引文格式:赵奕澎,姚宜斌,王鑫哲.利用ArUco码的单目VIO快速初始化算法[J].武汉大学学报(信息科学版),2026,51(5):979-986.DOI:10.13203/j.whugis20240174

Citation:ZHAO Yipeng,YAO Yibin,WANG Xinzhe.A Fast Initialization Algorithm for Monocular VIO Using ArUco Markers[J].Geomatics and Information Science of Wuhan University,2026,51(5):979-986.DOI:10.13203/j.whugis20240174

利用ArUco码的单目VIO快速初始化算法

赵奕澎[1] 姚宜斌[1,2] 王鑫哲[1]

1 武汉大学测绘学院,湖北 武汉,430079

2 武汉大学地球空间环境与大地测量教育部重点实验室,湖北 武汉,430079

摘 要:针对目前单目视觉惯性里程计(visual inertial odometry,VIO)算法初始化对设备运动状态和视觉观测条件要求较为严格,在工程应用中难以保证的问题,提出了一种利用ArUco码辅助的单目VIO快速初始化算法。该算法结合ArUco码已知的边长信息,通过边长约束恢复深度信息并获取相机与ArUco码之间的空间关系。同时结合惯性测量单元(inertial measurement unit,IMU)静止初始化流程与事先标定的相机与IMU之间的外参,实现全局坐标系与ArUco码之间空间关系的获取,完成初始化流程。实验结果表明,算法角点深度恢复误差平均值为0.75 mm,相比VINS-Fusion、OpenVINS及ORB-SLAM3的单目VIO方案的初始化算法,在效率及对后续VIO的收敛帮助上有着更好的表现,初始化后小范围2 m内VIO定位误差为0.009 m,姿态精度为0.4°,且对单目VIO整体轨迹精度带来一定程度的提升。

关键词:单目VIO;ArUco码;快速初始化

中图分类号:P228 **文献标识码**:A **收稿日期**:2025-02-11

DOI:10.13203/j.whugis20240174 **文章编号**:1671-8860(2026)05-0979-08

A Fast Initialization Algorithm for Monocular VIO Using ArUco Markers

ZHAO Yipeng[1] *YAO Yibin*[1,2] *WANG Xinzhe*[1]

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

2 Key Laboratory of Geospace Environment and Geodesy, Ministry of Education, Wuhan University, Wuhan 430079, China

Abstract: **Objectives**: Current monocular visual inertial odometry (VIO) algorithms impose stringent requirements on device motion states and visual observation conditions during initialization, which are often challenging to meet in real-world engineering applications. A fast and stable initialization method is thus essential. **Methods**: The proposed method utilizes the known edge length of ArUco markers to recover depth information through edge length constraints, thereby estimating the spatial relationship between the camera and the ArUco markers. Concurrently, the inertial measurement unit (IMU) is initialized under stationary conditions. By incorporating the pre-calibrated extrinsic transformation between the camera and IMU, the spatial relationship between the global coordinate system and the ArUco markers is established. **Results**: The accuracy of depth recovery at corner points was validated using profiles of the differences between the measured distance of two points and the recovered depth. Multiple experiments under both well-lit and low-light conditions showed that the mean error of this difference was 0.75 mm. The VIO positioning accuracy of the device within a certain spatial range after initialization was further evaluated, and several experiments demonstrated a positioning error of 0.009 m and an attitude error of 0.4°. Additional experiments were conducted to assess the stability and efficiency of the proposed method. Compared with the monocular VIO initialization methods used in VINS-Fusion, OpenVINS, and ORB-SLAM3, the proposed method achieved higher stability and efficiency, while also yielding a better overall trajectory in the subsequent VIO process. **Conclusions**: The proposed method demonstrates high efficiency and robustness in initialization, with nota-

基金项目:国家自然科学基金重点项目(42330105)。

第一作者:赵奕澎,硕士,研究方向为多传感器融合定位。2018302141345@whu.edu.cn

通信作者:王鑫哲,博士,助理研究员。Wxz_gnss@whu.edu.cn

ble improvements in positioning accuracy during subsequent VIO tracking.

