和形状分类等任务中，点云的点级特征和几何信息可能更为重要，而邻域关系可简单地使用*k*-NN计算，然后通过传统的CNN直接进行卷积. 这样不通过构造图的技术求解方式，类似于PointNet和PointNet++等基于点的方法，在计算效率上通常较高，因为它们无需构建和处理复杂的图结构，而在保持排列不变性等点云特性、拓扑结构处理以及信息传播等方面表现不如图问题求解.

总之，选择适合的方法取决于任务的特点和需求. 图求解方法在处理点云的关系和拓扑结构时表现优秀，而非图问题求解方法更适用于提取点云几何信息的任务.

6.5 GPC框架的通用性

不同场景的数据集涉及的任务不同，所面临的挑战也有所不同. 在室内场景中，由于通常存在较多的家具和结构细节，点云的稠密性和噪声较高，导致语义分割和目标检测任务更具挑战性. 而在室外场景中，点云可能较为稀疏，受到不同光照和天气条件的影响，因此需要更强大的特征提取和分析方法.

尽管不同场景数据集存在挑战的差异，本节所提出的GPC框架仍然具有一定的通用性. 该框架以图的思想处理点云数据，并通过直接或间接构建图结构来描述点云中的局部和全局关系. 这种通用性在于，无论是室内还是室外场景，无论涉及物体还是整个场景的任务，点云的图表示都可以被应用. 不同场景下的点云数据可以被转换为相应的图结构，并在这个统一的图表示下进行后续的任务处理，使得算法的设计和实现更具一致性. 在该框架下，需要根据不同场景和任务要求，再进一步决定构造图结构的具体策略和计算方式，采用何种图算法或图神经网络技术.

7 未来发展方向

GPC研究不同于传统的基于体素或基于点集的点云处理方法，它保留了点云的非结构化信息，显式定义了点云的空间拓扑关系[101]. 经过近年的不断发展，理论和实践2方面都证实了图技术特别是GNN模型对于非结构化的点云数据来说是一种有效表示和处理框架[112]. 作为图和点云交叉的新兴研究领域，GPC颠覆了以往3维识别领域的体素化或点阵变换的运行结构，拓展改造了GNN的场景架构设计. 无论是从点云底层数据处理、中层表示学习到高层次识别任务，还是从静态点云到动态点云，GPC逐步在全层次应用范围内取得创造性突破，相关研究逐年增加，模型性能逐步提高.

尽管已经取得了不俗的表现和广泛的适用性，目前GPC研究在图构造和优化方面的研究和实践仍然处于起步阶段. 由于非结构化特性带来的复杂拓扑关系和不规则空间分布，GPC面临着传统规则图像数据不曾遇到的巨大挑战，包括计算成本、稳定性和可扩展性等. 在训练大量样本数据或在实际应用场景下，GPC研究仍然存在许多不足和开放性问题，是目前尚未完善和亟待探索的研究关键点，亟需更

多的学者从事 GPC 的研究、探索和分析. 下面分类别探讨了未来的研究要点和发展趋势，希望能为不同层次任务以及不同数据类型的未来研究工作提供参考和启发.

7.1 多模态融合

随着传感器技术的进步，同时获取 2 维图像和点云数据等多源、多模态数据成为了可能. 图像是 3 维物体在 2 维平面的投影，分布清晰、有序；而点云完整保留 3 维物体的几何形状，无序不规则. 融合 2 类数据、利用 2 者的差异以及利用发挥多模态数据分析的优势，高清 2 维图像弥补了 LiDAR 点云数据稀疏的缺点，点云数据补充了图像数据的 3 维立体信息的损失. 融合多模态数据提高点云的分类、分割、检测等识别精度日益成为业界学界的迫切需求. 2 维图像与 3 维点云的多模态融合将成为点云识别领域的主要发展趋势之一[162-164]. 目前，已有研究表明规则图像可转化为图结构来表示，使相关不规则优化技术应用于视觉领域的模式识别[165-168]. 将图像关键特征通过图网络化，可直接与原始点云数据融合，能够避免将点云数据体素化或规则化导致离散损失，最大限度地保留点云蕴含的 3 维信息. 总之，采用非线性的图表示和 GNN 等相关技术，是提高点云与图像多模态融合性能、增强融合效果的一种新思路.

7.2 动态性点云

GPC 方法面临着 2 个方面的动态性，即点云数据的动态性和构建的图结构表示的动态性. 现实生活中道路上、游戏和动画应用中的视觉物体或人物常常是运动的；点云特征空间也会逐层变化，导致初始图结构与最终输出层特征空间拓扑结构具有分布差异[104]. 虽然已有少数研究关注动态点云中的空间结构，但由于动态点云数据的海量性、缺损程度高、数据集完善性差等原因，相关探索仍处于初级阶段，如何利用图结构的优势处理动态点云中的时间空间不规则采样模式，进一步提高后续任务的精度，目前仍是一个开放性科学问题. 另一方面，在逐层提取点云特征过程中，特征空间随着神经网络层数的加深而改变，点云的图结构表示也应该随之变化，否则其表示精度以及蕴含的信息将逐层损失，造成精度偏移. 不过，逐层更新图结构必将增高计算成本. 如何构造有效的图结构，在特征空间逐层变化时，仍然能够高效更新所蕴含的信息，是 GPC 方法未来发展的一个关键议题.

7.3 彩色点云

由于设备获取或人工标注等方式，带有 RGB 颜色属性的点云，即彩色点云，在应用领域更受关注，因为颜色信息能够辅助提高包括配准[169]、分割[170]和目标检测[171-172]等多个任务的精度. 色彩是人眼识别物体、区分物体的重要依据和参考，所以，彩色点云是未来 3 维视觉的重要研究对象，色彩信息的挖掘是点云数据挖掘的关键环节，在未来各行业应用中更具有应用价值，也更符合实际场景. 即便如此，颜色信息使点云数据维度更高，且与坐标信息具有本质差别，因此彩色点云的处理一直存在困难和挑战. 然而，GPC 模型能够同时容纳颜色信息和坐标信息，将所有信息均作为图的节点特征，一次性输入进图神经网络模型或进行图信号处理. 因此，GPC 框架是未来彩色点云处理的重要技术路线之一. 鉴于 GNN 和图信号处理理论对点云高维度信息挖掘的天然优势，以及彩色信息在 3 维识别中的重要作用，笔者认为，彩色点云是 GPC 未来应用的主要方向之一，是提高点云识别精度的必由之路.

