

数据融合

王逸舟

目录

0.1. 基于激光雷达点云与图像融合的车辆目标检测方法	1
1. 夜间	1

0.1. 基于激光雷达点云与图像融合的车辆目标检测方法. [?] 提出了一种基于 4 线激光雷达 (LADAR) 与摄像头融合的方案, 用于提高智能车辆对车辆目标的检测精度。首先调用卷积神经网络来识别图像中的目标, 然后将点云与图像数据进行空间匹配, 最后采用 R-Tree 算法快速配准检测框与相应的点云数据。利用点云的深度信息就能获得目标的准确位置。经过真实道路场景采集的图像与点云数据进行测试, 结果表明: 该融合算法将漏检概率 (FN) 从 Mask R-CNN 方法的 14.86% 降低到 8.03% ; 因而, 该融合算法能够有效的降低图像漏检的概率。

0.1.1. 前人的工作. 目前大多数研究集中于图像数据与多线激光雷达 (laser detection and ranging,LADAR), 比如文献 [5] 采用 64 线激光雷达点云数据与图像数据融合的方式, 利用点云的深度信息与插值算法获得深度图像 (RGB-D), 最终利用特征级融合提高了检测精度; 文献 [6] 也基于深度图像数据提出一种三维物体检测框架——Frustum PointNets。该方法直接在原始点云数据上操作, 分别利用成熟的二维目标检测和三维目标检测的深度学习网络实现定位, 实现了较好的检测性能; 文献 [7] 提出融合网络——多模态数据融合的多视图三维物体检测网络 (multi-view 3D, MV3D)。其利用卷积层分别从鸟瞰图 (aerial viewAV) 和前视图 (front view,FV) 提取点云特征, 并在鸟 R-Tree 算法, 快速地将检测框与框内的点云关联起来, 瞰图分支利用鸟瞰视角的点云产生高度精确的 3D 候选框, 然后将特征投射回 AV 和 FV 以及图像的特征层, 最后融合 3 个分支的特征。通过这种多视图的编码方案, 能够获得对稀疏 3D 点云更有效和紧凑的表达。激光雷达的线束越多, 提供的信息越多, 越有利于准确识别目标, 但往往计算资源的要求、方案的成本也越高, 不利于商业化。

0.1.2. 数据融合. 由图 1 可知: 利用点云信息的前提是将检测框与框内的点云进一步关联起来。常规的做法是循环遍历每个雷达数据点, 逐一判断其是否在每个检测框内。这种做法的缺陷在于效率低下。假设每张图像中有 M 个检测框和 N 个点云数据, 那么循环遍历的时间复杂度为 $O t (MN)$ [23], 并且对于不同交通场景下的物体的数量差别较大, 这意味测距效率不稳定, 进而会影响整个检测系统的效率。理论上, 图像上的检测框与点云数据的相对空间结构关系是固定的, 因此可以通过一次遍历所有检测框与数据点来锁定相应的位置关系, 其时间复杂度为 $O t (M+N)$ 。为了解决检测框与数据点匹配的效率问题, 引入 R-Tree 算法 [24] 。

3.1 R-Tree 算法 R-Tree 是一种按照对象的空间位置关系来组织数据的动态平衡树, 具有很强的灵活性和可调节性。如图 4a 所示, 从叶子结点开始, 每一个结点的父亲结点在空间上完全包含其所有的子结点。逐层上溯, 形成一个完整的、描述了不同层次的结点之间的空间包含关系的数据结构。

1.0.2. 应用激光雷达与相机信息融合的障碍物识别. 针对探测器在地外星体表面软着陆过程中的障碍物识别问题, 提出了一种融合三维点云数据与灰度图像数据进行精确障碍物识别的方法; 首先利用坐标转换将灰度图像与三维点云归一化到同一坐标系下, 实现传感器数据的融合; 然后采用改进 K 均值聚类算法对预处理后灰度图像进行图像分割, 生成光学障碍图; 最后利用开源库 PCL(pointcloudlibrary) 对激光雷达生成的三维激光点云数据进行处理, 采用随机采样一致性算法提取着陆区地形水平面, 对去除水平面后的点云数据进行点云分割, 分离出突起物、凹坑等障碍物, 并通过激光雷达与相机转换坐标系, 投影到像平面, 生成最终障碍图。图像预处理 通过分析发现星体表面灰度图像中的噪声主要包含加性噪声和随机干扰噪声, 为此首先采用中值滤波器对图像预处理, 滤除图像的噪声。同时为更好地适应光照对图像采集的影响, 采用直方图均衡化的方式, 对灰度图像进行增强” 1。

