

# 基于人工路标的易部署室内机器人全局定位系统\*

廖志涵, 吴 锋<sup>†</sup>

(中国科学技术大学 计算机科学与技术学院, 合肥 230000)

**摘要:** 在一些布局易变或存在较多动态障碍物的室内, 移动机器人的全局定位依然面临较大的应用挑战。针对这类场景, 实现了一种新的基于人工路标的易部署室内机器人全局定位系统。该系统将人工路标粘贴在不易被遮挡的天花板上作为参照物, 仅依赖一个摄像头即能实现稳定的全局定位。整个系统根据具体的功能分为地图构建和全局定位两个过程。在地图构建过程中, 系统使用激光 SLAM 算法所输出的位姿估计结果为基准, 根据相机对路标点的观测信息来自动估计人工路标点在全局坐标系中的位姿, 建立人工路标地图。而在全局定位过程中, 该系统则是根据相机对地图中已知位姿的人工路标点的观测信息, 结合里程计与 IMU 融合的预积分信息来对位姿进行实时估计。充分的实验测试表明, 机器人在该系统所部署范围内运行的定位误差稳定在 10 cm 以内, 且运行过程可以保证实时位姿输出, 满足典型实际室内移动机器人全局定位的应用需求。

**关键词:** 室内机器人全局定位; 天花板定位; 人工路标; 传感器融合

中图分类号: TP242.6

文献标志码: A

文章编号: 1001-3695(2021)09-033-2741-05

doi:10.19734/j.issn.1001-3695.2021.01.0033

## Man-made landmark based convenient indoor-robot global localization system

Liao Zhihan, Wu Feng<sup>†</sup>

(School of Computer Science & Technology, University of Science & Technology of China, Hefei 230000, China)

**Abstract:** Global localization algorithms for mobile robot still face many challenges when using in some indoor environments, where the layout changes frequently or there are many dynamic obstacles. Aiming at such scenarios, this paper proposed a new and convenient global localization system based on man-made landmark. The system localized based on multiple man-made landmarks attached on the unshaded ceiling, and could achieve stable pose estimation with only one camera. The whole system consisted of two main processes: map construction and global localization, according to their specific functions. During the map construction, it automatically estimated the pose of the man-made landmark in the global coordinate system, according to the camera's observation of the landmark and the pose estimation output by some laser SLAM algorithm. Given this, it then established the map of environment based on the pose of each landmark. During the global localization, it estimated the robot's pose in real time based on the camera's observation of the man-made landmarks in the map with known poses, combined with the pre-integrated information fused with the odometer and IMU. Sufficient experiments show that the localization error of the system is stable within 10 cm when running in the deployed environment, and the system can also ensure real-time pose output, which meets the application requirements of typical indoor mobile robot global localization.

**Key words:** indoor-robot global localization; ceiling-view localization; man-made landmark; sensor fusion

## 0 引言

近年来, 移动机器人相关产品逐渐出现在人们的日常生活中, 如扫地机器人、送餐机器人、仓储机器人等。这类机器人通常具备一定的自主移动能力, 可以通过对周围环境参照物的检测和多次观测来估算自身在环境中的位姿, 基于该定位信息来规划最优路线以及灵活避开环境中一些静态或动态障碍物。

作为重要的研究方向之一, 移动机器人定位技术得到了学者们长期的关注, 并发展出大量的定位算法。其根据所使用的位姿观测传感器种类和对应的数据关联方法, 可划分为基于激光的定位算法和基于视觉的定位算法。经过多年的研究, 两者均已在理论方法和框架实现上取得了较大进步。由于摄像头相比于激光雷达更为廉价且可以获得更为丰富的环境信息, 视觉定位方法得到了机器人应用开发者的青睐, 在实际中被大量使用。

然而, 现有的机器人视觉定位算法还存在一定的不足, 很难保证在某些实际应用环境中的稳定运行。究其原因, 主要是由于相关运行环境经常发生较大布局改变(如餐厅中随意摆放的桌椅), 以及环境中存在的一些动态障碍物(如熙熙攘攘

的人群)和光照影响(如早、中、晚间的光源变化)等因素。这会对视觉传感器以及相关算法造成较大影响, 从而使定位算法得到的位姿估计精度下降甚至产生巨大的偏差<sup>[1]</sup>。

针对以上问题, 学者们进行了大量的尝试, 并有一部分学者提出通过改变相机的观测视角来减小以上因素对相关定位算法的影响。其中, CV-SLAM<sup>[2]</sup>首次提出使用相机观察不易被遮挡的天花板中的路标信息进行位姿估计。同样是通过观测天花板中所含有的路标信息进行定位, se2lam<sup>[3]</sup>则是基于 ORB-SLAM<sup>[4]</sup>框架, 除了从连续的天花板观测图像中提取出 orb 特征点并进行跟踪外, 还在跟踪过程中融入了里程计信息, 从而使得算法输出的估计位姿更加精准且具备真实尺度。不同于 CV-SLAM 和 se2lam 从图像中提取特征点并对其进行跟踪解算位姿, 文献[5]则是针对含有较多栅格的天花板场景, 将栅格作为路标点, 从每次观测到的栅格信息推算出当前位姿。上述方法均是通过从对天花板的连续观测图像中提取有效自然路标点并对其进行跟踪来进行位姿解算。虽然该类方法容易部署, 但容易受到环境影响。特别是若天花板中纹理信息较少, 则算法的定位精度会受到较大影响, 甚至无法运行。

