

融合混合注意力实例分割的视觉同步定位与建图算法

江浩玮¹, 陈孟元^{1,2*}, 袁学超³¹安徽工程大学电气工程学院, 安徽 芜湖 241000;²高端装备先进感知与智能控制教育部重点实验室, 安徽 芜湖 241000;³芜湖市固高自动化技术有限公司, 安徽 芜湖 241000

摘要 针对视觉同步定位与建图算法在遮挡情况下易受到干扰而导致定位误差较大且闭环检测精度较低等问题, 提出一种融合混合注意力实例分割的视觉同步定位与建图算法。该算法能够动态调整被遮挡物识别权重, 在出现遮挡情况时提升对被遮挡物的特征提取与识别能力。同时采用概率去误匹配算法去除错误匹配点对, 增加位姿求解及关键帧选取精度, 从而更好地修正机器人位姿、提高系统构图的准确率。通过 KITTI 公开数据集和真实场景对所提算法进行测试, 结果表明, 所提算法在闭环准确率上与 ORB-SLAM2 算法相比约提高 10.7%, 平移误差约减小 27.6%, 体现了良好的构图能力。

关键词 遥感; 同步定位与地图构建; 注意力机制; 实例分割; 目标识别; 闭环检测

中图分类号 TP391

文献标志码 A

DOI: 10.3788/LOP213265

Visual Simultaneous Localization and Mapping Algorithm Combining Mixed Attention Instance Segmentation

Jiang Haowei¹, Chen Mengyuan^{1,2*}, Yuan Xuechao³¹College of Electrical Engineering, Anhui Polytechnic University, Wuhu 241000, Anhui, China;²Key Laboratory of Advanced Perception and Intelligent Control of High-End Equipment, Wuhu 241000, Anhui, China;³Wuhu Googol Automation Technology Co., Ltd., Wuhu 241000, Anhui, China

Abstract The visual simultaneous localization and mapping algorithm is easy to be interfered under occlusion, which leads to large positioning error and low closed-loop detection accuracy. In this paper, a visual simultaneous localization and mapping algorithm based on mixed attention instance segmentation is proposed, which can dynamically adjust the recognition weight of the occluded object and improve the feature extraction and recognition ability of the occluded object in the case of occlusion. At the same time, a probabilistic mismatching removal algorithm is used to remove the wrong matching point pairs and increase the accuracy of pose solution and key frame selection. In this way, the robot pose can be better corrected and the accuracy of system composition can be improved. The proposed algorithm is tested through KITTI open dataset and real scenes, and the results show that the closed-loop accuracy of the proposed algorithm is about 10.7% higher than ORB-SLAM2 algorithm, and the translation error is about 27.6% lower, reflecting good composition ability.

Key words remote sensing; simultaneous localization and mapping; attentional mechanism; instance segmentation; target recognition; closed-loop detection

1 引言

同步定位与建图(SLAM)是指移动机器人进入未知环境时, 在没有环境先验输入情况下, 对所处位置进行定位并同时形成环境地图, 从而实现后续的导航任务^[1-3]。视觉 SLAM 利用视觉传感器获取图像, 从图像

信息中提取移动机器人位姿信息, 具有应用场景广泛、耗能较低的特点。不过视觉 SLAM 也存在缺陷, 累积漂移误差的修正受外界环境影响仍存在精度不高的情况^[4-6]。

为了在多场景下消除累积漂移误差对系统的影响, 加入闭环检测来不断修正移动机器人位姿, 旨在通

收稿日期: 2021-12-17; 修回日期: 2022-02-11; 录用日期: 2022-03-01; 网络首发日期: 2022-03-11

基金项目: 国家自然科学基金(61903002)、安徽省高校协同创新项目(GXXT-2021-050)、芜湖市科技计划项目(2020yf59)、安徽工程大学-鸠江区产业协同创新专项基金(2021cyxtb8)、安徽工程大学中青年拔尖人才项目

通信作者: *mychen@ahpu.edu.cn

对对不同时刻场景的辨别来判断移动机器人当前位置是否为之前某个相同位置^[7-8]。Konolige 等^[9]提出 Random Sample Consensus(RANSAC)算法提取特征点周围三维信息进行闭环检测,然而该算法在几何结构相同率高的情况下效果不尽如人意。Olson^[10]提出相关性扫描匹配(CSM)算法,通过每一帧与另一帧之间的相似性进行匹配,相似度高的则认定为同一场景。这种算法虽无需局部地图的存储,但计算过于复杂、计算量过大。Cummins 等^[11]在假定图像只存在旋转角度的情况下,只检测图像特征的几何一致来提高算法效率,但存在过分堆叠环境下,只利用特征一致性存在区分度弱、提供信息不充分等问题。Botterill 等^[12]使用词袋模型思想将每幅图片表示为某些单词的集合,进行特征分类。Mur-Artal 等^[13]提出的 ORB-SLAM2 算法,融合 Oriented Fast and Rotated Brief(ORB)算法良好的特征计算不变性,有效解决基于词袋模型可能存在的定位丢失问题。但上述所有算法在存在遮挡的情况下均难以提取目标物体的部分或全部特征。Arandjelović 等^[14]通过视觉词典中的词汇提取语义信息,将图像语义与闭环检测环节相结合,在闭环检测时引入语义信息代替单纯的几何信息来实现鲁棒的场景识别,规避了点特征参与匹配时可能存在的误识别,使匹配结果准确度大大提高。迅速发展的深度学习又给人们提供了另一种思路。Detect-SLAM^[15]系统使用 SSD^[16]目标检测网络对关键帧进行动态物体检测,解决存在的动态遮挡问题,但所需时间相对较多。Wang 等^[17]利用 YOLOv3 算法检测特定运动物体,相较 Detect-SLAM 算法运行时间显著减少。但以上两种算法只对物体大致位置进行框选,忽略了物体的边缘轮廓,提取精度较差。Yu 等^[18]提出的 DS-SLAM 算法在 ORB-SLAM2 框架上添加 SegNet^[19]提取目标的像素级信息进行分割,解决物体遮挡问题,提高了提取精度。上述算法均是

