DOI:10.19651/j.cnki.emt.2417314

基于 Deeplabv3+和 LK 光流的动态视觉 SLAM 算法*

周若轩[1] 张瑞乾[1] 陈 勇[1,2] 袁旭浩[1] 秦慧军[1]
(1.北京信息科技大学机电工程学院 北京 100192;2.新能源汽车北京实验室 北京 100192)

摘 要:传统 ORB-SLAM3 系统在静态环境中表现优秀,但存在动态特征时会引入不必要的噪声,造成特征匹配出现错误,而现有动态 SLAM 算法难以完整判断潜在动态特征,出现漏检或误检导致定位精度下降。针对这些问题,将语义分割网络 Deeplabv3+与 LK 光流法融合进 ORB-SLAM3 的跟踪线程,首先将 Deeplabv3+的主干网络替换为 Mobilenetv3,提高语义分割的精度,然后使用语义分割获取潜在动态目标掩码,初步滤除动态特征点,对剩余特征点进行 LK 光流计算,将光流平均误差作为阈值防止剩余静态特征点过少导致位姿估计失败。相比于原 ORB-SLAM3,本文改进的算法在 TUM 高动态序列的定位精度平均提升了 47.92%,与现有优秀动态 SLAM 算法相比,本文算法在 TUM 数据集的 Walking_static 序列取得了最高的定位精度。

关键词:同时定位与地图构建;动态环境;语义分割;光流法
中图分类号:TP391.9;TN98 **文献标识码**:A **国家标准学科分类代码**:510.99

Dynamic visual SLAM algorithm based on deeplabv3+ and LK optical flow

Zhou Ruoxuan[1] Zhang Ruiqian[1] Chen Yong[1,2] Yuan Xuhao[1] Qin Huijun[1]
(1. School of Mechanical and Electrical Engineering, Beijing Information Science and Technology University, Beijing 100192, China; 2. New Energy Vehicle Research Lab, Beijing 100192, China)

Abstract: The traditional ORB-SLAM3 system demonstrates excellent performance in static environments, however, the presence of dynamic features introduces unnecessary noise, leading to errors in feature matching and inaccuracies in camera pose estimation. Existing dynamic SLAM algorithms face challenges in comprehensively identifying potential dynamic features, resulting in missed detections or false positives and consequently degrading localization accuracy. To tackle these issues, the semantic segmentation network Deeplabv3+ and the Lucas-Kanade optical flow method are incorporated into the tracking thread of ORB-SLAM3. Specifically, the backbone network of Deeplabv3+ is replaced with Mobilenetv3 to enhance the precision of semantic segmentation. Semantic segmentation is then used to obtain a mask of potential dynamic objects, which is employed to preliminarily filter out dynamic feature points. The remaining feature points undergo LK optical flow calculation, with the average optical flow error serving as a threshold to prevent the insufficient number of static feature points from causing pose estimation failure. In comparison to the original ORB-SLAM3, the improved algorithm in this study achieves an average localization accuracy improvement of 47.92% on the high-dynamic sequences of the TUM dataset. Furthermore, among existing advanced dynamic SLAM algorithms, the proposed method achieved the highest localization accuracy on the Walking_static sequence of the TUM dataset.

Keywords: SLAM; dynamic environment; semantic segmentation; feature matchin

0 引 言

同时定位与地图构建(simultaneous localization and mapping,SLAM)算法通过融合来自激光雷达、摄像头和惯性测量单元(inertial measurement unit,IMU)等多种传感器的数据,不仅可以提升自动驾驶系统的定位精度,还能为环境感知提供精准的数据支持。现有视觉 SLAM 算法根据相机位姿估计方式的不同分为直接法和特征点法。直接法利用图像中的亮度信息来估计相机的运动和场景结构,特征点法通过提取和匹配图像中的特征点来进行相机位姿估计。关于直接法,Forster 等[1]提出了一种视觉位姿计算方法 SVO,该方法采用直接法解决连续图像帧之间的像素

收稿日期:2024-11-08
*基金项目:国家自然科学基金面上项目(52077007)资助

位移问题，并利用小摄动模型对位姿和场点云进行优化，从而在未知环境下实现快速准确的定位。关于特征点法，Campos 等[2]推出了 ORB-SLAM3 算法，通过引入 IMU 和相机融合技术，首次实现了在算法各个阶段对高视差共视关键帧中所有先前信息的有效再利用。王喜红等[3]在 ORB-SLAM2 的基础上引入四叉树策略筛选划分后的网格金字塔，特征匹配时引入 GMS 算法剔除误匹配的特征点，提高了匹配的准确性。上述 SLAM 系统在静态环境中表现优秀，但存在动态特征（如行人、车辆等）会引入不必要的噪声，导致定位出现误差。

