

电子科技大学

专业学位研究生学位论文中期考评表

攻读学位级别： ☐ 博士 ☒ 硕士

培养方式： ☒ 全日制 ☐ 非全日制

专业学位类别及领域： 软件工程

学 院： 信息与软件工程学院

学 号： 202222090537

姓 名： 马逸逍

论 文 题 目： 室内动态场景下的

V-SLAM 算法研究

校内指导教师： 殷光强

校外指导教师： 梁熙

填 表 日 期： 2024 年 9 月 15 日

电子科技大学研究生院

一、已完成的主要工作

1. 开题报告通过时间： 2023 年 12 月 21 日

2. 课程学习情况

是否已达到培养方案规定的学分要求	☐ 是 ☒ 否
------------------	--

3. 论文研究进展

从理论分析或计算部分、实验实践（或实证）工作等方面进行总结（可续页）

视觉同步定位与建图（V-SLAM）在动态环境中通过识别运动的物体，减少这些物体对机器人定位的干扰。大部分 V-SLAM 算法假设周围环境是静态的，因此在固定场景下通常能够实现良好的定位与建图效果。然而，在动态场景中，摄像头可能会捕捉到运动物体，从而导致定位过程中对自身姿态的误判，甚至可能引发定位失败。

本文以室内动态场景为研究对象，设计了一个结合 YOLOv8 实例分割和几何信息（Geometric）的 GeoYOLO-SLAM 系统。该 SLAM 系统可以区分正处于运动状态的先验动态物体，以实现在复杂场景中准确检测和区分处于运动状态的先验动态物体。再通过给地图点增加动态概率属性，根据物体的运动状态更新地图点的动态概率，以便在 tracking 阶段只匹配静态地图点匹配的静态特征点从而解决动态特征点对相机位姿估计的干扰问题。具体架构如图 1。

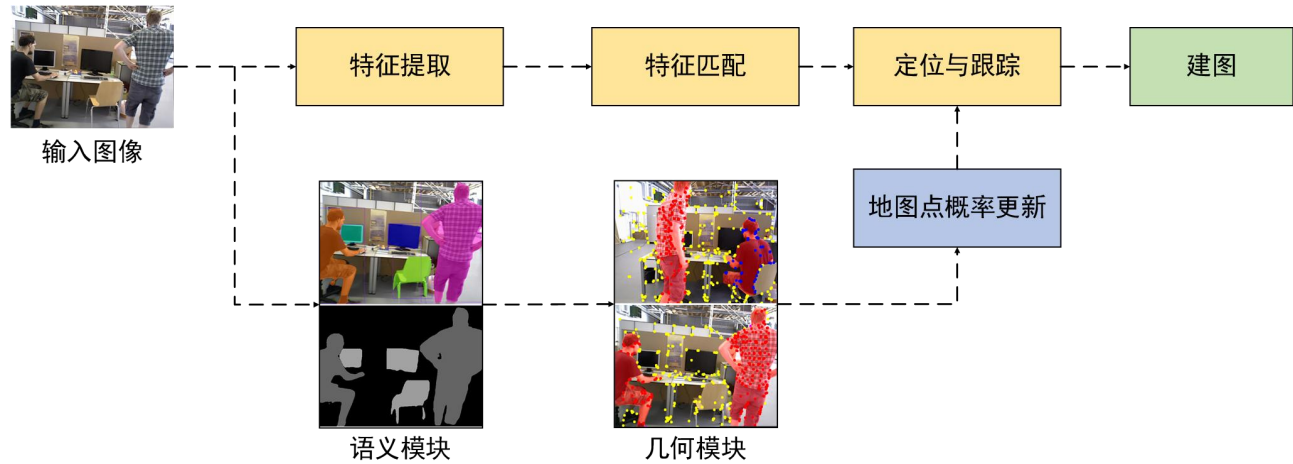


图 1 GeoYOLO-SLAM 系统结构图

一、语义模块研究进展

1. 理论分析

在机器人操作系统（ROS）的框架下，利用 ACTION 通讯机制构建了一个高效的数据处理与交互系统。该系统能够确保在每个视频流的关键帧上，自动地向服务端发送需要实例分割的图片。充分利用了 ROS 的分布式处理能力和异步通讯特性，确保了请求的及时发送与响应。实例分割请求与结果返回过程如下图 2 所示，客户端每收到两个关键帧 KF_i 和 KF_{i+1} ($i = 0, \dots, n - 1$ 其中 n 为关键帧的数量)就将其发送至服务端进行实例分割，服务端进行 YOLOv8 实例分割后返回这两帧的处理结果 S_i 和 S_{i+1} 。