7.4 自监督学习

点云的数据集场景复杂且规模庞大. 均衡数据集的采集一直以来都是点云处理领域的重要环节. 然而数据集的采集和标注耗时耗力，需要经年累月的人力维护. 况且，数据集的标注质量和样本均衡性也难以评估. 如何在已有的少量具有优质标注数据集、大量弱标注或无标注数据集的条件下，通过弱监督或自监督的方式完成各项任务，是未来重要发展趋势之一[99-101]. 自监督或弱监督 GPC 方法，可借助图结构推断作为先验知识，通过图匹配或 GNN 等方式从无标注数据中挖掘信息. 利用良好的先验知识和结构推断策略，实现接近于有监督条件下的模型表现性能，能够大大降低算法对数据集标注质量的依赖，推动 3 维点云底层处理乃至识别任务研究的进一步发展.

7.5 邻域搜索

邻域搜索机制是构建图结构以及图上节点查找或采样的重要环节. 点云的图表示构造需要计算每个节点与邻居节点的欧氏空间距离，据此进行 2 点间相连. 当点云的规模过于庞大时，可采用 FPS 采样部分节点作为中心节点，只计算它们的邻域，提高计算效率[173]. 如果点云数据随时间发生改变，或随着特征的逐层提取，节点的位置坐标或嵌入向量发生变化，导致坐标距离或特征关系近似度需要重新计算，这将产生新的邻域和局部图结构. 虽然 GPC 方法的图结构最大程度保留了点云原始的非结构信息，但点云原始数据本身不具备显式的连接信息，导致邻

域搜索机制对整体效率具有较大影响. 目前大多使用 k-NN 等计算代价昂贵的邻域点搜索机制，使得 GPC 方法计算复杂度偏高[37,161]. 此外，由于构建图结构的策略单一，图结构的强大表示能力在点云数据上的发挥受限，影响了后续特征提取的精度上限. 亟须提出更加灵活且具有突破性的 GPC 表示方法，其中如何高效构造邻域关系是未来提高 GPC 性能的关键.

8 总　　结

点云是一种立体图像，在平面图像研究了半个世纪的情况下，点云的出现，一方面是平面图像向 3 维空间扩展的结果，另一方面也是现代化快速 LiDAR 探测硬件技术长足发展后应运而生的结果. 点云的研究，在理论上，或许对平面图像研究是一个很好的补充，或许很多平面图像的理论可以扩展到 3 维图像. 点云图像理论的一个特例为平面图像理论，某种程度上是对平面图像研究的“降维打击”，使图像理论完成量变到质变的飞跃. 在实践上，点云可以极大提升自动驾驶的视觉纵深准确性，直接避免了使用多角度多个图像配合计算探测距离的问题，大大加快了计算速度，从而将自动驾驶真正推向实用. 目前，点云已开始在美团、百度、谷歌、特斯拉等公司自动驾驶中进行实验和使用，点云的研究在理论上滞后于实践，一些理论上的问题亟待解决. 点云的产生和发展，不仅是理论上的必然，也是应用上半个世纪的等待. 点云的研究，必将把人工智能推送到一个崭新的时代.

在点云研究中，GPC 的思想由来已久，很早就有借助图分割等传统的非机器学习的图算法来处理点云的非结构特性的研究. 当点云广泛用于 3 维重建与监控[174-176]、医学或文物影像[169,173]以及遥感[177-179]等领域，同时在处理非结构化特征问题上 GNN 技术取得巨大突破后，GPC 作为图与点云的一个交叉领域，展现出卓越的研究潜力和优势：

1）在底层处理方面，GPC 研究早期使用传统图算法处理静态点云，如图分割、图匹配等，近期主要用图卷积模型. 对于动态点云，GPC 研究主要集中在时空图的构建和优化过程. 然而，无论是静态点云还是动态点云的底层处理，现阶段的 GPC 研究都在构造图结构方面存在 k-NN 导致的效率限制. 此外，虽然时空图能够同时对时序和空间关系进行捕捉和建模，但优化过程计算成本较高，GPC 在时空图模型优化方面仍然有些路要走.

2）在表示学习方面，GPC 主要基于图卷积网络开展表示学习研究，或者采用无监督的图神经网络模型，如图生成对抗网络、图编码器等. 然而，无论是图卷积网络还是无监督情况下的 GPC 表示学习模型，均受限于邻域计算的效率问题. 此外，无监督、弱监督或自监督场景下的 GPC 表示学习研究仍处于起步阶段，模型的扩展性、可解释性都需要点云和图两大领域的研究人员更多合作和共同探索.

3）在识别任务方面，GPC 对于高层任务的研究主要集中在图卷积及池化操作的定义方面，其他图模型的研究成果目前仍然较少. 现有 GPC 的图卷积模型主要通过 k-NN 构建局部图结构来执行，但 k-NN 需要对逐个点对计算点间距离，存在效率上的限制性，并不能充分发挥图卷积的性能，需要进一步展开 GPC 在图构造方面的相关研究.

笔者认为，用 GPC 框架高效、高精度捕获动态、静态点云信息是大势所趋. 在 GNN 和点云研究备受瞩目、日行千里的当下，作为点云和图领域的新兴交叉研究领域，GPC 目前正在也未来必将受到越来越多研究者的探索和关注. 在未来研究的道路上，还可能涌现更多的不同图结构类型，如 2 部图、异质图，以及更加灵活的邻域搜索和更新的方案出现，通过层出不穷的 GNN 框架精准且高效率分析点云的不规则结构特性，实现计算效率与精度的平衡. 如何在充分满足点云数据特征的基础上，最大限度发挥图结构和图模型的优点，在精确挖掘特征的同时尽可能降低结构构造及优化成本，是未来 GPC 研究需要着手解决的问题.

作者贡献声明：梁循对论文框架提出指导意见；李志莹完成文献调研并撰写论文；蒋洪迅提出指导意见并修改论文.

参　考　文　献

[1] Amenta N, Bern M, Kamvysselis M. A new Voronoi-based surface reconstruction algorithm [C] //Proc of the 25th Int Conf on Computer Graphics and Interactive Techniques. New York: ACM, 1998: 415−422

[2] Boissonnat J D. Geometric structures for three-dimensional shape representation[J]. ACM Transactions on Graphics, 1984, 3(4): 266−286

[3] Bernardini F, Mittleman J, Rushmeier H, et al. The ball-pivoting algorithm for surface reconstruction[J]. IEEE Transactions on Visualization and Computer Graphics, 1999, 5(4): 349−359

[4] Schechter H, Bridson R. Ghost SPH for animating water[J]. ACM

Transactions on Graphics, 2012, 31(4): 1−8

[5] Riegler G, Ulusoy A O, Geiger A. OctNet: Learning deep 3D representations at high resolutions [C] //Proc of 2017 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2017: 6620−6629

[6] Qi C R, Su Hao, Niebner M, et al. Volumetric and multi-view CNNs for object classification on 3D data [C] //Proc of 2016 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2016: 5648−5656

