DOI:10.14188/j.2095-6045.20260249

文章编号:2095-6045(2026)05-0123-07

移动CT运行中的多传感器融合定位方法

高长天[1]　姚宜斌[1]　王鑫哲[1]

1 武汉大学测绘学院,湖北 武汉,430079

A Multi-sensor Fusion Positioning Method for Mobile CT in Operation

GAO Changtian[1]　*YAO Yibin*[1]　*WANG Xinzhe*[1]

1 School of Geodsy and Geomatics, Wuhan University, Wuhan 430079, China

摘　要:移动计算机断层扫描(computed tomography,CT)近年来在研究和实际应用中不断增多,但针对CT机架在运行过程中的定位问题,相关研究仍然不足。本文提出一种面向移动CT的高精度多传感器融合定位方法。该方法利用ArUco标记的已知边长提供尺度信息,用于单目视觉中的深度恢复,并进一步建立机架与床体之间的坐标变换关系。同时,提出一种静止条件下的视觉-惯性初始化方法,以保证系统能够高效、准确地完成初始化。通过对移动CT运行场景进行充分仿真,并开展系统性实验验证。该方法在深度恢复、初始化和定位精度方面均表现良好,尤其适用于传统单目VIO(visual-inertial odometry)算法因运动激励不足而难以获得可靠结果的场景。

关键词:多传感器融合;机器人定位;ArUco标记;VIO初始化;移动CT

中图分类号:P208;TP751;TP242

文献标志码:A

Abstract: Mobile CT systems have seen increasing adoption in both research and practical applications in recent years; however, research addressing positioning challenges during CT scanner operation remains insufficient. This paper proposes a high-precision multi-sensor fusion positioning method tailored for mobile CT systems. The method utilizes known side lengths annotated with ArUco markers to provide scale information for depth reconstruction in monocular vision, while establishing coordinate transformation relationships between the scanner frame and the patient table. Additionally, a visual-inertial initialization method under stationary conditions is introduced to ensure efficient and accurate system initialization. Comprehensive simulations of mobile CT operation scenarios and systematic experimental validation demonstrate that the proposed method exhibits excellent performance in depth reconstruction, initialization, and positioning accuracy, particularly effective in scenarios where traditional monocular VIO algorithms struggle to produce reliable results due to insufficient motion excitation.

Key words: multi-sensor fusion; robot localization; ArUco marking; VIO initialization; mobile CT

随着机器人技术的快速发展,各类移动机器人在医疗领域中的应用日益广泛[1]。计算机断层扫描(computed tomography,CT)是一种重要的医学成像技术,能够获取人体内部结构的详细断层图像[2],在脑部、胸部和腹腔相关疾病的诊断中具有重要作用。近年来,移动CT因其轻量化和便携性而受到越来越多的关注[3]。

与固定式CT不同,移动CT在工作过程中会发生机架运动。根据CT成像原理,为保证扫描结果的质量,设备需要提供高精度的实时位姿反馈。在移动机器人定位常用的各类传感器中,单目视觉能够获取物体形状、纹理等丰富的环境信息,同时具有成本低、设备轻便等优点[4]。单目视觉-惯性里程计(visual-inertial odometry,VIO)通过融合惯性测量单元(inertial measurement unit,IMU),能够结合相机与IMU各自的优势,从而在实际应用中提高系统鲁棒性[5]。目前,已有多种视觉-惯性融合算法被提出,并应用于不同场景[6-9]。此外,轮式编码器作为轮式机器人中最常见的传感器之一,也经常被用于多传感器融合定位算法[10,11]。

然而,移动CT的运行场景具有明显特殊性:设备在各轴方向上的运动激励并不充分,这会显著影响单目VIO算法的位姿估计精度[10,12],其中尺度不确定性问题尤为突出。虽然双目相机可以在一定程

引用格式:高长天,姚宜斌,王鑫哲. 移动CT运行中的多传感器融合定位方法[J]. 测绘地理信息,2026,51(5):123-129(GAO Changtian,YAO Yibin,WANG Xinzhe. A Multi-sensor Fusion Positioning Method for Mobile CT in Operation[J]. Journal of Geomatics,2026,51(5):123-129)

基金项目:国家重点研发计划(2023YFB3907100)。

度上缓解这一问题，但也会带来额外的硬件成本和计算开销[13]。OpenCV中的ArUco标记具有检测简单、快速且鲁棒性较强的特点，已被用于多个工业应用场景[14-17]。ArUco标记是基于二进制方形二维码的视觉标记，由黑色外框和内部黑白编码矩阵组成，是计算机视觉常用人工标记。由于ArUco标记的实际边长已知，其边长信息可以作为一种约束，为单目VIO算法提供重要的尺度信息。现有研究通常将ArUco标记集成到同步定位与地图构建(simultaneous localization and mapping，SLAM)系统中[15]。此外，ArUco标记也在多传感器融合定位中发挥了重要作用。Xing等[16]采用联邦卡尔曼滤波方法，将标记、光流、超声波和惯性传感器等多源信息进行融合，从而获得连续定位结果；Faria等[18]结合超宽带飞行时间(ultra-wideband time-of-flight，UWB ToF)技术与环境中布设的ArUco标记，开发了一种低成本的精确定位方案。

