



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2021
01
VOL.26

ISSN1006-8961
CN11-3758/TB





第26卷第1期 (总第297期)

2021年1月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿, 均视为同意在本刊网站及CNKI等全文数据库出版, 所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用, 所有内容, 未经许可, 不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院

主办单位 中国科学院空天信息创新研究院

中国图象图形学学会

北京应用物理与计算数学研究所

主 编 吴一戎

编辑出版 《中国图象图形学报》编辑出版委员会

通信地址 北京市海淀区北四环西路19号

邮 编 100190

电子信箱 jig@radi.ac.cn

电 话 010-58887035

网 址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218号

总 发 行 北京报刊发行局

订 购 全国各地邮局

海外发行 中国国际图书贸易集团有限公司

(邮政信箱: 北京399信箱 邮编: 100048)

印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian

Monthly, Started in 1996

Superintended by Chinese Academy of Sciences

Sponsored by Aerospace Information Research Institute, CAS

China Society of Image and Graphics

Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Wu Yirong

Editor, Publisher Editorial and Publishing Board of Journal of
Image and Graphics

Address No. 19, North 4th Ring Road West, Haidian District,
Beijing, P. R. China

Zip code 100190

E-mail jig@radi.ac.cn

Telephone 010-58887035

Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals

Domestic All Local Post Offices in China

Overseas China International Book Trading Corporation

(P.O.Box 399, Beijing 100048, P.R.China)

Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB

ISSN 1006-8961

CODEN ZTTXFZ

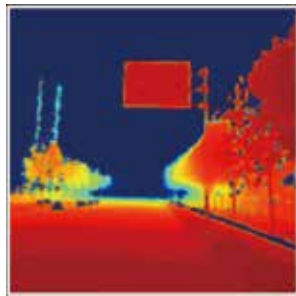
国外发行代号 M1406

国内邮发代号 82-831

国内定价 60.00元



面向智能驾驶的平行视觉感知：基本概念、框架与应用
(第0067页)



Edge-guided GAN: 边界信息引导的深度图像修复
(第0186页)



利用边缘计算的多车协同激光雷达SLAM(第0218页)

序言	王飞跃, 张云泉 I
编者按	II

综述

面向智能驾驶测试的仿真场景构建技术综述	
任秉韬, 邓伟文, 白雪松, 李江坤, 纵瑞雪, 朱冰, 丁娟	0001
自动驾驶软件测试技术研究综述	
冯洋, 夏志龙, 郭安, 陈振宇	0013
强化学习的自动驾驶控制技术研究进展	
潘峰, 鲍泓	0028
自动驾驶地图的数据标准比较研究	
詹骄, 郭迟, 雷婷婷, 屈宜琪, 吴杭彬, 刘经南	0036
视觉感知的端到端自动驾驶运动规划综述	
刘旖菲, 胡学敏, 陈国文, 刘士豪, 陈龙	0049

平行驾驶

面向智能驾驶的平行视觉感知：基本概念、框架与应用	
李轩, 王飞跃	0067
平行视觉的基本框架与关键算法	
张慧, 李轩, 王飞跃	0082

目标检测与跟踪

自动驾驶场景的尺度感知实时行人检测	
徐歆恺, 马岩, 钱旭, 张龔	0093
实时视觉目标跟踪与视频对象分割多任务框架	
李瀚, 刘坤华, 刘嘉杰, 张晓晔	0101
提升预测框定位稳定性的视频目标检测	
郝腾龙, 李熙莹	0113
道路结构特征下的车道线智能检测	
张翔, 唐小林, 黄岩军	0123
适用全速域大曲率路径的自动驾驶跟踪算法	
张龔, 郑颖, 鲍泓	0135
伪3D卷积神经网络与注意力机制结合的疲劳驾驶检测	
庄员, 戚湧	0143
基于眼部自商图—梯度图共生矩阵的疲劳驾驶检测	
潘剑凯, 柳政卿, 王秋成	0154

自动驾驶场景感知与仿真

结合局部平面参数预测的无监督单目图像深度估计	
周大可, 田径, 杨欣	0165
深度纯追随的拟人化无人驾驶转向控制模型	
单云霄, 黄润辉, 何泽, 龚志豪, 景民, 邹雪松	0176
Edge-guided GAN: 边界信息引导的深度图像修复	
刘坤华, 王雪辉, 谢玉婷, 胡坚耀	0186
引入概率分布的深度神经网络贪婪剪枝	
胡骏, 黄启鹏, 刘嘉昕, 刘威, 袁淮, 赵宏	0198

高精地图构建与SLAM

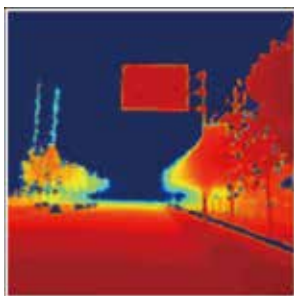
开放道路中匹配高精度地图的在线相机外参标定	
廖文龙, 赵华卿, 严骏驰	0208
利用边缘计算的多车协同激光雷达SLAM	
崔明月, 钟仕鹏, 刘思瑶, 李博洋, 吴成昊, 黄凯	0218

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



Parallel visual perception for intelligent driving: basic concept, framework and application (P0067)



Edge-guided GAN: a depth image inpainting approach guided by edge information (P0186)



Cooperative LiDAR SLAM for multi-vehicles based on edge computing (P0218)

Review

- Technologies of virtual scenario construction for intelligent driving testing
Ren Bingtao, Deng Weiwen, Bai Xuesong, Li Jiangkun, Zong Ruixue, Zhu Bing, Ding Juan 0001
- Survey of testing techniques of autonomous driving software
Feng Yang, Xia Zhilong, Guo An, Chen Zhenyu 0013
- Research progress of automatic driving control technology based on reinforcement learning
Pan Feng, Bao Hong 0028
- Comparative study on data standards of autonomous driving map
Zhan Jiao, Guo Chi, Lei Tingting, Qu Yiqi, Wu Hangbin, Liu Jingnan 0036
- Review of end-to-end motion planning for autonomous driving with visual perception
Liu Yifei, Hu Xuemin, Chen Guowen, Liu Shihao, Chen Long 0049

Parallel Driving

- Parallel visual perception for intelligent driving: basic concept, framework and application
Li Xuan, Wang Feiyue 0067
- The basic framework and key algorithms of parallel vision
Zhang Hui, Li Xuan, Wang Feiyue 0082