收稿日期: 2021-01-25; 修回日期: 2021-03-31 基金项目: 国家重点研发计划项目(2017YFB1002204); 国家自然科学基金资助项目(61603368)

作者简介: 廖志涵(1994-), 男, 硕士研究生, 主要研究方向为移动机器人定位、多传感器融合; 吴锋(1984-), 男(通信作者), 副教授, 主要研究方向为智能机器人(wufeng02@ustc.edu.cn)。

为了在弱纹理环境下依旧可以进行定位,文献[6]提出将多个存储有自身全局坐标的 QR 码作为路标粘贴在天花板上,再根据高频工业相机的观测信息直接推导出当前机器人在全局坐标系中的位姿。同样是基于已知自身全局坐标的 QR 码,文献[7]则是融入了里程计与 IMU 信息进行航位推算,并使用相关滤波器融合路标点的观测信息后输出估计位姿。上述方法均是通过在环境中部署已知自身坐标的人工路标并结合相机对其的观测信息来进行位姿估计。而人工路标在全局坐标系中的坐标则是通过人工测量或者动作捕捉获得,这无疑加大了这类系统的部署难度。

针对现有自然路标点易受环境影响和现有的人工路标点不易部署的问题,本文以人工路标点法为切入点,将计算效率更高且适用于长距离观测的 AprilTag<sup>[8]</sup> 作为路标点部署在天花板上来对环境进行描述,提出不易受环境影响且容易部署的室内机器人全局定位系统。整个系统可根据其具体功能划分为地图构建和全局定位两个部分。其中,地图构建过程使用了现有激光 SLAM 算法(如 cartographer<sup>[9]</sup>)所输出的估计位姿为参照,根据相机对路标点的观测信息估计人工路标点在全局坐标系中的位姿,建立人工路标地图。相对于已有人工路标点法中对路标点位姿的繁杂测量方式,本文方法在确保路标点坐标精度的同时,还极大减小了系统部署的难度。在全局定位过程中,本文则是基于在地图建立过程中所得到的路标点信息,使用优化的方法融合相机对路标点的观测、里程计和 IMU 信息输出估计位姿,从而减少相机观测误差对位姿估计的影响,实现实时输出定位。

综上,本文对于机器人定位研究的主要贡献如下:

- a) 结合相关激光 SLAM 算法的位姿估计结果和相机对人工路标点的观测信息,本文提出通过构建一系列残差并对其优化的方法确定人工路标点在全局坐标系中的位姿,解决了因激光位姿估计误差和相机观测误差所带来的误差累计,保证了人工路标点的位姿估计精度。
- b) 基于提出的方法,本文实现了一套容易部署的全局定位系统,该系统能够自动建图,无须对人工路标点的位姿进行繁琐的测量,且仅需要一个廉价的摄像头即能实现实时定位。
- c) 基于对本文系统进行的室内实验测试和分析,证实了本文系统所估计人工路标的全局位姿基本准确,全局定位输出的位姿也具备较小误差,输出精度和频率可以满足一般室内机器人的导航需求。

此外,使用本文方法确定全局位姿的人工路标还可作为其他算法<sup>[10]</sup>的“回环”检测与“回环”矫正参照物,由此更好地消除位姿估计的累积误差。同时,本文系统的全局定位算法输出的估计位姿也可用作其他算法的初始估计位姿,从而使得相关算法输出更精确地估计位姿。因此,本文系统不仅具有较高的应用价值,还具有一定的理论意义。

## 1 算法框架

### 1.1 符号说明

#### 1.1.1 传感器与相关坐标系

本文系统所使用的机器人平台搭载的相关传感器与计算资源如图 1(a)所示。基于图中所列出的传感器,本文定义如图 1(b)所示坐标系,其中  $C$  表示相机坐标系,  $B$  表示机器人底盘坐标系,  $I$  表示 IMU 坐标系,  $L$  表示激光坐标系,  $W$  表示世界坐标系,  $O$  表示里程计坐标系。

#### 1.1.2 位姿表示与坐标系转换

基于上述坐标系定义,本文使用式(1)表示在某种观测传感器获得第  $k$  次(对应时刻为  $t_k$ )有效观测(即该观测可用于位姿优化)时,  $B$  坐标系在  $W$  坐标系中的位姿。

$${}^W\xi = ({}^Wt_{B_k}, {}^W r), {}^Wt_{B_k} = (t_x, t_y, t_z)^T, {}^W r = (r_x, r_y, r_z)^T \quad (1)$$

其中:  ${}^Wt_{B_k}$  表示  $B$  坐标系在  $W$  坐标系中的三维坐标;  ${}^W r$  表示从  $B$  坐标系转换到  $W$  坐标系的三维旋转向量。该旋转向量所对应的四元数表示为  ${}^Wq$ , 同时可通过罗德里格斯公式<sup>[11]</sup>转换为其所对应的三维旋转矩阵  ${}^W R$ 。该三维位姿所对应的二维平面位姿为

$${}^W B_k v = P({}^W \xi) = ({}^W \rho_{B_k}, {}^W \theta) \\ {}^W \rho_{B_k} = (({}^W t_{B_k})_x, ({}^W t_{B_k})_y)^T, {}^W \theta = ({}^W r)_z \quad (2)$$

