

# 面向点云的三维物体识别方法综述

郝雯 王映辉 宁小娟 梁玮 石争浩

(西安理工大学计算机科学与工程学院 西安 710048)

**摘要** 随着三维扫描技术的快速发展,获取各类场景的点云数据已经非常简单快捷;加之点云数据具备不受光照、阴影、纹理的影响等优势,基于点云的三维物体识别已成为计算机视觉领域的研究热点。首先,对近年来面向点云数据的三维物体识别方法进行归纳和总结;然后,对已有方法的优势及缺点进行分析;最后,指出点云物体识别中所面临的挑战及进一步的研究方向。

**关键词** 点云数据,三维物体识别,特征提取,图匹配

**中图分类号** TP391.4 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2017.09.002

## Survey of 3D Object Recognition for Point Clouds

HAO Wen WANG Ying-hui NING Xiao-juan LIANG Wei SHI Zheng-hao

(Institute of Computer Science and Engineering, Xi'an University of Technology, Xi'an 710048, China)

**Abstract** With the rapid development of 3D scanning technology, it is convenient to obtain point clouds of different scenes. Since point clouds are not influenced by light, shadows and textures, recognizing 3D object from scene point clouds has become a research hotspot of computer vision. This paper first summarized the 3D object recognition methods from point clouds in recent years. Then the advantages and disadvantages of the existing methods were discussed. Finally, the challenges and further research directions of object recognition were pointed out.

**Keywords** Point clouds, 3D object recognition, Feature extraction, Graph matching

## 1 引言

随着三维成像技术的发展,结构光测量与激光扫描已经成为一种快速获取空间数据的重要手段,它们能够快速且精确地获得被测对象表面的三维坐标数据,为获取场景的点云数据提供了全新的技术手段。与传统测量技术相比,这些非接触式三维数据采集技术具有快速性、不接触性、实时性以及扫描精度高等诸多优点。点云数据就是利用结构光扫描仪或者三维激光扫描仪采集到的分布在三维空间中的离散点集,它对复杂场景以及物体的外形表达具有独特的优势,加之其获取的快速性和便捷性,已被广泛地应用在计算机图形学、CAD造型设计、文物古迹保护、树木测量等多个行业。

由于自动导航、自动检测、装配任务以及移动机器人代替人到复杂或危险的环境中执行探测、巡逻任务等大量现实应用的需要,对场景中物体的识别研究引起了人们极大的关注。人类具有强大的识别能力,可以从复杂场景中迅速地识别物体。但要使计算机能够像人类一样自动地识别出场景中的

物体,仍有很大的困难。

目前,基于图像的物体识别已经有很多研究成果<sup>[1-3]</sup>,但是二维图像的成像过程是从三维空间映射到二维空间,这个过程会丢失大量的信息;而且一个无法忽视的事实是,最好的视觉系统应该是面向三维世界的。随着自动化程度的提高,机器人导航、工业零件检测及抓取等众多领域对计算机视觉系统的要求越来越高,因此基于二维图像的物体识别已经无法满足人类的要求。而目前三维点云数据的获取已经非常快捷,同时三维点云数据的采集不受光照影响,规避了二维图像遇到的光照、姿态等问题,因此基于点云数据的三维物体识别也引起了人们的重视。

## 2 识别方法总结

已有的三维点云物体识别方法多是通过分析/提取物体的特征点几何属性、形状属性、结构属性或者多种属性的组合等特征进行比对、学习,从而完成物体的识别与分类。根据所利用特征的不同,已有的三维点云物体识别方法可以分为4类:基于局部特征的物体识别方法、基于全局特征的物体识别

到稿日期:2017-02-06 返修日期:2017-05-11 本文受国家自然科学基金(61602373, 61472319, 61401355),西安市碑林区科技项目(GX1615),陕西省自然科学基金(2015JZ015)资助。

郝雯(1986-),女,博士,讲师,CCF会员,主要研究方向为计算机图形学、模式识别, E-mail: haowensxf@163.com(通信作者);王映辉(1967-),男,教授,博士生导师,主要研究方向为空间物体识别、三维数据场可视化等;宁小娟(1982-),女,副教授,主要研究方向为空间物体识别、三维场景目标分割;梁玮(1985-),女,讲师,主要研究方向为模式识别与图像压缩;石争浩(1968-),男,副教授,主要研究方向为模式识别与图像处理。

方法、基于图匹配的物体识别方法以及基于机器学习的物体识别方法。

## 2.1 基于局部特征的物体识别方法

基于局部特征的物体识别方法主要是通过局部来识别整体<sup>[4]</sup>。该方法无需对处理数据进行分割,往往通过提取物体的关键点、边缘或者面片等局部特征并进行比对来完成物体的识别。其中,特征提取是物体识别中非常关键的一步,它将直接影响到物体识别系统的性能。