针对视觉 SLAM 在动态环境中定位精度降低的问题，学者们提出了多种解决方案，包括基于几何的传统动态物体剔除方法和基于深度学习的动态物体剔除方法。关于传统的动态剔除方法，Fan 等[4]研究了噪声模块的引入，提出了一种能够构建高精度点云地图的 SLAM-PCD 方法。Sun 等[5]中提出了一种运动目标去除算法，该算法首先通过稀疏光流法检测动态目标的轮廓，然后利用 Grab-Cut 算法进一步进行目标分割。但这种方法的有效性依赖于相机处于静止状态的假设，因此在实际应用中存在一定的限制。关于深度学习的动态剔除方法，Dyna-SLAM[6]是在 ORB-SLAM2 的基础上结合 Mask R-CNN 和多视图几何的方法进行运动分割；RDS-SLAM[7]在 ORB-SLAM3 的基础上添加了语义线程和基于语义的优化线程，利用移动概率来更新和传播语义信息。Zhang 等[8]提出了一种基于深度学习的语义视觉 SLAM（Sem-VSLAM）方法，首先对帧图像进行语义分割并修正分割结果，然后利用语义信息和运动一致性检测来去除动态目标的特征点，并基于剩余的静态特征点进行位姿估计。Sun 等[9]提出了一种基于 ORB-SLAM3 算法的新型并行处理系统 SOLO-SLAM。通过改进语义线程和设计新的动态点过滤策略，并行完成语义线程和 SLAM 线程，并且使用区域动态度和几何约束的组合进一步增强动态点的过滤效果。Yu 等[10]提出的基于 RGB-D 相机的 DS-SLAM 算法，通过使用 SegNet 网络[11]对动态对象进行语义分割，同时利用 RANSAC 方法计算相机位姿，并结合运动一致性检测技术，剔除了场景中的动态物体。不过，该算法对运动目标的识别仅限于场景中的人。方娟等[12]使用 YOLOv4 算法来提取目标的边界框，并应用最大类间方差算法对深度图像进行前景分割。Ye 等[13]则运用 YOLOv5 算法来处理动态对象检测中的特征点，保留静态特征点以进行姿态估计。然而，由于目标检测方法所提供的边界框无法实现像素级的精确度，去除动态特征点的做法有可能导致静态特征点的丢失。黄友锐等[14]通过加入 Ghost 模块和 SimAM 注意力机制优化 YOLOv5 模型，使用 LK 光流和动态特征点剔除方法剔除动态特征点。Zheng 等[15]提出了一种轻量级的 RLD-SLAM 算法，结合了目标检测与贝叶斯滤波，利用语义信息和运动信息来跟踪动态目标。Zhong 等[16]开发了一种 DynaTM-SLAM 算法，旨在同时优化相机的姿态和地图点信息。Li 等[17]结合了语义特征和混合高斯分布模型进行特征点分割，有效区分静态和动态特征点。上述 SLAM 系统主要在原算法基础上加入目标检测或图像分割滤除动态特征点，但在缺乏足够语义数据或训练数据的场景中，分割网络泛化能力不佳，导致图像目标边缘检测不精确，最终动态特征点的检测和去除不准确。

本文针对现有动态 SLAM 算法难以完整判断潜在动态特征，出现漏检或误检导致定位精度下降的问题，提出一种结合 Deeplabv3＋和 LK（lucas-kanade）光流的视觉 SLAM 系统。本文的贡献如下：

1）将 Deeplabv3＋的主干网络替换为 Mobilenetv3，能够提高语义分割的精度；

2）将 Deeplabv3＋与 LK 光流法融合进 ORB-SLAM3，能够有效滤除动态物体的特征点；

3）引入阈值法，使用光流平均误差作为阈值，可以防止剩余静态点过少导致位姿估计失败；

4）针对本文提出的算法在 TUM 数据集进行了实验验证，实验结果表明本文算法相比于原 ORB-SLAM3 和现有优秀动态 SLAM 算法具有更高的定位精度。

1　基于 deeplabv3＋的语义分割算法

语义分割网络 deeplabv3＋采用编码器-解码器结构。编码器利用主干网络提取浅层特征边，采用空间金字塔池化模块 ASPP 在不同尺度上对图像进行池化操作。ASPP 用以捕获不同尺度下的上下文信息；浅层特征边则被直接送入解码器。解码器将浅层特征与深层特征进行融合，经过一次普通卷积后进行上采样恢复到原图像尺寸，如图 1 所示。

