动态遮挡场景下基于改进Transformer实例分割的VSLAM算法

陈孟元[1,2,3],韩朋朋[1],刘金辉[1],张玉坤[1],江浩玮[1],丁陵梅[1]
(1. 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. 安徽工程大学产业创新技术有限公司,安徽芜湖 241000)

摘 要: 针对传统SLAM(Simultaneous Localization And Mapping)算法在动态遮挡场景下难以标记被遮挡物体,无法准确判断潜在物体运动状态以及剔除动态物体后特征点数量较少等问题,提出一种动态遮挡场景下基于改进Transformer实例分割的VSLAM算法(Improved Transformer instance segmentation under Dynamic occlusion VSLAM algorithm, ITD-SLAM). 本算法通过设计一种多注意力模块,引导模型关注被遮挡区域,同时改进相对位置编码优化被遮挡物体边界语义性,精确标记出潜在动态物体. 为减少动态物体对SLAM系统定位精度的影响,通过相机位姿估计、物体运动估计与物体运动判断三个步骤估计潜在动态物体运动状态,并剔除其中的动态物体. 根据网格流运动模型补全剔除区域的静态背景,并利用信息熵与交叉熵筛选修复区域特征点,补充高质量特征点用于相机位姿估计. 在公开数据集TUM和真实场景中进行验证,结果表明本文算法均方根误差与DynaSLAM相比减少22.94%,表现出了较好的构图能力.

关键词: 同时定位与地图构建;动态环境;物体遮挡;实例分割;运动判断;背景修复
基金项目: 国家自然科学基金(No.61903002);安徽省高校协同创新项目(No.GXXT-2021-050);安徽工程大学中青年拔尖人才项目;安徽工程大学引进人才科研启动基金
中图分类号: TP242.6 **文献标识码:** A **文章编号:** 0372-2112(2023)07-1812-14
电子学报URL:http://www.ejournal.org.cn **DOI:**10.12263/DZXB.20220310

Improved Transformer Instance Segmentation Under Dynamic Occlusion Based VSLAM Algorithm

CHEN Meng-yuan[1,2,3], HAN Peng-peng[1], LIU Jin-hui[1], ZHANG Yu-kun[1], JIANG Hao-wei[1], DING Ling-mei[1]
(1. *School of Electrical Engineering, Anhui Polytechnic University, Wuhu, Anhui* 241000, *China*;
2. *Key Laboratory of Advanced Perception and Intelligent Control of High-End Equipment(Ministry of Education), Wuhu, Anhui* 241000, *China*; 3. *Industry Innovation Technology Co., Ltd., Anhui Polytechnic University, Wuhu, Anhui* 241000, *China*)

Abstract: For traditional SLAM (Simultaneous Localization And Mapping) algorithms, it is difficult to mark occluded objects in dynamic scenes with occlusion, and is impossible to accurately judge the motion state of potential objects as well as the number of feature points after culling dynamic objects is small. This paper proposes a VSLAM algorithm based on improved transformer instance segmentation under dynamic occlusion (ITD-SLAM) in dynamic occlusion scenarios. By designing a multi-attention module, this algorithm guides the model to pay attention to the occluded area, and at the same time improves the relative position encoding to optimize the boundary semantics of occluded objects, and accurately mark potential dynamic objects. In order to reduce the influence of dynamic objects on the positioning accuracy of the SLAM system, the motion state of potential dynamic objects is estimated through three steps of camera pose estimation, object motion estimation and object motion judgment, and dynamic objects are eliminated. According to the grid flow motion model, the static background of the culled area is completed, and the feature points of the repair area are screened and repaired by information entropy, and the high-quality feature points are supplemented for camera pose estimation. Experimental results on the public datasets show that this algorithm has better composition ability with its root mean square error reduced by 22.94% when compared with DynaSLAM.

收稿日期:2021-03-25;修回日期:2022-10-11;责任编辑:李勇锋

Key words: simultaneous localization and mapping; dynamic environment; object occlusion; instance segmentation; motion judgment; background repair

Foundation Item(s): National Natural Science Foundation of China (No.61903002); Collaborative Innovation Projects in Anhui Universities (No.GXXT-2021-050); Anhui University of Engineering Young and Middle-aged Top Talent Project; Scientific Research Start-up Fund for Introducing Talents to Anhui University of Engineering

1 引言

同时定位与地图构建(Simultaneous Localization And Mapping, SLAM)是移动机器人在未知环境下,使用自身搭载的相机、激光雷达等传感器,建立局部环境地图,同时估计机器人自身所处位置[1~3]. 根据机器人使用传感器的不同,SLAM主要分为激光SLAM和视觉SLAM(Visual SLAM,VSLAM)两大类. 激光发生器在获取信息量、功耗、体积等方面的劣势限制了激光SLAM的应用,使得获取更丰富信息的VSLAM逐渐成为SLAM领域的一个热门研究方向[4~6].

目前较为成熟的视觉SLAM系统有ORB-SLAM(Oriented FAST and Rotated BRIEF SLAM)[7]和DSO(Direct Sparse Odometry)[8]等. 当前主流视觉SLAM系统都在大规模环境中实现了高精度的定位与构图,但其仅适用于静止不变的环境,当环境中存在运动物体时会被误认为相机存在运动,系统难以进行准确的位姿估计与构图[9]. 针对在动态场景中SLAM系统的定位问题,张慧娟等[10]利用特征点匹配对计算两帧图像初始变换矩阵评估线特征静态权重,最终使用静态直线特征进行视觉里程计环境探测,减小动态物体对SLAM系统定位的影响,该方法使用标准RANSAC优化剔除外点仅在静态环境下具有较大优势. 为进一步提高动态物体运动状态判断准确性,Sun等[11]提出了一种基于RGB-D相机的动态环境下SLAM去除运动物体算法,该算法利用具有静态假设的深度图像检测平面区域,将检测结果与使用重投影误差得到的粗略运动区域相结合,进一步细化物体运动区域. Zhang等[12]提出一种基于光流残差的动态分割与稠密融合的FlowFusion算法,该算法通过最小聚类的平均残差计算3D地图点的运动概率,实现对动态物体的移除. 利用几何约束实现动态点的剔除,能够使其在一定程度上适应动态环境,但其所构建的地图无法获取更高级的语义信息. 随着深度学习的不断发展,文献[13]提出的Mask-SLAM通过DeepLabV2算法分割图像并为每个像素分配对应的语义标签,根据获取的语义信息剔除其中的先验动态物体,但该方法未考虑被移动的物体(如椅子、书本). Xiao等[14]提出Dynamic-SLAM算法,该算法采用SSD进行先验目标检测,根据目标区域特征判断潜在动态物体是否运动,识别动态物体并剔除动态特征点,但该系统中难以精准分割动态区域. 文献[15]提出的Detect-SLAM算法在检测语义目标的同时对该目标特征点进行运动概率传播,在位姿估计时更新概率值并保留低于阈值的静态点,进一步提高动态物体的识别. 姚二亮等[16]提出一种鲁棒SLAM方法,通过YOLOv3算法初步提取动态物体所在区域,并添加边缘距离变换误差与光度误差细化动态物体区域,但该方法未能细化动态物体边界信息,易造成边缘特征点运动状态判断错误. Bescos等[17]提出利用实例分割网络Mask RCNN能够精确分割运动目标(人和车等)边界,并结合多视图几何检测潜在动态物体. 该算法依据两帧中同一关键点的变化角度判断该点是否属于动态点,但在特征跟踪过程中关键点变化角度易受到噪声影响,并且使用的实例分割网络面对遮挡环境分割效果较差. Yu等[18]提出基于RGB-D的DS-SLAM系统,该算法将SegNet语义分割网络与运动一致性相结合实现动态目标检测,显著提高SLAM系统在高动态环境下鲁棒性和稳定性,但该方法在剔除动态特征点后,易导致特征跟踪因特征点过少而失败.

综上所述,现有SLAM算法在动态遮挡环境中,存在实例分割网络无法准确标记被遮挡动态物体,难以准确判断潜在动态物体的运动以及剔除动态物体后特征点数量较少等问题,本文提出一种ITD-SLAM(Improved Transformer instance segmentation under Dynamic occlusion VSLAM algorithm)算法,本文实例分割网络通过融合多注意力机制[19~21]与相对位置编码[22~24]实现被遮挡动态物体的分割,将其中潜在动态物体采用潜在动态物体运动三步判断法以解决潜在运动物体的运动判断. 对剔除区域进行静态背景修复并提取该区域特征点,增加静态特征点数量. 在公开数据集TUM和真实场景中对本文算法进行验证. 实验结果表明ITD-SLAM算法与ORB-SLAM2[25],DS-SLAM,DynaSLAM算法相比,在动态物体判断方面有较大优势,表现出良好的构图能力.

