

单目视觉里程计位姿解算质量评估因素分析

李凯林 李建胜 王安成 李雪强 杨子迪

(信息工程大学地理空间信息学院 郑州 450000)

摘要:近年来,弹性PNT和综合PNT不断发展,多源传感器的弹性融合机制受到广泛关注。针对多源组合导航中视觉传感器的测量精度和可靠性受环境影响大,对精度的影响程度难以单独评估等问题,通过研究单目视觉里程计位姿解算质量的评估因素,体现多源传感器融合系统中视觉传感器的测量质量,提出基于李代数推导的反应空间点构型对相机位姿恢复精度影响的精度衰减因子 Visual-DOP,设计了单目视觉里程计位姿解算质量评估算法,并在视觉里程计权威数据集KITTI上做了验证,通过控制变量的方法解耦每一个评估因素与定位定姿精度的关系,结果显示采用的位姿解算质量评估因素能够正确反应单目视觉里程计的导航精度受视觉测量质量的影响程度。

关键词:弹性融合;视觉精度衰减因素;视觉里程计;位姿解算;视觉质量评估

中图分类号: TP391.41 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Analysis of pose calculation quality evaluation factors of monocular visual odometry

Li Kailin Li Jiansheng Wang Ancheng Li Xueqiang Yang Zidi

(School of Geospatial Information, Information Engineering University, Zhengzhou 450000, China)

Abstract: In recent years, with the continuous development of resilient PNT and integrated PNT, the elastic fusion mechanism of multi-source sensors has attracted wide attention. Aiming at the problem that the measurement accuracy and reliability of visual sensors in multi-source integrated navigation are greatly affected by the environment, and the degree of influence on the accuracy is difficult to be evaluated independently. The evaluation factors of the pose calculation quality of monocular visual odometry were analyzed to reflect the measurement quality of visual sensors in multi-source sensor fusion system. A precision attenuation factor Visual-DOP based on the influence of the configuration of the reaction space point on the pose recovery accuracy of the camera was proposed. Further, the pose calculation quality evaluation algorithm of monocular visual odometry was designed, and it was verified on the visual odometry authoritative data set KITTI. The connection between each evaluation factor and the positioning accuracy was decoupled by the method of control variables. The results indicate that the pose calculation quality evaluation factors can correctly reflect the influence of visual measurement quality on the navigation accuracy of monocular visual odometer.

Keywords: resilient integration; Visual-DOP; visual odometry; pose calculation; visual quality assessment

0 引言

随着综合导航与授时(positioning, navigation and time, PNT)体系^[1]和弹性PNT框架^[2]的发展,多源传感器的弹性融合机制成为了当下的研究热点,而在多数场景下,视觉传感器测量的精度变化较大,与其他传感器融合时权重难以确定,导致融合系统在挑战环境下的精度下降。因此,建立一套完善的视觉传感器测量质量评估指标来反应测量精度的变化,分析视觉传感器测量质量的评估因素,对

弹性融合导航系统的算法策略选择、测量协方差调整、以及在避免整体定位精度发散等方面具有重要意义。

视觉里程计(visual odometry, VO)是基于摄像头采集图像在连续运动过程中恢复相机位姿的定位方法。按照相机种类划分,常见的有单目视觉里程计、双目视觉里程计和三原色深度(red green blue-depth, RGB-D)视觉里程计,单目视觉里程计区别于双目和RGB-D视觉里程计的主要在于单目VO无法恢复真实尺度^[3],在多源传感器融合导航系统中,尺度信息通常由惯导等其他传感器提供。目前在

多源融合导航方向,关于视觉定位质量评估的研究较少,但也有少量文献涉及了部分研究,Won 等^[4]根据卫星和视觉特征的总体测量方程,提出了衡量卫星和空间点几何分布强度权重精度因子 (weighted dilution of precision, WDOP),并将 WDOP 值应用于一种选择性组合导航滤波系统,仿真和实验表明该系统能够有效提高定位精度。王高杰等^[5]利用相机与空间点之间的方向矢量的余弦构建系数阵,通过系数阵求解协方差矩阵精度因子 (dilution of precision, DOP)来反应位姿解算质量,但该协方差矩阵为 4 维,不能反应 6 自由度的位姿变量。

针对视觉传感器测量质量评估问题,本文用 n 点透视法 (perspective- n -point, PnP) 点数、平均重投影误差、历史点数量、视觉精度衰减因子 (visual-dilution of precision, Visual-DOP) 值 4 个评估因素来描述视觉定位定姿解算的质量好坏,其中 Visual-DOP 基于李代数推导的反应空间点构型对相机位姿恢复精度影响的精度衰减因素提出,此外,基于位姿真值模拟多源融合导航过程中其他传感器对视觉测量的辅助,设计了单目视觉里程计位姿解算质量评估算法,在视觉里程计权威数据集 KITTI 上做了测试,通过控制变量的研究方法对实验数据进行分析,实验结果表明,本文采用的评估因素能够较好地反映视觉导航过程中的精度变化,为多传感器融合导航算法的视觉传感器精度估计以及测量方程定权提供借鉴意义。

1 位姿解算评估因素

纯视觉导航的过程的核心是三角化空间点和 PnP 解算相机位姿两步,而视觉测量的质量往往体现在这两步的解算中,甚至在多源传感器融合的实时定位与建图 (simultaneous localization and mapping, SLAM) 过程中,视觉测量过程也只是多了整体的光束法平差 (bundle adjustment, BA) 和滑动窗口 BA 等步骤^[6],其优化所使用的初值也是基于三角化点坐标和 PnP 解算位姿的结果。目前,多源融合导航中,对视觉测量质量的把控停留在是否满足最小 PnP 解算点数的层面,即解决视觉定位的有无问题,对定位质量的评估考虑相对较少。