将原本的主干网络 Xception 替换为 Mobilenetv3，Mobilenetv3 重新设计了耗时层，分为 Large 结构和 Small 结构，虽然 Small 结构可以在不损失精度的情况下去掉一些多余的层，相比于 Large 结构其更加简洁，但其没有完整的提取图像信息，为了充分提取特征信息，采用 Large 结构作为网络的主干。

采用 Hard Swish（h-swish）激活函数，组合 Relu6 算子：

$$h\text{-}swish = x\frac{Relu6(x+3)}{6} \tag{1}$$

$Relu6$ 是一种常用的激活函数，公式为：

$$Relu6 = \min\{\max\{0,x\},6\} \tag{2}$$

使用深度可分离卷积作为主要的特征提取层，将标准卷积分为深度卷积和逐点卷积两个步骤，其中深度卷积使用 3×3 的卷积核对特征图进行卷积操作，逐点卷积使用 1×1 的卷积核对特征图进行升维，这样得到的特征通道数大于原有通道数，提高了信息的有效性。

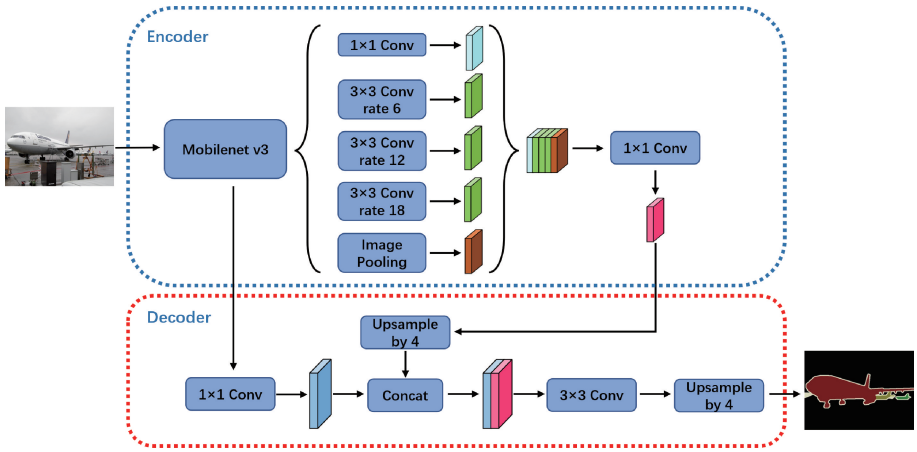


图 1　Deeplabv3+算法框架

Fig. 1　Deeplabv3+ algorithm framework

2　结合 LK 光流法和 Deeplabv3+的动态点滤除

通过结合语义分割来辅助光流估计,可以去除静态背景的对齐误差,从而提高动态特征点的检测和去除精度。语义分割的底层架构为 Deeplabv3+,使用 Mobilenetv3 作为主干网络,提高语义分割的精度;slam 底层算法为 ORB-SLAM3,在跟踪线程中加入动态特征点滤除模块,使用语义分割获取潜在动态目标掩码,初步滤除动态特征点,对剩余点进行 LK 光流计算,设定光流平均误差作为阈值防止剩余静态点过少。改进的 ORB-SLAM3 跟踪线程框架如图 2 所示。

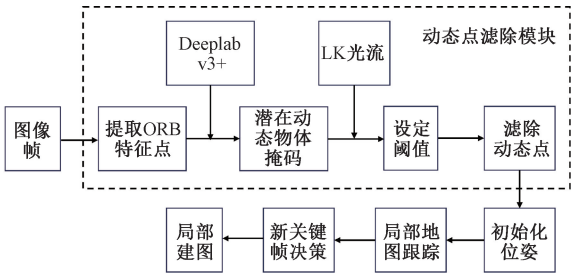


图 2　改进的 ORB-SLAM3 跟踪线程框架

Fig. 2　Improved ORB-SLAM3 tracking thread framework

2.1　LK 光流法

LK 光流法将三维空间物体的三维矢量投影到二维成像平面上,得到光流矢量,用来描述某一点的瞬时速度。对于像素 (x,y),在时间 t 和 $t+\Delta t$ 时刻的图像强度 $I(x,y,t)$ 和 $I(x+u\Delta t, y+v\Delta t, t+\Delta t)$,二者是相等的,其中 u 和 v 分别是像素水平和垂直方向上的光流。