2 系统整体框架

本文针对传统SLAM算法在动态遮挡场景下难以标记被遮挡物体,无法准确判断潜在物体运动状态以及剔除动态物体后特征点数量较少等问题,提出一种

改进 Transformer 实例分割的 ITD-SLAM 算法. 该算法包含潜在动态物体分割、潜在动态物体运动判断、背景修复与特征提取三个环节. 动态物体分割环节通过本文所提 APNET(the Attention mechanism and relative Position encoding into the instance NETwork)实例分割网络对图像信息中潜在动态物体进行实例分割,并提取图像特征点. 在潜在动态物体运动判断环节,通过潜在动态物体运动三步判断法获取动态物体并剔除动态点. 在背景修复与特征提取环节,利用先前图像帧信息对已剔除动态物体区域进行背景修复,并提取修复后区域特征点,为 SLAM 系统定位与构图提供更多信息. ITD-SLAM 算法框架图如图 1 所示.

3 潜在动态物体分割

为解决传统物体分割算法在进行动态物体分割时由于物体被遮挡造成分割失效或效果较差等问题,本文提出一种融合多注意力机制与相对位置编码的实例分割网络 APNET. 该网络包含提取图像局部特征的主干网络模块、增加被遮挡物体像素权重的多注意力模块以及提取全局特征和提高边界语义信息的融合相对位置编码 Transformer 模块. 主干网络模块在 ResNet-101 基础上增加特征金字塔网络 FPN[26]构建残差-特征金字塔网络,同时为更有效融合来自主干网络的不同类型特征,将具有较为丰富边界信息的特征图 P3 输入到融合相对位置编码 Transformer 模块,将具有丰富高维特征信息的特征图 $P4$、$P5$ 输入到多注意力模块. 多注意力模块使用通道与空间相结合的多重注意力机制,以加强被遮挡物体部分的像素权重,提高遮挡物体的识别率. 将相对位置编码融入 Transformer 网络构建融合相对位置编码 Transformer 模块,通过像素间距离重新分配像素权重,加强遮挡物体与被遮挡物体之间边界语义信息,提高被遮挡动态物体分割边界的精确度. 本文提出的 APNET 分割网络结构图如图 2 所示.

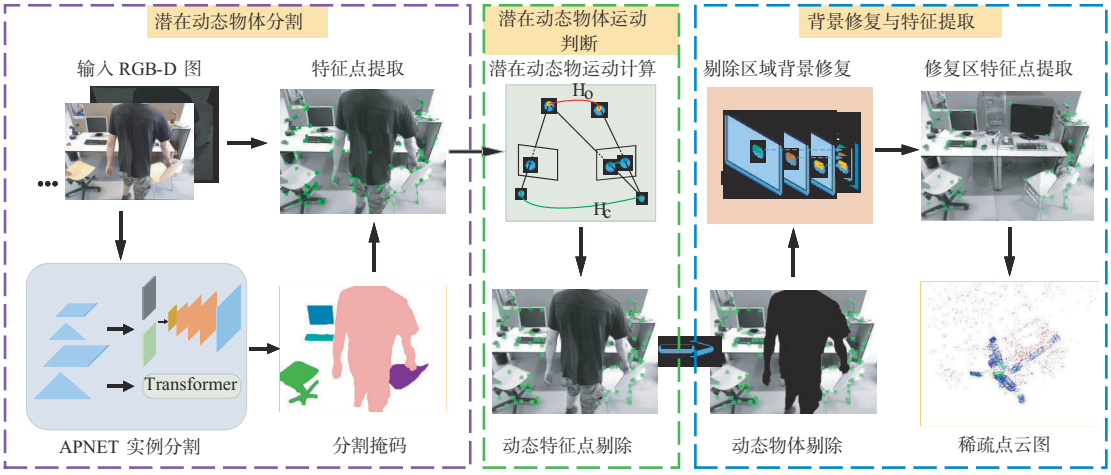


图 1 ITD-SLAM 算法系统框架图

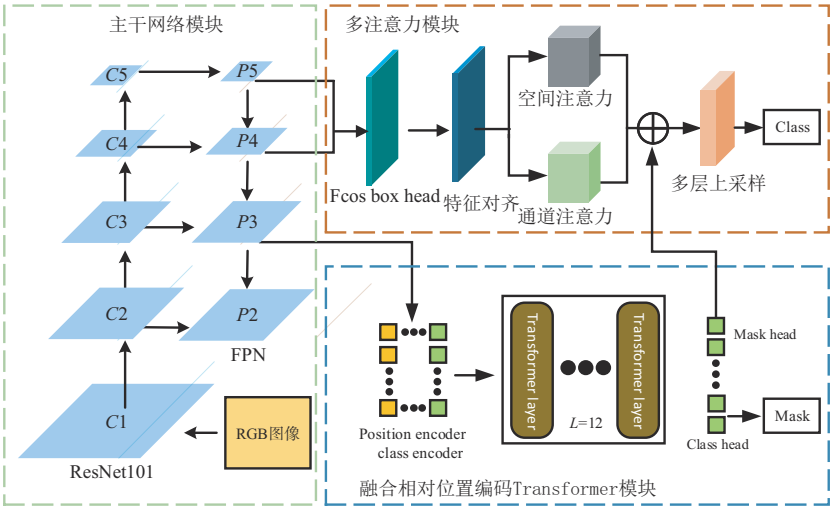


图 2 APNET 分割网络结构图

3.1 多注意力模块

视觉SLAM系统运行过程中被遮挡物体占据环境较大区域,若无法准确识别出其中潜在动态物体,将对SLAM系统稳定性造成干扰. 本文设计的多注意力模块由通道注意力机制和空间注意力机制两个子网络组成. 将特征图$\boldsymbol{F}$分别在通道维度和空间维度进行连接,并将获取的特征图$\boldsymbol{F}'$与特征图$\boldsymbol{F}''$进行concat融合得到输出$\boldsymbol{F}'''$. 多注意力模块利用通道与空间注意力机制结合生成更具有判别能力的特征表达,该模块不仅增大了感兴趣区域特征通道权重,还能获取感兴趣区域空间位置,充分突出被遮挡区域有效特征信息,同时避免了干扰信息影响. 如图3所示为多注意力模块结构图.

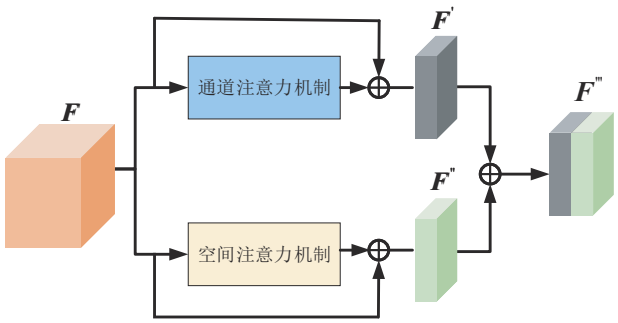


图3　多注意力模块结构图

如图4所示为本文通道注意力机制图. 通道注意力机制作用为将特征图中各层通道分配相应权重. 本算法将$H\times W\times C$特征图$\boldsymbol{F}$输入到通道注意力机制中,对特征图进行全局注意平均池化和最大池化操作,从而得到特征图每个通道的信息. 通过平均池化和最大池化获得的特征$\boldsymbol{F}_{\text{avg}}$与$\boldsymbol{F}_{\text{max}}$经过全连接层FC模块加强通道之间关联性,并对各通道权重进行重新分配,更好地对遮挡特征进行学习. 经过通道注意力机制获得的输出$\boldsymbol{f}_v$计算方式如式(1)所示.

$$\boldsymbol{f}_v=\sigma\left(\boldsymbol{P}\eta(\boldsymbol{F}_{\text{avg}}+\boldsymbol{F}_{\text{max}})\right) \tag{1}$$

其中,σ表示Sigmoid函数,η表示ReLU函数,$\boldsymbol{P}$为全连接层的参数,最后用$\boldsymbol{f}_v$对输入特征图$\boldsymbol{F}$进行逐层通道加权得到$\boldsymbol{F}'$.