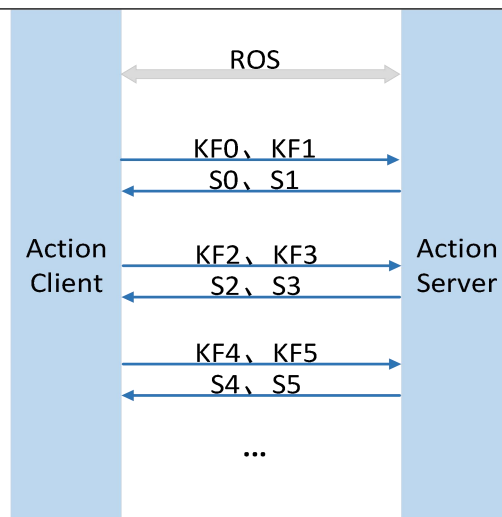


图 2 基于 ROS 的关键帧和实际结果的通信机制

对返回的结果进行筛选处理，提取出每个先验动态物体（例如人）的 **mask** 信息并标号区分开，再通过形态学扩张后得到先验动态物体的位置信息跟深度图的物体边界做交集得到先验动态物体的 **mask** 信息。每次向服务端发送两个关键帧进行实例分割并将结果返回到客户端后根据处理后的 **mask** 结果更新地图点的移动概率，之后直接通过静态地图点对应的特征点的匹配求得基础矩阵 F 。

2. 工作内容

如下是 TUM 数据集 `rgbd_dataset_freiburg3_walking_xyz` 中的一个关键帧中对象 **mask** 的处理结果。如下图 3 所示，首先关键帧通过 ROS 发送给服务端在 GPU 上进行 YOLOv8 后获得实例分割结果图。该图中，不同对象分别使用不同颜色标注出各自的 **mask** 信息。其次，对彩色图进行灰度处理，将每个像素由 RGB 转换成 0 到 255 之间的灰度值。其中每个对象的灰度值都各不相同，以便之后区分不同的对象。然后，通过对象的灰度值提取不同的先验动态物体的 **mask**。最后，将每个对象的 **mask** 信息结合深度图信息将 **mask** 对象边沿与深度图中对象深度变化明显的边沿取交集获得最后的对象 **mask** 信息。

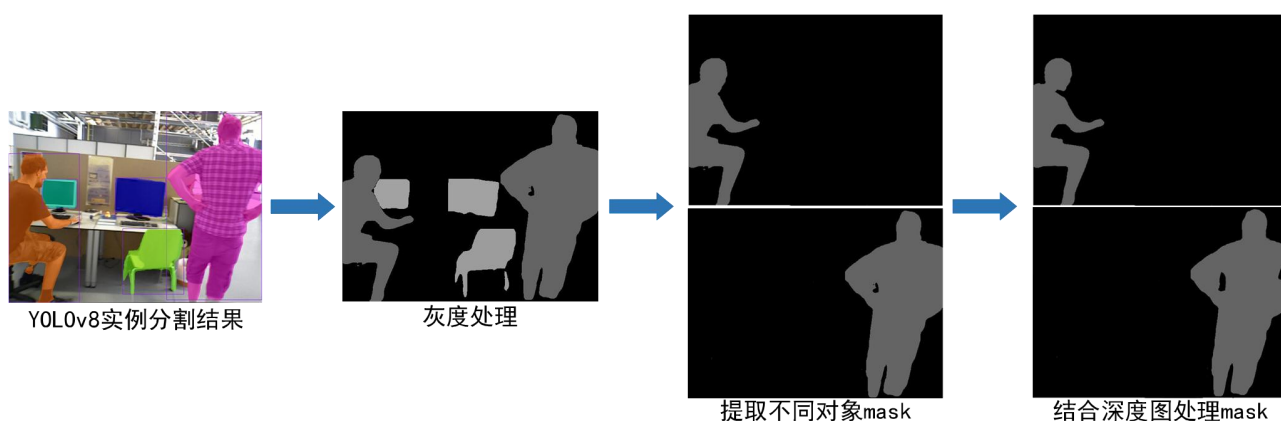


图 3 处理不同先验动态物体对象 **mask**

二、几何模块与地图点移动概率研究进展

1. 理论分析

先验动态物体在实际情况中不一定处于运动状态，例如室内静止不动的人。在纹理信息丰富的

情况下可以提取较多的特征点，如果仅依靠类别将先验动态物体上的特征点全部去除，虽然可能会误删静止物体上的特征点但仍然有大量特征点可用于匹配。然而在纹理信息较少的背景下，例如静止的人靠在室内白墙上，简单地按照类别删除特征点会导致丢失大量静态特征点从而导致特征点数量不足等问题。