Key words: monocular visual inertial odometry; ArUco marker; fast initialization

高精度且可靠的定位服务,是智能机器人在家庭、商场、医疗及工业等场景下完成作业任务的重要保障[1]。定位技术中的各种传感器技术中[2-3],单目视觉可以获得物体形状、纹理等丰富的环境信息,但光照变化、运动物体及绝对尺度未知等问题都会严重影响单目视觉定位精度[4]。融合惯性测量单元(inertial measurement unit, IMU)的单目视觉惯性里程计(visual inertial odometry, VIO)结合了两种传感器的优点,在不同的应用场景下具有更高的鲁棒性。但单目VIO系统高度非线性化,为了保证系统运行过程中的精度和鲁棒性,获取高精度的状态初始值,算法的初始化步骤至关重要。现有的文献研究主要将初始化步骤分为紧组合与松组合两种形式[5]。紧组合先利用陀螺仪测量和校准的外部参数来估计相机旋转,然后利用视觉和加速度计观测构建的约束方程来求解初始速度和重力矢量[6-7]。松组合结合视觉运动恢复结构法估计的相机姿态及利用IMU测量值来估计初始状态变量[8-10]。近年来也有研究将深度学习的算法应用到VIO系统的初始化之中[11-12]。但现有的算法初始化精度可能受视觉观测的影响,且对初始化过程中的运动状态往往有着复杂且严格的要求,在实际工程应用中这是难以保障的,因此设计一种快速稳定的初始化算法,更好地贴合实际工程应用场景,是十分重要的工作。

作为机器视觉领域最常见的地标之一,二维码已经被应用在很多机器人的定位导航系统之中。其中OpenCV中的ArUco码具有检测简单、快速且鲁棒性较强的特点,已有诸多研究针对ArUco码在定位系统中的应用展开了研究[13-14],本文借鉴现有的基于ArUco码的定位算法,利用ArUco码的已知尺寸进行辅助,提出了一种在仅使用一到两个ArUco码的情况下,针对单目VIO系统的快速初始化算法,解决了单目相机无法获取绝对尺度信息的问题。同时算法基于松组合的思想,完成了IMU初始状态的估计,并通过事先标定的外参,将视觉信息与IMU信息融合,实现整个VIO系统在静止状态下的快速初始化流程,主要流程如图1所示。

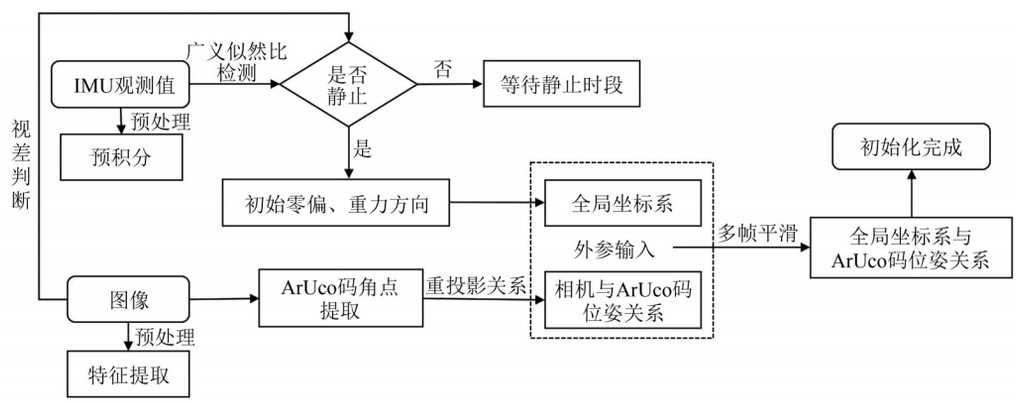


图1 利用ArUco码辅助的单目VIO初始化算法流程图

Fig. 1 Flowchart of Monocular VIO Initialization Algorithm with ArUco Marker Assistance

1 理论与方法

1.1 符号说明

为了便于描述和表达,首先定义一些符号和坐标系,本文涉及到的坐标系及转换关系如图2所示,图中g为全局坐标系(即世界坐标系),b为IMU自身坐标系,f为每一帧图像对应的相机坐标系,m代表人为定义的ArUco码坐标系。坐标系之间的转换关系表示为$T_{i,j}$,即代表由j坐标系到i坐标系的坐标变换矩阵,由旋转矩阵和平移向量组成:

$$T_{i,j}=\left[\boldsymbol{R}_{i,j}\quad \boldsymbol{t}_{i,j}\right] \tag{1}$$

式中,$\boldsymbol{R}_{i,j}$表示两个坐标系之间的旋转关系;$\boldsymbol{t}_{i,j}$表示两个坐标系之间的平移关系。

1.2 静止检测与IMU初始化

1.2.1 静止检测

在涉及到IMU的导航定位算法中,零速修正是常见的约束手段[15-16],其中涉及到的零速检测的算法主要包括:加速度幅值检测、加速度滑动方差检测、角速度幅值检测以及角速度滑动方差检测等。本文针对IMU的静止检测基于广义似