1.0.3. 基于 ROS 与三维点云图像的室内物体精准定位. 考虑到 ROS SLAM 构建的地图只能描述环境的二维信息, 三维点云图像只能描述物体独立的三维信息等特点,[?] 融合了 ROS SLAM 的 Gmapping 算法构建的室内二维地图与物体的三维点云图像信息, 提出了一种复合坐标定位系统。首先对不同室内进行分类, 进行一维坐标的标定, 其次通过对 Gmapping 算法构建好的地图等进行二三维坐标标定, 再结合空间信息构成外部坐标系 p, 最后通过对采集到的物体三维点云坐标进行仿射运算获得物体基于外部坐标三维坐标, 结合一维坐标, 对物体进行复合四维坐标定位。整个定位实验数据表明, 物体室内的位置平均测量误差只有 4.2cm, 其定位精度比起常见的超声波与红外线定位系统提高 6.7%, 比基于蓝牙角度测量的定位系统定位精度提高 20%, 比超宽带定位系统提高 72%。物体定位误差小, 定位精准。

1.0.4. 基于激光雷达和摄像头信息融合的车辆检测算法. 针对摄像头在无人驾驶系统车辆检测中易受环境干扰的问题, 通过激光雷达数据和摄像头图像进行融合, 提出了一种强鲁棒性实时车辆检测算法。首先, 将三维激光雷达点云通过深度补全方法转换为和图像具有相同分辨率的二维密集深度图。然后将彩色图像和密集深度图分别通过 YOLOv3 实时目标检测框架得到各自的车辆检测信息。最后, 提出了决策级融合方法将两者的检测结果进行融合, 得到了最终的车辆检测结果。在 KITTI 数据集上对算法进行评估, 实验结果表明该算法完全满足无人驾驶车辆所需的强鲁棒性、强实时性和高检测精度的要求。

整个系统由 3 部分组成, 分别为深度补全、车辆检测和决策级融合。系统整体结构如图 1 所示

1. 深度补全在深度补全之前, 需要先进行预处理操作, 将三维激光点云转换为二维稀疏深度图。在预处理过程中, 要将激光雷达和摄像头进行精确校准和联合标定, 从而可以精确地将每个三维激光雷达点云帧投影到二维彩色图像平面上, 形成稀疏的深度图。传感器之间的坐标转换关系如图 2 所示。雷达三维点云投影成二维深度图稀疏深度图补全为密集深度图 2 车辆检测本文选用 YOLOv3 进行车辆检测,YOLOv3 在 2 个训练集 (彩色图像和密集深度图) 上

分别进行训练, 最终得到 2 个训练好的模型。3. 决策级融合本部分依据深度图像和彩色图像在 YOLOv3 中的检测结果, 将得到的边界框信息和相应的置信度进行融合, 从而得到最终的检测结果。首先将边界框进行融合, 通过判断深度图像目标边界框和彩色图像目标边界框交并比 (IOU) 的大小, 选择不同的融合策略: 当交并比小于 0.5 时, 认为是 2 个独立的检测目标, 不进行融合;

1.0.5. 基于数据融合的目标测距方法研究. 单目视觉进行目标识别有着巨大优势, 但在目标测距方面存在精度不足且测量过程不稳定的问题, 一种基于 4 线激光雷达与摄像头融合的联合测距的方法被提出并改善这一问题。首先利用卷积神经网络检测图像中的目标, 得到相应的检测框; 与此同时, 通过标定相机内外参, 将三维的激光点云数据转换到二维平面, 得到 2 种数据对于检测环境的一致性表达。然后利用 R-Tree 算法快速配准检测框与相应的点云数据。此时, 利用点云的深度信息能获得目标在真实世界的位置, 并提出联合测距的方法来进一步提高测距精度。最终经实车采集的数据验证了所提算法的有效性。

1.0.6. 基于信息融合的无人驾驶图像数据处理方法. [?] 在无人驾驶领域的研究中, 激光雷达和图像识别始终是支撑其发展的两大热点。本研究利用 MobileNet 神经网络结合加速器对摄像头图像进行处理, 再针对激光雷达的云点数据采用聚类分析的方法进行处理, 最终将云点数据与图像数据进行叠加生成深度信息, 完成对障碍物的识别。

