

基于相机与激光雷达融合的目标定位与跟踪

张普^{1,2,3,4}, 刘金清^{1,2,3*}, 肖金超⁴, 熊俊峰⁴, 冯天伟⁴, 王忠泽⁴¹福建师范大学医学光电科学与技术教育部重点实验室, 福建 福州 350007;²福建省光子技术重点实验室, 福建 福州 350007;³福建省光电传感应用工程技术研究中心, 福建 福州 350007;⁴广州工业智能研究院, 广东 广州 511458

摘要 环境感知是无人驾驶的关键技术, 针对相机缺乏深度信息无法定位检测目标以及目标跟踪精度较差的问题, 提出一种基于相机与激光雷达融合的目标定位与跟踪算法。该算法通过图像检测框内的激光雷达点云簇在像素平面的面积比例大小获得检测目标的定位信息, 然后根据检测目标的轮廓点云在像素坐标系下的横向移动速度和纵向移动速度融合图像检测框中心坐标提高目标跟踪精度。实验结果表明: 所提目标定位算法正确率为 88.5417%, 且平均每帧处理时间仅为 0.03 s, 满足实时性要求; 图像检测框中心横坐标的平均误差为 4.49 pixel, 纵坐标的平均误差为 1.80 pixel, 平均区域重叠率为 87.42%。

关键词 传感器融合; 机器视觉; 3D 激光雷达; 目标定位; 目标跟踪

中图分类号 TP183

文献标志码 A

DOI: 10.3788/LOP231537

Target Localization and Tracking Method Based on Camera and LiDAR Fusion

Zhang Pu^{1,2,3,4}, Liu Jinqing^{1,2,3*}, Xiao Jinchao⁴, Xiong Junfeng⁴,
Feng Tianwei⁴, Wang Zhongze⁴¹Key Laboratory of Optoelectronic Science and Technology for Medicine of Ministry of Education,
Fujian Normal University, Fuzhou 350007, Fujian, China;²Fujian Provincial Key Laboratory of Photonic Technology, Fujian Normal University,
Fuzhou 350007, Fujian, China;³Fujian Provincial Engineering Technology Research Center of Photoelectric Sensing Application,
Fujian Normal University, Fuzhou 350007, Fujian, China;⁴Guangzhou Institute of Industrial Intelligence, Guangzhou 511458, Guangdong, China

Abstract Environmental perception is a key technology for unmanned driving. However, cameras often lack depth information to locate and detect targets and have poor tracking accuracy; therefore, a target localization and tracking algorithm based on the fusion of camera and LiDAR technologies is proposed. This algorithm obtains the positioning information of the detected target by measuring the proportion of the area of the LiDAR point cloud cluster in the pixel plane within the image detection frame. Subsequently, based on the horizontal and vertical movement speeds of the detected target's contour point cloud in the pixel coordinate system, the center coordinate of the image detection frame is fused to improve the target tracking accuracy. The experimental results show that the accuracy of the proposed target localization algorithm is 88.5417%, and the average processing time per frame is only 0.03 s, meeting real-time requirements. The average error of the horizontal axis of the image detection frame center is 4.49 pixel, the average error of the vertical axis is 1.80 pixel, and the average area overlap rate is 87.42%.

Key words sensor fusion; machine vision; 3D LiDAR; target localization; target tracking

收稿日期: 2023-06-15; 修回日期: 2023-07-17; 录用日期: 2023-08-01; 网络首发日期: 2023-08-18

基金项目: 国家自然科学基金委员会面上基金(62273332)、南沙区重点领域科技项目(2022ZD016)

通信作者: *jqliu8208@fjnu.edu.cn

1 引言

近些年来无人驾驶受到广泛关注,而可靠的环境感知系统是无人驾驶安全行驶的基础要求和前提条件。感知系统对动态目标进行定位和跟踪,为决策、规划系统提供足够的环境信息,因此对目标定位以及跟踪的研究具有重要意义。

对目标进行定位和跟踪的前提是检测到目标,由于图像数据拥有丰富的颜色和纹理信息,基于图像的深度学习已成为当下各个应用领域中目标检测任务的主流算法^[1],虽然上述算法可以实现目标检测,但是由于图像数据缺乏深度信息无法对目标进行定位^[2],而且相机容易受到光照变换的影响导致算法跟踪精度不高^[3]。

激光雷达点云有着丰富的三维空间信息,但是缺乏颜色信息并且具有稀疏性,当检测目标距离较远时,基于深度学习的目标检测算法效果将显著下降^[4-7]。综上所述,图像的分类信息优势和点云的距离信息优势的结合可以弥补彼此的缺点,有效地实现检测目标定位以及跟踪,从而提高无人驾驶的安全性。

结合两传感器的优点,众多学者对多传感器融合进行了研究。文献[7]通过图像检测框视锥体^[8]的范围来框选激光雷达数据,但是该文献中明确了每个三维锥形区域中有且仅有一个目标,不能存在任何干扰,且该文献的跟踪任务中只使用了相机数据,跟踪精度不高。文献[9]将激光雷达三维点云数据投影到二维像素平面,并根据相对像素面积大小的概念定义激光雷达检测目标物的置信度,提高了检测精度。文献[10]通