Object Detection and Tracking

- Scale-aware EfficientDet: real-time pedestrian detection algorithm for automated driving
Xu Xinkai, Ma Yan, Qian Xu, Zhang Yan 0093
- Multitask framework for video object tracking and segmentation combined with multi-scale inter-frame information
Li Han, Liu Kunhua, Liu Jiajie, Zhang Xiaoye 0101
- Video object detection method for improving the stability of bounding box
Hao Tenglong, Li Xiying 0113
- Intelligent detection of lane based on road structure characteristics
Zhang Xiang, Tang Xiaolin, Huang Yanjun 0123
- Autonomous vehicle tracking algorithm for high curvature path in full speed range
Zhang Yan, Zheng Ying, Bao Hong 0135
- Driving fatigue detection based on pseudo 3D convolutional neural network and attention mechanisms
Zhuang Yuan, Qi Yong 0143
- Fatigue driving detection based on ocular self-quotient image and gradient image co-occurrence matrix
Pan Jiankai, Liu Zhengqing, Wang Qiucheng 0154

Scene Perception and Simulation of Autonomous Driving

- Unsupervised monocular image depth estimation based on the prediction of local plane parameters
Zhou Dake, Tian Jing, Yang Xin 0165
- Human-like steering model for autonomous driving based on deep pure pursuit method
Shan Yunxiao, Huang Runhui, He Ze, Gong Zhihao, Jing Min, Zou Xuesong 0176
- Edge-guided GAN: a depth image inpainting approach guided by edge information
Liu Kunhua, Wang Xuehui, Xie Yuting, Hu Jianyao 0186
- Greedy pruning of deep neural networks fused with probability distribution
Hu Jun, Huang Qipeng, Liu Jiaxin, Liu Wei, Yuan Huai, Zhao Hong 0198

High-Definition Map Construction and SLAM

- Online extrinsic camera calibration based on high-definition map matching on public roadway
Liao Wenlong, Zhao Huaqing, Yan Junchi 0208
- Cooperative LiDAR SLAM for multi-vehicles based on edge computing
Cui Mingyue, Zhong Shipeng, Liu Siyao, Li Boyang, Wu Chenghao, Huang Kai 0218

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2021)01-0208-10

论文引用格式: Liao W L, Zhao H Q and Yan J C. 2021. Online extrinsic camera calibration based on high-definition map matching on public roadway. Journal of Image and Graphics, 26(01): 0208-0217 (廖文龙, 赵华卿, 严骏驰. 2021. 开放道路中匹配高精度地图的在线相机外参标定. 中国图象图形学报, 26(01): 0208-0217) [DOI:10.11834/jig.200432]

开放道路中匹配高精度地图的在线相机外参标定

廖文龙^{1,2}, 赵华卿², 严骏驰^{1*}

1. 上海交通大学, 上海 200240; 2. 安徽酷哇机器人有限公司, 芜湖 241010

摘要: 目的 相机外参标定是 ADAS(advanced driver-assistance systems)等应用领域的关键环节。传统的相机外参标定方法通常依赖特定场景和特定标志物,无法实时实地进行动态标定。部分结合 SLAM(simultaneous localization and mapping)或 VIO(visual inertia odometry)的外参标定方法依赖于点特征匹配,且精度往往不高。针对 ADAS 应用,本文提出了一种相机地图匹配的外参自校正方法。**方法** 首先通过深度学习对图像中的车道线进行检测提取,数据筛选及后处理完成后,作为优化问题的输入;其次通过最近邻域解决车道线点关联,并在像平面内定义重投影误差;最后,通过梯度下降方法迭代求解最优的相机外参矩阵,使得像平面内检测车道线与地图车道线真值重投影匹配误差最小。**结果** 在开放道路上的测试车辆显示,本文方法经过多次迭代后收敛至正确的外参,其旋转角精度小于 0.2° ,平移精度小于 0.2 m ,对比基于消失点或 VIO 的标定方法(精度为 2.2° 及 0.3 m),本文方法精度具备明显优势。同时,在相机外参动态改变时,所提出方法可迅速收敛至相机新外参。**结论** 本文方法不依赖于特定场景,支持实时迭代优化进行外参优化,有效提高了相机外参精确度,精度满足 ADAS 需求。

关键词: 外参标定;地图匹配;车道线;梯度下降;在线标定

Online extrinsic camera calibration based on high-definition map matching on public roadway

Liao Wenlong^{1,2}, Zhao Huaqing², Yan Junchi¹

1. Shanghai Jiao Tong University, Shanghai 200240, China; 2. Anhui COWAROBOT Co., Ltd., Wuhu 241010, China

Abstract: **Objective** Camera calibration is one of the key factors of the perception in advanced driver-assistance systems (ADAS) and many other applications. Traditional camera calibration methods and even some state-of-the-art calibration algorithms, which are currently widely used in factories, strongly rely on specific scenes and specific markers. Existing methods to calibrate the extrinsic parameters of the camera are inconvenient and inaccurate, and current algorithms have some obvious disadvantages, which might cause serious accidents, damage the vehicle, or threaten the safety of passengers. Theoretically, once calibrated, the extrinsic parameters of the camera, including the position and the posture of camera installation, will be fixed and stable. However, the extrinsic parameters of a camera change throughout the lifetime of a vehicle. Real-time dynamic calibration is useful in cases when vehicles are transported or when cameras are removed for maintenance or replacement. Other extrinsic parameter calibration methods solves the estimation by simultaneous localization and mapping or visual inertia odometry (VIO) technologies. These methods try to extract point features and match points with the same characters, and the spatial transformation of different frames is calculated accordingly from the matched point pairs.

收稿日期:2020-07-31;修回日期:2020-10-20;预印本日期:2020-10-27

* 通信作者:严骏驰 yanjunchi@sjtu.edu.cn

基金项目:科技创新 2030-重大项目(2020AAA0107600);国家自然科学基金项目(61972250)

Supported by: Science and Technology Innovation 2030(2020AAA0107600); National Natural Science Foundation of China(61972250)

