

移动机器人视觉里程计技术研究综述

陈明方 黄良恩 王 森 张永霞 陈中平

(昆明理工大学机电工程学院, 昆明 650500)

摘要: 随着移动机器人技术不断发展,里程计技术已经成为移动机器人实现环境感知的关键技术,其发展水平对提高机器人的自主化和智能化具有重要意义。首先,系统阐述了同步定位与地图构建(Simultaneous localization and mapping, SLAM)中激光SLAM和视觉SLAM的发展近况,阐述了经典SLAM框架及其数学描述,简要介绍了3类常见相机的相机模型及其视觉里程计的数学描述。其次,分别对传统视觉里程计和深度学习里程计的研究进展进行系统阐述。对比分析了近10年来各类里程计算法的优势与不足。另外,对比分析了7种常用数据集的性能。最后,从精度、鲁棒性、数据集、多模态等方面总结了里程计技术面临的问题,从提高算法实时性、鲁棒性等方面展望了视觉里程计的发展趋势:更加智能化、小型化新型传感器的发展;与无监督学习融合;语义表达技术的提高;集群机器人协同技术的发展。

关键词: 视觉里程计; 特征法; 直接法; 深度学习; 同步定位与地图构建; 数据集

中图分类号: TP242.6; TP391.41 **文献标识码:** A **文章编号:** 1000-1298(2024)03-0001-20

OSID:



Survey of Research on Visual Odometry Technology for Mobile Robots

CHEN Mingfang HUANG Liang'en WANG Sen ZHANG Yongxia CHEN Zhongping

(Faculty of Mechanical and Electrical Engineering, Kunming University of Science and Technology, Kunming 650500, China)

Abstract: With the continuous development of mobile robot technology, odometry technology has become a key technology for mobile robots to realize environmental perception, and its development level is of great significance to improving the autonomy and intelligence of robots. Firstly, the current development status of laser simultaneous localization and mapping (SLAM) and visual SLAM in simultaneous localization and mapping was systematically explained. The classic SLAM framework and its mathematical description were expounded, and the camera models of three common types of cameras and their mathematical description of visual odometry were briefly introduced. Secondly, the research progress of traditional visual odometry and deep learning odometry were systematically elaborated. The advantages and disadvantages of various mileage calculation methods in the past ten years were compared and analyzed. In addition, the performance of seven commonly used data sets was comparatively analyzed. Finally, the problems faced by odometry technology were summarized from the aspects of accuracy, robustness, data sets, and multi-modality, and five development trends of visual odometry were prospected from the aspects of improving the real-time performance and robustness of the algorithm. For the development of more intelligent and miniaturized new sensors, the integration with unsupervised learning, the improvement of semantic expression technology and the development of cluster robot collaboration technology were introduced.

Key words: visual odometry; feature method; direct method; deep learning; SLAM; dataset

0 引言

随着科学技术的不断发展,机器人技术发生了

重大变革。目前,机器人已经进入智能化时代,环境感知技术的发展迫在眉睫。

同步定位与地图构建技术(Simultaneous

收稿日期: 2023-07-15 修回日期: 2023-10-10

基金项目: 国家自然科学基金项目(51965029)、国家重点研发计划项目(2018YFB1306103)和云南省重大专项(202002AC080001)

作者简介: 陈明方(1975—),男,教授,博士,主要从事机器人控制理论和运动感知技术研究,E-mail: mfcchen_robot@126.com

通信作者: 王森(1983—),男,副教授,博士,主要从事视觉感知、测量和工业视觉检测研究,E-mail: wangsen0401@126.com

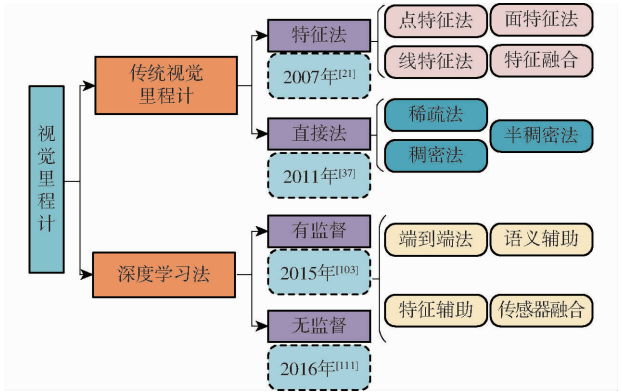
localization and mapping, SLAM) 的出现和发展,极大促进了机器人环境感知技术的发展。SMTH 等^[1]首先提出了 SLAM 的概念,并用于实现机器人的同时定位与地图构建^[2],其主要解决“我在哪里”、“我要去哪里”、“我应该怎么去”3 个问题^[3]。经过近 40 年的发展,SLAM 技术已取得突破性研究进展^[4-5],相关算法也成为移动机器人运动控制的重要研究内容^[6]。

里程计是 SLAM 技术的关键,直接决定 SLAM 技术的性能。目前,国外诸多学者已发表较多与里程技术相关的综述文献。KAZEBOURONI 等^[7]介绍了视觉 SLAM 的基本定义,综述了相关的视觉里程计方法。JIA 等^[8]根据 SLAM 系统的框架,介绍了视觉传感器、视觉里程计、后端优化及回环检测,并综述了关键技术。SERVIÈRES 等^[9]介绍了特征法、直接法和相关的评估数据集。GUPTA 等^[10]以无人机领域的 SLAM 技术为切入点,综述了无人机领域的 SLAM 技术和融合图像处理的 SLAM 技术。SHALAL 等^[11]详细论述了农业机器人中视觉传感器特征提取、数据处理及传感器融合的相关算法。AGUIAR 等^[12]针对果园场景 SLAM 建图技术,阐述了农林机器人视觉传感器上数据关联、处理及视觉 SLAM 建图策略。FASIOLO 等^[13]介绍了当前农业领域主流的移动机器人硬件平台,相关 SLAM 技术中主流的激光传感器和视觉传感器。GAO 等^[14]针对视觉传感器详述了 SLAM 技术中立体相机的使用方法,并比较了几种与立体视觉相关的 SLAM 开源框架。

国内对里程计技术的系统阐述较少,主要侧重于硬件设计或硬件性能方面。BAI 等^[15]介绍了农业机器人导航技术的研究进展,展望了各种视觉传感器的发展趋势。DING 等^[16]介绍了复杂非结构化农业环境中完成定位和建图的单一传感器技术、多传感器融合技术。陈世浪等^[17]总结了近年来视觉 SLAM 硬件平台研究取得的进展,对比分析了 ZED 双目相机、Event 相机、Kinect 一代结构光相机及二代 TOF 相机等传感器的性能。另外,当前国内能检索的相关文献,大多年代较远,缺乏对最新算法的比较和分析,无法反映里程计的最新研究进展^[18-20]。

本文将按照视觉里程计发展的时间顺序,将其分为传统视觉里程计和基于深度学习的视觉里程计,系统全面阐述各类视觉里程计相关技术研究现状。图 1 为视觉里程计发展的总体框架图。

首先,介绍 SLAM 系统及其数学描述,根据传感器不同分别介绍激光 SLAM 和视觉 SLAM。分析 SALM 算法的框架、数学描述及 3 种相机里程计的



数学描述。其次,将视觉里程计分为传统视觉里程计和深度学习视觉里程计,系统阐述各类里程计技术的研究现状、成果和算法的应用效果,对比总结各类算法的优势与不足。另外,介绍 7 种较为经典数据集的性能及应用情况。最后,探讨 3 类里程计方法的联系及不足,并提出视觉里程计性能的评价指标。展望视觉 SLAM 及其里程计技术的未来发展趋势,为相关研究工作提供参考。

1 经典 SLAM

1.1 SLAM 分类

根据获取环境信息的传感器不同,SLAM 技术可分为激光 SLAM 和视觉 SLAM。激光 SLAM 的传感器主要为 2D、3D 激光雷达。视觉 SLAM 的传感器主要为相机,主要有单目相机^[21]、双目相机^[22]、RGB-D 相机^[23]和事件相机等^[24]。表 1 为几种常见视觉传感器的优缺点和主要应用场景。

1.1.1 激光 SLAM

激光 SLAM 技术发展历史较为悠久,算法较为成熟^[25-27],代表性的有 Gmapping 算法^[28]、Gartographer 算法^[29]、LOAM (LiDAR odometry and mapping, LOAM) 算法^[30]和改进的 LOAM 算法等^[31-32]。Gmapping 和 Gartographer 算法通常采用单线程激光雷达,仅适用于室内。Cartographer 虽然同时支持 2D 和 3D 建图模式,但 3D 建图模式下构建的地图格式仍为 2D。LOAM 算法采用多线程激光雷达,可在室外运行,能构建 3D 点云地图^[33]。虽然激光 SLAM 技术较成熟,但探测范围有限且价格昂贵,导致技术推广困难。

1.1.2 视觉 SLAM

近年来,随着相机技术的发展,涌现出单目、双目及 RGB-D 等一大批产品,促使视觉 SLAM 迅猛发展。根据图像数据处理方式不同,视觉 SLAM 技

表 1 传感器性能

Tab.1 Sensor performance

传感器类型	优点	缺点	应用场景
激光雷达	鲁棒性好、稳定性较高；计算量低，对 CPU 的要求低于相机传感器	无法获取语义信息；不适用恶劣环境，如雨雾天气；无法获取透视体的深度信息	室内低速小范围场景
单目相机	结构简单；成本较低；标定和识别过程容易	无法确定单独图像的深度信息和物体的真实大小	室内、室外场景
双目相机	能确定物体真实尺度	计算量大；标定过程复杂；需要 GPU 或 FPGA 加速，耗费算力巨大	室内、室外小范围场景
RGB-D 相机	主动性较强，计算量小	视野窄，可测量范围小，易受光照影响，无法识别透射体	室内小范围场景
事件相机	低延迟；计算功耗小，算力要求低；高动态范围	数据稀疏性较强；冗余信息较多，有效信息少	高速、高动态场景

术可分为特征法、直接法和半直接法，近年来随着深度学习技术的发展，又出现了深度学习 SLAM 算法。特征法的代表性算法有适用于单目相机的 PTAM (Parallel tracking and mapping) 算法^[21,34]和适用于单目、双目、RGB-D 相机的 ORB_SLAM2 (Oriented features from accelerated segment test and rotated binary robust independent elementary feature - SLAM2) 算法^[35-36]。直接法的代表性算法有 DTAM (Density tracking and mapping)、LSD-SLAM (Large-scale direct monocular - SLAM)^[33,37-38]、DSO (Direct sparse odometry) 算法^[39]。半直接法主要是 SVO (Semi-direct monocular visual odo-metry) 算法^[33]。

虽然相机视野较大且价格便宜，可很好地弥补激光 SLAM 的不足，但相机对光照较敏感，不适用于光照较强的弱纹理场景和光照变化的场景。目前，两类 SLAM 技术均只在单一环境或相对稳定的环境中效果较好；但是在动态环境中，地图构建与定位的误差较大，甚至会出错。

1.2 经典 SLAM 框架

在国内外专家努力下，SLAM 已经形成了一套较成熟的框架，如图 2 所示。主要由前端（处理传感器数据）、后端（优化前端数据）和回环检测（提高系统鲁棒性）等环节构成。

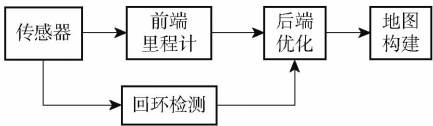


图 2 经典 SLAM 框架

Fig.2 Classic SLAM framework

传感器主要有激光和视觉两类，激光 SLAM 里程计主要实现关键帧的提取与匹配，视觉 SLAM 里程计主要用于估算相邻图像间相机的运动及局部地图。后端优化主要对里程计及回环检测进行统一优化，获得全局地图和轨迹。回环检测是重复定位的关键，当检测到回环时，后端将对回环信息进行处

理。地图构建则主要依赖于轨迹，根据机器人运动轨迹就可完成地图构建。

1.3 数学描述

移动机器人的 SLAM 问题即运动和观测问题。由于 SLAM 技术通过算法实现，需将 SLAM 问题数学描述，数学方程为运动方程和观测方程。

由于传感器以离散点的方式采集各时刻机器人的位置及其周围环境信息，因此需离散化机器人的运动，进而确定各离散时刻的位置和路标点。设机器人的位置为 x ，离散化时间 k 取正整数，机器人在 k 时刻的位置为 x_k ，同理， $k-1$ 时刻的位置为 x_{k-1} 。设传感器离散输入数据为 u ，则 k 时刻传感器的数据输入为 u_k 。设机器人运动过程中噪声为 w ，则 k 时刻的噪声为 w_k ，因此，机器人运动方程可抽象表示为

$$x_k = f(x_{k-1}, u_k, w_k) \tag{1}$$

机器人在运动过程中需确定周围存在的“特征体”，即需确定机器人的观测方程。以“路标点”作为被观察对象， y 表示路标点，则第 j 个路标点可表示为 y_j 。设在 k 时刻观测第 j 个路标点，观测时刻输入的噪声为 $v_{k,j}$ ，则机器人观测方程为

$$z_{k,j} = h(y_j, x_k, v_{k,j}) \tag{2}$$

式中 $z_{k,j}$ ——观测量

将 SLAM 过程用形象化的图形表示，则 SLAM 过程的运动方程和观测方程的位姿路标图可表示为图 3。

2 视觉里程计模型

激光 SLAM 传感器为 2D 或 3D 激光雷达。由于激光雷达的扫描范围较小、价格昂贵，导致推广困难。视觉 SLAM 因传感器具有智能化程度高、感知信息丰富等优点，得到了迅速发展和广泛的应用，其前端即为视觉里程计。