对遮挡物体类别进行识别,忽略了对被遮挡物特征的提取,使得整体可用特征较少,不利于 SLAM 算法的闭环,且检测到的物体可能是处于静止状态的,例如在 KITTI 数据集或停车场中,静态车辆可以用作 SLAM 算法闭环检测。若只考虑遮挡物类别,算法精度将会大大降低,甚至会导致闭环检测失败。

因此,所提算法改进实例分割算法 Mask-RCNN^[20]主干网络,并将两种软注意力(空间、通道)机制与视觉 SLAM 算法相结合,增大被遮挡物识别权重,提高 SLAM 算法闭环检测时对被遮挡物特征的提取与识别能力。同时,为避免因环境变化导致图像匹配错误率较高,而造成用于实例分割的关键帧冗余:引入概率去误匹配算法,计算每张图片匹配特征点对距离分布;去除极端分布匹配点对,减少选取关键帧时存在的误匹配得到更为优质的关键帧。改进后算法可以更好地实现相同场景下的识别与匹配,增强算法在遮挡环境下对被遮挡物体的识别能力,提升 SLAM 建图环节的稳定性。

2 算法结构

所提算法由混合注意力 Mask-RCNN、概率去误匹配算法以及定位建图等 3 个部分构成。将机器人探索未知环境时视觉传感器获取的图像信息输入融合混合注意力 Mask-RCNN 实例分割算法中对被遮挡物进行注意力加权,最终得出实例分割结果。概率去误匹配算法对输入图像进行特征提取与匹配,并根据匹配点对概率分布去除可能存在的误匹配。根据图像间位姿估计更新移动机器人轨迹地图,并与已有关键帧库中图像进行比对得出相似性评分,确定阈值将相似性较低的图片设定为关键帧进行闭环检测。若产生闭环,则用较前时刻相应位姿信息进行修正构图。系统框架图如图 1 所示,其中 ROI 为感兴趣区域的缩写。

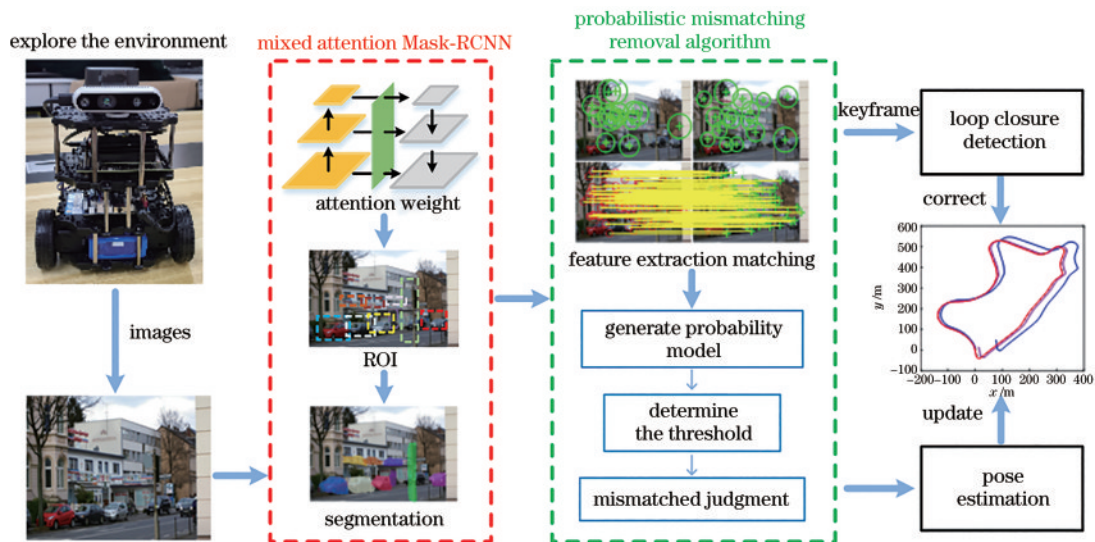


图 1 系统框架图

Fig. 1 System framework diagram

3 特征提取与匹配

3.1 融合混合注意力机制的 Mask-RCNN

基于深度学习的 Mask-RCNN 实例分割算法常被用来检测图像中存在的物体,并对已识别物体进行像素级分割,生成准确的图像语义信息。但在遮挡较为频繁的环境中,针对被遮挡物体的特征提取与识别的速度和准确度均会受到一定影响。本实验采用一种从空间、通道这 2 个不同维度软注意力相结合的混合注意力机制改进 Mask-RCNN 实例分割算法。注意力机

制源于人眼在观察事物时会本能地关注想关注的目标区域,因此视觉系统在处理人眼获取的图像信息时会投入更多的注意力去提取该区域的细节信息。使用该机制可以减少环境信息对识别分割结果的影响、降低算法识别过程中自然环境背景分配权重、提高被遮挡目标的分配权重,从而使算法更加聚焦于既定目标物体,避免局部混淆和部分遮挡所带来的检测准确度下降的问题,进而提高检测准确度。混合注意力 Mask-RCNN 算法框架图如图 2 所示。