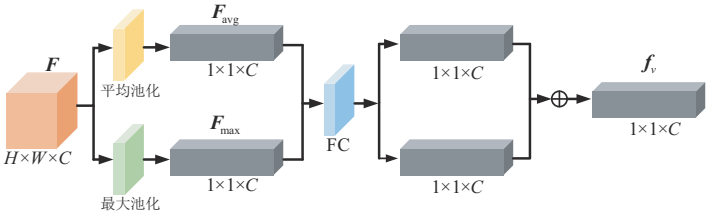


图4　通道注意力机制

本文使用的空间注意力机制如图5所示. 其主要作用为增加特征图中被遮挡位置像素值权重,经过学习不断调整各个权重值,进而引导网络关注遮挡部分所在区域. 将特征图$\boldsymbol{F}$输入空间注意力机制,通过平均池化和最大池化后进行concat融合形成$H\times W\times 2$特征图$\boldsymbol{f}_c$,再通过3×3×1卷积层和sigmoid函数得到空间注意图$\boldsymbol{f}_u$,其计算方式如式(2)所示. 将$\boldsymbol{f}_u$与输入特征图连接得到经空间注意力加权后的特征图$\boldsymbol{F}''$.

$$\boldsymbol{f}_u=\sigma(c(\boldsymbol{f}_c)) \tag{2}$$

其中,σ表示Sigmoid函数,c为3×3×1卷积网络. 主干网络与多注意力模块主要参数设置如表1所示.

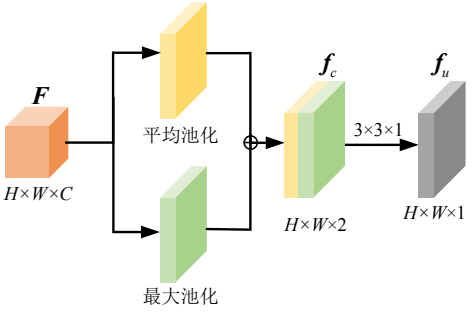


图5　空间注意力机制

3.2 融合相对位置编码Transformer模块

由于Transformer在特征提取和获取长距离像素间依赖关系方面有着优异表现,近期在计算机视觉领域受到极大关注. 传统的Transformer位置编码存在计算量过大、无法准确表征像素间相关性等问题. 本文引入从实数域映射到有限整数域的分段索引函数减少参数数量,降低计算成本. 同时,将改进相对位置编码融入Transformer网络,根据像素间距离重新分配像素权重,增加遮挡物体与被遮挡物体边界语义相关性,提高被遮挡物体分割准确性.

如图6所示为像素相对位置可视化图. 图像被分割成16×16的非交叉图像块,相同颜色对应位置共享相同编码. 本文采用D4棋盘法计算像素间相对距离,任意两个像素坐标$I_i(x_i,y_i)$, $I_j(x_j,y_j)$之间距离为$\text{D4}(I_i,I_j)=|x_i-x_j|+|y_i-y_j|$. 为减少位置编码计算量,本算法将像素间相对距离对应到有限集合中的整数,然后利用整数索引相对位置权重r_{ij},并在不同相对位置间共享权值. 本文从实数域映射到有限整数域的分段

表1 主干网络与多注意力模块主要参数设置

编号	层类型	区域划分	输出值
1	Conv$C1$	Resnet-101,$C1$	(64,H/4,W/4)
2	Conv$C2$	Resnet-101,$C2$	(256,H/4,W/4)
4	Conv$C3$	Resnet-101,$C3$	(512,H/8,W/8)
5	Conv$C4$	Resnet-101,$C4$	(1024,H/16,W/16)
6	Conv$C5$	Resnet-101,$C5$	(2048,H/32,W/32)
7	Conv$P2$	FPN,$P2$	(256,H/4,W/4)
8	Conv$P3$	FPN,$P3$	(512,H/8,W/8)
9	Conv$P4$	FPN,$P4$	(1024,H/16,W/16)
10	Conv$P5$	FPN,$P5$	(2048,H/32,W/32)
11	FC	Channel attention	(C,1,1)
12	Convolution Kernel Size	Spatial attention	(1,3,3)
13	Activation Function	Channel attention	Sigmoid
14	Activation Function	Spatial attention	Sigmoid

索引函数$g(x)$,计算方式如式(3)所示,通过分段索引函数极大减少计算参数数量,在相对距离较短的位置分配更多可学习的参数.

$$g(x)=\begin{cases}|x| & ,|x|\leqslant\alpha\\ \operatorname{sign}(x)\cdot\min\left(\beta,\left(\alpha+\dfrac{\ln(|x|/\alpha)}{\ln(\gamma/\alpha)}(\beta-\alpha)\right)\right) & ,|x|>\alpha\end{cases}\tag{3}$$

其中,sign(·)为符号函数.α决定分段点,β控制输出范围在$[-\beta,\beta]$,γ用于调整对数部分的曲率.

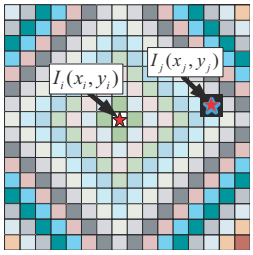


图6 像素相对位置可视化图

Transformer的核心是自注意力机制,它能建模输入列序之间关系,然而自注意力机制无法标记输入元素顺序,因此引入位置编码对Transformer尤为重要.目前Transformer在图像领域位置编码难以表征输入元素之间相关性,无法精确区分遮挡物体与被遮挡物体边界.本文提出一种相对位置编码算法,该算法使用点积计算输入元素之间相关性分数e_{ij},其计算方式如式(4)所示.

$$e_{ij}=\frac{(\boldsymbol{x}_i\boldsymbol{W}^Q)((\boldsymbol{x}_j\boldsymbol{W}^K)^{\mathrm{T}}+r_{ij}{}^{\mathrm{T}})}{\sqrt{d_z}}\tag{4}$$

其中,$r_{ij}\in\mathbb{R}^{d_z}$为二维相对位置权重,且与query交互;$\boldsymbol{W}^Q$、$\boldsymbol{W}^K$为可训练参数矩阵;$\boldsymbol{x}_i$、$\boldsymbol{x}_j$为图像块的输入,$d_z$表示输出矩阵维度.

采用softmax计算权重系数a_{ij},其计算方式如式(5)所示.

$$a_{ij}=\frac{\exp(e_{ij})}{\sum_{k=1}^{n}\exp(e_{ik})}\tag{5}$$

经过线性变换后,每个输出加权和z_i计算方式如式(6)所示.

$$z_i=\sum_{j=1}^{n}a_{ij}(\boldsymbol{x}_j\boldsymbol{W}^V+r_{ij}^V)\tag{6}$$

其中,$r_{ij}^V\in\mathbb{R}^{d_z}$为相对位置权重,$\boldsymbol{W}^V$为参数矩阵.

4 潜在动态物体运动判断

在潜在动态物体分割环节中使用实例分割网络仅能获取图像中潜在动态物体,而对于潜在动态物体是否运动缺少有效性判断.针对该问题本文引入潜在动态物体运动三步判断法,通过相机位姿估计、物体运动估计与物体运动判断三个步骤估计潜在动态物体运动状态,并剔除动态物体.潜在动态物体运动三步判断法示意图如图7所示.其具体步骤如下.

步骤1 相机位姿估计

机器人在实时运行过程中,在已知摄像机标定参数和特征点深度前提下,将空间中静态点m从参考帧F_{k-1}关联到后一帧F_k,其计算方式如式(7)所示.

$$m_k=\Delta(\boldsymbol{H}_c\Delta^{-1}I_{k-1})\tag{7}$$

其中,Δ和Δ^{-1}分别对应投影函数和反向投影函数,$\boldsymbol{H}_c\in\mathrm{SE}(3)$为相机姿态的相对变换矩阵;$I_{k-1}$为空间静态点投影到$F_{k-1}$中3D点$(m_{k-1},m_{z-1})$,其中$m_{k-1}$为该点在帧$F_{k-1}$中的2D像素坐标,$m_{z-1}$为该点在帧$F_{k-1}$中的深度;$m_k$为空间静态点投影到$F_k$中2D像素坐标.

最小化重投影误差$e(H_c)$计算方式如式(8)所示.

$$\begin{cases}e(H_c)=m_k'-\Delta(I_{k-1}\Delta^{-1}\boldsymbol{H}_c)\\ \boldsymbol{H}_c=\exp(\boldsymbol{h}_c)\end{cases}\tag{8}$$

其中,$\boldsymbol{h}_c\in\mathrm{SE}(3)$为相机姿态相对变换向量,$m_k'$为前一帧$F_{k-1}$中2D像素坐标$m_{k-1}$投影到当前帧坐标.