过决策级融合获得目标的类别信息以及距离信息,但是在目标受到遮挡时目标框的编号和距离准确度不高。文献[11]使用聚类后的点云簇与图像检测框的最大交并比获得目标点云数据,有良好的目标定位效果,但是该方法对相机视角内的所有点云数据进行聚类,运算时间较长。

针对以上问题,本文提出一种高效率、高精度的相机与激光雷达融合的目标定位与跟踪方法,创新地提出图像检测框内前景点云面积与背景点云面积比例阈值,可以有效对非检测目标点云进行滤波。针对单一相机对目标跟踪精度不高的问题,使用卡尔曼滤波器融合相机和雷达的观测数值,有效地提高了目标的跟踪精度。

2 算法原理

所提算法流程图如图 1 所示。通过对图像数据使用深度学习进行目标检测,获得检测目标的类别和方位信息。基于图像深度学习的网络模型分为速度较快的一阶段模型^[12-13]和精度较高的两阶段模型^[14-17],为满足实时性,采用一阶段模型 YOLOv5,然后根据图像检测框内切割雷达点云,对切割的雷达点云进行密度聚类(DBSCAN)^[18]得到点云簇,再将激光点云簇在像素平面的前景与背景的面积比值与设计的阈值大小进行比较获得检测目标的定位点云,最后将前后两帧得到的定位点云在像素坐标系的横向移动速度和纵向移动速度观测值与图像检测框的观测值进行融合,得到高精度的跟踪效果。

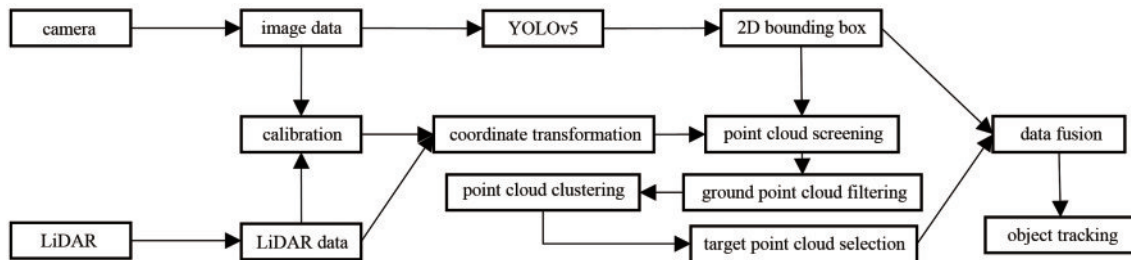


图1 相机与激光雷达融合算法流程

Fig. 1 The process of camera and LiDAR fusion algorithm

2.1 相机与激光雷达的时间与空间同步

不同传感器之间的工作频率不同,通常相机的工作频率为 30 Hz,激光雷达的工作频率为 10 Hz,若不进行时间同步将导致接收到相机数据的时刻与雷达数据的时刻不对应^[19],因此需要将不同时刻获得的同一目标观测数据转换到统一的融合时刻,即在时间上进行校准。时间同步分硬同步方法和软同步方法两种,所采用的 kitti 数据集使用硬触发的方式进行时间同步,激光扫描仪以每秒 10 帧的速度触发摄像机,雷达采集数据时,绕着竖直轴旋转扫描,只有当雷达旋转到与相机的朝向一致时才会触发相机采集图像。

由于相机与激光雷达的安装位置以及坐标系不同,同一个检测目标在两个传感器的坐标系下的位置不同,无法确定检测目标的具体位置,故需要将不同传感器的测量值变换到一个坐标系。统一两个传感器的坐标系,也叫作空间同步,须对其中一个坐标系进行旋转和平移,本实验将雷达坐标系下的数据经过旋转和平移统一到相机坐标系中,空间同步如图 2 所示。

旋转矩阵用 \mathbf{R} 表示,平移向量用 \mathbf{T} 表示,将激光雷达坐标系中任意的一个点 $\mathbf{P}_L(x_L, y_L, z_L)$ 转换到相机的像素坐标系下。首先,将点 \mathbf{P}_L 转换到相机坐标系下,转换后用点 $\mathbf{P}_C(x_C, y_C, z_C)$ 表示,其转换关系如下:

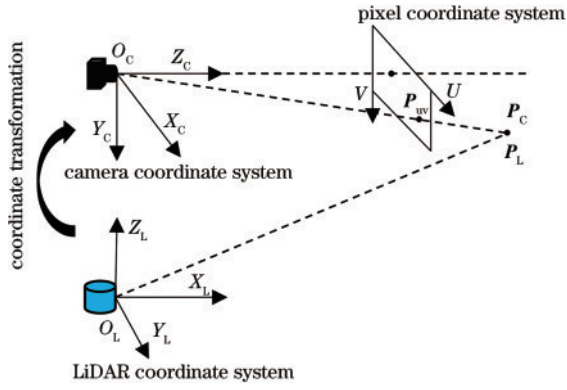


图 2 相机与激光雷达坐标系转换

Fig. 2 Camera and LiDAR coordinate system transformation

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_l \\ y_l \\ z_l \\ 1 \end{bmatrix}, \quad (1)$$