根据阿里巴巴达摩院公布的 2019 年十大科技趋势, 无人驾驶技术进入冷静的发展期, 无人驾驶技术的环境复杂性使得无人驾驶的研究方向不再仅仅局限于真正行驶在路上的汽车, 而是面向于自动公交, 无人车物流配送, 园区循环利用等相对固定的应用模式。无人驾驶目前更加专用化、小型化、嵌入式化。激光雷达有精确度高, 探测范围广, 受环境干扰小的优点, 图像识别结合深度学习, 可以自主学习, 具有优化能力。因此以激光雷达、基于深度学习技术的图像识别为基础的无人车传感系统将是一个理想的方案。本研究采用 MobileNet 的轻量化网络结合 Movidius 加速器对图像数据信息进行处理, 再采用 MK60 主控板对激光雷达数据进行采样, 将图像数据信息与激光雷达信息进行叠加生成深度图, 最终完成对障碍物的识别。

1 基于 MobileNet 的图片数据处理基于 MobileNet 的图像处理算法在小型化嵌入式系统中有着良好的应用, 其以低功耗为主要目标, 设计功耗低于 15 瓦。1.1 图像预处理图像信息采集中最根本的是图像的处理, 而采集到的图像往往是不准确的。首先, 由于存在大量干扰信息, 因此最初采集到的图像信息中往往含有错误信息, 需要对图像进行初步的处理。本研究首先对最初的图像做中值滤波的处理, 以减小噪声对图像质量的干扰。其次, 由于图像采集装置往往与水平地面存在一定的夹角, 因此对采集到的图像有一定畸变, 需要对滤过波的图像再采取畸变的分析与矫正, 以便采集的图像信息数据更为准确和方便以后使用。1.2 图像信息数据的处理结果经过多种网络模型的组合比较和试验, 决定采用卷积神经网络 (CNN) 和 MobileNet 模型来进行图像信息数据的处理。处理结果如图 1 所示, 原图经过预处理之后, 成功识别到图

中的人物信息,并返回坐标等信息。1.3 模型优化 MobileNet 模型配合使用 TensorFlow 深度学习框架,进行数值计算,可移植性更强。第一步,将数据集划分,通过代码转化,将采集和处理好的数据集分为训练数据集与测试数据集,再用 MobileNet 模型训练。第二步,当模型训练好后,储存测试数据集,并进行测试。如果模型没有训练成功,继续对模型训练,最终训练好模型。

2 在程度不同的相似性,找出一些能够度量云点或变量之间相似程度的统计量,以此为依据划分,把一些相似程度较大的指标聚合为一类,直到把所指标聚合完毕。目前常用的聚类分析方法有 K-Means、K-Means++ 以及 ISOData,为了使设备小型化、低功耗化,并且在不影响性能的前提下,决定采用较为简单的斜率分析法对激光雷达返回的数据进行分析。激光雷达的聚类分析 2.1 激光雷达与聚类分析激光雷达精度高,识别范围广,但其数据量大。所以,如何从海量的数据中挖掘有用的数据进行分析,是研究激光雷达的重中之重。聚类分析的基本思想,在于所研究的云点或变量之间存

2.2 数据的聚类单线激光雷达数据以无人车为中心,遵循一定方向与速度向四周未知环境进行扫描采集,经上位机处理后返回的数据为角度与距离,由此便可以直观地知道障碍物与无人车的位置关系,如图 2 所示。

如图 2 中所示,同一物体是由连续的点所组成,也就是雷达在物体表面所扫描出来的一系列点。以这些点为基准作散点图,如图 3 所示,同一类物体高度基本不变,所以点与点之间斜率基本为零,而两类物体之间存在距离的突变,临界两点间的斜率会有剧烈的波动。定义以下公式用于求数据的斜率: $\Delta d = \Delta n \pm \epsilon$, $n = 0, 1, 2 \dots (1)$ 其中 n 为角度, d_n 为对应角度下的距离, ϵ 为误差容忍度, Δn 是角度差, Δd 是距离差, $\Delta d / \Delta n$ 是所求得的斜率,处理后的数据图如图 4 所示。

由前所述可知,图片经 MobileNet 处理后返回一系列数据,包括已识别目标在图片上的位置坐标等相关信息,但是缺乏目标的深度信息,因此通过激光雷达来获取深度信息。如图 5 所示,摄像头与雷达在同一铅垂线上,已知摄像头的视角为 θ 并与雷达极坐标下的角度相对应,照片宽度为 1,则每一像素点对应角度为 θ / W ,通过目标的位置坐标可得到目标相对于摄像头的角度,也就得到了目标对应于雷达的角度。通过雷达聚类分析可知,在图像识别的目标处,雷达亦做好了物体的聚类,经上位机处理的雷达数据又包含角度与距离信息,因此可通过角度 θ 获得目标的深