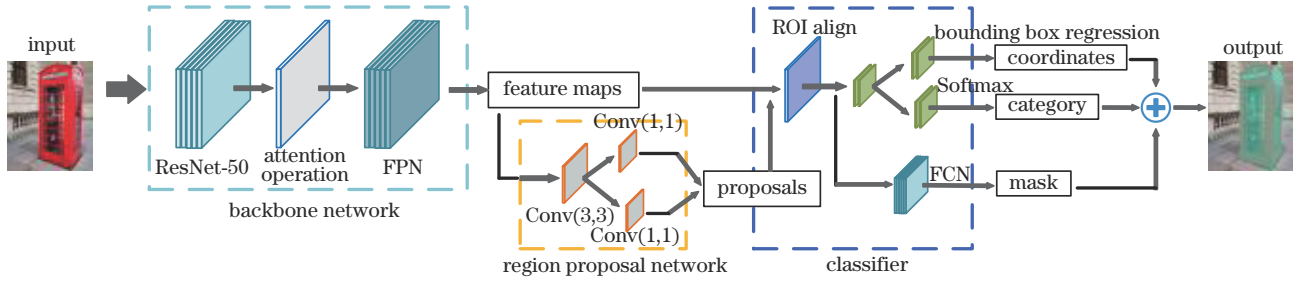


图 2 混合注意力 Mask-RCNN 算法框架图

Fig. 2 Framework diagram of mixed attention Mask-RCNN algorithm

所提实例分割算法 Mask-RCNN 的主干网络由深度残差网络(ResNet-50)及特征金字塔网络(FPN)^[21]构成。通过预训练得到特征图注意力权重分配法则,在空间和通道这 2 个层面对特征图所含特征信息进行注意力权重运算,从而实现空间和通道上的权重变化。使用特征提取网络 ResNet-50 进行特征提取,该网络对输入的视觉图像进行逐层卷积,获得含有不同维度特征的特征图 F_1, F_2, F_3, F_4, F_5 。考虑到在提高精度的同时减少引入注意力机制对算法运行速度的影响,只选取特征图 F_2, F_3, F_4, F_5 进行后续空间维度注意力运算得到特征图 W 和通道维度注意力运算生成具有不同维度权重分配的特征图 W' ,计算过程为

$$W_n = SA[\text{Conv}_{1 \times 1}(F_n)], \quad (1)$$

$$W'_n = CA(W_n), \quad (2)$$

式中: $SA(\cdot)$ 表示空间注意力权重分配; $CA(\cdot)$ 表示通道注意力权重分配。对得到的特征图 W'_5 使用 FPN 进行逐层上采样,并于对应特征图 W' 融合得到相应特征图 P ,生成各层特征图的计算过程为

$$P_n = f_u(P_{n+1}) \oplus W'_n, \quad (3)$$

式中: $f_u(\cdot)$ 表示上采样运算; \oplus 代表相应元素相加。同时,为便于后续对特征的标注适当缩小图像尺寸,对最深层特征图 P_5 进行池化运算得到 P_6 。FPN 对融合特征图 P 进行卷积去除上采样混叠,并将所有输出特征图输入后续网络完成主干网络工作。加入混合注意力机制后的所提主干网络结构如图 3 所示。

混合注意力机制首先使用空间注意力对输入特征图进行最大池化和平均池化,并将 2 次池化结果按通道融合生成 1 个有效特征描述符(表示整个通道的最

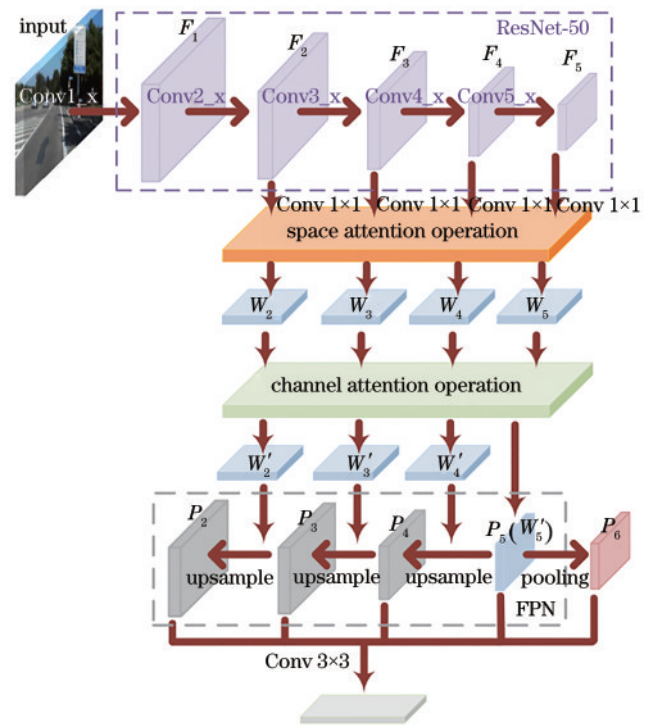


图 3 所提主干网络结构

Fig. 3 Proposed backbone network structure

大池化特征和平均池化特征),再通过卷积得到空间注意力权重 A_s 并与输入特征图连接实现空间注意力权重分配。生成公式如下:

$$A_s(F) = \sigma\left\{\text{Conv}_{7 \times 7}\left\{\left[\text{MaxPool}(F); \text{AvgPool}(F)\right]\right\}\right\}, \quad (4)$$

$$M_s(F) = A_s(F) \otimes F, \quad (5)$$

式中： F 为输入特征图； $\sigma(\bullet)$ 为激活函数层； $\text{MaxPool}(\bullet)$ 表示对特征图空间信息最大池化； $\text{AvgPool}(\bullet)$ 表示对

特征图空间信息平均池化； \otimes 表示按位相乘； $M_s(F)$ 为空间注意力结构输出。空间注意力结构如图 4 所示。

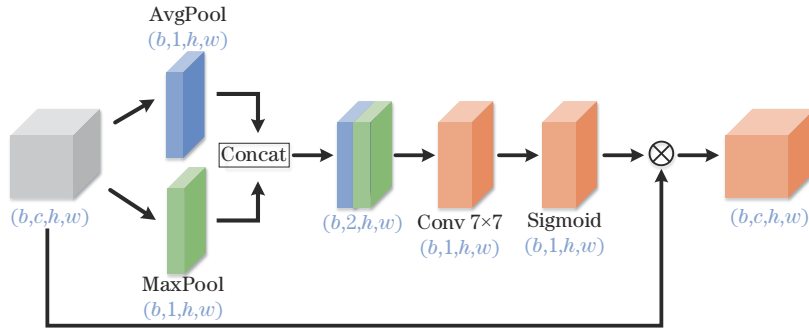


图 4 空间注意力结构

Fig. 4 Spatial attention structure

对具有空间注意力的特征图,使用通道注意力对特征图各特征通道分配相应注意力权重,通道注意力将输入特征图全局平均池化和最大池化的结果分别输入多层感知机中,得到通道注意力权重 A_c 。并与输入特征图相连接实现通道注意力权重分配,生成公式为

$$A_c(F) = \sigma\{\text{MLP}[\text{MaxPool}(F)] + \text{MLP}[\text{AvgPool}(F)]\}, \quad (6)$$

$$M_c(F) = A_c(F) \otimes F, \quad (7)$$

式中： F 为输入特征图； $\text{MLP}(\bullet)$ 代表多层感知机运算； $M_c(F)$ 为通道注意力结构输出。通道注意力结构如图 5 所示,改进后主干网络主要输出及参数设置如表 1 所示。