尽管ArUco标记能够提供丰富信息并用于估计设备位姿，但在现有研究中，当仅使用一个或两个ArUco标记时，其效果可能受到限制。为了提升ArUco标记在移动CT运行场景中的作用，有必要在充分利用视觉信息的同时，进一步融合IMU和轮式编码器数据。

从固定式CT的硬件结构来看，与上述移动CT不同的是，其工作方式较为特殊：运行过程中机架通常保持静止，而床体发生运动，因此高质量扫描主要依赖于对床体的精确控制。相比之下，移动CT面临的挑战在于机架运动控制更加复杂，且机架与床体之间的关系不再固定。因此，在运行过程中获取机架与床体之间的绝对位置关系变得十分必要，而这一关系的获得依赖于系统初始化过程。

由于多传感器定位系统具有较强非线性[5]，为了保证系统运行过程中的准确性和鲁棒性，获得高精度的状态初始值十分关键，这一过程也就是算法中的视觉-惯性初始化步骤。He等[19]将现有初始化方法分为绝对静止法、松耦合法和紧耦合法。其中，大多数紧耦合方法[20,21]和松耦合方法[6,22]都对设备在各方向上的运动激励提出了较高要求，而这一条件在移动CT运行场景中难以满足。绝对静止法通常将初始速度设为零，并利用IMU测量值计算重力矢量和陀螺仪零偏。该方法较为直接，但没有融合视觉信息，同时对IMU数据质量要求较高[23-27]。近年来，也有研究尝试将深度学习方法用于单目VIO算法的初始化[23,24]，但这类方法的适用场景相对有限，并且会消耗更多计算资源。

针对上述问题，本文提出一种面向移动CT的高精度多传感器融合定位方法。该方法不仅充分利用ArUco标记的边长信息，还结合了一种鲁棒的静止初始化方法。

1 面向移动CT的高精度多传感器融合定位方法

1.1 坐标系与符号定义

为便于后续描述，首先对本文涉及的坐标系和主要符号进行说明。本文用到的坐标系及其变换关系如图1所示。除相机坐标系和世界坐标系外，本体坐标系，即IMU所在坐标系，记为 $body$，简记为 b；人为定义的ArUco标记坐标系记为 mrs，简记为 m。从坐标系 j 到坐标系 i 的变换矩阵用 $\boldsymbol{T}_{ij}$ 表示，其形式为：

$$\boldsymbol{T}_{ij}=\begin{bmatrix}\boldsymbol{R}_{ij} & \boldsymbol{t}_{ij}\\ 0_{1\times 3} & 1\end{bmatrix} \tag{1}$$

其中，$\boldsymbol{R}_{ij}$ 为旋转矩阵；$\boldsymbol{t}_{ij}$ 为平移向量。相机观测到的ArUco标记角点记为 c_i^j，j 表示第 i 个标记上的第 j 个角点，角点按左下、左上、右上、右下的顺序编号。该角点在像素坐标系中表示为 $\left(u_i^j, v_i^j\right)$，在相机坐标系中为 $\left(x_i^j, y_i^j, z_i^j\right)$，在相机归一化坐标系中为 $\left(\hat{x}_i^j, \hat{y}_i^j, 1\right)$，在ArUco坐标系中记为 mc_i^j。

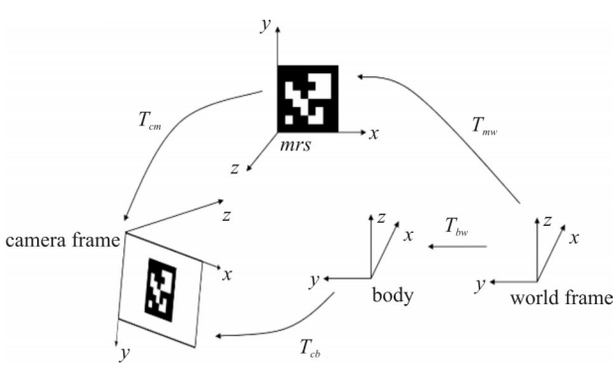


图1 本文方法中所定义的坐标系

Fig. 1 The Coordinate System Defined in the Method

1.2 方法总体框架

本文方法的整体流程如图2所示。算法采用滑动窗口优化，窗口内待估计的状态量定义为：

$$\begin{aligned}\boldsymbol{X}&=[x_1, x_2, \cdots, x_n, \rho_1^c, \rho_2^c, \cdots, \rho_m^c, \rho_1, \rho_2, \cdots, \rho_M]\\ \boldsymbol{x}_k&=[p_{b_k}^w, v_{b_k}^w, q_{b_k}^w, b_a, b_g], k\in[1, n]\end{aligned} \tag{2}$$