其中:使用  $(\cdot)_x$ 、 $(\cdot)_y$ 、 $(\cdot)_z$  分别表示某三维向量的  $x$ 、 $y$ 、 $z$  部分。基于该三维位姿表示,本文系统中将某一在  $W$  坐标系中三维坐标为  ${}^W t_{m_i}$  的路标点  $m_i$  转换到第  $k+1$  次  $B$  坐标系中的坐标  ${}^{B_{k+1}} t_{m_i}$  的过程表示为

$${}^{B_{k+1}} t_{m_i} = {}^{B_{k+1}} R^T ({}^W t_{m_i} - {}^W t_{B_{k+1}}) \quad (3)$$

为方便书写,本文中定义运算符  $(*)$ , 将式(3)简写为

$${}^{B_{k+1}} t_{m_i} = {}^{B_{k+1}} \xi^{-1} * {}^W t_{m_i} \quad (4)$$

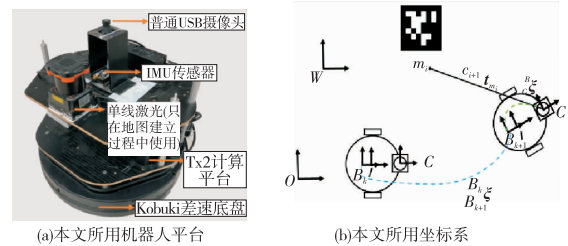


图 1 本文所用的机器人平台和坐标系  
Fig. 1 Robot and coordinate in our system

#### 1.1.3 相机观测模型

本文中所使用的相机观测模型为针孔相机观测模型。基于式(4)中的位姿转换方式,该相机观测模型将编号为  $m_i$  的 AprilTag 中的某一角点  $j$  在  $W$  坐标系中的坐标  ${}^W t_{m_i,j}$  投影为第  $k$  次有效图像观测中的某一像素  $u_{m_i,j}^k$  的过程写作:

$$u_{m_i,j}^k = \pi(\delta, {}^W \xi, {}^B \xi, {}^W t_{m_i,j}) \quad (5)$$

其中:  $\delta = (f_x, f_y, c_x, c_y, k_1, \dots, k_n)$  表示相机内参;  ${}^B \xi$  表示相机到底盘的外参转换,其通过文献[12]中所搭建的标定系统获得。

### 1.2 算法整体流程

本文算法的整体流程可划分为地图建立过程和全局定位过程,前者的功能是要对人工路标在世界坐标系中的坐标进行估计,而后者则是基于前者的估计结果,对底盘在全局坐标系中的位姿进行估计。两个部分的算法流程分别如图 2 和 3 所示。

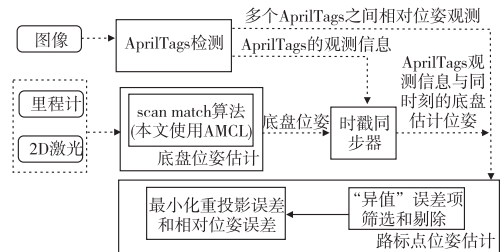


图 2 地图建立算法流程

Fig. 2 Flow chart of the mapping algorithm

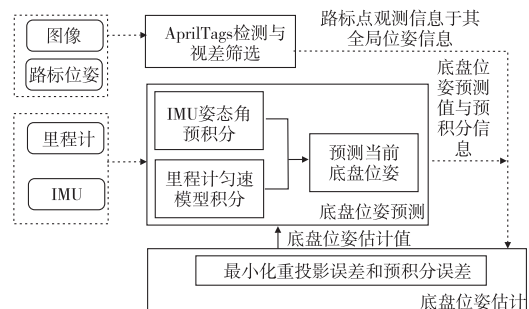


图 3 全局定位算法流程

Fig. 3 Flow chart of the global localization algorithm

在地图建立过程中,本文采用单线激光作为位姿观测传感器。在使用 cartographer 算法构建用于路径规划的二维平面地图的同时,本文还根据其底盘位姿估计结果和相机对路标点的观测信息,通过构建重投影误差和相对位姿误差并对其进行优化的方法来对路标点在全局坐标系中的位姿进行估计。

基于地图建立过程中所估计的路标点位姿,全局定位过程则是以上一次有效观测所对应时刻为积分起点,结合里程计与 IMU 进行积分所得的预积分信息,以及相机对已知自身坐标的路标点的观测信息,通过构建预积分误差和重投影误差并对其进行优化的方法来对当前的底盘位姿进行实时估计。

### 1.3 地图建立过程

#### 1.3.1 问题定义

在本文系统中,整个环境地图是由部署在天花板上的人工路标组成。

$$M = \{m_i | i \in N\} \quad (6)$$

其中:人工路标的观测信息是由其所对应的唯一编码和人工路标四个角点在图像中的坐标组成(本文使用文献[8]提供的开源库从图像中获取上述信息)。针对一张编码为  $m_i$ ,宽度为  $s$  且绝对平整的 AprilTag,本文以其中心为坐标原点,定义其四个角点在该人工路标坐标系中的坐标如图 4 所示。在第  $t$  时刻,该路标点的观测信息即其四个角点在图像中的像素坐标,表示为

$$w_{m_i}^t = \{u_{m_i,j}^t | j=0,1,2,3\} \quad (7)$$

同时将该时刻对应图像中观测到的所有路标点集合表示为

$$f^t = \{m_i | m_i \in M\} \quad (8)$$

地图建立过程既是要基于底盘位姿估计结果和上述路标点的观测信息,获取所有路标点在世界坐标系中的位姿。

$$X = ({}^W_{m_0}\xi, {}^W_{m_1}\xi, \dots, {}^W_{m_i}\xi, \dots, {}^W_{m_n}\xi) \quad m_i \in M \quad (9)$$