However, according to the absence of texture information such as when one is in an indoor environment, the accuracy of extrinsic parameters is not always satisfactory. The common situation is that the algorithm cannot obtain any feature from the existing frames or the features that are obtained are not enough to calculate the position. To solve this problem and achieve the requirement of ADAS, this paper proposes a self-calibrating method that is based on aligning the detected lanes by the camera with a high-definition (HD) map. **Method** Feature extraction is the first step of calibration. The most common feature extraction method is to acquire features from frames, calculate the gradient or other specific information of every single pixel, and select the pixels with the most significant values as the detected features. In this paper, we introduce a state-of-the-art algorithm that uses deep learning to detect lane points in the images grabbed from the camera. Some parts of the extrinsic parameters, including longitudinal translation, are unobservable when the vehicle is moving; thus, a data filtering and post-processing method is proposed. Images are classified into three classes: invalid frame, data frame, and key frame. The data filtering rule will efficiently divide the obtained frames into these three types according to the information the frame carries. Then, in the next step, the reprojection error (or loss) is defined in the imaging plane. The process consists of four steps: 1) The lane detected in the HD map is projected to the image plane, and the nearest neighborhood is associated with every detected lane point. This step is similar to feature matching, but it focuses only on the distance of the nearest potential match points. 2) The distance of points and the normal vectorial angle is calculated, and different weights are assigned based on different image types. 3) The geometric constraints of lanes in the image plane and frame of the camera are solved. The initial guess of the extrinsic parameter is determined; the guess is often imprecise and valid only in the cases when the lane is a straight line and the camera translation is known. 4) A gradient descent-based iterative optimal method is used to minimize the reprojection error, and the optimal extrinsic parameter could be determined at the same time. We use such methods to calibrate the camera extrinsic for the vehicles because of several reasons and advantages. The extrinsic parameters are calibrated by using gradient descent because extrinsic parameters are hypothesized to change slowly enough during the lifetime of a vehicle. Therefore, optimizing the extrinsic parameters by using gradient descent could maintain the accuracy of the current extrinsic parameters. Even when outliers occur, the system could remain stable for a period of time rather than have rapidly changing extrinsic parameters, which is considered dangerous when the vehicle is in motion. Deep learning is used to calibrate the extrinsic parameters because the lanes look different according to different road conditions. With any current method, losing some features of the lane points is common. However, the deep learning method does not have such problems; with enough training data, lanes in any extreme case could be used, even in totally different environments or in most cases of extreme weather. **Result** Experiments on an open road show that the designed loss function is meaningful and convex. With 250 iterations, the proposed method can converge to the true extrinsic parameter, and its rotation accuracy is 0.2° and the translation accuracy is 0.03 m. Compared with the VIO-based and another lane detection-based method, our approach is more accurate with the HD map information. Another experiment shows that the proposed method can quickly converge to the new true value when extrinsic parameters change dynamically. **Conclusion** With the use of lane detection, the proposed method does not depend on specific scenarios or marks. Through the matching of the detected lane and the HD map with numerical optimization, the calibration can be performed in real time, and it improves the accuracy of extrinsic parameters more significantly than other methods. The accuracy of the proposed method meets the requirements of ADAS, showing great value in the field of industry.

Key words: extrinsic parameter calibration; map alignment; lane; gradient descent; online calibration

0 引言

在 ADAS (advanced driving-assistance systems) 或自动驾驶领域,通常使用相机感知车辆周围的障碍物或道路标示 (Dabral 等, 2014; 周欣 等, 2003)。应用于相机图像上的检测及跟踪通常参考相机坐标

系,需要将其转换至车辆坐标系或世界坐标系,从而实现车辆的控制。相机内参以及安装参数(相机以及车辆坐标系之间的相对关系,亦或相机外参)显著影响了相机感知周围环境的精度,如何快速有效地对相机内外参进行标定是 ADAS 车辆量产以及维护的重要环节。业内已经对相机外参标定进行了广泛深入研究,但在 ADAS 应用领域仍然存在如下

挑战:

1) 在车辆维修、后装市场相机安装时,如何在开放道路场景下快速进行外参标定;

2) 在车辆长达 10 年以上的生命周期内,相机外参随着时间推移而逐渐变化,如何及时在线调整相机外参。

传统方法利用棋盘格等标定物对相机内参进行标定 (Tsai, 1987; Zhang, 2000), 该方法被广泛使用且足够精确。根据环境中的特征静止这一原则,一些新方法 (Civera 等, 2009) 先建立相同场景中不同视角的多帧图像的像平面特征点约束关系,并根据该约束关系进行相机内参标定。这类方法无需特定标定物,并可实时运用。由于相机在汽车制造以及维护环节通常作为一个独立的整体,其内参在产品生命周期内不会变化,本文采用传统方法进行内参标定。

为精确确定相机与车辆之间的坐标系关系,也即相机外参标定,一般做法是建立高精度标定场进行辅助标定。高精度标定场配备了位姿追踪设备以及特定标定物,并利用机器人手眼标定法 (Daniilidis, 1999) 确定相机外参。手眼标定法在外参的求解过程中需要标志物与相机之间的空间位姿关系,根据标定物的不同维度可分为 3D、2D 及 1D 等标定方法 (Zhang, 2004)。这类方法通常均依赖于满足特定的点、线或面等约束的地面标志,只适用于离线标定。但在实际车辆的整个生命周期内,由于维修以及结构形变等原因,相机外参将相对于出厂状态可能出现显著性变化。如何在脱离精密而昂贵的标定场的情况下,在线实时进行相机外参标定是该领域的一个难题。

Esquivel 等人 (2007), Carrera 等人 (2011), Yang 和 Shen (2017) 提出了在自然环境下结合视觉里程计 (visual odometry) 或 vSLAM (visual simultaneous localization and mapping) 实时进行相机外参标定,这类方法无需特定参考物,可在线实时对相机外参进行标定。但在车载相机应用中,由于道路空旷,图像特征点较少,无法精确建立特征点间的帧间约束。另外不同于广泛研究的 visual-inertial system (VINS) 或 visual-inertial odometry (VIO) 车载相机仅能在平面内运动,部分状态量无法通过有限时间的测量确定,其标定结果通常难以满足 ADAS 精度需求。一般来说,车载相机需检测 200 m 范围内的行

人、车辆等障碍物,其检测精度约 1 m,假定相机视场 (field of view, FOV) 为 57° ,根据几何关系可知相机外参精度要求为旋转角约 0.2° ,平移约 0.2 m。

Schoepflin 和 Dailey (2003) 以及李婵等人 (2019) 提出了一种基于车道线标定固定式交通监控相机的外参方法,其利用车道或行车轨迹的消失点约束对外参进行自动标定。在车道线为 3 维直线的基本假设下,这类方法解决了场景依赖,并可实时校正外参,具备较高的实用性。然而在城市开放场景下,道路通常为曲线 (包括起伏以及转向),消失点模型并不成立。此外,这类方法通常需手动确定相机高度,自动化程度较低。