式中，ρ_k^c 表示第 k 个角点的深度，ρ_k 表示第 k 个普通视觉特征点的深度；$p_{b_k}^w$ 和 $v_{b_k}^w$ 分别表示第 k 时刻机体系在世界坐标系下的位置和速度；$q_{b_k}^w$ 表示第 k 时刻机体系相对于世界坐标系的姿态；b_a 和 b_g 分别为加

速度计和陀螺仪的零偏。

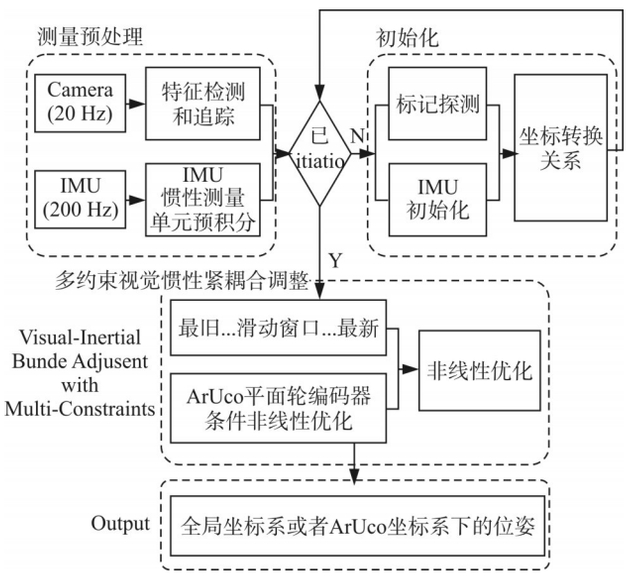


图2　本文方法的基本流程

Fig. 2　Basic Workflow of the Proposed Method

图3为该优化问题对应的因子图。其中，圆形节点表示待估计的状态量，方框表示观测因子。该因子图主要包含惯性因子 i、平面因子 p、速度因子 v、视觉因子 f 和角点因子 c。惯性因子和视觉因子的具体形式在已有文献中已有详细描述[6,9]，此处不再赘述。角点因子主要来源于ArUco标记角点的视觉重投影误差。与常规光流特征不同，本文直接在每一帧图像中检测ArUco标记的角点。此外，针对移动CT工作场景的特点，引入了平面约束以改善 z 轴方向及滚转、俯仰角的估计效果。速度因子则由轮式编码器提供，主要用于抑制尺度漂移。

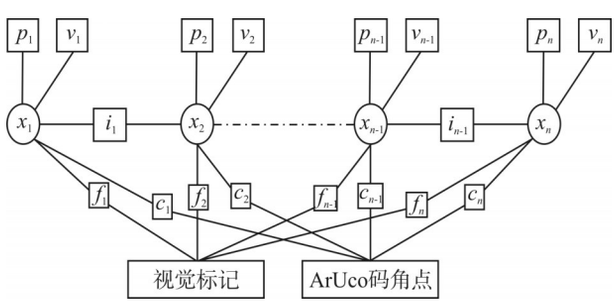


图3　本文方法的因子图

Fig. 3　Factor Diagram of the Proposed Method

1.3　ArUco标记检测与重投影关系

本文采用OpenCV中的ArUco标记检测模块提取角点[14]。为获得角点深度，利用标记的已知边长建立以下观测方程：

$$\begin{cases} S=\sqrt{(x_i^{j^*}-x_i^j)^2+(y_i^{j^*}-y_i^j)^2+(z_i^{j^*}-z_i^j)^2} \\ \sqrt{2}\,S=\sqrt{(x_i^{j^+}-x_i^j)^2+(y_i^{j^+}-y_i^j)^2+(z_i^{j^+}-z_i^j)^2} \end{cases} \tag{3}$$

其中，S 为ArUco标记的实际边长；j 与 j^* 为相邻角点；j 与 j^+ 为对角线角点。一个标记包含四个角点，由此可列出六个约束方程。求解后即可得到各角点的深度，进而计算出相机坐标系下带深度的角点坐标。若多个标记水平排列，还可将标记间距作为附加约束，原理与式(3)类似，本文不再展开。

在建立重投影关系前，需先定义ArUco坐标系：当一个或多个标记水平排列时，以图像左侧第一个标记的左下角点为原点建立坐标系，具体轴向关系见图1。以相机坐标系到ArUco坐标系的变换为例，每个角点满足：

$$mc_i^j=\boldsymbol{R}_{mc}\times\left(x_i^j,y_i^j,z_i^j\right)+t_{mc} \tag{4}$$

对应的残差为：

$$e_i^j=\boldsymbol{R}_{mc}\times\left(x_i^j,y_i^j,z_i^j\right)+t_{mc}-mc_i^j \tag{5}$$