LK 光流法对小窗口内的光流进行估计,使用最小二乘法最小化以下目标函数:

$$E(u,v)=\sum_{(x,y)\in W}(I_x u+I_y v+I_t)^2 \tag{3}$$

式中:W 是一个局部窗口,$I_x u$、$I_y u$、I_t 分别是图像在 x,y,t 方向的梯度。分别求解其对 u 和 v 的偏导数并设为 0。

$$\sum_{(x,y)\in W}(I_x^2 u+I_x I_y v+I_x I_t)=0 \tag{4}$$

$$\sum_{(x,y)\in W}(I_x I_y u+I_y^2 v+I_y I_t)=0 \tag{5}$$

将式(4)和(5)写成矩阵形式:

$$\begin{bmatrix}\sum\limits_{(x,y)\in W}I_x^2 & \sum\limits_{(x,y)\in W}I_x I_y\\ \sum\limits_{(x,y)\in W}I_x I_y & \sum\limits_{(x,y)\in W}I_y^2\end{bmatrix}\begin{bmatrix}u\\v\end{bmatrix}=\begin{bmatrix}-\sum\limits_{(x,y)\in W}I_x I_t\\-\sum\limits_{(x,y)\in W}I_y I_t\end{bmatrix} \tag{6}$$

简化得:

$$\boldsymbol{A}\begin{bmatrix}u\\v\end{bmatrix}=b \tag{7}$$

由式(7)可以得到 u 和 v 的最小二乘解:

$$\begin{bmatrix}u\\v\end{bmatrix}=\boldsymbol{A}^{-1}b \tag{8}$$

式中:u 和 v 分别为像素点在水平和垂直方向的运动速度,即为光流,$\boldsymbol{A}^{-1}$ 是矩阵 $\boldsymbol{A}$ 的逆矩阵。

光流法在动态特征估计方面具有一定的有效性,但在面对大尺度运动或运动速度较快等复杂场景时,光流的估计精度下降。通过结合语义分割来辅助光流估计,可以去除静态背景的对齐误差,从而提高动态特征点的检测和去除精度。

2.2　运动特征的处理

ORB-SLAM3 虽然能通过随机抽样一致算法(random sample consensus,RANSAC)来剔除少量动态特征点,但是运动物体占据画面较大或数量较多时,该算法仍不能很好的滤除这些动态特征。因此将 LK 光流法和 deeplabv3+结合,算法结构如图 3 所示。

结合 LK 光流和 deeplabv3+滤除动态特征点的具体流程如下:

1)原始图像进行 deeplabv3+语义分割处理,将人、车、动物等划归为潜在动态目标,并获取潜在动态目标掩码,

如图 4 所示。图 4(a)为 Tum 数据集的 fr3_walking_hs 序列的原图像，图 4(b)为该图像使用 deeplabv3＋检测后的掩码图，在 deeplabv3＋语义分割结果中，人所在类别的像素值为 15，将其设定为 255，使更明显。

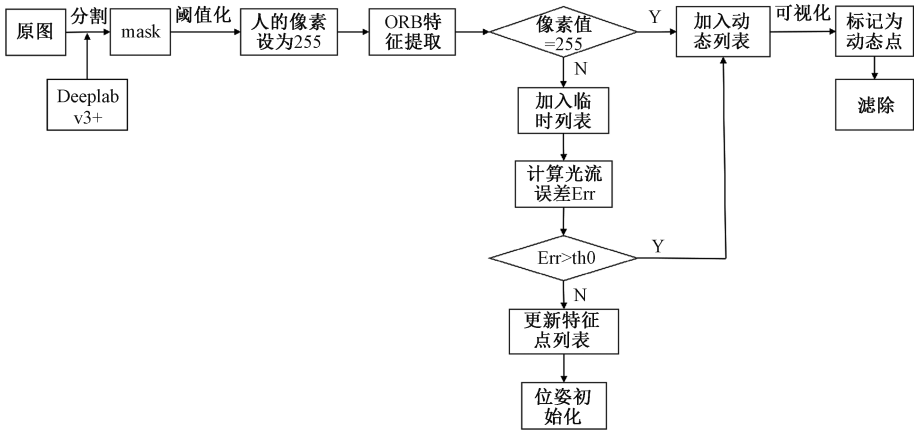
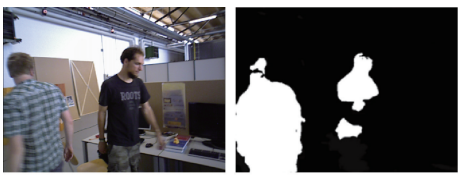