将$\tilde{\boldsymbol{h}}_c\in\mathbb{R}^6$定义为从SE(3)映射到$\mathbb{R}^6$的符号运算,最小二乘解$\tilde{\boldsymbol{h}}_c^*$如式(9)所示,通过高斯-牛顿算法优化求解得到相机位姿.

$$\tilde{\boldsymbol{h}}_c^*=\arg\min_{\tilde{h}_c}\boldsymbol{\Sigma}_i^n\rho_h(e(\boldsymbol{h}_c)^{\mathrm{T}}\boldsymbol{\Sigma}_p^{-1}e(\boldsymbol{h}_c))\tag{9}$$

其中,ρ_h为惩罚因子,$\boldsymbol{\Sigma}_p$为重投影误差的协方差矩阵,n为残差运算所需3D点投影至2D点数量.

步骤2 物体运动估计

依据相机运动估计物体位姿变换矩阵$\boldsymbol{H}_o\in\mathrm{SE}(3)$,将潜在动态对象建模为一个带有位姿变换矩阵$\boldsymbol{H}_o$的实体.将空间中动态点$\tilde{m}$从参考帧$F_{k-1}$关联到后一帧$F_k$,其计算方式如式(10)所示.

$$\begin{cases}\tilde{m}_k=\Delta(\boldsymbol{H}_c\boldsymbol{H}_o\Delta^{-1}I'_{k-1})\\ I'_{k-1}(\tilde{m}_{k-1},\tilde{m}_{z-1})\end{cases}\tag{10}$$

其中,$\tilde{m}_{k-1}$为深度图像帧中2D像素坐标,$\tilde{m}_{z-1}$为帧中坐标点深度.$\tilde{m}_k$为该点$\tilde{m}$在帧F_k中的2D点坐标.

同理,通过重投影误差与最小二乘法计算得到物体位姿变换矩阵$\boldsymbol{H}_o$,其计算公式如式(11)所示.

$$\begin{cases}e(\boldsymbol{H}_o)=\tilde{m}'_k-\Delta(\boldsymbol{H}_c\boldsymbol{H}_o\Delta^{-1}I'_{k-1})\\ \tilde{\boldsymbol{h}}_o^*=\underset{\tilde{h}_o}{\arg\min}\,\boldsymbol{\Sigma}_i^{n_b}\rho_h(e(\boldsymbol{h}_o)^{\mathrm{T}}\boldsymbol{\Sigma}_p^{-1}e(\boldsymbol{h}_o))\end{cases}\tag{11}$$

步骤3　物体运动判断

由于特征跟踪过程中存在噪声影响,仅通过物体运动变换矩阵$\boldsymbol{H}_o$难以判断物体是否运动.本文采用二维图像测量判断物体状态,假设特征点为静态点的投影点$\tilde{m}'_{k-1}$与其真实投影点$\tilde{m}_k$的像素距离d为动态视觉误差.鉴于图像中潜在动态物体含有丰富的像素点,因此计算图像潜在动态物体上像素点动态视觉误差d的中位数$\overline{L}$表示为物体动态视觉误差.$\overline{L}$的计算方式如式(12)所示.

$$\overline{L}=\mathrm{med}\{d\}=\mathrm{med}\{\|\tilde{m}'_{k-1},\tilde{m}_k\|\}\tag{12}$$

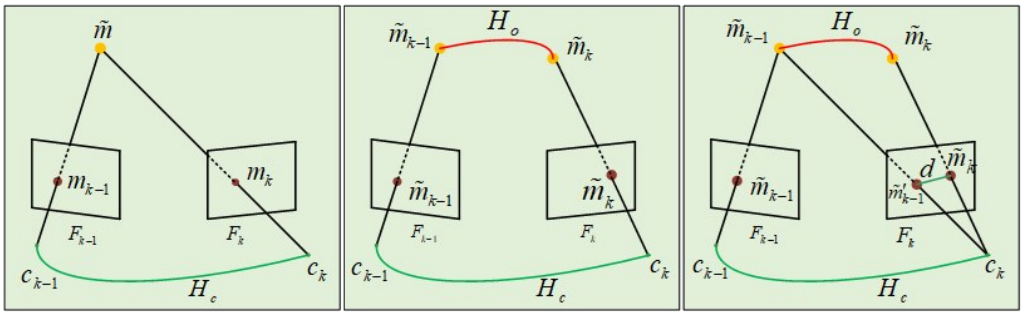
在非线性姿态优化阶段,利用高斯-牛顿迭代法可得到物体运动估计的不确定性误差为$\sum_x=(J^{\mathrm{T}}\boldsymbol{\Sigma}_r^{-1}J)^{-1}$.设定不确定性误差满足$k$维高斯分布,则它的微分熵$H(x_o)$计算方式如式(13)所示.

$$H(x_o)=\frac{1}{2}\log((2\pi e)^k|\boldsymbol{\Sigma}_{x_o}|)\tag{13}$$

微分熵可以被视为从光度残差最小化导出的姿态不确定性.具体来说,高熵值的三维运动观测将导致图像上某物体的较大移位,相反的是低熵值的观测将产生较小图像差异.基于此将物体动态偏差与一个由微分熵引导并随熵缓慢变大的动态阈值$\Delta d=f(H(x))$进行对比,若$\overline{L}>\Delta d$判断该物体为动态物体.潜在动态物体运动判断方法如算法1所示.

算法1 潜在动态物体运动判断

输入: 连续帧图像F_k
输出: 图像中动态物体与静态物体
参数:Δd
FOR new F_k available THEN
　标记潜在动态物体所在区域
　获取图像特征点
　　FOR 特征点属于图像 THEN
　　相机位姿估计与物体运动估计
　　计算潜在动态物体特征点动态偏差中位数
　　$\overline{L}=\mathrm{med}\{\|m'_{k-1},m_k\|\}$
　　计算特征点运动微分熵
　　$\Delta d=H(x_o)=\frac{1}{2}\log((2\pi e)^k|\boldsymbol{\Sigma}_{x_o}|)$
　　IF $\overline{L}>\Delta d$
　　特征点为动态物体
　　ELSE
　　特征点为静态物体
　　END IF
　　END FOR
END FOR



(a) 相机位姿估计　　(b) 物体运动估计　　(c) 物体运动判断

图7　物体运动判断

5　背景修复与特征提取

传统动态SLAM算法剔除动态物体后,用于位姿估计的特征信息随之减少,使得动态SLAM算法在动态物体较多的场景中位姿估计准确率较低,继而影响回环检测和完整全静态场景地图构建.因此,本文提出一种基于网格流模型的背景修复方法,借助先前帧静态信息完成动态区域剔除后当前帧RGB图像和深度图静态背景修复,并引入信息熵与交叉熵筛选修复区域高质量特征点.

基于网格流的背景修复示意图如图8所示,本算法以待修复关键帧F_t为起点,沿着箭头方向移动一个15帧的时间窗口.窗口内的关键帧图像根据两帧之间网格流依次与待修复关键帧图像对齐.当所有关键帧图像与待修复帧图像对齐的情况下,将待修复关键帧图像的缺失区域沿着箭头方向向前索引所对应像素.若

向前索引到一个对应像素，则直接进行缺失区域像素补全，若索引到多个对应像素值，则对索引到的像素取平均值再进行缺失区域像素补全.

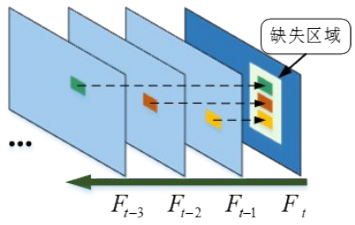


图8 基于网格流的背景修复

为避免修复区域特征点冗余，本文将修复区域分割成$N\times N$个子区域，并使用信息熵评价图像信息量，依据信息熵值判断显著特征区域，从ORB算法提取的特征点中筛选出高质量特征点. 本文设H_i为局部信息熵，$\overline{H}$为信息熵均值，若局部区域满足$H_i>\overline{H}$，则该区域为显著区域. 计算方式如式(14)所示.

$$\begin{cases} H_i=-\sum_{i=0}^{n-1} p_i\log_2 p_i \\ \overline{H}=\dfrac{1}{N^2}\sum_{i=0}^{n-1} H_i \end{cases} \tag{14}$$

其中，n为灰度级数，P_i为灰度值为i的像素占总像素点数的比值.