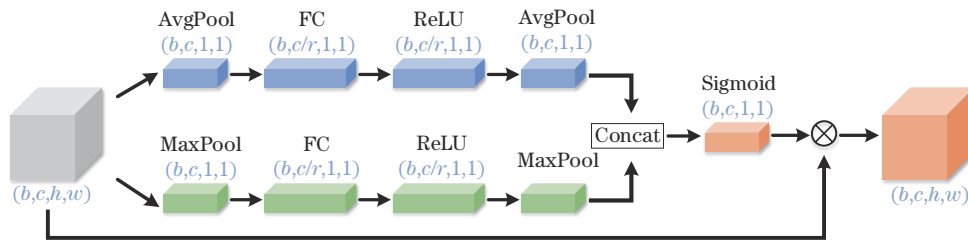


图 5 通道注意力结构

Fig. 5 Channel attention structure

表 1 混合注意力主干网络主要参数设置

Table 1 Main parameter of mixed attention backbone network

Number	Type	Area	Value
0	Conv1_x-Output	ResNet-50, Conv1_x	(64, H/4, W/4)
1	Conv2_x-Output	ResNet-50, Conv2_x	(256, H/4, W/4)
2	Conv3_x-Output	ResNet-50, Conv3_x	(512, H/8, W/8)
3	Conv4_x-Output	ResNet-50, Conv4_x	(1024, H/16, W/16)
4	Conv5_x-Output	ResNet-50, Conv5_x	(2048, H/32, W/32)
5	Upsample stride	FPN	2
6	Convolution kernel size	Spatial attention	7×7
7	Activation function	Spatial attention	Sigmoid
8	Activation function	Channel attention-MLP	ReLU
9	Activation function	Channel attention	Sigmoid

3.2 概率去误匹配算法

ORB 特征提取与匹配作为目前常用的特征提取算法之一,图像信息通过算法匹配时易受到特征相似性或外界环境干扰而产生大量误匹配,影响后续对关键帧的判断。由此,本实验引入概率学算法对产生的

误匹配进行剔除。去除误匹配流程图如图 6 所示。

通过对 2 帧图片匹配点对计算欧氏距离并分组统计,确定其概率学模型分布选取合理阈值从而剔除差异过大的匹配点对。将第 1 张图像 X 和第 2 张图像 X' 中提取的特征点定义为 2 个特征点集合 P 和 Q 。特征

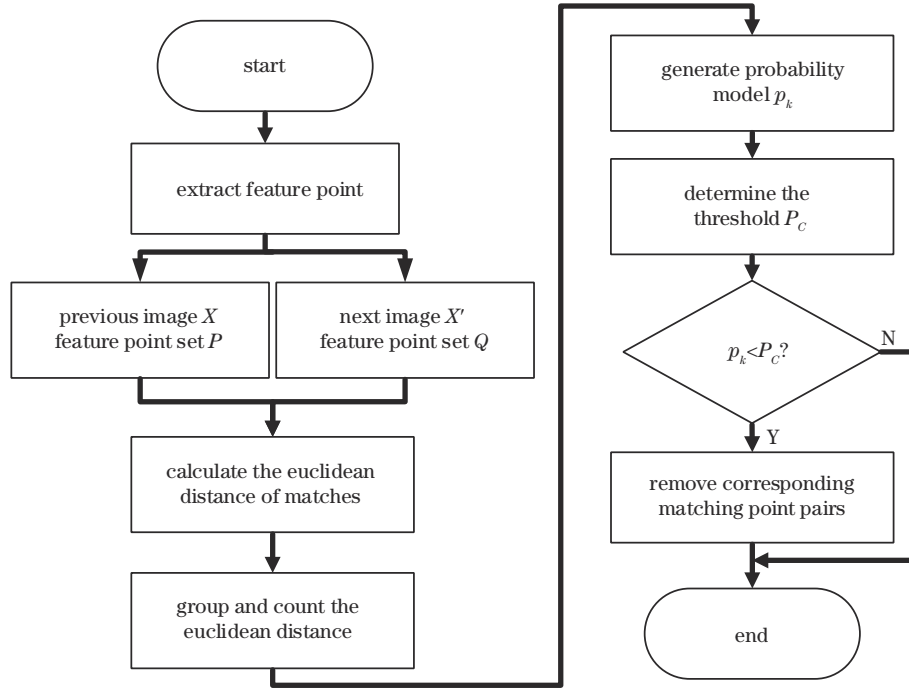


图 6 去除误匹配流程图

Fig. 6 Flow chart of mismatching remove

点集合表示如下:

$$P = \{X_1, X_2, X_3, \dots, X_{N-1}, X_N\}, \quad (8)$$

$$Q = \{X'_1, X'_2, X'_3, \dots, X'_{N-1}, X'_N\}, \quad (9)$$

式中: N 表示提取的特征点总个数。计算出相邻 2 帧图像中所有匹配点对欧氏距离 D_E , 根据选取的去除误匹配前匹配特征点总对数 N , 将全部欧氏距离分成 γ 个小组。 γ 的取值为

$$\gamma = \text{ceil}[\text{sqrt}(N)], \quad (10)$$

式中: $\text{ceil}(\cdot)$ 代表向上取整; $\text{sqrt}(\cdot)$ 代表根号运算。统计 D_E 中最大值 $D_{E\text{Max}}$ 和最小值 $D_{E\text{Min}}$, 则组距 L 表示为

$$L = \frac{D_{E\text{Max}} - D_{E\text{Min}}}{\gamma}, \quad 0 < D_{E\text{Min}} < L. \quad (11)$$