图 3　动态点滤除的算法结构

Fig. 3　Algorithm structure for dynamic point filtering



(a) 原图　(b) 掩码图

(a) Original image　(b) Mask diagram

图 4　语义分割结果图

Fig. 4　Semantic segmentation result diagram

2)遍历关键点列表 mvKeys 中的所有点，对于每个关键点，检查其在掩码图像 mask 中对应位置的像素值。如果像素值不等于 255，则将该关键点添加到临时列表_mvKeys 中，并将其描述符添加到临时描述符矩阵_mDescriptors 中。如果像素值等于 255，则将该关键点添加到动态关键点列表 mvKeysDynamic_express 中。

3)交换关键点列表 mvKeys 和临时列表_mvKeys，以及关键点描述符矩阵 mDescriptors 和临时描述符矩阵_mDescriptors，对动态关键点列表 mvKeysDynamic_express 中的点进行可视化处理，标记为红色。

4)对交换后的关键点列表 mvKeys 中的特征点进行光流提取，将光流的平均误差作为光流阈值，防止因剩余点不足而导致位姿估计失败。像素点的光流误差为前一帧和当前帧的欧几里德距离：

$$Err[i] = \| Cur[i] - Pre[i] \|^2 \tag{9}$$

式中：$Err[i]$为光流误差，$Cur[i]$是当前帧的点的预测位置，$Pre[i]$是前一帧对应点的预测位置。通过计算所有特征点的误差总和：

$$total = \sum_{i=0}^{N-1} Err[i] \tag{10}$$

由式(10)可得光流平均误差：

$$th0 = \frac{total}{N} \tag{11}$$

式中：$total$ 为误差总和，$th0$ 为平均误差，N 为特征点的数量。

光流误差 $Err[i]$超过平均误差 $th0$ 的所有像素点均视为动态点，标记为红色并从 mvKeys 特征点列表中删除。再次交换特征点列表 mvKeys 和临时列表_mvKeys，关键点描述符矩阵 mDescriptors 和临时描述符矩阵_mDescriptors。仅计算 ORB 提取的特征点的光流误差，而不对整幅图像进行光流误差计算，可以减少计算量，有效提升系统的处理效率。

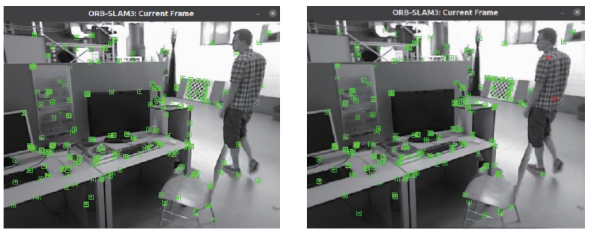
3　实验与分析

使用 TUM RGB-D 开源数据集计算本文算法在 RGB-D 场景下的定位精度，通过与原 ORB-SLAM3 算法以及现有优秀动态 SLAM 算法进行比较，验证本文算法在动态场景下的定位准确性与鲁棒性。实验在 Ubuntu20.04 操作系统下进行，处理器为 Intel i7-13700，显示适配器为 NVIDIAGeForce RTX 4060 GPU，Cuda 版本为 11.8。

分别对原 ORB-SLAM3 算法、仅加入 LK 光流法、仅加入 Deeplabv3＋、同时加入 LK 光流和 Deeplabv3＋进行特征点提取。如图 5 所示，绿色的点为进行位姿解算的特征点，红色为标记的动态特征点，不进行位姿解算，身穿格子衣服的行人为动态目标。可以看到，图 5(a)中所有特征点均设为静态点；图 5(b)中只识别了少量动态点；图 5(c)中行人身上还残留了少量特征点未被识别为动态；图 5(d)中行人身上的所有特征点均被识别为动态，证明同时加入 LK 光流和 Deeplabv3＋后识别动态点效果最好。

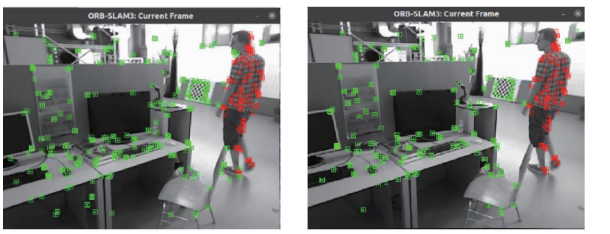
使用绝对轨迹误差的均方根误差(root mean square error，RMSE)作为 ORB-SLAM3 系统的评价指标。$RMSE$ 值越小，代表 SLAM 定位精度越高。

TUM 数据集的不同图像序列下消融实验的 RMSE 值比较如表 1 所示。加粗部分为同一图像序列使用不同算法取得的最低 RMSE 值。可以看到，除了 Sitting_rpy 序