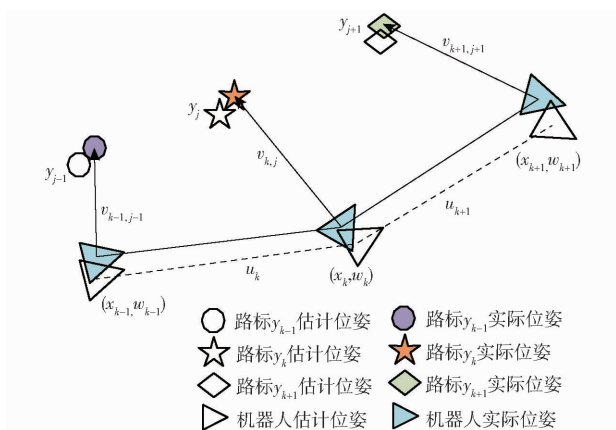


图3 位姿路标图

Fig. 3 Orientation roadmap

2.1 相机模型

视觉里程计是根据相机获取的信息估计相机运动。相机获取三维空间的环境信息并将其投影至二维平面,该过程简化为一个几何模型^[2],据此可得相机模型。

本文主要介绍3种较为成熟的相机模型,即单目、双目和RGB-D相机模型及其里程计。

2.1.1 单目相机

图4为单目相机成像模型,成像平面和相机平面的坐标系分别为 $o_1x_1y_1z_1$ 、 $oxyz$,相机焦距为 f 。

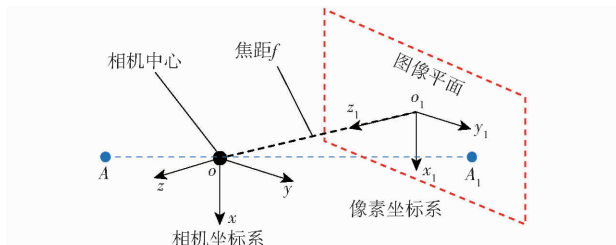


图4 单目相机成像模型

Fig. 4 Monocular camera phase model

根据空间点 A 在相机坐标系 $oxyz$ 下的坐标 A 、图像平面中点 A_1 在像素坐标系 $o_1x_1y_1z_1$ 下的坐标 A_1 和像素坐标 A_2 、相机的内参矩阵 B ,可得图像平面中点 A_1 的像素坐标为

$$zA_2 = BA \quad (3)$$

式中 z ——点 A 竖坐标

2.1.2 双目相机

单目相机模型中(图4),直线 AA_1 上全部点在图像平面的投影均为点 A_1 ,因此无法确定点 A 的空间位置。图5为双目相机模型,该相机由左右目相机构成,左右相机水平放置,根据视差可确定点 A 的空间位置。工作时,左右相机同时采集工作区域的图像,通过图像间的视差可得到图像的像素深度^[3]。图5中,根据三角形相似关系得

$$\frac{z}{z+f} = \frac{l_{C_L C_R}}{l_{A_L A_R}} \quad (4)$$

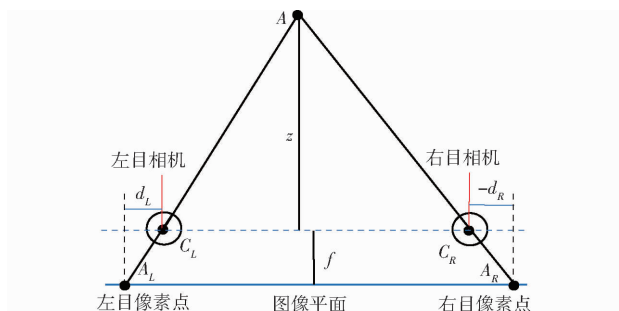


图5 双目相机模型

Fig. 5 Binocular camera model

相机视差可由左、右目相机视差求得,利用视差即可估算像素点到相机中心点距离。

2.1.3 RGB-D相机

RGB-D相机主动性较强,无需搭建中间桥梁,就可获取像素点的深度信息。RGB-D相机主要有2类:利用红外结构光原理,通过发射和接收测距,代表性产品有Kinect一代^[40];利用飞行时间原理,通过脉冲光的发射和接收测距,代表性产品有Kinect二代^[41]。

图6为一种常见的RGB-D相机,该相机采用乐视三合一感摄像头ASTRA PRO,配有彩色相机、红外相机、红外投影仪及USB接口。图7为该相机深度图像。

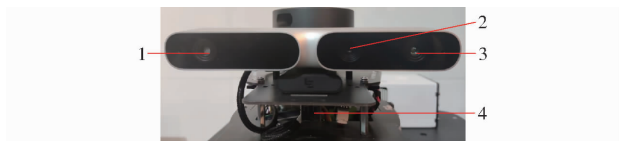


图6 RGB-D相机模型

Fig. 6 RGB-D camera model

1. 红外投影仪 2. 彩色相机 3. 红外相机 4. USB接口

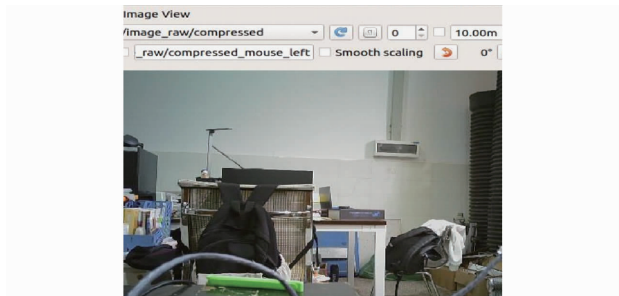


图7 RGB-D相机图像

Fig. 7 RGB-D camera image

2.2 相机位姿

通常将相机装在机器人上,跟随机器人同步运动。机器人运动时,相机在时刻 k 采集的外界场景图像记为 D_k ($k=1, 2, \dots, M$)^[42-43], M 为正整数。由于每一时刻的图像都对应特定的相机位姿,且相机与机器人同步运动,因此相机的位姿变换矩阵可由旋转和平移表示。将相机 $k-1$ 至 k 时刻的位姿

变换矩阵以特殊欧氏群表示为

$$T_{k,k-1} = \begin{bmatrix} \mathbf{R}_{k,k-1} & \mathbf{t}_{k,k-1} \\ 0 & 1 \end{bmatrix} \quad (\mathbf{R}_{k,k-1} \in \mathbf{R}^{3 \times 3}; \mathbf{t}_{k,k-1} \in \mathbf{R}^{3 \times 1}) \quad (5)$$

式中 $\mathbf{R}_{k,k-1}$ ——旋转矩阵 $\mathbf{t}_{k,k-1}$ ——平移向量

空间点 A 在相机坐标系 $oxyz$ 中的坐标可由其世界坐标在当前相机位姿下变换获得。设空间点 A 的世界坐标 $\mathbf{A}' = (x', y', z')$, 则点 A 在相机坐标系下的坐标 $\mathbf{A}$ 可由世界坐标 $\mathbf{A}'$ (易测量) 与相机位姿矩阵解算, 获得

$$\mathbf{A} = \mathbf{T}_{k,k-1} \mathbf{A}' \quad (6)$$

若已知相机 $k-1$ 至 k 时刻的位姿变换矩阵和空间点的世界坐标时, 便可根据式 (6) 解算像素坐标 $\mathbf{A}_2$ 。

设相机的位姿集为 $\mathbf{E}_n (n=0, 1, \dots, Q)$, Q 为正整数, 初始时刻的位姿为 $\mathbf{E}_0$, 则 k 时刻相机位姿可表示为

$$\mathbf{E}_k = \mathbf{E}_{k-1} \mathbf{T}_{k,k-1} = \mathbf{E}_0 \prod_{m=1}^k \mathbf{T}_{m,m-1} \quad (7)$$

3 传统视觉里程计

3.1 特征法

3.1.1 特征法概述

视觉里程计的传统算法主要有特征法和直接法, 其中特征法又可分为特征点法、特征线法和特征面法。特征点法和特征线法都需要提取、描述并匹配特征, 基本过程相似。面特征的提取则主要有基于图像和基于点云的方法, 其过程与点、线特征不尽相同, 面特征匹配大多根据面的法向量与坐标系原点到面的距离完成。本文重点对点特征提取、描述及匹配进行系统阐述。

特征点法中的特征点由关键点和描述子构成, 常见的特征点提取算法有 ORB 算法^[44-46]、SIFT (Scale invariant feature transform) 算法^[47-48]、SURF (Speed up robust features) 算法等^[49-51]。其中, ORB 特征的关键点为 Oriented FAST (Oriented features from accelerated segment test), 是在 FAST 基础上改进的角点, 其描述子为 BRIEF (Binary robust independent elementary feature) 描述子。ORB 特征提取首先需要检测 Oriented FAST, 并由 Oriented FAST 位置计算 BRIEF。其次是进行 BRIEF 匹配, 最简单的匹配方法为暴力匹配, 即测量相邻两时刻图像之间的 BRIEF 距离, 取距离最小的一对 BRIEF 作为匹配点。不同类型的 BRIEF 可取不同的距离度量, 浮点型 BRIEF 可由欧氏范数度量, 二进制 BRIEF 则由汉明距离度量。当完成关键点检测和特

征点匹配后, 就可大致估计相机的运动。图 8 为匹配点对的匹配结果, 匹配时选用汉明距离度量。由于未进行筛选, 匹配效果较差、误差较大、可信度较低。

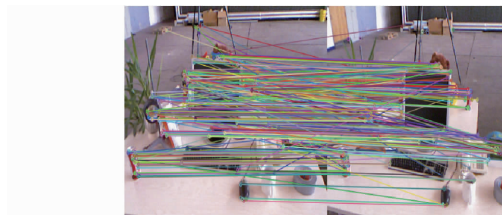


图 8 筛选前匹配点对匹配结果

Fig. 8 Matching point pair matching results before filtering

筛选前的匹配中存在大量误匹配, 需进一步筛选匹配点对。目前通常采用两倍最小 BRIEF 距离 (两倍最小距离) 作为误匹配的判断标准, 实现匹配点对的筛选。当 BRIEF 之间的距离大于 2 倍最小距离时, 判断该匹配点对为误匹配点对, 将其剔除; 当 BRIEF 之间的距离小于 2 倍最小距离时认为匹配无误, 保留该匹配点对。图 9 为筛选结果, 筛选后匹配效果较好、代表性强、准确率和可信度较高。

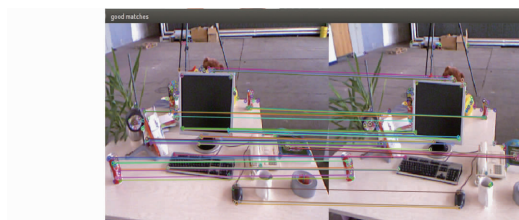


图 9 筛选后匹配点对匹配结果

Fig. 9 Matching point pair matching results after filtering

特征点法主要有以下两种方案:

(1) 当使用单目相机时, 像素坐标是平面坐标, 像素坐标的求解即为平面坐标求解。在求解两组平面点时, 对极几何本质矩阵的解算, 可通过经典八点法将本质矩阵视为线性方程。利用本质矩阵的内在特性, 将结果投影至本质矩阵的流线上^[52]。

(2) 当使用双目相机时, 可获得物体的距离, 通过空间点可直接估计相机的位置, 再根据迭代最近点来估计空间点的相机位姿。另外, 当相机采集的数据点一组为空间点, 另一组为平面点时, 解决方案也较多, 如直接线性变换法、P3P 法等。

3.1.2 相关算法

2007 年, 视觉 SLAM 技术取得突破性进展, KLEIN 等^[21]提出一种基于关键帧的单目视觉 SLAM 方法——PTAM。该方法第一次提出前端和后端, 界定了两者的区别, 提出基于关键帧的前端, 对后续基于特征法的里程计算法研究具有重要意义。

特征法里程计算法主要提供视觉传感器(相机)的大致运动情况,后端根据图像信息估计相机的位置。图像以矩阵的形式表达亮度和色彩信息,直接通过矩阵估计相机的运动情况,非常复杂,难以实现。为此,文献[53-54]提出对光线和动态物体敏感度低的特征点法。该算法具有一定的稳定性,但计算复杂度较高,很难保证视觉里程计的实时性。2011年,RUBLEE等^[55]提出一种基于BRIEF快速二进制描述符ORB,实验证明ORB比SIFT有效。原ORB特征检测算法由于阈值固定,在复杂光照场景中特征点数量急剧下降。为提高复杂光照场景特征检测算法的适应性,近年来国内外学者通过改变特征检测阈值提出诸多自适应ORB特征检测方法^[56-57]。YAO等^[58]针对复杂环境ORB特征分布不均问题,提出一种基于改进四叉树的ORB特征提取算法。算法根据采集图像的灰度不同,自适应选择阈值,实验结果表明该算法特征检测效果更稳定、快速。GUO等^[59]提出一种局部自适应阈值的ORB特征检测算法,测试结果显示,无论天气阴晴雨雪,该算法的稳定性均优于原ORB算法。YU等^[60]针对光照变化的复杂场景,提出一种改进自适应阈值的ORB-SLAM算法。在EuRoC数据集上的测试结果表明,该算法的稳定性与鲁棒性均高于原ORB算法。WU等^[61]提出一种适应不同光照条件的可变提取半径自适应ORB特征检测方法,实验结果表明该方法具有较好的鲁棒性、准确性及快速性,能满足不同光照场景的需求。自适应ORB特征检测算法在一定程度上避免了传统ORB算法因光照变化导致特征点数量急剧下降问题。与SIFT^[53]和SURF^[54]相比,自适应ORB特征检测算法的稳定性和实时性更好。