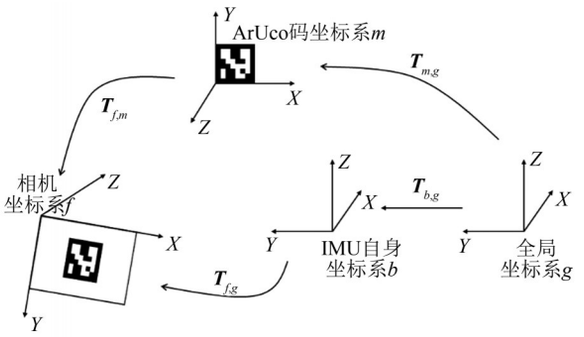


图 2　坐标系及转换关系

Fig. 2　Coordinate System and Transformation Relationship

然比检测算法[17]，同时由于视觉观测易受动态物体的影响。本文对视觉观测也进行了静止状态检测，利用图像提取得到的特征点，计算窗口内每一帧与最新帧的特征点视差平均值，具体表达式为：

$$V=\frac{1}{m}\sum_{i=0}^{m-1}\sum_{j=0}^{r-1}\left\| p_i^j-p_{m-1}^j \right\|^2 \tag{2}$$

式中，m代表滑动窗口内图像帧的总数；r代表匹配到的特征点总数；p_i^j表示第i帧匹配到的第j个特征点；‖ ‖表示计算模长。设置视觉静止检测阈值t_V，当计算得到的$V<t_V$时，认为视觉静止检测通过。当IMU静止检测与视觉静止检测均通过时，认为设备处于静止状态，继续执行后续初始化流程。

1.2.2　IMU状态初始化

IMU状态初始化的主要目的是为了获取IMU在起始阶段的朝向以及加速度计与陀螺仪零偏的初值。由于通常认为重力方向为垂直地面向下，因此根据重力向量与加速度观测量，获取IMU状态与重力方向的关系，即IMU初始状态与全局坐标系的旋转关系。具体使用到OpenVINS中施密特正交化的算法[18]，恢复出g坐标系到IMU初始状态的旋转矩阵$\boldsymbol{R}_{i_0,g}$，以及IMU陀螺零偏初值b_0^g以及IMU加速度计零偏初值b_0^a。

1.3　ArUco码角点检测与坐标关系获取

1.3.1　角点检测与深度恢复

本文使用了OpenCV中的ArUco码，并使用OpenCV中的ArUco模块进行角点的检测。利用检测到的角点坐标$\left(u_i^j,v_i^j\right)$，其中i表示第i个ArUco码，j表示第i个ArUco码的第j个角点。人为规定角点的排列顺序为左下、右下、右上、左上。与事先标定得到的相机内参矩阵$\boldsymbol{K}$，得到角点在f坐标系下的坐标$\left(\hat{x}_i^j,\hat{y}_i^j,1\right)$，此时的坐标尚不包含深度信息。为了获取角点坐标的深度信息，本文使用ArUco码已知的边长S，对其中一个AR码，设未知的角点的深度值$\boldsymbol{d}=\left[d_i^1\ d_i^2\ d_i^3\ d_i^4\right]$，则角点在$f$坐标系下的坐标$\left(x_i^j,y_i^j,z_i^j\right)=\boldsymbol{d}\left(\hat{x}_i^j,\hat{y}_i^j,1\right)$。理论上$f$坐标系下相邻角点间的距离即为ArUco码的边长：

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

式中，上标j^*与j即代表相邻的两个角点。同理，对角线上两个角点的距离可表示为：

$$\sqrt{2}\,S=\sqrt{(x_i^{j^+}-x_i^j)^2+(y_i^{j^+}-y_i^j)^2+(z_i^{j^+}-z_i^j)^2} \tag{4}$$

式中，上标j^+与j即代表位于对角线上的两个角点。对于一个ArUco码的4个角点，可由以上两种约束列出6个方程，联合求解后得到$\boldsymbol{d}$，然后得到角点在f坐标系下带有深度信息的坐标$\left(x_i^j,y_i^j,z_i^j\right)$。当存在多个ArUco码水平排列时，可以增加ArUco码之间的距离作为新的约束条件。由于与上述约束原理相同，本文不再赘述。

1.3.2　相机坐标系与ArUco码坐标系的转换

本文自定义了ArUco码坐标系，当一个或多个码水平排列时，定义在图像中左侧第一个ArUco码的左下角点为ArUco码坐标系原点，X轴与Y轴位于ArUco码平面，Z轴垂直于ArUco码平面，具体坐标轴空间关系见图2。对每个角点c_i^j，根据坐标转换关系，在m坐标系下表示为：

$$\boldsymbol{m}_i^j=\boldsymbol{R}_{m,f}\left(x_i^j,y_i^j,z_i^j\right)+\boldsymbol{t}_{m,f} \tag{5}$$