将所有角点的残差汇总，得到如下最小二乘：

$$\min F\left(\boldsymbol{R}_{mc},t_{mc}\right)=\min\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{4}\|e_i^j\|^2 \tag{6}$$

其中，N 表示ArUco标记总数。该最小二乘问题可利用Ceres库[25]，通过Gauss-Newton方法求解。对应的雅可比矩阵可写为：

$$\boldsymbol{J}=[-\boldsymbol{R}_{mc}\times(x_i^j,y_i^j,z_i^j)^{\wedge}\boldsymbol{I}_{3\times3}]^{\mathrm{T}} \tag{7}$$

其中，$(*)^{\wedge}$ 表示向量对应的反对称矩阵；$[*]^{\mathrm{T}}$ 表示矩阵转置。经过迭代优化后可得到 $T_{mc}=[\boldsymbol{R}_{mc}\,t_{mc}]$。同样，上述重投影关系也可建立在两个相机坐标系之间。

1.4　静止状态下的初始化

视觉-惯性算法通常需要先完成初始化。本文结合移动CT的实际工作特点，将整个系统的初始化设计为在设备静止状态下进行。系统参考Ground-Fusion方法[26]判断设备是否静止；只有当视觉、IMU和轮式编码器三类判据同时满足时，才认为设备处于静止状态。此时，初始位置和全局坐标系下的初始速度设为 $0_{3\times1}$，初始姿态和零偏则按照OpenVINS中的方法求得[27]。

前文已经提到，世界坐标系到ArUco坐标系的变换关系是机架定位中需要确定的重要量。基于前面建立的重投影问题，可在最初的 n 个静止阶段中求得相机坐标系到ArUco坐标系的变换，记为 $\left[T_{mc}^1,T_{mc}^2,\cdots,T_{mc}^n\right]$

随后，为了减小单次提取ArUco标记角点时产生的误差影响，对这些变换矩阵取平均：

$$\bar{T}_{mc}=\sum_{i=1}^{n}T_{mc}^i/n \tag{8}$$

其中，$\bar{T}_{mc}$ 为平均后的变换矩阵。这里的 $\sum$ 只是一个简写，并不是对矩阵元素直接求和；涉及姿态平均的部分，按文献[28]中的方法处理。然后结合相机到IMU的外参 T_{i_0c} 和前面求得的初始姿态 R_{i_0w}，世界坐标系到ArUco坐标系的变换 T_{mw} 写为：

$$T_{mw}=\bar{T}_{mc}\times T_{i_0c}^{\mathrm{T}}\times T_{i_0w} \tag{9}$$

其中，

$$T_{i_0w}=\begin{bmatrix}R_{i_0w} & 0_{1\times 3}\\ 0_{1\times 3} & 1\end{bmatrix}$$

这样，便完成了初始化。

1.5 约束项构建

视觉、平面以及 IMU 预积分约束在文献[6,9,10,22]中已经讨论得比较充分，这里不再重复。角点约束仍然从前面的重投影关系出发。除此之外，还需要计算静止阶段内角点坐标的平均因子：

$$[\bar{x}_i^j,\bar{y}_i^j,\bar{z}_i^j]=\sum_{k=1}^{n}[x_{ik}^j,y_{ik}^j,z_{ik}^j] \tag{10}$$

其中，n 是静止时间段内的观测次数。后续每一帧与初始阶段之间建立的重投影关系，作为角点约束加入优化问题。

单目 VIO 不可避免会出现尺度漂移，在匀速运动时这一问题更明显。因此，这里引入轮式因子，把它作为速度约束来减小尺度漂移的影响。具体做法是，将轮式编码器测得的速度看作机体系 x 方向上的一维速度观测，此时残差写为：

$$\boldsymbol{r}_v=\boldsymbol{v}_k^w-\boldsymbol{R}_{b_k}^w\times\boldsymbol{v}_k^o \tag{11}$$

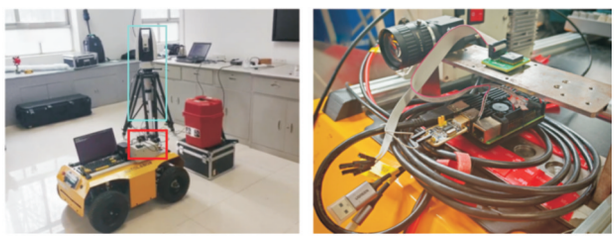
其中，$\boldsymbol{v}_k^w$ 为 k 时刻世界坐标系下的速度；$\boldsymbol{R}_{b_k}^w$ 为 k 时刻从机体系到世界坐标系的旋转矩阵；$\boldsymbol{v}_k^o$ 为 k 时刻轮式编码器的观测值。