以组距 L 设立区间 $[0, L], [L, 2L], [2L, 3L], \dots, [\gamma L, (\gamma + 1)L]$, 将所有欧氏距离 D_H 进行分组。不同区间内出现欧氏距离 D_i 的次数假设为离散型随机分布 Y , 可能的全部取值为 $y_1, y_2, \dots, y_\gamma, y_{\gamma+1}$, 因此概率分布函数为

$$p_k = P\{Y = y_k\} = \frac{y_k}{N}, \quad k = 1, 2, \dots, \gamma, \gamma + 1. \quad (12)$$

根据欧氏距离概率分布, 设定概率阈值 $P_c = \text{Max}(p_k)$, 则 y_k 频段内的点是正确匹配的特征点进行保留。当 $P_k < P_c$ 时, 则认为此时频段内的点为误匹配, 需将该频段内的所有特征点舍去。

4 实验结果与分析

实验平台计算机处理器采用 Inter i9-10850K 芯片, GPU 为 RTX 3060 显卡, 显存为 12 GB, 内存为

64 GB, 在 Ubuntu 系统 Python 3.7 环境下进行实验。

4.1 KITTI 公开数据集测试

4.1.1 实例分割算法实验

依次选取 KITTI 数据集^[22]02 序列第 126 帧及 07 序列第 710 帧, 使用改进前算法和所提算法进行实例分割, 其中图 7(a)、(c) 为改进前算法实例分割提取结果, 图 7(b)、(d) 为所提算法实例分割提取结果。由于动态行驶车辆的遮挡, 改进前算法无法识别被遮挡的静止车辆, 而融合混合注意力机制后的实例分割算法在同一场景下能增加被遮挡物体权重, 从而更好识别部分遮挡的关键目标, 有效减少了环境因素对目标物体识别的影响, 提高了算法的精确性和适应性。

表 2 使用平均精准率 (AP) 作为评估标准进行对

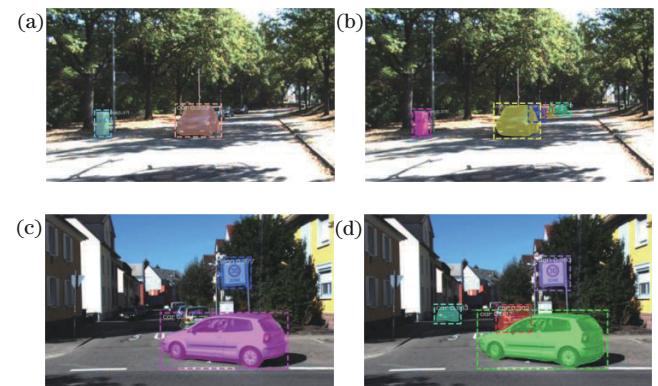


图 7 02 和 07 序列实例分割结果。(a)(c)改进前算法;

(b)(d)所提算法

Fig. 7 Instance segmentation results in 02 and 07 sequences.

(a)(c) Pre-improved algorithm; (b)(d) proposed algorithm

表 2 算法测试结果 AP 对比
Table 2 Comparison of algorithm test results in AP

Algorithm	Backbone	AP / %					
		AP	AP ₅₀	AP ₇₅	AP _s	AP _M	AP _L
Mask-RCNN	ResNet-50-FPN	33.4	54.9	35.3	14.7	35.2	50.1
Proposed algorithm	ResNet-50-MAM-FPN	34.9	57.5	36.9	15.3	36.9	52.5

比,因融合混合注意力机制后算法对物体边缘提取更为精准,使得所提算法 AP 较改进前算法提高约 4.6%。

4.1.2 概率去误匹配算法实验

在 KITTI 公开数据集 00 序列上将所提算法与其他算法进行对比实验,图 8(a)、(b)、(c) 分别为 Speeded Up Robust Features (SURF)、ORB 及所提算法特征匹配结果。其中,黄色实线为正确匹配点对,绿色虚线为错误匹配点对。从图 8 可以看出,与同类算法使用相同图片作为输入时,所提算法存在的误匹配

现象更少。

表 3 展示了 3 种算法在多个序列上提取的匹配特征点对数、有效匹配对数、有效匹配率、匹配所耗时间的对比情况。结果表明,所提算法在根据设定阈值去除不合理匹配点对后,在继承了 ORB 特征匹配算法匹配速度快的优势上,有效匹配率相较 ORB 算法有着大幅度提升。相比 SURF 算法,改进后算法在有效匹配率上几乎相似,但大大提高了匹配所需时间。

4.1.3 轨迹地图构建实验

为验证所提算法有效性,使用 KITTI 公开数据集对所提算法、ORB-SLAM2 和 DS-SLAM 这 3 种算法进行评估。图 9 为 KITTI 数据集不同序列中 3 种算法的运行结果及真实轨迹,表 4 为 3 种算法在不同序列下的平均距离误差、平均角度误差、闭环准确率。实验结果表明,所提算法相较其他 2 种算法在未闭环路径序列和闭环路径序列中均有着较小的运行轨迹误差。且在闭环路径序列的测试结果表明,所提算法的闭环准确率相较其他算法有较大提升。其中,07 序列集动态车辆干扰居多,遮挡情况出现更为频繁,本实验添加的注意力机制能更好地识别被遮挡物体,提取并分割被遮挡目标信息,提高识别与匹配精度,使得算法轨迹更接近于真实轨迹。

图 10 为 3 种算法在 07 序列上的每帧处理时间实验结果对比图。从图 10 中可以看出,由于对物体进行实例分割,所提算法相较 ORB-SLAM2 算法每帧处理时间有所增加,但每帧处理时间平均为 0.09 s,仍能满足系统实时性要求。

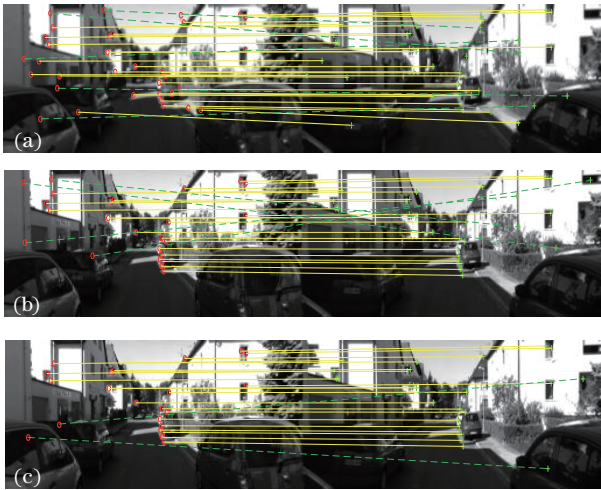


图 8 00 序列算法匹配结果。(a)SURF 特征匹配结果;(b) ORB 特征匹配结果;(c)所提算法特征匹配结果
Fig.8 Matching results in 00 sequence. (a) SURF feature matching results; (b) ORB feature matching results; (c) proposed algorithm feature matching results