其中:单个路标点的位姿表示为

$${}^W_{m_i}\xi = ({}^W_{t_{m_i}}, {}^W_{r_{m_i}}), {}^W_{t_{m_i}} = (x_i, y_i, z_i)^T, {}^W_{r_{m_i}} = (0, 0, \theta_i)^T \quad (10)$$

其中:估计变量为  $x_i, y_i, z_i, \theta_i$ 。

#### 1.3.2 重投影误差构建

在地图建立过程中,若激光 SLAM 算法在  $k$  次有效激光观测时(即算法使用该观测完成位姿优化),当前对应时刻的图像中可以观测到如图 5 所示路标点集合  $f^{tk}$ 。

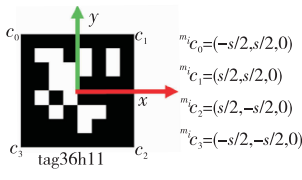


图 4 本文使用的 AprilTag

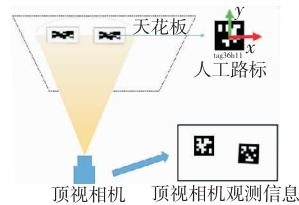


图 5 路标点观测信息

Fig. 4 AprilTag used in this paper Fig. 5 Landmark observation

则对  $f^{tk}$  中某一路标点  $m_i$  中某一角点  ${}^{m_i}c_k$  的重投影误差为

$$e_{m_i,j}^k(X) = \pi(\delta, E({}^W_{B_k}\mathbf{v}) {}^B_k \xi, {}^W_{m_i}\xi * {}^{m_i}c_j) - u_{m_i,j}^k \quad (11)$$

其中:  $E({}^W_{B_k}\mathbf{v})$  表示该二维位姿所对应的三维位姿  ${}^W_{B_k}\hat{\xi}$ 。

$${}^W_{B_k}\hat{\xi} = E({}^W_{B_k}\mathbf{v})$$

$${}^W_{t_{B_k}} = (({}^W_{\rho_{B_k}})_x, ({}^W_{\rho_{B_k}})_y, 0), {}^W_{r_{B_k}} = (0, 0, {}^W_{\theta_{B_k}}) \quad (12)$$

基于式(11)所构建的重投影误差,则整个地图建立过程中由重投影误差所构建的损失函数表示为

$$J_p(X) = \sum_k \sum_{m_i \in f^{tk}} \sum_{j=0}^3 e_{m_i,j}^k(X)^T \Omega_C^{-1} e_{m_i,j}^k(X) \quad (13)$$

其中: $k$  表示第  $k$  次有效激光观测; $f^{tk}$  表示第  $k$  次有效激光观测时刻所对应图像中的路标点观测信息。

#### 1.3.3 相对位姿误差构建

在地图建立过程中,若在某一时刻  $t$  中所对应图像检测出的路标点集合  $f^t$  中含有如图 6 所示多个路标点。

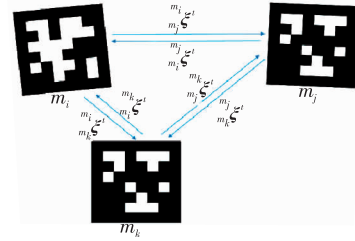


图 6 相对位姿观测

Fig. 6 Observation of relative position

则可以使用相关 PnP 算法求得  $f^t$  中路标点  $m_i$  和  $m_j$  在该时刻  $C$  坐标系中分别所对应的位姿  ${}^C_{m_i}\xi^t$  和  ${}^C_{m_j}\xi^t$ ,则在该时刻中,  $m_i$  和  $m_j$  的相对位姿观测为

$${}^{m_i}_{m_j}\xi^t = ({}^C_{m_i}\xi^t)^{-1} {}^C_{m_j}\xi^t, {}^{m_j}_{m_i}\xi^t = ({}^C_{m_j}\xi^t)^{-1} {}^C_{m_i}\xi^t \quad (14)$$

基于式(14)中所推导的相对位姿测量,若在构建重投影时所对应的图像观测中含有  $m_i$  和  $m_j$  两个路标点,则路标点  $m_i$  中某一角点  ${}^{m_i}c_j$  基于  $m_i$  和  $m_j$  之间的相对位姿观测误差为

$$e_{m_i,j}^k({}^{m_j}_{m_i}\xi^t, X) = \pi(\delta, E({}^W_{B_k}\mathbf{v}) {}^B_k \xi, {}^{m_j}_{m_i}\xi^t * {}^{m_i}c_j) - u_{m_i,j}^k \quad (15)$$

则整个地图建立过程中由相对位姿误差构建的损失函数为

$$J_r(X) = \sum_k \sum_{m_i, m_j \in F_{m_i, m_j}} \sum_{j=0}^3 e_{m_i,j}^k({}^{m_j}_{m_i}\xi^t, X)^T \Omega_C^{-1} e_{m_i,j}^k({}^{m_j}_{m_i}\xi^t, X) \quad (16)$$