早期的地图构建大多依赖价格昂贵的激光扫描仪,随着RGB-D相机的出现,提供了一种可替代激光扫描仪的地图构建新方案。2013年,ENDRES等^[62]提出了一种新的SLAM系统,无需借助其它传感器或视觉里程计,仅通过RGB-D相机便可稳健地生成高精度的三维地图。实验结果表明,该方法在相机快速移动及特征稀缺的场景中都具有较好的鲁棒性。随着低成本、低重量RGB-D产品的推出,该方法逐渐应用于小型家用机器人和飞行机器人。但是,该方法无回环检测。为此MUR-ARTAL等^[36]提出了基于特征点法的ORB-SLAM系统,系统中加入回环检测来消除累计误差的影响。ORB-SLAM系统可在小型、大型室内外环境中实时运行,通过当时最主流的27个数据序列进行评估,结果表明该系统实现了前所未有的性能,是基于点特征提

取的集大成者。为提高ORB-SLAM系统的精度,2016年BELTER等^[63]引入实际误差,系统精度显著提高。陈劭等^[64]则通过引入随机回环检测,提出一种用于移动机器人的RGB-D SLAM算法,降低了机器人定位累积误差。2017年,MUR-ARTAL等^[65]提出扩展型ORB-SLAM2系统,系统框架如图10所示。该系统有地图重用、环路闭合和重新定位等功能,可用于单目、立体和RGB-D相机完整的同步定位与建图。ORB-SLAM2系统后端是单目和立体观测系统,可实现精确轨迹估计。该系统的定位模式简单,SLAM效果较好,应用场景广泛。

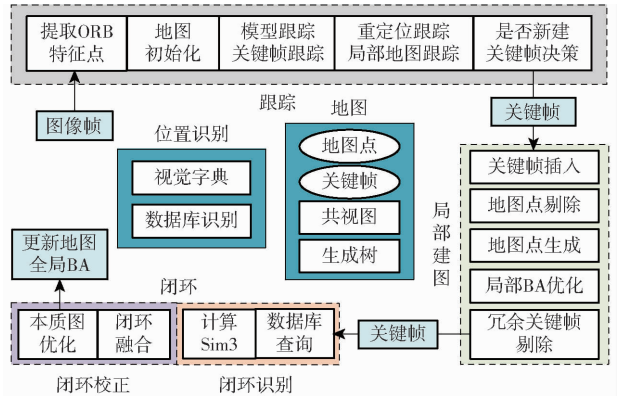


图 10 ORB-SLAM2 系统框图
Fig. 10 ORB-SLAM2 system block diagram

近年来,ORB-SLAM系列系统发展迅速。2019年,范永祥等^[66]利用装有RGB-D SLAM系统的手机实现了森林样地构建及样地属性计算。结果表明该系统对森林分类及样地属性估算精度较高。2021年CAMPOS等^[46]在ORB-SLAM2的基础上提出了一种全新的算法框架ORB-SLAM3,适用于鱼眼相机、针孔相机、RGB-D相机。相比ORB-SLAM和ORB-SLAM2,该算法具有更好的稳定性和精度。为适应复杂多变的环境,2022年LI等^[67]对连续采集的输入图像,将跟踪算法与目标检测相结合,改进ORB-SLAM2框架,提出一种目标跟踪和静态点检测相结合的视觉语义SLAM。董蕊芳等^[68]通过融合3D/2D特征、路标位置及视觉信息,提出一种改进型ORB-SLAM2系统,实现了机器人导航。森林公园中的实验结果表明该系统性能优于传统RGB-D SLAM。2023年,针对果园机器人定位与建图问题,丛佩超等^[69]通过优化FAST角点和描述子的阈值,利用图像金字塔法和高斯滤波剔除劣质ORB特征,提出一种改进型ORB-SLAM2算法。通过真实果园场景测试,完成了果园喷药机器人定位与建图,结果表明该算法的定位与轨迹精度均显著高于ORB-SLAM2。针对单一传感器建图精度、实时性较差等问题,刘庆运等^[70]提出了一种

融合激光雷达的增强型 ORB-SLAM2 算法, 结果表明融合算法建图精度较高, 可满足导航要求。

点特征提取的 SLAM 算法仅适用于特征明显、丰富的场景, 对特征不明显的场景一般提取场景的线、面特征^[71], 或者融合点线、点面特征。常见的线特征提取算法有 EDLines(Edge draw lines)算法^[72]、线二进制描述符算法 LBD (Line binary descriptor)^[73]及 LSD 提取算法^[74]等。

2014 年, LEE 等^[75]提出了一种线特征提取算法, 对于结构化场景的可靠性更强。2015 年, LU 等^[76]以 RGB-D 相机作传感器, 提出一种点线特征融合的视觉里程计系统, 该系统算法的鲁棒性较好, 与使用单一特征的算法相比, 相机运动估计量的不确定性更小, 实验结果显示该方法的平均平移误差比同类方法小 34.9%, 且适用恒定和变化的光照条件。但是, 当相机快速移动、旋转时, 特征法 SLAM 系统并不稳定。为此, 2018 年 WANG 等^[77]建立点、线投影模型并增加角度观测优化了线特征, 将运动估计模型作为相机的实际运动, 设计一种自动提取和匹配点、线特征的算法, 通过 EuRoC MAV 数据集验证该方法有效。针对低纹理环境、特征点较少场景, 2019 年, GOMEZ-OJEDA 等^[78]提出了一种适用于特征稀缺或分布不均场景的 PL-SLAM 系统。2020 年, LI 等^[79]利用弱曼哈顿假设, 基于直线和曲面法线估计相机的旋转。进一步根据点、线特征计算平移并优化姿态, 解决了低纹理场景单目 SLAM 效果较差的问题。2021 年, WANG 等^[80]提出了一种 LF-SLAM 系统。该系统通过预测和更新 3D 线段连续 2D 投影的线流, 并将 3D 线在时间维度上观测进行编码。对于模糊、特征被遮挡的场景, 该系统具有较好的 SLAM 效果。

由于室内空间限制且 GPS 不可用, 室内农业、畜牧业作业时, 无人机定位精度较低。2021 年, KRUL 等^[81]在温室大棚中利用单目相机进行 LSD-SLAM 和 ORB-SLAM 实验, 实验结果表明在室内农牧业方面 ORB-SLAM 算法的效果较好。针对复杂田间环境三维重建较为困难等问题, 2022 年, LI 等^[82]提出了一种基于点特征和线特征融合的 RGB-D SLAM 算法。作者优化点线特征关键帧, 并设计了一种点云滤波算法, 进一步基于融合的点线特征完成了三维场景重建。田间场景的实验结果表明该方法具有较高的鲁棒性。2023 年, LSLAM 等^[83]通过融合 LSD 线特征与 ORB 点特征, 提出了一种立体视觉 SLAM 算法, 在低光照农业数据集上进行测试, 结果表明所提方法在农业环境中具有较高的定位精度和建图精度。

3.1.3 算法对比

目前, 基于特征法里程计的 SLAM 方案较多且成熟。表 2 列出了近 10 年较为典型的特征法, 对比阐述了它们的优势和不足。

3.1.4 小结

虽然特征法里程计或 SLAM 算法研究取得较好的进展, 但仍存在以下问题:

(1) 对特征的依赖性强: 特征法的核心为环境特征提取, 对于环境纹理较差的场景, 尽管提出了一些改进方法, 但大多是通过增加算法复杂度或牺牲计算资源来实现。

(2) 易受光照变化和遮挡影响: 特征提取过程大多依赖于图像的外观, 若光照明暗变化较大或者相机视野被遮挡, 就会影响特征的提取和匹配, 从而影响定位和建图的精度。

(3) 固定场景假设: 特征法通常假设场景物体固定不动, 即物体的位置和姿态不发生变化。然而, 在现实中环境的物体可能会发生移动、旋转、变形等, 导致特征匹配失效, 从而使定位和建图失败。

3.2 直接法

3.2.1 直接法概述

特征点法仅使用了图像中的少数特征点, 特征提取不完整, 造成大量有用信息丢失^[89]。融合特征点线的 SLAM 算法, 仅适用于特征丰富、规则的场景。为解决特征法的不足, 提出了直接法。

图 11 为直接法简图, 设空间点 A_3 在世界坐标系中的坐标为 $A_3 = (x_w, y_w, z_w)$ 。

第 $k-1$ 时刻像素点 A_4 和第 k 时刻像素点 A_5 的表达式为

$$\begin{cases} A_4 = \begin{bmatrix} x_4 \\ y_4 \\ 1 \end{bmatrix} = \frac{BA_3}{h_1} \\ A_5 = \begin{bmatrix} x_5 \\ y_5 \\ 1 \end{bmatrix} = \frac{B(R_{k,k-1}A_3 + t_{k,k-1})}{h_2} \end{cases} \quad (8)$$

式中, h_1 为点 A_3 在相机 1 下的深度, h_2 为 A_3 在相机 2 下的深度。 $R_{k,k-1}$ 、 $t_{k,k-1}$ 为相机 2 相对于相机 1 在第 $k-1$ 至第 k 时刻的旋转矩阵和平移向量。

直接法源于稀疏光流算法 KLT, 由光流演变而来^[90-91]。特征法一般仅能构造系数地图, 而直接法不仅能获得稀疏地图, 还能建立半稠密、稠密地图。当采用部分图像点时, 将得到稀疏光流, 而使用全部图像点时, 则能得到稠密光流。其中稀疏光流以 Lucas-kanada 为代表, 稠密光流则以 Hord-schunck 为代表。直接法不涉及特征的匹配, 计算速

表 2 典型特征法
Tab.2 Typical feature method

算法	相机	特征点	特征线	发布时间	优势	不足
ORB-SLAM ^[36]	单目	✓		2015 年	室内外均可估计相机位置并构建地图；高效、鲁棒,适用复杂场景	相机标定要求较高,需要手动设置内参;算法的适用范围较窄
ORB-SLAM2 ^[65]	单目立体	✓		2017 年	支持单目、双目、RGB-D 相机;具有建图、跟踪、定位及回环检测等功能,效果较好	动态场景易出现漂移;对计算资源要求较高;低纹理区域,稳定性和精度不高
RTAB-MA ^[84]	立体 RGB-D	✓		2019 年	可进行多传感器融合、在线增量式学习;支持长时间运行;开源库,可修改、扩展	在扩展卡尔曼滤波中,受干扰的影响估计状态的精度降低;算法性能受环境影响较大
Fast ORB-SLAM ^[85]	RGB-D	✓		2020 年	高速;鲁棒性强;易于扩展	使用粗到细的策略,会丢失部分细节信息;低纹理场景效果不佳
ORB-SLAM3 ^[46]	单目立体	✓		2021 年	引入了语义特征,提高了建图和定位精度;具有多地图管理系统,不同场景的地图可叠加处理	算法复杂度较高,对计算资源要求较高;官方文档较少,不利于初学者使用
改进 ORB-SLAM2 ^[67]	RGB-D	✓		2022 年	结合目标跟踪和静态点检测算法,能精确捕捉动态场景中的物体	未探究算法的实时性和稳定性,系统整体性能的限制因素研究不足;仅考虑单一传感器数据输入
文献[75]		✓	✓	2014 年	提出一种基于线特征的 SLAM 方法,适用于光照、天气、季节变化等场景	不适合自然环境;需借助高精度地图和人工标注数据,建模难度和计算量大
文献[76]	RGB-D	✓	✓	2015 年	结合点和线特征,鲁棒性和稳定性更强;进行光照补偿和分割,解决了复杂环境中的遮挡和光照问题	可扩展性差;计算复杂度较高
PL-SLAM ^[86]	RGB-D	✓	✓	2017 年	稳健性强;定位精度高;实时性较好	不适合特征稀少场景;算法较复杂
文献[77]	立体	✓	✓	2018 年	鲁棒性强;精度高;可扩展性好	运行速度相对较慢;点线特征数量较少的复杂场景,定位精度较差
PL-VIO ^[87]	单目	✓	✓	2018 年	处理稀疏场景能力强;定位精度高	计算资源要求高;对相机内参敏感
PL-SLAM ^[78]	立体	✓	✓	2019 年	以双目相机提供几何信息,位置估计准确;实时性好,运行速度快	双摄像头间距离的影响较大,需进行相应的参数调整和校正
PL-VINS ^[88]	双目立体	✓	✓	2020 年	基于窗口的优化方法,避免了过度拟合噪声;预积分测量模型估计 IMU 的状态,提高了估计精度	对计算资源要求高;涉及 IMU 数据的预积分,系统需进行周期性重置或校准
文献[79]	单目	✓	✓	2020 年	算法简单高效,可实时定位和建图;使用 3D 结构约束大大降低漂移	在相机运动较快或多次反复穿越同一区域的情况下,仍会出现漂移
文献[80]	全景	✓	✓	2021 年	改进线特征描述;结合斜率因素与长度信息,稳定性好;超像素和平面分割提高线流检测和跟踪效果	复杂环境检测和跟踪较困难;计算资源需求大,需考虑线流影响;不适合嵌入式和移动设备