式中： x_l, x_c 的下标 L、C 分别表示激光雷达坐标系下的数据和相机坐标系下的数据。

其次，将相机坐标系的数据转换到像素坐标系下：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{Z_c} \begin{bmatrix} \frac{1}{dx} & 0 & 0 & u_0 \\ 0 & \frac{1}{dy} & 0 & v_0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix}, \quad (2)$$

式中： f 为相机的焦距； dx, dy 为图像坐标系到像素坐标系的量化值； u_0, v_0 为光心 o 在像素坐标系下的坐标； u, v 为激光雷达坐标系中的点 (x_l, y_l, z_l) 在像素坐标系下的坐标。

2.2 检测框内图像数据框选

通过旋转矩阵 \mathbf{R} 与平移向量 \mathbf{T} 将激光雷达数据转移到相机坐标系中，故此时可根据相机的成像模型将图像上任意一点在相机坐标系中相对于 Z_c 轴水平角度和垂直角度计算出来，根据水平角度和垂直角度找到对应的激光雷达点云数据。基于 YOLOv5 的图像目标检测会返回检测框的两个顶点信息，如图 3 中的左上角 (u_1, v_1) 和右下角 (u_2, v_2) 所示，根据相机的视锥体模型，通过像素坐标系中的检测框两个顶点相对于 Z 轴的垂直角度以及水平角度来框选出相机坐标系下的雷达点云。图 3 中，相机光心为 o ，焦距为 f ， (u_0, v_0) 为光心 o 在像素坐标系下的坐标。以像素坐标系下检测框左上角顶点 (u_1, v_1) 、相机坐标系下的激光雷达点云 (x_c, y_c, z_c) 为例，检测框左上角顶点相对于 Z_c 轴的垂直角度 ϕ_v 与水平角度 β_v 如式 (3)、(4) 所示，相机坐标系下的激光雷达点云相对于 Z_c 轴的垂直角度 ϕ_y 与水平角度 β_x 如式 (5)、(6) 所示，筛选出雷达点云中 ϕ_y 与 β_x 大于图像检测框左上角和小于右下角的 ϕ_v 与 β_v 的所有点，即可得到检测框内的雷达点云数据。

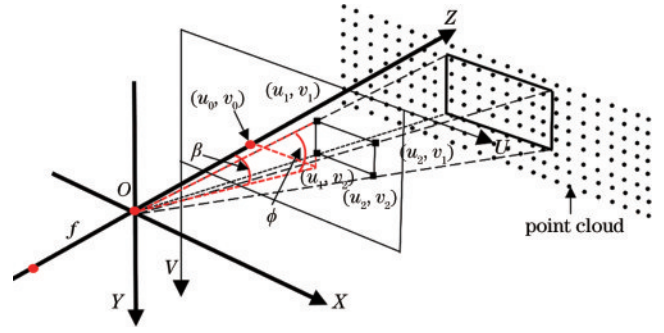


图 3 框选点云

Fig. 3 Point cloud screening

$$\phi_v = \arctan \frac{v_1 - v_0}{\sqrt{f^2 + (u_1 - u_0)^2}}, \quad (3)$$

$$\beta_v = \arctan \frac{u_1 - u_0}{f}, \quad (4)$$

$$\phi_y = \arctan \frac{y_c}{\sqrt{x_c^2 + z_c^2}}, \quad (5)$$

$$\beta_x = \arctan \frac{x_c}{z_c}. \quad (6)$$

2.3 雷达数据处理

2.3.1 地面点云滤波

针对检测目标的定位信息而言，地面返回的雷达点云属于干扰信息需要将其滤除，假设车辆在水平地面运动，根据地面点云的特点采用随机采样一致 (RANSAC)^[20] 算法对地面数据进行拟合，该方法简单，通过迭代的方式从一组包含局外点的样本中高精度地拟合出数学模型参数^[21]。

2.3.2 点云聚类

根据激光雷达无法穿透物体的特性，不同物体雷达点云的坐标值会发生跳变，如检测框内的障碍物点云簇的平均坐标值要小于检测目标点云簇，背景点云簇的平均距离值要大于检测目标点云簇，根据雷达点云距离值跳变的特点构建距离角度直角坐标系，横轴为相机坐标系下的激光雷达点云与相机坐标系 Z_c 轴之间的水平角度 β_x ，纵轴为相机坐标系下的激光点云与相机坐标系 X_c 轴的垂直距离。

距离角度直角坐标系中的激光雷达数据可通过密度聚类算法得到不同距离的点云簇 $P_i (i = 1, 2, \dots, n)$ ，与其他聚类算法相比，该算法不需要指定聚类结果的种类数，通过自适应的方式获取结果。

2.3.3 基于像素坐标系下矩形面积比较的障碍物点云滤波算法

基于深度学习的图像目标检测在检测目标受到一定程度的遮挡时仍然能够识别，但受到矩形检测框的限制，检测框无法完全将检测目标的轮廓勾勒出来，导致引入障碍物点云和背景点云，如图 4 所示。检测框内的点云组成成分 4 种情况，如表 1 所示。



图 4 激光雷达点云投影

Fig. 4 LiDAR point cloud projection

表 1 检测框内点云组成部分

Table 1 Point cloud components within the 2D box

Case 1	Case 2	Case 3	Case 4
Target point cloud	Target point cloud; obstacle point cloud	Target point cloud; background point cloud	Target point cloud; obstacle point cloud; background point cloud