[7] Malik J, Shimada S, Elhayek A, et al. HandVoxNet++: 3D hand shape and pose estimation using voxel-based neural networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(12): 8962−8974

[8] Su Hang, Maji S, Kalogerakis E, et al. Multi-view convolutional neural networks for 3D shape recognition [C] //Proc of 2015 IEEE Int Conf on Computer Vision (ICCV). Los Alamitos, CA: IEEE Computer Society, 2015: 945−953

[9] Liebelt J, Schmid C. Multi-view object class detection with a 3D geometric model [C] //Proc of 2010 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2010: 1688−1695

[10] Li Lei, Zhu Siyu, Fu Hongbo, et al. End-to-end learning local multi-view descriptors for 3D point clouds [C] //Proc of 2020 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2020: 1916−1925

[11] Chen Xiaozhi, Ma Huimin, Wan Ji, et al. Multi-view 3D object detection network for autonomous driving [C] //Proc of 2017 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2017: 6526−6534

[12] Wang Nanyang, Zhang Yinda, Li Zhuwen, et al. Pixel2Mesh: Generating 3D mesh models from single RGB images [C] //Proc of the 15th European Conf on Computer Vision (ECCV). Berlin: Springer, 2018: 55−71

[13] Guo Kan, Zou Dongqing, Chen Xiaowu. 3D mesh labeling via deep convolutional neural networks[J]. ACM Transactions on Graphics, 2015, 35(1): 1−12

[14] Qi C R, Su Hao, Mo Kaichun, et al. PointNet: Deep learning on point sets for 3D classification and segmentation [C] //Proc of 2017 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2017: 77−85

[15] Qi C R, Yi Li, Su Hao, et al. PointNet++: Deep hierarchical feature learning on point sets in a metric space [C] //Proc of 2017 Advances in Neural Information Processing Systems. New York: ACM, 2017: 5100−5109

[16] Li Shanglin, Xie Wenjun, Li Lin, et al. Computer rapid modeling methods of architecture: A review[J]. Chinese Journal of Computers, 2019, 42(9): 1966−1990 (in Chinese)
(李尚林, 谢文军, 李琳, 等. 计算机快速建筑建模方法综述[J]. 计算机学报, 2019, 42(9): 1966−1990)

[17] Lu Weixin, Zhou Yao, Wan Guowei, et al. L3-Net: Towards learning based LiDAR localization for autonomous driving [C] //Proc of 2019 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2019: 6382−6391

[18] Geiger A, Lenz P, Urtasun R. Are we ready for autonomous driving? The KITTI vision benchmark suite [C] //Proc of 2012 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2012: 3354−3361

[19] Meyer G P, Laddha A, Kee E, et al. LaserNet: An efficient probabilistic 3D object detector for autonomous driving [C] //Proc of 2019 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2019: 12669−12678

[20] Das S, Patibandla H, Bhattacharya S, et al. TMCOSS: Thresholded multi-criteria online subset selection for data-efficient autonomous driving [C] //Proc of 2021 IEEE Int Conf on Computer Vision (ICCV). Los Alamitos, CA: IEEE Computer Society, 2021: 6321−6330

[21] Marinoiu E, Zanfir M, Olaru V, et al. 3D human sensing, action and emotion recognition in robot assisted therapy of children with autism [C] //Proc of 2018 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2018: 2158−2167

[22] Kasaei S H, Sock J, Lopes L S, et al. Perceiving, learning, and recognizing 3D objects: An approach to cognitive service robots [C] //Proc of the 32nd AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2018: 596−603

[23] Im S, Ha H, Choe G, et al. High quality structure from small motion for rolling shutter cameras [C] //Proc of 2015 IEEE Int Conf on Computer Vision (ICCV). Los Alamitos, CA: IEEE Computer Society, 2015: 837−845

[24] Behley J, Garbade M, Milioto A, et al. SemanticKITTI: A dataset for semantic scene understanding of LiDAR sequences [C] //Proc of 2019 IEEE Int Conf on Computer Vision (ICCV). Los Alamitos, CA: IEEE Computer Society, 2019: 9296−9306

[25] Frueh C, Zakhor A. Constructing 3D city models by merging ground-based and airborne views [C] //Proc of 2003 IEEE Computer Society Conf on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2003: II-562

[26] Papon J, Abramov A, Schoeler M, et al. Voxel cloud connectivity segmentation - Supervoxels for point clouds [C] //Proc of 2013 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2013: 2027−2034

[27] Lee D, Lee J, Lee J, et al. Regularization strategy for point cloud via rigidly mixed sample [C] //Proc of 2021 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2021: 15895−15904

[28] Lei Huan, Akhtar N, Mian A. Octree guided CNN with spherical kernels for 3D point clouds [C] //Proc of 2019 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2019: 9623−9632

[29] Zhao Hengshuang, Jiang Li, Fu Chiwing, et al. Pointweb: Enhancing local neighborhood features for point cloud processing [C] //Proc of 2019 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2019: 5560−5568

[30] Zhou Haoran, Chen Honghua, Feng Yidan, et al. Geometry and learning co-supported normal estimation for unstructured point cloud [C] //Proc of 2020 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2020: 13235−13244

[31] Behl A, Paschalidou D, Donne S, et al. Pointflownet: Learning representations for rigid motion estimation from point clouds[C]//Proc of 2019 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2019: 7954−7963

[32] Meng Hsienyu, Gao Lin, Lai Yukun, et al. VV-Net: Voxel VAE net with group convolutions for point cloud segmentation[C]//Proc of 2019 IEEE Int Conf on Computer Vision (ICCV). Los Alamitos, CA: IEEE Computer Society, 2019: 8499−8507

[33] Casajus P H, Ritschel T, Ropinski T. Total denoising: Unsupervised learning of 3D point cloud cleaning[C]//Proc of 2019 IEEE Int Conf on Computer Vision (ICCV). Los Alamitos, CA: IEEE Computer Society, 2019: 52−60

[34] Landrieu L, Simonovsky M. Large-scale point cloud semantic segmentation with superpoint graphs[C]//Proc of 2018 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2018: 4558−4567

[35] Simonovsky M, Komodakis N. Dynamic edge-conditioned filters in convolutional neural networks on graphs[C]//Proc of 2017 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2017: 29−38

[36] Li Guohao, Muller M, Thabet A, et al. DeepGCNs: Can GCNs go as deep as CNNs[C]//Proc of 2019 IEEE Int Conf on Computer Vision (ICCV). Los Alamitos, CA: IEEE Computer Society, 2019: 9266−9275

[37] Xu Qiangeng, Sun Xudong, Wu Choying, et al. Grid-GCN for fast and scalable point cloud learning[C]//Proc of 2020 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2020: 5660−5669