2 实验与分析

2.1 实验设计

由于移动CT的实际工作场景比较特殊，本文搭建了图4所示的模拟实验模型，并在此基础上进行实验。图4(a)展示了实验场景，相关设备安装在移动底盘上，位置见图中的红色框。设备之间的固定方式如图4(b)所示，主要性能参数列于表1。

实验所用底盘采用Ackermann模型，并带有轮式编码器输出。图4(a)中的蓝色框为Leica Absolute Tracker AT960，下文简称AT960。AT960能够实时跟踪目标，并给出高精度定位结果。实验中，本文使用AT960记录底盘运动时的参考轨迹。随后，从深度恢复、初始化成功率与效率，以及定位结果三个方面对本文方法进行分析。



(a) (b)

图4 实验场景与设备布置

Fig. 4 Experimental Scenario and Equipment Layout

表1 实验所用设备列表

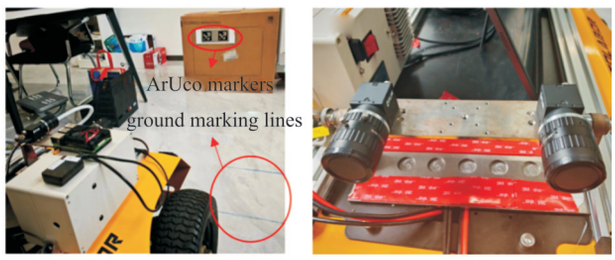
Tab. 1 List of Equipment Used in the Experiment

设备	型号	主要参数
相机	FLIR BFS－U3－50S5	分辨率：1 224×720 频率：20 Hz 角随机游走：$0.2^{\circ}/\sqrt{\mathrm{h}}$
IMU	ADIS16505－2	速度随机游走： $0.012\ \mathrm{m/s}/\sqrt{\mathrm{h}}$ 频率：200 Hz
计算单元	Raspberry Pi 4B	内存：4 GB

2.2 深度结果验证

相机光心的准确位置很难直接标定出来，因此AT960不能直接给出相机光心到ArUco标记之间的真实深度。为此，这里采用一种间接验证方式来检查本文方法的深度恢复精度：首先分别计算两个位置处算法输出的深度，再取这两个深度的差值；同时，用AT960测量这两个位置之间的距离，用作对照，实验布置如图5(a)所示。实验中使用两个边长为7 cm的ArUco标记，两者间隔5 cm；地面上设置两条蓝色标线，间距约30 cm，并使其与ArUco标记所在平面保持垂直。在每个位置，设备静止测量3 s，并取这段时间内算法输出深度的平均值。此外，还搭建了一套双目相机系统作为对比。该系统由两台FLIR BFS-U3-50S5相机组成，分辨率为1 224×720，基线长度为16 cm，通过三角测量计算深度[29]，如图5(b)所示。同时，使用EPnP方法[30]求取ArUco标记在相机坐标系下的深度。最后，对三种方法得到的深度差进行比较。

图6为深度恢复实验得到的误差结果。横轴为每组20次实验中两个静止点到ArUco标记平面的距离中位数。柱状图给出了三种方法在各组实验中的最小误差，其中橙色、青色和紫色分别对应本文方法、三角测量方法和EPnP方法。每个柱形顶部的黑色线段表示该组实验的最大误差。相关统计数据



(a)　　(b)

图5　深度恢复实验布置与双目相机结构

Fig. 5　Layout of the Depth Restoration Experiment and Structure of the Binocular Camera

列于表2。可以看出,本文方法的深度恢复精度高于三角测量法和 EPnP 方法。就本实验采用的 ArUco 标记而言,距离在 1.2～1.5 m 之间时,本文方法的表现最好。

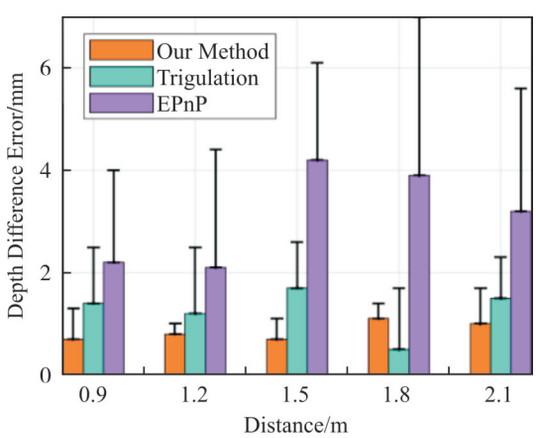


图6　深度恢复差值范围

Fig. 6　Range of Depth Restoration Differences

2.3　初始化结果分析

这一节主要看初始化的成功率和用时,并在移动 CT 的模拟工作场景中,把本文方法和几种基线方法放在一起比较。实验场景里,底盘的运动被限制在前进和后退两个方向上。表 3 列出了 5 组实验结果。VINS-Mono[6] 和 ORB-SLAM3[22] 在静止阶段不能完成初始化,OpenVINS[27] 则可以。基于这一点,序列 1 和序列 2 中,底盘在开始运动前先静止 1 s;序列 3 和序列 4 中静止 2 s;序列 5 中静止 3 s。表中给出的时间,是从底盘开始运动到初始化完成之间的时间。从结果看,本文方法在初始化成功率和用时上与 OpenVINS 接近;而且静止阶段一结束,初始化就可以完成。