由于聚类后的点云簇中存在非检测目标点云数据,为过滤掉非检测目标点云数据,根据图像检测框内的障碍物点云、检测目标点云以及背景点云的面积占比以及各点云簇之间的平均距离值的大小不同,创新地提出通过前景点云簇与背景点云簇面积比阈值 σ 来过滤掉非检测目标点云的方法。

对聚类后的点云簇 $P_i(i=1, 2, \dots, n)$ 按平均距离数值从小到大排序,取出距离值最小的一簇点云作为前景点云簇,如图 5(a)中紫色矩形框所示,根据式(2)将相机坐标系下的点云投影到像素坐标系下,如图 5(b)中紫色矩形框所示,将前景点云簇与剩下的点云簇组为背景点云簇并投影到像素坐标系下,如图 5(b)中红色矩形框

所示,通过矩形面积比值判断当前簇点云是否为检测目标点云,根据式(7)、(8)计算面积比,若小于面积比阈值 σ ,则将前景点云簇删除,取下一簇点云作为前景点云簇进行投影,如图 5(d)中橙色矩形框所示,将前景点云簇与剩余点云簇投影组为背景点云,如图 5(e)中橙色矩形框所示,将背景点云投影到像素坐标系下,如图 5(f)中红色矩形框所示,计算面积比,若大于面积比阈值 σ ,则前景点云为检测目标点云。经过多次实验,将面积比阈值 σ 设置为 2/3,检测框内点云面积的计算公式、面积比计算公式分别为

$$S(P_i) = |\max(u) - \min(u)| \times |\max(v) - \min(v)|, \quad (7)$$

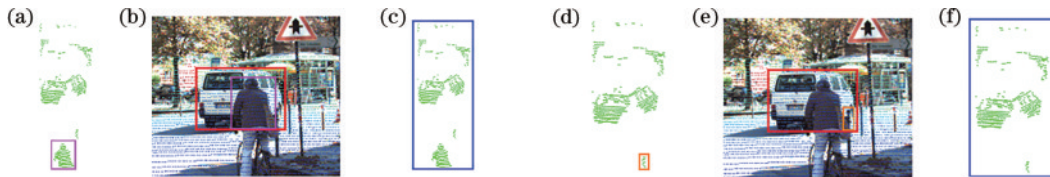


图 5 目标点云选取。(a)当前点云簇选取;(b)当前点云簇投影;(c)剩余点云簇选取;(d)当前点云簇选取;(e)当前点云簇投影;(f)剩余点云簇选取

Fig. 5 Target point cloud selection. (a) Current point cloud cluster selection; (b) current point cloud cluster projection; (c) selection of remaining point cloud clusters; (d) current point cloud cluster selection; (e) current point cloud cluster projection; (f) selection of remaining point cloud clusters

$$Q_i = \frac{S(P_i)}{S(P_i + P_{i+1} + \dots + P_n)}, \quad (8)$$

式中: u, v 分别为当前点云簇 P_i 的每个点云投影到像素坐标系横轴与纵轴的坐标值; Q_i 为前景点云簇 P_i 与背景点云簇 $P_i(i=1, 2, \dots, n)$ 在像素坐标系下的面积比值。算法具体步骤如图 6 所示。

2.3.4 基于相机与激光雷达融合的目标追踪

基于图像数据的目标跟踪存在不够精准的问题,在 DeepSORT 目标跟踪算法中引入图像检测框内目

标点云边缘的横向移动速度和纵向移动速度,其状态量为

$$D_k = [x_k^c, y_k^c, v_{xk}^l, v_{yk}^l, \omega_k^c, h_k^c], \quad (9)$$

式中: c, l 分别表示为相机数据和雷达数据; x_k^c 为像素坐标系下第 k 帧图像的检测框几何中心的横坐标; y_k^c 为像素坐标系下第 k 帧图像的检测框几何中心的纵坐标; v_{xk}^l 为第 k 帧图像检测框内的检测目标定位点云轮廓在像素坐标系的 X 轴方向移动速度; v_{yk}^l 为第 k 帧图像检测框内的检测目标定位点云轮廓在像素坐标系的

Algorithm 1: Obstacle point cloud filtering algorithm

Input :
P: clustered point cloud clusters
 σ : area ratio threshold

Method :
 1) if len(**P**)<2:
 2) **P** is the detection target point cloud
 3) else:
 sort point cloud clusters **P** by distance from smallest to largest
 4) for i in range len(**P**):
 5) calculate the area of point cloud cluster P_i
 6) calculate the area of point cloud cluster $P_{i+...+P_{len(P)}}$
 7) calculate the area ratio Q_i
 8) if $Q_i > \sigma$:
 9) the i -th cluster point cloud is the detection target point cloud
 10) else:
 11) delete the i -th cluster point cloud

Output :
 the detection target point cloud

图 6 基于面积比较的障碍物点云滤波算法流程

Fig. 6 Process of obstacle point cloud filtering algorithm based on area comparison

Y轴方向移动速度; w_k^c 为像素坐标系下第 k 帧图像检测框的宽高比; h_k^c 为像素坐标系下第 k 帧图像检测框的高。

检测目标定位点云轮廓的横向移动速度 v_{xk}^c 由点在像素坐标系下的几何中心相对于检测框的几何中心的偏移来决定:居中时,点云在像素坐标系的横向移动速度为左右两侧速度之和的一半;偏左时,说明右侧受到遮挡,故选取左侧速度作为点云移动速度,右侧同理。