单目视觉里程计的位姿解算由 PnP 过程完成,三角化过程解算的空间点位置精度直接影响到 PnP 结果,而在三角化过程中,不能保证空间点的位置解算完全准确,存在随机误差或粗差,从而传递到的 PnP 解算的结果中。其次,在 PnP 过程中,如果跟踪的图像像素点、相机中心、空间点不能完全在一条直线上,也会造成 PnP 解算的误差,这种图像点与空间点之间的配准程度也存在随机误差或粗差。根据随机误差和粗差的性质,可以预测,1)PnP 点数的增加能在一定程度上弱化随机误差。2)平均重投影误差越小,点对的配准程度以及空间点的坐标准确度越高。3)如果空间点坐标被优化的次数越多,点的定位精度更高,这种误差的影响越小。4)与卫星导航理论类似,空间点的空间分布

越好,误差的弱化程度越好。因此,将三角化空间点和 PnP 解算过程中可能评估位姿解算质量的因素提取出来,以 PnP 点数、平均重投影误差、历史点数量、Visual-DOP 值为位姿解算评估因素,并验证这些因素对评估单目视觉里程计位姿解算质量的正确性。

本文对视觉测量质量评估因素进行了分析,提出基于李代数推导的视觉精度衰减因子 Visual-DOP 来衡量特征点空间构型的好坏,其次,针对 4 种评估因素基于真值模拟外部传感器对视觉传感器的尺度辅助,设计单目视觉里程计位姿解算质量评估算法,最后基于 KITTI 数据集用控制变量法进行实验解耦评估因素对位姿精度的影响,技术路线如图 1 所示。

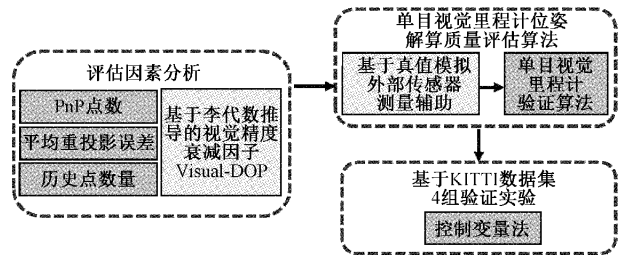


图 1 技术路线

1.1 PnP 点数

PnP 是求解 3D 到 2D 点对运动的方法。所需的最少点数一般取决于所使用的 PnP 方法种类,如 3 点透视法 (perspective-3-point, P3P) 方法至少需要 3 对 3D-2D 点,直接线性变换 (direct linear transformation, DLT) 方法至少需要 6 对 3D-2D 点,此外,常见的 PnP 方法还有高效透视 n 点算法 (efficient perspective- n -point, EPnP)、非标定透视 n 点算法 (uncalibrated perspective- n -point, UPnP) 方法、最小化重投影误差迭代法等,它们利用更多信息,而且使用迭代的方式对相机位姿进行优化。其中 EPnP 方法^[7]是 SLAM 和视觉导航过程中常用的 PnP 方法,能够充分利用点位约束,优化多点信息,将 PnP 问题转化为绝对定向问题,获得高效稳定的闭式解。EPnP 方法构造了一组虚拟的空间控制点 $\mathbf{p}_j^w (j = 1, 2, 3, 4)$,将图像检测到的空间点的坐标 $\mathbf{y}_i^c (i = 1, 2, \dots, n)$ 表示为这些虚拟控制点坐标的加权和:

$$\mathbf{y}_i^w = \sum_{j=1}^4 \lambda_{ij} \mathbf{p}_j^w, \sum_{j=1}^4 \lambda_{ij} = 1 \quad (1)$$

$$\mathbf{y}_i^c = \sum_{j=1}^4 \lambda_{ij} \mathbf{p}_j^c, \sum_{j=1}^4 \lambda_{ij} = 1 \quad (2)$$

式中:上标 w 表示世界坐标系,上标 c 表示相机坐标系, λ_{ij} 为加权系数。

解出 \mathbf{p}_j^w 后,将 PnP 问题转化为绝对定向误差最小化问题:

$$\min \sum_{j=1}^4 \|\mathbf{p}_j^c - (\mathbf{R}\mathbf{p}_j^w + \mathbf{t})\|^2 \quad \text{s. t. } \mathbf{R} \in SO(3) \quad (3)$$

式中: \mathbf{R}, \mathbf{t} 分别为世界坐标系到相机坐标的旋转矩阵和平

移向量,也是待求的定位结果。

PnP 点数反映位姿解算约束方程的个数,可以在一定程度上约束测量过程中的偶然误差,得到更加准确的定位结果。

1.2 平均重投影误差

重投影误差是利用求解的相机位姿将空间测量点重新投影到图像平面上,再与像平面观测到的空间点图像坐标作差得到的结果。考虑 n 个空间点世界系坐标齐次形式 $\mathbf{P}_i = [X_i, Y_i, Z_i, 1]^T (i = 1, 2, \dots, n)$, 在图像平面的投影坐标 $\mathbf{p}_i = [u_i, v_i, 1]^T (i = 1, 2, \dots, n)$, 其中解算的世界系到相机系的旋转矩阵和平移向量分别为 \mathbf{R}, \mathbf{t} , 存在如下关系:

$$s_i \mathbf{p}_i = \mathbf{K} \mathbf{T} \mathbf{P}_i \quad (4)$$

$$\mathbf{T} = [\mathbf{R} \mid \mathbf{t}] \quad (5)$$

式中: $s_i (i = 1, 2, \dots, n)$ 为尺度因子, \mathbf{K} 为相机内参矩阵, \mathbf{T} 为坐标转换矩阵。

则平均重投影误差定义如下:

$$error_{\text{reproject}} = \sqrt{\frac{1}{n} \sum_i \left\| \mathbf{p}_i - \frac{1}{s_i} (\mathbf{K} \mathbf{T} \mathbf{P}_i) \right\|^2} \quad (6)$$