[38] Zhang Yingxue, Rabbat M. A graph-CNN for 3D point cloud classification[C]// Proc of 2018 IEEE Int Conf on Acoustics, Speech, and Signal Processing (ICASSP). Piscataway, NJ: IEEE, 2018: 6279−6283

[39] Landrieu L, Boussaha M. Point cloud oversegmentation with graph-structured deep metric learning[C]//Proc of 2019 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2019: 7432−7441

[40] Wang Chu, Samari B, Siddiqi K. Local spectral graph convolution for point set feature learning[C]//Proc of the 15th European Conf on Computer Vision (ECCV). Berlin: Springer, 2018: 56−71

[41] Li Ruihui, Li Xianzhi, Heng Phengann, et al. PointAugment: An auto-augmentation framework for point cloud classification[C]//Proc of 2020 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2020: 6377−6386

[42] Lin Xuanxiang, Chen Ke, Jia Kui. Object point cloud classification via poly-convolutional architecture search[C]//Proc of the 29th ACM Int Conf on Multimedia (MM). New York: ACM, 2021: 807−815

[43] Wang Lei, Huang Yuchun, Hou Yaolin, et al. Graph attention convolution for point cloud semantic segmentation[C]//Proc of 2019 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2019: 10288−10297

[44] Li Rongkang, Zhang Yumeng, Niu Dongmei, et al. PointVGG: Graph convolutional network with progressive aggregating features on point clouds[J]. Neurocomputing, 2021, 429: 187−198

[45] Xu Yusheng, Tuttas S, Hoegner L, et al. Geometric primitive extraction from point clouds of construction sites using VGS[J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(3): 424−428

[46] Chen Yang, Xu Yaming, Xing Yin, et al. DGCNN network architecture with densely connected point pairs in multiscale local regions for ALS point cloud classification[J]. IEEE Geoscience and Remote Sensing Letters, 2022, 19: 1−5

[47] Chen Siheng, Tian Dong, Feng Chen, et al. Fast resampling of three-dimensional point clouds via graphs[J]. IEEE Transactions on Signal Processing, 2018, 66(3): 666−681

[48] Chen Chaomei. Science mapping: A systematic review of the literature[J]. Journal of Data and Information Science, 2017, 2(2): 1−40

[49] Chen Chaomei, Song Ming. Visualizing a field of research: A methodology of systematic scientometric reviews[J]. Plos One, 2019, 14(10): e0223994

[50] Guo Yulan, Wang Hanyun, Hu Qingyong, et al. Deep learning for 3D point clouds: A survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 43(12): 4338−4364

[51] Fei Ben, Yang Weidong, Chen Wenming, et al. Comprehensive review of deep learning-based 3D point cloud completion processing and analysis[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(12): 22862−22883

[52] Thanou D, Chou P A, Frossard P. Graph-based compression of dynamic 3D point cloud sequences[J]. IEEE Transactions on Image Processing, 2016, 25(4): 1765−1778

[53] DeQueiroz R L, Chou P A. Motion-compensated compression of dynamic voxelized point clouds[J]. IEEE Transactions on Image Processing, 2017, 26(8): 3886−3895

[54] Liu Xingyu, Yan Mengyuan, Bohg J. Meteornet: Deep learning on dynamic 3D point cloud sequences[C]//Proc of 2019 IEEE Int Conf on Computer Vision (ICCV). Los Alamitos, CA: IEEE Computer Society, 2019: 9245−9254

[55] Guerrero P, Kleiman Y, Ovsjanikov M, et al. PCPNet: Learning local shape properties from raw point clouds[J]. Computer Graphics Forum, 2018, 37(2): 75−85

[56] University College London. PCPNet dataset[EB/OL]. [2023-01-04]. http://geometry.cs.ucl.ac.uk/projects/2018/pcpnet/

[57] Demarsin K, Vanderstraeten D, Volodine T, et al. Detection of closed sharp edges in point clouds using normal estimation and graph theory[J]. CAD Computer Aided Design, 2007, 39(4): 276−283

[58] Wang Yue, Sun Yongbin, Liu Ziwei, et al. Dynamic graph CNN for learning on point clouds[J]. ACM Transactions on Graphics, 2019, 38(5): 1−12

[59] Shen Yiru, Feng Chen, Yang Yaoqing, et al. Mining point cloud local structures by kernel correlation and graph pooling[C]//Proc of 2018 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2018: 4548−4557

[60] Adoni H W Y, Nahhal T, Krichen M, et al. A survey of current challenges in partitioning and processing of graph-structured data in parallel and distributed systems[J]. Distributed and Parallel Databases, 2020, 38(2): 495−530

[61] Xiang Yingzhuo, Tan Juxian, Han Jiesi, et al. Survey of graph matching algorithms [J]. Computer Science, 2018, 45(6): 27−31, 45 (in Chinese)
(项英倬, 谭菊仙, 韩杰思, 等. 图匹配技术研究[J]. 计算机科学, 2018, 45(6): 27−31, 45)

[62] Jurić A, Kendeš F, Marković I, et al. A comparison of graph optimization approaches for pose estimation in SLAM [C] //Proc of the 44th Int Convention on Information, Communication and Electronic Technology (MIPRO). Piscataway, NJ: IEEE, 2021: 1113−1118

[63] Gori M, Monfardini G, Scarselli F. A new model for learning in graph domains [C] //Proc of 2005 IEEE Int Joint Conf on Neural Networks (IJCNN). Piscataway, NJ: IEEE, 2005: 729−734

[64] Bruna J, Zaremba W, Szlam A, et al. Spectral networks and locally connected networks on graphs [C/OL] //Proc of the 2nd Int Conf on Learning Representations (ICLR). 2014[2023-01-04]. https://openreview.net/forum?id=DQNsQf-UsoDBa

[65] Shuman D I, Narang S K, Frossard P, et al. The emerging field of signal processing on graphs[J]. IEEE Signal Processing Magazine, 2013, 30(3): 83−98

[66] Defferrard M, Bresson X, Vandergheynst P. Convolutional neural networks on graphs with fast localized spectral filtering [C] //Proc of 2016 Advances in Neural Information Processing Systems. New York: ACM, 2016: 3837−3845

[67] Levie R, Monti F, Bresson X, et al. CayleyNets: Graph convolutional neural networks with complex rational spectral filters[J]. IEEE Transactions on Signal Processing, 2019, 67(1): 97−109

[68] Hamilton W, Ying Rex, Leskovec J. Inductive representation learning on large graphs [C] //Proc of 2017 Advances in Neural Information Processing Systems. New York: ACM, 2017: 1024−1034

[69] Chen Jie, Ma Tengfei, Xiao Cao. FastGCN: Fast learning with graph convolutional networks via importance sampling [C/OL] //Proc of the 6th Int Conf on Learning Representations (ICLR). 2018[2023-01-04]. https://openreview.net/forum?id=rytstxWAW