为进一步筛选征点，本文通过该区域内两点之间交叉熵与设定阈值的大小筛选区域特征点，交叉熵越小相似度越高，若两点交叉熵小于设定的阈值则舍弃该点. 交叉熵的计算公式如式(15)所示.

$$D(g,h)=-\sum_{i,j=1}^{u} g(p,i)\log_2(h(p',j)) \tag{15}$$

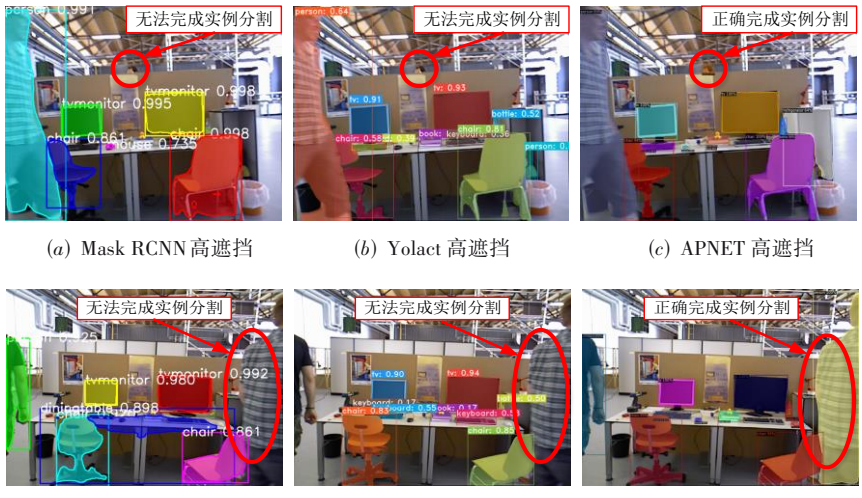
其中，u为描述符的维数，$g(p,i)$为特征点p的描述符，$h(p',j)$为特征点p'的描述符.

6 实验结果与分析

本文实验所用平台软硬件配置：CPU为Inter i9-12900K处理器，主频3.2 GHz，内存16 GB，GPU为RTX3090显卡，显存24 GB，系统为Ubuntu18.04.

6.1 潜在动态物体分割算法实验

为了验证本文所提算法APNET在高程度与低程度遮挡情况下实例分割性能，本文在TUM数据集中选择所需场景序列. Yolact、Mask RCNN和APNET三种算法在TUM据集下不同序列实例分割结果对比如图9所示，红色圈为本文所作辅助标记. 其中，图9(a)~(c)所示为场景中存在严重遮挡情况. 通过3幅图实例分割效果对比可知，由于Yolact与Mask RCNN缺少对严重遮挡目标区域特征增强网络层，导致在严重遮挡情况下易造成实例分割失效. 而APNET与前两种算法相比，APNET上层设计了一种多注意力模块，使APNET专注被遮挡信息同时抑制无用信息，从而提升实例分割效果. 如图9(d)~(f)为低遮挡环境下效果图. 由3幅图像对比可知，低程度遮挡所占区域较大，Yolact与Mask RCNN在网络层中只使用了卷积网络，而卷积网络仅针对局部特征有效，对于所提情况效果较差. APNET下层采用融合相对位置编码Transformer模块，其中相对位置编码根据像素间距离重新分配权重，增加被遮挡物体边界语义性. Transformer网络可以捕获全局像素间依赖关系，提高所占较大区域物体识别与分割准确率，因此APNET高程度与低程度遮挡情环境有突出效果.



(a) Mask RCNN高遮挡 (b) Yolact高遮挡 (c) APNET高遮挡

(d) Mask RCNN低遮挡 (e) Yolact低遮挡 (f) APNET低遮挡

图9 TUM数据集3种算法实例分割对比图

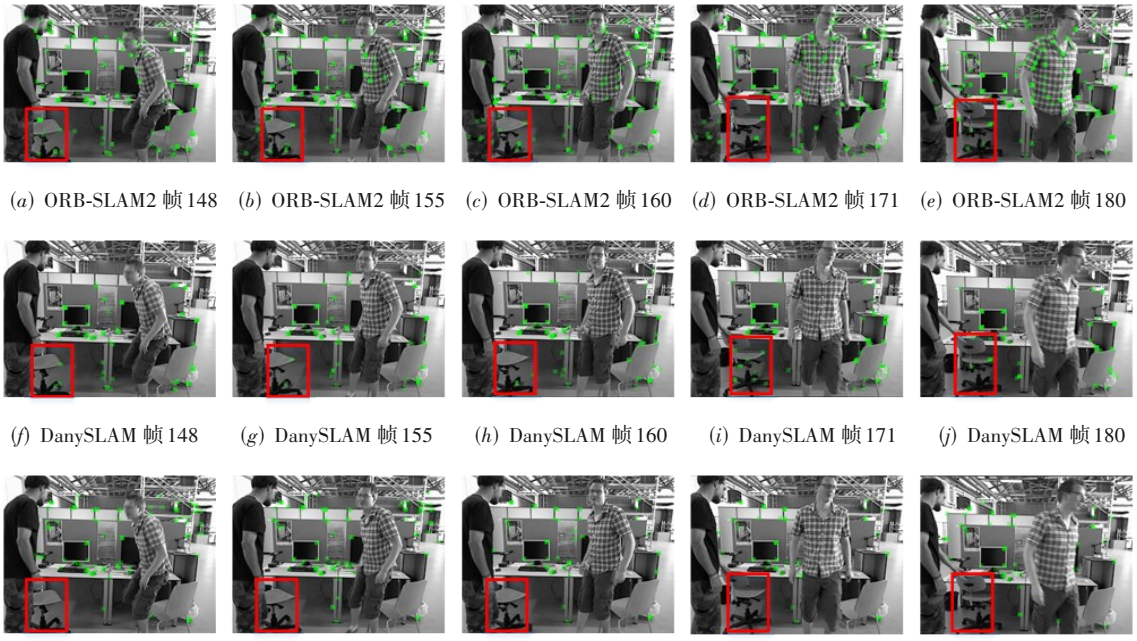
表2为3种算法平均精准率AP(Average Precision)对比表. 从表中数据可知,APNET算法平均精确度与Yolact算法相比平均精准率提高39.30%,与Mask RCNN相比提高20.12%.

表2　Yolact、Mask RCNN与APNET算法AP值对比　　单位:%

Algorithm	AP	AP_{50}	AP_{75}	AP_S	AP_M	AP_L
Yolact	31.2	50.6	32.8	12.1	33.3	47.1
Mask RCNN	35.7	58.0	37.8	15.5	38.1	52.4
APNET	39.5	60.1	42.2	16.7	58.4	70.1

6.2　潜在动态物体运动判断实验

本文选取TUM数据集中高动态序列对潜在动态物体状态判断与动态点剔除效果进行评估. 如图10所示为ORB-SLAM2、DynaSLAM和本文所提ITD-SLAM 3种算法剔除动态点效果对比图,其中红色矩形框为本文辅助框. 图10(*a*)~(*e*)所示为ORB-SLAM2提取特征点图,由于ORB-SLAM2算法使用了BA优化算法,可将少量外点剔除,而在高动态场景下,ORB-SLAM2算法明显无法继续剔除动态点. 图10(*f*)~(*j*)所示为DynaSLAM剔除动态点效果图,该算法依据两帧中同一关键点的变化角度判断该点是否属于动态点,但在特征跟踪过程中关键点变化角度易受到光线、角度等影响,所以DynaSLAM未能完全剔除被推动椅子上动态点. 图10(*k*)~(*o*)为ITD-SLAM剔除动态点效果图,该算法通过APNET确定潜在动态区域,并利用二维图像中动态视觉误差与微分熵相结合判断潜在动态点状态,因此本文算法能对潜在动态物体进行精准判断并剔除动态物体.