平均重投影误差反应了由空间点对解算位姿的拟合程度,也能够一定程度上反应空间点的三角化精度。

1.3 历史点数量

历史点是指在多帧观测中重复出现,被连续观测,且被多次三角化,定位精度更高的空间三维点。对应在多源融合 SLAM 中被持续优化的关键帧共视点坐标,被优化的次数越多,与整体地图的拟合程度越好。而在纯视觉导航中,我们采取对这类点进行多次三角化,取共视帧公共点三角化坐标的平均值作为该点的最终坐标,并用于后续的 PnP 解算,来提高定位精度。假设某空间点的世界系坐标为 $\mathbf{P}_1 = [X_1, Y_1, Z_1]^T$, 在当前帧被三角化后的结果为 $\mathbf{P}_2 = [X_2, Y_2, Z_2]^T$, 则该点新的世界系坐标 \mathbf{P} 取:

$$\mathbf{P} = \frac{1}{2} (\mathbf{P}_1 + \mathbf{P}_2) \quad (7)$$

当该点坐标被优化的次数大于预设值,则被认定因为一个历史点,若新的三角化结果与旧三角化结果之间的空间距离超过一定阈值,则判断为不稳定点,取消该点坐标值,认为该点未被三角化。通过统计一次 PnP 解算中历史点的数量可以在一定程度上判断解算所用空间点坐标的精度好坏。历史点标记如图 2 所示,红紫色表示历史点,颜色越趋近红色表示被优化的次数越多,蓝色表示非历史点。



图 2 历史点标记图

1.4 Visual-DOP

除三点解算 P3P 方法外,EPnP 等方法在本质上是利用相机观测模型建立测量方程,求最优解的过程。我们将不便于表示成变量的待求位姿矩阵 \mathbf{T} 转换成李代数形式,建立测量方程,通过系数矩阵得到表达空间点几何构型的因素 Visual-DOP 值。将空间点世界坐标系坐标 \mathbf{P}_i^w 变换到相机坐标系坐标 \mathbf{P}_i^c :

$$\mathbf{P}_i^c = (\mathbf{T} \mathbf{P}_i^w)_{1:3} = [X_i^c, Y_i^c, Z_i^c] \quad (8)$$

$$\mathbf{T} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{bmatrix} \quad (9)$$

式中: $[X_i^c, Y_i^c, Z_i^c]$ 为空间点相机坐标系下坐标, \mathbf{R}, \mathbf{t} 分别为世界坐标系到相机坐标的旋转矩阵和平移向量。

由相机投影模型得到测量值:

$$\mathbf{z}_i = \begin{bmatrix} u_i \\ v_i \end{bmatrix} = \begin{bmatrix} f_x \frac{X_i^c}{Z_i^c} + c_x \\ f_y \frac{Y_i^c}{Z_i^c} + c_y \end{bmatrix} \quad (10)$$

式中: \mathbf{z}_i 为测量值, (u_i, v_i) 为空间点投影在图像像素下的坐标, f_x, f_y 为相机内参矩阵中的 x, y 方向焦距, (c_x, c_y) 为图像平面像主点坐标。

对位姿矩阵 \mathbf{T} 对应的李代数 \mathbf{X} 左乘扰动 $\delta \mathbf{X}$, 由链式法则和李代数扰动求导模型:

$$\frac{\delta \mathbf{z}_i}{\delta \mathbf{X}} = \frac{\delta \mathbf{z}}{\delta \mathbf{P}_i^c} \frac{\delta \mathbf{P}_i^c}{\delta \mathbf{X}} = - \begin{bmatrix} \frac{f_x}{Z_i^c} & 0 & -\frac{f_x X_i^c}{Z_i^{c2}} \\ 0 & \frac{f_y}{Z_i^c} & -\frac{f_y Y_i^c}{Z_i^{c2}} \end{bmatrix} [\mathbf{I}, -\mathbf{P}_i^{c \wedge}] \quad (11)$$

式中: \mathbf{I} 为单位阵, 上标 \wedge 表示向量取反对称阵, 展开得到 2×6 的雅可比矩阵。

$$\frac{\delta \mathbf{z}_i}{\delta \mathbf{X}} = \mathbf{H}_i = - \begin{bmatrix} \frac{f_x}{Z_i^c} & 0 & -\frac{f_x X_i^c}{Z_i^{c2}} & -\frac{f_x X_i^c Y_i^c}{Z_i^{c2}} & f_x + \frac{f_x X_i^{c2}}{Z_i^{c2}} & -\frac{f_y Y_i^c}{Z_i^{c2}} \\ 0 & \frac{f_y}{Z_i^c} & -\frac{f_y Y_i^c}{Z_i^{c2}} & -f_y - \frac{f_y Y_i^{c2}}{Z_i^{c2}} & \frac{f_y X_i^c Y_i^c}{Z_i^{c2}} & \frac{f_y X_i^c}{Z_i^{c2}} \end{bmatrix} \quad (12)$$

由 n 点组成总测量方程:

$$\delta \mathbf{z} = \mathbf{H} \delta \mathbf{X} + \mathbf{v} \quad (13)$$

$$\mathbf{H} = [\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_n]^T \quad (14)$$