其中: $k$  表示第  $k$  次有效激光观测;  $F_{m_i, m_j}$  表示运行过程中由所有存在相对位姿观测的  $m_i$  和  $m_j$  构成的组合对集合;  $\chi$  表示存在对  $m_i$  和  $m_j$  之间相对位姿观测的时刻集合。

#### 1.3.4 估计参数初始值与“异值”数据剔除

基于式(13)(16)所推导的损失函数,在地图建立过程中需要优化的问题为

$$X = \arg \min_X (J_p(X) + J_r(X)) \quad (17)$$

本文使用全局离线优化的方式对式(17)进行优化。即使用实验平台收集一段时间数据后,再构造上述损失函数进行优化。其中,所有估计参数的初始值都是根据其首次观测信息推导得到,如在第  $k$  次有效激光观测的底盘估计位姿为  ${}^W_{B_k}\mathbf{v}$ ,且首次观测到路标点  $m_i$ ,则可根据 PnP 算法解算得到  $m_i$  在当前时刻相机中的位姿为  ${}^{m_i}_{B_k}\xi$ ,则将  $m_i$  位姿的初始估计值设置为

$${}^W_{m_i}\hat{\xi} = E({}^W_{B_k}\mathbf{v}) {}^B_k \xi {}^B_k \xi {}^{m_i}_{B_k} \xi \quad (18)$$

同时,由于底盘位姿的估计精度可能会受运动学畸变<sup>[13]</sup>等原因影响而下降,故在优化迭代过程中本文会将一些重投影误差高于某一阈值的误差函数从优化过程中剔除。而针对路标点之间的相对位姿观测,本文则是基于文献[14],构建有向图。该图中节点为路标点位姿,边为两个路标点的相对位姿观测,边的权重为该边两端所连节点的相对位姿误差。在优化迭代过程中,本文会将该图中一些相对位姿误差较大的边从优化过程中剔除,确保每个节点之间只有两条误差最小且方向相反的有向边连接。

### 1.4 全局定位过程

#### 1.4.1 问题定义

基于在地图建立过程中所得地图  $M$  中各个路标点的位姿信息以及底盘在  $W$  坐标系中的初始位姿。全局定位过程则是以里程计和 IMU 融合的位姿信息作为底盘位姿初始估计值,结合相机对  $M$  中路标点的观测信息以及上一估计位姿之间的预积分信息来对当前时刻的位姿信息进行优化求解。故在全局定位过程中,需要优化的变量为

$$X_k = ({}^W_{B_k}\xi, b_k^g) \quad (19)$$

其中: $k$  表示第  $k$  次有效图像观测(即在对应时刻的图像中可以检测出路标且满足视差检测要求);  ${}^W_{B_k}\xi$  表示在第  $k$  次有效

图像观测时刻所对应的底盘位姿(后文记为第  $k$  次位姿);  $\mathbf{b}_g^k$  表示同时刻 IMU 角速度偏差。

### 1.4.2 里程计与 IMU 融合

基于第  $k$  次位姿的估计结果  $\mathbf{X}_k$ , 里程计与 IMU 融合的目的既是要融合两者的相关测量信息, 在有新的有效图像观测时, 使用两者的信息积分出第  $k+1$  次位姿估计  $\mathbf{X}_{k+1}$  的在相应时刻的位姿估计初始值, 其融合过程如图 7 所示。

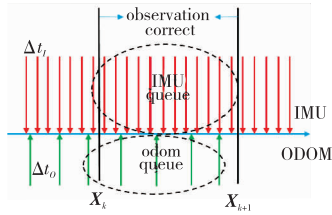


图 7 里程计与 IMU 融合  
Fig.7 Fusion of odometry and IMU

本文中里程计与 IMU 的融合形式主要是基于文献[15], 基于  $\mathbf{X}_k$  的姿态角信息, 使用 IMU 的角速度测量信息积分出  $\mathbf{X}_{k+1}$  中的姿态角信息。而针对  $\mathbf{X}_{k+1}$  中的坐标信息, 则是假设底盘在第  $k$  次位姿估计和第  $k+1$  次位姿估计之间以匀速运动的方式移动, 故使用里程计信息和 IMU 所解算的姿态角信息解算出第  $k$  次位姿估计和第  $k+1$  次位姿估计之间的速度再结合其间隔时间即可得到  $\mathbf{X}_{k+1}$  中的坐标信息。

$$\begin{aligned} {}^W \mathbf{t}_{B_{k+1}} &= {}^W \mathbf{t}_{B_k} + \Delta \hat{p}_{B_k B_{k+1}}, {}^{B_{k+1}} \hat{\mathbf{R}} = {}^{B_k} \mathbf{R} \Delta B_{k+1}^k \hat{\mathbf{R}} \\ {}^W \mathbf{V}_{B_{k+1}} &= \frac{{}^{B_{k+1}} \mathbf{R}^O {}^{B_{k+1}} \mathbf{R}^T ({}^O \mathbf{t}_{B_k} - {}^O \mathbf{t}_{B_{k+1}})}{t_k - t_{k+1}} \end{aligned}$$

$$\Delta \hat{p}_{B_k B_{k+1}} = {}^W \mathbf{V}_{B_{k+1}} \cdot (t_{k+1} - t_k), \Delta B_{k+1}^k \hat{\mathbf{R}} = \prod_{i=t_k}^{t_{k+1}} \exp(\mathbf{w}_i - \mathbf{b}_g^k - \mathbf{n}_g^k) \cdot \Delta t \quad (20)$$