综上所述,在特定场合下外参标定问题已有完整解决方案,但针对前述的 ADAS 应用领域的挑战,已有方法较难直接适用。

随着 ADAS 的基础设施逐渐成熟,道路高精度地图作为 ADAS 必备的要素之一 (Ghallabi 等, 2018) 也逐步完善,在实际应用中其可通过测绘车或行驶轨迹测量获得 (Knoop 等, 2017), 亦可通过导航电子地图供应商处购买获得。另外,ADAS 车辆通常配备组合导航高精度定位系统,低成本解决方案下 RMS (root mean square) 可达 20 mm (Li 等, 2017)。

车辆行驶于高精度地图覆盖范围时,车道线信息是普遍存在且特征明显的,利用车道线进行外参标定是自然且直接的方法。结合车辆位姿以及高精度地图的车道线信息,本文提出了一种相机地图匹配的在线外参标定方法,该方法首先假设某一外参 T ,并根据该外参将世界坐标系 W 下的车道点 P_w 投影至相机坐标系 C 得到 3 维点 P_c ,合理设计误差函数 L 评价 P_c 以及相机检测的车道点 D_c 的投影误差 $L(T_c^e)$,采用最小化车道线曲线至像平面重投影误差 (bundle adjustment) 的思想 (Triggs 等, 1999) 求解最优相机外参 $\hat{T}_v^e$ 。

本文方法采用深度学习对图像中的车道线进行检测并与地图进行匹配,无需设置特定的标志物,适用于实际车辆行驶场景。相对于现有方法,本文方法具备如下特点:

1) 将外参标定问题转换为相机—地图车道线重投影误差最小化问题,外参平移及旋转部分均能在线迭代优化;

2) 根据高精度地图与相机匹配,不假设车道线

为平面或直线等形式,提高了外参标定精度;

3)将高精度地图车道线投影至像平面并采用最近邻域关联相机观测,有效解决了车道线无法进行特征匹配以及投影误差高、计算复杂度等问题;

4)采用像平面距离以及法向量夹角重新定义投影误差,该投影误差与外参单调下降,避免了局部极小;

5)设计简单可行的数据筛选规则,避免在道路环境下,外参纵向平移部分不可观而导致的外参发散。

实验表明,在实际 ADAS 场景中,本文方法可有效迭代求解相机外参,其旋转角精度小于 0.2° ,平移小于 0.1 m ,对比其他方法精度显著提升。同时,在外参改变时,本文方法能够快速收敛至变化后外参,满足 ADAS 应用需求。

1 问题分析

如图 1 所示建立坐标系,相机外参标定即确定相机坐标系 C 与车辆坐标系 V 之间的坐标系变换 T_v^c 。 T_v^c 由旋转角 $(\theta_{\text{roll}}, \theta_{\text{pitch}}, \theta_{\text{yaw}})$ 以及平移 (T_x, T_y, T_z) 构成, θ_{roll} 为滚转角, θ_{pitch} 为俯仰角, θ_{yaw} 为偏航角。

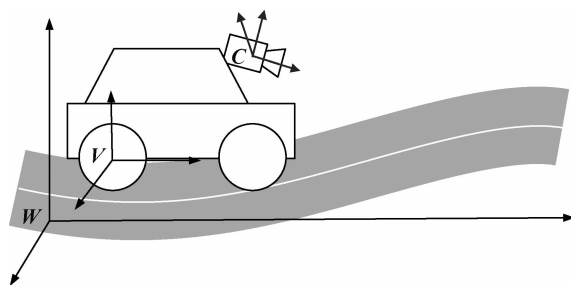


图 1 坐标系关系(相机坐标系 C , 车辆坐标系 V 以及地图坐标系 W)

Fig. 1 Coordinate system relationship (camera coordinate system C , vehicle coordinate system V and map coordinate system W)

考虑到车辆控制时通常以后轮中心为控制原点,在 ADAS 领域通常将车辆坐标系建立于车辆后轮中心在理想地面的投影为原点,以车辆正前方为坐标系 X 轴。由于该坐标系为虚拟坐标系,机械结构上较难确定后轮中心以及车辆正前方,直接关联车辆运动与相机检测从而确定相机外参能够有效减小误差。

记图像畸变校正后的像平面内的车道线点为 $(u, v)^T$,根据小孔成像模型有如下关系

$$z_c \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \mathbf{M} \mathbf{P}_c = \mathbf{M} \mathbf{T}_v^c \mathbf{T}_w^v \mathbf{P}_w \quad (1)$$

式中, z_c 为车道线点 $\mathbf{P}_c$ 到相机的深度, $\mathbf{M}$ 为相机内参矩阵, $\mathbf{T}_w^v$ 为世界坐标系 W 与车辆坐标系 V 之间的坐标变换,即车辆当前的位姿。

定义车道线观测点以及地图参考点的重投影误差为损失,则有损失函数为

$$L(\mathbf{T}_v^c) = \int \left\| \left(\frac{\mathbf{M} \mathbf{T}_v^c \mathbf{T}_w^v \mathbf{P}_w}{z_c} - (u, v, 1)^T \right) \right\| d\mathbf{P}_w \quad (2)$$

式中, $\mathbf{P}_w$ 为高精度地图中车道线在世界坐标系下的位置, $\mathbf{T}_w^v$ 通过全球定位系统(global positioning system, GPS)、惯性测量单元(inertial measurement unit, IMU)以及轮速等信息融合获得。由于 $\mathbf{P}_w$ 以及 $\mathbf{T}_w^v$ 已知,则损失为 $\mathbf{T}_v^c$ 的函数。

车辆在正常行驶过程中会逐渐遍历部分或全部车道,联立车辆在不同位姿下的损失,则相机的外参标定问题可以等价为最小化损失的优化问题,即

$$\hat{\mathbf{T}}_v^c = \operatorname{argmin} L(\mathbf{T}_v^c) \quad (3)$$

由于车道线在车辆纵向上无明显纹理特征,因此无法建立 $\mathbf{P}_w$ 与 $(u, v)^T$ 之间的一一映射。故式(3)无法直接求解,为简化处理,将上式中的点到点的误差转换为点集到点集的误差,则有

$$L(\mathbf{T}_v^c) = \int \left\| \frac{\mathbf{M} \mathbf{T}_v^c \mathbf{T}_w^v \mathbf{P}_w}{z_c} - (u, v, 1)^T \right\| dt \quad (4)$$

2 标定方法

由于式(4)难以表示为数学解析形式,本文采用数值解法对外参进行估计,其过程可分为车道线检测、数据筛选以及数值优化 3 个部分。

首先在图像上检测车道线获得在像平面内的车道线特征,为防止过优化以及提升计算效率对检出的特征进行了筛选,当车道线特征数量(图像帧数)大于一定阈值时采用梯度下降方法对相机外参进行优化。该过程可在车道线特征收集的同时进行优化,从而实时迭代进行。