点签名法<sup>[5]</sup>(Point Signature, PS)是最早基于点深度信息进行特征描述的方法,其通过物体中每点所在曲线的法向量与参考矢量定义旋转角度。点签名是一个与该旋转角度有关的有向距离(描述向量),通过匹配场景与模型的点签名实现识别。该方法易受噪声的影响,而且一个特征点可能拥有多个描述向量。另外,一些学者提出利用直方图来表示关键点的局部特征,根据所利用属性的不同,其可以分为:基于点空间分布的直方图方法和基于几何属性的直方图方法。

### (1) 基于点空间分布的直方图方法

基于点空间分布的直方图方法通过统计关键点局部邻域的空间分布,生成直方图来描述关键点的局部几何特征。

旋转图像法<sup>[6]</sup>(Spin Image, SI)也是一种典型的基于点的特征描述方法,它是 Johnson 等人于 1999 年提出的一种利用二维数据表征三维特征的自旋图方法。旋转图像的主要思想是一个图像绕法向量旋转  $360^\circ$ ,以图像中每个像素栅格所遇到的点云数作为其灰度值。文献<sup>[7-8]</sup>提出基于旋转图像法的三维点云场景物体识别方法,将每个点用一幅二维旋转图像进行表示。Date<sup>[7]</sup>通过对点云数据进行法向量平均化以及均匀点采样来减小点密度分布不均以及噪声的影响,然后通过比对待识别物体与场景的旋转图像完成目标物体的识别与提取。该方法主要用于识别场景中比较规则的物体,如 CAD 模型等;对于复杂的室外场景物体,计算量会非常大。另外,旋转图像法要求点云数据均匀分布,满足旋转和平移变换不变性,却不满足尺度变换不变性。Frome<sup>[9]</sup>提出三维形状上下文方法(3D Shape Context, 3DSC)来完成物体的识别。三维形状上下文方法将特征点邻域划分为三维球形栅格,通过统计栅格内的点云数据量得到三维形状上下文特征信息,通过比对特征信息完成物体的识别。该方法在特征匹配时需要将场景特征描述向量绕 Z 轴旋转并计算与模型描述特征向量的相关度,计算量较大。随后,Tombari<sup>[10]</sup>对三维形状上下文方法进行了扩展,提出了唯一形状上下文(Unique Shape Context, USC)描述符。该方法避免了对同一个关键点的多个描述,减少了特征匹配的二义性。Zhong<sup>[11]</sup>提出了一种三维形状描述符(Intrinsic Shape Signatures, ISS)来描述点云的局部特征,通过主成分分析法计算得到特征矢量并构建直角参考坐标系,采用类似于文献<sup>[9]</sup>的方法获得特征向量 ISS,通过矩阵计算得到姿态的放置和变换,并将其直接代入识别检索表中作为识别正确性的参数值,以避免由只定义 Z 轴带来的方位模糊。该方法容易受到噪声以及点云数据分布不均的影响,导致特征提取的鲁棒性不强。Guo<sup>[12]</sup>提出一种名为旋转投影统计(Rotational Projection Statistics, RoPS)的局部形状描述符来完成三维点云物体的识别,该描述符是将局部

曲面分别投影到 XOY, XOZ, YOZ 3 个坐标平面后分别统计特征点的分布情况,并利用直方图进行表示。后来,Guo<sup>[13]</sup>又提出一种名为 TriSI (Tri-Spin-Image)的局部形状描述符,该描述符可以有效地表示在柱面坐标系下点的分布情况。该方法首先提出一种改进的局部参考坐标系(Local Reference Frame),利用不同的权重组合选择最佳参数,然后生成 TriSI 特征,最后利用分层特征匹配策略识别场景中的物体。Guo<sup>[14]</sup>也对各类常见的局部描述符进行了比较,并测试了各类局部描述符在噪声、遮挡、网格分辨率变化等情况下的鲁棒性。

### (2) 基于几何属性的直方图方法

基于几何属性的直方图方法是根据关键点局部邻域的几何属性(例如法向量、曲率等微分几何信息)生成直方图来描述关键点的局部几何特征。

Tombari<sup>[15]</sup>和 Salti<sup>[16]</sup>提出了一种能够表示拓扑特征的直方图签名方法(Signature of Histograms Of Orientation, SHOT),该方法具有旋转不变性并对噪声具有鲁棒性。与选择固定的几何特征不同,Taati<sup>[17]</sup>提出一种可变维的局部形状描述符(Variable-Dimensional Local Shape Descriptor, VD-LSD)用于识别深度图像中的物体。该方法首先计算每个点的一系列局部特征(位置、方向和分散性等),然后利用优化过程制定最优性能的可变特征描述符识别物体。另外,Prakhya<sup>[18]</sup>对直方图签名方法<sup>[16]</sup>进行改进,通过将实值向量转化为二值向量,提出了名为 B-SHOT (Binary Signatures of Histograms of Orientations)的三维特征描述符。