其中:  $\mathbf{w}_i$  为 IMU 在第  $i$  时刻的角速度测量信息;  $\Delta t$  为 IMU 角速度测量频率;  $\mathbf{b}_g^k$  为角速度测量值的白噪声。

### 1.4.3 相机观测与位姿优化

基于第  $k+1$  次有效观测图像所对应里程计与 IMU 的融合预测底盘位姿  ${}^{B_{k+1}} \hat{\xi}$ , 位姿优化过程则是将该融合所得底盘位姿作为优化初始值, 结合相机与底盘之间的外参  ${}^B c$  和路标点观测信息  $f^{k+1}$  以及相关路标点的位姿信息, 可得观测  $f^{k+1}$  中某一路标点  $m_i$  的某一角点  $m_i^c$  的重投影误差为

$$e_{m_i^c}(X_{k+1}) = \pi(\delta, {}^{B_{k+1}} \hat{\xi}, {}^B c, m_i^c) - u_{m_i^c}^k \quad (21)$$

基于式(21)中的重投影误差, 可知由图像观测信息  $f^{k+1}$  所构建的损失函数为

$$J_p(X_{k+1}) = \sum_{m_i \in f^{k+1}} \sum_{j=0}^3 e_{m_i^c}(X_{k+1})^T \Omega_c^{-1} e_{m_i^c}(X_{k+1}) \quad (22)$$

同时结合 3.4.2 节中所推导的里程计与 IMU 融合信息和第  $k$  次有效观测图像所对应的估计位姿  ${}^{B_k} \hat{\xi}$ , 可以得到里程计与 IMU 融合的预积分误差为

$$\begin{aligned} e_p(X_k, X_{k+1}) &= {}^W \mathbf{t}_{B_{k+1}} - {}^W \mathbf{t}_{B_k} - \Delta \hat{p}_{B_k B_{k+1}} \\ e_r(X_k, X_{k+1}) &= 2({}^W \mathbf{q}^{-1} \otimes {}^{B_k} \hat{\mathbf{q}} \otimes ({}^{B_k} \hat{\mathbf{q}})^{-1})_{xyz} \\ e_{b_g}(X_k, X_{k+1}) &= \mathbf{b}_g^{k+1} - \mathbf{b}_g^k \end{aligned} \quad (23)$$

则由里程计与 IMU 融合测量构造的损失函数为

$$\begin{aligned} J_l(X_k, X_{k+1}) &= e_p(X_k, X_{k+1})^T \Omega_p^{-1} e_p(X_k, X_{k+1}) + \\ & [e_r(X_k, X_{k+1}), e_{b_g}(X_k, X_{k+1})]^T \Omega_r^{-1} \cdot \\ & [e_r(X_k, X_{k+1}), e_{b_g}(X_k, X_{k+1})] \end{aligned} \quad (24)$$

其中:  $\Omega_p^{-1}, \Omega_r^{-1}$  为协方差矩阵,  $\Omega_p^{-1}$  的推导参考了文献[16],  $\Omega_r^{-1}$  的推导方式则是参考了文献[15]。

基于式(22)(24)所推导的损失函数, 在全局定位过程中, 需要优化求解的问题为

$$\mathbf{X}_{k+1} = \arg \min_{\mathbf{X}_{k+1}} (J_p(\mathbf{X}_{k+1}) + J_l(\mathbf{X}_k, \mathbf{X}_{k+1})) \quad (25)$$

在对式(25)进行优化求解后, 本文会及时对相关参数进行对应更新, 以确保后续底盘位姿估计结果的精度。同时基于该底盘位姿估计结果, 本文使用里程计和 IMU 信息进行积分实时输出  $k+1$  和  $k+2$  估计位姿之间某一时刻的底盘位姿信息。

## 2 实验与分析

### 2.1 实验配置

本文所用实验平台与相关传感器如图 1(a) 所示, 其中实验平台为 Turtlebot2, 算法运行平台为 NVIDIA 公司的边缘计算平台 JETSON TX2, 使用 64 位 Ubuntu 操作系统。使用的相关传感器为: 具备较高精度的 Hokuyo 二维激光传感器、分辨率为  $640 \times 480$  的普通相机模组、普通 IMU 模组。实验中所用的里程计数据由 Turtlebot2 底盘输出。基于上述传感器和实验平台, 本文算法中的地图建立过程使用了激光传感器和相机传感器, 而全局定位过程则是基于地图建立过程中所得的路标点位姿, 使用了相机传感器和 IMU 传感器。

### 2.2 实际场景测试设计

由于本实验所设计系统并无相关公开测试数据集, 故本算法主要通过如图 8 所示  $9 \text{ m} \times 5 \text{ m}$  的室内环境中进行有效性验证。同时, 由于实验平台在运行过程中的底盘真值位姿难以获取, 故本实验主要将本文算法以及对比算法所输出的底盘位姿估计结果和同时运行的 AMCL 算法输出的底盘位姿估计结果进行对比来对本文系统的有效性进行评估。本文使用宽度为  $0.2 \text{ m}$  并印制在平整纸板上的 AprilTag 作为人工路标点, 将其粘贴在天花板上对机器人的实际运行环境进行描述。其中, 所使用的人工路标点数目为 23 个, 且其编码均不重复, 两路标点之间最小部署间隔距离为  $1 \text{ m}$ , 最大部署间隔距离为  $2 \text{ m}$ 。