表 3 KITTI 数据集有效匹配率和匹配用时对比
Table 3 Comparison of effective matching rate and matching time on KITTI

Sequence	SURF				ORB				Proposed algorithm			
	Matching pairs	Effective matching pairs	Effective matching rate / %	Matching time / s	Matching pairs	Effective matching pairs	Effective matching rate / %	Matching time / s	Matching pairs	Effective matching pairs	Effective matching rate / %	Matching time / s
00	1235	1002	81.1	0.1156	512	395	77.1	0.0089	494	392	79.4	0.0115
01	1254	1030	82.1	0.1172	490	381	77.8	0.0084	489	396	81.0	0.0097
06	1560	1264	81.0	0.1405	607	457	75.3	0.0094	524	424	80.9	0.0122
07	1438	1196	83.2	0.1281	530	405	76.4	0.0090	507	426	84.0	0.0119
09	1320	1088	82.4	0.1261	507	390	76.9	0.0086	501	412	82.2	0.0102
10	1480	1210	81.8	0.1364	552	424	76.8	0.0092	514	424	82.5	0.0121
Average	1381	1132	81.9	0.1273	533	409	76.7	0.0089	505	412	81.7	0.0113
Variance	14376	9476	0.57	0.083 × 10 ⁻³	4568	648	0.58	0.115 × 10 ⁻⁶	140	190	2.1	0.937 × 10 ⁻⁶

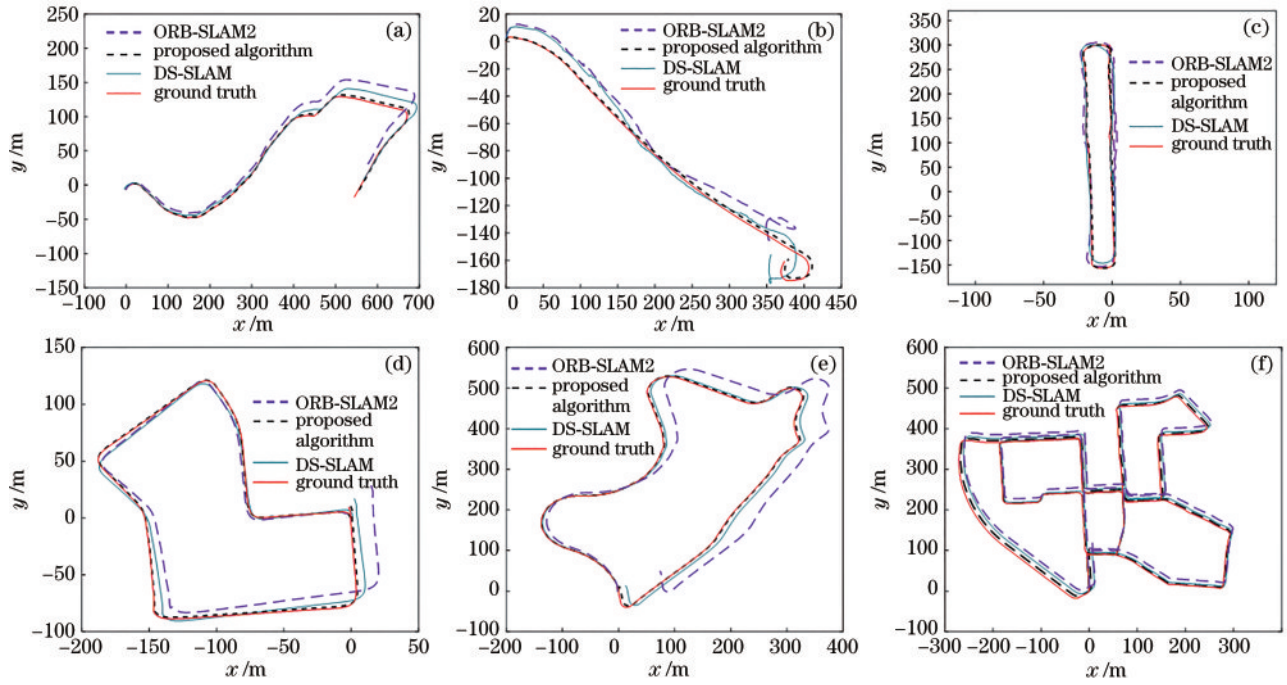


图 9 KITTI数据集不同序列运行轨迹。(a) 10序列;(b) 01序列;(c) 06序列;(d) 07序列;(e) 09序列;(f) 00序列
Fig. 9 Operating trajectories in different sequences on KITTI. (a) 10 sequence; (b) 01 sequence; (c) 06 sequence; (d) 07 sequence; (e) 09 sequence; (f) 00 sequence

表 4 KITTI数据集运行结果对比

Table 4 Comparison of operating results on KITTI

Sequence	ORB-SLAM2			DS-SLAM			Proposed algorithm		
	Average distance	Average angle	Precision rate of loop	Average distance	Average angle	Precision rate of loop	Average distance	Average angle	Precision rate of loop
	Error /m	Error /m	detection /%	Error /m	Error /m	detection /%	Error /m	Error /m	detection /%
10	3.15	1.55		2.62	0.94		2.01	0.82	
01	3.26	1.39		3.01	0.88		2.32	0.79	
06	2.99	1.57	77.9	2.51	0.79	82.3	2.38	0.73	86.4
07	3.05	1.30		2.72	0.61		2.53	0.50	
09	3.11	1.43		2.87	0.85		2.14	0.72	
00	3.64	1.24	76.6	2.94	0.97	80.4	2.54	0.87	84.7