2.4　定位精度实验

最后,对几种方法的定位精度进行了统计。图 7 给出了两次实验的定位轨迹:图 7(a)是在光照充足条件下得到的结果,对应图 7(c)中的实验 1;图 7(b)则是在光照较差条件下得到的结果,对应图 7(d)中的实验 2。蓝色实线为 AT960 提供的参考轨迹,橙色轨迹为本文方法的输出。两种光照条件下,本文方法得到的轨迹都与参考轨迹保持了较高的重合度。

另外,还将本文方法与基线方法进行了对比,定位结果列于表 4。结合图 8 可以看出,本文方法在精度和鲁棒性上都更好。尤其是在实验 2 中,OpenVINS 由于视觉观测不稳定而出现失效;相比之下,本文方法通过紧耦合方式融合轮式编码器和视觉测量,因此仍能保持较稳定的定位结果。

为观察轮式编码器对定位的影响,实验分别测试了使用和不使用轮式编码器两种情况。图 9 中,橙色实线为参考轨迹,深棕色实线表示未使用轮式编码器时的定位轨迹,绿色实线表示加入轮式编码器后的定位轨迹。从图中能看出,不加入轮式编码器时,视觉-惯性算法的定位结果存在较明显的尺度偏差;加入轮式编码器后,尺度偏差明显减小。理论上,编码器精度越高,定位结果也会相应更准确。

3　结束语

本文围绕移动 CT 的运行特点,提出了一种高精度多传感器融合定位算法。针对这类场景中运动受限、初始化困难以及尺度不稳定等问题,本文主要完成了以下工作。

一方面,利用 ArUco 角点的已知尺寸进行深度恢复,并通过重投影关系求得不同帧下带有实际尺度的角点坐标。另一方面,设计了静止状态下的初

表2　各距离段的深度恢复误差(红色斜体为最优)/mm

Tab. 2　Depth Recovery Errors for Each Distance Segment (Red Bold Indicates the Optimal Result)/mm

	本文方法			三角测量法[29]			EPnP[30]		
	最小值	最大值	平均值	最小值	最大值	平均值	最小值	最大值	平均值
0.9 m	*0.70*	*1.30*	*0.95*	1.40	2.50	1.95	2.20	4.00	3.13
1.2 m	*0.80*	*1.00*	*0.84*	1.20	2.50	1.88	2.10	4.40	3.46
1.5 m	*0.70*	*1.10*	*0.89*	1.70	2.60	2.05	4.20	6.10	4.92
1.8 m	1.10	*1.40*	*1.28*	*0.50*	1.70	1.40	3.90	7.00	5.15
2.1 m	*1.00*	*1.70*	*1.35*	1.50	2.30	1.85	3.20	5.60	3.77

表 3 不同方法的初始化用时

Tab. 3 Initialization Time for Different Methods

序列	本文方法	VINS-mono[6]	ORB-SLAM3[22]	OpenVINS[27]
1	<0.1 s	失败	失败	<0.1 s
2	<0.1 s	3.2 s	4.4 s	<0.1 s
3	<0.1 s	失败	失败	<0.1 s
4	<0.1 s	失败	失败	<0.1 s
5	<0.1 s	5.8 s	失败	<0.1 s