对于垂直方向速度 v_{yk}^c ,由于车辆在道路上行驶时,垂直方向上受到遮挡的情况极少,故选取点云在像素坐标系的纵向移动速度为上下两侧速度之和的一半。

假设车辆为一个匀速运动的模型,由于车辆在运动过程中存在加速、减速以及转向等动作,故在模型中加入过程噪声 N_k :

$$N_k = \left[\frac{1}{2} a_x \Delta t^2, \frac{1}{2} a_y \Delta t^2, a_x \Delta t, a_y \Delta t, \Delta w, \Delta h \right]^T, \quad (10)$$

式中: $\frac{1}{2} a_x \Delta t^2$ 、 $\frac{1}{2} a_y \Delta t^2$ 为检测框横纵向的位移噪声; $a_x \Delta t$ 、 $a_y \Delta t$ 为检测框横纵向速度噪声; Δw 、 Δh 为检测框宽高变化噪声。

最终,系统的状态预测方程为

$$\hat{D}_k = F \hat{D}_{k-1} + N_k, \quad (11)$$

式中: F 为预测矩阵,也叫运动估计矩阵,用于从 \hat{D}_{k-1} 预测 \hat{D}_k 。

协方差矩阵 P_k 的预测方程为

$$P_k = F P_{k-1} F^T + Q_k, \quad (12)$$

式中: Q_k 为处理误差协方差矩阵。

观测变量 Z_k 与状态变量 D_k 的维度相同,故观测矩阵 H 的表达式为

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (13)$$

卡尔曼增益 K_k' 的更新方程为

$$K_k' = P_k H^T (H P_k H^T + R_k)^{-1}, \quad (14)$$

式中:矩阵 R_k 为观测值 Z_k 的协方差矩阵,表示传感器的精确程度,其矩阵维度跟 Z_k 相同。卡尔曼增益,表示量测值与预测值对最后估计结果的重要程度。增益越大,最后估计结果接近量测值的概率越大;反之,越接近预测值。

结合观测值 Z_k 、预测值 \hat{D}_k 、卡尔曼增益 K_k' 对当前状态进行估计:

$$\hat{D}_k' = \hat{D}_k + K_k' (Z_k - H \hat{D}_k). \quad (15)$$

更新协方差矩阵 P_k :

$$P_k = P_k - K_k' H P_k. \quad (16)$$

3 实验结果与分析

实验采用 kitti 数据集,激光雷达为 Velodyne HDL-64E 机械旋转式激光雷达,相机为彩色相机 FL2-14S3C-C,算法运行平台为高性能笔记本 R9000P, Windows 10 操作系统, AMD Ryzen7 5800H 处理器。

3.1 面积比阈值选取

不同的面积比阈值 σ ,对检测框的抗干扰能力及定位准确率有影响,过高的阈值会导致目标定位到背景点云,过低的阈值会造成抗干扰能力的下降,为获得最佳面积比阈值,选取 1/2、7/12、2/3、3/4 等 4 个阈值进行定位实验,计算各阈值的平均准确率,对比结果如表 2 所示。

表 2 阈值对比

Table 2 Threshold comparison

σ	1/2	7/12	2/3	3/4
Accuracy / %	81.1250	85.4167	88.5417	74.4791

可以看出,阈值 $\sigma = 2/3$ 时的定位准确率最高,故将其作为本研究的实验阈值。

3.2 目标定位算法的可行性验证

对算法的可行性进行测试,原始点云数据用红色表示,在检测框内的检测目标激光雷达点云赋予绿色,用 3D 包围盒框出。图 7(a)的检测框中含有两个障碍物信息,图 7(b)为激光点云在距离角度笛卡儿坐标系中的分布图,图 7(c)为聚类后的图,其中,圆点为未聚类成功的点,图 7(d)为最终定位结果。