度信息,即 L 值,无人车与障碍物的直线距离,通过极坐标与直角坐标间的转换关系,也可求出障碍物相对对于无人车前进道路上的距离。获得目标的深度信息后,无人车能够判断车与物体的距离,从而更好地进行避障等功能。

由这两部分数据进行固定角度于像素之间的叠加之后就可以在一张 2D 图片的基础上形成一张具有特定识别物体深度信息的深度图片,得到此图片后对于后续的无人车路径规划,避障与跟随等功能都可以很方便地实现,叠加效果如图 6 所示。

1.0.7. 基于融合感知的场景数据提取技术研究. [?] 驾驶场景数据是智能网联汽车自动驾驶技术研发的重要基础, 是自动驾驶算法开发以及产品测试的核心资源。提出一种基于视觉与激光雷达融合感知技术的场景数据提取方法, 通过视觉感知技术处理场景采集的视频数据, 获取车辆周围道路、基础设施、行人、车辆等场景动态和静态元素, 并与激光雷达探测结果进行匹配, 从而从采集平台采集到的传感器数据中提取出符合 OpenScenario 格式的驾驶场景数据。并且以中汽中心驾驶场景数据采集平台所采集到的数据为实验对象, 通过与单目视觉场景采集平台采集的数据进行对比, 验证该方法的准确性。

1.0.8. 基于图像和激光雷达点云数据融合的非结构化道路识别. 无人驾驶汽车作为智能车最终的发展方向, 它的安全行驶依赖对周围环境的准确理解。现实环境中, 存在大量道路特征不明显的非结构化道路区域, 基于道路模型的道路识别方法具有一定的局限性。基于可通行区域的像素级别的区域分割可以很好避免模型假设带来的局限性。通过对获得的数据进行像素级别的语义分割, 划分出图像中的可通行区域, 对于道路识别尤其是非结构化道路识别具有重要意义, 可以为智能车提供准确的可通行区域, 提高智能车行驶的安全性。本文提出了一种基于深度学习和数据融合的非结构化道路识别算法, 主要目的是提取出道路的可通行范围, 具体的过程为: (1) 对 KITTI 数据集进行预处理, 得到训练用的图片和标签, 对点云进行多余点筛选和空间投影。(2) 建立 FCN-8s 模型, 利用处理后的数据训练模型, 得到用于分割的深度学习模型并进行分割和特征提取。(3) 利用处理后的点云, 根据高度信息和横向信息分割出落在路面上的点云, 之后利用 alphaShapes 算法提取出点云的边界, 通过判断像素是否在边界内来获得道路分割。(4) 利用深度学习得到的数据和点云的分割结果, 构建马尔科夫随机场模型, 并利用迭代条件峰值算法求解能量函数的全局最小值, 能量最小值对应的像素状态就是最终的分割结果, 并利用行业内常用的指标对深度学习分割结果、马尔科夫随机场模型分割结果进行评价和对比。

1.0.9. 机载激光点云与影像数据融合方法的研究. [?] 影像信息可以反映综合地表纹理、光谱信息及空间属性信息。结合高分辨率的影像信息, 机载激光点云数据即可成为获取 DEM 及地形特征的良好数据源。讨论了数据融合的两种方式, 结果表明, 基于多元数据拥有丰富的物方信息, 其融合数据在 DEM 过滤及特征提取方面有优势。

LiDAR 数据的分割, 是将一组具有近似特征的点分成一块。对野外的激光扫描数据, 通常对其进行同类分割 (例如屋顶、广场和裸露地表)。对于 DEM 提取和特征提取, 许多研究需要这样的属性信息。因此, 出现一系列分割算法, 其主要基于 LiDAR 数据。例如, Tovar 博士和 Pfeifer 的强健内插分割 [7], 分割的结果与点密度相关联。对于低密度点云数据, 处理比较简单, 但是缺少空间及纹理信息导致寻找分割边界困难。对于高密度点云数据, 处理较复杂, 且需要将其分成多个小块处理。LiDAR 数据提供高精度的三维点, 但缺少纹理信息及特征线。相反, 高空间分辨率的光学影像可以提供特征线及丰富的纹理信息 [8]。因此, 本文将激光数据与影像数据结合在一起即可互补各自的缺点, 充分发挥各自的优势。所采用的联合分割算法是基