Rusu 提出了一种基于几何特征的特征直方图<sup>[19]</sup>(Point Feature Histograms, PFH)描述符,该描述符计算任意两个顶点间的法向量夹角和距离等特征,并映射到直方图中获取统计信息。该算法比较耗时,没法达到实时的要求。随后,Rusu<sup>[20]</sup>在保留 PFH 特征识别能力的基础上,提出了快速点特征直方图(Fast Point Feature Histograms, FPFH)描述符。该描述符记录了某点法向量与其 k 邻域点法向量间的关系信息,不再计算邻域内任意两点间的统计信息,相比 PFH,提高了计算效率。

表 1 对典型的局部三维描述符进行了总结,并对不同方法的性能进行了比较。

表 1 典型的局部 3D 描述符

| 局部描述符  | 作者  | 算法性能                           |
|--------|---|--------------------------------|
| PS     | Chua <sup>[5]</sup>                             | 受噪声影响,特征可能不唯一                  |
| SI     | Johnson <sup>[6]</sup>                          | 要求点云均匀分布                       |
| 3DSC   | Frome <sup>[9]</sup>                            | 优于 Spin Image                  |
| PFH    | Rusu <sup>[19]</sup>                            | 计算量大,耗时                        |
| ISS    | Zhong <sup>[11]</sup>                           | 优于 Spin Image                  |
| FPFH   | Rusu <sup>[20]</sup>                            | 优于 PFH                         |
| USC    | Tombari <sup>[10]</sup>                         | 优于 3DSC                        |
| VD-LSD | Taati <sup>[17]</sup>                           | 优于 Spin Image                  |
| RoPS   | Guo <sup>[12]</sup>                             | 优于 Spin Image, SHOT            |
| SHOT   | Tombari <sup>[15]</sup> 和 Salti <sup>[16]</sup> | 优于 Point Signature, Spin Image |
| TriSI  | Guo <sup>[13]</sup>                             | 优于 Spin Image, SHOT            |
| B-SHOT | Prakhya <sup>[18]</sup>                         | 与 SHOT 相比,速度快且需要的存储空间小         |

## 2.2 基于全局特征的物体识别方法

基于全局特征的方法需要从背景中将目标物体分割出

来,通过描述和比对三维物体整体形状中的全部或者最显著的几何特征来完成物体的识别。这类方法被广泛地应用于3D物体的表示匹配和分类中<sup>[21-22]</sup>。

Rusu<sup>[23]</sup>在FPFH快速点特征直方图描述符<sup>[20]</sup>的基础上提出视点特征直方图(View Feature Histogram, VFH),该方法在计算相对法线间的夹角中加入视点信息,可以区分物体的不同位姿,同时保持旋转缩放不变性。Aldoma<sup>[24]</sup>于2011年对视点特征直方图进行扩展,提出聚类视点特征直方图(Clustered View Feature Histogram, CVFH)描述符。该方法首先根据曲率值去除噪声点或者边界点,然后利用区域增长的方法对点云物体的光滑区域进行分割,最后计算每一个分割区域的视点特征直方图。随后, Aldoma<sup>[25]</sup>通过插值对CVFH进行改进,提出OUR-CVFH(Oriented, Unique and Repeatable Clustered Viewpoint Feature Histogram)描述符,基于该描述符的物体识别准确率有了很大的提高。

另外, Rusu<sup>[26]</sup>首先利用M估计抽样一致性算法对点云场景进行分割,然后基于FPFH快速点特征直方图描述符<sup>[20]</sup>计算每个分割点云簇的局部特征,并利用条件随机场将分割后的点云簇分类为不同的基本形状。在分类的基础上,提出全局快速点特征直方图(Global Fast Point Feature Histograms, GFPPH)描述符,该描述符记录了整个模型中局部几何区域间的关系,并在此基础上,利用支持向量机完成三维模型类型的标记与识别。Marton<sup>[27]</sup>在GFPPH描述符<sup>[26]</sup>的基础上去除点分类这一操作并向形状基元添加几何解,提出全局基于半径的曲面描述符(Global Radius-based Surface Descriptor, GRSD)。该方法首先对点云数据进行栅格化并对曲面进行分类,然后根据基本形状类别,利用RSD计算特征,最后将分类的点云簇归入已有的类别中。