针对上述问题，首先利用先验静态特征点的匹配结果来计算出较为可靠的基础矩阵 F 。接着利用对极几何原理，逐一检查先验动态物体中的匹配特征点，检查它们到各自极线的距离。如果这一距离超出了设定的阈值 θ ，就将这些特征点视为不符合极线几何约束的离群点，认为这通常是因为动态物体的移动所致。如果一个先验动态物体内的离群点所占比例超过一定的阈值，则认为这个先验动态物体是真实处于运动状态的。下图所示显示了两个连续的关键帧 KF_1 于 KF_2 之间的对极几何约束。

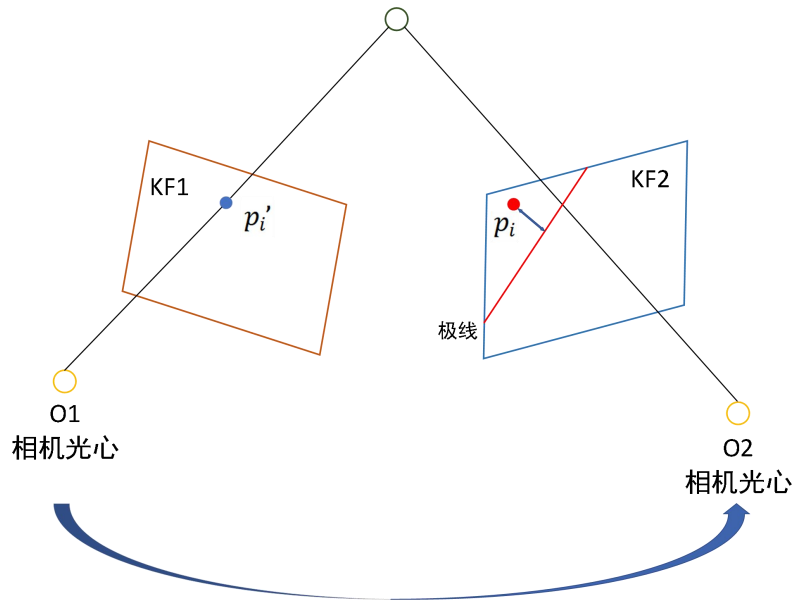


图 4 对极几何判断特征点运动状态

设 p_i' 和 p_i 分别表示前一关键帧和当前帧中匹配的关键点。

$$p_i' = [x_{p_i'}, y_{p_i'}, 1] \quad (1)$$

$$p_i = [x_{p_i}, y_{p_i}, 1] \quad (2)$$

其中， x 与 y 为匹配关键点的像素坐标，极线 $l_{p_i'}$ 的公式如下：

$$l_{p_i'} = \begin{bmatrix} X_{p_i'} \\ Y_{p_i'} \\ Z_{p_i'} \end{bmatrix} = F(p_i')^T = F \begin{bmatrix} x_{p_i'} \\ y_{p_i'} \\ 1 \end{bmatrix} \quad (3)$$

其中 $X_{p_i'}$, $Y_{p_i'}$, $Z_{p_i'}$ 表示极线向量的坐标， F 表示基本矩阵。则匹配点 p_i 到极线 $l_{p_i'}$ 的距离表示为：

$$d(p_i, l_{p_i'}) = \frac{|p_i F(p_i')^T|}{\sqrt{\|X_{p_i'}\|^2 + \|Y_{p_i'}\|^2}} \quad (4)$$

当该关键帧中先验动态物体范围内有超过一定阈值的匹配特征点到极线的距离超过 θ ，则认为该先验动态物体的对象处于真实运动状态。否则，认为该先验动态物体处于静止状态。进而更新这