注:✓表示具有该特征。

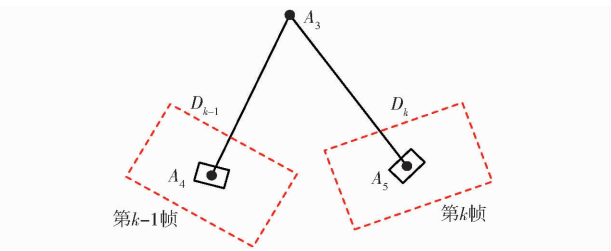


图 11 直接法简图
Fig. 11 Direct method diagram

度较快。在缺少特征或特征模糊的场景中,直接法明显优于特征法^[43,92]。

单层直接法来自稀疏的关键点,使用的像素点

为几百到上千,经拓展后可变成金字塔式多层直接法^[2]。图 12 为多层直接法追踪图,这种稀疏法避免了计算描述子,速度响应性较好,但只能完成稀疏计



图 12 直接法光流追踪
Fig. 12 Direct optical flow tracking

算和构建。半稠密直接法则仅对有梯度的像素进行记录,舍弃没有梯度或者梯度较小的像素。稠密直接法考虑图像中的所有像素,全部的像素均纳入计算范围,计算量非常大。

3.2.2 相关算法

2011 年,NEWCOMBE 等^[37]提出了 DTAM,是直接法的第一次应用。与特征法相比,直接法对快速运动场景下密集模型的跟踪性能较好,能有效地实现物理增强现实中的实时场景交互,但计算量较大、应用场景有限。为解决 DTAM 计算量较大的问题,2014 年,ENGEL 等^[93]第一次提出了一种用于单目相机的直接法,可完成半稠密地图构建。其中,最具代表性的 LSD-SLAM 算法允许构建大规模、环境一致的地图。为进一步扩大直接法的应用范围,2015 年,ENGEL 等^[94]又提出了一种用于立体相机的直接法。该算法可鲁棒地处理各帧之间亮度的剧烈变化,大大提高了算法性能。在 ENGEL 研究的基础上,CARUSO 等^[95]在全向鱼眼相机中采用直接法,可对视野超过 180° 的中央成像设备进行建模,还统一了全向鱼眼相机图像的对准公式,可在失真的图像上快速、准确地实现立体增强。

由于稠密、半稠密地图仍采用较多的像素点,计算量较大,且采集图像的区域大多为边界,稳定性较差。因此,2017 年,ENGEL 等^[96]提出了直接稀疏里程计(Direct sparse odometry, DSO)的视觉里程计算法,该算法是直接法的集大成者,完整地集成了光度校准,是一种新颖、高精度的稀疏视觉里程计方法。该方法不依赖于关键点的检测和描述子计算,可自然地从高梯度图像区域采集像素。在不同数据集上进行实验,结果表明该方法在各类现实环境中,跟踪精度和鲁棒性都显著优于直接法和间接法。尽管 DSO 的性能较好,但它不是一个完整的 SLAM 系统。为此 2018 年 GAO 等^[97]将 DSO 应用到具有闭环检测和位姿图优化的单目视觉 SLAM 系统中,利用具有足够强度梯度的任何图像像素,即使在没有特征的区域也具有较好的鲁棒性。在公开数据集上的评估结果表明,该算法的总体性能可媲美最先进的特征法。为进一步提升系统的鲁棒性、实时性,2021 年 MO 等^[98]提出了不依赖特征提取、匹配的 DSV-SLAM,在保持较强鲁棒性和实时性的条件下,扩展了单目直接法稀疏里程计。2022 年,CHEN 等^[99]提出了一种闭环直接法稀疏里程计,通过融合图像语义信息来降低算法计算量。

近年来,国内学者将直接法应用于闭环检测和位姿优化的研究工作也较多。2019 年,高成强等^[100]针对动态环境提出了一种融合运动检测算法

的半直接法,实验结果表明该算法可有效提高系统的鲁棒性和实时性。2021 年,李伟等^[101]提出一种改进 M-ORB(Mature-ORB)的视觉 SLAM 直接-闭环检测算法,实验结果表明该算法的准确性和实时性较好。直接法里程计应用愈加广泛,2021 年,为提高农业环境中自动驾驶汽车或机器人的避障能力,SHU 等^[102]提出了一种将稀疏单目直接法与多视图立体重建算法相结合的 SLAM 算法,评估结果表明该算法可完成各种农业任务的重建工作。为解决非结构化农业环境中视觉 SLAM 效果较差的问题,2021 年 YU 等^[103]通过优化 Stereo-DSO,设计一种立体直接视觉里程计系统。由于农业机器人通常是沿作物运动,因此作者在滑动窗口中保留了一个较远处的关键帧。在 Thrive 数据集和 Rosario 数据集上对算法进行测评,结果均证明算法有效。2023 年,针对农业环境中光照变化和动态问题,YU 等^[104]提出一种用于农业场景的立体直接法里程计,通过公共农业数据集测试,结果表明该算法在复杂农业环境中性能良好。

3.2.3 小结

虽然直接法在某些方面优于特征法,但仍然存在以下不足或需进一步提高之处:

(1) 灰度不变假设并不容易满足,当光照变化,算法的有效性将会降低。

(2) 要求相机的运动较慢或较高的采样频率,不适合大尺度运动场景。

(3) 在实时性要求的场景中,需要高速运算来完成大量的优化和重投影操作,现有的计算机性能无法满足,需要更高性能的硬件支持。

(4) 直接法具有较强的鲁棒性和适应性,但精度方面有待提升,例如对光照和遮挡等场景变化的适应性不足,以及对周围环境的边缘、角等细节特征的识别和处理能力有待提升。

(5) 长时间运行易导致误差积累和漂移等问题,需更加完善的算法设计和优化策略,确保系统的稳定性和精度。

(6) 直接法需要大量的数据进行训练和测试,但当前的数据集仅针对一些特定的需求,不够全面、充分。

4 基于深度学习的视觉里程计

传统视觉里程计以几何模型的方式表示,具有计算量较大、适用场景有限、要求光度不变等缺点,复杂外界场景的适应性较差。

2006 年以来,深度学习在图像分割和目标检测方面的应用日益成熟。越来越多的专家、学者将深

度学习应用于视觉里程计中。根据深度学习有、无监督,可将视觉里程计分为有监督和无监督学习的视觉里程计。

4.1 有监督学习里程计

KENDALL 等^[105]第一次提出一个单目相机的 6 自由度实时定位系统,该系统的卷积神经网络(Convolutional neural networks, CNN)以端到端的方式,从单一 RGB 图像确定 6 自由度相机的姿态,无需额外的工作或图形优化。图 13 为 CNN 网络结构,该系统可在室内外实时运行,每帧耗时 5 ms。在大型户外场景中,该系统的精度约为 2 m 和 3°,在室内的精度约为 0.5 m 和 5°。

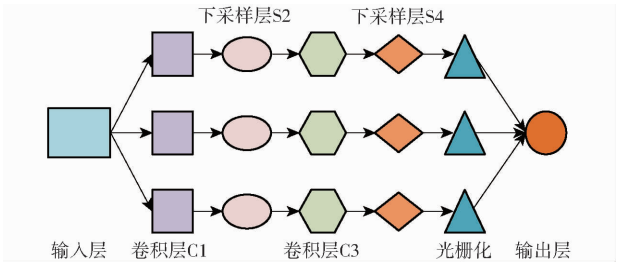


图 13 CNN 网络结构简图
Fig. 13 CNN network structure diagram

虽然单目里程计技术性能较好,但设计过程复杂。为此,2017 年 WANG 等^[106]提出深度视觉里程计,利用深度递归卷积神经网络设计一个单目视觉里程计端到端的框架,深度学习效果较好,弥补了传统里程计的不足。但深度卷积神经网络是以均方误差最小作期望值进行训练,存在收敛慢和局部最优解较差等问题。为此,2018 年 FU 等^[107]引入一种间距逐渐增加的深度离散化策略,将深度网络学习化为有序回归问题。

单目视图的深度估计忽略了确定深度信息的重要特征——运动。为此,2019 年 WANG 等^[108]提出一种学习多视图密集深度图的里程计估计方法。该方法利用多视图图像重投影和前后向流的一致性损失进行训练,尽管在精度和速度方面效果较好,但与传统里程计相比仍存在差距。因此可见,尽管深度神经网络在视觉 SLAM 中的应用取得一定进展,但仍无法取代传统方法。为增强传统视觉 SLAM 系统,BRUNO 等^[109]以深度学习描述环境特征,并结合几何法构建一个名为 LIFT-SLAM 的视觉 SLAM 系统,其性能与当前最先进的视觉 SLAM 性能相当。

目前,有监督深度学习里程计技术在农业领域应用愈加广泛。2021 年,AGUIAR 等^[110]提出一种有监督学习葡萄树干特征检测算法。实验结果表明该算法可完成葡萄树干特征的检测工作,其检测精度高达 84.16%。2023 年,RAUF 等^[111]通过设计卷

积、循环神经网络,提出了一个有监督学习的里程计算法。通过棉田数据集训练网络,并利用数据集对训练后的网络进行测试,测试结果表明,算法效果符合预期结果。农业场景中不同田间条件对机器人自主导航的影响不同,导致作物检测效果不尽相同。为使机器人在不同田间场景均具有较好的适应性,2023 年,DE SILVA 等^[112]提出了一种 U-Net 深度学习网络。通过有监督学习方式搭建网络并训练样本集,训练后再进行田间条件测试,测试结果显示算法的作物检测性能较好。

有监督学习 SLAM 算法存在以下不足或需进一步改进之处:

- (1) 数据依赖性强:有监督学习需大量有精确位姿信息的高质量数据集进行训练,当使用新传感器、新设备或新场景时,需重新训练。
- (2) 局限性:训练过程中只是对已知环境进行学习和优化,忽略未知环境中的复杂情形,因此在处理未知环境或者不常见场景时表现欠佳。
- (3) 泛化能力不足:只能生成与训练数据集相似的结果,难以适应不同的环境和场景变化,泛化能力不足。

- (4) 实时性不足:算力有限,基于有监督学习更倾向于离线处理,在实时性场景下无法满足需求。

4.2 无监督学习里程计

当前大多数深度卷积神经网络需利用大量手动标记的数据来训练,耗时较长,无法保证 SLAM 算法的实时性。

2016 年,GARG 等^[113]第一次提出用于单目视图深度预测的无监督学习深度卷积网络。网络训练所用数据无需手动注释,也无需校准深度传感器,在 KITTI 数据集上训练后的性能与当时最先进的有监督学方法相当。针对在非结构化视频序列中 3D 几何和相机估计等问题,2017 年 ZHOU 等^[114]提出一种无监督学习框架,直接估计输入集的相机姿态参数,使用单目视图推断参考视图的深度。2018 年,LI 等^[115]提出一种新的单目视觉里程计,通过深度神经网络估计单目相机的姿态及视野深度,系统如图 14 所示。实验结果表明,该系统在姿态精度方面表现良好。2020 年,LI 等^[116]提出一种无监督学习深度 SLAM 系统,以立体图像代替数据集的标注。

上述研究更多考虑算法的正确性和精确性,而忽略了算法的效率。为此,2021 年 LI 等^[117]提出一种多特征提取网络 and 自监督学习的实时视觉 SLAM 系统,设计一个简化的卷积神经网络用于特征点检测和描述,取代传统的几何方法,尽管效率较高,但准确性有所降低。为此,2022 年 AN 等^[118]利用多

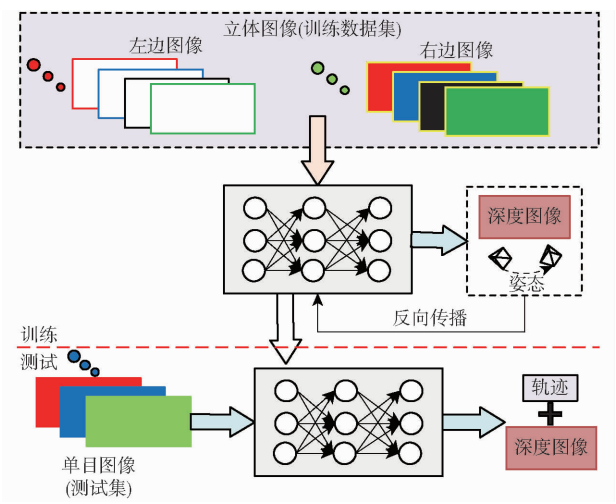


图 14 深度神经网络系统
Fig. 14 Deep neural network system

传感器互补的特点,将相机和激光雷达数据进行融合,提出一种多传感器融合的无监督 SLAM 方法,兼顾了算法的效率和准确性。

虽然有监督学习对田间杂草检测分类效果较好,但创建大型农业数据集需要耗费大量时间。2018 年,BAH 等^[119]提出一种无监督学习的新颖方法,结果显示,菠菜田杂草检测误差为 1.5%,豆田杂草检测误差也不超过 6%。2021 年,HUANG 等^[120]设计一种端到端学习的苹果园农业机器人跟踪系统。机器人可自动行驶并采集数据,无需手动

标记训练样本。测试结果显示,该视觉导航系统能针对不同场景引导机器人调整位姿,并成功穿越果园。2023 年,吴雄伟等^[121]针对温室机器人作业时存在尺度不确定问题,提出一种无监督的视觉里程计方法。在种植番茄的温室场景开展了实验研究,实验结果显示,无监督的里程计精度较高。

目前无监督学习仍存在以下不足或需进一步改进之处:

- (1) 精度问题:通过自学习构建地图并给机器人定位,需进一步提高算法的精度。
- (2) 数据要求高:需有大规模、高质量的数据集,否则很难获得精确的结果。
- (3) 训练时间长:与其他算法相比,无监督学习需更长时间的训练才能达到最佳性能。
- (4) 实践应用基础薄弱:虽然无监督学习在一些领域应用较成功,但仍存在诸多问题。例如,如何处理动态场景、建立精确地图以及实时定位等问题。