由于相机的测量特性,将像平面内的 2 维特征直接恢复至 3 维世界坐标系较为困难,因此本文将车道线真值投影至像平面,并在像平面内设置损失

函数,合理设定初值,并采用对损失函数进行反向求导迭代更新相机外参,使得地图以及检测的车道线特征重投影误差最小。

2.1 车道线检测

本文采用深度学习中的像素分割技术对车道线进行实时检测。U-Net ++ (Zhou 等, 2018) 是图像分割领域常用的一种网络结构,常用于医学图像处理等领域,基于该网络本文定义了左车道线以及右车道线两个语义要素,并经过训练实现车道线检测,其检测结果如图 2 所示。

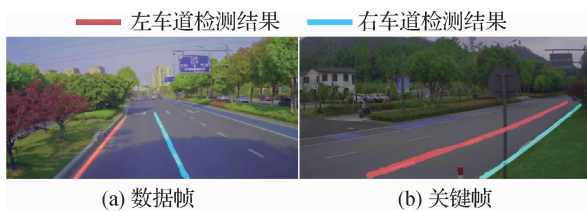


图 2 车道线检测结果

Fig. 2 Result of lane detection

((a) data frame; (b) key frame)

2.2 数据筛选

车道线通常由曲线以及直线构成,无论是曲线或直线,其曲率均较小。极限情况下,在车道线曲率为 0 且车辆不频繁变道时,对于任意形式的 $L(T)$ 有

$$\frac{\partial L}{\partial T_x} \equiv 0 \quad (5)$$

即车道线对于标定纵向平移 T_x 并不提供有用的信息,这与汽车驾驶常识一致。因此对于 T_x 的标定,需选择车辆转向、并道和掉头等明显偏航的场景。

将车载相机采集的图像 I 分为无用帧、数据帧 I_D 以及关键帧 I_K , 设置如下规则进行帧筛选:

1) 当图像中检出的车道线像素个数小于一定阈值时作为无用帧,以避免车辆在通过路口以及交通拥堵时图像内无明显车道线;

2) 距离上一关键帧车辆行驶距离以及车辆偏航角度均小于一定阈值时作为无用帧,以避免采集重复车道线信息;

3) 在不满足规则 1) 和 2), 且车辆与车道线真值(地图)夹角大于一定阈值时作为关键帧;

4) 其他情形所采集图像作为数据帧。

由于无用帧不包含车道线信息,或仅包含已统

计的车道线信息,因此不参与 $L(T_v^c)$ 优化。

记数据帧以及关键帧的数量分别为 N_D 和 N_K , 显然 $N_D \gg N_K$, 故不区分关键帧以及数据帧将导致外参对 $(T_y, T_z, \theta_{yaw}, \theta_{pitch}, \theta_{roll})$ 过优化,而 T_x 陷入局部极小。因此,本文设计如下更新规则:

当 $N_D + N_K > C_1$ 且 $N_K \geq C_2$ 时,已收集足够的关键帧,故对 $(T_x, T_y, T_z, \theta_{yaw}, \theta_{pitch}, \theta_{roll})$ 进行优化。

当 $N_D + N_K > C_1$ 且 $N_K < C_2$ 时,所收集关键帧较少,仅对 $(T_y, T_z, \theta_{yaw}, \theta_{pitch}, \theta_{roll})$ 进行优化。

C_1 与 C_2 为预设阈值。显然,该阈值越大,所优化外参越趋近于真值,但过大的阈值将使得数据收集过程变长且解优化复杂度上升。另外,在车辆实际行驶过程中,其通常平行于车道线,所收集的图像关键帧相对于数据帧数量较少。

2.3 重投影误差

对于 $\hat{T}_v^c$ 求解的关键在于重投影误差 L 的定义以及数值优化方法的选择。重投影误差 L 是衡量在某一确定的外参 T_v^c 下相机所检测的车道线与地图中车道线真值的偏离程度。

重投影误差可以定义在像平面坐标系或车辆坐标系下,后者误差通常为欧氏距离,有明显的物理含义。考虑到实际道路的高低起伏(包括上下坡、排水等因素),地面通常为 3D 曲面,此时将单目相机检测的车道线反向投影至车辆坐标系计算复杂度较高,因此本文在像平面坐标系定义重投影误差,记像平面内检出的车道线点位置为 (u_i, v_i) , 对应该车道点的法线方向为 φ_i , 则式(4)可转换为

$$L = \sum_i^j (\omega_1 \|(u_i - u_j^w, v_i - v_j^w)\| + \omega_2 \|\varphi_i - \varphi_j^w\|) \quad (6)$$

式中,法线方向 φ_i 具体计算参考 Ouyang 和 Feng (2005) 的方法。 (u_j^w, v_j^w) 为地图中车道线点在像平面内的投影 $P_j^{c'}$ 。如前文所述,为解决车道线点的点匹配问题,本文采用最近邻域方法 (Muja 和 Lowe, 2014) 生成每一个检测的车道线点 P_i^c 对应的真实车道线在像平面内的投影 $P_j^{c'}$, 即重投影误差计算过程可表述为如下步骤:

1) 基于相机外参 T_v^c 以及车辆位姿 T_w^v 将地图中的车道线点集(距离车辆 200 m 范围内)投影至相机坐标系;

2) 根据相机内参矩阵将已进行坐标系变换点

集投影至像平面;

3) 计算已投影地图车道线点集以及检出的车道线点集法线方向;

4) 根据最近邻域原则确定地图车道线点以及检出车道线点的关联;

5) 根据式(6)确定重投影误差。

根据式(5), 为防止纵向平移 T_x 发散或陷入局部极小, 需提高关键帧 I_K 的重投影误差, 故

$$(\omega_1, \omega_2) = \begin{cases} (\omega_1^c, \omega_2^c) & I \in I_D \\ \frac{N_D}{N_K}(\omega_1^c, \omega_2^c) & I \in I_K \end{cases} \quad (7)$$

式中, (ω_1^c, ω_2^c) 为超参数。

2.4 非线性优化

最小化重投影误差优化方法通常包括最速下降法、Newton 法、Gauss-Newton 法以及 Levenberg-Marquadt 方法 (Triggs 等, 1999)。其中, 最速下降法计算简单, 迭代过程中每一步均能保证损失函数下降, 稳定性较高。本文采用最速下降法对式(2)进行非线性优化, 优化步骤可分为初值估计、梯度下降以及收敛判定。

2.4.1 初值估计

正确的初值可加快 $\hat{T}_v^c$ 的收敛, 并避免陷入局部极小或发散。根据 ADAS 前视相机一般安装规则以及道路建设规范, 可做如下简化:

1) 车辆所处的地面为平面;

2) 车载相机检测到的车道线为直线, 且左右车道线相互平行;

3) 车载相机正向安装且安装纵向位置 T_x 以及安装高度 T_z 已知。

记在相机坐标系下的直线方程为

$$\begin{cases} x = A_1 z + B_1 \\ y = A_2 z + B_2 \end{cases} \quad (8)$$

值得指出的是, 式(8)在极端情况下会失效, 但不影响本文后续推导及结论。代入式(1)有

$$\begin{cases} u = u_0 + f_x \left(A_1 + \frac{B_1}{z} \right) \\ v = v_0 + f_y \left(A_2 + \frac{B_2}{z} \right) \end{cases}, \quad M = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (9)$$

根据式(9)知像平面内车道线为直线, 可在像平面内对直线进行拟合, 拟合后方程为 $u = Av + B$ 。

考虑车道线远处逐渐相交的约束, 通过直线方程计算左右车道线的交点 $z \rightarrow \infty$ 时 (u, v) 坐标, 有

$$\begin{aligned} u_\infty &= \lim_{z \rightarrow \infty} u = A_1 f_x + u_0 \\ v_\infty &= \lim_{z \rightarrow \infty} v = A_2 f_y + v_0 \end{aligned} \quad (10)$$

由于相机安装高度 T_z 已知, 忽略俯仰角 θ_{pitch} 的影响, 则相机检测的最近车道点在相机坐标系下 $y = T_z$, 联立式(9)(10)及 $y = T_z$, 可求得相机坐标系下的车道线方程。

将地图中车道线投影至车辆坐标系, 并拟合直线方程。对上述两直线进行配准即可获得初始相机安装参数 $(T_x, T_y, \theta_{yaw}, \theta_{pitch}, \theta_{roll})$ 。

2.4.2 梯度下降

沿着梯度的负方向可迭代寻找使式(6)最小的变量值, 即

$$T_k = T_{k-1} - \lambda \nabla L(T_{k-1}) \quad (11)$$

式中, λ 为步长。步长决定了梯度下降收敛速度以及精度, 本文选择步长为应用需求精度上限的四分之一。由于式(4)无法解析表示, 本文采用离散数值方法计算梯度。

3 实验

采用实际路测车辆在浙江某地开放道路进行算法验证。测试车辆配备多种传感器, 如图3所示, 其配置安装了 NovAtel 组合导航定位系统、32 线激光雷达以及多个焦距为 6 mm 的分辨率为 $1\,920 \times 1\,080$ 像素的车载相机。计算资源方面采用了配备 Intel i7 7700 CPU, 8 GB DDR3 RAM 以及 NVIDIA 1080TI 显卡的工业控制计算机。

测试车在浙江德清公开道路正常行驶, 行驶路径如图4所示, 图中点集为车辆行驶路径, 背景为谷歌地球。其总路程约 11.7 km, 该路程中包括了路口、变道、并线、掉头以及左右转等种场景。

通过测试车的车载激光雷达收集道路点云数据, 经过运动补偿、坐标变换(由组合导航定位系统提供)、人工语义标注及后处理可获得高精度地图数据如图5所示。随机选择 50 个车道线端点, 对比高精度 RTK (real time kinematic) 可知该地图精度(标准差)约为 10 cm。

3.1 重投影误差与外参误差关系

采用 $(100\text{ m}, 2^\circ)$ 的采样间隔, 分析函数 $L(T_v^c)$



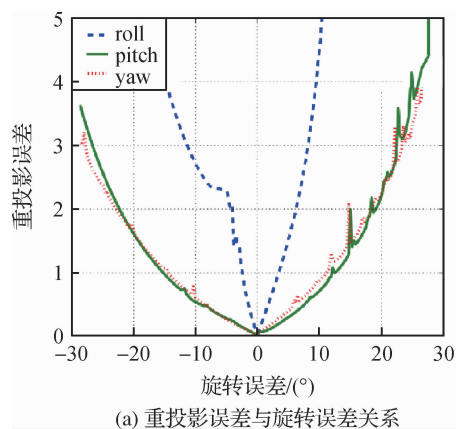
图3 实验车辆配置

Fig. 3 Vehicle configuration for experiment((a) vehicle for experiment; (b) lidar and cameras; (c) front camera; (d) RTK + IMU positioning system)

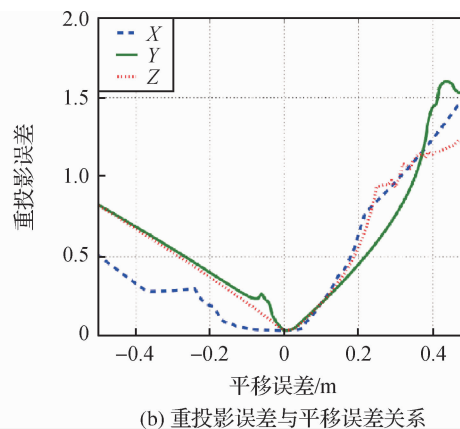


图4 谷歌地球上的实验车辆行驶路径

Fig. 4 Driving path of vehicle on Google Earth



(a) 重投影误差与旋转误差关系



(b) 重投影误差与平移误差关系

图6 重投影误差与外参误差关系

Fig. 6 The relationship between error and extrinsic

((a) relationship between error and rotation; (b) relationship between error and translation)

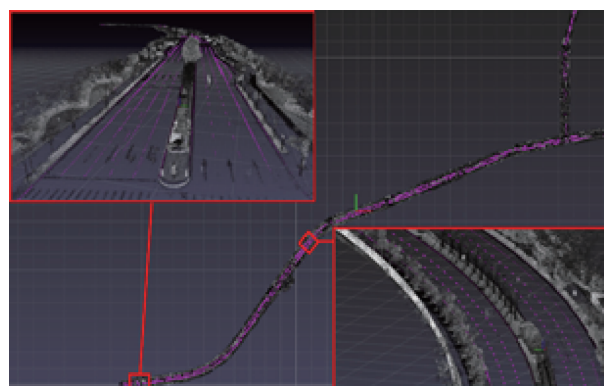


图5 测试路段高精度地图(UTM 坐标系下的3 维点云, 每个点具备属性信息, 浅红色点为车道线点)

Fig. 5 Hd map of the test road(each 3D point cloud in the UTM coordinate system has attribute information, the light red point is the lane point)