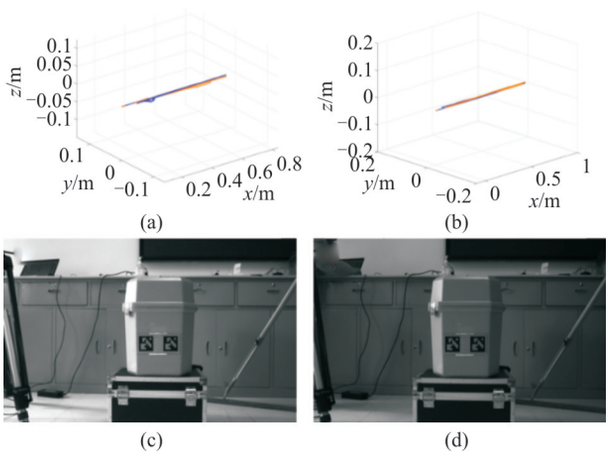


图 7 定位轨迹与光照条件

Fig. 7 Localization Trajectory and Light Condition

表 4 两组实验的 ATE RMSE/cm

Tab. 4 ATE RMSE for both Experimental Groups /cm

序列	本文方法	OpenVINS (with TrackAruco)
实验 1	0.69	2.04
实验 2	1.02	失败

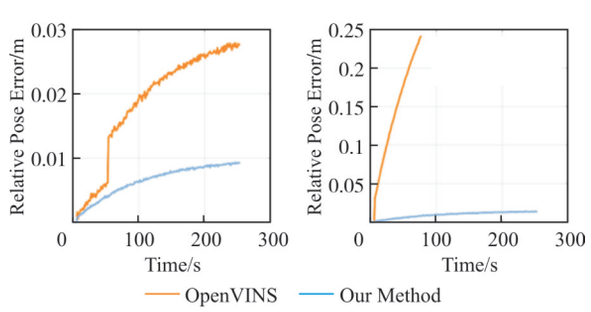


图 8 实验中的相对位姿误差

Fig. 8 Relative Pose Error in the Experiment

始化方法，使系统能够在较短时间内完成较准确的初始化。另外由实验 2 可知，OpenVINS 由于视觉观测不稳定而出现失效的问题在本文方法中得以解决，此为本文方法优势所在模拟场景中的实验结果表明，该方法可以满足移动 CT 场景下的定位需求。

不过，要将该方法进一步用于真实移动 CT 运行环境，还需要继续改进初始化后 ArUco 标记不可见时的定位精度。另外，本文采用了因子图方法，这种方法实时性较差，但是从实际情况看，一次检测过

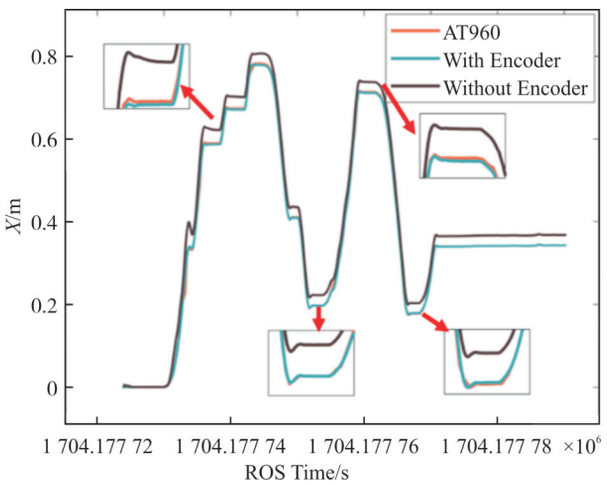


图 9 加入轮式编码器后的定位精度改善

Fig. 9 Improvement in Positioning Accuracy with the Addition of a Rotary Encoder

程仅需 2～3 min，因此其带来的影响可忽略不计。双目相机在这一场景中的作用也值得进一步考察。随着移动 CT 产品开发的推进，后续还需要在真实手术场景中开展实验，以检验方法在实际应用中的表现。

参 考 文 献

[1] Dupont P E, Nelson B J, Goldfarb M, et al. A Decade Retrospective of Medical Robotics Research from 2010 to 2020[J]. Science Robotics, 2021, 6(60): eabi8017

[2] Hermena S, Young M. CT-Scan Image Production Procedures [M]//StatPearls. Treasure Island: StatPearls Publishing, 2023. https://www.ncbi.nlm.nih.gov/books/NBK574519/

[3] Andersson H, Tamaddon A, Malekian M, et al. Comparison of Image Quality Between a Novel Mobile CT Scanner and Current Generation Stationary CT Scanners[J]. Neuroradiology, 2023, 65(3): 503-512

[4] Abaspur Kazerouni I, Fitzgerald L, Dooly G, et al. A Survey of State-of-the-Art on Visual SLAM[J]. Expert Systems with Applications, 2022, 205: 117734

[5] Huang Guoquan. Visual-Inertial Navigation: A Concise Review[C]//2019 International Conference on Robotics and Automation (ICRA). 2019: 9 572-9 582

[6] Qin Tong, Li Peiliang, Shen Shaojie. VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator[J]. IEEE Transactions on Robotics, 2018, 34(4): 1 004-1 020

[7] Sun Fengming, Liu Zuoyuan, Zhong Ruiyang, et al. Toward Large-Scale PSD-Based Visible Light Positioning with Visual-Inertial Fusion[J]. Measurement Science and Technology, 2025, 36(2): 026311

[8] Wu Jianfeng, Xiong Jian, Guo Hang. Improving Ro-

bustness of Line Features for VIO in Dynamic Scene [J]. Measurement Science and Technology, 2022, 33(6): 065204

[9] Forster C, Carlone L, Dellaert F, et al. On-Manifold Preintegration for Real-Time Visual-Inertial Odometry [J]. IEEE Transactions on Robotics, 2017, 33(1): 1-21

[10] Wu K J, Guo C X, Georgiou G, et al. VINS on Wheels [C]//2017 IEEE International Conference on Robotics and Automation (ICRA). May 29-June 3, 2017, Singapore. New York: IEEE Press, 2017: 5 155-5 162