(*a*) ORB-SLAM2 帧148　(*b*) ORB-SLAM2 帧155　(*c*) ORB-SLAM2 帧160　(*d*) ORB-SLAM2 帧171　(*e*) ORB-SLAM2 帧180

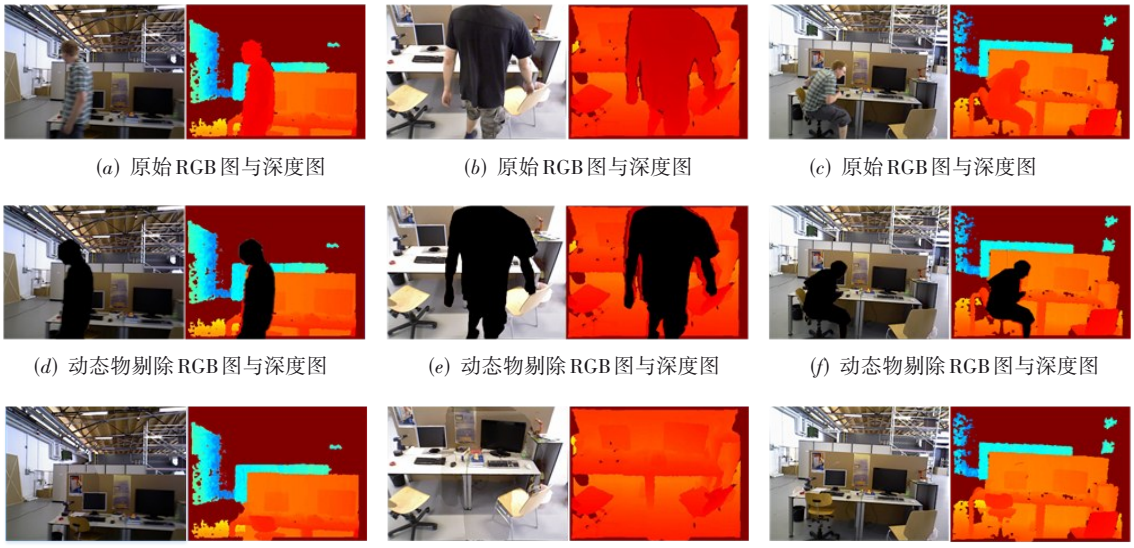
(*f*) DanySLAM 帧148　(*g*) DanySLAM 帧155　(*h*) DanySLAM 帧160　(*i*) DanySLAM 帧171　(*j*) DanySLAM 帧180

(*k*) ITD-SLAM 帧148　(*l*) ITD-SLAM 帧155　(*m*) ITD-SLAM 帧160　(*n*) ITD-SLAM 帧171　(*o*) ITD-SLAM 帧180

图10　3种不同算法动态点剔除效果对比图

6.3　背景修复与特征提取

为得到用于位姿估计与构图的完整静态场景,本文采用网格流模型对剔除区域进行补全. 背景修复过程效果图如图11所示,其中图11(*a*)~(*c*)所示为系统输入包含动态人物原始RGB图与原始深度图. 图11(*d*)~(*f*)所示为动态物剔除RGB图与深度图,本文算法将其中判断为动态的物体剔除,并根据先前帧信息修复图像缺失区域,从而建立完整全静态图像. 图11(*g*)~(*i*)所示为经过背景修复后RGB图像与深度图像,修复后图像只包含场景中原始静态背景. 本文选取TUM数据集验证本文算法ITD-SLAM在动态场景下特征提取效果. 如图12所示为TUM数据集上ORB-SLAM2、DynaSLAM和ITD-SLAM算法3种算法特征提取对比结果. 图12(*a*)~(*c*)为ORB-SLAM2特征提取效果,该算法未将动态物体上特征点剔除,如行人身上的特征点,导致算法受动态环境干扰严重,位姿估计与构图效果较差. 图12(*d*)~(*f*)为DynaSLAM特征提取效果,DynaSLAM剔除场景中动态物体的特征点,但未进行背景修复,可用静态特征点较少. 图12(*g*)~(*i*)为本文算法特征提取效果,本文算法通过剔除动态区域并修复剔除区域的静态背景,从而提取到更为丰富静态特征点,能构建更准确的轨迹图.



(*a*) 原始RGB图与深度图 (*b*) 原始RGB图与深度图 (*c*) 原始RGB图与深度图

(*d*) 动态物剔除RGB图与深度图 (*e*) 动态物剔除RGB图与深度图 (*f*) 动态物剔除RGB图与深度图

(*g*) 修复RGB图与深度图 (*h*) 修复RGB图与深度图 (*i*) 修复RGB图与深度图

图11 背景修复效果图

(*a*) ORB-SLAM2特征提取 (*b*) ORB-SLAM2特征提取 (*c*) ORB-SLAM2特征提取

(*d*) DynaSLAM特征提取 (*e*) DynaSLAM特征提取 (*f*) DynaSLAM特征提取

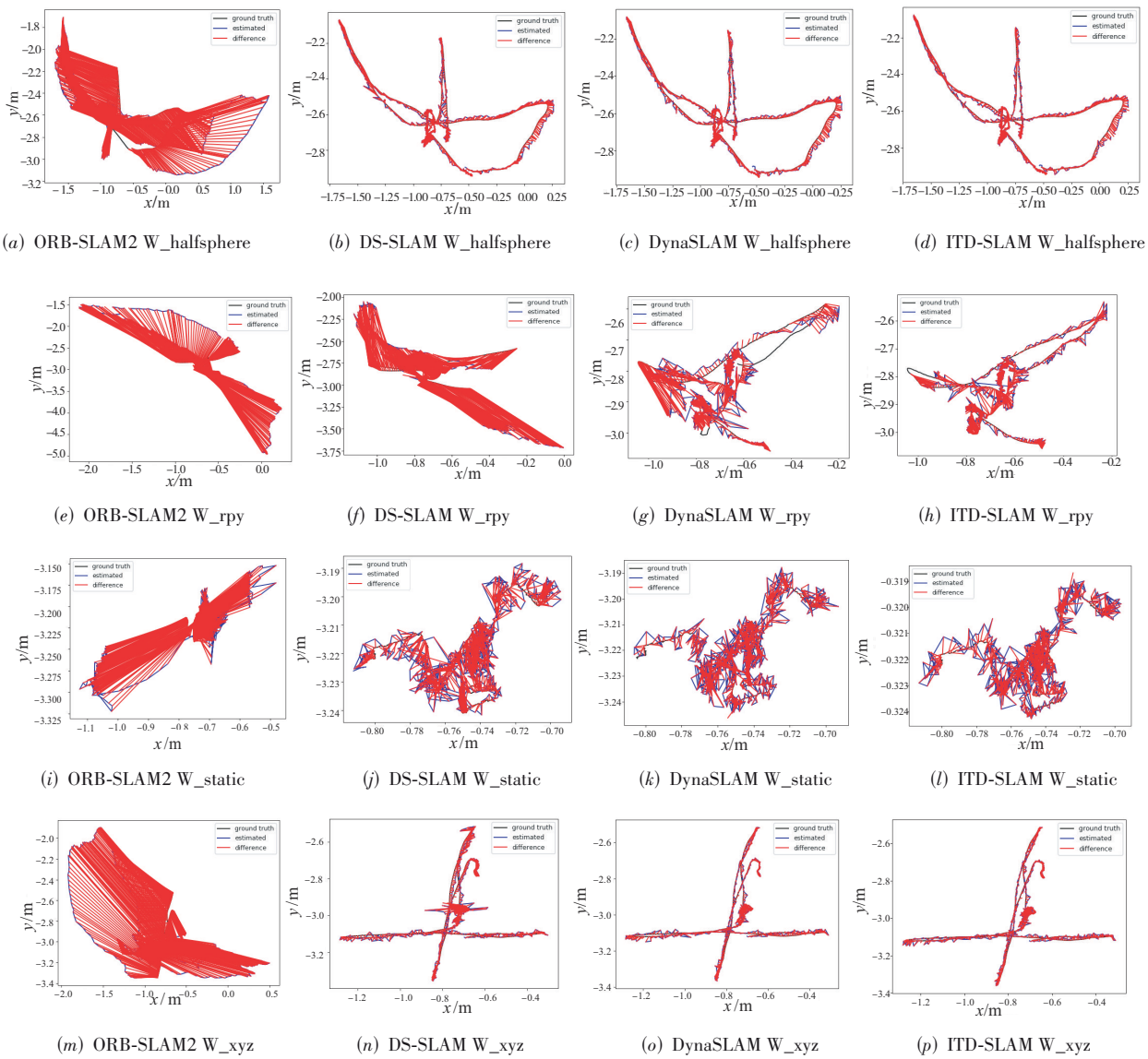
(*g*) ITD-SLAM特征提取 (*h*) ITD-SLAM特征提取 (*i*) ITD-SLAM特征提取

图12 3种不同算法特征提取结果

6.4 SLAM系统评估

本文选取TUM数据集中含有动态场景的W_half-sphere,W_rpy,W_xyz和W_static序列验证本文算法有效性. 图13分别为ORB-SLAM2,DS-SLAM,DynaSLAM和

本文所提ITD-SLAM 4种算法在不同场景下构建的轨迹图,图中黑色线表示相机运动真实轨迹,蓝色线为SLAM算法估计的相机运动轨迹,红色线为轨迹误差. 从图13可以看出,由于ORB-SLAM2算法无法识别环境中动态物体,构图效果较差. DS-SLAM和DynaSLAM算法通过实例分割对动态物体所在区域进行标记并剔除环境中动态物体,减小动态物体对全局一致性构图影响. 但无法在高遮挡环境下准确识别潜在动态物体且难以对潜在动态物体运动状态进行判断,同时这两种算法未进行动态物体剔除后的背景修复,用于位姿估计和构图的静态特征点较少,影响算法定位精度. 本文提出的ITD-SLAM算法由于融入APNET实例分割算法,能够识别环境中的高遮挡动态物体并剔除,且利用背景修复算法补全剔除动态物体后的静态背景,筛选出其中的高质量特征点用于位姿估计和构图,减少了高动态物体对构图的影响. 因此本文算法轨迹误差更小,构建的轨迹图更加接近真实轨迹.