[70] Wang Chun, Pan Shirui, Long Guodong, et al. MGAE: Marginalized graph autoencoder for graph clustering [C] //Proc of the 26th ACM on Conf on Information and Knowledge Management. New York: ACM, 2017: 889−898

[71] Hasanzadeh A, Hajiramezanali E, Narayanan K, et al. Semi-implicit graph variational auto-encoders [C] //Proc of 2019 Advances in Neural Information Processing Systems. New York: ACM, 2019: 10711−10722

[72] Wang Hongwei, Wang Jia, Wang Jialin, et al. GraphGAN: Graph representation learning with generative adversarial nets [C] //Proc of the 32nd AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2018: 2508−2515

[73] Ding Ming, Tang Jie, Zhang Jie. Semi-supervised learning on graphs with generative adversarial nets [C] //Proc of the 27th ACM Int Conf on Information and Knowledge Management. New York: ACM, 2018: 913−922

[74] Lee J B, Rossi R, Kong Xiangnan. Graph classification using structural attention [C] //Proc of the 24th ACM SIGKDD Int Conf on Knowledge Discovery & Data Mining. New York: ACM, 2018: 1666−1674

[75] Knyazev B, Taylor G W, Amer M. Understanding attention and generalization in graph neural networks [C] //Proc of 2019 Advances in Neural Information Processing Systems. New York: ACM, 2019: 4204−4214

[76] Perozzi B, Al-Rfou R, Skiena S. DeepWalk: Online learning of social representations [C] //Proc of the 20th ACM SIGKDD Int Conf on Knowledge discovery and data mining. New York: ACM, 2014: 701−710

[77] Grover A, Leskovec J. Node2vec: Scalable feature learning for networks [C] //Proc of the 22nd ACM SIGKDD Int Conf on Knowledge Discovery and Data Mining. New York: ACM, 2016: 855−864

[78] Tu Cunchao, Liu Han, Liu Zhiyuan, et al. CANE: Context-aware network embedding for relation modeling [C] //Proc of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Stroudsburg, PA: ACL, 2017: 1722−1731

[79] Li Ziyao, Zhang Liang, Song Guojie. SepNE: Bringing separability to network embedding [C] //Proc of the 33rd AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2019: 4261−4268

[80] Zhang Shunli, Xu Yanzhi, Zhou Mingquan, et al. Registration of point clouds based on matching of general adaptive neighborhood[J]. Chinese Journal of Computers, 2019, 42(9): 2114−2126 (in Chinese)
(张顺利, 徐艳芝, 周明全, 等. 基于自适应邻域匹配的点云配准方法[J]. 计算机学报, 2019, 42(9): 2114−2126)

[81] Xu Siyu, Zhu Jihua, Tian Zhiqiang, et al. Stepwise refinement approach for registration of multi-view point sets[J]. Acta Automatica Sinica, 2019, 45(8): 1486−1494 (in Chinese)
(徐思雨, 祝继华, 田智强, 等. 逐步求精的多视角点云配准方法[J]. 自动化学报, 2019, 45(8): 1486−1494)

[82] Miao Yongwei, Liu Jiazong, Sun Yuliang, et al. Point cloud completion of indoor scenes based on category-instance segmentation[J]. Chinese Journal of Computers, 2021, 44(11): 2189−2202 (in Chinese)
(缪永伟, 刘家宗, 孙瑜亮, 等. 基于类别-实例分割的室内点云场景修复补全[J]. 计算机学报, 2021, 44(11): 2189−2202)

[83] Cai Qinyi, Chen Zhonggui, Cao Juan. High-quality point cloud resampling method based on optimal transport theory[J]. Chinese Journal of Computers, 2022, 45((1):): 135−147 (in Chinese)
(蔡钦镒, 陈中贵, 曹娟. 基于最优传输理论的高质量点云重采样方法[J]. 计算机学报, 2022, 45(1): 135−147)

[84] Gu Shuai, Hou Junhui, Zeng Huanqiang, et al. 3D point cloud attribute compression via graph prediction[J]. IEEE Signal Processing Letters, 2020, 27: 176−180

[85] Zhang Congyi, Wei Zizhuang, Xu Haowen, et al. Scale variable fast global point cloud registration[J]. Chinese Journal of Computers, 2019, 42(9): 1939−1952 (in Chinese)
(张琮毅, 魏子庄, 徐昊文, 等. 尺度可变的快速全局点云配准方法[J]. 计算机学报, 2019, 42(9): 1939−1952)

[86] Zeng Jin, Cheung G, Ng M, et al. 3D point cloud denoising using graph Laplacian regularization of a low dimensional manifold

model[J]. IEEE Transactions on Image Processing, 2020, 29: 3474−3489

[87] Hu Wei, Hu Qianjiang, Wang Zehua, et al. Dynamic point cloud denoising via manifold-to-manifold distance[J]. IEEE Transactions on Image Processing, 2021, 30: 6168−6183

[88] Fu Zeqing, Hu Wei. Dynamic point cloud inpainting via spatial-temporal graph learning[J]. IEEE Transactions on Multimedia, 2021, 23: 3022−3034

[89] Xu Yiqun, Hu Wei, Wang Shanshe, et al. Predictive generalized graph Fourier transform for attribute compression of dynamic point clouds[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2021, 31(5): 1968−1982

[90] Bayram E, Frossard P, Vural E, et al. Analysis of airborne LiDAR point clouds with spectral graph filtering[J]. IEEE Geoscience and Remote Sensing Letters, 2018, 15(8): 1284−1288

[91] Wang Feipeng, Xiao Jun, Wang Ying. Efficient rock-mass point cloud registration using n-point complete graphs[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(11): 9332−9343

[92] Lu Bin, Fan Xiaoming. Research on 3D point cloud skeleton extraction based on improved adaptive k-means clustering[J]. Acta Automatica Sinica, 2022, 48(8): 1994−2006 (in Chinese)
(鲁斌, 范晓明. 基于改进自适应 k 均值聚类的三维点云骨架提取的研究[J]. 自动化学报, 2022, 48(8): 1994−2006)

[93] Lu Min, Guo Yulan, Zhang Jun, et al. Recognizing objects in 3D point clouds with multi-scale local features[J]. Sensors (Switzerland), 2014, 14(12): 24156−24173

[94] Rusu R B, Blodow N, Marton Z C, et al. Aligning point cloud views using persistent feature histograms [C] //Proc of 2008 IEEE/RSJ Int Conf on Intelligent Robots and Systems (IROS). Piscataway, NJ: IEEE, 2008: 3384−3391

[95] Rusu R B, Blodow N, Beetz M. Fast point feature histograms (FPFH) for 3D registration [C] //Proc of 2009 IEEE Int Conf on Robotics and Automation. Piscataway, NJ: IEEE, 2009: 3212−3217