[11] Liang Bowen, Tao Yourui, Zhu Huabo, et al. Advancing Robust State Estimation of Wheeled Robots in Degenerate Environments: Harnessing Ground Manifold and Motion States[J]. Measurement Science and Technology, 2024, 35(4): 046308

[12] Martinelli A. Visual-Inertial Structure from Motion: Observability and Resolvability[C]//2013 IEEE/RSJ International Conference on Intelligent Robots and Systems. November 3-7, 2013, Tokyo, Japan. New York: IEEE Press, 2013: 4 235-4 242

[13] Yang Yulin, Huang Guoquan. Observability Analysis of Aided INS with Heterogeneous Features of Points, Lines, and Planes[J]. IEEE Transactions on Robotics, 2019, 35(6): 1 399-1 418

[14] Bradski G. The OpenCV Library[J]. Dr. Dobb's Journal: Software Tools for the Professional Programmer, 2000, 25(11): 120-123

[15] Liu Yuanpeng, Zhou Jun, Li Yida, et al. A High-Accuracy Pose Measurement System for Robotic Automated Assembly in Large-Scale Space[J]. Measurement, 2022, 188: 110426

[16] Xing Boyang, Zhu Quanmin, Pan Feng, et al. Marker-Based Multi-sensor Fusion Indoor Localization System for Micro Air Vehicles[J]. Sensors, 2018, 18(6): 1706

[17] Roos-Hoefgeest S, Garcia I A, Gonzalez R C. Mobile Robot Localization in Industrial Environments Using a Ring of Cameras and ArUco Markers[C]//IECON 2021 —47th Annual Conference of the IEEE Industrial Electronics Society. October 13-16, 2021, Toronto, ON, Canada. New York: IEEE Press, 2021: 5 081-5 086

[18] Faria S D N. Sensor Fusion for Mobile Robot Localization Using UWB and ArUco Markers[D]. Lisbon: Universidade Nova de Lisboa, 2021

[19] He Yijia, Xu Bo, Ouyang Zhanpeng, et al. A Rotation-Translation-Decoupled Solution for Robust and Efficient Visual-Inertial Initialization [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 17-24, 2023, Vancouver, BC, Canada. New York: IEEE Press, 2023: 739-748

[20] Martinelli A. Closed-Form Solution of Visual-Inertial Structure from Motion [J]. International Journal of Computer Vision, 2014, 106(2): 138-152

[21] Domínguez-Conti J, Yin Jianfeng, Alami Y, et al. Visual-Inertial SLAM Initialization: A General Linear Formulation and a Gravity-Observing Non-Linear Optimization[C]//2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). October 16-20, 2019, Munich, Germany. New York: IEEE Press, 2019: 37-45

[22] Campos C, Elvira R, Rodríguez J J G, et al. ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial, and Multimap SLAM [J]. IEEE Transactions on Robotics, 2021, 37(6): 1 874-1 890

[23] Zhou Yunwen, Kar A, Turner E, et al. Learned Monocular Depth Priors in Visual-Inertial Initialization [C]//Computer Vision—ECCV 2022. Cham: Springer, 2022: 552-570

[24] Almalioglu Y, Turan M, Saputra M R U, et al. Self-VIO: Self-Supervised Deep Monocular Visual——Inertial Odometry and Depth Estimation[J]. Neural Networks, 2022, 150(C): 119-136

[25] McCord T B, Sotin C. Ceres: Evolution and Current State[J]. Journal of Geophysical Research: Planets, 2005, 110(E5): 2004JE002244

[26] Yin Jie, Li Ang, Xi Wei, et al. Ground-Fusion: A Low-Cost Ground SLAM System Robust to Corner Cases[PP/OL]. 2024: arXiv: 2402.14308. https://arxiv.org/abs/2402.14308

[27] Geneva P, Eckenhoff K, Lee W, et al. OpenVINS: A Research Platform for Visual-Inertial Estimation[C]//2020 IEEE International Conference on Robotics and Automation (ICRA). May 31-August 31, 2020. Paris, France. IEEE, 2020: 4 666-4 672

[28] Markley F L, Cheng Yang, Crassidis J L, et al. Averaging Quaternions[J]. Journal of Guidance, Control, and Dynamics, 2007, 30(4): 1 193-1 197

[29] Thurmond V A. The Point of Triangulation[J]. Journal of Nursing Scholarship, 2001, 33(3): 253-258

[30] Lepetit V, Moreno-Noguer F, Fua P. EPnP: An Accurate O(n) Solution to the PnP Problem[J]. International Journal of Computer Vision, 2009, 81(2): 155-166

修回日期:2026-09-03

第一作者:高长天,硕士,研究方向:大地测量学与测量工程。E-mail:2023202140036@whu.edu.cn

通信作者:姚宜斌,博士,教授。现主要从事全球导航卫星系统(GNSS)数据处理、对流层与电离层建模、卫星大地测量学理论与方法及其应用。E-mail:ybyao@whu.edu.cn