Wohlkinger<sup>[28]</sup>提出形状功能集合描述符(Ensemble of Shape Functions, ESF),该描述符包括角度、点距离和面积3个特征。从三维模型中随机选择3个点,计算点间的距离、三角形面积以及两线段间的角度形成全局描述符,以用于相似形状的检索。Shang<sup>[29]</sup>提出一种名为PWSE(Potential Well Space Embedding)的物体识别方法,该方法是通过比较物体不同姿势下的曲面误差完成物体的识别,基于误差表面最小值确定一系列特征向量作为待识别物体及其姿态的值。Chen<sup>[30]</sup>提出一种新的全局傅里叶直方图描述符,用于识别场景中的物体。该描述符不仅利用了柱面极坐标,而且实现了绕垂直轴旋转无关。Drost<sup>[31]</sup>提出一种新的面向点对特征(Oriented Point Pair Feature)的全局描述符,并利用快速投票机制完成模型的局部匹配,该描述符包括模型中每个点对的特征,代表从点对特征空间到模型的映射,模型上类似的特征会映射到同一地方。该方法对稀疏的点云场景能获得较好的识别效果。

基于全局特征的物体识别方法从整体上完成物体的识别,对三维模型细节的描述不足,对具有相似形状的物体的识别效果并不好,并且该类方法对噪声和遮挡比较敏感<sup>[32]</sup>。

### 2.3 基于图匹配的物体识别方法

基于图匹配的物体识别方法通常将点云数据分解成基本形状,用一个抽象的点代表这些基本形状,并利用拓扑图表示

形状之间的邻近关系,利用图对三维点云物体进行表示,通过图匹配完成对物体的识别。

早期的物体识别多集中在从深度图像中识别规则物体(例如CAD模型)。程义民<sup>[33]</sup>首先对物体的深度图像进行区域分割,然后计算曲面的高斯曲率平均值、平均曲率平均值、曲率直方图和曲率的熵等几何信息。基于这些几何属性,利用一个属性关系图(Attribute Relational Graph, ARG)来描述物体,最后通过与模型库中的模型ARG图进行优化匹配来完成物体的识别。该方法对识别零部件等人造物体有较好的效果。

近年来,已经有学者利用图匹配完成点云场景中物体的识别。Schnabel<sup>[34]</sup>首先利用随机抽样一致性算法(Random Sample Consensus, RANSAC)将点云数据分解成基本形状,将每个基本形状用一个抽象的点表示,并利用拓扑图表示它们之间的邻近关系,通过图匹配完成对物体的识别。该方法用于识别场景中相同的物体,不能识别同类但大小不同的物体;另外,其用到的查询图和约束条件必须经过手工定义。Nieuwenhuisen<sup>[35]</sup>通过创建点云注释形状图,应用于图匹配的方法完成机器人对场景中物体的抓取。Berner<sup>[36]</sup>通过改进其原来的方法<sup>[35]</sup>,将轮廓线与形状信息相结合,完成场景中物体的探测以及姿势的估计。文献<sup>[35-36]</sup>中的实验数据多是由几个纸盒(长方体)和罐子(圆柱体)构成的简单场景,若使用这些方法从室内场景中识别出由复杂曲面构成的物体,会显得力不从心。Zhao<sup>[37]</sup>提出一种基于显著性的特定物体探测方法,该方法首先利用区域增长的方法提取场景中的几何图元(如直线、平面等),然后分析场景中的常见物体(汽车、电线杆等)由哪些几何图元组成,最后利用图匹配的方法完成场景中特定物体的探测。Hao<sup>[38]</sup>首先对不同类型的平面间的连接类型进行定义,通过分析场景中常见物体(楼梯、汽车、书柜等)的结构,构造拓扑结构图,同时记录常见物体的结构编码,通过比对结构编码来完成物体的识别。该方法可以用来识别具有不同尺寸的同类物体,但其只适用于由平面组成的物体识别,对于结构复杂、不能由基本形状表示的自由曲面物体的识别具有一定的局限性。

基于图匹配的物体识别方法多适用于由基本形状组成的物体,对于不能由基本形状表示的自由曲面物体以及由于数据缺失导致结构不完整的物体的识别具有一定的局限性。

### 2.4 基于机器学习的物体识别方法

基于机器学习的物体识别方法通过提取、学习样本的特征,利用分类器模型完成场景中物体的分类与识别。随着计算机视觉技术和认知技术的快速发展,对场景中物体的识别逐渐从对特定物体的识别(如车辆、树木、建筑物等)过渡到对多类物体的识别。马尔可夫随机场<sup>[39]</sup>、支持向量机<sup>[40]</sup>、随机森林<sup>[41]</sup>和条件随机场<sup>[42-43]</sup>等模型的使用增强了场景信息的关联,有效地提高了从场景中识别物体的速度和精度。根据所处理场景的不同,其可以分为:室外场景物体识别方法和室内场景物体识别方法。