残差可表示为$\boldsymbol{e}_i^j=\boldsymbol{R}_{m,f}\left(x_i^j,y_i^j,z_i^j\right)+\boldsymbol{t}_{m,f}-\boldsymbol{m}_i^j$，将所有角点的误差方程列出，构建最小二乘：

$$\min\left(\boldsymbol{R}_{m,f},\boldsymbol{t}_{m,f}\right)=\min\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{4}\left\|\boldsymbol{e}_i^j\right\|^2 \tag{6}$$

式中，N代表ArUco码的个数。对$\boldsymbol{R}_{m,f}$利用泰勒展开公式求导，可以求出每个误差项对状态变量的导数：

$$\begin{cases}\dfrac{\partial \boldsymbol{e}_i^j}{\partial \boldsymbol{R}_{m,f}}\approx-\boldsymbol{R}_{m,f}\left(x_i^j,y_i^j,z_i^j\right)^{\cdot}\\[2ex]\dfrac{\partial \boldsymbol{e}_i^j}{\partial \boldsymbol{t}_{m,f}}=\boldsymbol{I}\end{cases} \tag{7}$$

式中，$(\ )^{\cdot}$代表向量的反对称矩阵；$\boldsymbol{I}$为单位矩阵。根据导数的计算，雅可比矩阵$\boldsymbol{J}_i^j$和参数$\Delta\boldsymbol{x}$可表示为：

$$\begin{cases}\boldsymbol{J}_i^j=\left[\dfrac{\partial \boldsymbol{e}_i^j}{\partial \boldsymbol{R}_{m,f}}\ \ \dfrac{\partial \boldsymbol{e}_i^j}{\partial \boldsymbol{t}_{m,f}}\right]^{\mathrm{T}}\\[2ex]\Delta\boldsymbol{x}=\left[\Delta\boldsymbol{R}_{m,f}\ \ \Delta\boldsymbol{t}_{m,f}\right]^{\mathrm{T}}\end{cases} \tag{8}$$

再根据高斯牛顿法，构建增量方程[19]：

$$\boldsymbol{J}(x)\boldsymbol{J}^{\mathrm{T}}(x)\Delta\boldsymbol{x}=-\boldsymbol{J}(x)f(x) \tag{9}$$

式中，$f(x)$为目标函数；$\boldsymbol{J}(x)$为目标函数对应的雅可比矩阵。

本文使用Ceres非线性优化库[20]，通过多次迭代运算，完成最小二乘问题的求解，得到最终

的参数结果 $T_{m,f}=\left[\boldsymbol{R}_{m,f}\ \ \boldsymbol{t}_{m,f}\right]$。

进一步地,在设备保持静止的时间段内,由上述步骤可以得到初始状态 n 帧的相机到ArUco码坐标转换矩阵序列 $\left[T_{m,f}^{1}\ T_{m,f}^{2}\cdots\ T_{m,f}^{n}\right]$,对序列取平均:

$$T_{m,f}=\sum_{i=1}^{n}\frac{T_{m,f}^{i}}{n} \tag{10}$$

此处的求和为广义的求和,针对旋转矩阵可以将其转换为四元数的形式后再求和,得到初始状态相机坐标系与ArUco码坐标系的转换矩阵,同时利用事先标定得到的外参矩阵 $T_{i_0,f}$,与之前求得的IMU初始状态与 g 坐标系的旋转关系 $\boldsymbol{R}_{i_0,g}$,得到全局坐标系与ArUco码坐标系的转换矩阵:

$$T_{m,g}=T_{m,f}T_{i_0,f}^{\mathrm{T}}\boldsymbol{R}_{i_0,g} \tag{11}$$

1.3.3 *初始化在后续VIO中的应用*

在完成上述步骤后,已获得了IMU的 b_0^g、b_0^a 和 $\boldsymbol{R}_{i_0,g}$,以及ArUco码角点的深度以及 g 坐标系与ArUco码坐标系之间的转换关系,本文称该状态即为完成初始化状态。但事实上,在该状态下,除了ArUco码角点外的视觉特征点仍未包含深度信息。因此,在初始化完成后刚开始的VIO阶段中,仍需ArUco码角点的辅助以获得更多视觉特征点的深度值。

在设备开始运动后,利用相机帧与帧之间的视觉特征点三角化技术,可以得到包含ArUco码角点在内的所有视觉特征点的相机坐标系下坐标,但此时的深度信息并不包含尺度意义,这里对ArUco码角点的深度值进行约束,对每个角点 c_i^j,将利用三角化得到的深度值记为 $\hat{z}_i^j$,在前文中已经获得了其带有尺度信息的深度 z_i^j,尤其可以对每个角点计算出一个尺度因子 s_i^j:

$$s_i^j=\frac{z_i^j}{\hat{z}_i^j} \tag{12}$$

对所有ArUco码角点的尺度因子取平均,得到平均尺度因子 $\bar{s}$,最后将该尺度因子应用到全部视觉特征点,得到所有特征点带有尺度信息的深度值,在后续VIO算法框架中使用到的视觉特征点坐标均为带有尺度信息的坐标。