(a) ORB-SLAM3　(b) 仅加入LK光流

(a) ORB-SLAM3　(b) Only adding LK optical flow



(c) 仅加入Deeplabv3+　(d) 加入LK光流和Deeplabv3+

(c) Only adding Deeplabv3+　(d) Adding LK optical flow and Deeplabv3+

图 5　特征点提取对比结果

Fig. 5　Feature point extraction comparison result

列,同时加入 LK 光流法和 Deeplabv3+在其余序列下都取得了最低的 RMSE 值,Sitting_rpy 序列中人始终保持坐姿,动态行为较少,因此不同算法在此序列下取得的 RMSE 值相差不大。

表 1　不同图像序列下消融实验的 RMSE 值对比

Table 1　Comparison of RMSE values from ablation experiments under different image sequences　m

序列	ORB-SLAM3	LK 光流	Deeplabv3+	本文算法
Walking_hs	0.139 5	0.092 9	0.100 3	**0.082 9**
Sitting_rpy	0.025 5	0.025 2	**0.024 1**	0.024 9
Walking_rpy	0.513 6	0.443 9	0.421 4	**0.330 1**
Walking_static	0.012 9	0.008 9	0.008 4	**0.006 2**
Walking_xyz	0.098 7	0.077 4	0.062 3	**0.036 1**

不同算法的 RMSE 提升率对比如表 2 所示,加粗部分为同一序列使用不同算法取得的最高提升率。Walking_xyz 序列中动态行为较多,行人处于持续走动状态,本文算法在此序列上的 RMSE 取得了最高的提升率,而 Sitting_rpy 序列中由于人物动态行为不多,相应的 RMSE 提升率不高。因此本文算法在动态场景中可以达到较好的效果。

表 2　改进算法的 RMSE 提升率对比

Table 2　Comparison of RMSE improvement rates of the improved algorithm　%

序列	LK 光流	Deeplabv3+	本文算法
Walking_hs	33.40	28.10	**40.57**
Sitting_rpy	1.18	**5.49**	2.35
Walking_rpy	13.57	17.95	**35.73**
Walking_static	31.01	34.88	**51.94**
Walking_xyz	51.58	36.88	**63.42**

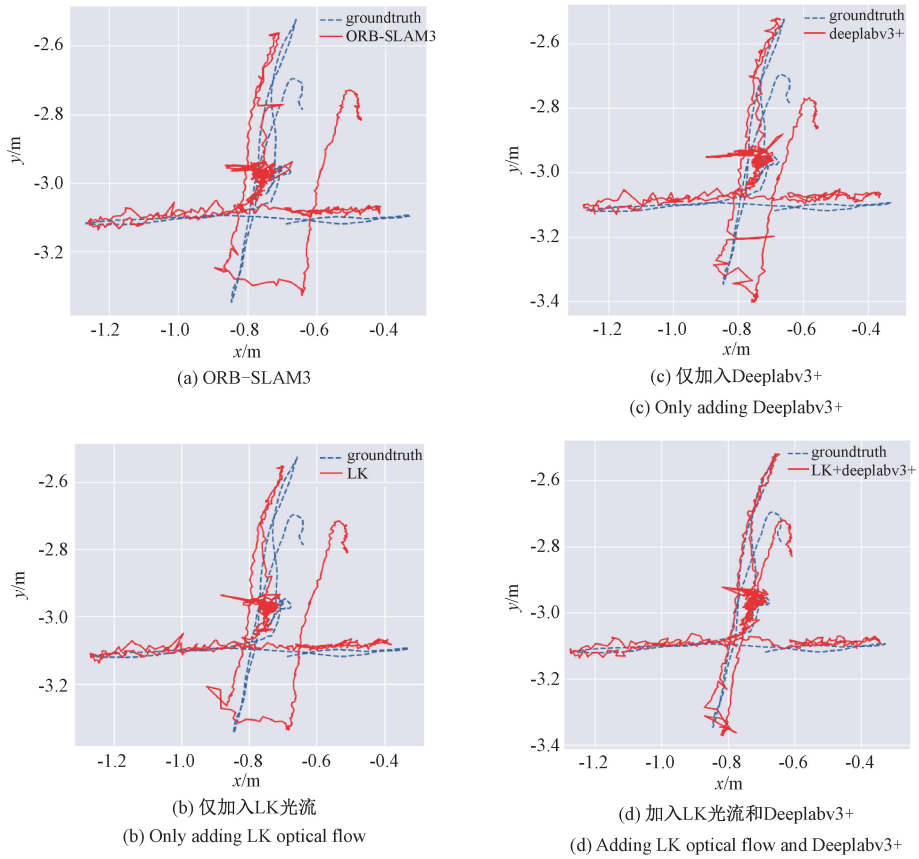
本文算法与现有优秀动态 SLAM 算法的绝对轨迹误差的 RMSE 值对比如表 3 所示。本文选取了 4 个动态 SLAM 算法进行对比,其中,Dyna-SLAM 是在 ORB-SLAM2 的基础上结合 Mask R-CNN 和多视图几何的方法进行运动分割;RDS-SLAM 在 ORB-SLAM3 的基础上添加了语义线程和基于语义的优化线程,利用移动概率来更新和传播语义信息;文献[18]在 ORB-SLAM3 基础上利用多帧间的位姿变换区分高动态和地动态物体,并结合 YOLOv8-Seg 实例分割剔除动态物体;Object-SLAM 在 ORB-SLAM3 基础上采用目标检测算法检测动态物体并剔除动态锚框内所有特征点。在 Walking_hs/rpy/static/xyz 四个高动态序列中进行对比,结果表明,本文改进算法在 Walking_static 序列取得了最高的定位精度;在 Walking_rpy 序列的定位精度还有不足,原因是该序列中相机的运动幅度较大,在提取光流过程中会造成一定程度的误检测,导致原本处于静态的特征点由于相机的运动而被光流法标记为动态;其余两个序列的定位精度差别不大。