式中: \mathbf{v} 表示测量噪声, \mathbf{H} 是总的系数矩阵。

DOP 值通常由观测方程系数矩阵构造的协因数矩阵的迹的平方根得到, 导出 Visual-DOP 的表达式:

$$DOP_{\text{Visual}} = \sqrt{\text{trace}((\mathbf{H}^T \mathbf{H})^{-1})} \quad (15)$$

类似于卫星几何精度衰减因子 (geometric dilution of precision, GDOP) 对卫星和定位点之间的几何构型的表达, Visual-DOP 值也表征了测量的空间点与待求相机位置之间的几何构型。

2 单目视觉里程计位姿解算质量评估算法

以上所构建的影响因素的作用体现在空间点三角化和 PnP 解算相机位姿两个方面,为了验证以上影响因素是否能够反应视觉测量值的好坏,设计了基于单目视觉里程计的位姿解算质量评估算法,利用权威数据集 KITTI 进行验证实验。该算法主要分为图像处理、初始运动结构恢复 (structure from motion, SfM)、后续帧导航和关键帧三角化 4 部分,基于给定的真值位姿模拟多源传感器融合过程中其他传感器对单目视觉尺度缺失的辅助,重点分析在单目视觉导航过程中的定位定姿质量,因此本文算法只针对视觉测量质量评估,不做实际导航用途。算法框架如图 3 所示。

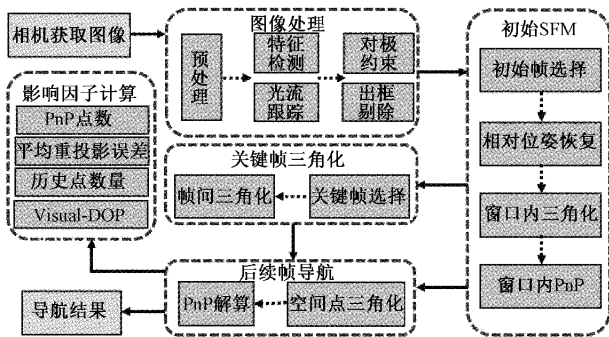


图 3 算法框架图

2.1 图像处理

相机获取的图像会经过灰度化和自适应直方图均衡^[8]两步进行预处理,自适应直方图均衡可以防止图像过于曝光或弱曝光。如图 4 所示,算法在视觉前端采用 Shi-Tomasi 角点法^[9]检测特征点,利用 LK 光流法进行特征点跟踪^[10],光流法跟踪相比于其他特征点检测匹配算法在处理速度上更快,在实际的导航应用中能够获得更好的实时性。为了避免错误跟踪和跟踪出框,参考冉峰等^[11]、任彬等^[12]的单目视觉里程计模型对跟踪点进行基于随机抽样一致性(random sample consensus, RANSAC)方法的对极约束剔除外点并进行出框点剔除,确保良好的跟踪和检测效果。设置检测点之间的最小距离,避免小区域内的密集检测,同时当图像框内特征点数量不够时,补充检测新的特征点,确保一定的特征点数量。

2.2 初始 SfM

初始 SfM 流程图如图 5 所示。初始 SfM 的目的是为了选择初始关键帧,并且三角化一定数量的空间点,作为后续导航过程的基础^[13]。

选取一定的窗口长度进行初始 SfM,把窗口的末尾帧作为一个关键帧,往前搜索,当某一帧满足视差和共视点条件时选取为初始关键帧,利用对极几何求解本质矩阵,从而恢复两个关键帧之间的旋转矩阵 R 和平移向量 t ,由于 t 是不具有尺度特征的单位向量,因此,模拟多源融合导航其



图 4 特征点检测与光流跟踪效果

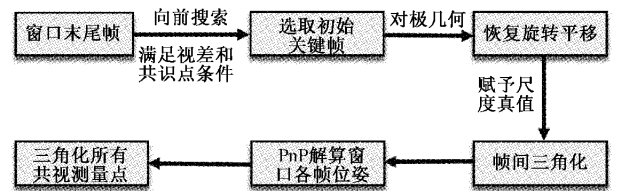


图 5 初始 SfM 流程图

他传感器对视觉导航的尺度辅助,赋予 t 尺度真值。以初始关键帧所在的相机坐标系为世界坐标系,则世界坐标系到两个关键帧的位姿 T_{k_0} 和 T_{k_1} 已知:

$$T_{k_0} = \begin{bmatrix} I & 0 \\ 0 & 1 \end{bmatrix}, T_{k_1} = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \quad (16)$$

式中: I 为单位阵。

对两帧之间的共视点进行三角化,再利用三角化点坐标 PnP 解算窗口内各相机帧位姿,最后再三角化所有共视测量点,初始 SfM 观测和构成如图 6 所示。

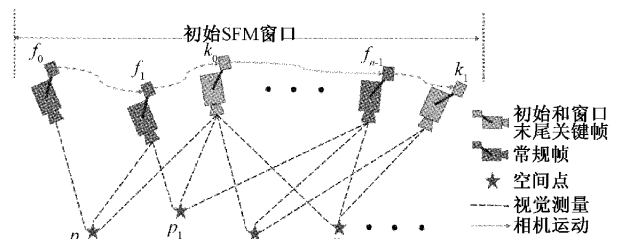


图 6 初始 SfM 观测图

2.3 后续帧导航

如图 7 所示,后续帧的导航分为三角化图像测量点和 PnP 解算相机位姿两步。算法对当前帧测量点进行监测,如果某测量点被当前帧的上一帧和一些历史的图像帧所共视,当共视关系满足一定的里程长度,则对该点进行三角化,为了避免单目视觉里程计尺度发散,模拟多源融合导航过程中惯导等其他传感器对视觉的尺度辅助,在空间点的三角化过程中所使用的相机位姿用真值坐标代替,确保单目视觉里程计位姿解算质量评估算法全过程不产生尺度发散。若该点已经被三角化,则取旧三角化坐标与新三角化