在不同外参下的下降梯度。由于外参维数较高,采用单一变量法进行简单分析,可得全路段的 $L(T_v^e)$ 与外参误差关系如图 6 所示。

由图 6 可知,在旋转误差 $\theta \in [-30^\circ, 30^\circ]$ 且平移误差 $T_n \in [-0.5 \text{ m}, 0.5 \text{ m}]$ 范围内 $L(T)$ 无局部极小,且随着外参误差逼近 0,损失值也单调逼近全局极小。因此,所定义 $L(T)$ 合理有效,采用梯度下降方法可快速收敛至最优外参。

3.2 梯度下降优化

设定 $(50 \text{ m}, 1^\circ)$ 的采样间隔,超参数 $C_1 = 20$, $C_2 = 20$,对上述数据进行优化。采用 pytorch 进行梯度下降,单次梯度下降约为 50 ms。如图 7 所示,损失在求解过程中稳步下降,且在第 231 次迭代处收敛,收敛后相机旋转角精度小于 0.2° ,相机平移

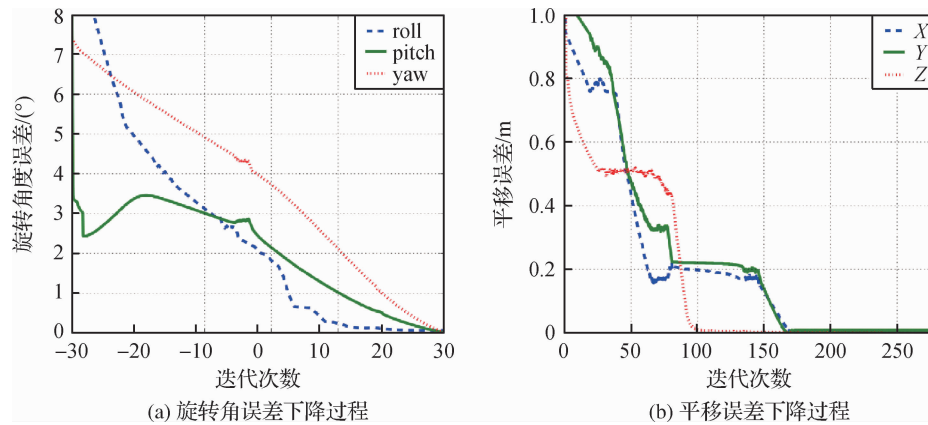


图7 外参误差收敛过程

Fig.7 Convergence process of extrinsic error((a) rotation; (b) translation)

小于0.1 m。

由于相机外参会随着时间改变而改变,为模拟相机安装位姿的实时变化,在 $t = 64$ 时刻切换相机输入。随着数据帧或关键帧图像获取,实时滚动对外参进行优化。在 $t + 1$ 时刻,由于大量图像数据于外参变换前采集,而对应于正确外参的图像数据此时仅为1帧,因此算法收敛至错误的外参。如图8

所示,随着新的有效图像逐渐参与优化,逐渐逼近新的外参,在 $t + 1 + C_1$ 后,所有采集的图像均对应一个新的真实外参,其精度恢复至(0.1 m,0.1°)。由于滚动优化的进行,外参迭代初值与最优外参误差较小,即梯度下降迭代次数较小。在 $t + 1 + C_1$ 时刻,梯度下降迭代次数为21次,优化时间为1 016 ms,满足实时应用。

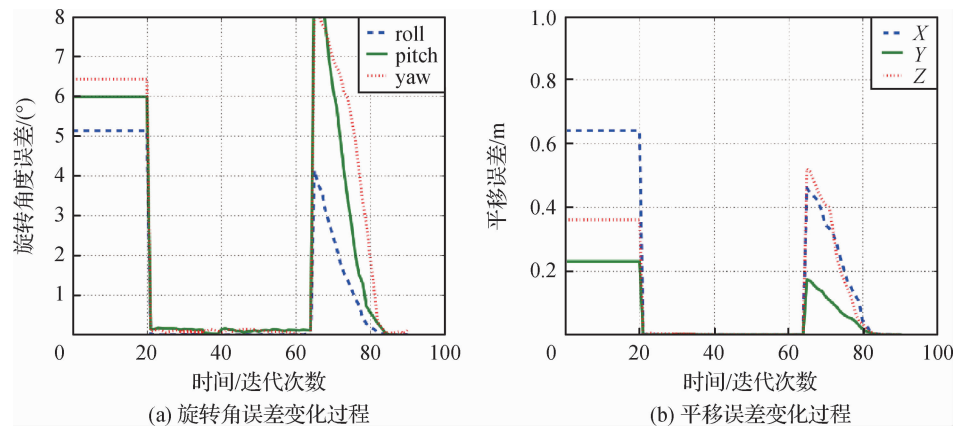


图8 外参突变时收敛过程

Fig.8 Convergence process of extrinsic when groundtruth change((a) rotaion; (b) translation)

3.3 对比实验

确定外参的方法较多,但与本文所研究问题完全相似的较少。为证明所提方法的有效性,本文将所提方法与基于VINS标定方法(Yang和Shen, 2017)以及基于消失点标定方法(李婵等,2019)进行外参标定精度对比,如表1所示。VINS标定方法输出的是相机与IMU之间的坐标系关系,在已知IMU与车辆坐标系间的坐标系变化时,该标定问题与求解 $\hat{T}_v^c$ 等价。基于消失点标定方法输出的是相

机与世界之间的坐标系关系,在已知车辆实时位姿

表1 不同方法的精度对比

Table 1 Accuracy comparison of different methods

	方法		
	VINS	消失点	本文
角度误差/(°)	2.2	6	0.2
平移误差/m	0.34	N/A	0.03

注:N/A为不适用(not applicable)。

时,该标定问题也与求解 $\hat{T}_v^c$ 等价。由于该方法依赖于相机安装高度已知,且仅标定滚转角和偏航角,故不对外参平移部分进行对比。

对比可知,本文方法以及 VINS 标定法外参精度明显优于消失点标定法,这是因为消失点标定法仅利用了车道线信息,而前者利用了更多场景的车道线以及车辆运动信息,从信息论角度出发,消失点标定法先天精度较差。

相对于 VINS 标定法,本文方法在外参角度以及平移精度上更优。其原因主要有两个方面:1) 通过深度学习提取车道线特征,并将其与地图真值匹配,其消除了相机匹配相机时的观测误差;2) 本文通过合理设计数据筛选规则,避免了由于车辆运动曲率较小、部分外参不可观从而导致估计误差较大的缺陷。