表 3　不同 SLAM 算法的绝对轨迹误差的 RMSE 值对比

Table 3　Comparison of RMSE values of absolute trajectory errors for different SLAM algorithms　m

序列	Dyna-SLAM	RDS-SLAM	文献[18]	Object-SLAM	本文算法
Walking_hs	0.029 2	0.024 8	0.029 1	0.097 1	0.082 9
Walking_rpy	0.046 9	0.040 9	0.038 3	—	0.330 1
Walking_static	0.010 8	0.034 3	0.009 2	0.013 4	0.006 2
Walking_xyz	0.031 5	0.010 4	0.026 1	0.042 9	0.036 1

在三维坐标系的 xy 平面,Tum 数据集的 walking_xyz 图像序列的真实轨迹与估计轨迹对比如图 6 所示,蓝色虚线为相机真实轨迹,红色实线为加入不同方法的估计轨迹。可以看到,同时加入 LK 光流法和 Deeplabv3+,算法的估计轨迹与真实轨迹重合度达到最高,算法的准确性和鲁棒性达到最好。



(a) ORB-SLAM3

(c) 仅加入Deeplabv3+

(c) Only adding Deeplabv3+

(b) 仅加入LK光流

(b) Only adding LK optical flow

(d) 加入LK光流和Deeplabv3+

(d) Adding LK optical flow and Deeplabv3+

图 6　walking_xyz 数据集轨迹对比图

Fig. 6　Trajectory comparison graph for the walking_xyz dataset

4　结　　论

提出的基于语义分割和 LK 光流的动态点滤除策略能完整识别环境中先验动态物体，通过滤除动态特征点有效减少了错误的特征匹配，采用阈值法防止剩余静态点过少，提升了 SLAM 系统的定位精度和鲁棒性。

本文改进的算法在 TUM 数据集高动态序列的测试结果表明，相比于原 ORB-SLAM3 算法，单独加入 LK 光流法的定位精度平均提升了 32.39%，单独加入 Deeplabv3+的定位精度平均提升了 29.43%，同时加入 LK 光流和 Deeplabv3+的定位精度平均提升了 47.92%。在动态较少的序列中定位精度提升不大。

相较于现有优秀动态 SLAM 算法，本文改进的算法在 Walking_static 序列取得了最高的定位精度，在 Walking_rpy 序列的定位精度还有不足，在其他序列的定位精度相当。

本文算法在相机运动幅度较大的场景中还存在不足，后续将进行相机运动补偿相关工作，结合 IMU 数据预积分技术，对视觉信息进行联合优化，提高位姿估计的精度和鲁棒性。

参考文献

[1] FORSTER C, PIZZOLI M, DAVIDE S. SVO: Fast semi-direct monocular visual odometry [C]. IEEE International Conference on Robotics & Automation. IEEE, 2014.

[2] CAMPOS C, ELVIRA R, RODRÍGUEZ J J G, et al. ORB-SLAM3: An accurate open-source library for visual, visual-inertial, and multimap SLAM[C]. IEEE Transactions on Robotics, 2021, 37(6): 1874-1890.