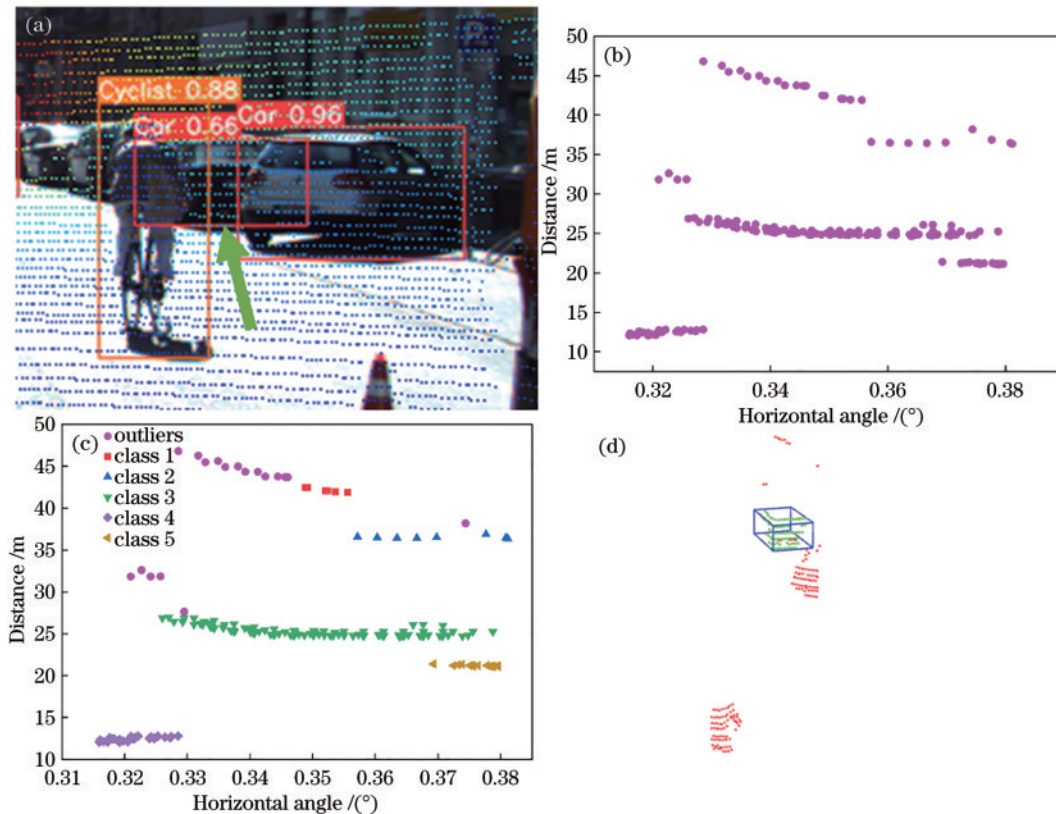


图 7 kitti 数据集测试结果。(a) 图像识别; (b) 聚类前; (c) 聚类后; (d) 3D 包围框

Fig. 7 The result of kitti dataset. (a) Image detection; (b) before clustering; (c) after clustering; (d) 3D bounding box

由图 7 的实验结果可看出,所提算法在检测目标受到严重遮挡时仍然能够对目标进行定位,且不受障碍物个数的限制。

3.3 目标定位算法准确率与实时性验证

为验证算法的可靠性,随机选取 kitti 数据集中的 100 帧点云和对应的图像,在 100 张图像中手动标注 192 个识别框,通过统计激光雷达点云筛选出的点云

数据与图像检测框中的实际物体进行对比,判断算法的正确率,并与文献[7]、文献[11]算法进行目标定位对比。识别框内的检测目标根据被遮挡程度以及距离分为简单、中等和困难等 3 类,简单为没有任何遮挡,中等为小部分遮挡,困难为遮挡部分较大且检测目标距离较远。对比实验结果如图 8 以及表 2 所示。

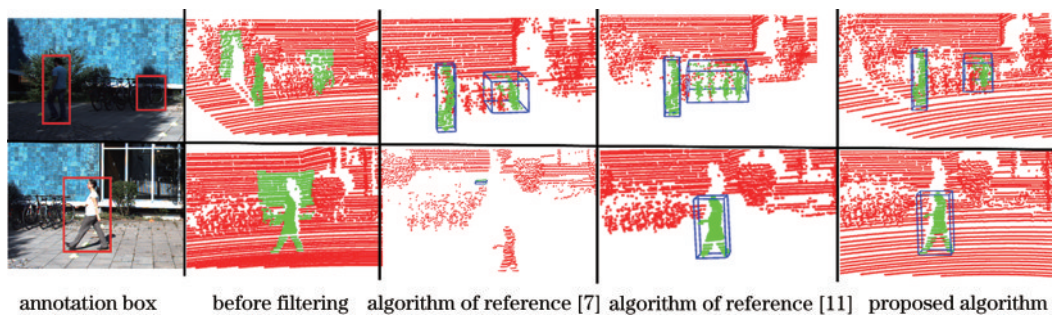


图 8 目标定位对比

Fig. 8 Target positioning comparison

文献[7]算法容易受到背景点云中的小物体的点云影响,导致定位出错。文献[11]中的算法在检测目标离其他物体较近时容易造成定位错误,而所提算法在以上情况下均能够进行准确的定位。100 帧数据的对比结果如表 3 所示。

相对于文献[7]和文献[11]算法,所提算法在简单、中等以及困难检测目标上的定位准确率都有提高,

总体定位准确率相对文献[7]提升 10.4167 百分点,相对于文献[11]算法提升 8.3334 百分点。由于所提算法对图像检测框内的点云进行聚类处理,大大降低了点云数量,降低了每帧点云的处理时间。且因为激光以射线的形式射出,点云具有稀疏性,点云数据不可能对应到每一个像素块,对于距离较远或较小的检测目标容易定位到背景点云。

表 3 算法对比

Table 3 Algorithm comparison

Algorithm	Easy / %	Moderate / %	Hard / %	Accuracy / %	Time / s
Algorithm of reference [7]	82.6667	86.9565	58.3333	78.125	0.1218
Algorithm of reference [11]	84	95.6521	52.0833	80.2083	0.1299
Proposed algorithm	92	97.1014	70.8333	88.5417	0.0314