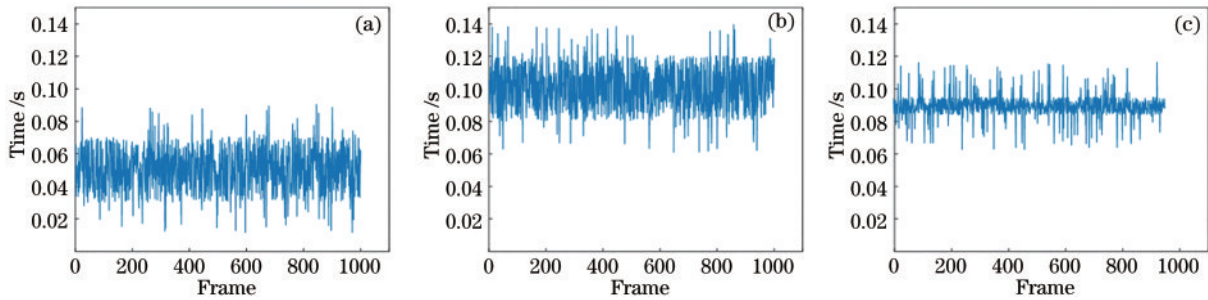


图 10 3种算法每帧处理时间。(a) ORB-SLAM2;(b) DS-SLAM;(c)所提算法
Fig. 10 Processing time per frame on three algorithms. (a) ORB-SLAM2; (b) DS-SLAM; (c) proposed algorithm

4.2 真实场景测试

本实验移动机器人使用 TurtleBot3 Burger, 其硬件外观如图 11 所示。携带深度相机为 Inter D435i。TurtleBot3 Burger 主要运行参数如表 5 所示。

选取学校实验室为室内实验地点, 大小为

10 m × 5 m, 真实场景如图 12(a) 所示, 平面布局图如图 12(b) 所示。其中, 黄色实体部分为会议桌, 阴影部分为桌子等其他障碍物。移动机器人围绕会议桌按“A→B→C→D”做“8”字形运动实时获取图像信息进行处理。



图 11 TurtleBot3 Burger
Fig. 11 TurtleBot3 Burger

4. 2. 1 实例分割算法实验

图 13 为移动机器人第 1 次运行至五角星位置处获取的一帧图像,其中图 13(a)、(b)分别为改进前算法和所提算法实例分割结果。从图 13 可以看出,由于遮挡情况出现较少,2 种算法分割结果相差无异。图 14 为第 2 次运行至五角星位置处时选取的图像,由于出

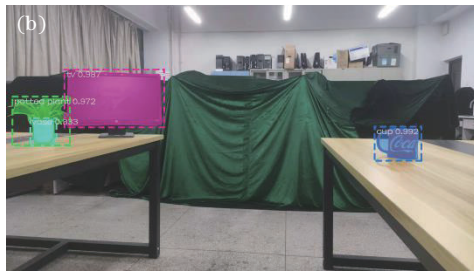


图 13 第 1 次五角星位置图像。(a)改进前实例分割结果;(b)所提算法实例分割结果

Fig. 13 Image of pentacle position for the first time. (a) Instance segmentation result of pre-improved algorithm; (b) instance segmentation result of proposed algorithm

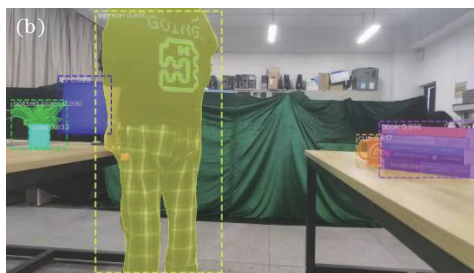
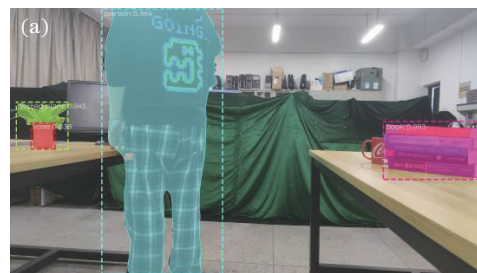


图 14 第 2 次五角星位置图像。(a)改进前实例分割结果;(b)所提算法实例分割结果

Fig. 14 Image of pentacle position for the second time. (a) Instance segmentation result of pre-improved algorithm; (b) instance segmentation result of proposed algorithm

4. 2. 2 轨迹地图构建实验

所提算法运行轨迹如图 15 所示。从图中运行轨迹可以看出,移动机器人运行到 D 点并向 B 点运动时,由于累计误差较大轨迹出现较为明显的漂移。随后从

表 5 运行参数设置
Table 5 Operation parameters setting

Variable	Parameter	Value
Running velocity	V_s	0.15 m/s
Rotating velocity	V_θ	2.1 rad/s
Range of directional	θ	$[0, 2\pi]$
Camera sampling frequency	H	30 frame/s

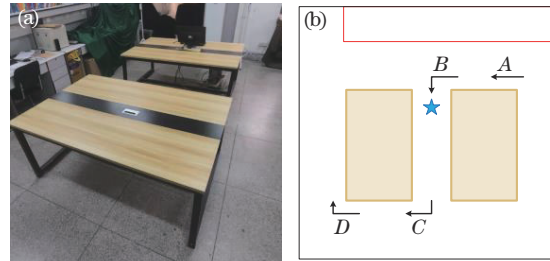


图 12 真实实验环境场景。(a)真实场景;(b)平面布局图
Fig. 12 Real experimental environment scene. (a) Real scene; (b) layout plan