基于上述环境部署, 本文多次使用如图 1(a) 所示的硬件平台进行相关测试数据收集。按照本文系统中各个过程的功能进行划分, 本文所搜集的数据可分为用于测试地图建立过程的数据和用于测试全局定位过程的数据。其中, 用于测试地图建立过程的数据记录了激光数据信息、里程计信息和图像信息, 该数据运行轨迹较长且行驶速度较慢, 可以确保部署在环境中的每一个部署的路标点都可以被相机观测到。基于地图建立过程中所建立的二维平面地图, 本文使用文献[9]所提多分辨率地图匹配算法获取并保存底盘全局定位初始位姿作为本文全局定位算法的初始底盘位姿, 除底盘初始全局位姿信息之外, 该测试数据还会记录里程计信息、IMU 信息、图像信息以及作为真值进行对比的 AMCL 算法所输出的底盘估计位姿。

### 2.3 实际场景测试结果

基于用于测试地图建立过程的数据, 本文首先使用地图建立过程中所提方法构建相关残差并优化后, 得到如图 9 所示的二维平面地图和人工路标点位姿组成的路标点地图, 该人工路标点位姿所构成的地图基本符合其实际部署布局。



图 8 室内实验环境  
Fig.8 Experiment environment



图 9 环境地图(方块为人工路标点的位姿)  
Fig.9 Environment map(square is the position of AprilTag landmark)

基于上述人工路标点位姿所构建的环境地图, 本文使用了三个用于定位的测试数据对本文方法进行测试。同时, 本文还将文献[17]作为对比方法进行比较, 为使该方法更适应本文所设计的实验场景和相关设备, 本文对其进行了一些改进, 其中包括: 使用里程计基于上一次位姿估计输出进行积分输出估计位姿的方法解决该算法在无人路标观测时就无法进行

位姿估计的问题、对短时间内该算法的位姿估计结果进行平均后作为最新时刻估计位姿输出的方法来缓解相机成像质量对位姿估计的影响。在三个测试数据集上分别运行相关算法,所得到的 AMCL 算法、本文算法、对比算法以及只使用里程计积分所得到的轨迹对比如图 10 所示。

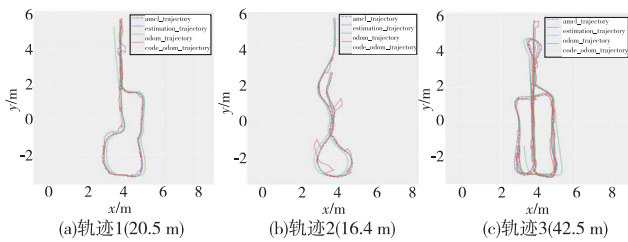


图 10 实际场景中的测试结果

Fig. 10 Experimental results in real environment

图 10 中,虚线部分代表 AMCL 算法所估计的轨迹结果,蓝色部分代表本文所提全局定位算法所估计的轨迹结果,绿色部分则代表只使用里程计积分所得到的轨迹结果,红色部分则是对比算法所估计的轨迹结果(见电子版)。若将 AMCL 算法所估计的位姿作为真实位姿值,基于文献[18]计算绝对轨迹误差,得到本文算法、对比算法以及只使用里程计进行积分所估计轨迹的绝对轨迹误差比较结果如表 1 所示。

表 1 绝对轨迹误差比较

Tab. 1 Comparison of the absolute trajectory error (ATE)

轨迹名	绝对轨迹误差的均方根误差 RMSE/m		
	里程计积分	对比方法	本文全局定位算法
轨迹 1	0.144 1	0.163 4	0.055 8
轨迹 2	0.150 7	0.201 5	0.059 6
轨迹 3	0.183 2	0.166 9	0.087 1

根据图 10 中的轨迹对比结果以及表 1 中所列出的数据,可以看到只使用里程计积分所得的位姿估计精度会随着运动距离的变长而不断累积增大,而对比方法虽然可以将位姿估计精度控制在一定范围内,但其位姿估计结果易受相机成像质量的影响而出现跳动现象。因此无法满足一些较高精度要求的低成本导航需求。而本文算法则是通过基于优化的多传感器融合的方式,有效缓解了单个传感器自身所存在的缺点,从而使得位姿估计结果更加精确,输出的轨迹也更加平滑,从而可以更好地适用于一些较高精度要求的低成本导航需求。虽然实验平台在运行过程中可能由于传感器噪声影响,以及由于路标点布局不合理或者图像受运动模糊影响而无法识别出路标点来对位姿进行即时矫正导致位姿累积误差变大,从而使使得本文算法所估计的轨迹相对 AMCL 算法所估计的轨迹产生了漂移,但是在多次成功观测到路标点,该累积误差可以得到有效减小。以上结果均说明了本文系统的有效性。

在时间效率上,本文算法使用单线程在 TX2 平台上运行。主要耗时模块为人工路标点检测模块和位姿优化模块。根据表 2 中的所列出的模块平均计算耗时,可知本文系统针对一次观测进行位姿优化的耗时为 3 ms 左右,能够满足相关应用的实时位姿输出要求。