3.4 相机与激光雷达融合的目标跟踪算法结果

为验证算法的跟踪效果,与DeepSORT算法以及粒子滤波(PF)算法进行对比实验,图9为所提算法的跟踪效果图,淡蓝色框为目标真实框,红色框为所提算法的跟踪位置,深蓝色框为文献[3]算法的跟踪位置,

酒红色框为PF算法的跟踪位置。为定量分析其追踪效果,采用X轴平均误差 e_x 、Y轴平均误差 e_y 、平均区域重叠率 o_r 来评价跟踪算法。方向跟踪效果如图10所示,误差以及重叠率见表4。



图9 受遮挡目标追踪效果图

Fig. 9 Occluded target tracking rendering

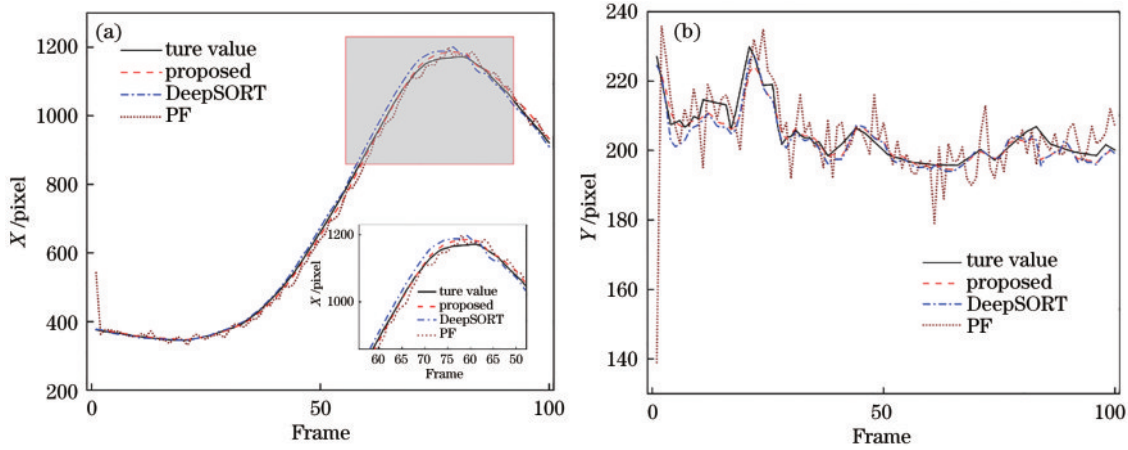


图10 目标追踪结果。(a) X方向追踪轨迹;(b) Y方向追踪轨迹

Fig. 10 Target tracking results. (a) X-trace; (b) Y-trace

表4 平均误差对比

Table 4 Comparison of average errors

Algorithm	e_x / pixel	e_y / pixel	o_r / %	Time / s
PF	14.7	5.87	23.4	0.1175
DeepSORT	9.45	2.17	87.06	0.0817
Proposed algorithm	4.49	1.80	87.42	0.0841

所提算法通过相邻两帧之间的时间间隔以及检测目标外围轮廓点云的位移获得检测目标的移动速度,将移动速度与图像检测框的中心坐标融合,跟踪效果良好。实验数据中,检测目标在43帧到56帧以

及66帧到95帧发生遮挡,通过图10(a)、(b)两图可以看出融合算法能够很好地拟合真实值,在X轴上的跟踪平均误差为4.49 pixel,在Y轴上的跟踪平均误差为1.80 pixel,平均区域重叠率为87.42%,平均每帧耗时为0.0841 s。

4 结 论

提出一种基于相机与激光融合的目标跟定位与跟踪方法,对检测框内的地面点云进行滤波,大大降低需要滤波的点云数量,提高处理速度,再利用图像检测框内检测前景与背景的面积比关系对非检测目标的点云

信息进行滤除,使得检测目标受到遮挡时有一定程度的抗干扰能力,提高了定位准确率。平均每帧运行时间为 0.03 s,满足实时性要求。再通过检测框内的定位点云在像素坐标系下的横向移动速度、纵向移动速度与图像检测框的高、宽高比、中心横、纵坐标数据的融合对目标进行跟踪,有效提高了跟踪精度,由检测到定位再到追踪的平均每帧耗时为 0.0841 s,可实现实时追踪。

由于所提算法根据图像检测框内前景与背景的面积比设计面积比阈值,当检测框内障碍物的面积超过面积比阈值时会造成点云数据误判,因此,后续的研究内容为检测框内识别目标的雷达点云自适应提取。