[3] 王喜红，雷斌，李园园，等. 基于 ORB-SLAM2 的改进特征匹配与稠密地图算法[J]. 电子测量技术，2024，47(18)：54-62.
WANG X H, LEI B, LI Y Y, et al. Improved feature matching and dense mapping algorithm based on ORB-SLAM2 [J]. Electronic Measurement Technology, 2024, 47(18): 54-62.

[4] FAN Y CH, ZHANG Q CH, LIU SH F, et al. A semantic SLAM with more accurate point cloud map in dynamic environments [J]. IEEE Access, 2020, 8: 112237-112252.

[5] SUN Y X, LIU M, MENG Q H. Invisibility: A moving-object removal approach for dynamic scene

modelling using RGB-D camera [C]. 2017 IEEE International Conference on Robotics and Biomimetics (ROBIO),2017:50-55.

[6] BESCOS B, FÁCIL J M, CIVERA J, et al. DynaSLAM: Tracking, mapping, and inpainting in dynamic scenes[J]. IEEE Robotics and Automation Letters, 2018, 3(4): 4076-4083.

[7] LIU Y B, MIURA J. RDS-SLAM: Real-time dynamic SLAM using semantic segmentation methods [J]. IEEE Access, 2021,9: 23772-23785.

[8] ZHANG K, DONG CH Y, GUO H F, et al. A semantic visual SLAM based on improved mask R-CNN in dynamic environment[J]. Robotica, 2024,42(10):3570-3591.

[9] SUN L X, WEI J Y, SU SH J, et al. SOLO-SLAM: A parallel semantic slam algorithm for dynamic scenes[J]. Sensors,2022, 22(18):6977.

[10] YU CH, LIU Z X, LIU X J, et al. DS-SLAM:A semantic visual SLAM towards dynamic environments[C]. IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway:IEEE, 2018:1168-1174.

[11] BADRINARAYANAN V, KENDALL A, CIPOLLAR. SegNet: A deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence,2017,39(12): 2481-2495.

[12] 方娟, 方振虎. 基于目标检测网络的动态场景下视觉 SLAM 优化[J]. 北京工业大学学报, 2022, 48(5): 466-475.
FANG J, FANG ZH H. Visual SLAM optimization in dynamic scenes based on object detection network. [J]. Journal of Beijing University of Technology, 2022, 48(5): 466-475.

[13] YE Q F, DONG CH Y, LIU X Y, et al. A visual odometry algorithm in dynamic scenes based on object detection[C]. 2022 5th International Conference on Pattern Recognition and Artificial Intelligence(PRAI), 2022:465-470.

[14] 黄友锐,王照锋,韩涛,等.结合轻量化 YOLOv5s 的动态视觉 SLAM算法[J].电子测量技术,2024,47(11):59-68.
HUANG Y R, WANG ZH F, HAN T, et al. Dynamic visual SLAM algorithm combined with lightweight YOLOv5s [J]. Electronic Measurement Technology, 2024, 47(11):59-68.

[15] ZHENG Z R, LIN SH F, YANG CH G. RLD-SLAM: A robust lightweight VI-SLAM for dynamic environments leveraging semantics and motion information [C]. IEEE Transactions on Industrial Electronics,2024,71(11): 14328-14338.

[16] ZHONG M L, HONG CH Y, JIA ZH Q, et al. DynaTM-SLAM: Fast filtering of dynamic feature points and object-based localization in dynamic indoor environments[J]. Robotics and Autonomous Systems, 2024, 174:104634.

[17] LI M R, HE J M, JIANG G G, et al. DDN-SLAM: Real-time dense dynamic neural implicit slam with joint semantic encoding [J]. ArXiv preprint arXiv: 2401.01545, 2024.

[18] 符强,曾凡治,纪元法,等.针对室内高低动态环境的视觉 SLAM 算法研究[J].电子测量技术,2024,47(21):159-167.
FU Q, ZENG F ZH, JI Y F, et al. Research on visual SLAM algorithms for indoor environments with high and low dynamics [J]. Electronic Measurement Technology, 2024, 47(21):159-167.

作者简介

周若轩,硕士研究生,主要研究方向为智能网联汽车。
E-mail:ZhouRX921@163.com

张瑞乾(通信作者),博士,副教授,主要研究方向为智能制造工艺与装备开发和汽车车身设计与工程分析。
E-mail:ZRQ99@163.com

陈勇,博士,教授,主要研究方向为电动车辆系统动力学和新能源汽车线控底盘系统集成与控制。
E-mail:chenyong_jz@126.com

袁旭浩,硕士研究生,主要研究方向为智能网联汽车。
E-mail:15634516297@163.com

秦慧军,硕士研究生,主要研究方向为智能网联汽车。
E-mail:1792695946@qq.com