于多层数据组, 包括基于相邻的几何拓扑关系 (例如高度、坡度和平面向量) 的点云数据, 以及从高分辨率影像中获取的空间属性。图 3 为分割结果。不同的分割块以不同颜色显示, 聚类物以不同的层进行划分, 并且地表不同部分以各自断裂线划分。图 3 是对原始激光数据进行分类后的结果对比, 在分类后的图像中, 不同类地物可以明显分辨出来。例如, 图中地面点被分割成一个几乎覆盖整个测区的分割块, 而建筑屋顶都被分割为较小的独立分割块。

1.0.10. 图像语义分割辅助的车载激光点云道路提取方法. [?] 针对车载激光扫描系统获得的点云数据量大, 难以获得有效特征进行分割分类提取道路的现状, 提出一种深度学习图像语义分割辅助的激光点云道路提取方法。采用**二维图像语义分割、数据融合配准粗分类、三维霍夫变换点云平面分割拟合和局部优化点云细分类**的四步工作流程对车载激光点云进行道路提取。在 2 段不同的城市道路点云数据中进行提取与评测, 获取的道路数据正确率与完整率均达到 99% 以上, 提取质量优异, 可满足实际应用需求。经实验分析, 该方法可有效提取不同道路情况的道路点云, 对点云数据的原始条件约束较少, 相比其他方法在普适性和鲁棒性上都有大幅的提升。

1.3 三维霍夫变换平面分割与拟合被赋类别的点云结构多样, 利用三维霍夫变换从点云数据中提取空间平面找到和道路点云结构相符的点云信息。首先找到经过这一点的所有平面, 使用球坐标表示平面方程, 即

1.0.11. 虚拟现实技术的三维图像重建系统. 为了提高三维图像质量, 设计基于虚拟现实技术的三维图像重建系统。首先, 提取三维图像点云信息, 通过虚拟现实和帧同步技术融合点云数据, 利用中值滤波技术去除图像噪声; 然后, 采用虚拟现实技术对降噪后图像开展点云优化, 自动检测图像点云的特点, 过滤掉噪声点, 有效覆盖图像表面微小矩形面片集, 提升图像三维重建完整度, 通过匹配前景得到点云位置; 最后, 重建过程中调整点云数据为高斯正太分布项目, 有效缩短图像前景和背景间重建差异, 实现三维图像重建。实验结果表明, 该系统的帧跟踪能力强, 重建三维图像清晰、分辨率高。

1.0.12. 基于视觉和激光数据融合的 3D 多目标跟踪. 多目标跟踪是自动驾驶领域的一个重要研究课题。通过精准和有效的跟踪, 自主车辆可以获知视野内车辆的速度并做出相应的运动规划。不同于大多数单独使用视觉或 3D 激光雷达数据的方法, 致力于融合当前自动驾驶车辆上标准配置的相机和激光雷达获得的视觉和 3D 点云信息, 从而达到跟踪感知物体的目的。首先, 使用**匈牙利算法**作为基本模型来关联相邻帧间的每一个物体的 3D 点云。之后, 使用 RGB 图像中的**外观特征**和 3D 点云的**几何特征**来纠正由于物体相近导致的目标索引 (ID) 互换。在新公开的 BLVD 数据集上进行算法评估, 结果表现出了良好的跟踪性能。

使用位置关联矩阵的匈牙利算法是基于在行车中帧间距离更小的物体对有更大的可能性属于同一个物体的假设。因为多目标跟踪的主要任务是进行对象关联, 所以没有把工作的重点放在对象检测模块上, 而是以三维对象的真值作为输入。

1.0.13. 基于面特征和 *SIFT* 特征的 *LiDAR* 点云与航空影像配准. 针对 *LiDAR* 数据与航空影像融合中的配准问题, 提出一种将面特征与点特征相结合的配准方法, 先由 *LiDAR* 点云生成深度影像, 对深度影像和航空影像提取面特征, 在此基础上采用 *SIFT* 算子提取点征, 完成 *LiDAR* 点云与航空影像的配准。文中方法采取了由面特征到 *sIFT* 特征的配准策略, 减少了面特征配准的数据量和 *sIFT* 算法的计算量。从 *1sPRS* 提供的数据集中选取了 3 组数据进行实验, 实验结果明该方法能有效减少 *SIFT* 算子的特征描述符的数量, 减少寻找正确匹配点的时间, 在保证配准精度的情首特特表况下提高配准的效率, 适用于城市地区等包含大量面特征地区的 *LiDAR* 点云与航空影像配准。