#### (1) 室外场景物体识别方法

建筑物、汽车和树木是室外场景中的常见物体。程健<sup>[40]</sup>将三维点云数据与二维栅格相结合,提取反射强度概率分布、

纵向高度轮廓分布和位置姿态 3 个特征,利用支持向量机分类器完成车辆的实时探测。Yao<sup>[41]</sup>提取 LiDAR 数据的光谱、几何和空间上下文属性,然后利用 AdaBoost 分类器来提取城市环境中的树木。Niemeyer<sup>[45]</sup>将随机森林分类器集成到条件随机场框架中,利用反射强度、反射强度的变化、主曲率等特征完成 Lidar 点云场景中建筑物的识别。

上述方法是对场景中的单个物体进行识别,很多学者利用机器学习对场景中的多个物体进行识别,这些方法需要预先对场景进行分割,在单个物体被分割和提取的基础上,通过样本特征学习完成多类物体的分类与识别。Golovinskiy<sup>[46]</sup>首先利用图割的方法对点云场景进行分割,然后提取每个物体的位置、离最近路边的距离等特征信息,最后利用支持向量机完成车辆、路灯等物体的分类和识别。该方法在分割前需要人工对物体进行定位。2012 年,Velizhev<sup>[47]</sup>提出利用隐式形状模型(Implicit Shape Model, ISM)代替支持向量机来识别场景中的汽车、路灯、杆状物等物体,并与文献[46]中的方法进行了比较。Zhao<sup>[48]</sup>和 Zhang<sup>[49]</sup>都首先对点云场景进行分割,然后提取每个分割面片的特征,利用支持向量机进行学习,将场景分为建筑物、地面和树木。Zhao<sup>[48]</sup>首先利用扫描线算法对点云场景进行分割,然后提取每个分割面片的最大/最小高度、法向量分布方差等特征,利用支持向量机进行学习,将场景分为建筑物、地面、树木和行人等。基于扫描线的分割方法在沿扫描剖面进行处理时,只能提取出沿扫描剖面方向曲率变化大的特征点,不能顾及到物体的细节特征。Lehtomäki<sup>[50]</sup>在点云数据栅格化的基础上去除场景中的路面以及建筑物,并利用 CC 标记方法<sup>[51]</sup>对剩余的点云数据进行分割,完成单个物体的提取;接着计算每个分割簇的旋转图像、局部描述符直方图(Local Descriptor Histograms, LDHs)和点分布等几何特征,通过支持向量机识别道路中的树木、路灯杆、交通标志、汽车、行人和围墙等物体。Wang<sup>[52]</sup>将霍夫森林扩展到三维空间,利用霍夫森林完成复杂城市场景中目标物体的识别。该方法首先去除路面点,然后通过栅格化将点云数据进行过分割,将每个栅格及其邻域栅格归为一个局部点云块,接着提取局部点云块的结构和反射特征,利用霍夫森林完成三维点云场景中汽车、路灯和交通标识牌的识别。Wang<sup>[53]</sup>提出了一种基于多尺度层次框架的物体分类方法,首先对点云数据进行多尺度采样,然后将每个尺度的采样点归为多个层次的点云簇,最后基于三层贝叶斯概率模型并利用 AdaBoost 分类器识别场景中的树木、建筑物、车辆以及人。

### (2) 室内场景物体识别方法

室内场景与人类的生活息息相关,也是人类感知和认识这个世界的重要客体。一些学者提出针对室内点云场景的点云物体识别方法。Kim<sup>[54]</sup>认为室内场景通常包含高密度的重复物体,如桌子、椅子和显示器等。根据场景中包含物体的重复性,将每类物体看作是基本形状的组合,利用随机抽样一致性算法提取场景中的基本形状,并计算每类物体在不同姿态下的特征,通过马尔可夫随机场进行特征学习,完成场景中物体的识别。Nan<sup>[55]</sup>提出了一种名为“Search-Classify”的室内点云场景识别方法,该方法首先将场景过分割,然后通过“搜索-分类”的思想将点云场景分割为单个对象(如桌子、凳

子等),利用随机决策森林完成桌子、凳子等室内物体的识别,最后利用模板匹配的方法完成室内场景的重建。庄严<sup>[56]</sup>将室内三维点云数据转化为二维 Bearing Angle 图,利用区域扩张算法提取场景中的平面,将物体碎片及其相对于物体中心的位置作为特征,利用一种基于 Gentleboost 算法的有监督学习方法完成对室内场景中椅子、办公桌、沙发、显示器等物体的认知。

基于机器学习的物体识别方法通常是建立在已经知道场景中包含哪些物体的基础上,通过对样本特征的学习,从场景中识别相应的物体。该类方法处理的场景所包含的物体类别较为固定,具有一定的规律性。