表 2 系统耗时

Tab. 2 Time cost of the system /ms

操作	平均耗时
路标点提取	1.781
位姿优化	1.234

### 3 结束语

为减少布局改变以及环境因素对相关传感器和定位算法的影响,本文提出并实现了基于人工路标的室内机器人全局定位系统。该系统分为地图构建和全局定位两个过程。在地图构建过程中,本文首次提出结合更为精准的激光传感器和相关算法所输出的位姿估计结果和相机对人工路标的多次观测信

息来确定部署在天花板上的路标在全局坐标系中的坐标。与现有相关方法相比,本文方法在保证全局定位精度估计精度的同时,还极大的减少了系统的部署难度。基于地图建立过程中所得到的人工路标位姿,本文在全局定位过程中,融合相机对人工路标的观测信息、IMU 和里程计信息输出底盘的位姿估计结果。经实验结果证明,在本文系统所部署范围内运行可以将定位误差控制在 10 cm 以内,并且可以在移动平台上对图像以及位姿优化进行实时处理。下一步工作计划结合地图建立中的位姿估计结果和人工路标观测信息,实现相机外参的在线标定,从而进一步减少系统的部署难度。同时,针对人工路标的布局方式进行建模,在保证位姿估计精度的前提下,使用更少人工路标点的个数对环境进行描述,从而进一步减小系统部署的成本。

### 参考文献:

- [1] Cadena C, Carlone L, Carrillo H, et al. Past, present, and future of simultaneous localization and mapping: toward the robust-perception age[J]. *IEEE Trans on Robotics*, 2016, 32(6): 1309-1332.
- [2] Jeong W Y, Lee K M. CV-SLAM: a new ceiling vision-based SLAM technique[C]//Proc of IEEE/RSJ International Conference on Intelligent Robots and Systems. 2005: 3195-3200.
- [3] Zheng Fan, Liu Yunhui. Visual-odometric localization and mapping for ground vehicles using SE(2)-XYZ constraints[C]//Proc of International Conference on Robotics and Automation. 2019: 3556-3562.
- [4] Mur-Artal R, Tardós J D. ORB-SLAM2: an open-source SLAM system for monocular, stereo, and RGB-D cameras[J]. *IEEE Trans on Robotics*, 2017, 33(5): 1255-1262.
- [5] Ribacki A, Jorge V A M, Mantelli M, et al. Vision-based global localization using ceiling space density[C]//Proc of IEEE International Conference on Robotics and Automation. 2018: 3502-3507.
- [6] Cavanini L, Cimini G, Ferracuti F, et al. A QR-code localization system for mobile robots: application to smart wheelchairs[C]//Proc of European Conference on Mobile Robots. 2017: 1-6.
- [7] Nazemzadeh P, Fontanelli D, Macii D, et al. Indoor localization of mobile robots through QR code detection and dead reckoning data fusion[J]. *IEEE/ASME Trans on Mechatronics*, 2017, 22(6): 2588-2599.
- [8] Olson E. AprilTag: a robust and flexible visual fiducial system[C]//Proc of IEEE International Conference on Robotics and Automation. 2011: 3400-3407.
- [9] Hess W, Kohler D, Rapp H, et al. Real-time loop closure in 2D LiDAR SLAM[C]//Proc of IEEE International Conference on Robotics and Automation. 2016: 1271-1278.
- [10] 李月华, 朱世强, 于亦奇. 工厂环境下改进的视觉 SLAM 算法[J]. *机器人*, 2019, 41(1): 95-103. (Li Yuehua, Zhu Shiqiang, Yu Yiqi. Improved visual SLAM algorithm in factory environment[J]. *Robot*, 2019, 41(1): 95-103.)
- [11] Barfoot T D. State estimation for robotics[M]. Cambridge: Cambridge University Press, 2017: 189.
- [12] Chen Guangda, Cui Guowei, Jin Zhongxiao, et al. Accurate intrinsic and extrinsic calibration of RGB-D cameras with GP-based depth correction[J]. *IEEE Sensors Journal*, 2018, 19(7): 2685-2694.
- [13] Zhang Ji, Singh S. Low-drift and real-time lidar odometry and mapping[J]. *Autonomous Robots*, 2017, 41(2): 401-416.
- [14] Muñoz-Salinas R, Martín-Jiménez M J, Yeguas-Bolívar E, et al. Mapping and localization from planar markers[J]. *Pattern Recognition*, 2018, 73: 158-171.
- [15] Forster C, Carlone L, Dellaert F, et al. IMU preintegration on manifold for efficient visual-inertial maximum-a-posteriori estimation [EB/OL]. (2015). <https://doi.org/10.15607/RSS.2015.XI.006>.
- [16] Thrun S. Probabilistic robotics[J]. *Communications of the ACM*, 2002, 45(3): 52-57.
- [17] Zhang Huijuan, Zhang Chengning, Yang Wei, et al. Localization and navigation using QR code for mobile robot in indoor environment[C]//Proc of IEEE International Conference on Robotics and Biomimetics. 2015: 2501-2506.
- [18] Sturm J, Engelhard N, Endres F, et al. A benchmark for the evaluation of RGB-D SLAM systems[C]//Proc of IEEE/RSJ International Conference on Intelligent Robots and Systems. 2012: 573-580.
- [19] Kim Y G, Park T H. Localization of mobile robots from full detection of ceiling outlines[C]//Proc of IEEE International Conference on Informatics and Automation. 2016: 1515-1520.