(a) ORB-SLAM2 W_halfsphere　(b) DS-SLAM W_halfsphere　(c) DynaSLAM W_halfsphere　(d) ITD-SLAM W_halfsphere

(e) ORB-SLAM2 W_rpy　(f) DS-SLAM W_rpy　(g) DynaSLAM W_rpy　(h) ITD-SLAM W_rpy

(i) ORB-SLAM2 W_static　(j) DS-SLAM W_static　(k) DynaSLAM W_static　(l) ITD-SLAM W_static

(m) ORB-SLAM2 W_xyz　(n) DS-SLAM W_xyz　(o) DynaSLAM W_xyz　(p) ITD-SLAM W_xyz

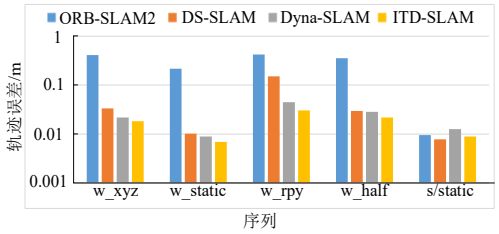
图13　4种不同算法轨迹图

本文使用均方根误差(RMSE)衡量绝对轨迹误差和相对轨迹误差验证ITD-SLAM的鲁棒性. 均方根数值越低代表系统鲁棒性越高. 图14为4种方法均方根误差对比结果. 由图14(*a*)可以看出,ITD-SLAM绝对轨迹误差RMSE相比于ORB-SLAM2、DS-SLAM、DynaSLAM分别减少77.87%、32.90%、22.94%. 图14(*b*)可知,ITD-SLAM相对平移误差RMSE相比于ORB-SLAM2、DS-SLAM、DynaSLAM分别减少77.08%、34.00%、24.65%. 图14(*c*)为相对旋转误差RMSE对比,ITD-SLAM相比于ORB-SLAM2、DS-SLAM、DynaSLAM

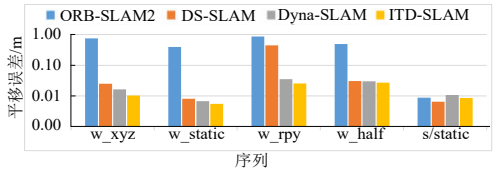
分别减少73.29%、19.98%、5.49%.

6.5 真实场景测试

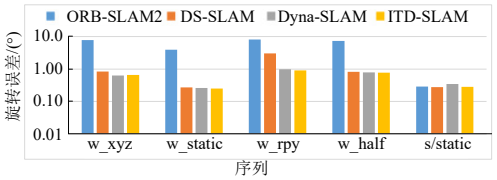
在真实动态场景中对本文算法有效性进行验证，在算法中语义分割作为系统的输入是由PC端离线生成. 实验平台为Husky轮式移动机器人，其硬件配置为：CPU为i7-10875H处理器，内存8 GB，GPU为GTX1080，操作系统为Ubuntu18.04. 机器人硬件外观如图15(*a*)所示，主要参数设置如表3所示.



(*a*) 绝对轨迹误差RMSE



(*b*) 相对平移误差RMSE



(*c*) 相对旋转误差RMSE

图14 4种不同算法对比结果

真实实验环境如图15(*b*)所示，大小为12 m × 5 m. 图15(*c*)为真实场景平面布局，其中黄色部分为工作台，A-B段为机器人运动路线，E-F段和C-D段为行人往返运动路线.

表3 主要参数设置

变量名称	参数	数值
运行速度	Vs	1 m/s
转速	V_θ	0.48 rad/s
方向变化范围	Θ	[0,2π]
相机采样频率	H	30 fps

6.5.1 潜在动态物体分割算法实验

图16所示为选取移动机器人运行过程中获取的某两帧图像信息，其中图16(*a*)为低遮挡场景中实例分割效果图，图16(*b*)为高遮挡场景下的实例分割效果. 从图中可以看出，本文算法APNET在低遮挡和高遮挡场景中均可对潜在动态物体准确分割，验证了本文算法在遮挡场景中对潜在动态物体分割的有效性.

6.5.2 潜在动态物体运动判断

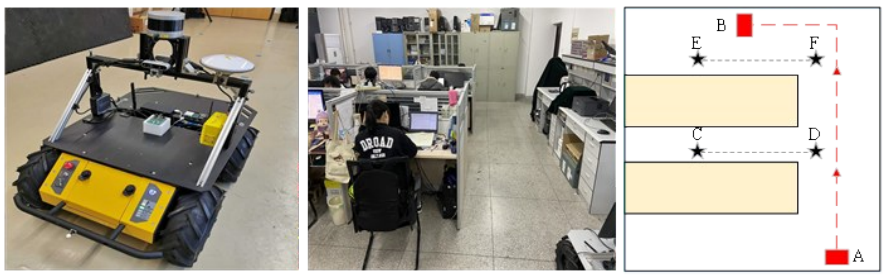
图17所示为被遮挡物体运动判断结果，其中，红色框为本文所作辅助框. 移动机器人在动态场景中运动时，场景中的动态物体会对系统位姿估计产生较大影响，而本文算法通过潜在动态物体所在区域进行分割，并对潜在动态物体进行运动判断，剔除场景中动态人物与被人推动的椅子，增加系统在动态场景中鲁棒性.

6.5.3 背景修复与特征点提取

ITD-SLAM算法特征提取结果如图18所示. 由于在静态特征点较少的情况下，剔除动态物体后可用静态点过少，导致SLAM算法位姿估计与构图效果较差. 而ITD-SLAM算法由于加入剔除区域背景修复，并对修复后的区域通过筛选高质量特征点，提高了机器人位姿估计精度与构图准确性.

6.5.4 轨迹地图构建

如图19所示为本文算法与ORB-SLAM2算法定位轨迹对比图. 由图19可以看出，在动态遮挡环境下本文算法ITD-SLAM轨迹相较于ORB-SLAM2轨迹较接近真实轨迹. 本文算法由于加入ATNET实例分割算法检测场景中潜在动态物体，并利用潜在运动物体三步判



(*a*) Husky轮式移动机器人 (*b*) 真实实验环境 (*c*) 平面布局图

图15 实验平台及真实实验环境场景

断法估计出其中的动态物体,剔除动态物体区域采用网格流模型修复并提取特征点,增加了用于位姿估计与构图特征点数量. 因此本文算法在动态遮挡环境下具有较高的鲁棒性.