2 实验与测试分析

2.1 实验设备及场景

实验的图像采集使用FLIR BFS-U3-50S5黑白相机,设置相机自动曝光,图像分辨率设置为1 224×720像素,图像采集的频率设置为20 Hz。MEMS级IMU选用了ADIS16505-2,采样频率设置为200 Hz,陀螺零偏稳定性为2.7 °/h,加表零偏稳定性为50 mg,角度随机游走为0.2 °/$\sqrt{\mathrm{h}}$,速度随机游走为0.012 m/s/$\sqrt{\mathrm{h}}$。相机与IMU之间的固定方式如图3所示,并利用Kalibr工具[21]对相机坐标系与IMU坐标系之间的外参进行了标定。

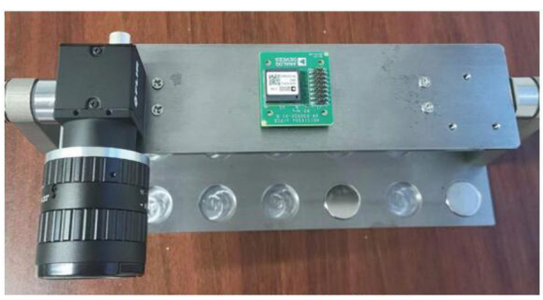


图3 相机与IMU之间的固定方式

Fig. 3 Fixation Between the Camera and IMU

实验中,使用了两个边长为7 cm的ArUco码,其间距为5 cm,同时参考了文献[22]对使用ArUco码定位定姿的研究,选择将ArUco码贴在距离设备初始位置约1.5 m处的正前方,以便进行ArUco码的观测与角点提取。为了充分验证本文算法的精度,实验使用Leica TS60测量机器人(测距、定位精度为1 mm,简称TS60),将Leica 360°棱镜放置在移动底盘之上,以便TS60进行测量跟踪,作为实验测距与定位的参考真值。其中移动底盘、ArUco码与TS60之间的位置关系如图4所示。



图4 移动底盘、ArUco码与Leica TS60之间的位置关系

Fig. 4 Placement Between the Mobile Chassis, ArUco Markers and the Leica TS60

2.2 实验结果分析

使用实测数据评估本文所提出的单目VIO初始化算法的性能与精度。为了验证提出的方案在光线较暗情况下的鲁棒性,实验针对光线充

足与光线较暗两种情况分别进行了数据采集与处理，两种情况下的典型场景如图 5 所示。



(a) 光线较暗　　(b) 光线充足

图 5　典型场景光线对比

Fig. 5　Typical Scenes Comparison of Lighting

由于相机光心位置无法准确辨别，即无法利用 TS60 直接获取相机光心与 ArUco 平面之间的准确深度值，因此实验选择间接验证本文算法深度恢复的精度。先人为在地面上标记出间距 30 cm 的两个点，并确保两点之间连线与 ArUco 码平面尽可能垂直。随后让设备多次停留在两点之间，对比 TS60 测量所得两点之间距离（简称 TS60 深度差值）与算法深度恢复出两点间的深度差值（简称算法深度差值）。图 6 给出了光线较暗的 20 组实验下前 18 组实验 TS60 深度差值与本文算法深度差值的结果对比图，空心圆虚线表示 TS60 深度差值与本文算法深度差值之间的差值，最后两组实验结果与前 18 组实验接近，TS60 深度差值与本文算法深度差值之间的差值分别为 0.91 mm 和 0.87 mm。光线较暗和光线充足情况下的精度统计值如表 1 所示，在两种情况下的平均深度差值平均值为 0.75 mm。