结果的平均值作为最终三角化结果。再利用所有当前帧测量的三角化点进行 PnP 解算恢复相机位姿, PnP 解算的初值为上一帧的相机位姿。

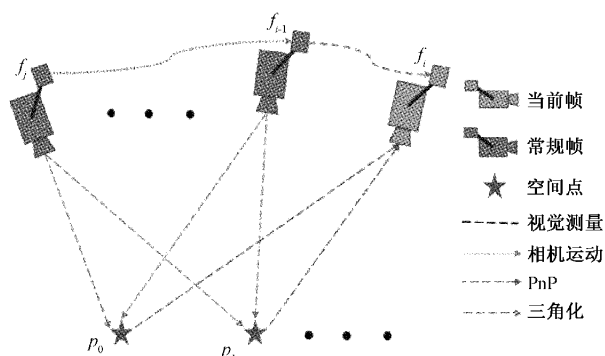


图7 后续帧导航示意图

针对在纯旋转运动或基线过小情况下存在的三角化失败问题,为了保证轨迹的平滑,当三角点过少无法进行 PnP 解算时,采用经典的八点法^[14],结合 RANSAC 方法求解本质矩阵来恢复相邻帧的旋转平移。

2.4 关键帧三角化

为了尽可能地使三角化空间点的坐标更加准确,从初始的 SFM 窗口关键帧开始,根据视差和共视点条件不断判断和选择新的关键帧,考虑当前帧与前一关键帧之间的共

视点数量为 n , 其中第 i 个共视点的在当前帧和前一关键帧的图像坐标分别为 $p_i^1 = [u_i^1, v_i^1]^T$ 和 $p_i^2 = [u_i^2, v_i^2]^T$ 。当平均视差与共视点数量满足条件则将当前帧选择为关键帧:

$$\begin{cases} \sqrt{\frac{1}{n} \sum_{i=1}^n \| (p_i^1 - p_i^2) \|^2} > \alpha \\ n > \beta \end{cases} \quad (17)$$

式中: α 和 β 分别为平均视差和共视点数量的经验阈值。

选取关键帧后,对相邻关键帧进行帧间三角化,与后续帧导航过程的三角化一致,三角化所用相机位姿同样取真值位姿,避免尺度发散对三角化点坐标带来的误差。

3 实验验证与评估

算法轨迹与真实轨迹的拟合程度是评估因素验证分析的基础,只有当定位结果保持全局的稳定才能保证影响因素分析的正确性。由于本文算法针对视觉位姿解算评估因素做分析和验证,同时模拟多源融合算法中惯导等传感器对视觉导航的尺度辅助,因此采用位姿真值做三角化解算,避免单目视觉里程计尺度发散,用来做视觉质量评估,而非实际导航算法。选用 KITTI 视觉里程计数据集^[15]的 00、01、02、03 段对本文算法进行测试,真值轨迹由卫星惯导组合导航结果得到。将算法得到的轨迹与真值轨迹,进行对比,测试结果如图 8 所示。

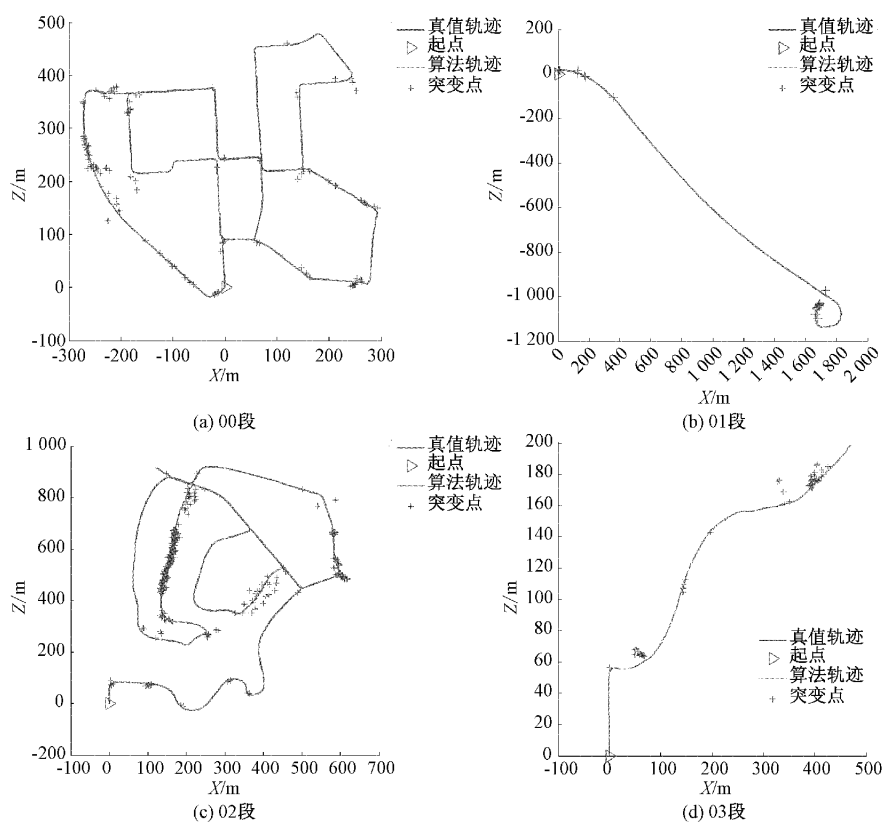


图8 评估算法与真值轨迹对比图

图 8 显示轨迹中存在一些定位误差过大或存在跳变导致轨迹不连续的突变点,是由于本算法在三角化过程中存在一些偏差较大的空间点,因此导致了位姿解算的不稳定,存在较大误差和轨迹突变。这些点误差过大,影响了评估的渐变性质,不纳入统计范围。表 1 显示了 4 个数据