些特征点对应的地图点的动态概率。

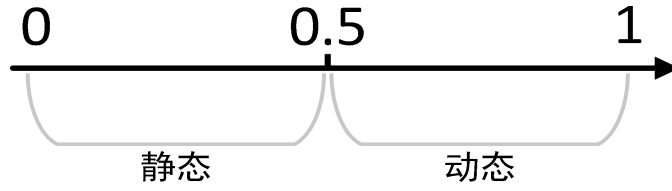


图 5 地图点的移动概率

定义移动概率 $p(m_t^i)$ 每个地图点 i 在当前时刻的移动概率，如图 5 所示。

如果地图点的移动概率接近于 1，那么它的状态就更有可能是动态的。如果地图点更接近于零，那么它就越静态。每个地图点有动态和静态两种状态，初始概率设为 0.5 ($\text{bel}(m_0)$)。考虑到语义分割不是 100% 准确，定义了观察移动概率：

$$p(z_t = d|m_t = d) = \alpha \quad (5)$$

$$p(z_t = s|m_t = d) = 1 - \alpha \quad (6)$$

$$p(z_t = s|m_t = s) = \beta \quad (7)$$

$$p(z_t = d|m_t = s) = 1 - \beta \quad (8)$$

其中 z 表示 YOLOv8 实例分割的结果， m 表示实际地图点的运动状态。在实验中，将 α 和 β 值设为 0.9。那么当前时刻地图点的移动概率 $\text{bel}(m_t)$ 依赖于之前时刻的实例分割结果 $z_{1:t}$ 以及地图点的初始状态 m_0 。那么移动概率的更新问题可以表示为：

$$\begin{aligned} \text{bel}(m_t) &= p(m_t|z_{1:t}, m_0) \\ &= \mu p(z_t|m_t, z_{1:t-1}, m_0) p(m_t|z_{1:t-1}, m_0) \\ &= \mu p(z_t|m_t) p(m_t|z_{1:t-1}, m_0) \\ &= \mu p(z_t|m_t) \overline{\text{bel}}(m_t) \end{aligned} \quad (9)$$

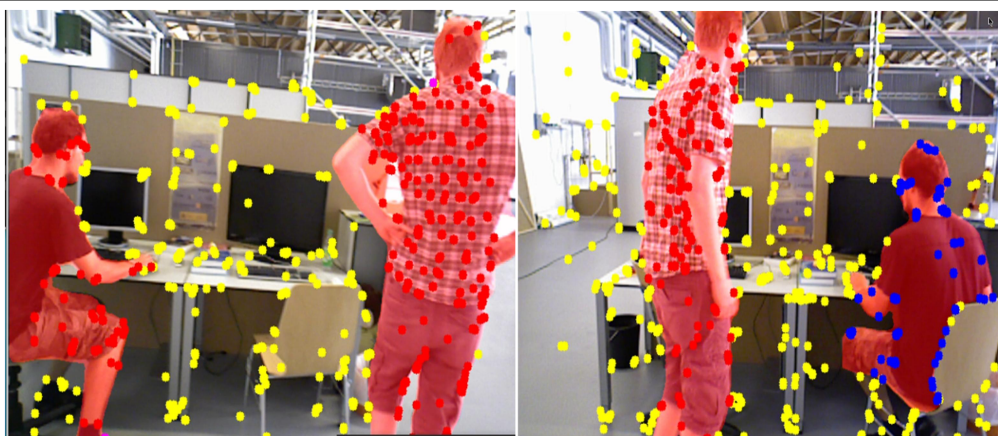
可以转换成当前时刻地图点的移动概率 $\text{bel}(m_t)$ 依赖于预测的地图点的移动概率 $\overline{\text{bel}}(m_t)$ 。由于无法观测到物体的突然变化，假设状态转移概率 $p(m_t = d|m_{t-1} = s) = 0$ ， $p(m_t = d|m_{t-1} = d) = 1$ ， $\mu = \text{bel}(m_t = d) + \text{bel}(m_t = s) / 2$ 。因此地图点的移动概率为：

$$\overline{\text{bel}}(m_t = d) = p(m_t = d|m_{t-1} = d) \text{bel}(m_{t-1} = d) \quad (10)$$

在更新地图点移动概率之后，将概率值小于 0.5 的地图点认为是静态物体对应的地图点，概率值大于 0.5 的地图点认为是动态物体对应的地图点。在跟踪阶段就只处理静态地图点对应的特征点，从而忽略动态特征点匹配。

2. 工作内容

经过 YOLOv8 实例分割出的先验动态物体处理后的 mask 如图 6 红色部分。如图 6(1)所示，经过对极几何判断后两个对象都真正处于运动状态，动态地图点对应特征点如左图红色特征点。如图 6(2)所示，坐在位置上处于静止状态的人在经过对极几何判断之后被标注为静止状态，静态地图点对应特征点如右图坐下的人上的蓝色特征点。