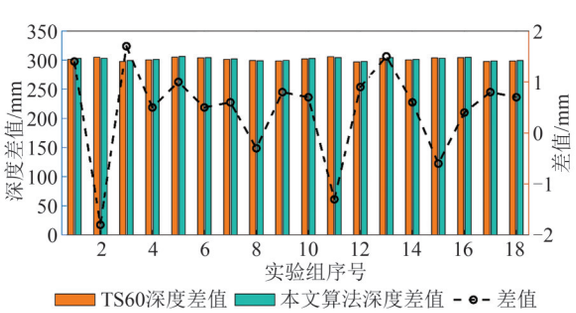


图 6　多组实验深度差值对比

Fig. 6　Comparison of Depth Differences in Multiple Experiments

表 1　深度差值实验精度统计值

Table 2　Depth Difference Experimental Accuracy Statistics

光线情况	实验组数/组	深度差值平均值/mm	深度差值标准差/mm
光线充足	20	0.62	0.39
光线较暗	20	0.89	0.46

为了证明本文算法的稳定性与可用性，将 VINS-Fusion 框架的初始化模块修改为本文算法，对比本文算法与原始 VINS-Fusion 单目 VIO、OpenVINS 单目 VIO、ORB-SLAM3 单目 VIO 与双目 VIO 5 种情况下的初始化情况与定位轨迹。共进行了 5 组实验，5 组实验的初始化环境及典型场景如图 7 所示，其中设备在静止启动后提供了充分的旋转及平移，以辅助 VINS-Fusion 及 ORB-SLAM3 单目 VIO 方案的初始化，最终初始化情况统计如表 2 所示。在实验中，本文算法与 OpenVINS 的 IMU 静止初始化窗口时间均设置为 1 s，且两种方法在窗口结束后均可立即完成初始化。尽管运动底盘能够提供充足的运动激励条件，VINS-Fusion 和 ORB-SLAM3 仍出现了初始化失败的情况，且该问题在室外环境中尤为明显。进一步地，5 组实验的平面轨迹如图 8 所示，从左至右依次为 5 组实验的轨迹，其中，以 ORB-SLAM3 双目 VIO 的轨迹（粉红色轨迹）为参考，本文算法对应的轨迹（淡紫色轨迹）整体表现出色，相比其他 3 种单目 VIO 方案，整体轨迹的漂移现象最小，这也侧面印证了本文算法在初始化阶段为 VIO 系统提供了出色的状态初值。但由于本文所进行的实验处于纹理不丰富场景，视觉条件恶劣，因此在长时间的运行中，依旧出现了较为明显的尺度偏差，未来将重点研究在此类挑战环境下单目 VIO 精度及稳定性提升的算法，以更好地实现单目 VIO 算法在各种工程环境下的应用。

为了更直观地体现出本文算法的初始化精度，利用 TS60 测量机器人对设备启动状态进行跟踪，重点分析不同 VIO 方案在初始化完成后小范围内（Y 方向前进值＜2 m）的定位精度；同时利用超核电子 CH104M 姿态测量模块（姿态角精度 0.2°）作为姿态参考结果，分析姿态变化精度。



(a) 室内1 (b) 室内2 (c) 室内+室外 (d) 室外1 (e) 室外2

图7 5组实验的初始化场景及典型场景

Fig. 7 Initial Conditions and Representative Scenarios of the Five Experiments

表2 5组实验初始化情况统计表

Table 2 Statistical Table of the Initialisation in Five Experiments

场景	本文算法		VINS-Fusion 单目VIO		OpenVINS 单目VIO		ORB-SLAM3 单目VIO		ORB-SLAM3 双目VIO	
	结果	耗时/s	结果	耗时/s	结果	耗时/s	结果	耗时/s	结果	耗时/s
室内1	√	1	√	2.7	√	1	√	4.4	√	0.4
室内2	√	1	√	3.1	√	1	√	4.2	√	0.3
室内+室外	√	1	√	2.5	√	1	√	5.1	√	0.4
室外1	√	1	×	—	√	1	√	4.7	√	0.5
室外2	√	1	×	—	√	1	×	—	√	0.5