[96] Achlioptas P, Diamanti O, Mitliagkas I, et al. Learning representations and generative models for 3D point clouds [C] // Proc of the 35th Int Conf on Machine Learning (ICML). New York: ACM, 2018: 67−85

[97] Bronstein M M, Bruna J, Lecun Y, et al. Geometric deep learning: Going beyond Euclidean data[J]. IEEE Signal Processing Magazine, 2017, 34(4): 18−42

[98] Scarselli F, Gori M, Tsoi A C, et al. The graph neural network model[J]. IEEE Transactions on Neural Networks, 2009, 20(1): 61−80

[99] Hassani K, Haley M. Unsupervised multi-task feature learning on point clouds [C] //Proc of 2019 IEEE Conf on Computer Vision (ICCV). Los Alamitos, CA: IEEE Computer Society, 2019: 8159−8170

[100] Chen Siheng, Duan Chaojing, Yang Yaoqing, et al. Deep unsupervised learning of 3D point clouds via graph topology inference and filtering[J]. IEEE Transactions on Image Processing, 2020, 29: 3183−3198

[101] Valsesia D, Fracastoro G, Magli E. Learning localized representations of point clouds with graph-convolutional generative adversarial networks[J]. IEEE Transactions on Multimedia, 2021, 23: 402−414

[102] Zhang Cheng, Chen Hao, Wan Haocheng, et al. Graph-PBN: Graph-based parallel branch network for efficient point cloud learning[J]. Graphical Models, 2022, 119: 101120

[103] Zhang Kuangen, Hao Ming, Wang Jing, et al. Linked dynamic graph CNN: Learning on point cloud via linking hierarchical features [J]. arXiv preprint, arXiv: 1904.10014, 2019

[104] Cui Yiming, Liu Xin, Liu Hongmin, et al. Geometric attentional dynamic graph convolutional neural networks for point cloud analysis[J]. Neurocomputing, 2021, 432: 300−310

[105] Yang Yaoqing, Feng Chen, Shen Yiru, et al. FoldingNet: Point cloud auto-encoder via deep grid deformation [C] //Proc of 2018 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2018: 206−215

[106] Valsesia D, Fracastoro G, Magli E. Learning localized generative models for 3D point clouds via graph convolution [C/OL] //Proc of the 7th Int Conf on Learning Representations (ICLR). 2019[2023-01-04]. https://openreview.net/forum?id=SJeXSo09FQ

[107] Grilli E, Menna F, Remondino F. A review of point clouds segmentation and classification algorithms[J]. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2017, XLII-2/W3: 339−344

[108] Chen Chaofan, Qian Shengsheng, Fang Quan, et al. HAPGN: Hierarchical attentive pooling graph network for point cloud segmentation[J]. IEEE Transactions on Multimedia, 2021, 23: 2335−2346

[109] Wang Weiyue, Yu R, Huang Qiangui, et al. SGPN: Similarity group proposal network for 3D point cloud instance segmentation [C] //Proc of 2018 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2018: 2569−2578

[110] Zhou Yin, Tuzel O. VoxelNet: End-to-end learning for point cloud based 3D object detection [C] //Proc of 2018 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2018: 4490−4499

[111] Ren Shaoqing, He Kaiming, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [C] //Proc of 2015 Conf on Neural Information Processing Systems (NeurIPS). Cambridge, MA: MIT, 2015: 91−99

[112] Shi Weijing, Rajkumar R. Point-GNN: Graph neural network for 3D object detection in a point cloud [C] //Proc of 2020 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2020: 1708−1716

[113] Pan Guanghua, Wang Jun, Ying Rendong, et al. 3DTI-Net: Learn inner transform invariant 3D geometry features using dynamic GCN [J]. arXiv preprint, arXiv: 1812.06254, 2018

[114] Chen Chao, Li Guanbin, Xu Ruijia, et al. ClusterNet: Deep hierarchical cluster network with rigorously rotation-invariant representation for point cloud analysis [C] //Proc of 2019 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2019: 4989−4997

[115] Liu Jinxian, Ni Bingbing, Li Caiyuan, et al. Dynamic points

agglomeration for hierarchical point sets learning [C] //Proc of 2019 IEEE Int Conf on Computer Vision (ICCV). Los Alamitos, CA: IEEE Computer Society, 2019: 7545−7554

[116] Shan Xuanyang, Sun Zhanli, Zeng Zhigang. RFNet: Convolutional neural network for 3D point cloud classification [J]. Acta Automatica Sinica, 2023, 49(11): 2350−2359 (in Chinese)
(单铉洋, 孙战里, 曾志刚. RFNet: 用于三维点云分类的卷积神经网络[J]. 自动化学报, 2023, 49(11): 2350−2359)

[117] Guo Rui, Zhou Yong, Zhao Jiaqi, et al. Point cloud classification by dynamic graph CNN with adaptive feature fusion[J]. IET Computer Vision, 2021, 15(3): 235−244

[118] Fan Zhaoxin, Liu Hongyan, He Jun, et al. A graph-based one-shot learning method for point cloud recognition[J]. Computer Graphics Forum, 2020, 39(7): 313−323

[119] Li Zhuqiang, Zhang Liqiang, Tong Xiaohua, et al. A three-step approach for TLS point cloud classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(9): 5412−5424

[120] Jiang Li, Zhao Hengshuang, Liu Shu, et al. Hierarchical point-edge interaction network for point cloud semantic segmentation [C] //Proc of 2019 IEEE Int Conf on Computer Vision (ICCV). Los Alamitos, CA: IEEE Computer Society, 2019: 10432−10440

[121] Khan S A, Shi Yilei, Shahzad M, et al. FGCN: Deep feature-based graph convolutional network for semantic segmentation of urban 3D point clouds [C] // Proc of 2020 IEEE Conf on Computer Vision and Pattern Recognition Workshops (CVPR Workshops). Los Alamitos, CA: IEEE Computer Society, 2020: 778−787

[122] Li Ying, Ma Lingfei, Zhong Zilong, et al. TGNet: Geometric graph CNN on 3D point cloud segmentation[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(5): 3588−3600

[123] Sun Yuliang, Miao Yongwei, Chen Jiazhou, et al. PGCNet: Patch graph convolutional network for point cloud segmentation of indoor scenes[J]. The Visual Computer, 2020, 36(10): 2407−2418

[124] Previtali M, Díaz-Vilariño L, Scaioni M. Indoor building reconstruction from occluded point clouds using graph-cut and ray-tracing[J]. Applied Sciences, 2018, 8(9): 1529