参 考 文 献

- [1] 罗逸豪, 孙创, 邵成, 等. 基于深度学习的水面无人艇目标检测算法综述[J]. 数字海洋与水下攻防, 2022, 5(6): 524-538.
Luo Y H, Sun C, Shao C, et al. Overview of target detection algorithms for surface unmanned boats based on deep learning[J]. Digital Ocean & Underwater Warfare, 2022, 5(6): 524-538.
- [2] 胡杰, 安永鹏, 徐文才, 等. 基于激光点云的深度语义和位置信息融合的三维目标检测[J]. 中国激光, 2023, 50(10): 1010003.
Hu J, An Y P, Xu W C, et al. 3D object detection based on deep semantics and position information fusion of laser point cloud[J]. Chinese Journal of Lasers, 2023, 50(10): 1010003.
- [3] Wojke N, Bewley A, Paulus D. Simple online and realtime tracking with a deep association metric[C]//2017 IEEE International Conference on Image Processing (ICIP), September 17-20, 2017, Beijing, China. New York: IEEE Press, 2018: 3645-3649.
- [4] Zhou Y, Tuzel O. VoxelNet: end-to-end learning for point cloud based 3D object detection[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 4490-4499.
- [5] Shi W J, Rajkumar R. Point-GNN: graph neural network for 3D object detection in a point cloud[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 1708-1716.
- [6] Lang A H, Vora S, Caesar H, et al. PointPillars: fast encoders for object detection from point clouds[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2020: 12689-12697.
- [7] 黄远宪, 李必军, 黄琦, 等. 融合相机与激光雷达的目标检测、跟踪与预测[J/OL]. 武汉大学学报(信息科学版): 1-8[2023-08-09]. <https://doi.org/10.13203/j.whugis.20210614>.
Huang Y X, Li B J, Huang Q, et al. Camera-LiDAR fusion for object detection, tracking and prediction[J/OL]. Geomatics and Information Science of Wuhan University: 1-8[2023-08-09]. <https://doi.org/10.13203/j.whugis.20210614>.
- [8] 宋一凡, 张鹏, 宗立波, 等. 改进的基于冗余点过滤的 3D 目标检测方法[J]. 计算机应用, 2020, 40(9): 2555-2560.
Song Y F, Zhang P, Zong L B, et al. Improved redundant point filtering-based 3D object detection method[J]. Journal of Computer Applications, 2020, 40(9): 2555-2560.
- [9] 张堡瑞, 肖宇峰, 郑又能. 基于激光雷达与视觉融合的水面漂浮物检测[J]. 应用激光, 2021, 41(3): 619-628.
Zhang B R, Xiao Y F, Zheng Y N. Detection of floating objects on water surface based on fusion of lidar and vision[J]. Applied Laser, 2021, 41(3): 619-628.
- [10] 李研芳, 黄影平. 基于激光雷达和相机融合的目标检测[J]. 电子测量技术, 2021, 44(5): 112-117.
Li Y F, Huang Y P. Target detection based on the fusion of lidar and camera[J]. Electronic Measurement Technology, 2021, 44(5): 112-117.
- [11] 崔善尧, 贾述斌, 黄劲松. 图像与激光点云融合的实时目标定位算法[J]. 导航定位学报, 2023, 11(2): 99-105.
Cui S Y, Jia S B, Huang J S. Real-time target positioning algorithm based on image and laser point cloud fusion[J]. Journal of Navigation and Positioning, 2023, 11(2): 99-105.
- [12] 周晓彦, 王珂, 李凌燕. 基于深度学习的目标检测算法综述[J]. 电子测量技术, 2017, 40(11): 89-93.
Zhou X Y, Wang K, Li L Y. Review of object detection based on deep learning[J]. Electronic Measurement Technology, 2017, 40(11): 89-93.
- [13] 许德刚, 王露, 李凡. 深度学习的典型目标检测算法研究综述[J]. 计算机工程与应用, 2021, 57(8): 10-25.
Xu D G, Wang L, Li F. Review of typical object detection algorithms for deep learning[J]. Computer Engineering and Applications, 2021, 57(8): 10-25.
- [14] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 580-587.
- [15] 曹诗雨, 刘跃虎, 李辛昭. 基于 Fast R-CNN 的车辆目标检测[J]. 中国图象图形学报, 2017, 22(5): 671-677.
Cao S Y, Liu Y H, Li X Z. Vehicle target detection based on Fast R-CNN[J]. Journal of Image and Graphics, 2017, 22(5): 671-677.
- [16] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [17] 路齐硕. 基于深度学习的目标检测方法研究[D]. 北京: 北京邮电大学, 2020: 4-6.
Lu Q S. Research on target detection method based on deep learning[D]. Beijing: Beijing University of Posts and Telecommunications, 2020: 4-6.
- [18] 周治国, 李怡瑶, 曹江微, 等. 基于 3D 激光雷达的水面目标检测算法研究[J]. 激光与光电子学进展, 2022, 59

- (18): 1815006.
Zhou Z G, Li Y Y, Cao J W, et al. Research on water target detection algorithm based on 3D lidar[J]. Laser & Optoelectronics Progress, 2022, 59(18): 1815006.
- [19] 张煌, 何佳洲, 王景石, 等. 基于全景图像的无人艇激光雷达杂波滤除方法[J]. 光学学报, 2022, 42(18): 1810001.
Zhang H, He J Z, Wang J S, et al. Clutter filtering method of unmanned boat lidar based on panoramic image [J]. Acta Optica Sinica, 2022, 42(18): 1810001.
- [20] Fischler M A, Bolles R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography[M]//Fischler M A, Firschein O. Readings in computer vision. Amsterdam: Elsevier, 1987: 726-740.
- [21] 左勇, 任阳, 杜志华, 等. 基于 LP-RANSAC 算法的路面点云快速移除算法[J]. 激光与光电子学进展, 2023, 60(14): 1428009.
Zuo Y, Ren Y, Du Z H, et al. Rapid removal algorithm of road surface point cloud based on LP-RANSAC algorithm[J]. Laser & Optoelectronics Progress, 2023, 60(14): 1428009.