注:√表示初始化成功,×表示初始化失败。

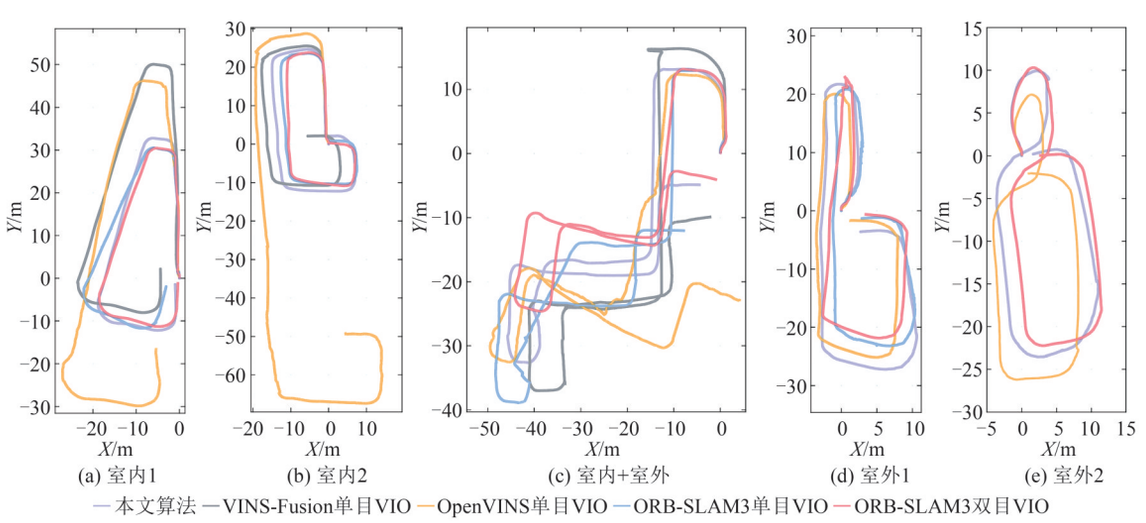


图8 5组实验定位轨迹平面图

Fig. 8 Plan View of Positioning Trajectory Under Five Experiments

对5组实验进行平均后,具体统计结果如表3所示。本文算法初始化成功后小范围内表现出色。定位误差和姿态误差均优于其他3种单目VIO方案,与ORB-SLAM3双目VIO初始化精度接近。

3 结 语

本文提出了一种利用ArUco码辅助的单目VIO快速初始化算法,利用已知尺寸的ArUco码恢复ArUco码角点深度信息,并构建ArUco码与相机坐标系间的坐标关系;融合IMU静止初始化结果与事先标定的相机与IMU之间的外参,完成全局初始化。经实验验证,本文所提出的算法角点深度恢复不同光线条件下多组实验的平均深度差值平均值为0.75 mm,且将本文算法与VINS-Fusion单目VIO、OpenVINS单目VIO、ORB-SLAM3单目VIO进行对比,本文算法在初始化性能及后续VIO算法收敛上均有着出色的表现,与ORB-SLAM3双目VIO初始化精度接近。为了更直观地反映初始化的具体精度,选择对初始化成功2 m范围内的位姿精度进行分

表 3　不同算法初始化成功后位姿精度分析

Table 3　Sitional Accuracy Analysis After Successful Initialisation of Different Algorithms

算法	定位误差/m	姿态误差/(°)
本文算法	0.009	0.4
VINS-Fusion 单目 VIO	0.032	1.2
OpenVINS 单目 VIO	0.021	0.6
ORB-SLAM3 单目 VIO	0.013	1.1
ORB-SLAM3 双目 VIO	0.005	0.3

析，本文算法的位置和姿态误差分别为 0.009 m 及 0.4°，也优于其他 3 种单目 VIO 初始化方案，从而为单目 VIO 的整体表现带来一定程度上的提升。

参　考　文　献

[1] ZHAO J W, WU Y B, BAO W, et al. Indoor Robot Mapping and Navigation System Based on Cyber-Physical Systems: Integration of SLAM Algorithm and Visual Information [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2025, 1-11.

[2] 崔晓珍，周琪，武东杰，等. GNSS/IMU 与里程计紧-松耦合的因子图融合定位方法[J]. 武汉大学学报(信息科学版)，2024，49(10)：1911-1921.
CUI Xiaozhen, ZHOU Qi, WU Dongjie, et al. Factor Graph Fusion Localization Method with Tight and Loose Coupling of GNSS/IMU and Odometry [J]. *Geomatics and Information Science of Wuhan University*, 2024, 49(10): 1911-1921.

[3] 辜声峰，戴春齐，何成鹏，等. 面向城市车载导航的多系统 PPP-RTK/VIO 半紧组合算法性能分析[J]. 武汉大学学报(信息科学版)，2021，46(12)：1852-1861.
GU Shengfeng, DAI Chunqi, HE Chengpeng, et al. Analysis of Semi-Tightly Coupled Multi-GNSS PPP-RTK/VIO for Vehicle Navigation in Urban Areas [J]. *Geomatics and Information Science of Wuhan University*, 2021, 46(12): 1852-1861.

[4] 廖健驰，李星星，冯绍权. GVIL：基于图优化的 GNSS PPP/视觉/惯性/激光雷达紧组合算法[J]. 武汉大学学报(信息科学版)，2023，48(7)：1204-1215.
LIAO Jianchi, LI Xingxing, FENG Shaoquan. GVIL: Tightly-Coupled GNSS PPP/Visual/INS/LiDAR SLAM Based on Graph Optimization [J]. *Geomatics and Information Science of Wuhan University*, 2023, 48(7): 1204-1215.

[5] HE Y J, XU B, OUYANG Z P, et al. A Rotation-Translation-Decoupled Solution for Robust and Efficient Visual-Inertial Initialization [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, Canada, 2023.

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

[7] DOMÍNGUEZ-CONTI J, YIN J F, ALAMI Y, et al. Visual-Inertial SLAM Initialization: A General Linear Formulation and a Gravity-Observing Non-Linear Optimization [C]//IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Munich, Germany, 2018.