[125] Sedlácek D, Zara J. Graph cut based point-cloud segmentation for polygonal reconstruction [C] //Proc of the 5th Int Symp on Advances in Visual Computing: Part II. Berlin: Springer, 2009: 218−227

[126] Pan Rongjiang, Taubin G. Automatic segmentation of point clouds from multi-view reconstruction using graph-cut[J]. The Visual Computer, 2016, 32(5): 601−609

[127] Luo Huan, Wang Cheng, Wen Chenglu, et al. Patch-based semantic labeling of road scene using colorized mobile LiDAR point clouds[J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(5): 1286−1297

[128] Ben-Shabat Y, Avraham T, Lindenbaum M, et al. Graph based over-segmentation methods for 3D point clouds[J]. Computer Vision and Image Understanding, 2018, 174: 12−23

[129] Te Gusi, Hu Wei, Zheng Amin, et al. RGCNN: Regularized graph CNN for point cloud segmentation [C] //Proc of the 26th ACM Int Conf on Multimedia (MM). New York: ACM, 2018: 746−754

[130] Xu Sheng, Wang Ruisheng, Wang Hao, et al. An optimal hierarchical clustering approach to mobile LiDAR point clouds[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(7): 2765−2776

[131] Golovinskiy A, Kim V G, Funkhouser T. Shape-based recognition of 3D point clouds in urban environments [C] //Proc of 2009 IEEE Int Conf on Computer Vision (ICCV). Los Alamitos, CA: IEEE Computer Society, 2009: 2154−2161

[132] Nieuwenhuisen M, Stückler J, Berner A, et al. Shape-primitive based object recognition and grasping [C/OL] //Proc of the 7th German Conf on Robotics (ROBOTIK). 2012: 468−472.[2023-01-04]. https://api.semanticscholar.org/CorpusID:14393050

[133] Berner A, Li Jun, Holz D, et al. Combining contour and shape primitives for object detection and pose estimation of prefabricated parts [C] //Proc of 2013 IEEE Int Conf on Image Processing. Piscataway, NJ: IEEE, 2013: 3326−3330

[134] Hao Wen, Wang Yinghui. Structure-based object detection from scene point clouds[J]. Neurocomputing, 2016, 191: 148−160

[135] Sun Xiaopeng, Li Sihui, Wang Lu, et al. Shape feature matching algorithm of ear point cloud using path following[J]. Journal of Software, 2015, 26(5): 1251−1264 (in Chinese)
(孙晓鹏, 李思慧, 王璐, 等. 耳廓点云形状特征匹配的路径跟随算法[J]. 软件学报, 2015, 26(5): 1251−1264)

[136] Chen Jintai, Lei Biwen, Song Qingyu, et al. A hierarchical graph network for 3D object detection on point clouds [C] //Proc of 2020 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2020: 389−398

[137] Zhang Yanan, Huang Di, Wang Yunhong. PC-RGNN: Point cloud completion and graph neural network for 3D object detection [C] //Proc of the 35th AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2021: 3430−3437

[138] Yin Junbo, Shen Jianbing, Gao Xin, et al. Graph neural network and spatiotemporal transformer attention for 3D video object detection from point clouds [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 45(8): 9822−9835

[139] Feng Mingtao, Gilani S Z, Wang Yaonan, et al. Relation graph network for 3D object detection in point clouds[J]. IEEE Transactions on Image Processing, 2021, 30: 92−107

[140] Wu Zhirong, Song Shuran, Khosla A, et al. 3D ShapeNets: A deep representation for volumetric shapes [C] //Proc of 2015 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2015: 1912−1920

[141] Princeton University. Princeton ModelNet project dataset [EB/OL]. [2023-01-04]. http://modelnet.cs.princeton.edu/

[142] Chang A X, Funkhouser T, Guibas L, et al. ShapeNet: An information-rich 3D model repository [J]. arXiv preprint, arXiv: 1512.03012, 2015

[143] Chang A X, Funkhouser T, Guibas L, et al. ShapeNet dataset [EB/OL]. [2023-01-04]. https://shapenet.org/

[144] Armeni I, Sener O, Zamir A R, et al. 3D semantic parsing of large-scale indoor spaces [C] //Proc of 2016 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2016: 1534−1543

[145] Stanford University. S3DIS dataset [EB/OL]. [2023-01-04]. http://buildingparser.stanford.edu/

[146] Hua B S, Pham Q H, Nguyen D T, et al. SceneNN: A scene meshes dataset with annotations [C] //Proc of the 4th Int Conf on 3D Vision (3DV). Piscataway, NJ: IEEE, 2016: 92−101

[147] Hua B S, Pham Q H, Nguyen D T, et al. SceneNN dataset [EB/OL]. [2023-01-04]. http://www.scenenn.net/

[148] Mccormac J, Handa A, Leutenegger S, et al. SceneNet RGB-D: Can 5M synthetic images beat generic ImageNet pre-training on indoor segmentation [C] //Proc of 2017 IEEE Int Conf on Computer Vision (ICCV). Los Alamitos, CA: IEEE Computer Society, 2017: 2697−2706

[149] Mccormac J, Handa A, Leutenegger S, et al. SceneNet dataset [EB/OL]. [2023-01-04]. https://robotvault.bitbucket.io/scenenet-rgbd.html

[150] Dai A, Chang A X, Savva M, et al. ScanNet: Richly-annotated 3D reconstructions of indoor scenes [C] //Proc of 2017 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2017: 2432−2443

[151] Dai A, Chang A X, Savva M, et al. ScanNet dataset [EB/OL]. [2023-01-04]. http://www.scan-net.org/

[152] Hackel T, Savinov N, Ladicky L, et al. Semantic3D. Net: A new large-scale point cloud classification benchmark[J]. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2017, IV-1/W1: 91−98

[153] Hackel T, Savinov N, Ladicky L, et al. Semantic3D dataset [EB/OL]. [2023-01-04]. http://www.semantic3d.net/

[154] Geiger A, Lenz P, Urtasun R. KITTI dataset [EB/OL]. [2023-01-04]. http://www.cvlibs.net/datasets/kitti

[155] Caesar H, Bankiti V, Lang A H, et al. NuScenes: A multimodal dataset for autonomous driving [C] //Proc of 2020 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2020: 11618−11628

[156] Caesar H, Bankiti V, Lang A H, et al. NuScenes dataset [EB/OL]. [2023-01-04]. https://nuscenes.org/

[157] Sun Pei, Kretzschmar H, Dotiwalla X, et al. Scalability in perception for autonomous driving: Waymo open dataset [C] //Proc of 2020 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2020: 2443−2451

[158] Sun Pei, Kretzschmar H, Dotiwalla X, et al. Waymo dataset [EB/OL]. [2023-01-04]. http://www.waymo.com/open