4.3 典型里程计对比

特征法里程计主要提取环境障碍物的点、线、面等特征;直接法无需特征的提取和匹配。深度学习方法对环境障碍物进行分类训练,训练后以类进行识别和划分。表 3 列举了特征法、直接法及深度学习法的优缺点和应用场景。

表 3 典型里程计方法对比
Tab.3 Comparison of typical odometer methods

方法	有无监督	优点	缺点	应用场景
特征法 SLAM		较为成熟;算法稳定性高;对光照不敏感,动态场景仍有较好的适应性	需提取场景特征,较直接法耗时,弱纹理场景的鲁棒性低;无法完成半稠密、稠密地图的重建	特征明显、纹理丰富的规则场景
直接法 SLAM		无需匹配和计算描述子,速度快;对纹理稀疏场景仍有较好鲁棒性;易完成稠密建图	基于灰度不变假设,易受光照变化影响	光照不变的场景
深度学习 SLAM	有监督	网络结构简单、效率高	需提前获取大量已标注的数据集	较复杂的场景
	无监督	避免大量数据集的获取和标注工作	网络的结构较复杂、效率低	较复杂的场景

5 常用数据集性能比较

目前已有较多评价算法性能的数据集,本文取几种常见的数据集对其性能进行分析。

5.1 KITTI 数据集

2013 年,德国卡尔斯鲁厄理工学院、马克斯·普朗克智能科学研究所和美国丰田技术研究所提出一个从大众车上捕获的新数据集,即 KITTI 数据集。该数据集主要用于移动机器人和自动驾驶研究领域^[122],可用于测评立体图像、光流、视觉测距、3D 物体检测和 3D 跟踪等技术的性能,具有较强的适用性。

数据集是通过各种传感器在频率 10 ~ 100 Hz 下连续记录 6 h 的交通场景获取的。交通场景主要为农村地区的高速公路,具有静态、动态特征点的市中心场景。测量传感器为高分辨率彩色和灰度立体相机、Velodyne 3D 激光扫描仪和高精度 GPS/IMU 惯性导航系统等^[122]。数据集主要包括城市区、住宅区、马路、校园及有行人的场景。

国内外学者提出的里程计算法大多在 KITTI 数据集上得到了较好验证。李鹏等^[123]为解决里程计鲁棒性差、漂移大等问题提出一种融合几何位姿与深度学习的视觉里程计。在 KITTI 数据集 00 至 10 序列进行测评,其最大平移误差为 5.146%,最大旋

转误差为 0.499° ,最大绝对轨迹误差为 22.67 ,最大相对位置误差为 0.14 m ,最大相对姿态误差为 0.077° 。ALMALIOGLU 等^[124]利用对抗性训练和自适应视觉-惯性传感器相融合提出一种自监督学习的视觉惯性里程计和深度图的恢复方法。利用 KITTI 数据集进行测评,最大位置误差为 1.95% ,最大姿态误差为 1.30° ,最大绝对轨迹误差为 0.009 m 。SHINDE 等^[125]提出一种端到端的多模态单目视觉惯性里程计,经 KITTI 数据集验证,平移误差为 12.47% ,转动误差为 3.79% 。

5.2 EuRoC 数据集

2016 年,EuRoC 数据集由苏黎世联邦理工学院公布^[126]。数据集拍摄相机的基线长约 11 cm ,拍摄频率为 20 Hz ,共获得 11 个双目序列。根据 MAV 速度、光照及纹理场景,11 个双目序列分为 3 种。

数据集包含 2 个不同的场景,即 2 批数据集,主要包括同步立体图像、IMU 测量和精确的地面实况等。其中,第 1 批在工业环境中获得,是由激光跟踪系统获取的地面实况,第 2 批数据集由运动捕捉系统在室内记录获得。图 15a 为第 1 批数据集获取场景,第 2 批数据获取场景如图 15b 所示。

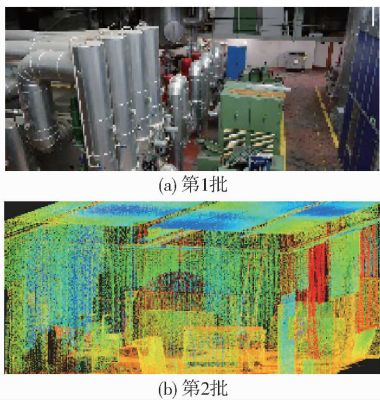


图 15 获取数据集场景
Fig. 15 Obtaining dataset scenarios

EuRoC 数据集常用于验证深度学习视觉惯性里程计的有效性。ASLAN 等^[127]利用 Savitzky Golay 技术去噪,同时对两帧间的惯性数据归一化处理,将数值转换为图像,在 EuRoC 数据集上验证,其绝对轨迹误差为 0.05 m 。ALMALIOGLU 等^[124]提出一种自监督学习的视觉惯性里程计和深度图像恢复的方法,在 EuRoC 数据集上与当时最先进的视觉惯性里程计做比较,证明该算法的性能较好,每 0.3 m ,算法平移误差为 25% ,每 15 ms 帧偏差为 5.43% 。余洪山等^[128]提出一种改进 Super Point 网络的单目视觉惯性系统 (Visual inertial slam, VINS),即 CNN-VINS,提升了恶劣环境下 VINS 系统的鲁棒性。与 VINS-Mono 系统相比,CNN-VINS 系统在光照剧

烈变化的恶劣场景中,定位精度提升 15% ;在光照变化缓慢的场景,绝对轨迹误差均值保持在 $0.067\sim 0.069\text{ m}$ 之间。

5.3 TUM RGB-D 数据集

TUM RGB-D 数据集由德国慕尼黑工业大学计算机科学系提出^[129],是通过 Kinect 传感器记录办公和工厂环境,涵盖场景丰富。数据集囊括了 39 个不同场景的序列,主要包括手持 SLAM、机器人 SLAM、动态物体、三维重建、验证集和标定文件等,是一种评估 RGB-D SLAM 系统的新基准。

数据集中彩色、深度图像的帧速率均在 30 Hz 以下,传感器分辨率为 $640\text{ 像素}\times 480\text{ 像素}$ 。另外,TUM 数据集还带有标准轨迹,非常适合做研究使用。

TUM RGB-D 数据集还含有动态数据集,动态 SLAM 研究通常使用其中 Dynamic Objects 和 Validation Files 对应的数据集。

验证动态场景里程计算法的效果时,TUM 数据集具有独特优势。王恩宝等^[130]针对动态场景,提出一种深度学习视觉里程计,通过 TUM 数据集验证,与传统 ORB-SLAM2 系统相比,绝对轨迹误差和相对位姿误差大幅降低,标准误差降低高达 36.6% 。李博等^[131]结合 ORB-SLAM2 和 YOLO v4 提出面向室内动态环境的视觉里程计,通过 TUM 数据集验证,绝对轨迹误差和相对位姿误差相比传统 ORB-SLAM2 系统降低 90% 以上。

5.4 Cityscapes 数据集

没有数据集能充分表征复杂的现实世界城市场景。为解决这个问题,2015 年德国奔驰公司引入 Cityscapes 数据集^[132]。Cityscapes 由 50 个不同城市街道上录制的立体声视频序列组成,每幅图像像素均为 $1\,024\text{ 像素}\times 2\,038\text{ 像素}$,其中 5 000 幅图像为高质量像素级别的注释,20 000 幅图像为粗略注释。Cityscapes 数据集量大、注释丰富,场景的可变性和复杂性都超越其他数据集。图 16 为 Cityscapes 数据集的语义图像序列。

Cityscapes 数据集在语义 SLAM 方面应用较广。



图 16 语义图像序列
Fig. 16 Semantic image sequence

LOU 等^[133]融合语义图像与低分辨率 3D 激光雷达点云,生成稠密语义深度图,提出一种融合激光雷达和单目视觉的改进 SLAM 和语义重建方法。通过 Cityscapes 数据集验证,所提方法优于动态场景下的 ORB-SLAM2 系统。为解决动态场景下的漂移和误差累积问题,MA 等^[134]提出一种直接法稀疏里程计,在 Cityscapes 数据集上进行验证,效果优于传统的直接法稀疏里程计。LI 等^[135]提出一种融合深度神经网络的单目半稠密里程计,在 Cityscapes 数据集上进行实验验证,语义分割过程仅需 1 024 ms 左右。

5.5 ETHI 数据集

为测试所提方法在真实数据上的照明不变性,2017 年瑞士苏黎世联邦理工大学计算机学院研究人员记录了适用于复杂光照场景评估的 ETHI 数据集^[136]。该数据集是由研究人员利用第一代 Kinect 相机在室内复杂光照场景记录的 RGB-D 序列,序列具有由 VICON 系统提供的地面实况,可用以验证视觉里程计方法在复杂光照场景中的定位性能。ETHI 数据集由 3 个序列构成,按照图像变化过程包括全局、局部和闪存 3 部分。图 17 为部分 ETHI 数据集序列。

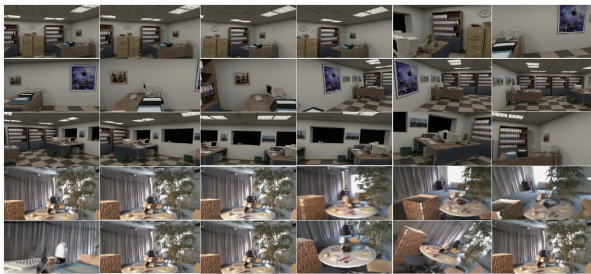


图 17 部分 ETHI 数据集序列

Fig. 17 Partial ETHI dataset sequence

ETHI 数据集主要用于室内复杂光照场景下视觉里程计算法定位性能的评估工作。复杂多变的光照使得计算机视觉研究变得困难,2019 年,HUANG 等^[137]提出了一种可同时提取两种图像特征的多特征提取算法。作者在跟踪线程之前增加了图像预处理步骤,以应对极端黑暗的条件。通过 ETHI 数据集测评结果显示,结合多种特征的方法能够在不影响实时性的前提下,提高算法的鲁棒性。2022 年,SUN 等^[138]基于点和线特征提出了一种适用于光照变化、低纹理和遮挡复杂场景的鲁棒 RGB-D 视觉里程计。该算法首先在视觉里程计系统中加入 EDLine 线特征。其次,使用一系列点描述符描述 EDLine 线。此外,还利用深度测量信息校正姿态误差,构建了一种新的双线重投影误差优化模型。最后,将线特征集成至 ORB-SLAM2 系统以改进机器

人姿态估计。ETHI 数据集测试结果显示,该算法性能较好。为进一步提高视觉 SLAM 在复杂光照场景的性能,2022 年,YU 等^[60]提出一种基于自适应快速阈值和图像增强的改进 ORB-SLAM 算法。通过一种改进的截断自适应伽马校正与反锐化掩模相结合,减小了由于照明差异引起的的影响。为验证算法性能,通过 ETHI 数据进行测评,结果表明,该算法能在复杂光照环境下获得较高的平均定位精度。

5.6 TartanAir 数据集

TartanAir 数据集由卡内基梅隆大学的机器人学院和英国剑桥大学的机器人学实验室共同创建^[139]。该数据集首次发布于 2020 年,是一个用于视觉 SLAM 和机器学习研究的数据集。TartanAir 数据集包含了丰富的室内、室外场景的 RGB-D 图像及精确的真实运动信息。这些数据可用于开发和 SLAM 算法评估、机器学习模型及视觉导航系统验证。数据集的独特之处在于它提供了高度复杂化和多样化的环境,旨在挑战深度学习、自主导航、场景理解等各种机器学习任务,是一种可用于探索视觉 SLAM 极限的数据集。图 18 为室内外各种场景下的部分 TartanAir 数据集序列。



图 18 部分 TartanAir 数据集序列

Fig. 18 Partial TartanAir dataset sequence

TartanAir 数据集主要用于视觉和机器学习领域的研究,特别是无人机和机器人领域。该数据集包含大量的图像和深度信息,适用于自主导航、视觉里程计、物体识别、语义分割、机器学习和深度学习算法的开发等场景或任务。2021 年,GU 等^[140]基于深度二进制的局部描述符设计了 DBLD-SLAM 单目 SLAM 系统。该算法通过 TartanAir 数据集进行验证,结果显示 DBLD-SLAM 系统的性能较好。在纯旋转、运动模糊和纹理差等极端条件时,ORB-SLAM3 系统的姿态估计效果不佳。2023 年,LIU 等^[141]提出了一种鲁棒补偿方法,该方法使用 DROID-VO 生成丢失的轨迹并将其准确地插入 ORB-SLAM3 中。在 TartanAir 数据集上测试结果表明,所提方法提高了 Hizher 的覆盖率,能够适应更多场景。2023 年,WOFK 等^[142]通过集成单目深

度估计和视觉惯性里程计,提出了一个集成的模块化视觉惯性深度估计系统,可与各种单目深度估计模型兼容。在 TartanAir 数据集上的测试结果表明,该系统仅需提取 50 个稀疏深度点,在低密度时具有更强的竞争力。

5.7 Bonn 数据集

Bonn 数据集由波恩大学于 2019 年发布,是一款在现实世界中收集的高动态性 RGB - D 数据集^[143]。该数据集通过华硕 Xtion Pro LIVE 获取 RGB - D 图像,使用 Optitrack Prime13 动作捕捉系统获取摄像机在地面的真实轨迹。Bonn 数据集含有 24 个高度动态场景,人们在其中执行不同的任务,例如搬运箱子和玩气球。动态场景占据该数据集的大部分,这对 SLAM 算法的鲁棒性提出了极大考验。图 19 为部分 Bonn 数据集的场景序列。