段对应的总里程、数据组数、正常数据组数和以正常数据计算的绝对位置均方根误差 (root-mean-squared error, RMSE),结果显示本文基于位姿真值三角化的单目视觉里程计位姿解算质量评估算法得到的轨迹能够拟合真实轨迹,满足评估条件。

表 1 KITTI 数据集测试数据

数据段	场景类型	总里程/m	总数据组数/组	正常数据/组	绝对位置误差 RMSE/m
00	街区道路	3 724.190	4 541	4 389	0.561
01	城际公路	2 452.200	1 101	1 061	0.632
02	城镇道路	5 065.700	4 661	4 169	0.707
03	乡镇道路	559.926	801	735	0.792

对场景相对复杂且拟合度相对较高的 00 段中 4 389 组数据进行汇总统计,利用近似控制变量的方法研究 PnP 点数、平均重投影误差、历史点数量、Visual-DOP 值 4 个评估因素与绝对位姿精度的关系,采用表 2 确定指标过滤数据,进行 4 组实

验。注:表中符号“—”表示该量为待统计的变量,其他量取数据中的常规值,确保每组实验的数据量尽可能多,同时确保其他 3 个因素稳定在一定区间或不变。由于历史点数量受 PnP 点数限制,表中“*”指取对应 PnP 点数的 60%~70%。

表 2 控制变量实验各因素取值

序号	PnP 点数/个	平均重投影误差/pixel	历史点数量/个	Visual-DOP	数据个数
实验 1	—	1.0~1.5	—	0.05~0.15	356
实验 2	50~60	—	35~40	0.05~0.15	333
实验 3	40	1.0~1.5	—	0.05~0.15	50
实验 4	8	1.0~1.5	3~6	—	67

1)实验 1 的目的是验证 PnP 点数对单目视觉里程计位姿解算质量的评估效果,控制除 PnP 点数外的变量,特殊地,由于历史点数量受 PnP 点数的限制,采取了固定百分比的形式,使得历史点百分比稳定在一定区间内。分别用多源传感器中具有代表性的最小化重投影误差迭代法和 EPnP 方法进行实验,由于一个 PnP 点数对应多组数据,因此对每一个 PnP 点数对应的绝对位置和绝对姿态误差求 RMSE,得到结果如图 9 所示。图 9 反应了 PnP 点数与绝对位姿误差的关系,统计了 PnP 点数为 6 点以上时的位姿精度,不论是 minimized 重投影误差迭代法还是 EPnP 方法,在点数少于 10 点时位姿误差较大。结果显示,两种方法在 PnP 点数较多时具有较好的一致性,当 PnP 点数为 12 个时出现姿态精度拐点,当 PnP 点数为 13 个时出现位置精度拐点,当 PnP 点数大于 40 个后,绝对姿态 RMSE 逐渐趋于稳定值 0.003 37 rad,绝对位置误差逐渐趋于稳定值 0.138 m,随着 PnP 点数增加,位姿误差逐渐减小,定位精度不断提高。实验结果表明,PnP 点数的增加可以弱化三角化过程的偶然误差,增加位姿解算的约束量,达到更高的位姿解算精度,并且在渐变过程中存在拐点,可以作为一个视觉位姿解算精度好坏的评估因素。

2)实验 2 的目的是验证平均重投影误差对单目视觉里程计位姿解算质量的评估效果,平均重投影误差反应了

视觉观测与空间点坐标的拟合程度,也能在一定程度上反应三角化点的坐标精度。统计得到 333 个数据点,如图 10 所示。结果表明,对数据点进行拟合得到斜率为正的回归直线,说明位姿误差和平均重投影误差存在正相关关系,平均重投影误差集中分布在 0.8~2.5 pixel 以内,此时的绝对姿态误差集中在 0.005 rad 以下,绝对位置精度集中在 0.8 m 以下,且拟合误差图表明,随着平均重投影误差的增大,位姿精度逐渐发散。可见,平均重投影误差能够对视觉测量的好坏参与评估。

3)实验 3 的目的是验证历史点数量对单目视觉里程计位姿解算质量的评估效果,历史点本身是其空间坐标被多次优化的点,坐标精度相对于常规空间点更高,用来 PnP 恢复相机位姿效果理应越好,实验中将历史点的判断阈值设置为 3,取 40 个 PnP 点,结果如图 11 所示,历史点数量总体与位姿精度正相关,当 PnP 点数取定为 40 时,历史点数量大于 32,即超过总点数的 80%时,绝对姿态精度集中在 0.006 rad 以下,绝对位置精度集中在 0.25 m 以内。结果表明,当 PnP 点数一定时,历史点数量越多,位姿精度越高。

4)实验 4 的目的是验证 Visual-DOP 对单目视觉里程计位姿解算质量的评估效果,Visual-DOP 由基于位姿矩阵李代数的测量方程推导,反应了空间点到相机中心的空间