### 3 算法总结及面临的挑战

综上所述,可以发现从三维点云数据中识别物体已经成为当前的研究热点,但其仍旧面临着许多挑战。

(1) 目前提出的基于局部特征的物体识别方法往往需要从大量的点云数据中提取物体的局部特征。由于点云场景数据具有大范围、大尺度和海量性的特性,使得每个物体包含大量的局部特征,每个局部特征对应一个高维描述向量,从而导致计算量大、计算效率低下等问题,因此,如何根据人类的视觉特性,快速从海量数据中提取描述性强的特征,同时减小计算量,是一个值得研究的方向。

(2) 目前多数三维物体识别方法多是针对点云分布较为均匀或者包含物体较少的场景进行物体识别,并没有一种比较统一的与数据密度分布无关的三维点云模型特征计算方法,但实测三维点云场景数据易受噪声的影响,且存在密度分布不均匀和无法保证法向量、曲率等微分几何特征计算的准确性的缺点。因此,提出一种比较统一的与数据密度分布无关的三维点云模型特征计算方法是一个值得研究的课题。

(3) 目前已有的基于图匹配的物体识别方法多是根据物体的形状以及形状间的拓扑关系来完成物体的识别,但是在实际扫描过程中,由于遮挡、噪声的干扰或者单侧扫描等因素,单个物体的完整数据很难获得,导致物体的拓扑结构不完整,已有的方法无法保证物体识别的准确性;而且已有的方法多是针对由基本形状组成的物体,对于结构复杂、不能由基本形状表示的自由曲面物体的识别具有一定的局限性。因此,如何在点云数据缺失以及遮挡严重的情况下提高物体的识别率是值得研究的。另外,如何“化繁为简”,将复杂形状分解为多个基本形状(如平面、圆柱等),从而完成由复杂形状组成物体的识别,也是一个值得研究的课题。

**结束语** 三维点云场景中物体的识别是一个具有重要意义和富有挑战的问题,近年来吸引了越来越多的研究人员的关注。本文以点云场景中物体的识别为主要目标,重点分析和总结了现有的基于点云的三维物体识别方法,讨论了不同类方法的优势与不足,最后指出点云物体识别技术所面临的挑战,并提出了引入视觉感知信息,从人类的视觉特性出发,在快速去除无用信息的基础上有效地从海量点云场景中提取出描述性强且能满足计算量限制的显著特征的解决思路。同时强调,对于无法由基本形状表示的物体,将复杂形状分解成简单的基本形状,并分析形状间的拓扑关系,从而完成场景中物体的识别,也是一个重要的研究方向。