图 19 部分 Bonn 数据集序列

Fig. 19 Part of Bonn dataset sequence

Bonn 数据集主要用于动态场景下 SLAM 算法的评估,尤其适用于室内动态场下的评估工作。2022 年,YAN 等^[144]为提高 SLAM 系统在动态环境下的鲁棒性和效率,提出一种基于几何信息和语义信息组合的动态 RGBD SLAM 算法。通过公共的 Bonn 数据集和 TUM 数据集进行多次评估,结果表明动态 RGBD SLAM 比当时最先进的动态 RGB - D SLAM 系统具有更高的鲁棒性和速度。为提高基于点特征视觉 SLAM 在室内动态场景的适应性,2022 年,WANG 等^[145]提出一种线特征和面特征融合的 RGB - D SLAM 算法。通过 Bonn 数据集进行实验评估,结果表明该算法准确性和鲁棒性均得到了提高。2023 年,CHENG 等^[146]基于 ORB - SLAM2 框架提出一个适用于动态场景的 SG - SLAM 系统。在 Bonn 数据集、TUM 数据集及 OpenLORIS - Scen 数据集进行实验测评,结果表明该算法不仅是动态场景中最实时、最准确、最强大的系统之一,还可以创建语义度量图。2023 年,LU 等^[147]提出一种室内动态元素剔除并对静态特征进行三维重建的方法。通过 Bonn 数据集和 TUM 数据集对所提方法进行测试,根据三维重建结果验证了所提方法的正确性。

5.8 小结

迄今为止,已出现较多数据集用于验证深度学

习里程计算法的性能,上述 7 类数据集最为典型。KITTI、Cityscapes 数据集主要适用场景为室外,可用于自动驾驶中计算机技术性能评估。EuRoC、TUM 数据集主要应用场景为室内,可用于三维重建技术性能评估,基本信息如表 4 所示。

表 4 7 种数据集信息

Tab. 4 Seven types of dataset information

数据集	相机	场景	网址
KITTI	立体	室外	http://www.cvlibs.net/datasets/kitti/eval_object.php https://projects.asl.ethz.ch/datasets/doku.php?id=
EuRoC	双目	室内	https://vision.in.tum.de/data/datasets/rgbd-dataset/download
TUM	深度	室内	https://www.cityscapes-dataset.com/
Cityscapes	立体	室外	http://cvg.ethz.ch/research/illumination-change-robust-dslam
ETHI	RGB - D	室内	http://theairlab.org/tartanair-dataset/
TartanAir	单目、立体	室外	http://www.ipb.uni-bonn.de/data/rgbd-dynamic-dataset
Bonn	深度	室内	

6 面临的问题与发展趋势

6.1 面临的问题

尽管 SLAM 技术已有近 40 年历史,里程计相关的研究也取得了大量成果,但现有技术仍存在诸多亟待解决的问题。综上所述,当前视觉里程计面临的问题如下:

(1) 精度问题:视觉里程计是通过图像序列计算相机的运动轨迹和场景的深度信息。当前视觉里程计技术的精度仍不够高,尤其对于复杂场景或光照变化较大的情况,往往难以准确获取相机的运动轨迹。

(2) 鲁棒性问题:鲁棒性是指算法对噪声、遮挡、姿态变化等因素的稳定性。在实际应用中,场景的复杂度和多样性都会影响视觉里程计的鲁棒性。

(3) 实时性问题:视觉里程计需及时处理图像,输出相机运动轨迹,因此实时性是重要考量因素,当前算法的速率需大幅提高。

(4) 数据集问题:视觉里程计基于数据驱动,但可用数据集的数量和质量有限,限制了算法的精度和鲁棒性。

(5) 多模态问题:视觉里程计常用单目或双目相机获得图像信息,而激光雷达、GPS 等传感器,可提供更多的场景深度信息,如何将多种传感器数据

进行有效融合,是当前视觉里程计急需解决的问题。

(6) 数据量不足:深度学习等基于数据驱动的方法,需要大量的数据进行训练,但目前可用的数据集有限。

(7) 噪声和漂移问题:在实际场景中,由于光照变化、物体运动等因素的影响,图像中的噪声和漂移很难避免,直接影响视觉里程计的精度和鲁棒性。

6.2 发展趋势

里程计是机器人环境感知的第一步,感知效果主要取决于传感器。随着 SLAM 技术的不断发展,未来里程计将会向以下几个方向发展:

(1)更加智能化、集成化的新型传感器

从单目、双目到 RGB-D 相机,相机的主动性越来越强、智能化程度越来越高,但仍存在诸多局限性。例如,单目相机无法获得深度信息,需要重新设计算法才能完成深度信息的提取;双目相机能够获得深度信息,但计算量大,算法效率较低;RGB-D 相机主动性较强、能够获取丰富的环境信息,但不适用于室外场景,对深色物体的识别率较低、成本较高。

目前,SLAM 技术多用于完成机器人的环境感知。若设计一种更加智能化、集多种传感器功能于一体的新型传感器,则可减轻机器人负重,使机器人具有更大的负载空间和更强负载能力,对机器人的运动控制具有重大的意义。因此,智能化、集成化的传感器技术是未来的一个发展方向。

(2)小型化新型传感器

虽然多传感器融合 SLAM 技术能够弥补单一传感器的不足,但不适用于所有对象。对于小型机器人,若安装传感器过多,将会增加机器人的负荷,不利于机器人作业。因此,小型化的新型传感器将是未来 SLAM 技术的必经之路。

(3)里程计技术与无监督学习的进一步融合

深度学习避免了复杂的计算过程,实现了端到端的 SLAM 技术。随计算机技术的发展,深度学习与里程计技术进一步融合,设计性能更强的里程计,对里程计技术发展具有重要意义。

(4)语义表达技术的进一步发展

当前语义表达仅以低级的几何方式表达,不具备人类的高级语义表达。例如,Bag-wods 方法存在丧失空间位置和语义表达准确率较低等问题。

进一步发展语义表达技术,当机器人进入未知环境时,可快速、准确理解区域的要素,提高机器人的智能化程度和自主能力。因此,寻求更高级的语义表达方式,将其与里程计融合,也将是里程计的发展方向之一。

(5)集群机器人协同的里程计技术

随着 SLAM 技术向大尺度复杂场景“进军”,单一机器人系统能力受限,并不能适应大尺度复杂场景的需求。相反,集群机器人系统由多台简单机器人系统协调而成,通过机器人间有效的协同合作可以很好地适用于大尺度复杂场景。

集群机器人系统涉及多个机器人传感器的融合,需进一步提高多传感器融合技术的稳定性和鲁棒性。

在集群机器人协作建图中,各机器人存在信息的传递与交流,因此,需进一步开发分布式的里程计技术,在协作过程中共享地图信息,实现机器人间的协作和通信,从而建立一致有效的地图,并完成有效的数据共享。另外,当机器人数量较多时,机器人间存在相互干扰问题,系统性能下降,因此,需合理分配各机器人的资源。例如,在协作建图时,里程计算法设计需考虑协调各机器人资源,使系统规则地实现 SLAM 过程。

综上,能实现多传感器融合、具有分布式功能并能协调分配各机器人资源的里程计技术将是其发展方向之一。

参 考 文 献

[1] SMITH R C, CHEESEMAM P. On the repressentation and estimation of spatial uncertainty [J]. The International Journal of Robotics Research, 1986, 5(4): 56-68.

[2] 高翔,张涛,刘毅,等.视觉 SLAM 十四讲从理论到实践[M].北京:电子工业出版社,2019.

[3] 周彦,李雅芳,王冬丽,等.视觉同时定位与地图创建综述[J].智能系统学报,2018,13(1):97-106.

[3] ZHOU Yan, LI Yafang, WANG Dongli, et al. A survey of VSLAM [J]. CAAI Transactions on Intelligent Systems, 2018, 13(1): 97-106. (in Chinese)

[4] SMITH R, SELF M, CHEESEMAM P. Estimating uncertain spatial relationships in robotics[J]. Machine Intelligence & Pattern Recognition, 1988, 4(5): 435-461.

[5] LEONARD J J, DURRANT-WHYTE H F. Mobile robot localization by tracking geometric beacons [J]. IEEE Transactions on Robotics & Automation, 1991, 7(3): 376-382.

[6] CADENA C, CARLONE L, CARRILLO H, et al. Past, present, and future of simultaneous localization and mapping: toward the robust perception age[J]. IEEE Transactions on Robotics, 2016, 32(6): 1309-1332.

[7] KAZEROUNI I A, FITZGERALD L, DOOLY G, et al. A survey of state-of-the-art on visual SLAM[J]. Expert Systems with Applications, 2022, 205(11): 1-17.

- [8] JIA G, LI X, ZHANG D, et al. Visual-SLAM classical framework and key techniques: a review[J]. *Sensors*, 2022, 22(12): 1–29.
- [9] SERVIÈRES M, RENAUDIN V, DUPUIS A, et al. Visual and visual-inertial SLAM: state of the art, classification, and experimental benchmarking[J]. *Journal of Sensors*, 2021, 2021: 1–26.
- [10] GUPTA A, FERNANDO X. Simultaneous localization and mapping (SLAM) and data fusion in unmanned aerial vehicles: recent advances and challenges[J]. *Drones*, 2022, 6(4): 1–35.
- [11] SHALAL N, LOW T, MCCARTHY C, et al. A review of autonomous navigation systems in agricultural environments[J]. *SEAg 2013: Innovative Agricultural Technologies For a Sustainable Future*, 2013(9): 1–16.
- [12] AGUIAR A S, DOS SANTOS F N, CUNHA J B, et al. Localization and mapping for robots in agriculture and forestry: a survey[J]. *Robotics*, 2020, 9(4): 97.
- [13] FASIOLO D T, SCALERA L, MASET E, et al. Towards autonomous mapping in agriculture: a review of supportive technologies for ground robotics[J]. *Robotics and Autonomous Systems*, 2023, 169(11): 1–34.
- [14] GAO B, LANG H, REN J. Stereo visual slam for autonomous vehicles: a review[C]// *International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2020: 1316–1322.
- [15] BAI Y, ZHANG B, XU N, et al. Vision-based navigation and guidance for agricultural autonomous vehicles and robots: a review[J]. *Computers and Electronics in Agriculture*, 2023, 205: 107584.
- [16] DING H, ZHANG B, ZHOU J, et al. Recent developments and applications of simultaneous localization and mapping in agriculture[J]. *Journal of Field Robotics*, 2022, 39(6): 956–983.
- [17] 陈世浪, 吴俊君. 基于 RGB-D 相机的 SLAM 技术研究综述[J]. *计算机工程与应用*, 2019, 55(7): 30–39, 126.
CHEN Shilang, WU Junjun. RGB-D SLAM: a survey[J]. *Computer Engineering and Applications*, 2019, 55(7): 30–39, 126. (in Chinese)
- [18] 丁文东, 徐德, 刘希龙, 等. 移动机器人视觉里程计综述[J]. *自动化学报*, 2018, 44(3): 385–400.
DING Wendong, XU De, LIU Xilong, et al. Review on visual odometry for mobile robots[J]. *Acta Automatica Sinica*, 2018, 44(3): 385–400. (in Chinese)
- [19] 祝朝政, 何明, 杨晟, 等. 单目视觉里程计研究综述[J]. *计算机工程与应用*, 2018, 54(7): 20–28, 55.
ZHU Chaozheng, HE Ming, YANG Sheng, et al. Review of monocular visual odometry[J]. *Computer Engineering and Application*, 2018, 54(7): 20–28, 55. (in Chinese)
- [20] 李宇波, 朱效洲, 卢惠民, 等. 视觉里程计技术综述[J]. *计算机应用研究*, 2012, 29(8): 2801–2805, 2810.
LI Yubo, ZHU Xiaozhou, LU Huimin, et al. Review on visual odometry technology[J]. *Application Research of Computers*, 2012, 29(8): 2801–2805, 2810. (in Chinese)
- [21] KLEIN G, MURRAY D. Parallel tracking and mapping for small AR workspaces[C]// *IEEE and ACM International Symposium on Mixed and Augmented Reality*. IEEE, 2007: 225–234.
- [22] LI X, LI Y, ÖRNEK E P, et al. Coplanar parametrization for stereo-SLAM and visual-inertial odometry[J]. *IEEE Robotics and Automation Letters*, 2020, 5(4): 6972–6979.
- [23] ZHANG T, ZHANG H, LI Y, et al. Flow fusion: dynamic dense RGB-D SLAM based on optical flow[C]// *International Conference on Robotics and Automation (ICRA)*. IEEE, 2020: 7322–7328.
- [24] GALLEGÓ G, DELBRÜCK T, ORCHARD G, et al. Event-based vision: a survey[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 44(1): 154–180.
- [25] LU F, MILIOS E. Globally consistent range scan alignment for environment mapping[J]. *Autonomous Robots*, 1997, 4(4): 333–349.
- [26] BAILEY T, DURRAT-WHYTE H. Simultaneous localization and mapping (SLAM): Part II[J]. *IEEE Robotics & Automation Magazine*, 2006, 13(3): 108–117.
- [27] DURRANT-WHYTE H, BAILEY T. Simultaneous localization and mapping: Part I[J]. *IEEE Robotics & Automation Magazine*, 2006, 13(2): 99–110.
- [28] GRISETTI G, STACHNISS C, BURGARD W. Improved techniques for grid mapping with Rao-Blackwellized particle filters[J]. *IEEE Transactions on Robotics*, 2007, 23(1): 34–46.
- [29] HESS W, KOHLER D, RAPP H, et al. Real time loop closure in 2D LiDAR SLAM[C]// *IEEE International Conference on Robotics and Automation*, 2016: 1271–1278.
- [30] ZHANG J, SINGH S. LOAM: LiDAR odometry and mapping in realtime[C]// *Robotics: Science and Systems Conference*, 2014: 1–9.
- [31] SHAN T, ENGLÖT B. LEGO-LOAM: lightweight and ground-optimized LiDAR odometry and mapping on variable terrain[C]// *International Conference on Intelligent Robots and Systems*, 2018: 4758–4765.
- [32] SHAN T, ENGLÖT B, MEYERS D, et al. LIO-SAM: tightly-coupled LiDAR inertial odometry via smoothing and mapping[C]// *International Conference on Intelligent Robot and Systems (IROS)*, 2020: 5135–5142.
- [33] 张虎. 机器人 SLAM 导航核心技术与实践[M]. 北京: 机械工业出版社, 2022: 223–273.
- [34] 温珍强. 结合单目视觉的 RGB-D SLAM 算法及其在云机器人下的实现[D]. 南昌: 华东交通大学, 2017.
WEN Zhenqiang. The RGB-D SLAM combined with monocular vision and its implementation under cloud robot[D]. Nanchang: East China Jiaotong University, 2017. (in Chinese)
- [35] ENDRES F, HESS J, ENGELHARD N, et al. An evaluation of the RGB-D SLAM system[C]// *International Conference on Robotics and Automation*. IEEE, 2012: 1691–1696.
- [36] MUR-ARTAL R, MONTIEL J M M, TARDOS J D. ORB-SLAM: a versatile and accurate monocular SLAM system[J]. *IEEE Transactions on Robotics*, 2015, 31(5): 1147–1163.