现人为设置的遮挡情况,改进前算法无法提取被物体遮挡的显示器和水杯,而所提算法仍能对被遮挡物体进行分割,提高了闭环检测的准确率。

B 点向 C 点运行时,所提算法提取被遮挡物体检测到产生闭环,开始修正当前位姿。真实场景实验均方根误差、运行时间以及闭环准确率如表 6 所示。

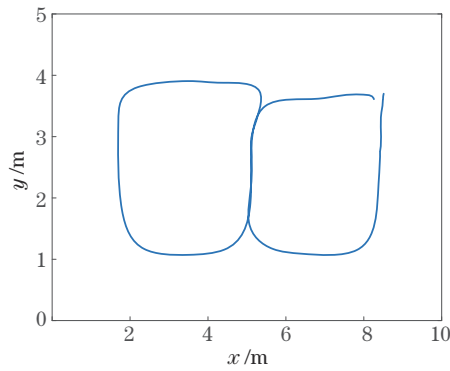


图 15 真实场景运行轨迹

Fig. 15 Operating trajectory in real scene

表 6 真实场景运行结果对比

Table 6 Comparison of running results in real scene

Index	RMSE /m	Time /s	Precision rate of loop detection /%
Proposed algorithm	0.52	270	85.3

5 结 论

移动机器人在复杂环境中的定位和闭环精度对其使用场景及性能有着较大影响,本实验提出一种融合混合注意力实例分割的视觉同步定位与建图算法。引入混合注意力机制动态为图像中被遮挡物体分配权重,增加遮挡场景下对被遮挡物体识别能力。并采用概率去误匹配算法对图像匹配点对进行筛选,去除不合理匹配,增加闭环检测所需关键帧的选取精度,提高闭环检测准确率,从而使移动机器人能够在遮挡场景中进行更为准确定位与建图。实验结果表明,采用所提算法的移动机器人在定位建图过程中能够更好地对被遮挡物体进行识别,修正运动轨迹,构建更为精准的轨迹图。后期在本实验研究结果之上,拟采用更高灵敏度传感器提高移动机器人运行过程中遇到动态障碍物的鲁棒性,最终实现机器人自主导航。

参 考 文 献

- [1] 张翠军, 张玉河. 基于 HHO 算法的 SLAM 闭环检测方法研究[J]. 激光与光电子学进展, 2021, 58(12): 1215006.
Zhang C J, Zhang Y H. Research on SLAM loop closure detection method based on HHO algorithm[J]. Laser & Optoelectronics Progress, 2021, 58(12): 1215006.
- [2] 徐雪松, 曾昱. 基于动态目标检测的视觉同步定位与地图构建算法[J]. 激光与光电子学进展, 2021, 58(16): 1615003.
Xu X S, Zeng Y. Visual simultaneous localization and mapping algorithm based on dynamic target detection[J]. Laser & Optoelectronics Progress, 2021, 58(16): 1615003.
- [3] 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.

- [4] 卢金, 刘宇红, 张荣芬. 面向动态场景的语义视觉里程计[J]. 激光与光电子学进展, 2021, 58(6): 0611001.
Lu J, Liu Y H, Zhang R F. Semantic-based visual odometry towards dynamic scenes[J]. Laser & Optoelectronics Progress, 2021, 58(6): 0611001.
- [5] 张峻宁, 苏群星, 刘鹏远, 等. 一种自适应特征地图匹配的改进 VSLAM 算法[J]. 自动化学报, 2019, 45(3): 553-565.
Zhang J N, Su Q X, Liu P Y, et al. An improved VSLAM algorithm based on adaptive feature map[J]. Acta Automatica Sinica, 2019, 45(3): 553-565.
- [6] 陈兴华, 蔡云飞, 唐印. 一种基于点线不变量的视觉 SLAM 算法[J]. 机器人, 2020, 42(4): 485-493.
Chen X H, Cai Y F, Tang Y. A visual SLAM algorithm based on line point invariants[J]. Robot, 2020, 42(4): 485-493.
- [7] Deschaud J E. IMLS-SLAM: scan-to-model matching based on 3D data[C]//2018 IEEE International Conference on Robotics and Automation, May 21-25, 2018, Brisbane, QLD, Australia. New York: IEEE Press, 2018: 2480-2485.
- [8] 占浩, 朱振才, 张永合, 等. 基于残差网络的图像序列闭环检测[J]. 激光与光电子学进展, 2021, 58(4): 0411003.
Zhan H, Zhu Z C, Zhang Y H, et al. Loop-closure detection using image sequencing based on ResNet[J]. Laser & Optoelectronics Progress, 2021, 58(4): 0411003.
- [9] Konolige K, Agrawal M. FrameSLAM: from bundle adjustment to real-time visual mapping[J]. IEEE Transactions on Robotics, 2008, 24(5): 1066-1077.
- [10] Olson E B. Real-time correlative scan matching[C]//2009 IEEE International Conference on Robotics and Automation, May 12-17, 2009, Kobe, Japan. New York: IEEE Press, 2009: 4387-4393.
- [11] Cummins M, Newman P. Appearance-only SLAM at large scale with FAB-MAP 2.0[J]. The International Journal of Robotics Research, 2011, 30(9): 1100-1123.
- [12] Botterill T, Mills S, Green R. Bag-of-words-driven, single-camera simultaneous localization and mapping[J]. Journal of Field Robotics, 2011, 28(2): 204-226.
- [13] Mur-Artal R, Tardós 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.
- [14] Arandjelović R, Zisserman A. Visual vocabulary with a semantic twist[M]//Cremers D, Reid I, Saito H, et al. Computer vision-ACCV 2014. Lecture notes in computer science. Cham: Springer, 2015, 9003: 178-195.
- [15] Zhong F W, Wang S, Zhang Z Q, et al. Detect-SLAM: making object detection and SLAM mutually beneficial [C]//2018 IEEE Winter Conference on Applications of Computer Vision, March 12-15, 2018, Lake Tahoe, NV, USA. New York: IEEE Press, 2018: 1001-1010.
- [16] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9905: 21-37.

- [17] Wang Z M, Zhang Q, Li J S, et al. A computationally efficient semantic SLAM solution for dynamic scenes[J]. *Remote Sensing*, 2019, 11(11): 1363.
- [18] Yu C, Liu Z X, Liu X J, et al. DS-SLAM: a semantic visual SLAM towards dynamic environments[C]//2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), October 1-5, 2018, Madrid, Spain. New York: IEEE Press, 2018: 1168-1174.
- [19] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481-2495.
- [20] He K M, Gkioxari G, Dollár P, et al. Mask R-CNN [C]//2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2980-2988.
- [21] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 936-944.
- [22] 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.