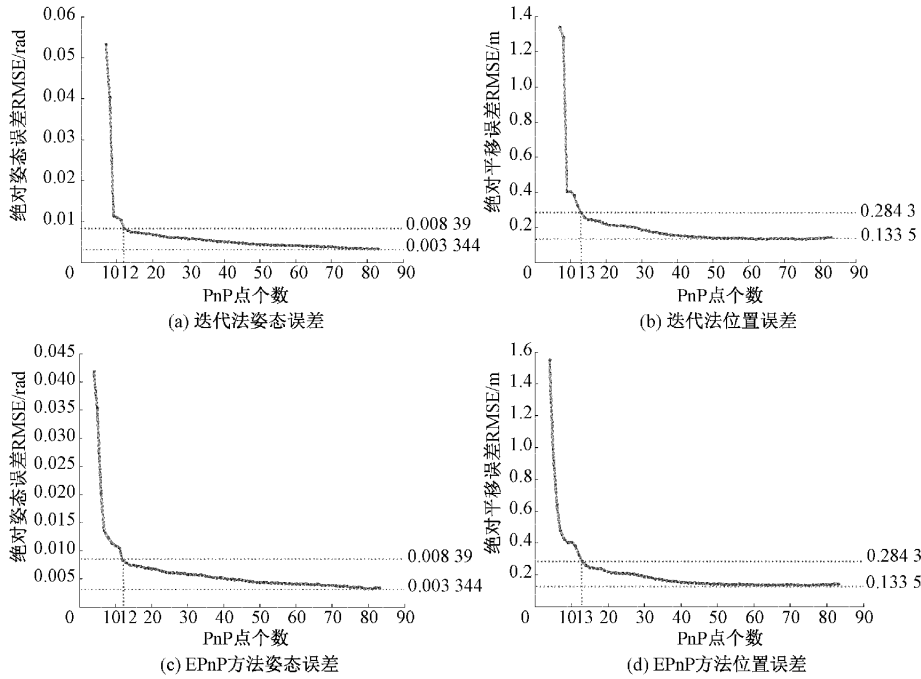


图9 PnP点数与位姿精度的关系

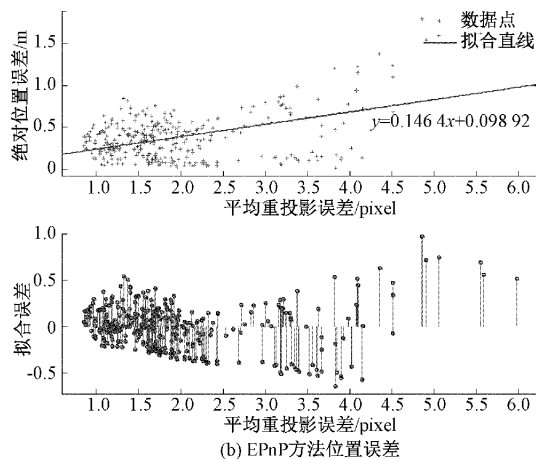
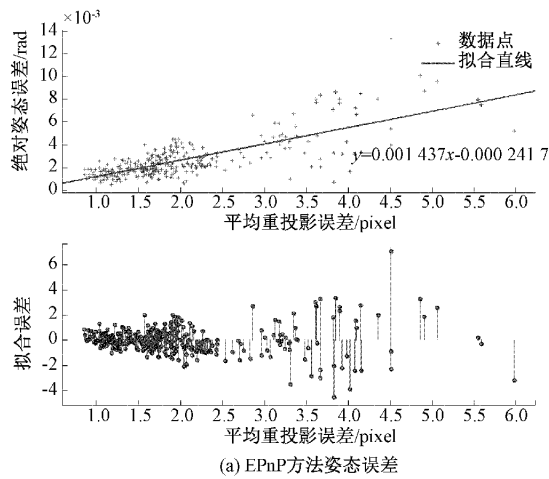


图10 平均重投影误差和位姿精度的关系

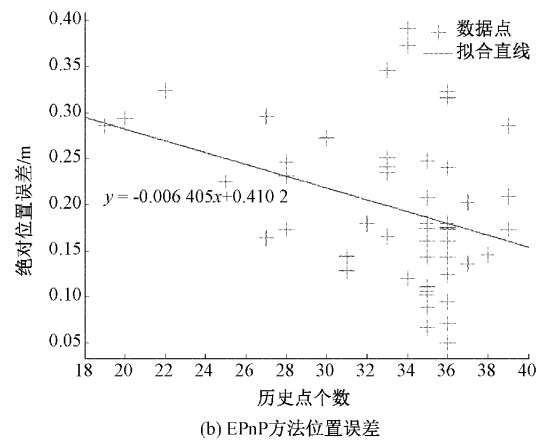
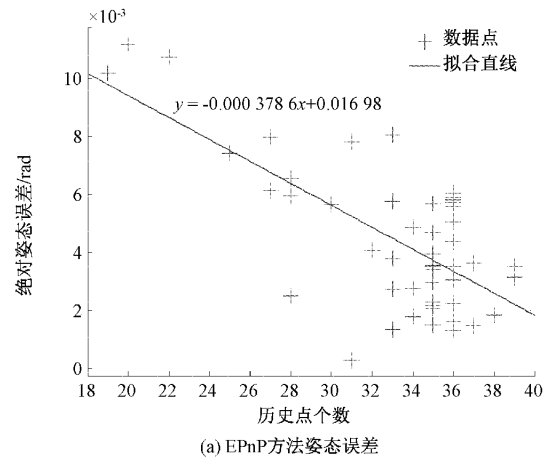


图11 历史点数量和位姿精度的关系

构型,当 PnP 点较多时,空间构型较分散,数据之间的对比度不明显,为了拓宽 Visual-DOP 的值域,取 PnP 点数为 8 点的数,这些点依图像场景随机分布。得到 67 组数据,实验结果如图 12 所示。结果显示,当 Visual-DOP 小于 0.15 时,Visual-DOP 在定位精度的影响因素中不占主导作用,而当 Visual-DOP 大于 0.15 时,定位误差基本呈线性增长,DOP 值为 0.4 时定位误差接近 2 m,Visual-DOP 的增大使得定位精度变差。本文所提通过李代数推导的以测量方程系数阵构建的 Visual-DOP 值能够较好地反应其对单目视觉里程计测量精度的影响。