- [37] NEWCOMBE R A, LOVEGROVE S J, DAV-ISON A J. DTAM: dense tracking and mapping in real-time[C] // International Conference on Computer Vision. IEEE, 2011: 2320 – 2327.
- [38] 黄琼. 面向 LSD – SLAM 算法的实时占用栅格图研究[D]. 西安: 西安科技大学, 2019.
HUANG Qiong. Research on real-time occupancy grid mapping for LSD – SLAM[D]. Xi'an: Xi'an University of Science and Technology, 2019. (in Chinese)
- [39] 王煜文. 基于 DSO 的动态环境语义地图构建技术研究[D]. 成都: 电子科技大学, 2022.
WANG Yuwen. Research on dynamic environment semantic map construction technology based on DSO [D]. Chengdu: University of Electronic Science and Technology, 2022. (in Chinese)
- [40] ZHANG Z. Microsoft kinect sensor and its effect[J]. IEEE Multimedia, 2012, 19(2): 4 – 10.
- [41] FANKHAUSER P, BLOESCH M, RODRIGUEZ D, et al. Kinect V2 for mobile robot navigation: evaluation and modeling[C] // Proceedings of International Conference on Advanced Robotics, 2015: 388 – 394.
- [42] 慈文彦, 黄影平, 胡兴. 视觉里程计算法研究综述[J]. 计算机应用研究, 2019, 36(9): 2561 – 2568.
CI Wenyan, HUANG Yingping, HU Xing. Review of visual odometry algorithm [J]. Application Research of Computers, 2019, 36(9): 2561 – 2568. (in Chinese)
- [43] 马科伟, 张镔石, 康宇航, 等. 移动机器人中视觉里程计技术综述[J]. 计算机工程, 2021, 47(11): 1 – 10.
MA Kewei, ZHANG Qieshi, KANG Yuhang, et al. Overview of visual odometry technology in mobile robots[J]. Computer Engineering, 2021, 47(11): 1 – 10. (in Chinese)
- [44] CHEN S B, ZHOU B D, JIANG C H, et al. A LiDAR/visual SLAM backend with loop closure detection and graph optimization [J]. Remote Sensing, 2021, 13(14): 2720.
- [45] GAN Y, ZHANG J H, CHEN K Q, et al. A dynamic detection method to improve SLAM performance[J]. Optoelectronics Letters, 2021, 17(11): 693 – 698.
- [46] CAMPOS C, ELVIRA R, RODRIGUEZ J J G, et al. ORB – SLAM3: an accurate open source library for visual, visual-inertial, and multi map SLAM[J]. IEEE Transactions on Robotics, 2021, 37(6): 1874 – 1890.
- [47] LOWE D G. Object recognition from local scaleinvariant features [C] // Proceedings of the Seventh IEEE International Conference on Computer Vision. IEEE, 1999: 1150 – 1157.
- [48] LOWE D G. Distinctive image features from scaleinvariant key points[J]. International Journal of Computer Vision, 2004, 60(2): 91 – 110.
- [49] BAY H, ESS A, TUYTELAARS T, et al. Speeded up robust features (SURF) [J]. Computer Vision and Image Understanding, 2008, 110(3): 346 – 359.
- [50] YANG X, CHENG K T. Accelerating surf detector on mobile devices [C] // Proceedings of the 20th ACM International Conference on Multimedia, 2012: 569 – 578.
- [51] HUANG H Y, YE H Y, SUN Y X, et al. Incorporating learnt local and global embeddings into monocular visual SLAM[J]. Autonomous Robots, 2021, 45(6): 789 – 803.
- [52] HARTLEY R I. In defense of the eight-point algorithm[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(6): 580 – 593.
- [53] NG P C, HENIKOFF S. SIFT: predicting amino acid changes that affect protein function[J]. Nucleic Acids Research, 2003, 31(13): 3812 – 3814.
- [54] BAY H, TUYTELAARS T, VAN G L. Surf: speeded up robust features[J]. Lecture Notes in Computer Science, 2006, 3951: 404 – 417.
- [55] RUBLEE E, RABAU D V, KONOLIGE K, et al. ORB: an efficient alternative to SIFT or SURF [C] // IEEE International Conference on Computer Vision, 2011: 2564 – 2571.
- [56] YANG H, LI H, CHEN K, et al. Feature points extraction based on improved ORB – SLAM [C] // 2019 Chinese Automation Congress (CAC). IEEE, 2019: 936 – 940.
- [57] 宗泽华, 张海军, 张劲锋. 基于自适应阈值的 ORB 特征提取与跟踪[J]. 电光与控制, 2022, 29(3): 110 – 114.
ZONG Zehua, ZHANG Haijun, ZHANG Jinfeng. ORB feature extraction and tracking based on adaptive threshold [J]. Electronics Optics & Control, 2022, 29(3): 110 – 114. (in Chinese)
- [58] YAO J, ZHANG P, WANG Y, et al. An adaptive uniform distribution ORB based on improved quadtree[J]. IEEE Access, 2019, 7: 143471 – 143478.
- [59] GUO J L, XU Y H, DUAN J W, et al. ORB – SLAM method based on local adaptive threshold extraction feature points[J]. Bulletin of Surveying and Mapping, 2021(9): 32.
- [60] YU L, YANG E, YANG B. AFE – ORB – SLAM: robust monocular v-SLAM based on adaptive fast threshold and image enhancement for complex lighting environments[J]. Journal of Intelligent & Robotic Systems, 2022, 105(2): 26.
- [61] WU X, SUN C, CHEN L, et al. Adaptive ORB feature detection with a variable extraction radius in Roi for complex illumination scenes[J]. Robotics and Autonomous Systems, 2022, 157: 104248.
- [62] ENDRES F, HESS J, STURM J, et al. 3-D mapping with an RGB – D camera[J]. IEEE Transactions on Robotics, 2013, 30(1): 177 – 187.
- [63] BELTER D, NOWICKI M, SKRZYPCZYŃSKI P. Improving accuracy of feature-based RGB – D SLAM by modeling spatial uncertainty of point features[C] // IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2016: 1279 – 1284.
- [64] 陈劭, 郭宇翔, 高天啸, 等. 移动机器人 RGB – D 视觉 SLAM 算法[J]. 农业机械学报, 2018, 49(10): 38 – 45.
CHEN Shao, GUO Yuxiang, GAO Tianxiao, et al. RGB – D visual SLAM algorithm for mobile robots[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(10): 38 – 45. (in Chinese)

- [65] MUR-ARTAL R, TARDOS J D. ORB-SLAM2: an open-source slam system for monocular, stereo, and RGB-D cameras[J]. IEEE Transactions on Robotics, 2017, 33(5): 1255-1262.
- [66] 范永祥,冯仲科,陈盼盼,等. 基于 RGB-D SLAM 手机的森林样地调查系统研究[J]. 农业机械学报, 2019, 50(8): 226-234.
FAN Yongxiang, FENG Zhongke, CHEN Panpan, et al. Research on forest plot survey system based on RGB-D SLAM mobile phone[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(8): 226-234. (in Chinese)
- [67] LI G H, CHEN S L. Visual SLAM in dynamic scenes based on object tracking and static points detection[J]. Journal of Intelligent & Robotic Systems, 2022, 104(2): 33.
- [68] 董蕊芳,王宇鹏,阚江明. 基于改进 ORB_SLAM2 的机器人视觉导航方法[J]. 农业机械学报, 2022, 53(10): 306-317.
DONG Ruifang, WANG Yupeng, KAN Jiangming. Visual navigation method for robot based on improved ORB_SLAM2[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(10): 306-317. (in Chinese)
- [69] 丛佩超,崔利营,万现全,等. 基于改进 ORB-SLAM2 的果园喷药机器人定位与稠密建图算法[J]. 农业机械学报, 2023, 54(7): 45-55.
CONG Peichao, CUI Liying, WAN Xianquan, et al. Localization and dense mapping algorithm for orchard spraying robot based on improved ORB-SLAM2[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(7): 45-55. (in Chinese)
- [70] 刘庆运,杨华阳,刘涛,等. 基于激光雷达与深度相机融合的 SLAM 算法[J]. 农业机械学报, 2023, 54(11): 29-38.
LIU Qingyun, YANG Huayang, LIU Tao, et al. SLAM algorithm based on fusion of LiDAR and depth camera[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(11): 29-38. (in Chinese)
- [71] 沈跃,肖鑫桦,刘慧,等. 果园机器人 LiDAR/IMU 紧耦合实时定位与建图方法[J]. 农业机械学报, 2023, 54(11): 20-28, 48.
SHEN Yue, XIAO Xinhua, LIU Hui, et al. Real-time localization and mapping method for agricultural robot in orchards based on LiDAR/IMU tight-coupling[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(11): 20-28, 48. (in Chinese)
- [72] ZHAO S, ZHANG T, WEI H. Fast visual inertial odometry with point-line features using adaptive EDLines algorithm[J]. Measurement Science and Technology, 2022, 33(10): 105401.
- [73] HANG L, KOCH R. An efficient and robust line segment matching approach based on LBD descriptor and pairwise geometric consistency[J]. Journal of Visual Communication & Image Representation, 2013, 24(7): 794-805.
- [74] RAFAEL G V G, JEREMIE J K B, JEANMICHEL M, et al. LSD: a line segment detector[J]. Image Processing on Line, 2012(2): 35-55.
- [75] LEE J H, LEE S, ZHANG G, et al. Outdoor place recognition in urban environments using straight lines[C] // International Conference on Robotics and Automation (ICRA). IEEE, 2014: 5550-5557.
- [76] LU Y, SONG D. Robust RGB-D odometry using point and line features[C] // Proceedings of the IEEE International Conference on Computer Vision, 2015: 3934-3942.
- [77] WANG R, DI K, WAN W, et al. Improved point-line feature based visual SLAM method for indoor scenes[J]. Sensors, 2018, 18(10): 1-20.
- [78] GOMEZ-OJEDA R, MORENO F A, ZUNIGA-NOËL D, et al. PL-SLAM: a stereo SLAM system through the combination of points and line segments[J]. IEEE Transactions on Robotics, 2019, 35(3): 734-746.
- [79] LI Y, BRASCH N, WANG Y, et al. Structure SLAM: low-drift monocular slam in indoor environments[J]. IEEE Robotics and Automation Letters, 2020, 5(4): 6583-6590.
- [80] WANG Q, YAN Z, WANG J, et al. Line flow based simultaneous localization and mapping[J]. IEEE Transactions on Robotics, 2021, 37(5): 1416-1432.
- [81] KRUL S, PANTOS C, FRANGULEA M, et al. Visual SLAM for indoor livestock and farming using a small drone with a monocular camera: a feasibility study[J]. Drones, 2021, 5(2): 1-15.
- [82] LI Q, WANG X, WU T, et al. Point-line feature fusion based fieldreal-time RGB-D SLAM[J]. Computers & Graphics, 2022, 107: 10-19.
- [83] LSLAM R, HABIBULLAH H, HOSSAIN T. AGRIS-LAM: a real-timestereo visual SLAM for agricultural environmentl[J]. Autonomous Robots, 2023, 47(6): 649-668.
- [84] LABBÉ M, MICHAUD F. RTAB-Map as an open-source LiDAR and visual simultaneous localization and mapping library for large-scale and long-term online operation[J]. Journal of Field Robotics, 2019, 36(2): 416-446.
- [85] FU Q, YU H, WANG X, et al. Fast ORB-SLAM without keypoint descriptors[J]. IEEE Transactions on Image Processing, 2021, 31(12): 1433-1446.
- [86] PUMAROLA A, VAKHITOV A, AGUDO A, et al. PL-SLAM: real-time monocular visual SLAM with points and lines[C] // IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2017: 4503-4508.
- [87] HE Y, ZHAO J, GUO Y, et al. PL-VIO: tightly coupled monocular visual-inertial odometry using point and line features[J]. Sensor, 2018, 18(4): 1159-1175.
- [88] FU Q, WANG J, YU H, et al. Pl-vins: real-time monocular visual-inertial SLAM with point and line features[J]. arXiv Preprint, arXiv: 2009.07462, 2020.
- [89] 胡凯,吴佳胜,郑翥,等. 视觉里程计研究综述[J]. 南京信息工程大学学报(自然科学), 2021, 13(3): 269-280.
HU Kai, WU Jiasheng, ZHENG Fei, et al. Review of visual odometer research[J]. Journal of Nanjing University of Information Technology (Natural Science), 2021, 13(3): 269-280. (in Chinese)
- [90] SUHR J K. Kanade-lucas-tomasi feature tracker[J]. Computer Vision (EEE6503), 2009: 9-18.
- [91] LEE C, KOSTA A K, ZHU A Z, et al. Spike-flownet: event-based optical flow estimation with energy-efficient hybrid neural