(*a*) 低遮挡实例分割效果　(*b*) 高遮挡实例分割效果

图 16　实例分割效果图

(*a*) 低遮挡运动判断效果图　(*b*) 高遮挡运动判断效果图

图 17　物体运动判断结果

(*a*) 低遮挡特征提取效果图　(*b*) 高遮挡特征提取效果图

图 18　特征提取效果

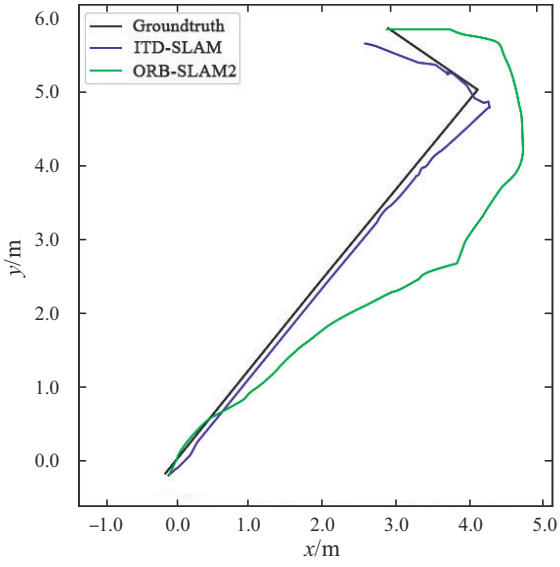


图 19　两种算法轨迹图对比

7　结论

为提高移动机器人在动态场景中鲁棒性,本文提出了一种ITD-SLAM算法,该算法具有以下优点.(1)针对现有实例分割网络在具有遮挡场景下难以正确标记被遮挡物体,提出一种融合多注意机制与相对位置编码的实例分割网络APNET,提高被遮挡动态物体的实例分割能力.(2)引入潜在动态物体运动三步判断法解决了潜在动态物体运动状态判断问题.(3)采用网格流与运动模型补全剔除区域图像,并提取该区域高质量特征点,为系统位姿估计和稀疏点云地图构建提供更多特征信息. 使用公开数据集TUM对本文所提算法进行验证,结果表明本文所提算法与ORB-SLAM2、DS-SLAM、DynaSLAM算法相比,在定位精度方面有较大优势,并体现出了良好的构图能力. 下一步将在本文研究的基础上,融合IMU(惯性测量单元)数据,为低纹理环境下的相机位姿求解添加约束项,进一步提升算法精度和鲁棒性.

参考文献

[1] 陈孟元, 丁陵梅, 张玉坤. 基于改进关键帧选取策略的快速PL-SLAM算法[J]. 电子学报, 2022, 50(3): 608-618.
CHEN M Y, DING L M, ZHANG Y K. Fast PL-SLAM algorithm based on improved keyframe extraction strategy [J]. Acta Electronica Sinica, 2022, 50(3): 608-618. (in Chinese)

[2] 李博洋, 刘思健, 崔明月, 等. 基于最小回环检测的多车协同SLAM框架[J]. 电子学报, 2021, 49(11): 2241-2250.
LI B Y, LIU S J, CUI M Y, et al. Multi-vehicle collaborative SLAM framework for minimum loop detection[J]. Acta Electronica Sinica, 2021, 49(11): 2241-2250. (in Chinese)

[3] 高兴波, 史旭华, 葛群峰, 等. 面向动态物体场景的视觉SLAM综述[J]. 机器人, 2021, 43(6): 733-750.
GAO X B, SHI X H, GE Q F, et al. A survey of visual SLAM for scenes with dynamic objects[J]. Robot, 2021, 43 (6): 733-750. (in Chinese)

[4] 曹剑飞, 余金城, 潘尚杰, 等. 采用双视觉里程计的SLAM位姿图优化方法[J]. 计算机辅助设计与图形学学报, 2021, 33(8): 1264-1272.
CAO J F, YU J C, PAN S J, et al. A SLAM pose graph optimization method using dual visual odometry[J]. Journal

of Computer-Aided Design & Computer Graphics, 2021, 33(8): 1264-1272. (in Chinese)

[5] 陈宝华, 邓磊, 陈志祥, 等. 基于即时稠密三维重构的无人机视觉定位[J]. 电子学报, 2017, 45(6): 1294-1300.
CHEN B H, DENG L, CHEN Z X, et al. Instant dense 3D reconstruction based UAV vision localization [J]. Acta Electronica Sinica, 2017, 45(6): 1294-1300. (in Chinese)

[6] 罗会兰, 陈鸿坤. 基于深度学习的目标检测研究综述[J]. 电子学报, 2020, 48(6): 1230-1239.
LUO H L, CHEN H K. Survey of object detection based on deep learning[J]. Acta Electronica Sinica, 2020, 48(6): 1230-1239. (in Chinese)

[7] MUR-ARTAL R, MONTIEL J M M, TARDÓS J D. ORB-SLAM: A versatile and accurate monocular SLAM system [J]. IEEE Transactions on Robotics, 2015, 31(5): 1147-1163.

[8] ENGEL J, KOLTUN V, CREMERS D. Direct sparse odometry[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 40(3): 611-625.

[9] 刘剑锋, 孙力帆, 普杰信, 等. 基于刚性约束的双移动机器人协同定位[J]. 电子学报, 2020, 48(9): 1777-1785.
LIU J F, SUN L F, PU J X, et al. Cooperative localization in a team of two mobile robots based on rigid constraints [J]. Acta Electronica Sinica, 2020, 48(9): 1777-1785. (in Chinese)

[10] 张慧娟, 方灶军, 杨桂林. 动态环境下基于线特征的RGB-D视觉里程计[J]. 机器人, 2019, 41(1):75-82.
ZHANG H J, FANG Z J, YANG G L. RGB-D visual odometry in dynamic environments using line features[J]. Robot, 2019, 41(1): 75-82. (in Chinese)

[11] SUN Y X, LIU M, MENG M Q H. Motion removal for reliable RGB-D SLAM in dynamic environments[J]. Robotics and Autonomous Systems, 2018, 108: 115-128.

[12] ZHANG T W, ZHANG H Y, LI Y, et al. FlowFusion: dynamic dense RGB-D SLAM based on optical flow[C]// 2020 IEEE International Conference on Robotics and Automation (ICRA). Piscataway: IEEE, 2020: 7322-7328.

[13] KANEKO M, IWAMI K, OGAWA T, et al. Mask-SLAM: Robust feature-based monocular SLAM by masking using semantic segmentation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway: IEEE, 2018: 371-3718.

[14] XIAO L H, WANG J G, QIU X S, et al. Dynamic-SLAM: Semantic monocular visual localization and mapping based on deep learning in dynamic environment[J]. Robotics and Autonomous Systems, 2019, 117: 1-16.

[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 (WACV). Piscataway: IEEE, 2018: 1001-1010.

[16] 姚二亮, 张合新, 宋海涛, 等. 基于语义信息和边缘一致性的鲁棒SLAM算法[J]. 机器人, 2019, 41(6): 751-760.
YAO E L, ZHANG H X, SONG H T, et al. Robust SLAM algorithm based on semantic information and edge consistency[J]. Robot, 2019, 41(6): 751-760. (in Chinese)

[17] 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.

[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). Piscataway: IEEE, 2019: 1168-1174.

[19] Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision. Cham: Springer International Publishing, 2018: 3-19.

[20] GEHRING J, AULI M, GRANGIER D, et al. Convolutional sequence to sequence learning[C]//Proceedings of the 34th International Conference on Machine Learning - Volume 70. New York: ACM, 2017: 1243-1252.

[21] CHEN L, ZHANG H W, XIAO J, et al. SCA-CNN: Spatial and channel-wise attention in convolutional networks for image captioning[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2017: 6298-6306.

[22] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers[C]//Computer Vision—ECCV 2020. Cham: Springer International Publish-

ing, 2020: 213-229.

[23] LIU Z, LIN Y T, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2022: 9992-10002.

[24] CHEN Z S, XIE L X, NIU J W, et al. Visformer: the vision-friendly transformer[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2022: 569-578.

[25] 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.

[26] 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 (CVPR). Piscataway: IEEE, 2017: 936-944.

作者简介

陈孟元　男,1984年1月生于安徽芜湖. 现为安徽工程大学电气工程学院教授,硕士生导师. 获安徽省科学技术奖一等奖. 主要研究方向为移动机器人SLAM、目标跟踪及路径规划.
E-mail: mychen@ahpu.edu.cn

韩朋朋　男,1992年5月出生于安徽省阜阳市. 现为安徽工程大学电气工程学院硕士研究生. 研究方向为移动机器人视觉SLAM算法.
E-mail: 1421384659@qq.com

刘金辉　男,1997年5月出生于安徽省亳州市. 现为安徽工程大学电气工程学院硕士研究生. 研究方向为移动机器人视觉SLAM算法.
E-mail: m17855336477@163.com

张玉坤　男,1995年5出生于安徽省亳州市. 现为安徽工程大学电气工程学院硕士研究生. 研究方向为移动机器人视觉SLAM算法.
E-mail: 1728443478@qq.com

江浩玮　男,1997年7月出生于安徽省安庆市. 现为安徽工程大学电气工程学院硕士研究生. 研究方向为移动机器人视觉SLAM算法.
E-mail: 576992617@qq.com

丁陵梅　女,1994年12月出生于江苏省泰州市. 现为安徽工程大学电气工程学院硕士研究生. 研究方向为移动机器人视觉SLAM算法.
E-mail: 1425288603@qq.com