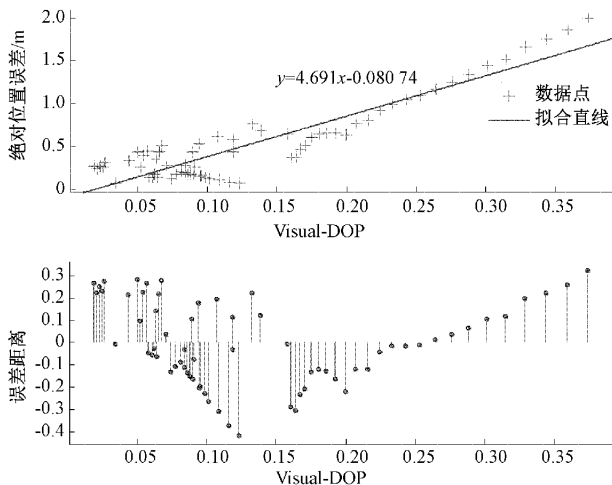


图 12 Visual-DOP 值与绝对位置精度的关系

综合以上实验,在本文单目视觉里程计位姿解算质量的评估中,结论如下:1)PnP 点数与位姿精度关系中,存在拐点,拐点前,PnP 点数影响程度较大,拐点后较平缓,达到一定数量后趋于稳定。2)平均重投影误差在一定阈值内误差比较集中,而随着该误差的不断变大,视觉测量质量明显变差。3)PnP 点数一定时,历史点数量越多,位姿精度趋于更高。4)PnP 点数一定时,Visual-DOP 在一定阈值内对定位精度不起主导作用,随着 DOP 值增大,定位精度明显变差。以 PnP 点数、平均重投影误差、历史点数量、Visual-DOP 值为要素的评估因素能够在一定程度上反应单目视觉里程计位姿解算质量的好坏。

4 结 论

弹性 PNT 和综合 PNT 研究的推进将会推动未来的导航手段向着精度更高、可靠性更好、成本更低的趋势不断发展,需要具备在不同场景和挑战环境下的应对能力,单一传感器的测量质量反应了这种环境的变化,从室内到室外,从少特征、弱纹理环境到特征丰富的场景,视觉定位精度是不断变化的,而如何反应这种精度变化并作用于多源传感器融合导航过程中是解决变化场景下弹性导航定位的关键,本文提供的以 PnP 点数、平均重投影误差、历史点数量、Visual-DOP 值为要素的单目视觉里程计位姿解算

质量评估因素可以反应单独的视觉测量中精度的变化,将为多源传感器融合导航领域的研究人员在视觉条件变化下的导航算法策略选择以及视觉测量协方差调整提供参考和解决思路。

参考文献

- [1] 杨元喜. 综合 PNT 体系及其关键技术[J]. 测绘学报, 2016,45(5):505-510.
- [2] 杨元喜. 弹性 PNT 基本框架[J]. 测绘学报, 2018, 47(7):893-898.
- [3] AQEL M O A, MARHABAN M H, SARIPAN M I, et al. Adaptive-search template matching technique based on vehicle acceleration for monocular visual odometry system[J]. IEEE Transactions on Electrical and Electronic Engineering, 2016,11(6):739-752.
- [4] WON D H, LEE E, HEO M, et al. Selective integration of GNSS, vision sensor, and INS using weighted DOP under GNSS-challenged environments[J]. Transactions on Instrumentation and Measurement, 2014, 63(9): 2288-2298.
- [5] 王高杰,郝向阳,欧阳文,等. 定位点分布对相机位姿解算精度影响分析[J]. 导航定位学报, 2020, 8(6): 21-28,70.
- [6] CAMPOS C, ELVIRA R, RODRÍGUEZ J J G, et al. ORB-SLAM3: An accurate open-source library for visual, visual-inertial and multi-map SLAM[J]. IEEE Transactions on Robotics, 2021, 37(6): 1874-1890.
- [7] LEPETIT V, MORENO-NOGUER F, FUA P. Epnp: An accurate on solution to the pnp problem[J]. International Journal of Computer Vision, 2009, 81(2): 155-166.
- [8] PIZER S M, AMBURN E P, AUSTIN J D, et al. Adaptive histogram equalization and its variations[J]. Computer Vision, Graphics, and Image Processing, 1987, 39(3): 355-368.
- [9] SHI J B, TOMASI C. Good features to track[C]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1994: 593-600.
- [10] LUCAS B D, KANADE T. An iterative image registration technique with an application to stereo vision [C]. Proceedings of DARPA Image Understanding Workshop, 1981:121-130.
- [11] 冉峰,李天,季渊,等. 基于改进 SURF 算法的单目视觉里程计[J]. 电子测量技术, 2017, 40(5): 185-188,200.
- [12] 任彬,宋海丽,赵增旭,等. 基于 RANSAC 的视觉里程计优化方法研究[J]. 仪器仪表学报, 2022, 43(6): 205-212.

- [13] QIN T, LI P, SHEN S. Vins-mono: A robust and versatile monocular visual-inertial state estimator[J]. IEEE Transactions on Robotics, 2018, 34 (4): 1004-1020.
- [14] HARTLEY R I. In defense of the eight-point algorithm[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(6): 580-593.
- [15] GEIGER A, LENZ P, URTASUN R. Are we ready for autonomous driving? The KITTI vision benchmark suite[C]. Conference on Computer Vision and Pattern Recognition. Providence, RI, USA: IEEE, 2012: 3354-3361.

作者简介

李凯林, 硕士研究生, 主要研究方向为机器视觉、图像处理与智能应用。

E-mail: 1341918478@qq.com