4 结 论

针对在开放道路上的相机外参标定问题,本文提出了一种基于车道线特征的相机地图重投影误差最小化的在线标定方法。其主要内容包括:1) 在可行的计算复杂度内定义了重投影误差,提高了外参标定精度;2) 提出了数据更新规则,并对纵向平移以及其他外参部分分别进行梯度下降优化,有效避免了纵向不可观以及优化过程发散问题。对比实验表明,本文方法的估计精度明显优于基于 VINS 的优化方法以及基于特征解算的消失点方法,另外当外参改变时,本文方法能够迅速收敛至突变后的外参,满足在线实时应用。

值得注意的是,车辆位姿误差以及高精度地图误差在实际应用场景中不可避免,其对外参标定结果的影响本文尚未探讨。在后续工作中,将根据定位概率分布以及地图概率分布建立投影误差统计学模型,从理论上分析外参收敛依据以及误差范围。

参考文献 (References)

- Carrera G, Angeli A and Davison A J. 2011. SLAM-based automatic extrinsic calibration of a multi-camera rig//Proceedings of 2011 IEEE International Conference on Robotics and Automation. Shanghai, China; IEEE: 2652-2659 [DOI: 10.1109/ICRA.2011.5980294]
- Civera J, Bueno D R, Davison A J and Montiel J M M. 2009. Camera self-calibration for sequential Bayesian structure from motion//Proceedings of 2009 IEEE International Conference on Robotics and Automation. Kobe, Japan; IEEE: 403-408 [DOI: 10.1109/ROBOT.2009.5152719]
- Dabral S, Kamath S, Appia V, Mody M, Zhang B Y and Batur U. 2014. Trends in camera based automotive driver assistance systems (ADAS)//The 57th IEEE International Midwest Symposium on Circuits and Systems (MWSCAS). College Station, USA; IEEE: 1110-1115 [DOI: 10.1109/MWSCAS.2014.6908613]
- Daniilidis K. 1999. Hand-eye calibration using dual quaternions. The International Journal of Robotics Research, 18(3): 286-298 [DOI: 10.1177/02783649922066213]
- Esquivel S, Woelk F and Koch R. 2007. Calibration of a multi-camera rig from non-overlapping views//The 29th DAGM Symposium on Pattern Recognition Symposium. Heidelberg, Germany; Springer: 82-91 [DOI: 10.1007/978-3-540-74936-3_9]
- Ghallabi F, Nashashibi F, El-Haj-Shhade G and Mittet M A. 2018. LIDAR-based lane marking detection for vehicle positioning in an HD map//Proceedings of the 21st International Conference on Intelligent Transportation Systems (ITSC). Maui, USA; IEEE: 2209-2214 [DOI: 10.1109/ITSC.2018.8569951]
- Knoop V L, De Bakker P F, Tiberius C C J M and Van Arem B. 2017. Lane determination with GPS precise point positioning. IEEE Transactions on Intelligent Transportation Systems, 18(9): 2503-2513 [DOI: 10.1109/TITS.2016.2632751]
- Li C, Song H S, Wu F F, Wang W and Wang X. 2019. Auto-calibration of the PTZ camera on the highway. Journal of Image and Graphics, 24(8): 1391-1399 (李婵, 宋煥生, 武非凡, 王伟, 王璇. 2019. 高速公路云台相机的自动标定. 中国图象图形学报, 24(8): 1391-1399) [DOI: 10.11834/JIG.180599]
- Li T, Zhang H P, Niu X J and Gao Z Z. 2017. Tightly-coupled integration of multi-GNSS single-frequency RTK and MEMS-IMU for enhanced positioning performance. Sensors, 17(11): #2462 [DOI: 10.3390/s17112462]
- Muja M and Lowe D G. 2014. Scalable nearest neighbor algorithms for high dimensional data. IEEE Transactions on Pattern Analysis and Machine Intelligence, 36(11): 2227-2240 [DOI: 10.1109/TPAMI.2014.2321376]
- Ouyang D S and Feng H Y. 2005. On the normal vector estimation for point cloud data from smooth surfaces. Computer-Aided Design, 37(10): 1071-1079 [DOI: 10.1016/j.cad.2004.11.005]
- Schoepflin T N and Dailey D J. 2003. Dynamic camera calibration of roadside traffic management cameras for vehicle speed estimation. IEEE Transactions on Intelligent Transportation Systems, 4(2): 90-98 [DOI: 10.1109/TITS.2003.821213]
- Triggs B, McLauchlan P F, Hartley R I and Fitzgibbon A W. 1999. Bundle adjustment—a modern synthesis//Proceedings of the International Workshop on Vision Algorithms: Theory and Practice.

Greece; Springer: 298-372 [DOI: 10.1007/3-540-44480-7_21]

Tsai R. 1987. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. IEEE Journal on Robotics and Automation, 3(4): 323-344 [DOI: 10.1109/JRA.1987.1087109]

Yang Z F and Shen S J. 2017. Monocular visual-inertial state estimation with online initialization and camera-IMU extrinsic calibration. IEEE Transactions on Automation Science and Engineering, 14(1): 39-51 [DOI: 10.1109/TASE.2016.2550621]

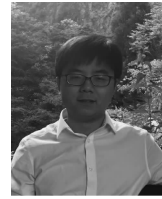
Zhang Z. 2000. A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11): 1330-1334 [DOI: 10.1109/34.888718]

Zhang Z Y. 2004. Camera calibration with one-dimensional objects. IEEE Transactions on Pattern Analysis and Machine Intelligence, 26(7): 892-899 [DOI: 10.1109/TPAMI.2004.21]

Zhou X, Huang X Y and Li Y. 2003. Lane keeping and distance measurement based on monocular vision. Journal of Image and Graphics, 8(5): 590-595 (周欣, 黄席樾, 黎昱. 2003. 基于单目视觉的高速公路车道保持与距离测量. 中国图象图形学报, 8(5): 590-595) [DOI: 10.3969/j.issn.1006-8961.2003.05.017]

Zhou Z W, Siddiquee M M R, Tajbakhsh N and Liang J M. 2018. UNet ++: a nested u-net architecture for medical image segmentation//Proceedings of the 4th International Workshop on Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Granada, Spain; Springer: 3-11 [DOI: 10.1007/978-3-030-00889-5_1]

作者简介



廖文龙, 1989年生, 男, 博士研究生, 主要研究方向为机器学习与智能驾驶。

E-mail: igoliao@sjtu.edu.cn



严骏驰, 通信作者, 男, 副教授, 主要研究方向为机器学习及计算机视觉。

E-mail: yanjunchi@sjtu.edu.cn

赵华卿, 女, 工程师, 主要研究方向为人工智能与机器视觉。

E-mail: zhaohuaqing@cowarobot.com