[159] Mao Jiageng, Niu Minzhe, Jiang Chenhan, et al. One million scenes for autonomous driving: ONCE dataset [J]. arXiv preprint, arXiv: 2106.11037, 2021

[160] Mao Jiageng, Niu Minzhe, Jiang Chenhan, et al. ONCE dataset [EB/OL]. [2023-01-04]. https://once-for-auto-driving.github.io/index.html

[161] Xu Xiang, Shuai Hui, Liu Qingshan. Octant convolutional neural network for 3D point cloud analysis[J]. Acta Automatica Sinica, 2021, 47(12): 2791−2800 (in Chinese)
(许翔, 帅惠, 刘青山. 基于卦限卷积神经网络的 3D 点云分析[J]. 自动化学报, 2021, 47(12): 2791−2800)

[162] Wang Jun, Zhu Li. 3D Building facade reconstruction based on image matching-point cloud fusing[J]. Chinese Journal of Computers, 2012, 35(10): 2072−2079 (in Chinese)
(王俊, 朱利. 基于图像匹配-点云融合的建筑物立面三维重建[J]. 计算机学报, 2012, 35(10): 2072−2079)

[163] Qi C R, Chen Xinlei, Litany O, et al. ImVoteNet: Boosting 3D object detection in point clouds with image votes [C] //Proc of 2020 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2020: 4403−4412

[164] Cui Yaodong, Chen Ren, Chu Wenbo, et al. Deep learning for image and point cloud fusion in autonomous driving: A review[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(2): 722−739

[165] Felzenszwalb P F, Huttenlocher D P. Efficient graph-based image segmentation[J]. International Journal of Computer Vision, 2004, 59(2): 167−181

[166] Boykov Y, Funka-Lea G. Graph cuts and efficient N-D image segmentation[J]. International Journal of Computer Vision, 2006, 70(2): 109−131

[167] Chen Zhaomin, Wei Xiushen, Wang Peng, et al. Multi-label image recognition with graph convolutional networks [C] //Proc of 2019 IEEE Conf on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2019: 5172−5181

[168] Hong Danfeng, Gao Lianru, Yao Jing, et al. Graph convolutional networks for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(7): 5966−5978

[169] Zhang Fan, Huang Xianfeng, Fang Wei, et al. Non-rigid registration of mural images and laser scanning data based on the optimization of the edges of interest[J]. SCIENTIA SINICA Informationis, 2012, 42(4): 493−503 (in Chinese)
(张帆, 黄先锋, 方伟, 等. 基于兴趣边缘优化的壁画影像与激光扫描数据非刚性配准[J]. 中国科学：信息科学, 2012, 42(4): 493−503)

[170] Wu Gang, Li Bin, Zhu Qibing, et al. Using color and 3D geometry features to segment fruit point cloud and improve fruit recognition accuracy[J]. Computers and Electronics in Agriculture, 2020, 174: 105475

[171] Quintana B, Prieto S A, Adan A, et al. Door detection in 3D coloured point clouds of indoor environments[J]. Automation in Construction, 2018, 85: 146−166

[172] Ouyang Zhenchao, Dong Xiaoyun, Cui Jiahe, et al. PV-EncoNet: Fast object detection based on colored point cloud[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(8): 12439−12450

[173] Zhang Lingming, Zhao Yue, Li Pengcheng, et al. Dental model segmentation network based on local attention mechanism[J]. Acta Electronica Sinica, 2022, 50(3): 681−690 (in Chinese)
(张凌明, 赵悦, 李鹏程, 等. 基于局部注意力机制的三维牙齿模型分割网络[J]. 电子学报, 2022, 50(3): 681−690)

[174] Li Shirui, Li Qi, Li Haiyang, et al. Real-time accurate 3D reconstruction based on Kinect v2[J]. Journal of Software, 2016, 27(10): 2519−2529 (in Chinese)
(李诗锐, 李琪, 李海洋, 等. 基于 Kinect v2 的实时精确三维重建系统[J]. 软件学报, 2016, 27(10): 2519−2529)

[175] Wang Xiaoting, Wang Lu, Meng Xiangxu. Multi-layers surface reconstruction method for point set with holes[J]. Journal of Software, 2016, 27(10): 2642−2653 (in Chinese)

(王筱婷, 王璐, 孟祥旭. 带洞点云多层同步表面重建方法[J]. 软件学报, 2016, 27(10): 2642−2653)

[176] Deng Lei, Chen Baohua, Lai Weiliang, et al. Interactive calibrating 3D surveillance system based on 3D reconstruction[J]. Acta Electronica Sinica, 2017, 45(3): 527−533 (in Chinese)
(邓磊, 陈宝华, 赖伟良, 等. 三维监控系统中基于三维重构的交互式标定[J]. 电子学报, 2017, 45(3): 527−533)

[177] Yu Yufeng, Zhao Huijing. Off-road localization using monocular camera and nodding LiDAR[J]. Acta Automatica Sinica, 2019, 45(9): 1791−1798 (in Chinese)
(俞毓锋, 赵卉菁. 基于相机与摇摆激光雷达融合的非结构化环境定位[J]. 自动化学报, 2019, 45(9): 1791−1798)

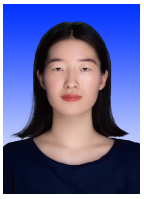
[178] Zhang Junning, Su Qunxing, Liu Pengyuan, et al. An improved VSLAM algorithm based on adaptive feature map[J]. Acta Automatica Sinica, 2019, 45(3): 553−565 (in Chinese)
(张峻宁, 苏群星, 刘鹏远, 等. 一种自适应特征地图匹配的改进VSLAM算法[J]. 自动化学报, 2019, 45(3): 553−565)

[179] Guo Fusheng, Gao Wei, Hu Zhanyi. Automatic generation of large scale true ortho-image from UAV images[J]. SCIENTIA SINICA Informationis, 2013, 43(11): 1383−1397 (in Chinese)
(郭复胜, 高伟, 胡占义. 无人机图像全自动生成大比例尺真正射影像方法[J]. 中国科学: 信息科学, 2013, 43(11): 1383−1397)



Liang Xun, born in 1965. PhD, professor, PhD supervisor. His main research interests include neural networks, social computing, and natural language processing.

梁 循, 1965 年生. 博士, 教授, 博士生导师. 主要研究方向为神经网络、社会计算、自然语言处理.



Li Zhiying, born in 1996. PhD candidate. Her main research interests include point cloud analysis, graph neural networks, and deep learning.

李志莹, 1996 年生. 博士研究生. 主要研究方向为点云分析、图神经网络、深度学习.



Jiang Hongxun, born in 1974. PhD, associate professor. His main research interests include information system engineering, network finance, and social computing.

蒋洪迅, 1974 年生. 博士, 副教授. 主要研究方向为信息系统工程、网络金融、社会计算.