## 参考文献

- [1] SEIDENARI L, SERRA G, BAGDANOV A D, et al. Local pyramidal descriptors for image recognition[J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2014, 36(5): 1033-1040.
- [2] TIMOFTE R, ZIMMERMANN K, VAN GOOL L. Multi-view traffic sign detection, recognition, and 3d localisation[J]. *Machine Vision and Applications*, 2014, 25(3): 633-647.
- [3] LU H, FENG X, LI X, et al. Superpixel level object recognition under local learning framework[J]. *Neurocomputing*, 2013, 120: 203-213.
- [4] GUO Y, BENNAMOUN M, SOHEL F, et al. 3D Object Recognition in Cluttered Scenes with Local Surface Features: A Survey[J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2014, 36(11): 2270-2287.
- [5] CHUA C S, JARVIS R. Point Signatures: A New Representation for 3D Object Recognition [J]. *International Journal of Computer Vision*, 1997, 25(1): 63-85.
- [6] JOHNSON A E, HEBERT M. Using spin images for efficient object recognition in cluttered 3d scenes[J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 1999, 21(5): 433-449.
- [7] DATE H, KANETA Y, HATSUKAIWA A, et al. Object Recognition in Terrestrial Laser Scan Data using Spin Images[J]. *Computer-Aided Design and Applications*, 2012, 9(2): 187-197.
- [8] LIU Y, MA J, ZHAO J, et al. Three dimensional automatic target recognition based on spin-images[J]. *Infrared and Laser Engineering*, 2012, 41(2): 543-548. (in Chinese)  
刘瑶, 马杰, 赵季, 等. 基于自旋图的三维自动目标识别[J]. *红外与激光工程*, 2012, 41(2): 543-548.
- [9] FROME A, HUBER D, KOLURI R, et al. Recognizing objects in range data using regional point descriptors[C]// *European Conference on Computer Vision*. Springer Berlin Heidelberg, 2004: 224-237.
- [10] TOMBARI F, SALT I S, DI STEFANO L. Unique shape context for 3D data description[C]// *Proceedings of the ACM workshop on 3D object retrieval*. ACM, 2010: 57-62.
- [11] ZHONG Y. Intrinsic shape signatures: A shape descriptor for 3d object recognition[C]// *2009 IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops)*. IEEE, 2009: 689-696.
- [12] GUO Y, SOHEL F, BENNAMOUN M, et al. Rotational projection statistics for 3D local surface description and object recognition[J]. *International Journal of Computer Vision*, 2013, 105(1): 63-86.
- [13] GUO Y, SOHEL F, BENNAMOUN M, et al. A novel local surface feature for 3D object recognition under clutter and occlusion [J]. *Information Sciences*, 2015, 293: 196-213.
- [14] GUO Y, BENNAMOUN M, SOHEL F, et al. A comprehensive performance evaluation of 3D local feature descriptors[J]. *International Journal of Computer Vision*, 2016, 116(1): 66-89.
- [15] TOMBARI F, SALT I S, DI STEFANO L. Unique signatures of histograms for local surface description [C] // *European Conference on Computer Vision*. Springer Berlin Heidelberg, 2010: 356-369.
- [16] SALT I S, TOMBARI F, DI STEFANO L. SHOT: unique signatures of histograms for surface and texture description[J]. *Computer Vision and Image Understanding*, 2014, 125: 251-264.
- [17] TAATI B, GREENSPAN M. Local shape descriptor selection for object recognition in range data[J]. *Computer Vision and Image Understanding*, 2011, 115(5): 681-694.
- [18] PRKAHYA S M, LIU B, LIN W. B-SHOT: A binary feature descriptor for fast and efficient keypoint matching on 3D point clouds[C]// *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2015: 1929-1934.
- [19] RUSU R B, BLODOW N, MARTON Z C, et al. Aligning point cloud views using persistent feature histograms [C] // *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2008: 3384-3391.
- [20] RUSU R B, BLODOW N, BEETZ M. Fast point feature histograms (FPFH) for 3D registration [C] // *IEEE International Conference on Robotics and Automation*, 2009 (ICRA '09). IEEE, 2009: 3212-3217.
- [21] BUSTOS B, KEIM D A, SAUPE D, et al. Feature-based similarity search in 3D object databases[J]. *ACM Computing Surveys (CSUR)*, 2005, 37(4): 345-387.
- [22] BRONSTEIN A M, BRONSTEIN M M, KIMMEL R. Numerical geometry of non-rigid shapes [M]. New York: Springer, 2008.
- [23] RUSU R B, BRADSKI G, THIBAUX R, et al. Fast 3d recognition and pose using the viewpoint feature histogram[C]// *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2010: 2155-2162.
- [24] ALDOMA A, VINCZE M, BLODOW N, et al. CAD-model recognition and 6DOF pose estimation using 3D cues[C]// *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*. IEEE, 2011: 585-592.
- [25] ALDOMA A, TOMBARI F, RUSU R B, et al. OUR-CVFH-oriented, unique and repeatable clustered viewpoint feature histogram for object recognition and 6DOF pose estimation[M]// *Pattern Recognition*. Springer Berlin Heidelberg, 2012: 113-122.
- [26] RUSU R B, HOLZBACH A, BEETZ M, et al. Detecting and segmenting objects for mobile manipulation [C] // *2009 IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops)*. IEEE, 2009: 47-54.
- [27] MARTON Z C, PANGERCIC D, BLODOW N, et al. Combined 2D-3D categorization and classification for multimodal perception systems [J]. *The International Journal of Robotics Research*, 2011, 30(11): 1378-1402.
- [28] WOHLKINGER W, VINCZE M. Ensemble of shape functions for 3d object classification[C]// *2011 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2011: 2987-2992.
- [29] SHANG L, GREENSPAN M. Real-time object recognition in sparse range images using error surface embedding[J]. *International Journal of Computer Vision*, 2010, 89(2/3): 211-228.
- [30] CHEN T, DAI B, LIU D, et al. Performance of global descrip-