- networks[C]//European Conference on Computer Vision. Cham: Springer International Publishing, 2020: 366–382.
- [92] 梁健. 基于直接法的单目同时定位与地图构建技术研究[D]. 上海: 东华大学, 2020.
LIANG Jian. Research on monocular simultaneous localization and mapping based on direct method[D]. Shanghai: Donghua University, 2020. (in Chinese)
- [93] ENGEL J, SCHPS T, CREMERS D. LSD-SLAM: large-scale direct monocular SLAM[C]//European Conference on Computer Vision. Cham: Springer International Publishing, 2014: 834–849.
- [94] ENGEL J, STÜCKLER J, CREMERS D. Large-scale direct SLAM with stereo cameras[C]//International Conference on Intelligent Robots and Systems (IROS). IEEE, 2015: 1935–1942.
- [95] CARUSO D, ENGEL J, CREMERS D. Large-scale direct SLAM for omnidirectional cameras[C]//International Conference on Intelligent Robots and Systems (IROS). IEEE, 2015: 141–148.
- [96] ENGEL J, KOLTUN V, CREMERS D. Direct sparse odometry[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 40(3): 611–625.
- [97] GAO X, WANG R, DEMMEL N, et al. LDSO: direct sparse odometry with loop closure[C]//International Conference on Intelligent Robots and Systems (IROS). IEEE, 2018: 2198–2204.
- [98] MO J, ISLAM M J, SATTAR J. Fast direct stereo visual SLAM[J]. IEEE Robotics and Automation Letters, 2021, 7(2): 778–785.
- [99] CHEN Y, NI J, MUTABAZI E, et al. A variable radius side window direct SLAM method based on semantic information[J]. Computational Intelligence and Neuroscience, 2022, 2022: 1–18.
- [100] 高成强, 张云洲, 王晓哲, 等. 面向室内动态环境的半直接法 RGB-D SLAM 算法[J]. 机器人, 2019, 41(3): 372–383.
GAO Chengqiang, ZHANG Yunzhou, WANG Xiaozhe, et al. Semi-direct RGB-D SLAM algorithm for dynamic indoor environment[J]. Robot, 2019, 41(3): 372–383. (in Chinese)
- [101] 李伟, 任孟瀚, 黄威豪, 等. 基于改进 M-ORB 的视觉 SLAM 直接-闭环检测算法[J]. 智能科学与技术学报, 2021, 3(4): 482–491.
LI Wei, REN Menghan, HUANG Weihao, et al. Improved M-ORB based direct-loop closure detection algorithm for visual SLAM[J]. Chinese Journal of Intelligence Science and Technology, 2021, 3(4): 482–491. (in Chinese)
- [102] SHU F, LESUR P, XIE Y, et al. SLAM in the field: an evaluation of monocular mapping and localization on challenging dynamic agricultural environment[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2021: 1761–1771.
- [103] YU T, ZHOU J, WANG L, et al. Accurate and robust stereo direct visual odometry for agricultural environment[C]//2021 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2021: 2480–2486.
- [104] YU T, YU X, LIU W L, et al. Scale-aware stereo direct visual odometry with online photometric calibration for agricultural environment[J]. Advanced Robotics, 2023, 37(6): 433–446.
- [105] KENDALL A, GRIMES M, CIPOLLA R. PoseNet: a convolutional network for real-time 6-dof camera relocalization[C]//Proceedings of the IEEE International Conference on Computer Vision. Santiago: IEEE, 2015: 2938–2946.
- [106] WANG S, CLARK R, WEN H, et al. Deepvo: towards end-to-end visual odometry with deep recurrent convolutional neural networks[C]//International Conference. IEEE, 2017: 2043–2050.
- [107] FU H, GONG M, WANG C, et al. Deep ordinal regression network for monocular depth estimation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 2002–2011.
- [108] WANG R, PIZER S M, FRAHM J M. Recurrent neural network for (un-) supervised learning of monocular video visual odometry and depth[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 5555–5564.
- [109] BRUNO H M S, COLOMBINI E L. LIFT-SLAM: a deep-learning feature-based monocular visual SLAM method[J]. Neuro Computing, 2021, 455: 97–110.
- [110] AGUIAR A S, MONTEIRO N N, SANTOS F N, et al. Bringing semantics to the vineyard: an approach on deep learning-based vine trunk detection[J]. Agriculture, 2021, 11(2): 131.
- [111] RAUF A, MUHAMAMD W, MEHMOOD Z, et al. Simultaneous localization and mapping of UAV for precision agricultural application[C]//2023 International Conference on Energy, Power, Environment, Control, and Computing (ICEPECC). IEEE, 2023: 1–6.
- [112] DE SILVA R, CIELNIAK G, WANG G, et al. Deep learning-based crop row detection for infield navigation of agrirobots[J]. Journal of Field Robotics, 2023(7): 1–23.
- [113] GARG R, BG V K, CARNEIRO G, et al. Unsupervised CNN for single view depth estimation: geometry to the rescue[C]//European Conference on Computer Vision. Springer, Cham, 2016: 740–756.
- [114] ZHOU T, BROWN M, SNAVELY N, et al. Unsupervised learning of depth and egomotion from video[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1851–1858.
- [115] LI R, WANG S, LONG Z, et al. Undeepvo: monocular visual odometry through unsupervised deep learning[C]//International Conference on Robotics and Automation (ICRA). IEEE, 2018: 7286–7291.
- [116] LI R, WANG S, GU D. Deep SLAM: a robust monocular SLAM system with unsupervised deep learning[J]. IEEE Transactions on Industrial Electronics, 2020, 68(4): 3577–3587.
- [117] LI G, YU L, FEI S. A deep-learning real-time visual SLAM system based on multi-task feature extraction network and self-supervised feature points[J]. Measurement, 2021, 168: 108403.
- [118] AN Y, SHI J, GU D, et al. Visual-LiDAR SLAM based on unsupervised multi-channel deep neural networks[J]. Cognitive Computation, 2022, 14(4): 1496–1508.

- [119] BAH M D, HAFIANE A, CANALS R. Deep learning with unsupervised data labeling for weed detection in line crops in UAV images[J]. *Remote Sensing*, 2018, 10(11): 1690.
- [120] HUANG P, ZHU L, ZHANG Z, et al. An end-to-end learning-based row-following system for an agricultural robot in structured apple orchards[J]. *Mathematical Problems in Engineering*, 2021, 2021: 1–14.
- [121] 吴雄伟,周云成,刘峻淳,等. 面向温室移动机器人的无监督视觉里程估计方法[J]. *农业工程学报*, 2023, 39(10): 163–174.
WU Xiongwei, ZHOU Yuncheng, LIU Juntao, et al. Unsupervised visual odometry method for greenhouse mobile robots[J]. *Transactions of the CSAE*, 2023, 39(10): 163–174. (in Chinese)
- [122] GEIGER A, LENZ P, STILLER C, et al. Vision meets robotics: the kitti dataset[J]. *The International Journal of Robotics Research*, 2013, 32(11): 1231–1237.
- [123] 李鹏,刘琪,何宸宇,等. 融合深度学习和几何定姿的光流视觉里程计研究[J]. *电机与控制学报*, 2020, 24(12): 142–151.
LI Peng, LIU Qi, HE Chenyu, et al. Research on optical flow visual odometer combined deep learning and geometric attitude determination[J]. *Electric Machines and Control*, 2020, 24(12): 142–151. (in Chinese)
- [124] ALMALIOGLU Y, TURAN M, SAPUTRA M R U, et al. Self-supervised deep monocular visual-inertial odometry and depth estimation[J]. *Neural Networks*, 2022, 150: 119–136.
- [125] SHINDE K, LEE J, HUMT M, et al. Learning multiplicative interactions with bayesian neural networks for visual-inertial odometry[J]. *arXiv Preprint, arXiv: 2007.07630*, 2020.
- [126] BURRI M, NIKOLIC J, GOHL P, et al. The EuRoC micro aerial vehicle datasets[J]. *The International Journal of Robotics Research*, 2016, 35(10): 1157–1163.
- [127] ASLAN M F, DURDU A, SABANCI K. Visualinertial image-odometry network (VIIONet): a Gaussian process regression-based deep architecture proposal for UAV pose estimation[J]. *Measurement*, 2022, 194: 111030.
- [128] 余洪山,郭丰,郭林峰,等. 融合改进 Super Point 网络的鲁棒单目视觉惯性 SLAM[J]. *仪器仪表学报*, 2021, 42(1): 116–126.
YU Hongshan, GUO Feng, GUO Linfeng, et al. Robust monocular visual-inertial SLAM based on the improved Super Point network[J]. *Chinese Journal of Scientific Instrument*, 2021, 42(1): 116–126. (in Chinese)
- [129] STURM J, ENGELHARD N, ENDRES F, et al. A benchmark for the evaluation of RGB–D SLAM systems[C] // *International Conference on Intelligent Robots and Systems*, 2012: 573–580.
- [130] 王恩宝,周亚丽,张奇志. 基于深度学习的视觉里程计改进与分析[J]. *计算机应用研究*, 2020, 37(增刊 2): 324–326, 329.
- [131] 李博,段中兴. 室内动态环境下基于深度学习的视觉里程计[J]. *小型微型计算机系统*, 2023, 44(1): 49–55.
LI Bo, DUAN Zhongxing. Visual odometry based on deep learning in dynamic indoor environment[J]. *Journal of Chinese Computer Systems*, 2023, 44(1): 49–55. (in Chinese)
- [132] CORDTS M, OMRAN M, RAMOS S, et al. The cityscapes dataset for semantic urban scene understanding[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 3213–3223.
- [133] LOU L, LI Y, ZHANG Q, et al. SLAM and 3D semantic reconstruction based on the fusion of LiDAR and monocular vision[J]. *Sensors*, 2023, 23(3): 1502.
- [134] MA P, BAI Y, ZHU J, et al. DSOD: DSO in dynamic environments[J]. *IEEE Access*, 2019, 7: 178300–178309.
- [135] LI X, AO H, BELAROUSSI R, et al. Fast semi-dense 3D semantic mapping with monocular visual SLAM[C] // *International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2017: 385–390.
- [136] PARK S, SCHÖPS T, POLLEFEYS M. Illumination change robustness in direct visual SLAM[C] // *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2017: 4523–4530.
- [137] HUANG J, LIU S. Robust simultaneous localization and mapping in low-light environment[J]. *Computer Animation and Virtual Worlds*, 2019, 30(3–4): e1895.
- [138] SUN C, QIAO N, GE W, et al. Robust RGB–D visual odometry using point and line features[C] // *2022 41st Chinese Control Conference (CCC)*. IEEE, 2022: 3826–3831.
- [139] WANG W, ZHU D, WANG X, et al. TartanAir: a dataset to push the limits of visual SLAM[C] // *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2021: 4909–4916.
- [140] GU X, WANG Y, MA T. DBLD–SLAM: a deep-learning visual SLAM system based on deep binary local descriptor[C] // *2021 International Conference on Control, Automation and Information Sciences (ICC–AIS)*. IEEE, 2021: 325–330.
- [141] LIU K, FAN Y, SU J, et al. A compensation method for lost trajectories of ORB–SLAM3 with DROID–VO[C] // *2023 3rd International Conference on Computer, Control and Robotics (ICCCR)*, 2023: 309–316.
- [142] WOFK D, RANFTL R, MÜLLER M, et al. Monocular visual-inertial depth estimation[J]. *arXiv Preprint, arXiv:2303.12134*, 2023.
- [143] PALAZZOLO E, BEHLEY J, LOTTES P, et al. Re-Fusion: 3D reconstruction in dynamic environments for RGB–D cameras exploiting residuals[C] // *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019: 7855–7862.
- [144] YAN L, HU X, ZHAO L, et al. DGS–SLAM: a fast and robust RGBD SLAM in dynamic environments combined by geometric and semantic information[J]. *Remote Sensing*, 2022, 14(3): 795.
- [145] WANG Y, XU K, TIAN Y, et al. DRG–SLAM: a semantic RGB–D SLAM using geometric features for indoor dynamic scene[C] // *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022: 1352–1359.
- [146] CHENG S, SUN C, ZHANG S, et al. SG–SLAM: a real-time RGB–D visual SLAM toward dynamic scenes with semantic and geometric information[J]. *IEEE Transactions on Instrumentation and Measurement*, 2023: 1–12.
- [147] LU Q, PAN Y, HU L, et al. A method for reconstructing background from RGB–D SLAM in indoor dynamic environments[J]. *Sensors*, 2023, 23(7): 3529.