[8] 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): 1874-1890.

[9] QIN T, LI P L, SHEN S J. VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator [J]. *IEEE Transactions on Robotics*, 2018, 34(4): 1004-1020.

[10] 张晨阳，杨健. 一种自适应点线特征和 IMU 耦合的视觉 SLAM 方法[J]. 武汉大学学报(信息科学版)，2025，50(10)：2048-2063.
ZHANG Chenyang, YANG Jian. A Visual SLAM Method Coupled with Adaptive Point-Line Features and IMU [J]. *Geomatics and Information Science of Wuhan University*, 2025, 50(10): 2048-2063.

[11] 史殿习，高云琦，宋林娜，等. Deep-Init：基于深度学习的视觉惯性里程计非联合初始化方法[J]. 计算机科学，2024，51(7)：327-336.
SHI Dianxi, GAO Yunqi, SONG Linna, et al. Deep-Init: Non Joint Initialization Method for Visual Inertial Odometry Based on Deep Learning [J]. *Computer Science*, 2024, 51(7): 327-336.

[12] ZHOU Y W, KAR A, TURNER E, et al. Learned Monocular Depth Priors InVisual-Inertial Initialization [C]// Computer Vision-ECCV, Switzerland, 2022.

[13] 牛小骥，王庭蔚，葛雯斐，等. 一种基于视觉二维码/惯性融合的室内高精度定位定姿方法[J]. 中国惯性技术学报，2023，31(11)：1067-1075.
NIU Xiaoji, WANG Tingwei, GE Wenfei, et al. A High Precision Indoor Positioning and Attitude Determination Method Based on Visual Two-Dimensional Code/Inertial Information[J]. *Journal of Chinese Inertial Technology*, 2023, 31(11): 1067-1075.

[14] 邢伯阳，潘峰，冯肖雪. ArUco-SLAM：基于 ArUco 二维码阵列的单目实时建图定位系统[J]. 北京

理工大学学报, 2020, 40(4): 427-433.
XING Boyang, PAN Feng, FENG Xiaoxue. ArUco-SLAM: A Monocular SLAM System Based on ArUco Landmark Array[J]. *Transactions of Beijing Institute of Technology*, 2020, 40(4): 427-433.

[15] 王甫红, 栾梦杰, 程雨欣, 等 城市环境下智能手机车载GNSS/MEMS IMU紧组合定位算法[J]. 武汉大学学报(信息科学版), 2023, 48(7): 1106-1116.
WANG Fuhong, LUAN Mengjie, CHENG Yuxin, et al. Smartphone GNSS/MEMS IMU Tightly-Coupled Integration Positioning Method for Vehicular Navigation in Urban Conditions[J]. *Geomatics and Information Science of Wuhan University*, 2023, 48(7): 1106-1116.

[16] ZHANG Q Y, XU X B, SHI F Y, et al. An Integrated Positioning Method with IMU/UWB Based on Geometric Constraints of Foot-to-Foot Distances[J]. *Measurement*, 2025, 242: 115771.

[17] YIN J, LI A, XI W, et al. Ground-Fusion: A Low-Cost Ground SLAM System Robust to Corner Cases[C]//IEEE International Conference on Robotics and Automation (ICRA), Yokohama, Japan, 2024.

[18] GENEVA P, ECKENHOFF K, LEE W, et al. OpenVINS: A Research Platform for Visual-Inertial Estimation[C]//IEEE International Conference on Robotics and Automation (ICRA), Paris, France, 2020.

[19] 高翔, 张涛, 刘毅, 等. 视觉SLAM十四讲: 从理论到实践[M]. 北京: 电子工业出版社, 2017.
GAO Xiang, ZHANG Tao, LIU Yi, et al. Fourteen Lectures on Visual SLAM: From Theory to Practice[M]. Beijing: Publishing House of Electronics Industry, 2017.

[20] AGARWAL S, MIERLE K. Ceres Solver[EB/OL]. (2023-01-01)[2024-08-11]. http://ceres-solver.org/tutorial.html.

[21] FURGALE P, REHDER J, SIEGWART R. Unified Temporal and Spatial Calibration for Multi-Sensor Systems[C]//IEEE/RSJ International Conference on Intelligent Robots and Systems, Tokyo, Japan, 2013.

[22] 葛雯斐, 牛小骥, 蒋郡祥, 等. 智能手机利用二维码路标进行定位定姿的方法研究[J]. 传感技术学报, 2019, 32(12): 1816-1823.
GE Wenfei, NIU Xiaoji, JIANG Junxiang, et al. Research on Position and Posture Determination of SmartPhone Using 2D Code Landmark[J]. *Chinese Journal of Sensors and Actuators*, 2019, 32(12): 1816-1823.