- tors for velodyne-based urban object recognition [C] // 2014 IEEE Intelligent Vehicles Symposium Proceedings. IEEE, 2014: 667-673.
- [31] DROST B, ULRICH M, NAVAB N, et al. Model globally, match locally: Efficient and robust 3D object recognition [C] // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2010: 998-1005.
- [32] LAM J. Object recognition by registration of repeatable 3d interest segments [D]. Queen's University, Canada, 2015.
- [33] CHENG Y M, DING H X, WANG Y X, et al. Curved Object Recognition Based on Geometrical Features [J]. Journal of Image and Graphics, 2000, 5(7): 573-579. (in Chinese)  
程义民, 丁红侠, 王以孝, 等. 基于几何特征的曲面物体识别 [J]. 中国图像图形学报, 2000, 5(7): 573-579.
- [34] SCHNABEL R, WESSEL R, WAHL R, et al. Shape recognition in 3d point-clouds [C] // The 16-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision. 2008.
- [35] NIEUWENHUISEN M, STÜCKLER J, BERNER A, et al. Shape-primitive based object recognition and grasping [C] // 7th German Conference on Robotics, Proceedings of ROBOTIK 2012. VDE, 2012: 1-5.
- [36] BERNER A, LI J, HOLZ D, et al. Combining contour and shape primitives for object detection and pose estimation of prefabricated parts [C] // 2013 IEEE International Conference on Image Processing. IEEE, 2013: 3326-3330.
- [37] ZHAO Y, HE M, ZHAO H, et al. Computing object-based saliency in urban scenes using laser sensing [C] // 2012 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2012: 4436-4443.
- [38] HAO W, WANG Y. Structure-based object detection from scene point clouds [J]. Neurocomputing, 2016, 191: 148-160.
- [39] AGRAWAL A, NAKAZAWA A, TAKEMURA H. MMM-classification of 3D Range Data [C] // IEEE International Conference on Robotics and Automation, 2009 (ICRA'09). IEEE, 2009: 2003-2008.
- [40] CHENG J, XIANG Z Y, YU H B, et al. Real-time vehicle detection using 3D lidar under complex urban environment [J]. Journal of Zhejiang University (Engineering Science), 2014, 48(12): 2101-2106. (in Chinese)  
程健, 项志宇, 于海滨, 等. 城市复杂环境下基于三维激光雷达实时车辆检测 [J]. 浙江大学学报(工学版), 2014, 48(12): 2101-2106.
- [41] SUN J, LAI Z L. Airborne LiDAR feature selection for urban classification using random forests [J]. Geomatics and Information Science of Wuhan University, 2014, 39(11): 1310-1313. (in Chinese)  
孙杰, 赖祖龙. 利用随机森林的城区机载 LiDAR 数据特征选择与分类 [J]. 武汉大学学报, 2014, 39(11): 1310-1313.
- [42] NIEMEYER J, ROTTENSTEINER F, SOERGEL U. Conditional random fields for lidar point cloud classification in complex urban areas [J]. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2012, 1(3): 263-268.
- [43] ZHUANG Y, LIU Y, HE G, et al. Contextual classification of 3D laser points with conditional random fields in urban environments [C] // 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2015: 3908-3913.
- [44] YAO W, WEI Y. Detection of 3-D individual trees in urban areas by combining airborne LiDAR data and imagery [J]. IEEE Geoscience and Remote Sensing Letters, 2013, 10(6): 1355-1359.
- [45] NIEMEYER J, ROTTENSTEINER F, SOERGEL U. Contextual classification of lidar data and building object detection in urban areas [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2014, 87: 152-165.
- [46] GOLOVINSKIN A, KIM V G, FUNKHOUSER T. Shape-based recognition of 3D point clouds in urban environments [C] // 2009 IEEE 12th International Conference on Computer Vision. IEEE, 2009: 2154-2161.
- [47] VELIZHEV A, SHAPOVALOV R, SCHINDLER K. Implicit shape models for object detection in 3D point clouds [C] // ISPRS Congress. 2012: 179-184.
- [48] ZHAO H, LIU Y, ZHU X, et al. Scene understanding in a large dynamic environment through a laser-based sensing [C] // 2010 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2010: 127-133.
- [49] ZHANG J X, LIN X G. Object-based classification of urban airborne lidar point clouds with multiple echoes using SVM [J]. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, 2012, 3: 135-140.
- [50] LEHTOMÄKI M, JAAKKOLA A, HYYPPÄ J, et al. Object classification and recognition from mobile laser scanning point clouds in a road environment [J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(2): 1226-1239.
- [51] AWAN S, MUHAMAD M, KUSEVIC K, et al. Object class recognition in mobile urban lidar data using global shape descriptors [C] // 2013 International Conference on 3DTV-Conference. IEEE, 2013: 350-357.
- [52] WANG H, WANG C, LUO H, et al. Object detection in terrestrial laser scanning point clouds based on Hough forest [J]. IEEE Geoscience and Remote Sensing Letters, 2014, 11(10): 1807-1811.
- [53] WANG Z, ZHANG L, FANG T, et al. A multiscale and hierarchical feature extraction method for terrestrial laser scanning point cloud classification [J]. IEEE Transactions on Geoscience and Remote Sensing, 2015, 53(5): 2409-2425.
- [54] KIM Y M, MITRA N, YAN D, et al. Acquisition of 3D Indoor Environments with Variability and Repetition [J]. ACM Transactions on Graphics, 2012, 31(6): 138.
- [55] NAN L, XIE K, SHARF A. A search-classify approach for cluttered indoor scene understanding [J]. ACM Transactions on Graphics (TOG), 2012, 31(6): 137.
- [56] ZHUANG Y, LU X B, LI Y H, et al. Mobile Robot Indoor Scene Cognition Using 3D Laser Scanning [J]. Acta Automatica Sinica, 2011, 37(10): 1232-1240. (in Chinese)  
庄严, 卢希彬, 李云辉, 等. 移动机器人基于三维激光测距的室内场景认知 [J]. 自动化学报, 2011, 37(10): 1232-1240.