(1)

(2)

图 6 结合实例分割与对极几何的对象运动状态处理

使用 TUM 室内情况下 Dynamic Objects 类别下的 5 个数据集评估跟踪精度，通过与最先进的 V-SLAM 方法进行比较来展示实时性能。在这里使用绝对轨迹误差 (Absolute Trajectory Error, ATE) 的均方根误差 (Root Mean Square Error, RMSE) 作为评估指标。

表 1 在 TUM 数据集上与最先进的 V-SLAM 比较所得 ATE 的 RMSE

Sequences	GeoYOLO-SLAM	ORB-SLAM3	DynaSLAM (Stereo)
	ATE	ATE	ATE
walking_static	0.011	0.376	0.007
walking_xyz	0.374	0.545	0.016
walking_half	0.174	0.657	0.030
walking_rpy	0.298	0.987	0.035
sitting/static	0.009	0.012	0.011

相较于 ORB-SLAM3 来说，在有动态物体的情况下 ATE 相对减少甚多，能较好适应动态情况。但是还有待改进。

4. 阶段性研究成果

按《研究生学位论文撰写格式规范》的格式要求分类填写与学位论文相关的阶段性研究成果，例如期刊论文、会议论文、科研获奖、专利、制定标准等，限填第一作者或导师为第一作者时的第二作者成果，其中已录用、已投稿或拟投稿的在括号内注明（可续页）

二、存在的主要问题和解决办法

1. 未按开题计划完成的研究工作，研究工作存在的原理性、技术性难题以及在实验条件等方面的限制（可续页）

技术性问题：

1、地图点动态概率更新不完全

在更新某一帧的地图点动态概率时，处于局部地图点中属于该帧的地图点会进行动态概率的更新。但是还有些地图点不在没有被该帧观测到，就导致这些地图点的动态概率没有被更新。在后期使用地图点动态概率的时候就可能出错。

2、关联遗漏的动态物体检测

目前只是处理了先验动态物体的运动状态。与先验动态物体相接触的关联遗漏的动态物体检测可能也会影响 V-SLAM 的位姿估计。

2. 针对上述问题采取何种解决办法，对学位论文的研究内容及所采取的理论方法、技术路线和实施方案的进一步调整，以及下一步的研究研究计划（可续页）

解决方法：

1、地图点动态概率更新不完全

使用历史观测数据，对于当前帧未观测到的地图点，可以利用这些地图点在过去帧中的观测数据来更新其动态概率。具体来说，可以维护一个历史观测列表，记录每个地图点被观测到的帧号及对应的概率或置信度。在每次更新时，遍历这些历史数据，并基于一定的策略来更新这些点的动态概率。

2、关联遗漏的动态物体检测

建立一个模型来模拟动态物体之间的相互作用。当检测到一个新的动态物体时，可以检查它是否与已知的先验动态物体有潜在的相互作用（如碰撞、接触等）。如果存在这样的相互作用，则更新这些物体的动态状态。

下一步的研究计划：

起止年月	完成内容
2024. 09-2024. 11	完成关联遗漏物体检测并解决问题优化算法，完成专利 1 篇。
2024. 11-2025. 01	优化模型，调整参数，验证实验效果，并设计系统。
2025. 01-2025. 02	完善整个系统，并撰写毕业论文初稿。
2025. 02-2025. 06	完善学位论文，完成学术论文 1 篇。

三、中期考评审查意见

1.导师对工作进展及研究计划的意见：

进展符合预期，研究计划合理可行。

校内导师（组）签字：

2024 年 9 月 20 日

校外导师签字：

2024 年 9 月 20 日

2. 中期考评专家组意见

考评日期	2024.09.20	考评地点	国际创新中心 B 栋一楼会议室 B105 腾讯会议：479-687-638
考评专家	杨远望、庄杰、李耶		
考评成绩	合格__3__票 基本合格__0__票 不合格__0__票		
结 论	☒通过 ☐原则通过 ☐不通过 通 过：表决票均为合格 原则通过：表决票中有 1 票为基本合格或不合格，其余为合格和基本合格 不 通 过：表决票中有 2 票及以上为不合格		

对学位论文工作进展以及下一步研究计划的建议，是否适合继续攻读学位：

研究工作进展正常，计划可行，适合继续攻读学位。

专家组签名：

年 月 日

3.学院意见：

负责人签名：

年 月 日